BIBLIOGRAFIA

1. E. A. Mendrela, *Uzwojenie silnika prądu przemiennego*, Patent PRL, Nr 92 674.
2. E. A. Mendrela, *Silnik indukcyjny*, Patent PRL, Nr 215 081.
3. G. Kamiński, *Pole magnetyczne, parametry i charakterystyki silnika asynchronicznego o ruchu posuwisto-obrotowym*, Rozprawy Elektrotechniczne, z. 4, 1980, ss. 963—986.
4. A. N. Miljach, *Osnowy teorii elektrodinamiceskich sistiem s triemia stiepieniami swobody dwiżenija*. Izd. AN USSR, Kijew 1956.
5. A. Miljach i in., *Triechstiepiennyje elektriceskije masiny*, Izd. Naukowa Dumka, Kijew 1979.
6. E. A. Mendrela, *Silnik indukcyjny obrotowo-liniowy*, Rozprawy Elektrotechniczne, z. 4, 1976, ss. 383—406.
7. J. Fleszar, E. A. Mendrela, *Twin-armature rotary-linear induction motor*, IEE Proc., Vol. 130, Pt. B, No. 3, May 1983, ss. 186—192.
8. J. Turowski, *Metody obliczania pola i parametrów bieżni indukcyjnych silników liniowych*, Rozprawy Elektrotechniczne, z. 2, 1973, ss. 371—396.
9. E. A. Mendrela, *Analityczna metoda wyznaczania pola i sił w dwutwornikowym indukcyjnym silniku obrotowo-liniowym z ekranowanym wirnikiem*, Rozprawy Elektrotechniczne, z. 2, 1979, ss. 385—400.
10. K. Komęza, *Obliczanie pola elektromagnetycznego i sił w silniku liniowym metodą elementów skończonych*, Rozprawa doktorska, ITMiAE, Pol. Łódzka, 1983.
11. S. A. Nasar, L. de Cid, *Certain approaches to the analysis of single-sided linear induction motors*, IEEE Transaction au PAS, Vol. 120, No 4, April 1973, ss. 477—483.
12. S. Yamamura, *Theory of Linear Induction Motors*, New York, J. Wiley and Sons, 1972.
13. M. Poloujadoff, *The Theory of Linear Induction Machines*, Oxfort University Press, 1980.
14. E. Gierczak, E. A. Mendrela, *Magnetic flux, current, force and power loss distribution in twin armature rotary-linear induction motor*, Scientia Electrica, w druku.
15. E. A. Mendrela i in., *A method of determination of the distance between fictitious primaries in computational model of linear induction motor used in Fourier series technique*, Archiv fur Elektrotechnik 66, 1983, ss. 151—156.
16. E. A. Mendrela, E. Gierczak, *Performance of rotary-linear induction motor with rotating-travelling field*, Electric Machines and Electromechanics, w druku.
17. E. A. Mendrela, *Silniki indukcyjne o dwóch stopniach swobody mechanicznej*, Zesz. Nauk. Pol. Świętokrzyskiej, Elektryka, 1984, monografia, w druku.
18. E. A. Mendrela, E. Gierczak, *Własności elektromechaniczne silnika indukcyjnego o wirniku kulistym*, Rozprawy Elektrotechniczne, z. 4, 1984, w druku.
19. E. A. Mendrela i in., *Stanowisko do badań silników obrotowo-liniowych*, Zesz. Nauk. Pol. Świętokrzyskiej, Elektryka, 7, 1980, ss, 121—129.
20. E. A. Mendrela, J. Turowski, *Rotary-linear duction motor*, Paper No A 781 091-1, IEEE PES Winter Meeting, New York, ss. 1—8, Jan. 29 — Feb. 3, 1978, Catalog No 78 CH 1295-5-PWR.
21. J. Fleszar, *Analiza pracy silnika indukcyjnego obrotowo-liniowego*, Rozprawa doktorska, Pol. Łódzka 1981.
22. M. Kant, Y. Porcher, *Etude d'un entrainement electromecanique a deux degres de liberte*, Publication de L'UTC Compiegne, Francja, 1976.
23. G. Kamiński i in., *Study of the linear rotating*, Proceedings of ICEM, Bruksela 1978.
24. G. Kamiński, *Silnik o ruchu posuwisto-obrotowym*, Patent PRL nr 111 337.
25. G. Kamiński, W. Przyborowski, *Silnik łukowy o ruchu posuwisto-obrotowym*, Patent PRL nr 109428.
26. G. Kamiński, *Pole magnetyczne, parametry i charakterystyki silnika asynchronicznego o ruchu posuwisto-obrotowym*, Rozprawa doktorska, Warszawa 1978.
27. J. Sikora, *Numeryczna analiza przestrzennych obwodów magnetycznych przy wymuszeniu harmonicznych na przykładzie silnika o ruchu posuwisto-obrotowym*, Rozprawa doktorska, Pol. Warszawska, 1979.

28. A. Virolleau, *D'un moteur hybride tournant lineaire*, UTC Compiegne, 1977.
29. E. W. Dwojnych, *Elektromagnitnyj moment elektrodwigatiela s striemia stiepieniami swobody dwiżenia,* Probl. Techn. Elektrodynamiki, Kiew, 1977, wyp. 35, ss. 32—36.
30. I. Kulikowski, *Induktiwnyje izmieritieli pierriemieszczeniej*, Gesonergoizdat, Moskwa 1961.
31. W. A. Barabanow, *O wazmożnosti sozdanija giroskopa s szarowym ratorom na magnitohidrodinamiczeskom podwiesie*, Probl. Techn. Elektrodinamiki, Kiew, 1971, wyp. 29, ss. 160—168.
32. W. Sonelski, *Wybrane zagadnienia z teorii asymetrycznych przetworników elektromechanicznych*, Pol. Śląska, Gliwice.
33. G. Kamiński i in., *Elektryczny silnik kulisty*, Patent PRL/FRANCE nr 784 102.
34. G. Kamiński, W. Przyborowski, *Charakteristic parameters of the complex motion motor*, Proceedings of IMACSTC-1 U.C. DE LOUVAIN, 1984.
35. G. Kamiński, *Badania nad silnikami o ruchu złożonym z wirnikiem kulistym*, Opracowanie nr 043/501/027/6, IME Pol. Warszawaska, 1983.
36. E. A. Mendrela, J. Fleszar, *Equivalent circuit of induction motors wit two degrees of mechanical freedom*, Acta Technica CSAV, No 6, 1983, ss. 697—715.
37. A. Virolleau, *Actionneur a quatre degres de liberte*, Publication de L'UTC Compiegne, Francja, 1977.

E. A. MENDRELA, G. KAMIŃSKI

ELECTRIC MOTORS WITH TWO DEGREES OF MECHANICAL FREEDOM

Summary

The article reviews literature on electric motors with two degrees of mechanical freedom. Main emphasis is put on Polish publications. Using geometrical structure as the foundation of their investigation, the authors identify three main groups of electric machines: planar electric motors, cylindrical electric motors and spherical electric motors. Schematic diagrams of the construction of each type are provided in the paper. With the help of an ideal model of a motor of infinite dimensions elements of a theory of the induction motor with two degrees of mechanical freedom are worked out. In the latter part of the article mechanical and electromechanical properties of the model are discussed.

E. A. MENDRELA, G. KAMIŃSKI

MOTEURS ÉLECTRIQUES À DEUX DEGRÉS DE LIBERTÉ MÉCANIQUE

Résumé

Dans l'étude on a passé en revue les travaux du domaine des moteurs à deux degrés de liberté mécanique, menés chez nnus ainsi qu'à l'étranger, en attachant surtout attention aux premiers. Du point de vue de leur structure géometrique, on a distingué, dans la préface, trois groupes de machines: moteurs plats du type *X-Y*, moteurs à structure cylindrique-rotative-linéaire et moteurs sphériques, en présentant d'une manière schématique leur construction. Ensuite, à partir de l'exemple d'un moteur idéalisé aux dimensions infinies, on a présenté les éléments de la thèorie du moteur d'induction à deux degrés de liberté mécanique. Dans la deuxième partie de l'étude on a traité les problèmes de construction et les propriétés électromécaniques des modèles des moteurs construits en métal.

E. A. MENDRELA, G. KAMIŃSKI

ELEKTROMOTOREN MIT ZWEI MECHANISCHEN FREIHEITSGRADEN

Zusammenfassung

Der Artikel bringt eine Übersicht über die im In- und Ausland geführten Untersuchungen im Bereich der Motoren mit zwei mechanischen Freiheitsgraden unter besonderer Berücksichtigung der inländischen Versuche. In der Einführung zum Thema wurden — vom Gesichtspunkt der geometrischen Struktur ausgehend — drei Maschinengruppen genannt: Flachmotoren Typ *X-Y*, Motoren mit zylindrisch-rotations-linearer Struktur, kugelförmige Motoren. Beispiele derer Konstruktionen wurden schematisch dargestellt. Im weiteren wurden anhand eines idealisierten Motors unendlicher Abmessungen Theorieelemente eines Induktionsmotors mit zwei mechanischen Freiheitsgraden geschildert. Im zweiten Teil des Artikels wurden Konstruktionsprobleme erörtert, und elektromechanische Eigenschaften der in Metall angefertigten Motorenmodelle gebracht.

Е. А. МЕНДРЕЛЯ, Г. КАМИНЬСКИ

ЭЛЕКТРИЧЕСКИЕ ДВИГАТЕЛИ С ДВУМЯ СТЕПЕНЯМИ МЕХАНИЧЕСКОЙ СВОБОДЫ

Резюме

Представлен обзор отечественных и зарубежных работ в области двигателей с двумя степенями механической свободы, в особенности этих первых. Во введении на основании геометрической структуры, выделены три групы машин: плоское двигатели типа *X–Y*, цилиндрические двигатели — линейно поворотные и сферические двигатели. Одновременно приведены примеры конструкции этих двигателей. На примере идеализированного двигателя с бесконечными размерами представлены элементы теории асинхронного двигателя с двумя степенями механической свободы. Во второй части работы рассмотрены вопросы конструкции и представлены электромеханические свойства исполненных в металле моделей двигателей.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 405—422

621.314.21

A Network Approach to the Solution of Stray Field Problems in Large Transformers

JANUSZ TUROWSKI (ŁÓDŹ),

Instytut Maszyn Elektrycznych i Transformatorów Politechniki Łódzkiej

MAREK TUROWSKI (ŁÓDŹ)

Instytut Elektroniki Politechniki Łódzkiej

The paper received 27.7.1983

Important and topical problem of stray flux heating, electrodynamic forces and additional losses can be effectively and quickly investigated by means of method of equivalent reluctance network. The computer program *MSR-2* [5], together with the reluctance network model [6] of the stray field make it possible to investigate the stray field distribution taking into account the complicated geometry and screening materials in the stray field zone. Using the simpler model (Fig. 1a) with 46 nodes, from the paper [6] a list (Table 1) of parameters (reluctances) has been developed for the model with 89 nodes (Fig. 1b). It has been shown that despite this difference, the results obtained by means of either model are very similar (Fig. 3 and 5 and curves *A* in Fig. 4). On the basis of this observation the 9 constructional modifications have been studied: *A*, *B* and *C* (Fig. 7) — using the 89-node model and *D*, *E*, *F*, *G*, *H* and *I* (Fig. 8 and 9) — the 46-node one.

The variants studied are not produced in this form, but they illustrate very well the effectiveness and sensivity of the method to any change in the structure and design of stray field zone. The short computation time — about 30 seconds for any variant — enables one to investigate in a short time many modifications from the point of view of local overheating risk, electrodynamic forces etc.

1. INTRODUCTION

Both earlier [1] and new own experience as well as the last conclusions of the CIGRE Transformer Commitee [2] continually confirm that „Stray flux heating is a very significant concern especially in the large power transformers..." and that „It is becoming increasingly evident that failures in the large transformers are very expensive..." ([2], pp. 44—45). That is why it is so important to have at the disposal a quick and simple method of the evaluation of stray field concentration and ivestigation of influence of a screening and metallic parts geometry on leakage field distribution. The equivalent network model proposed in the papers [3] and [4] seems to be one of the most suitable and effective for this application.

The computer program *MSR-2* [5] and equivalent reluctance network model presented

in the paper [6], together with the table of reluctances (Table 1) make it possible to draw quick technical conclusions in these questions.

The paper presents several such examples. Some of them are not necessarily normally used, but they illustrate well the sensivity and application of the network method to structures with complicated boundary conditions.

The method of equivalent reluctance network is based on the „magnetic Ohm's and Kirchhoff's laws" (Fig. 1)[1]

$$u_{\mu i} = R_{\mu i}\Phi_i = H_i l_i; \quad R_{\mu i} = \frac{l_i}{\mu_i s_i}; \quad \sum_{i=1}^{n} \Phi_i = 0; \quad \sum_{i=1}^{m} u_{\mu i} = 0, \tag{1}$$

where: $i = xz$ (Fig. 1) — N° of branch or space element; $u_{\mu i} = z_i l_i$ — magnetic voltage or magnetomotive force (*m.m.f.*) of excitation coil with z_i turns; $R_{\mu i}$ — reluctance of i-th element; $\Phi_i = B_i s_i$ — magnetic flux in i-th branch; H_i, B_i — magnetic field strength and flux density; l_i, s_i — length and cross-section area of reluctance element.

In the case of transformer the *m.m.f.* of high- (*HV*) and low-voltage (*LV*) winding equals $\hat{F} = \sqrt{2} \cdot I_H z_H = \sqrt{2} \cdot I_L z_L$ and the elementary *m.m.f.* $\hat{U}_p$ of greater loops of the Fig. 1a in the interwinding gap equals

$$\hat{U}_p = \frac{\hat{F}}{n + \frac{n_1}{2}} = \frac{\sqrt{2} \cdot I_H z_H}{n + \frac{n_1}{2}}, \tag{2}$$

where, according to Fig. 1a, $n = 6$ — the amount of the greater loops along the X-axis on the whole height h_u of windings; $n_1 = 4$ — the amount of smaller loops at the top and bottom of windings; $p = 14, 34, 54$ — numbers of loops mentioned above.

The *m.m.f.* of smaller loops on Fig. 1a: $\hat{U}_{p1} = \frac{1}{2}\hat{U}_p$, $(p1 = 74, 94)$.

The theory and the method of calculating particular reluctances for the simpler model (Fig. 1a) have been given in the [6]. It has been shown in the next chapter of this paper that in many cases it is sufficient, for the first rough conclusions, to use the simpler model from Fig. 1a, by which one can claculate a constructional modification during the very short time of about 30 seconds.

2. INFLUENCE OF MESH DENSITY (AMOUNT OF NODES) ON THE ACCURACY OF COMPUTATION

The choice of the amount of nodes in the network model (Fig. 1) can be significant not only considering the computer time, which after all is always very short (30 to 80 seconds), but rather — from the point of view of time necessary for the data preparation and for post-elaboration of results, obtained from the computer.

[1] This analogue is used only in the sense of formal mathematical equivalence. In fact the correct physical analogue of magnetic flux is electric flux and not electric current. Reluctance at the physical analogue should be taken ,,as the inverse of ,,magnetic capacitance" rather than ,,magnetic resistance" but ,,this restriction does not apply if the analogue is only a mathrmatical equivalence" (P. Hammond — 1971, page 171).

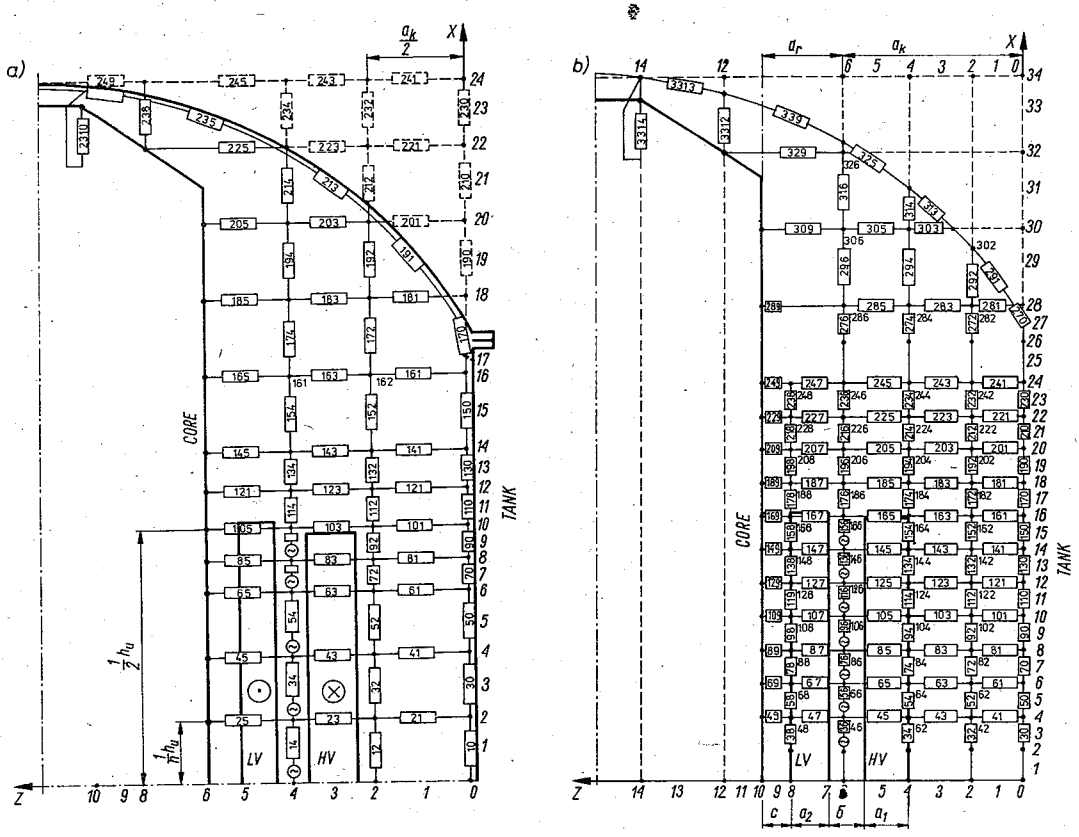


Fig. 1. A network model of stray field of power transformer
a) with 46 or 51 nodes, b) with 89 or 91 nodes

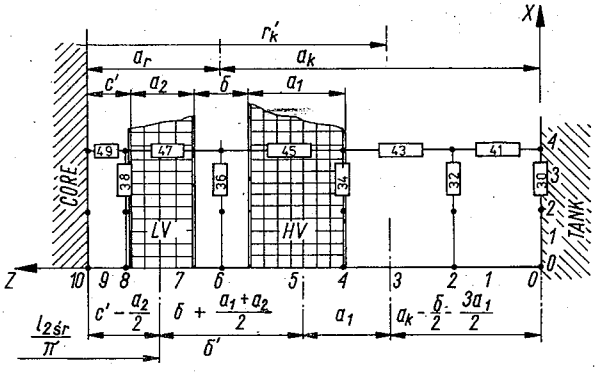


Fig. 2. The assumed axial layer dimensions for the calculation of reluctances of denser ,,89"-model (Fig. 1b)

The parameters (reluctances) of the simpler model (Fig. 1a) calculated and given in the form of Table in the paper [6], can be reused in the denser model (Fig. 1b) by the simple proportional recalculation, according to the new dimensions given in the Fig. 2.

All parameters have been expressed in per unit values (pu v.) related to data of the interwinding gap δ.

2.1. Recalculation of the reluctances from the simpler network (Fig. 1a) to the densor one (Fig. 1b) — Table 1

a) Axial reluctances R_s ($s = 30$ and 270) of the tank wall (Fig. 1b) have the same values as reluctances R_{sa} for the greater loops of Fig. 1a. The tank wall reluctances R_{s1} of smaller loops are the same in both models. The reluctances of oblique cover have been taken roughly as $R_{s2} \approx \sqrt{2} \cdot R_s$ ($s2 = 291$, 313, 325, 339 and 3313 — Fig. 1b).

In order to take into account the induced current reaction, the reluctances of all solid conductive elements have been expressed [6] as complex ones. The nonlinearity of the permeability of steel towards the depth of metal has been considered with the help of linearization coefficients a_p and a_q [6]. The permeability along the X-axis has been taken as constant, but it is possible to take the changes of μ into account by using the iteration method. The field strength H_s^r and H_{s1}^r can be caluculated with the help of the same formulas as in model from Fig. 1a ([6], table 1).

b) Radial reluctances R_k at the tank wall (Fig. 1b, Table 1), at the assumed division of layers according to Fig. 2, have been recalculated from the reluctances R_{ka} [6] of the Fig. 1a in the relation (r — means „relative", i.e. pu value):

$$R_k^r = R_{ka}^r \frac{\frac{1}{2}\left(a_k - \frac{\delta}{2} - \frac{3a_1}{2}\right)}{\frac{1}{2}a_k} \frac{h_u \Big/ \left(n + \frac{n_1}{2}\right)}{h_u/2\left(n + \frac{n_1}{2}\right)} = R_{ka}^r \left(2 - \frac{\delta + 3a_1}{a_k}\right) \approx R_{ka}^r = 1{,}0891 \text{ pu v.} \tag{3}$$

where n and $\frac{n_1}{2}$ according to (2); $k = 61$, 81, 101, 121, 141, 161, 181, 201, 221.

The normal component of flux density B_k^r on a surface of tank wall, in pu v. (for smaller loops in Fig. 1b) can be expressed with the same formula as for Fig. 1a, i.e.

$$B_k^r = 2k_b \Phi_k^r = 1{,}6496 \Phi_k^r \quad (k = 61, 81, 101, 121, 141, 161, 181, 201, 221). \tag{4}$$

For the edge branches it was assumed roughly that

$$R_{k1}^r \approx \frac{2}{3} R_k^r = 0{,}72607 \text{ pu v.}, \quad B_{k1}^r \approx \frac{2}{3} B_k^r = 1{,}1 \Phi_{k1}^r, \quad (k1 = 41, 241) \tag{5}$$

and

$$R_{281}^r \approx \frac{1}{2} R_k^r = 0{,}54455 \text{ pu v.}, \quad B_{281}^r \approx \frac{1}{2} B_k^r = 0{,}8248 \Phi_{281}^r \text{ pu v.} \tag{6}$$

c) Axial reluctances of oil layer close to the tank wall have been recalculated in the relation

$$R_x^r = R_{xa}^r \frac{a_k - \frac{\delta + a_1}{2}}{a_k - \frac{\delta + 3a_1}{2}} \frac{n + \frac{n_1}{2}}{2n + n_1} = 0{,}27016 \text{ pu v.} \tag{7}$$

where $x = 52, 72, 92, 112, 132, 152, 172, 192, 212, 232$; and

$$R_{x1}^r = 2R_x^r = 0{,}4648 \text{ pu v.} \quad (x1 = 32, 272), \tag{8}$$

$$R_{292}^r \approx 1{,}5 R_x^r = 0{,}3486 \text{ pu v.} \tag{9}$$

d) The second column of the radial reluctances towards the tank can be evaluated by the recalculation

$$R_{k3}^r = R_{ka} \frac{\frac{1}{2}\left(a_k - \frac{\delta}{2} - \frac{3a_1}{2}\right) + \frac{a_1}{2}}{\frac{1}{2}a_k} \frac{h_u \Big/ \left(n + \frac{n_1}{2}\right)}{h_u/(2n + n_1)} = R_{ka}\left(2 - \frac{\delta + 2a_1}{a_k}\right) = 1{,}3881 \text{ pu v.} \tag{10}$$

where $k3 = 63, 83, 103, 123, 143, 163, 183, 203, 223$; and

$$R_{k4}^r = \frac{2}{3} R_{k3}^r = 0{,}9254 \text{ pu v.} \quad (k4 = 43, 243), \tag{11}$$

$$R_{283}^r \approx \frac{1}{2} R_{k3} = 0{,}69405 \text{ pu v.}, \tag{12}$$

$$R_{303}^s \approx R_{k3} = 1{,}3881 \text{ pu v.} \tag{13}$$

e) Axial reluctances at the *HV* winding:

$$R_{x2}^r = R_{xa}^r \frac{a_k - \frac{\delta + a_1}{2}}{a_1} \frac{n + n_1/2}{2n + n_1} = 0{,}4958 \text{ pu v.} \tag{14}$$

where $x2 = 54, 74, 94, 114, 134, 154, 174, 194, 214, 234$,

$$R_{x3}^r = 2R_{x2}^r = 0{,}85288 \text{ pu v.} \quad (x3 = 34, 274, 294, 314). \tag{15}$$

f) Radial reluctances in the *HV* winding:

$$R_{k5}^r = R_{ka}^r \frac{\frac{\delta}{2} + a_1}{\frac{1}{2}a_k} \frac{\frac{h_u}{n + \frac{n_1}{2}}}{\frac{h_u}{2n + n_1}} \frac{D + 2a_r + a_k}{D + 2a_r + \delta + a_1} = R_{ka}^r \frac{2\delta + 4a_1}{a_k} \frac{D + 2a_r + a_k}{D + 2a_r + \delta + a_1} =$$

$$= 1{,}8694 \text{ pu v.}, \tag{16}$$

where $k5 = 65, 85, 105, 125, 145, 165, 185, 205, 225,$

$$R^r_{k6} \approx \frac{2}{3} R^r_{k5} = 1{,}2431 \text{ pu v.} \quad (k6 = 45, 245), \tag{17}$$

$$R^r_{k7} = \frac{1}{2} R^r_{k5} = 0{,}9323 \text{ pu v.} \quad (k7 = 285, 305). \tag{18}$$

A radial component of the flux density in the *HV* winding in pu v.:

$$B^r_{k5} = 2k_B \Phi^r_{k5} = 1{,}6496 \Phi^r_{k5}; \quad B^r_{k6} \approx \frac{2}{3} B^r_{k5} = 1{,}0997 \Phi^r_{k6}, \tag{19}$$

$$B^r_{k7} \approx \frac{1}{2} B^r_{k5} = 0{,}8248 \Phi^r_{k7}; \quad (k7 = 285, 305). \tag{20}$$

g) Axial reluctances and *m.m.f.* E in the interwinding gap and above (Fig. 1b):

$$R^r_p = 1{,}0 \text{ pu v.}, \quad E^r_p = 1{,}0 \text{ pu v.}, \quad B^r_p = 1{,}0 \text{ pu v.} \ (p = 36), \tag{21}$$

$$R^r_{p1} = 0{,}5 \text{ pu v.}, \tag{22}$$

where $p1 = 56, 76, 96, 116, 136, 156, 176, 196, 216, 236;$

$$E^r_{p11} = 0{,}5 \text{ pu v.}, \quad B^r_{p11} = 1{,}0 \text{ pu v.}, \tag{23}$$

where $p11 = 56, 76, 96, 116, 136, 156,$

$$R^r_{p2} \approx R^r_{36} \frac{\delta'}{\delta' + \frac{a_2}{2} + c'} = 0{,}7644 \text{ pu v.} \quad (p2 = 276, 296, 316). \tag{24}$$

h) Radial reluctances and flux densities towards core, in the *LV* winding

$$R^r_r = R^r_{ra} \frac{\frac{\delta}{2} + \frac{a_2}{2} + \frac{c' + a_2/2}{2}}{a_r} \frac{h_u/(n + n_1/2)}{h_u/(2n + n_1)} = R^r_{ra} \frac{\delta + \frac{3}{2} a_2 + c'}{a_r} = 1{,}6272 \text{ pu v.}, \tag{25}$$

$$B^r_r = \frac{4\pi r'_k}{l_{2sr}} k_B \Phi^r_r = 2{,}6592 \Phi^r_r \tag{26}$$

where $r = 67, 87, 107, 127, 147, 167, 187, 207, 227;$

$$R^r_{r1} \approx \frac{2}{3} R^r_r = 1{,}0848 \text{ pu v.}, \quad B^r_{r1} = \frac{2}{3} B^r_r = 1{,}7728 \Phi^r_{r1} \ (r1 = 47.247). \tag{27}$$

i) Axial reluctances close to the core surface:

$$R^r_{x8} = R^r_p \frac{\delta'}{c' + \frac{a_2}{2}} = 1{,}6224 \text{ pu v.} \tag{28}$$

where $x8 = 58, 78, 98, 118, 138, 158, 178, 198, 218, 238;$

$$R^r_{38} = 2R^r_{x8} = 3{,}2448 \text{ pu v.} \tag{29}$$

j) Radial reluctances and flux densities close to the core surface, for the smaller loops:

$$R^r_{r2} = R^r_{ra} \frac{\frac{1}{2}\left(c' + \frac{a_2}{2}\right)}{a_r} \cdot 2 = 0{,}4294 \text{ pu v.}, \tag{30}$$

$$B^r_{r2} = \frac{4\pi r'_k}{l_{2śr}} k_B \Phi^r_{r2} = 2{,}6592\,\Phi^r_{r2}, \tag{31}$$

where $r2 = 69, 89, 109, 129, 149, 169, 189, 209, 229$;

$$R^r_{49} \approx R^r_{249} \approx \frac{2}{3} R^r_{r2} = 0{,}2863 \text{ pu v.}, \quad B^r_{49} \approx B^r_{249} \approx \frac{2}{3} B^r_{r2} = 1{,}7728\,\Phi^r_{49}, \tag{32}$$

$$R^r_{289} = R^r_{ra} = 1{,}0283 \text{ pu v.}, \quad B^r_{289} = 1{,}3296\,\Phi^r_{289}. \tag{33}$$

k) Radial reluctances and flux densities close to the core surface, for the greater loops:

$$R^r_{309} = R^r_{ra} = 1{,}0283 \text{ pu v.}, \quad B^r_{309} \approx 1{,}3296\,\Phi^r_{309}, \tag{34}$$

$$\begin{gathered} R^r_{329} \approx 1{,}4 \cdot R^r_{ra} = 1{,}4396 \text{ pu v. (see Fig. 1a)}, \\ B^r_{329} \approx 1{,}4 \cdot 1{,}3296\,\Phi^r_{329} = 1{,}86\,\Phi^r_{329}, \end{gathered} \tag{35}$$

$$R^r_{3312} \approx R^r_{238a} = 0{,}4762 \text{ pu v.}, \quad B^r_{3312} \approx 0{,}5832\,\Phi^r_{3312}, \tag{36}$$

$$R^r_{3314} \approx R^r_{2310a} = 0{,}1905 \text{ pu v.}, \quad B^r_{3314} \approx 0{,}5832\,\Phi^r_{3314}. \tag{37}$$

2.1. Data introduction and results of computation

The computation of field distribution has been conducted with the help of both models (Fig. 1), using the parameters given in the paper [6] for the simpler model (Fig. 1a). The parameters (reluctances) for the denser model (Fig. 1b) are listed in the Table 1.

For the complex reluctances R_{si} of the solid metalic parts a „magnetic inductance", as a formal mathematical equivalence in the form $L_{si} = \frac{1}{\omega} \operatorname{Im}\{R_{si}\}$ has been introduced.

Comparison of the results computed with the help of both models shows (Fig. 3, 4, 5 and 6) that, in spite of a great difference between the amount of nodes (46 versus 89) [2] in the models, a quite satisfactory agreement of the results has been obtained.

Fig. 3 and the modification *A* in the Fig. 4 show the practically acceptable agreements of graphs obtained from the 46- and 89-nodes model. The discrepancies in the graphs *B* and *C* on Fig. 4 are caused by the fact that the length of the top magnetic shunt *S* in the „46"-model is taken by 15,2% longer than in the „89"-model.

Fig. 4. shows at the same time how effectively the top magnetic shunt *S* decreases the losses in the tank wall. Especially if it is taken into account that the stray losses in the tank wall are proportional to the $\hat{H}_k^2$ or more [1].

[2] Complex reluctances of solid metalic parts use one additional node to divide (separate) the real and imaginary part of reluctance.

Table 1

The reluctances of the network model with 89 nodes (Fig. 1b) of stray field of 240 MVA transformer

Part	Reluctance		No in the Fig. 1b	Magnetic field
	Symbol	pu value		
				$\|H^r\| =$
Tank wall	R_s^r	0,09321 + $+j$0,1512 a) 63 b) 0	$L_s = 4{,}8128 \cdot 10^{-4}$ s s = 30,270; a) When screen Cu 4 mm b) When screen Fe, $\mu = \infty$	$2{,}0472 \cdot \mathrm{Re}(\Delta U_s^r)$ a) $1{,}0741 \cdot \Delta U_s^r$ b) 0
—,,—	R_{s1}^r	0,04661 + $+j$0,0756 a) 31,5 b) 0	$L_{s1} = 2{,}4064 \cdot 10^{-4}$ s $s1$ = 50, 70, 90, 110, 130, 150, 170, 190, 210, 230 a) When screen Cu 4 mm b) When screen Fe, $\mu = \infty$	$4{,}0945 \cdot \mathrm{Re}(\Delta U_{s1}^r)$ a) $2{,}1482 \Delta U_{s1}^r$ b) 0
Cover	R_{s2}^r	0,1318 + $+j$0,2138 a) 91,25 b) 0	$L_{s2} = 6{,}8035 \cdot 10^{-4}$ s $s2$ = 291, 313, 325, 339, 3313 a) When screen Cu 4 mm b) When screen Fe, $\mu = \infty$	$1{,}4476\ \mathrm{Re}(\Delta U_{s2}^r)$ a) $0{,}7595 \cdot \Delta U_{s2}^r$ b) 0
				$\|B^r\| =$
Radial to tank	R_k^r	1,0891	k = 61, 81, 101, 121, 141, 161, 181, 201, 221	$1{,}6496\ \Phi_k^r$
—,,—	R_{k1}^r	0,72607	$k1$ = 41, 241	$1{,}0997\ \Phi_{k1}^r$
—,,—	R_{281}^r	0,54455	281	$0{,}8248\ \Phi_{281}^r$
Axial to tank	R_x^r	0,27016	x = 52, 72, 92, 112, 132, 152, 172, 192, 212, 232	—
	R_{x1}^r	0,4648	$x1$ = 32, 272	—
	R_{292}^r	0,3486	292	—
Radial to tank	R_{k3}^r	1,3881	$k3$ = 63, 83, 103, 123, 143, 163, 183, 203, 223, 303	—
	R_{283}^r	0,69405	283	—
Axial at *HV* winding	R_{x2}^r	0,4958	$x2$ = 54, 74, 94, 114, 134, 154, 174, 194, 214, 234	—
	R_{x3}^r	0,85288	$x3$ = 34, 274, 294, 314	—

c.d. Table 1

Radial in *HV* winding	R^r_{k5}	1,8694	$k5 = 65, 85, 105, 125, 145, 165, 185, 205, 225$	$B^r_{k5} = 1{,}6496\ \Phi^r_{k5}$
	R^r_{k6}	1,2431	$k6 = 45, 245$	$B^r_{k6} = 1{,}0997\ \Phi^r_{k6}$
	R^r_{k7}	0,9323	$k7 = 285, 305$	$B^r_{k7} = 0{,}8248\ \Phi^r_{k7}$
In gap and over	R^r_p; E^r_p	1; 1	$p = 36$	$B^r_p = 1$; $H^r_p = 1$
	R^r_{p1}	0,5	$p1 = 56, 76, 96, 116, 136, 156, 176, 196, 216, 236$	—
	E^r_{p11}	0,5	$p11 = 56, 76, 96, 116, 136, 156$	$B^r_{p11} = 1$; $H^r_{p11} = 0{,}5$
	R^r_{p2}	0,7644	$p2 = 276, 296, 316$	—
Radial in *LV* winding	R^r_r	1,6272	$r = 67, 87, 107, 127, 147, 167, 187, 207, 227$	$B^r_r = 2{,}6592\ \Phi^r_r$
	R^r_{r1}	1,0848	$r1 = 47, 247$	$B^r_{r1} = 1{,}7728\ \Phi^r_{r1}$
Axial by core	R^r_{x8}	1,6224	$x8 = 58, 78, 98, 118, 138, 158, 178, 198, 218, 238$	—
	R^r_{38}	3,2448	38	—
Radial by core	R^r_{r2}	0,4294	$r2 = 69, 89, 109, 129, 149, 169, 189, 209, 229$	$B^r_{r2} = 2{,}6592\ \Phi^r_{r2}$
	R^r_{r3}	0,2863	$r3 = 49, 249$	$B^r_{r3} = 1{,}7728\ \Phi^r_{r3}$
	R^r_{289}	1,0283	289	$B^r_{289} = 1{,}3296\ \Phi^r_{289}$
By the core yoke	R^r_{309}	1,0283	309	$B^r_{309} \cong 1{,}3296\ \Phi^r_{309}$
	R^r_{329}	1,4396	329	$B^r_{329} \cong 1{,}86\ \Phi^r_{329}$
	R^r_{3312}	0,4762	3312	$B^r_{3312} \cong 0{,}5832\ \Phi^r_{3312}$
	R^r_{3314}	0,1905	3314	$B^r_{3314} \cong 0{,}5832\ \Phi^r_{3314}$

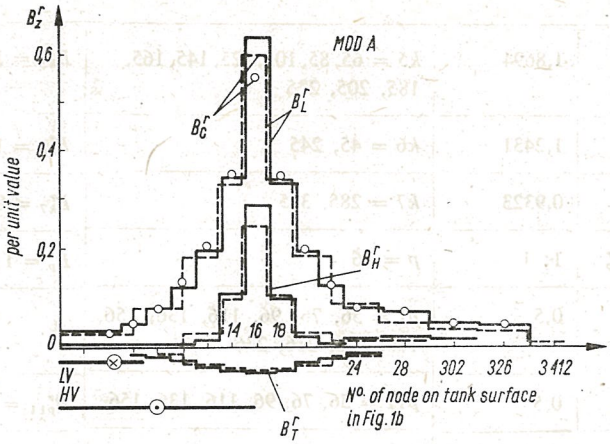


Fig. 3. Influence of the reluctance network density on the results of computation of radial component of stray flux density in 240 MVA transformer without any screening or shunts (MODification *A*)

--- network with 46 nodes (Fig. 1a); — network with 89 nodes (Fig. 1b); B_T^r — on the tank surface, B_H^r — on the axis of *HV* winding, B_L^r — on the axis of *LV* winiding, B_C^r — on the surface of the core (ooo — for Fig. 1b)

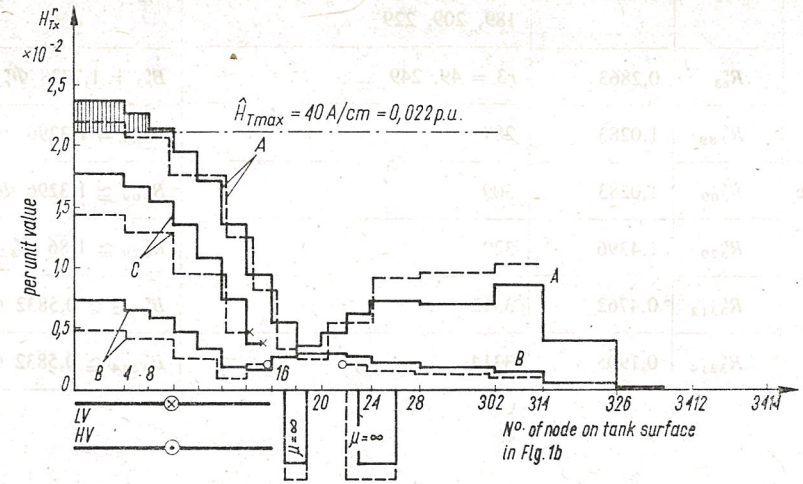


Fig. 4. Influence of the reluctance network density on the results of computation of a tangent component of magnetic field strength H_k^r on the tank surface

--- 46 nodes (Fig. 1a); — 89 nodes (Fig. 1b); *A* — transformer without any screening or shunts, *B* — with the top and bottom magnetic shunts under yoke beams (Fig. 5) — with the relucatnces $R_{125}^r = R_{123}^r = 0$ (Fig. 1a) and $R_{189}^r = R_{187}^r = R_{185}^r = 0$ (for Fig. 1b), *C* — like *B* but with higher poistion of shunt, with reluctances $R_{165}^r = R_{163}^r = 0$ (for Fig. 1a) and $R_{249}^r = R_{247}^r = R_{245} = 0$ (for Fig. 1b)

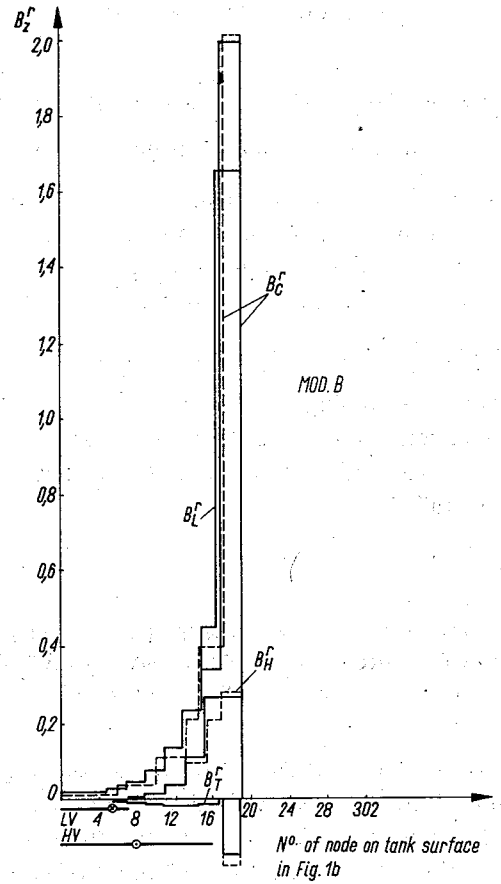


Fig. 5. Influence of the reluctance network density on the results of computation of radial component of stray flux density B_z^r in 240 MVA transformer with the lower position of magnetic shunt (MODification *B*)

Symbols as in Fig. 3 and 4

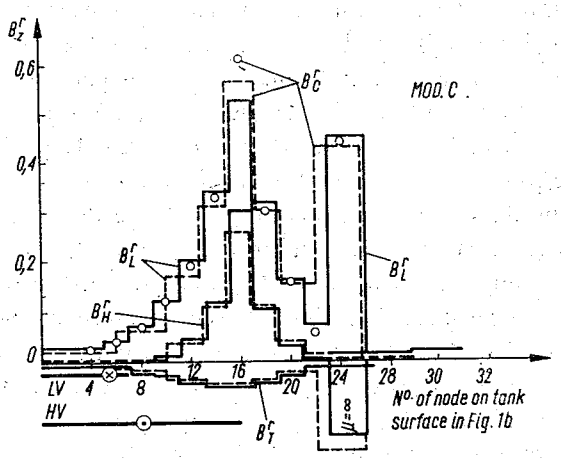


Fig. 6. Influence of reluctance network density on the results of computation of radial component B_z^r of stray flux density in 240 MVA transformer, with the higher position of magnetic shunt (MODification *C*)

Symbols as in the Fig. 3 and 4

It can also be seen from the Fig. 4, that the tangent component of magnetic field strength $\hat{H}_k$ on surface of the tank does not exceed much in any place the lowest value $\hat{H}_{s,max} = 40\frac{A}{cm}$, permissible from the point of view of local overheating hazard (see [1], page 9). This permissible value depends on the acceptable hot spot temperature, way of field excitation and cooling. Considering the last CIGRE opinions [2] that the acceptable hot spot temperature of metal parts, being in contact with oil, can reach at the overload 140°C and even 180°C, the permissible field strength $\hat{H}_{s,max}$ can reach 60 to $150\frac{A}{cm}$, depending on way of field excitation and cooling.

On the basis of positive conclusion from the comparison of both models from Fig. 1, a further analysis of the design modlfications *D* to *I* has been carried out with the help of a simpler model (Fig. 1a).

3. INFLUENCE OF THE DESIGN AND SCREENING ON THE STRAY FIELD, ELECTRODYNAMIC FORCES AND LOCAL OVERHEATING HAZARD IN LARGE POWER TRANSFORMERS

Using the computer program *MSR-2* [5] and both kinds of computation models (Fig. 1) the 9 various design modifications of stray field zone in large transformers have been studied. Namely 3 modifications of the denser „89"-model: *A*, *B*, *C* (Fig. 7) and 6 modifications of the simpler „46"-model: *D*, *E*, *F*, *G*, *H* and *I* (Fig. 8 and 9).

As can be seen from Fig. 7A the drawing of the stepped graphs is not necessary becaues the continuous line connecting the computed points goes through the centers of the steps. Only the top „step" should be used for the extrapolation of top value on the principle of equal areas, as shown in circle (Fig. 7A).

The basic modification *A* (Fig. 3, 4 and 7A) has no screens or shounts. The modifications *B* and *C* (Fig. 5, 6 and 7) have been calculated on the assumption that the magnetic shounts *S* are made of laminated transformer iron, with magnetic permeability $\mu = \infty$ and conductivity $\gamma = 0$ ($R' = 0$). The big influence of the position of this shunt on the local overheating of core surface is evident (Fig. 7B and C). The shunts also influence on axial electrodynamic forces and additional losses in windings.

Apart from the insulation requirements, one can state that placing shunt *S* close to the top of windings (Fig. 7B) removes extreme radial flux density out of the windings. Owing to this, the maximum axial forces in *LV* winding decrease by ca. 45% and in *HV* winding by ca. 30%. At the same time additional losses and local heating on the edge of windings decrease too. However this shunt causes simultaneously significant increase of the risk of local overheating of external packages of magnetic core.

The higher position of shunt *S* (Fig. 7C) not only restores the previous situation from the Fig. 7A, but even causes additional superfluous stray field concentration on the surface of core and tank. The optimization of this position will be the subject of further investigation.

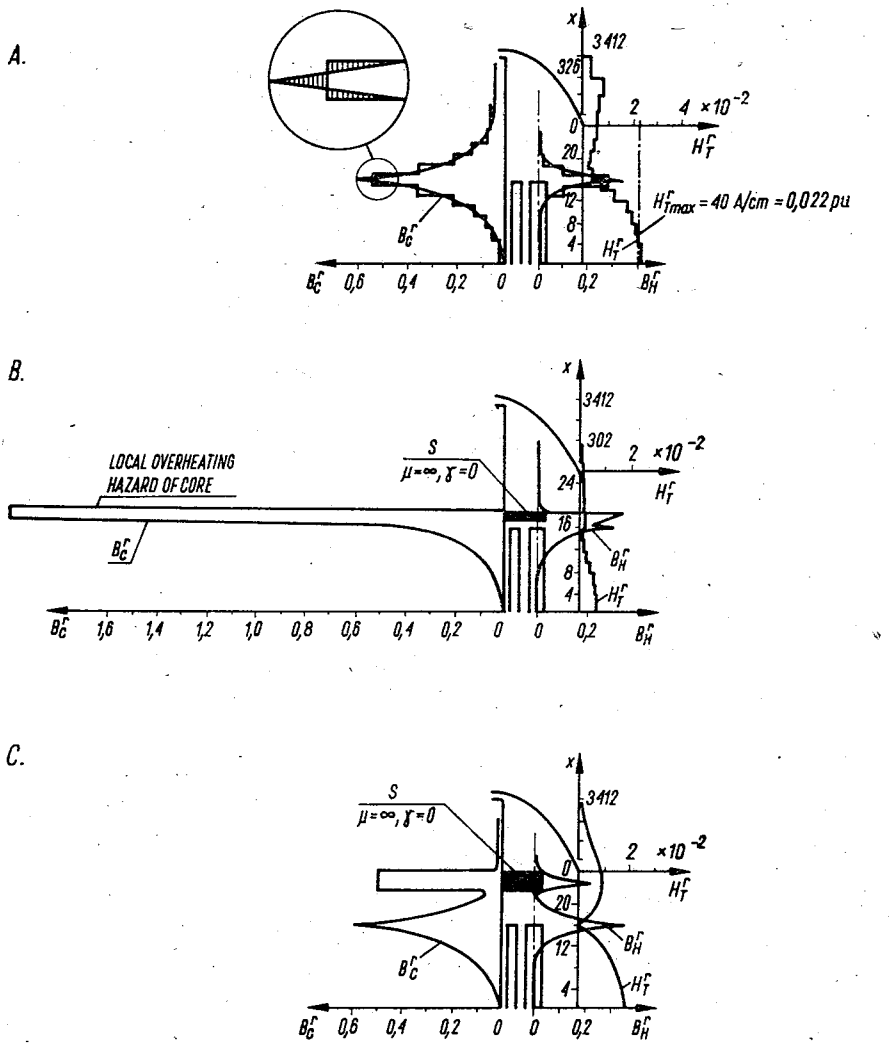


Fig. 7. Influence of magnetic shunt S at the top and bottom of windings on a local overheating hazard of the tank wall and/or core

Model from Fig. 1b. Symbols as in Fig. 3 and 4

The existence of cover and tank flanged joints at the top part of tank (Fig. 8 and 9) has practically not important influence on the stray field distribution, apart from this spot. The introduction of idealised ($\mu = \infty$, $\gamma = 0$, $R = 0$) magnetic screen on the tank wall (Fig. 8E) causes — as expected — the strong dercease of tank losses and some increase of axial forces and additional losses in HV winding. Fig. 8F demonstrates the field and eddy current distribution on the edge of electromagnetic (Cu) screens. Despite the big increase of H_T^r, this value is much lower than the permissible one, from the point of view of local heating risk. According to paper [1] (page 9) this value shold not exceed $\hat{H}_{s,\,max} = 4530\sqrt{d_m} = 4530\sqrt{4 \cdot 10^{-3}} = 286{,}5\,\frac{\text{A}}{\text{cm}}$, i.e. $18{,}8 \cdot 10^{-2}$ per unit value.

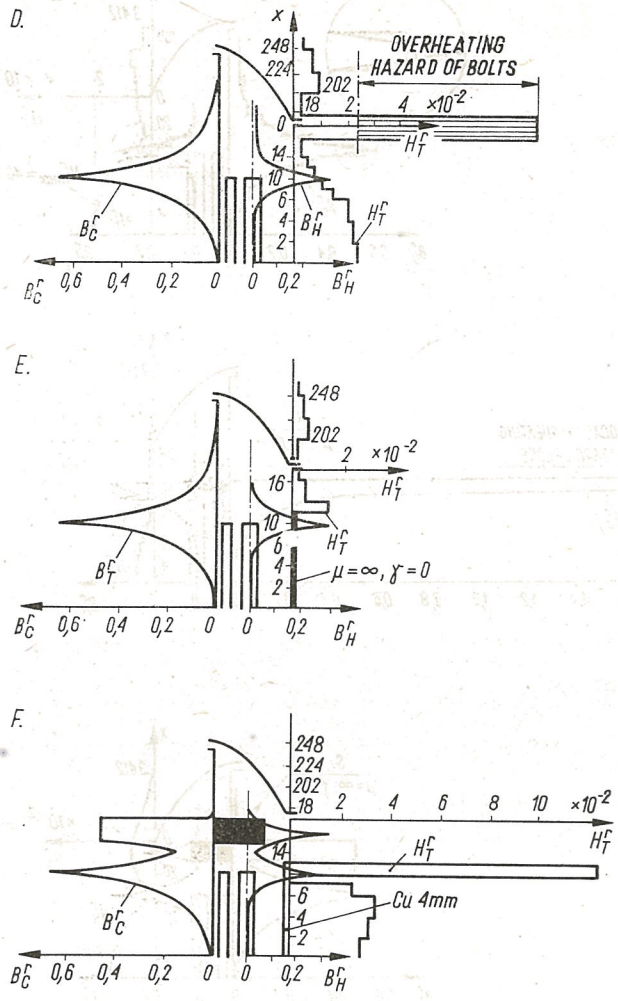


Fig. 8. Influence of tank screening and position of the cover and tank flanged joint on the local overheating hazard of tank wall and/or core

D — with joint but without screening, *E* — with joint and magnetic screening of tank wall, *F* — with joint, magnetic top shun and electromagnetic (Cu 4 mm) screening of the tank wall. Model from Fig. 1a. Symbols as in Fig. 3 and 4

Electromagnetic screens (Fig. 8F and 9) as well as the yoke beams made of solid steel (Fig. 9H and I) have neglible influence on the stray field distribution within windings however on the solid steel yoke beams appears new risk of local overheating, because the field H_B^r on its surface is greater than the permissible value.

The problem of determination of permissible value $\hat{H}_{s,max}$, taking into account the last conclusions from the Boston Transformer Colloquium of CIGRE [2], will be presented in another paper.

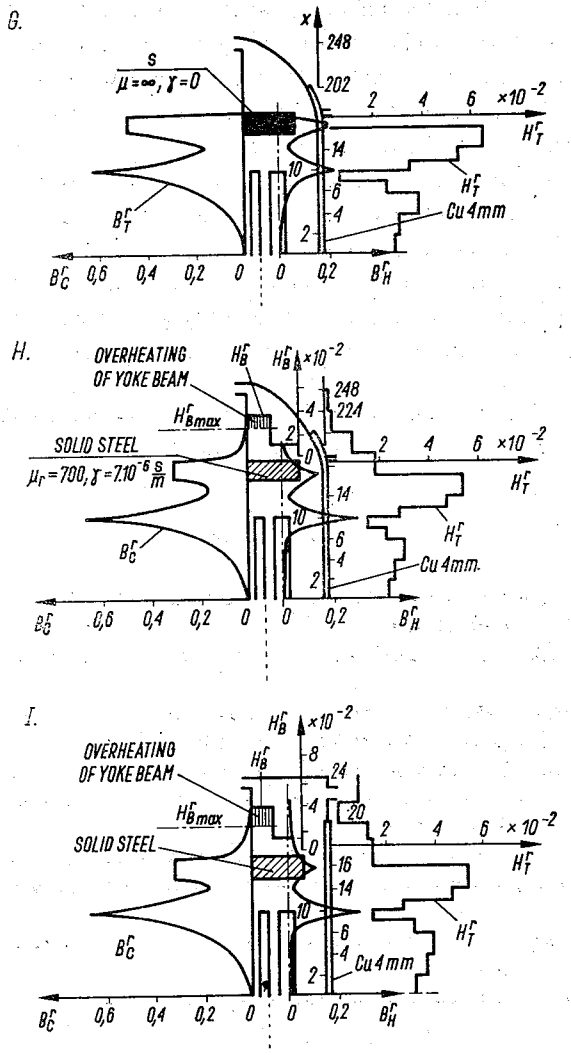


Fig. 9. Influence of electromagnetic (Cu 4 mm) screens and solid steel yoke beams on the stray field di stribution and local overheating hazard

4. CONCLUSIONS

The method of equivalent reluctance network appeared to be one of the most convenient and not too sophisticated way of analysis and optimisation of a stray field zone in large power transformers. It makes it possible to resolve and investigate the design variants with complicated geometry and boundary conditions. The accuracy of computation can be easily increased by a simple multiplication of the amount of nodes and loops of the model, but for the first rough evaluations the very simple model (Fig. 1a), with the time of computation of about 30 seconds, can be used.

In order to make sure that the half of a winding height (Fig. 1) is the satisfactory model of a transformer's stray field, the whole-height model has been computed additionaly, i.e. containing upper and low half from Fig. 1a. Results in corresponding branches of network were exactly the same and there was no current in horizontal OZ-axis. It fully confirms the correctness of the one-half model.

Figures 7 to 9 show not only qualitative, but also quantitative effects of any change in the geometry, material and dimensions of screens in transformer. The method is very sensitive to any change of these parameters.

The method presented enables one also to take into account the finite dimensions and asymmetry of windings [6], nonlinear permeability [5] as well as the three-phase nature and three-dimensional structure of the stray field in transformers [3].

Acknowledgement. This work has been funded by the Central Research Program 05.5A for the fundamental investigations in electromagnetic fields.

REFERENCES

1. M. Kozłowski, J. Turowski, *Stary losses and local overheating hazard in transformers*, Paris CIGRE, Session 1972, Report 12—10.
2. L. W. Long, *Overloading and loading limitations of large transformers*, ELECTRA, CIGRE, No. 86, 1983, pp. 33—51.
3. J. Turowski, *Network model of stray field in a three-phase transformer* (in Polish), Zeszyty Nauk. Politechniki Łódzkiej „Elektryka", No. 66, 1979, pp. 105—122.
4. J. Turowski, *Trends in calculation and investigation of stray losses and local overheating in large transformers*, International Colloquium on Transformers, CIGRE Study Committe No. 12 (Transformers), Cambridge, USA, 2—8 Sept. 1981.
5. M. Turowski, M. Kopeć, *The program for fast stray-field analysis in power transformers with the aid of a method of equivalent reluctance network* (in Polish), Rozprawy Elektrotechniczne, No. 2, 1985.
6. J. Turowski, *Reluctance model of stray field in transformer* (in Polish), Rozprawy Elektrotechniczne, No. 4, 1984, pp. 1121—1144.
7. J. Turowski, *Technical Electrodynamics* (in Polish), WNT, Warszawa, 1968.

J. TUROWSKI, M. TUROWSKI

ZASTOSOWANIE METODY SIECI RELUKTANCYJNYCH DO ROZWIĄZYWANIA PROBLEMÓW ROZPROSZENIOWYCH W TRANSFORMATORACH

Streszczenie

Bardzo ważny i aktualny problem lokalnych nagrzań, sił elektrodynamicznych i strat dodatkowych, powodowanych przez strumień rozproszenia w transformatorze może być efektywnie i szybko badany za pomocą metody równoważnych sieci reluktancyjnych. Program komputerowy *MSR-2* [5], łącznie z modelem sieci reluktancyjnej [6] dla pola rozproszenia, umożliwia badanie rozkładu strumienia rozproszenia z uwzględnieniem skomplikowanej geometrii i różnych materiałów ekranów, stosowanych w strefie rozproszeniowej transformatora.

Wykorzystując prostszy model (rys. 1a) na 46 węzłów, z pracy [6], opracowano tabelę (tabl. 1) parametrów (reluktancji) dla modelu o 89 węzłach (rys. 1b). Wykazano, że mimo tych różnic wyniki otrzymane za pomocą obu modeli są bardzo zbliżone (rys. 3 i 5 oraz krzywe A na rys. 4). Na tej podstawie przeba-

dano 9 modyfikacji konstrukcyjnych, z czego *A*, *B* i *C* (rys. 7) — za pomocą modelu o 89 węzłach oraz *D*, *E*, *F*, *G*, *H* i *I* (rys. 8 i 9) — z wykorzystaniem modelu o 46 węzłach. Przebadane warianty nie są produkowane w praktyce, lecz służą do zilustrowania skuteczności i czułości metody na wszelkie zmiany układu i konstrukcji strefy rozproszeniowej. Krótki czas obliczeń maszynowych — około 30 sekund na każdy wariant — umożliwia szybkie badanie wielu modyfikacji konstrukcyjnych z punktu widzenia lokalnych przegrzań, sił elektrodynamicznych itd.

J. TUROWSKI, M. TUROWSKI

APPLICATION DE LA MÉTHODE DU RÉSEAU DES RÉLUCTANCES ÉQUIVALENT POUR RÉSOUDRE LES PROBLÈMES DU CHAMP DE FUITES DES TRANSFORMATEURS

Résumé

Le problème très important et actuel des chauffements locaux, des forces électrodynamiques et des pertes supplémentaires, causées par le champ de fuites des transformateurs, peut être facilement et rapidement examiné par méthode du réseau des réluctances équivalent.

Le programme de l'ordinateur *MSR*-2 [5] conjointement au modèle du réseau des réluctances [6] pour le champ de fuites, rend possible l'examen de la répartition du flux de fuites, ayant égard à la forme géométrique compliquée et à la diversité des matériaux des écrans, utilisés dans la zone de fuites du transformateur.

En appliquant le modèle plus simple (fig. 1a) à 46 noeuds de l'étude [6], on a élaboré le tableau (tabl. 1) des paramètres (des réluctances) pour le modèle à 89 noeuds (fig. 1b). On a montré que malgré ces différences les résultats obtenus avec les deux modèles sont tout proches (fig. 3 et 5 et les courbes *A* de la fig. 4). Se basant sur ces résultats on a examiné 9 constructions différentes, dont *A*, *B* et *C* (fig. 7) à l'aide du modèle à 89 noeuds et *D*, *E*, *F*, *G*, *H* et *I* (fig. 8 et 9) à l'aide du modèle à 46 noeuds. Les variations examinées ne sont pas produites en pratique et ne servent qu'à illustrer l'effectivité et la sensibilité de la méthode correspondant aux changements du système et à la construction de la zone de fuites. La courte durée des calculs à l'aide de l'ordinateur — 30 secondes à peu près pour chaque variation — permet d'examiner rapidement beaucoup de modifications de construction du point de vue des chauffements locaux, des forces électrodynamiques etc.

J. TUROWSKI, M. TUROWSKI

ANWENDUNG DER NETZMETHODE ZUR LÖSUNG VON STREUFLUßPROBLEMEN GROßER TRANSFORMATOREN

Zusammenfassung

Wichtige und aktuelle Probleme der durch Streufluß hervorgerufenen Erwärmung, der elektrodynamischen Kräfte sowie Zusatzverluste können wirksam und schnell mit Hilfe einer äquivalenten Reluktanz-Netzmethode untersucht werden. Ein *MSR-2*-Komputerprogramm [5] ermöglicht gemeinsam mit einem Reluktanz-Netzmodell des Streuflusses [6] die Untersuchung der Streufeldverteilung unter Berücksichtigung komplizierter Geometrie und des Abschirmungsmaterials in der Streufeldzone. Mit Hilfe des vereinfachten Modells (Fig. 1a) mit 46 Knoten wurde nach Artikel [6] die Tafel (Table 1) der Parameter (der Reluktanzen) für ein Modell mit 89 Knoten (Fig. 1b) berechnet. Es wurde nachgewiesen, daß trotz dieser Knotenzahlverschiedenheit die Ergebnisse, die mittels der beiden Modelle errechnet werden, sich wenig voreinander unterscheiden (Fig. 3 und 5 und Kurve *A* gemäß Fig. 4). Auf Grund dieser Feststellungen sind 9 konstruktive Modifikationen untersucht worden: *A*, *B* und *C* (Fig. 7) — mit Hilfe des 89-Knoten-Modells und *D*, *E*, *F*, *G*, *H* und *I* (Fig. 8 und 9) — mittels des 46-Knoten-Modells.

Die dargestellten Varianten werden in dieser Form nicht erzeugt, sie illustrieren jedoch gute Wirksamkeit und Empfindlichkeit der Netzmethode bezüglich jeder Konstruktionsänderung in der Streufeldzone. Die kurze Berechnungszeit ermöglicht die Durchführung von Untersuchungen vieler Modifikationen im kurzen Zeitraum hinsichtlich der Gefahr lokaler Überwärmung, elektrodynamischer Kräfte u.dgl.

Я. ТУРОВСКИ, М. ТУРОВСКИ

ПРИМЕНЕНИЕ МЕТОДА ЦЕПЕЙ К РЕШЕНИЮ ПРОБЛЕМ ПОЛЯ РАССЕЯНИЯ В МОЩНЫХ ТРАНСФОРМАТОРАХ

Р е з ю м е

Так собственный опыт [1] как и последние выводы Трансформаторного Комитета СИГРЭ [2] подтверждают постоянно большое значение проблем мастных нагревов вызываемых потоками рассеяния. Компютерная программа *MSR*-2 [5], вместе с моделью релюктансной цепи [6] позволяют быстро и эффективно исследовать распределение поля рассеяния и его концентрации при сложной геометрии и таких граничных условиях, которые не могли быть учтены при применяемых до сих пор аналитических методах расчета. На основании упрощенной модели на 46 узлов из работы [6] (рис. 1а) разработана таблица (Таб. 1) магнитных сопротивлений для модели на 89 узлов (рис. 16). Таблица 1 указывает на примерах способ ввода данных в программу *MSR*-2 и обработки результатов. Показано, что результаты расчета поля, несмотря на 2-кратное увеличение числа узлов, не на много отличаются (рис. 3 и 5 и кривая *A* на рис. 4). На этом основании было исследовано распределение поля рассеяния для 9-ти конструкционных модификаций, из чего три *A*, *B*, *C* (рис. 7) — на модели с 89 узлами и шесть *D*, *E*, *F*, *G*, *H*, *I* (рис. 8 и 9) — на модели с 46 узлами. Рисунки эти не представляют непосредственных решений применяемых на практике, но показывают эффективность и чувствительность метода на модели изменения материалов и геометрии зоны рассеяния, а также эффективность экранирования. Время расчета одного варианта — около 30 секунд — позваляет на быстрое исследование многих модификаций, с точки зрения электромагнитных усилий, добавочных потерь и местных парегревов.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 423—436

621.365

Siły elektrodynamiczne w ciekłym metalu mieszanym indukcyjnie w procesach półciągłego lub ciągłego odlewania wlewków cylindrycznych

CZESŁAW SAJDAK (KATOWICE)

Instytut Metalurgii Politechniki Śląskiej

Otrzymano 25.4.1983

Praca zawiera analizę pola elektromagnetycznego w cylindrycznym mieszadle indukcyjnym, wytwarzającym pole biegnące, stosowanym przy odlewaniu wlewków stalowych i z metali nieżelaznych. Przedstawiono związki bezwymiarowe określające m.in. zależność składowych siły elektrodynamicznej, mocy czynnej oraz sprawności od podstawowych wymiarów i parametrów mieszadła.

Podano przykłady obliczeniowe. Stwierdzono m.in. duży wpływ parametrów ω_{r1} oraz α_r na wielkość siły elektrodynamicznej oraz mocy czynnej we wlewku.

1. WSTĘP

W hutnictwie, przy ciągłym lub półciągłym odlewaniu wlewków stalowych i z metali nieżelaznych coraz częściej stosuje się mieszanie indukcyjne, bezpośrednio w krystalizatorze (w strefie chłodzenia pierwotnego) lub poza nim (w strefie chłodzenia wtórnego) [1—4]. Przy mieszaniu w krystalizatorze praktycznie cały wlewek ma postać ciekłą, natomiast w strefie chłodzenia wtórnego mieszaniu podlega tylko ciekły rdzeń, a warstwa zewnętrzna wlewka jest już zakrzepnięta. W zależności od kształtu wsadu stosuje się wzbudniki płaskie lub cylindryczne, wytwarzające pole biegnące lub wirujące [5].

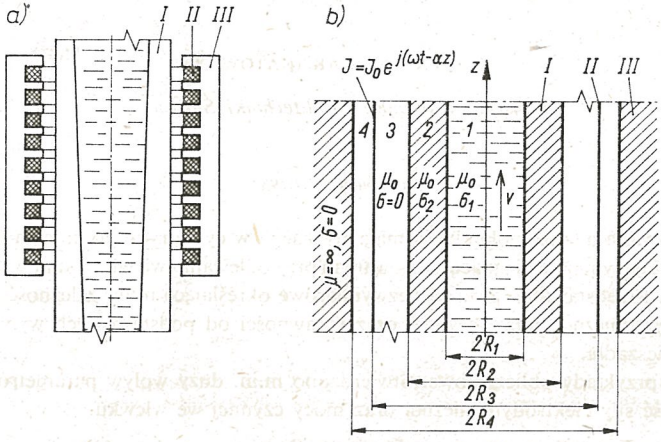
Wprawienie ciekłego metalu w ruch pozwala uzyskać właściwą strukturę wlewka (lepsze własności powierzchni, mniejsza segregacja, równomierne rozmieszczenie wtrąceń), a w rezultacie wyższą jakość wyrobów gotowych. W większym stopniu można również uniezależnić się od zmieniających się warunków technologicznych odlewania: szybkości, temperatury metalu, intensywności chłodzenia pierwotnego i wtórnego [6—8].

Zjawiska towarzyszące mieszaniu elektromagnetycznemu przy odlewaniu są bardzo złożone i stosunkowo mało poznane od strony teoretycznej. Ich kompleksowy opis wymaga rozwiązania zagadnienia elektromagnetycznego, hydrodynamicznego i termokinetycznego, a następnie badań wpływu parametrów odlewania na proces krystalizacji, strukturę i własności wlewka.

Niewątpliwie podstawowe znaczenie dla właściwego przebiegu odlewania mają cechy konstrukcyjne oraz warunki zasilania mieszadła. Chodzi głównie o dobór częstotliwości

prądu, liczby faz i mocy wzbudnika, rodzaju mieszania (w polu wirującym lub biegnącym), miejsca położenia mieszadła oraz długości strefy mieszania.

Analizie pola elektromagnetycznego w mieszadłach poświęcono ostatnio wiele prac. Rozpatruje się w nich zwykle jednowymiarowe przypadki padania cylindrycznej fali biegnącej lub wirującej na powierzchnię jednorodnego, nieskończenie długiego wlewka nieruchomego (np. [9—11]), lub poruszającego się ze stałą prędkością (np. [12, 13]). Znane są również rozwiązania równań pola w prostych modelach mieszadeł z wzbudnikiem o ograniczonej wysokości (np. [14, 15]).



Rys. 1. Mieszadło indukcyjne
a) przekrój podłużny, b) model; I — wlewek, II — uzwojenie, III — rdzeń magnetyczny

Praca jest przyczynkiem do obliczania cylindrycznych mieszadeł elektromagnetycznych z polem biegnącym (rys. 1a), umieszczanych w strefie chłodzenia wtórnego (bez krystalizatora). Przedstawione zależności oraz przykłady liczbowe stanowią rozszerzenie dotychczasowych badań. Mogą być wykorzystane przy projektowaniu wzbudników mieszadeł (dobór konstrukcji i zasilania).

2. MODEL MIESZADŁA

Dla uproszczenia rozwiązania równań pola elektromagnetycznego rzeczywisty układ: mieszadło — wlewek cylindryczny (rys. 1a) zastąpiono modelem pokazanym na rys. 1b. Wlewek reprezentowany jest w nim przez dwuwarstwowy, stały wsad cylindryczny, nieskończenie rozległy wzdłuż osi z. Warstwa wewnętrzna (modelująca ciekły rdzeń) jest obszarem izotropowym o przenikalności magnetycznej $\mu_1 = \mu_0$ i konduktywności σ_1.

Z charakteru pola sił elektrodynamicznych wynika, że w rzeczywistym urządzeniu wektor prędkości mieszanego metalu ma dwie składowe: $\mathbf{v} = v_z(\varrho, z)\mathbf{1}_z + v_\varrho(\varrho, z)\mathbf{1}_\varrho$. W celu ułatwienia analizy i zwiększenia praktycznej przydatności metody obliczeniowej uwzględnia się tutaj tylko ruch metalu w kierunku osi $z(\mathbf{v} = v_z\mathbf{1}_z)$, przy czym zakłada się, że $v_z \neq f(\varrho, z)$. Przyjęcie $v_\varrho = 0$, podobnie jak to uczyniono np. w pracach [9, 11, 12],

jest uzasadnione, ponieważ zwykle $v_z \gg v_\varrho$ [9], szczególnie w obszarze dużych gęstości sił elektrodynamicznych, a więc w warstwach metalu leżących blisko powierzchni zewnętrznej ciekłego rdzenia wlewka. Ocena roli zmiennego rozkładu prędkości wzdłuż promienia wlewka ($v = v_z(\varrho)$) na siły elektrodynamiczne będzie przedmiotem osobnego opracowania.

Pomija się ponadto wpływ wymuszonego mechanicznie ruchu całego wlewka na wielkości pola ($v_w = 0$). Warstwa zewnętrzna, odpowiadająca zakrzepniętej powierzchni wlewka, jest więc nieruchoma. Temperatury warstwy zewnętrznej ($\vartheta_2 \approx$ const) i ciekłego rdzenia ($\vartheta_1 \approx$ const) są różne i dlatego $\sigma_2 \neq \sigma_1$.

Wlewek znajduje się wewnątrz nieskończenie cienkiego i rozległego wzbudnika wytwarzającego biegnące pole magnetyczne. Rozpatruje się tylko podstawową harmoniczną gęstości liniowej prądu wzbudnika

$$J = J_0 \exp[j(\omega t - \alpha z)], \tag{1}$$

gdzie:

$\alpha = \dfrac{\pi}{\tau}$,

J_0 — wartość skuteczna gęstości liniowej prądu,

τ — podziałka biegunowa wzbudnika,

$\omega = 2\pi f$,

f — częstotliwość prądu.

Wówczas potencjał wektorowy

$$A = A_0 \exp[j(\omega t - \alpha z)]. \tag{2}$$

Przyjęcie do rozważań nieskończenie długiego wzbudnika oznacza pominięcie podłużnego efektu brzegowego co oddala znacznie model od rzeczywistego obiektu. W przypadku długich wzbudników (w stosunku do promienia wlewka) mieszadeł indukcyjnych ($L = 0{,}5 \div 1$ m) założenie takie jest jednak dopuszczalne dla metod przybliżonych i stosowane np. w praktyce obliczeniowej urządzeń podobnego typu (pompy indukcyjne) [16].

Pomiędzy wzbudnikiem i wlewkiem (rys. 1a) nie występują ekrany przewodzące, co ma miejsce w mieszadłach stosowanych w strefie chłodzenia pierwotnego. Urządzenia tego typu analizuje się w pracach [17, 18], w których uwzględniono efekt ekranowania ścianek krystalizatora.

3. POTENCJAŁ WEKTOROWY

W środowisku izotropowym, poruszającym się z prędkością v, potencjał wektorowy $\mathbf{A}$ opisany jest następującym równaniem różniczkowym

$$\Delta \mathbf{A} + \mu\sigma(\mathbf{v} \times \operatorname{rot} \mathbf{A}) = \mu\sigma \frac{\partial}{\partial t} \mathbf{A}. \tag{3}$$

Jeżeli uwzględni się, że w analizowanym modelu (rys. 1b) wektory $\mathbf{A}$ i v mają tylko po jednej składowej: $\mathbf{A} = A\mathbf{1}_\Theta$ oraz $\mathbf{v} = v\,\mathbf{1}_z$, wtedy równanie (3) w układzie współrzędnych

walcowych uprości się do

$$\frac{\partial^2 A}{\partial \varrho^2}+\frac{1}{\varrho}\frac{\partial A}{\partial \varrho}+\frac{\partial^2 A}{\partial z^2}-\left(\frac{1}{\varrho^2}+\mu\sigma v\frac{\partial}{\partial z}+\mu\sigma\frac{\partial}{\partial t}\right)A=0.$$

Dla nadania rezultatom obliczeń większej uniwersalności wygodnie jest przedstawić wszystkie parametry modelu mieszadła w postaci bezwymiarowej. Po transformacji

$$\varrho_r=\frac{\varrho}{R_3};\quad z_r=\frac{z}{R_3};\quad \tau_r=\frac{\tau}{R_3};\quad \alpha_r=\alpha R_3;$$

$$\omega_r=\omega\mu_i\sigma_i R_3^2;\quad R_{r1}=\frac{R_1}{R_3};\quad R_{r2}=\frac{R_2}{R_3};\quad R_{r4}=\frac{R_4}{R_3};$$

$$A_r=\frac{A_0}{\mu_0 J_0 R_3};\quad t_r=\omega t$$

oraz wykorzystaniu zależności (2), otrzymuje się:

$$\frac{\partial^2 A_r}{\partial \varrho_r^2}+\frac{1}{\varrho_r}\frac{\partial A_r}{\partial \varrho_r}-\left(\frac{1}{\varrho_r^2}+\beta_r^2\right)A_r=0, \tag{4}$$

gdzie:

$\beta_r=\sqrt{\alpha_r^2+j\omega_r s}$, s — poślizg, $s=1-\frac{v}{v_s}$, v_s — prędkość synchroniczna pola biegnącego, $v_s=2f\tau$.

Rozwiązaniem ogólnym równania (4) w i-tym obszarze obliczeniowym ($i=1, 2, 3, 4$) jest kombinacja liniowa zmodyfikowanych funkcji Bessela

$$A_{ri}(\varrho_r, z_r, t_r)=[C_i I_1(\beta_{ri}\varrho_r)+D_i K_1(\beta_{ri}\varrho_r)]\exp[j(t_r-\alpha_r z_r)]. \tag{5}$$

Stałe całkowania C_i i D_i wyznacza się z warunków brzegowych dla potencjału wektorowego na granicach obszarów i. Pełną postać układu równań określających stałe C_i i D_i podano w tabl. 1. W szczególności

$$C_1=\frac{1}{R_{r1}R_{r2}}\frac{M_1}{M},$$

$$D_1=0,$$

$$C_2=\frac{1}{R_{r2}}\frac{M_1 M_4}{M},$$

$$D_2=\frac{1}{R_{r2}}\frac{M_1 M_3}{M},$$

przy czym

$$M_1=I_1(\alpha_r)K_0(\alpha_r R_{r4})+K_1(\alpha_r)I_0(\alpha_r R_{r4}),$$

$$M_3=\beta_{r2}I_0(\beta_{r2}R_{r1})I_1(\beta_{r1}R_{r1})-\beta_{r1}I_1(\beta_{r2}R_{r1})I_0(\beta_{r1}R_{r1}),$$

$$M_4=\beta_{r1}K_1(\beta_{r2}R_{r1})I_0(\beta_{r1}R_{r1})+\beta_{r2}K_0(\beta_{r2}R_{r1})I_1(\beta_{r1}R_{r1}),$$

$$M_5=M_4 I_1(\beta_{r2}R_{r2})+M_3 K_1(\beta_{r2}R_{r2}),$$

$$M_6=M_4 I_0(\beta_{r2}R_{r2})-M_3 K_0(\beta_{r2}R_{r2}),$$

Tablica 1

Układ równań określający stałe całkowania C_i i D_i

$$
\begin{vmatrix}
I_1(\beta_{r1}R_{r1}) & -I_1(\beta_{r2}R_{r1}) & -K_1(\beta_{r2}R_{r1}) & 0 & 0 & 0 & 0 \\
\beta_{r1}I_0(\beta_{r1}R_{r1}) & -\beta_{r2}I_0(\beta_{r2}R_{r1}) & \beta_{r2}K_0(\beta_{r2}R_{r1}) & 0 & 0 & 0 & 0 \\
0 & I_1(\beta_{r2}R_{r2}) & K_1(\beta_{r2}R_{r2}) & -I_1(\alpha_r R_{r2}) & -K_1(\alpha_r R_{r2}) & 0 & 0 \\
0 & \beta_{r2}I_0(\beta_{r2}R_{r2}) & -\beta_{r2}K_0(\beta_{r2}R_{r2}) & -\alpha_r I_0(\alpha_r R_{r2}) & \alpha_r K_0(\alpha_r R_{r2}) & 0 & 0 \\
0 & 0 & 0 & I_1(\alpha_r) & K_1(\alpha_r) & -I_1(\alpha_r) & -K_1(\alpha_r) \\
0 & 0 & 0 & \alpha_r I_0(\alpha_r) & -\alpha_r K_0(\alpha_r) & -\alpha_r I_0(\alpha_r) & \alpha_r K_0(\alpha_r) \\
0 & 0 & 0 & 0 & 0 & I_0(\alpha_r R_{r4}) & -K_0(\alpha_r R_{r4})
\end{vmatrix}
\cdot
\begin{vmatrix} C_1 \\ C_2 \\ D_2 \\ C_3 \\ D_3 \\ C_4 \\ D_4 \end{vmatrix}
=
\begin{vmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{vmatrix}
$$

$$M_7 = \alpha_r M_5 I_0(\alpha_r R_{r2}) - \beta_{r2} M_6 I_1(\alpha_r R_{r2}),$$
$$M_8 = \alpha_r M_5 K_0(\alpha_r R_{r2}) + \beta_{r2} M_6 K_1(\alpha_r R_{r2}),$$
$$M = M_8 I_0(\alpha_r R_{r4}) - M_7 K_0(\alpha_r R_{r4}),$$
$$\beta_{r1} = \sqrt{\alpha_r^2 + j\omega_{r1} s},$$
$$\beta_{r2} = \sqrt{\alpha_r^2 + j\omega_{r2}},$$
$$\omega_{r1} = \omega\mu_0\sigma_1 R_3^2,$$
$$\omega_{r2} = \omega\mu_0\sigma_2 R_3^2.$$

Po wprowadzeniu konduktywności względnej $\sigma_{r2} = \frac{\sigma_2}{\sigma_1}$ do parametru ω_{r2} otrzymuje się

$$\omega_{r2} = \omega_{r1}\sigma_{r2},$$

a wtedy

$$\beta_{r2} = \sqrt{\alpha_r^2 + j\omega_{r1}\sigma_{r2}}.$$

4. SIŁA ELEKTRODYNAMICZNA, MOC CZYNNA I SPRAWNOŚĆ

Dla przebiegów sinusoidalnie zmiennych średnia za okres wartość wektora gęstości objętościowej siły elektrodynamicznej (N/m³), działającej na ruchome środowisko przewodzące, wynosi

$$\mathbf{f} = \frac{1}{2}\operatorname{Re}\mathbf{J}_m \times \mathbf{B}_m^*,$$

gdzie:

J — wektor gęstości prądu indukowanego

$$\mathbf{J} = \sigma\left[-\frac{\partial \mathbf{A}}{\partial t} + \mathbf{v} \times \operatorname{rot}\mathbf{A}\right],$$

B* — wektor zespolony sprzężony indukcji magnetycznej.

W rozpatrywanym przypadku

$$\mathbf{J} = J\mathbf{1}_\Theta = -\sigma\left[\frac{\partial A}{\partial t} + v\frac{\partial A}{\partial z}\right]\mathbf{1}_\Theta,$$
$$\mathbf{B} = B_\varrho \mathbf{1}_\varrho + B_z \mathbf{1}_z,$$

przy czym:

$$B_\varrho = -\frac{\partial A}{\partial z}, \qquad B_z = \frac{1}{\varrho}\frac{\partial(\varrho A)}{\partial \varrho}.$$

Tak więc

$$\mathbf{f} = f_\varrho \mathbf{1}_\varrho + f_z \mathbf{1}_z.$$

W ciekłym metalu (obszar 1)

$$f_{\varrho 1} = \mathrm{Re} J_1(\varrho, z) B^*_{z1}(\varrho, z),$$
$$f_{z1} = -\mathrm{Re} J_1(\varrho, z) B^*_{\varrho 1}(\varrho, z).$$

W postaci bezwymiarowej

$$J_r = \frac{R_3}{J_0} J,$$

$$B_{r(\varrho, z)} = \frac{1}{\mu_0 J_0} B_{(\varrho, z)},$$

$$f_{r(\varrho, z)} = \frac{R_3}{\mu_0 J_0^2} f_{(\varrho, z)}.$$

Ostatecznie, składowe wektora gęstości objętościowej siły wynoszą

$$f_{r\varrho 1} = 2\omega_{r1} s \left(\frac{M_1}{R_{r1} R_{r2} |M|} \right)^2 \mathrm{Re}[-j\beta^*_{r1} I_1(\beta_{r1} \varrho_r) I_0(\beta^*_{r1} \varrho_r)], \tag{6}$$

$$f_{rz1} = 2\omega_{r1} s \alpha_r \left(\frac{M_1}{R_{r1} R_{r2} |M|} \right)^2 |I_1(\beta_{r1} \varrho_r)|^2. \tag{7}$$

Zależności (6) i (7) są podstawą do orientacyjnego (ze względu na duży stopień uproszczenia modelu obliczeniowego) określenia optymalnych parametrów konstrukcyjnych oraz zasilania mieszadła. Mogą być również wykorzystane do rozwiązania równania Naviera-Stokesa i znalezienia rozkładu prędkości ciekłego metalu w rdzeniu wlewka.

Celowe jest obliczenie składowej całkowitej siły elektrodynamicznej, wywołującej ruch metalu w kierunku osi z

$$F_{z1} = \int_V f_{z1} dV.$$

Wartość siły F_{z1} działającej na odcinek wlewka równy długości rzeczywistego wzbudnika $L = 2p_b\tau$, gdzie: p_b — liczba par biegunów wzbudnika, wynosi

$$F_{z1} = \int_0^{R_1} \int_0^{2\pi} \int_0^{L} f_{z1} \varrho dz d\Theta d\varrho = 2\pi L \int_0^{R_1} f_{z1} \varrho d\varrho. \tag{8}$$

Ostatecznie, po wprowadzeniu wielkości bezwymiarowych

$$F_{rz1} = \frac{F_{z1}}{L R_3 \mu_0 J_0^2} = \frac{2\pi\alpha_r}{R_{r1}} \left(\frac{M_1}{R_{r2}|M|} \right)^2 \mathrm{Im}[\beta_{r1} I_0(\beta_{r1} R_{r1}) I_1(\beta^*_{r1} R_{r1})].$$

Podobną postać do zależności (8) ma składowa siły elektrodynamicznej w osi ϱ

$$F_{\varrho 1} = 2\pi L \int_0^{R_1} f_{\varrho 1} \varrho d\varrho.$$

Moc czynna wydzielana w ciekłym rdzeniu wlewka związana jest z siłą elektrodynamiczną F_{z1} następująco

$$P_1 = \frac{\omega s}{\alpha} F_{z1} + v F_{z1} = P_{e1} + P_{m1}. \tag{9}$$

Moc P_{e1} reprezentuje w równaniu (9) straty mocy w ciekłym metalu (ciepło Joule'a), natomiast P_{m1} — moc związaną z ruchem metalu (moc mechaniczna).

Część mocy przekazywanej od wzbudnika do wlewka wydziela się w jego zewnętrznej zakrzepniętej warstwie (obszar 2)

$$P_{e2} = \int_V p_{v2}\,dV = \int_{R_1}^{R_2}\int_0^{2\pi}\int_0^{L} p_{v2}\,\varrho\,dz\,d\Theta\,d\varrho = \frac{2\pi l}{\sigma_2}\int_{R_1}^{R_2} |J_2|^2\varrho\,d\varrho,$$

gdzie p_{v2} — gęstość objętościowa mocy czynnej.

W postaci bezwymiarowej

$$P_{er(1,2)} = \frac{\sigma_{(1,2)}}{J_0^2 L} P_{e(1,2)}.$$

Sprawność elektryczną i mechaniczną mieszadła indukcyjnego, przy pominięciu strat mocy w uzwojeniu i rdzeniu magnetycznym wzbudnika, określa się następująco

$$\eta_e = \frac{P_1}{P_1 + P_{e2}} = \frac{\frac{\omega_{r1}}{\alpha_r} F_{rz1}}{\frac{\omega_{r1}}{\alpha_r} F_{rz1} + \frac{1}{\sigma_{r2}} P_{er2}},$$

$$\eta_m = (1-s)\eta_e.$$

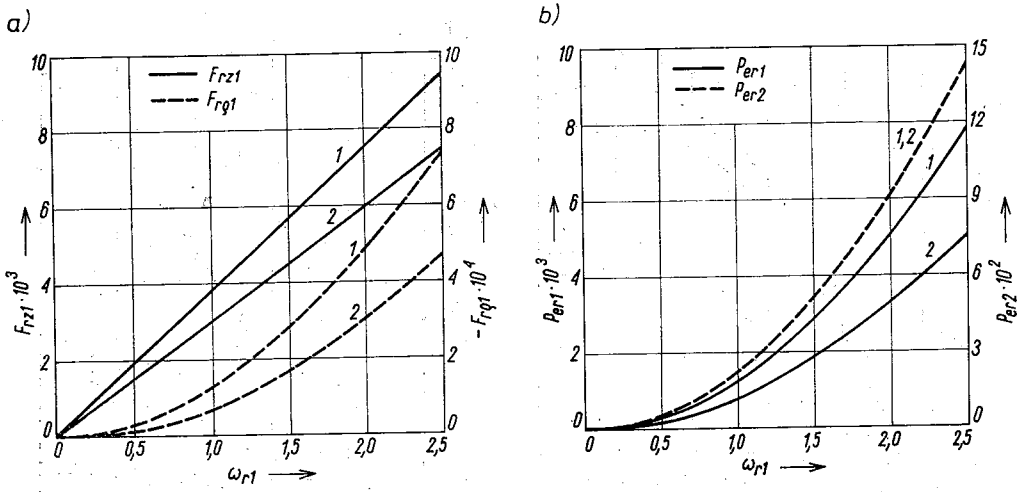
5. PRZYKŁADY OBLICZENIOWE

Dla zilustrowania możliwości wykorzystania zależności z p. 4 przeprowadzono przykładowe obliczenia dla parametrów bezwymiarowych zestawionych w tabl. 2. Wyniki przedstawiono w formie wykresów na rys. 2÷7. Oznaczenia krzywych są zgodne z numeracją wariantów danych obliczeniowych w tabl. 2. Wartości, dla których wyznaczono krzywe *1* i *2* z rys. 2÷4 mogą odpowiadać np. mieszadłu o podziałce biegunowej $\tau = 0{,}1$

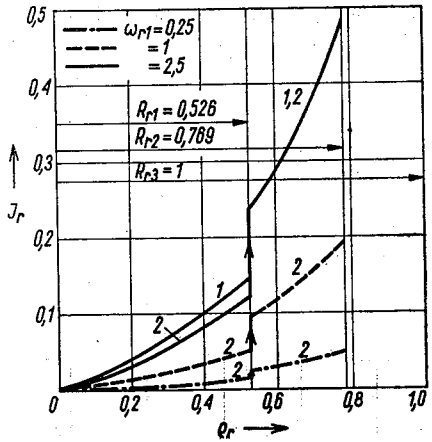
Tablica 2

Dane obliczeniowe mieszadeł indukcyjnych

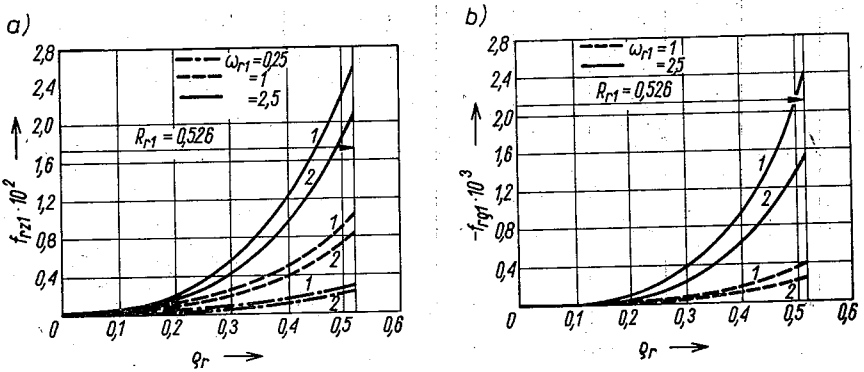
Nr krzywej	R_{r1}	R_{r2}	R_{r4}	α_r	σ_{r2}	s
1	0,526	0,789	1,105	2,98	1,57	1
2	0,526	0,789	1,105	2,98	1,57	0,8
3	0,414	0,586	1,138	2,28	3,59	1
4	0,414	0,586	1,138	2,28	3,59	0,9
5	0,414	0,586	1,138	1,52	3,59	1
6	0,414	0,586	1,138	1,52	3,59	0,9
7	0,276	0,586	1,138	2,28	3,59	1
8	0,276	0,586	1,138	2,28	3,59	0,9
9	0,276	0,586	1,138	1,52	3,59	1
10	0,276	0,586	1,138	1,52	3,59	0,9



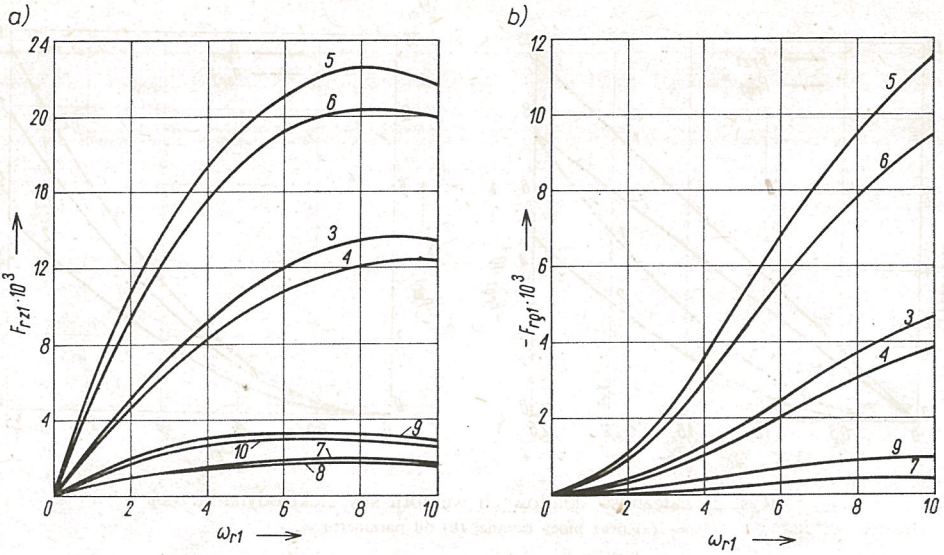
Rys. 2. Zależność składowych wektora siły elektrodynamicznej
(a) oraz mocy czynnej (b) od parametru ω_{r1}



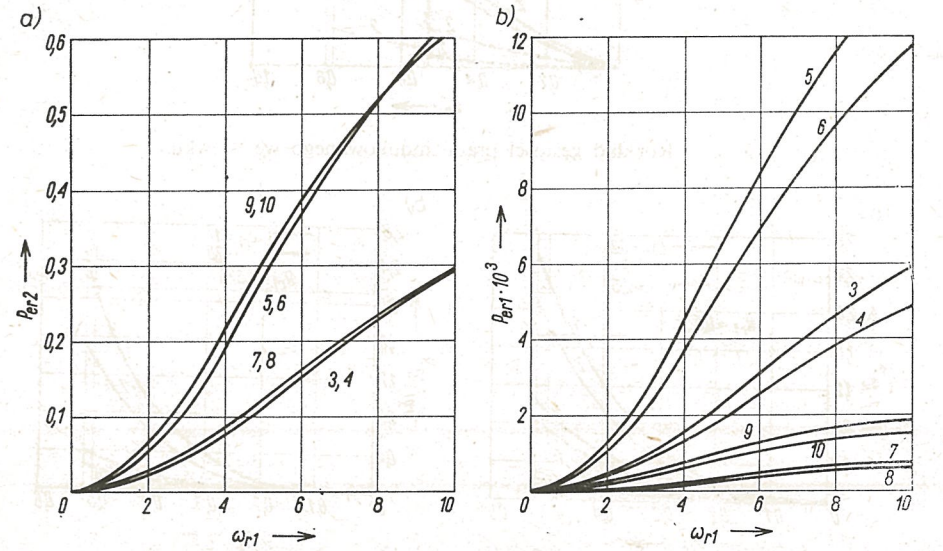
Rys. 3. Rozkład gęstości prądu indukowanego we wlewku



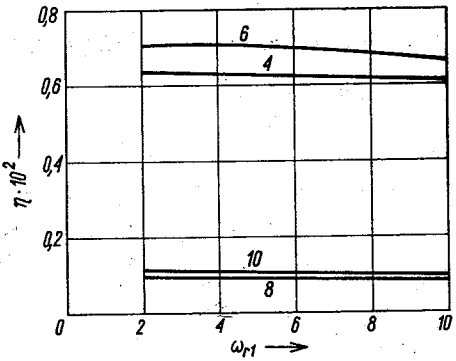
Rys. 4. Rozkłady składowych wektora gęstości siły elektrodynamicznej w obszarze ciekłym wlewka



Rys. 5. Zależność składowych wektora siły elektrodynamicznej od parametru ω_{r1}



Rys. 6. Zależność mocy czynnej w warstwie zakrzepniętej (a) i ciekłej (b) wlewka od parametru ω_{r1}



Rys. 7. Zależność sprawności mechanicznej mieszania indukcyjnego od parametru ω_{r1} przy stałym poślizgu s

m, do wlewków stalowych o średnicy $2R_2 = 0{,}15$ m; natomiast dane dla charakterystyk $3 \div 10$ — np. mieszadłu o podziałkach biegunowych $\tau = 0{,}1$ i $0{,}15$ m, do wlewków z metali nieżelaznych o średnicy $2R_2 = 0{,}085$ m. Obliczenia wykonano dla dużych poślizgów $s = 0{,}8 \div 1$, gdyż z takimi ma się do czynienia przy mieszaniu indukcyjnym ciekłych metali.

Dla projektowania mieszadeł indukcyjnych szczególnie interesujące są charakterystyki F_{rz1}, $F_{r\varrho 1} = f(\omega_{r1})$ (rys. 2 i 5) oraz f_{rz1}, $f_{r\varrho 1} = f(\varrho_r)$ (rys. 4). Widoczna jest bardzo duża zależność składowych wektora siły elektrodynamicznej od parametru ω_{r1}. Dla dużych wartości ω_{r1} (rys. 5) i małej grubości warstwy zakrzepniętej (krzywe $3 \div 6$) występuje dosyć wyraźne maksimum funkcji $F_{rz1} = f(\omega_{r1})$. Wzrost grubości warstwy zewnętrznej wlewka powoduje zmniejszenie siły elektrodynamicznej działającej na ciekły metal (krzywe $7 \div 10$) oraz poszerzenie zakresu optymalnych wartości ω_{r1}. Widoczna jest również duża zależność F_{rz1} i $F_{r\varrho 1}$ od podziałki biegunowej uzwojenia wzbudnika (np. krzywe 3 i 5).

Dla oceny stopnia oddziaływania wlewka na pole elektromagnetyczne wytwarzane przez wzbudnik oraz wpływu ekranowania zakrzepniętej warstwy zewnętrznej wlewka określono współczynniki k_z i k_ϱ, zdefiniowane następująco

$$k_z = \left| \frac{B_{rz1}}{B^0_{rz}} \right|_{\varrho_r = R_{r1}},$$

$$k_\varrho = \left| \frac{B_{r\varrho 1}}{B^0_{r\varrho}} \right|_{\varrho_r = R_{r1}},$$

gdzie B^0_{rz} i $B^0_{r\varrho}$ są składowymi wektora indukcji magnetycznej dla pustego wzbudnika (bez wlewka)

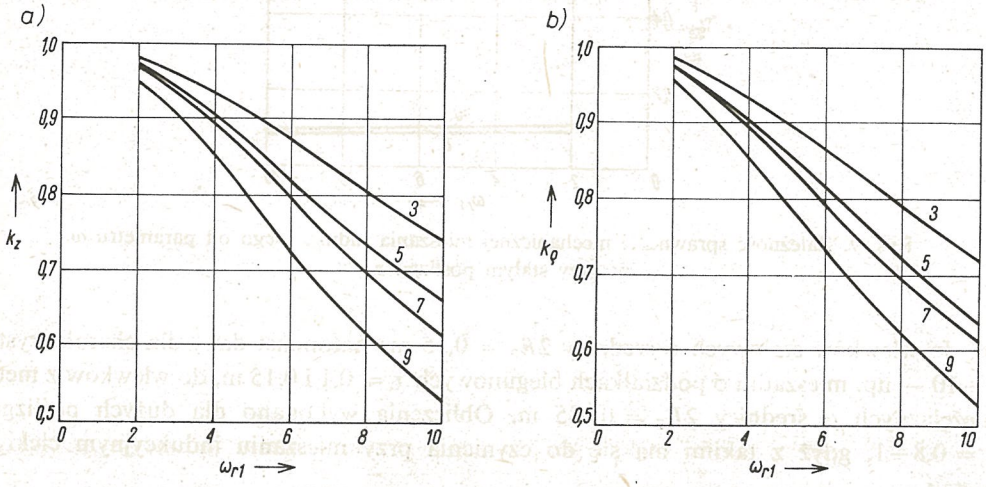
$$B^0_{rz} = \frac{\alpha_r X}{I_o(\alpha_r R_{r4})} I_0(\alpha_r \varrho_r),$$

$$B^0_{r\varrho} = \frac{j\alpha_r X}{I_0(\alpha_r R_{r4})} I_1(\alpha_r \varrho_r),$$

przy czym

$$X = I_1(\alpha_r)K_0(\alpha_r R_{r4}) + K_1(\alpha_r)I_0(\alpha_r R_{r4}).$$

Zależności k_z, $k_\varrho = f(\omega_{r1})$ pokazano na rys. 8.



Rys. 8. Zależność współczynników k_z i k_ϱ od parametru ω_{r1}

6. PODSUMOWANIE

W pracy rozwiązano zagadnienie elektromagnetyczne w cylindrycznym mieszadle indukcyjnym, stosowanym przy ciągłym i półciągłym odlewaniu stali i metali nieżelaznych. Mimo daleko idącego uproszczenia modelu obliczeniowego uwzględniono w nim zasadnicze wymiary i parametry wzbudnika oraz wlewka.

Przedstawione zależności w formie bezwymiarowej mogą być podstawą do skonstruowania charakterystyk uniwersalnych, przydatnych wprost do projektowania mieszadeł analizowanego typu. Podane przykłady obliczeniowe wskazują na możliwość właściwego doboru parametrów konstrukcyjnych i zasilania wzbudnika (przede wszystkim podziałki biegunowej i częstotliwości prądu) dla uzyskania optymalnych warunków pracy mieszadła.

BIBLIOGRAFIA

1. K. Schwerdtfeger, *Metallurgische Probleme beim Stranggiessen von Stahl*, Stahl u. Eisen, t. 98, 1978, nr 6, ss. 225—235.
2. B. Hanas, *Improving metallurgical processes by inductive stirring*, 9th International Congress UIE, Cannes, 1980.
3. S. Kollberg, *Electromagnetic stirrers in the secondary cooling of continuous casting machines*, 9th International Congress UIE, Cannes, 1980.
4. B. A. Sokunow, i inni, *Wlijanije elektromagnitnogo pieriemieszziwanija na strukturu slitkow pri poluniepriery wnoj otliwkie*, Magnitnaja gidrodinamika, nr 3, 1977, ss. 127—130.
5. H. S. Marr, *Electromagnetic strirring: stepping stone to improved continuously cast products*, Iron and Steel International, February 1979, ss. 29—41.

6. N. Delassus i inni, *Elektromagnetische Rührwerke in der sekundären Kühlung beim Strangguss von Knüppeln und Vorblöcken*, Fachber. Hüttenpraxis, 1977, nr 10, ss. 838—851.
7. H. Takeuchi i inni, *Quality improvement of continuously cast stainless steel blooms through electromagnetic stirring*, Trans. Iron Steel Inst. Jap., t. 18, 1978, nr 6, ss. 352—360.
8. R. Alberny, J. P. Birat, *Brassage électromagnetique et qualité des produits de coulée continue*, Circ. Inf. Techn., t. 34, 1977, nr 4, ss. 925—944.
9. J. Szekely, C. W. Chang, *Turbulent electromagnetically driven flow in metals processing*, Ironmaking and Steelmaking, 1977, nr 3, ss. 190—204.
10. J. Szekely, *Fluid flow phenomena in metals processing*, Academic Press, 1979, New York.
11. K. Nakanishi i in., *Possible relationship between energy dissipation and agitation in steel processing operation*, Ironmaking and Steelmaking, 1975, nr 3, ss. 193—197.
12. J. Szekely, S. Asai, Transactions ISIJ, 1975, nr 15, ss. 270—275.
13. K. H. Tacke, K. Schwerdtfeger, *Rührgeschwindigkeit in Rundstrangguss bei Rührung mit elektromagnetischen Drehfeldern*, Stahl u. Eisen, t. 99, 1979, nr 1, ss. 7—12.
14. M. J. Stołow, *Elektromagnitnoje pole i ponderomotornyje siły w ustrojstwach elektromagnitnogo pieriemieszywanija cilindriczieskogo tipa*, Issledowanija w obłasti promyszlennogo elektronagriewa, wyp. 4, Moskwa 1970, ss. 192—199.
15. F. Fikus i in., *Obliczanie pola elektromagnetycznego i sił elektrodynamicznych w cylindrycznych mieszadłach do ciekłych metali*, Archiwum Elektrotechniki, t. XXVIII, z. 1, 1979, ss. 189—201.
16. J. K. Krumin, *Wzaimodiejstwije bieguszcziego magnitnogo polja s prowodjaszczej sredoj*, Zinatne, Riga 1969.
17. W. Haubitzer, *Das elektromagnetische Feld, die Druck- und Geschwindigkeitsverteilung im Rundstrangguss bei Drehfeld erregten Rührspulen*, Elektrowärme International, t. 37, 1979, B. 6, ss. B297—B300.
18. Cz. Sajdak, *The calculation of electromagnetic forces acting upon a billet stirred inductively in the process of casting*, Revue Roum. Sci. Techn.-Electrotechn. et Energ. (w druku).

C. SAJDAK

ELECTROMAGNETIC FORCES IN LIQUID METAL STIRRED INDUCTIVELY IN THE PROCESS OF SEMI-CONTINUOUS AND CONTINUOUS CASTING OF INGOTS

Summary

The paper presents an analysis of the electromagnetic field in a cylindrical induction stirrer which produces a traveling magnetic field and can be applied to casting steel and non-ferrous ingots. The presentation focuses on dimensionless relationships determining the dependence of the components of the electromagnetic force, active power and efficiency on the overall dimensions and parameters of the stirrer. A calculation example is given. For instance, a significant influence of the parameters ω_{r1} and α_r on the electrodynamic force and the active power in the ingot are noted.

C. SAJDAK

LES FORCES ÉLECTRODYNAMIQUES DANS LE MÉTAL LIQUIDE MALAXÉ PAR MÉTHODE À INDUCTION DANS LE PROCESSUS DE LA COULÉE SEMI-CONTINUE ET CONTINUE DES LINGOTS CYLINDRIQUES

Résumé

L'étude contient l'analyse du champ électromagnétique de l'agitateur cylindrique à induction, produisant le champ magnétique courant, dans la coulée des lingots en acier et en métaux non ferreux. On a présenté les rapports, non dimensionnels, qui définissent la dépendence des composantes de la force électrodynamique, de la puissance active et de l'efficacité, suivant le mesurage de base et les paramètres de l'agitateur. On a indiqué des exemples des calculs. On a constaté, entre autres, une grande influence des paramètres ω_{r1} et α_r sur la force électrodynamique et sur la puissance active dans le lingot.

C. SAJDAK

ELEKTRODYNAMISCHE KRÄFTE IM INDUKTIVEN GEMISCHTEN FLÜSSIGEN METALL IN HALB- BZW. STRANGGIEßVERFAHREN VON ZYLINDRISCHEN BLÖCKEN

Zusammenfassung

Die Bearbeitung enthält eine Analyse des elektromagnetischen Feldes im zylindrischen induktiven Rührwerk mit Wanderfeld, das beim Gießen von Stahlblöcken und Zainen angewandt wird. Es wurden dabei die dimensionalen Zusammenhänge dargestellt, die u.a. die Abhängigkeit der Komponenten der elektrodynamischen Kraft, der Wirkleistung und des Wirkungsgrades von den Grundabmessungen sowie den Parametern des Rührwerks bestimmen. Berechnungsbeispiele wurden angegeben. Unter anderem wurde der große Einfluß der Parameter ω_{r1} und α_r auf die Größe der elektrodynamischen Kraft sowie der Wirkungsleistung im Block festgestellt.

Ч. САЙДАК

ЭЛЕКТРОДИНАМИЧЕСКИЕ СИЛЫ В ЖИДКОМ МЕТАЛЛЕ ИНДУКЦИОННО ПЕРЕМЕШИВАЕМОМ В ПРОЦЕССАХ ПОЛУНЕПРЕРЫВНОГО И НЕПРЕРЫВНОГО ЛИТЬЯ ЦИЛИНДРИЧЕСКИХ СЛИТКОВ

Резюме

Приведен анализ электромагнитного поля в цилиндрической индукционной мешалке, производящей бегущее магнитное поле, применяемой при литье слитков стальных и из цветных металлов. Представлены безразмерные соотношения определяющие зависимость составляющих электродинамической силы, активной мощности а также к.п.д. от основных размеров и параметров мешалки. Поданы расчетные примеры. Констатировано большое влияние параметров ω_{r1} а также α_r на величину электродинамической силы и активной мощности в слитке.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 437—451

621.3.014.3

Wybrane zagadnienia analizy mikroskopowej śladów łuku na katodzie

ANDRZEJ SOBIESZCZUK (LUBLIN)

Instytut Przetwarzania i Użytkowania Energii Elektrycznej Politechniki Lubelskiej

Otrzymano 30.9.1983

Mikroskopowa analiza śladów łuku na katodzie od wielu lat była wykorzystywana do wyznaczania gęstości prądu w plamce katodowej i analizy zjawisk zachodzących na katodzie łuku elektrycznego. Autor, w oparciu o własne mikroskopowe badania śladów łuku na katodzie oraz dane zaczerpnięte z literatury, podejmuje próbę szacowania parametrów plamek elementarnych, szczególnie w zakresie średniej gęstości prądu w plamce elementarnej. W analizie mikroskopowej śladów łuku na katodzie zwrócono uwagę na trudności i wątpliwości w jednoznacznej ich interpretacji oraz na szereg dotychczas niedostatecznie uwzględnianych zjawisk występujących przy ich powstawaniu i zanikaniu.

1. WSTĘP

Większość informacji o strukturze budowy plamki katodowej łuku elektrycznego dostarczają obserwacje jej śladów na katodzie. Ślad pozostawiony przez plamkę na powierzchni katody od dawna wykorzystywany był do analizy zjawisk zachodzących w katodowej stopie łuku. Podczas obserwacji i pomiarów śladu łuku nie ma możliwości analizowania rzeczywistej powierzchni plamki, lecz jedynie zniszczenia powierzchni katody dokonane przez łuk. Powiększenia śladu łuku nieruchomego rzędu kilkunastu, kilkudziesięciu razy pozwalają najczęściej stwierdzić, że ślad łuku charakteryzuje się zwartą budową. Początkowo przyjmowano, że plamka katodowa jest jednolita w budowie. Wnioski, jakie wyciągano z tych obserwacji, odpowiadały teorii budowy plamki katodowej opartej na ciągłym w czasie przepływie prądu w katodowej stopie łuku i są one powszechnie znane z literatury.

Badania śladów łuku ruchomego oraz wprowadzenie pojęcia plamki cząstkowej i grupowej stworzyły nowy jakościowo obraz śladu łuku. Ślad łuku na katodzie oraz plamka katodowa przestają być traktowane jako jednolite [3, 16].

W śladach łuku, zwłaszcza ruchomego, obserwuje się niejednorodności, rozszczepienia na szereg gałęzi równoległych, obszary nieciągłości oraz odrębne zgrupowania śladu o nieregularnych kształtach. Powstają trudności w jednoznacznym określaniu granicy plamki

katodowej jako całości, a następnie granic plamek cząstkowych lub grupowych. Bardzo wygodne i często nawet nieodzowne jest przyjmowanie granic nadtopień (chociaż nie zawsze ścisłe i możliwe) jako granicy plamki katodowej i plamek cząstkowych [3, 23, 25, 26, 33].

Bezpośrednie obserwacje śladów łuku pod mikroskopem przy powiększeniach rzędu 100, 1000 razy pozwoliły stwierdzić, że ślad łuku zawiera szereg mikrowgłębień w kształcie kraterów.

Zaproponowany w [4] impulsowy model katodowej stopy łuku uwzględniający cząstkową budowę plamki katodowej oraz wprowadzenie pojęcia plamki elementarnej, której ślady przyjęto identyfikować z zaobserwowanymi kraterami na powierzchni katody, zmieniły zakres prowadzonych badań.

Zastosowanie do badań na początku lat siedemdziesiątych mikroskopu skaningowego rozszerzyło możliwości analizowania śladów łuku poprzez stosowanie znacznie większych powiększeń, rzędu 1000, 10000 i więcej razy. Uzyskane znacznie lepszej jakości zdjęcia śladów łuku w porównaniu ze zdjęciami otrzymywanymi ze zwykłych mikroskopów optycznych dostarczyły dodatkowych informacji, dotychczas nie wystarczająco uwzględnianych w analizie śladów łuku.

Obserwowany ślad łuku jest stanem końcowym palenia się łuku i stanowi zbiór kraterów wzajemnie się nakładających, powstałych w wyniku przypadkowego powstawania i zanikania plamek elementarnych. Różnorodność średnic kraterów i ich głębokości związane są ze: stanem powierzchni (głębokością i rodzajem tlenków, stopniem gładkości), czasem palenia się łuku, rodzajem materiału katody (temperaturą topnienia, własnościami cieplno-fizycznymi), wartością prądu łuku, szybkością poruszania się łuku, wartością prądu płynącego przez plamkę elementarną oraz jego gęstością.

W śladach łuku sfotografowanych mikroskopem skaningowym obserwujemy oprócz śladów w postaci kraterów również struktury krystaliczne, występujące przy krzepnięciu stopionego metalu. Przypadkowe pojawianie się i znikanie plamek elementarnych powoduje powstanie w katodzie niejednorodnego pola temperaturowego, które po zgaśnięciu łuku kształtuje przebieg procesów krystalizacyjnych na powierzchni katody. Na przebieg tych procesów mają wpływ następujące czynniki [36, 37]:

a) stopień przechłodzenia ciekłego metalu,
b) stopień czystości,
c) nierozpuszczone resztki krystaliczne zawarte w ciekłym metalu w temperaturze krzepnięcia,
d) ciśnienie zewnętrzne,
e) wstrząsy mechaniczne,
f) warunki krzepnięcia.

Od stopnia przechłodzenia zależy przede wszystkim liczba zarodków krystalizacji i liniowa szybkość krystalizacji. Obie te wielkości mają decydujący wpływ na wielkość ziarna.

Napięcie powierzchniowe ciekłego metalu przeciwstawia się siłom skupiającym atomy w rozrastających się kryształach, które sprzyjają powstawaniu płaskich powierzchni i ostrych krawędzi. Jeżeli napięcie powierzchniowe ma wartość wyższą od sił skupiających, to powstałe wtedy kryształy przyjmują kształt bardziej zaokrąglony.

Zanieczyszczenia w metalach mogą być dwojakiego rodzaju: metaliczne i niemetalicz-

ne. Pierwsze z nich rozpuszczają się w ciekłych metalach a nawet bardzo często w stałych i nieznacznie wpływają na krystalizację i wielkość ziarna. Zanieczyszczenia niemetaliczne, które istnieją w ciekłym metalu powiększają podczas krzepnięcia liczbę zarodków krystalicznych i w ten sposób przyczyniają się do rozdrabniania budowy. Zanieczyszczenia metaliczne i niemetaliczne wpływają na nieregularność kryształów oraz tworzenie się pęcherzyków gazowych, których ślady widoczne są w śladach łuku.

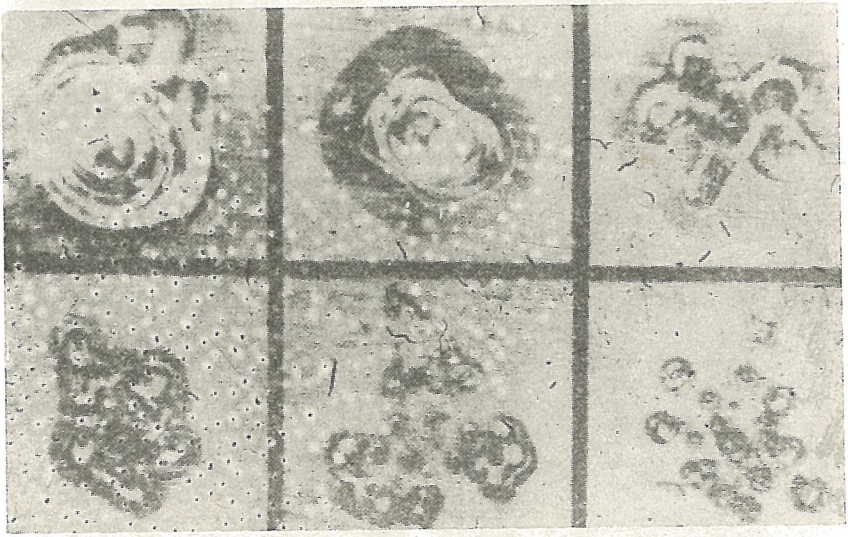
Wstrząsy metali ciekłych, poddanych krzepnięciu sprzyjają wytworzeniu się licznych zarodków krystalicznych, a więc struktury bardziej drobnoziarnistej. Jednocześnie następuje usuwanie z cieczy gazów i wtrąceń niemetalicznych.

Przebieg krystalizacji ma decydujący wpływ na wielkość, kształt i orientację krystalograficzną kryształów. Procesy krystalizacyjne zacierają ślady wytworzone przez plamki cząstkowe i elementarne, deformują więc ostateczny obraz śladu łuku, który dopiero po zakończeniu tych procesów podlega badaniom.

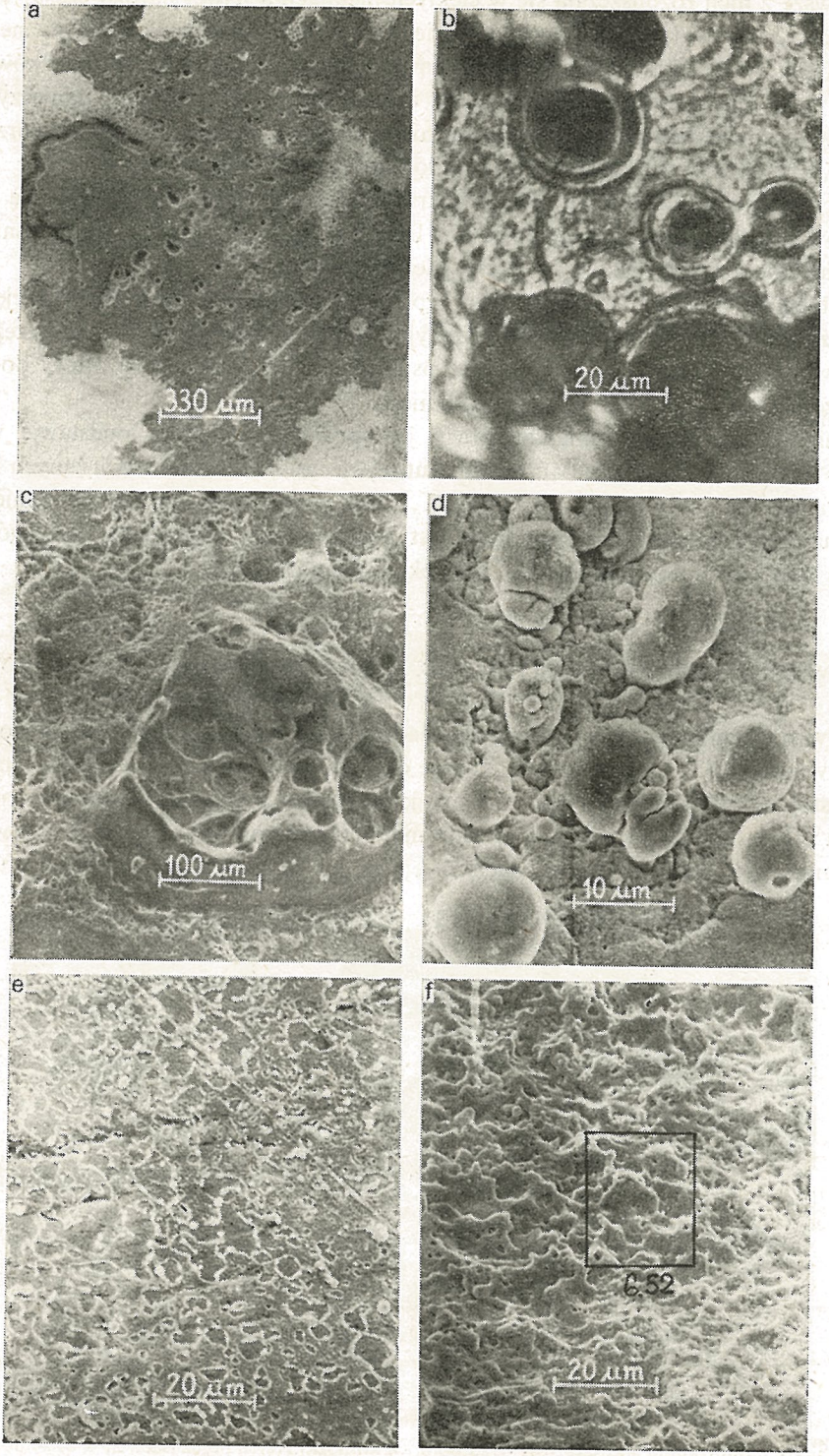
Zamierzeniem autora niniejszej pracy jest przedstawienie, na podstawie własnych badań i znanych wyników badań innych autorów, ogółu zagadnień związanych z mikroskopową analizą śladów łuku na katodzie. Autor wskazuje również na istnienie problemów, na które w dalszym ciągu brak jest jednoznacznej odpowiedzi, szczególnie gdy dotyczą one określenia parametrów plamek elementarnych.

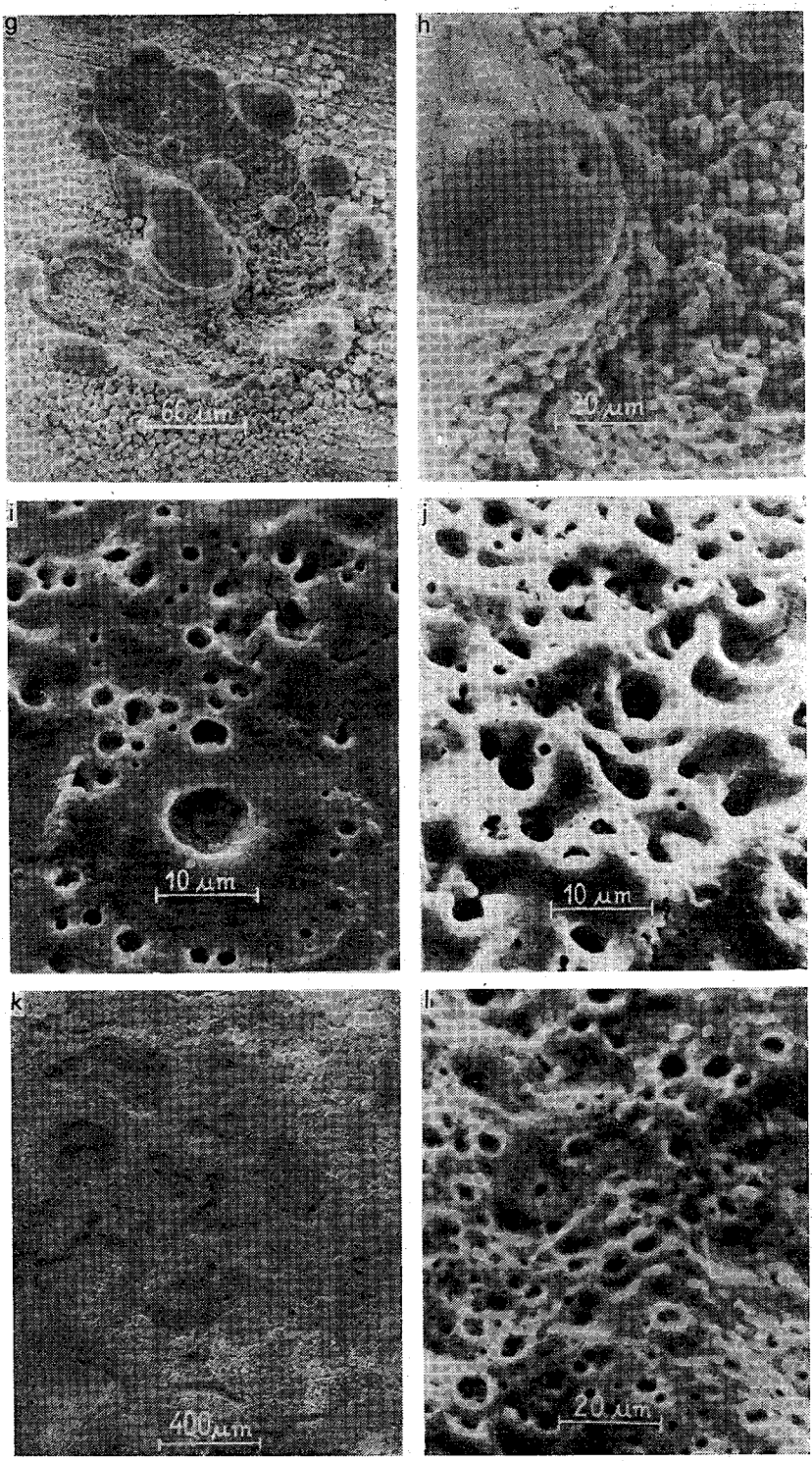
2. WYNIKI DOTYCHCZASOWYCH BADAŃ NAD ŚLADAMI ŁUKU NA KATODZIE

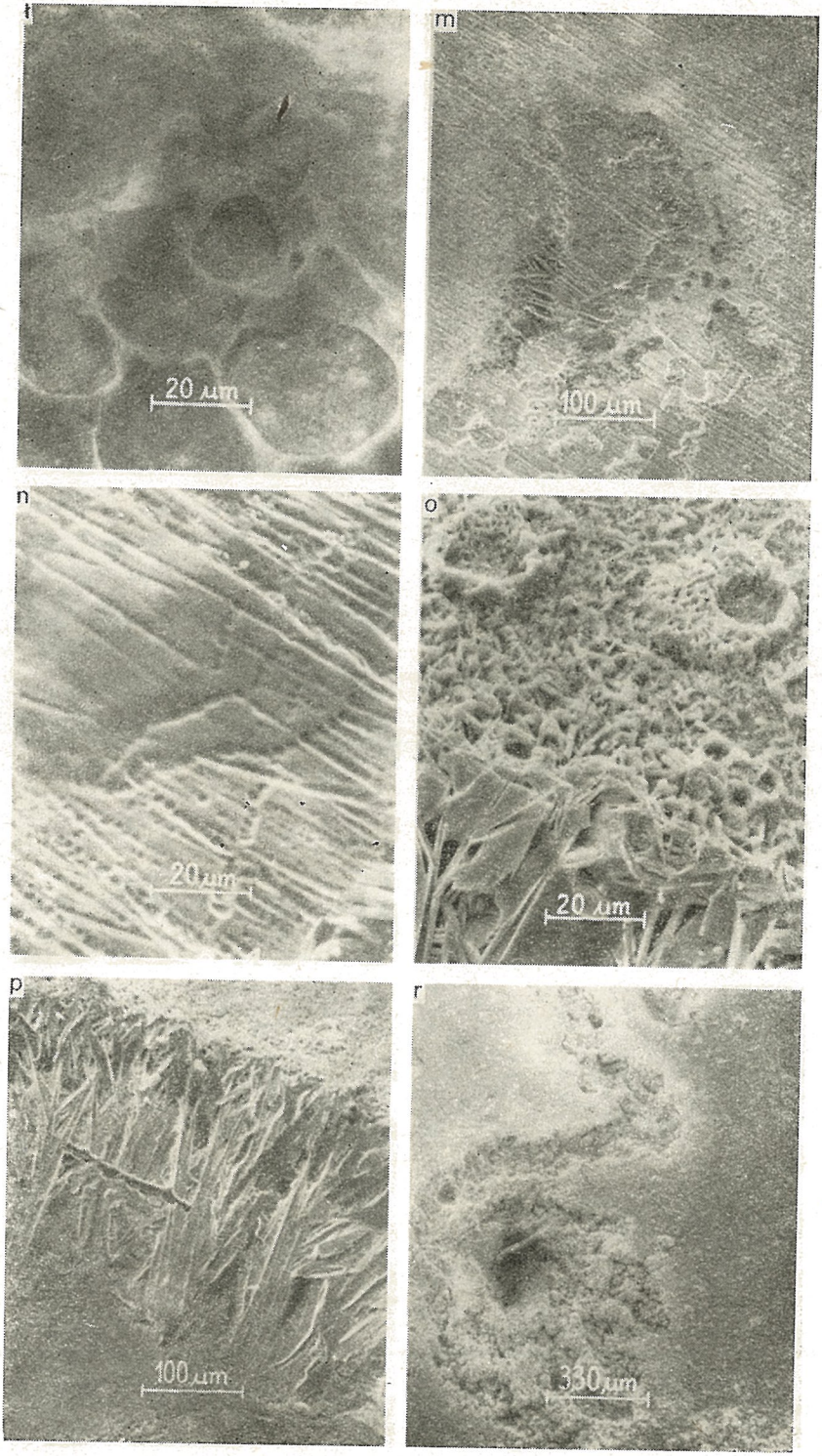
Pierwsze wyniki badań i obserwacji śladów łuku na katodzie przedstawione przez autora w pracy [33] wykonane były jedynie dla katody miedzianej i wykorzystane zostały jako wstępne informacje dla poparcia przyjętych założeń nowego wówczas impulsowego modelu stopy katodowej łuku. Zastosowanie mikroskopu skaningowego pozwoliło na wykonanie zdjęć śladów łuku o znacznie lepszej jakości i bardziej plastycznych, a także



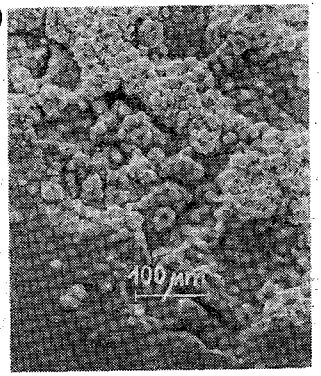
Rys. 1. Ślady wyładowania sfotografowane za pomocą komórki Kerra po czasach 2, 4, 8, 20 30, 50, μs [30]


a
330 µm
b
20 µm
c
100 µm
d
10 µm
e
20 µm
f
6.52
20 µm

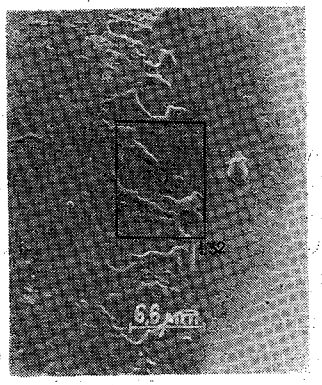

g
60 μm
h
20 μm
i
10 μm
j
10 μm
k
400 μm
l
20 μm


k
20 μm
m
100 μm
n
20 μm
o
20 μm
p
100 μm
r
330 μm

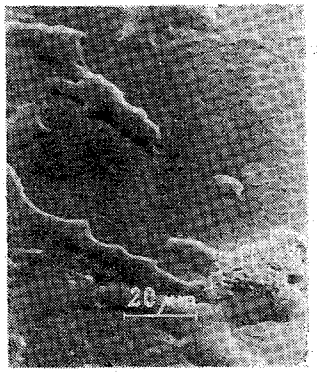
s)



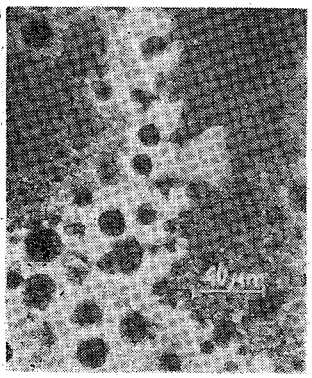
t)



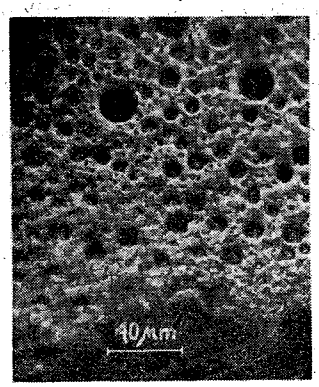
u)



w)



z)



na zaobserwowanie wspomnianych już struktur krystalicznych. Dla łuku pozornie nieruchomego i wolnoporuszającego się otrzymane ślady łuku potwierdzają obserwacje innych autorów [5, 12, 14, 15, 18, 24, 28, 38], którzy stwierdzili, że ślad łuku zawiera szereg mikrowgłębień w kształcie kraterów.

Przedstawione przykładowo na rys. 2 zdjęcia śladów łuku w zakresie powiększeń $50 \div 3000$ razy, wykonane przez autora przy różnych wartościach prądu łuku i czasach palenia się stosunkowo długich, potwierdzają ten fakt w większości obserwowanych śladów. Wzajemne położenie kraterów jest zupełnie przypadkowe. Średnice kraterów są zróżnicowane, niemniej dokładna analiza (szczególnie bezpośrednie obserwacje pod mikroskopem optycznym) wykazuje, że na dnie dużych kraterów występują kratery mniejsze, które wymiarami odpowiadają kraterom występującym w innych miejscach na powierzchni katody (rys. 2a, b, f, k). Skłania to do wysunięcia przypuszczenia, że duże kratery tworzone są przez powstające i zanikające na danym obszarze zgrupowania plamek elementarnych.

W przebiegu tworzenia się miejscowego śladu odpowiadającego zgrupowaniu plamek elementarnych można znaleźć analogie w wynikach badań wyładowań krótkotrwałych Sommerville'a (rys. 1) [30]. W miejscach znacznego zgrupowania plamek elementarnych lub w miejscach gdzie łuk pali się dłuższy czas nieruchomo, ślady plamek elementarnych mogą być trudno zauważalne a nawet niewidoczne. Silne przegrzanie powierzchni katody oraz wzajemne nakładanie się kraterów prowadzi do deformacji śladu i bezwładnego rozpływu stopionej warstwy metalu poddanemu dodatkowo działaniu ciśnień, których źródłem jest stopa katodowa i strumienie plazmy z anody. Stygnąca katoda po przemieszczeniu się plamki w inne miejsce lub po zgaszeniu łuku zniekształca i wyrównuje ślady plamek elementarnych. Kratery najłatwiej jest zauważyć w miejscach, gdzie metal nie był jeszcze stopiony lub temperatura była zbliżona do temperatury topnienia, a więc na obrzeżach śladu i krawędziach zgrupowań plamek elementarnych.

Przeprowadzone w latach sześćdziesiątych badania śladów krótkotrwałych wyładowań łukowych w powietrzu i próżni wykazały że ślady tych wyładowań na katodzie składają się z dużej liczby kraterów o średnicy $(1 \div 150)$ μm, przy czym każdemu kraterowi odpowiada eksplozja materiału katody [23, 25, 30, 9, 38]. Każdy krater jest śladem po plamce katodowej (nie używano wówczas terminu plamka elementarna), której czas istnienia w jednym miejscu oceniano na $(5 \cdot 10^{-4} \div 4 \cdot 10^{-6})$s [23, 35, 38].

Rys. 2. Obrazy śladów łuku na katodzie

a) $I = 5$ A, $t_l = 1$ s Cu; b) $I = 500$ A, $t_l = 10$ ms Cu;
c) $I = 20$ A, $t_l = 1$ s Cu; d) $I = 5$ A, $t_l = 1$ s Cu;
e) $I = 70$ A, $V_l = 3$ m/s Cu; f) $I = 50$ A, $V_l = 0$ Cu;
g) $I = 1$ A, $t_l = 0{,}2$ s Fe; h) $I = 1$ A, $t_l = 0{,}2$ s Fe;
i) $I = 50$ A, $t_l = 0{,}2$ s Fe; j) $I = 50$ A, $t_l = 0{,}2$ s Fe;
k) $I = 20$ A, $t_l = 0{,}2$ s Ni; l) $I = 20$ A, $t_l = 0{,}2$ s Ni;
ł) $I = 20$ A, $t_l = 0{,}2$ s Ni; m) $I = 1$ A, $t_l = 0{,}2$ s Mo;
n) $I = 1$ A, $t_l = 0{,}2$ s Mo; o) $I = 1$ A, $t_l = 1$ s Mo;
p) $I = 1$ A, $t_l = 1$ s Mo; r) $I = 1$ A, $t_l = 0{,}2$ s Sn;
s) $I = 1$ A, $t_l = 0{,}2$ s Sn; t) $I = 50$ A, $V_l = 0{,}1$ m/s W;
u) $I = 50$ A, $V_l = 0{,}1$ m/s W; w) $I = 100$ A, $t_l = 0{,}2$ s W;
z) $I = 100$ A, $t_l = 0{,}2$ s W

Tablica 1

Pomierzone najmniejsze i największe wartości średnic karterów d_e oraz odpowiadających im powierzchni S_e

Materiał katody	d_e	S_e
	$\times 10^{-6}$ m	m^2
miedź	$5 \div 15$	$2 \cdot 10^{-11} \div 2 \cdot 10^{-10}$
stal	$2 \div 10$	$3 \cdot 10^{-12} \div 8 \cdot 10^{-11}$
nikiel	$10 \div 30$	$8 \cdot 10^{-11} \div 7 \cdot 10^{-10}$
molibden	$2 \div 10$	$3 \cdot 10^{-12} \div 8 \cdot 10^{-11}$
wolfram	$1 \div 20$	$1 \cdot 10^{-12} \div 3 \cdot 10^{-10}$

W tablicy 1 zestawiono pomierzone przez autora najmniejsze i największe wartości średnic kraterów d_e oraz odpowiadające im powierzchnie S_e ze zdjęć śladów łuku, których przykłady przedstawiono na rys. (2a ÷ 2z). W katodzie molibdenowej (rys. 2o, p) oprócz kraterów widoczne są kryształy powstałe po zgaszeniu łuku.

Interesujący przykład śladów plamek elementarnych zaobserwowano na wolframie (rys. 2t, u); widoczne jaśniejsze krążki bez oznak nadtopień potwierdzają założenie zawarte w impulsowym modelu stopy katodowej, że nie każda plamka elementarna musi pozostawiać ślad w postaci krateru lub nadtopienia.

Wiele wątpliwości występuje przy interpretacji zdjęć śladów na katodzie żelaznej (w rzeczywistości stalowej). Widoczne na rys. 2i i 2j kratery i głębokie otwory mogą być zarówno śladami plamek elementarnych jak i pozostałościami po wydobywających się pod wpływem wysokiej temperatury cząsteczkach domieszek innych pierwiastków (np. węgla) lub uwolnionych pęcherzyków gazów. Ostatnie uwagi muszą być uwzględnione również przy interpretacji śladów łuku dla innych materiałów katody, a szczególnie przy określaniu wszelkiego rodzaju wgłębień jako śladów plamek elementarnych. Obraz śladu łuku jest obrazem powierzchni katody po zgaszeniu łuku więc procesy chemiczne, np. utlenianie, powstawanie związków metalicznych i niemetalicznych zachodzące w czasie chłodzenia przegrzanej warstwy materiału katody kształtują obraz śladu łuku. Obecność w śladach łuku struktur o nieokreślonym składzie szczególnie widoczna na rys. 3, 4, 2c, d, e, f, g, h, i, j, l, p, r, u, w, z, nie pozwala jednoznacznie interpretować śladów plamek elementarnych, a tym bardziej formułować precyzyjnych wniosków odnośnie ich wymiarów.

W ostatnich latach w badaniach Guila, Hitchcocka, Seckera [18, 19, 20, 21, 22, 31] przeprowadzonych przy krótkich czasach łukowych nawet rzędu 10^{-9} s i przy różnych grubościach tlenków pokrywających katodę $(2 \div 300)$ nm stwierdzono obecność kraterów o wymiarach $(0{,}1 \div 1)$ μm. Zastosowane przez autorów warunki badań znacznie odbiegały od normalnych warunków pracy łączników, podczas których łuk pali się przez kilka lub kilkanaście ms lub nawet dłużej, jak to ma miejsce w zakresie prądów krytycznych. Otrzymane natomiast przez autora obrazy śladów łuku dotyczą czasów łukowych dostatecznie długich, aby na powierzchni katody wytworzyła się warstwa stopionego metalu, na której powstają i zanikają plamki elementrarne.

Wpływ prędkości łuku na obraz śladu łuku jest szeroko omówiony w literaturze. Różne prędkości łuku oraz odmienne formy zniszczeń połukowych tłumaczy się właściwościami fizycznymi materiałów elektrod [1, 8, 9, 10, 24, 26, 32]. Wielu badaczy stwierdziło również, że prędkość ruchu łuku zależy od stanu powierzchni katody (tlenki, warstwy obce, chropowatość) [2, 8, 11, 24, 26, 27].

Analiza śladów łuku ruchomego przeprowadzona pod kątem obecności kraterów dowiodła, że ze wzrostem prędkości łuku obserwuje się zanik kraterów [5, 6]. Przykładowo przedstawione fragmenty śladu łuku na katodzie miedzianej dla łuku nieruchomego (rys. 2f) i łuku ruchomego $V_l = 3$ m/s (rys. 2e) wykazują, że przy łuku nieruchomym i ruchomym widoczne są struktury charakterystyczne dla tlenków i wgłębienia wynikające z nałożenia dużej liczby kraterów. Natomiast przy łuku ruchomym występują jedynie płytkie powierzchniowe zniszczenia o nieregularnych kształtach. Podobne stwierdzenia odnoszą się i do innych materiałów.

Z punktu widzenia impulsowego modelu stopy katodowej nie każdy impuls energii docierający do powierzchni katody winien doprowadzić do wybuchowego parowania mikroobjętości materiału. O efekcie wyładowania elementarnego w postaci śladu decyduje: wielkość i gęstość energii zawartej w impulsie, czas spłynięcia ładunku, tzw. podstawowa temperatura na powierzchni katody, rodzaj materiału katody [4, 5, 28, 29, 33, 34, 38]. Najdogodniejszymi miejscami do powstania kraterów są te, w których czas pozostawania plamki lub stopy katodowej jest dostatecznie długi, aby temperatura podstawowa zbliżona była do temperatury topnienia materiału katody.

Badania śladów łuku przedstawione w niniejszej pracy dotyczą: łuku palącego się w powietrzu o ciśnieniu atmosferycznym, zakresu prądów łukowych $(1 \div 100)$ A i czasów łukowych rzędu $(0{,}1 \div 1)$ s. Pomierzone przez autora zakresy średnic kraterów dla rozpatrywanych materiałów katody wraz z przedstawionymi uwagami odnośnie czynników kształtujących proces ich powstawania stanowią jeden z elementów pozwalających jedynie szacować parametry plamki elementarnej.

Gebauer [14, 15] badając liczbę pojedynczych plamek występujących w łuku elektrycznym określił prąd przypadający na pojedynczą plamkę $(1 \div 3)$ A, natomiast Kesajew [23] podaje, że minimalny prąd, tzw. prąd progowy, przy którym łuk pali się stabilnie i plamka katodowa nie ulega podziałowi, zależy od temperatury wrzenia materiału katody oraz przewodności cieplnej materiału katody, dla miedzi nie przekracza 3 A. Biorąc pod uwagę rozważania przedstawione w pracach [4,33, 34] przyjęto, że prąd w plamce elementarnej zawiera się w granicach kilku amperów.

Zakładając prąd w plamce elementarnej, np. $(1 \div 5)$ A oraz dzieląc wartość prądu 1 A przez największą pomierzoną powierzchnię krateru i wartość prądu 5 A przez najmniejszą powierzchnię krateru (tablica 1), otrzymano następujące przedziały średnich gęstości prądu w plamce elementarnej (tablica 2).

W tablicy 2 przedstawiono również przedziały średnich gęstości prądu w plamkach elementarnych otrzymane na podstawie obliczeń pól temperaturowych plamek elementarnych. Pełny opis metody i założeń przyjętych do obliczeń zamieszczono w pracach [3, 33]. W obliczeniach założono między innymi, że plamki elementarne powstają i zanikają na powierzchni katody o temperaturze zbliżonej do temperatury topnienia metalu. Zanik plamki elementarnej następuje w chwili, kiedy temperatura na jej krawędzi osiąga tempe-

Tablica 2

Gęstość prądu w plamkach elementarnych wyznaczona na podstawie pomiarów średnic kraterów i obliczeń pól temperaturowych plamek elementarnych

Materiał katody	j_e A/m² z pomiarów średnic i powierzchni kraterów	j_e A/m² z obliczeń pól temperaturowych plamek elementanych
miedź	$1 \cdot 10^{10} \div 3 \cdot 10^{11}$	$1 \cdot 10^{11} \div 5 \cdot 10^{11}$
stal	$1 \cdot 10^{10} \div 2 \cdot 10^{12}$	$5 \cdot 10^{9} \div 5 \cdot 10^{10}$
nikiel	$1 \cdot 10^{9} \div 1 \cdot 10^{10}$	$5 \cdot 10^{9} \div 5 \cdot 10^{10}$
wolfram	$3 \cdot 10^{9} \div 1 \cdot 10^{12}$	$5 \cdot 10^{10} \div 5 \cdot 10^{11}$

raturę wrzenia metalu, po czym następuje eksplozyjne parowanie cząstek par metalu i powstaje ślad w postaci krateru.

Z [23, 38] wynika, że czas trwania eksplozji mikroobjętości materiału katody w krótkotrwałych wyładowaniach łukowych i iskrowych jest rzędu $(10^{-8} \div 10^{-6})$ s, również rozważania przedstawione w pracach [33, 35] skłaniają do przyjęcia tego przedziału czasowego jako czasu „życia" plamek elementarnych.

Zgodność średnich wartości gęstości prądu w plamce elementarnej (tablica 2) wyznaczonych przy opisanych wyżej założeniach oraz danych zaczerpniętych z literatury można uznać za zadowalającą w zakresie rzędów wielkości. Rozpiętość górnej wartości gęstości prądu j_e dla żelaza ($5 \cdot 10^{10}$ A/m² i $2 \cdot 10^{12}$ A/m²) można tłumaczyć uwagami przytoczonymi odnośnie rysunków 2i i 2j. Należy mieć również na uwadze fakt, że średnice kraterów pomierzono na powierzchni katody, wykonanej z metalu o wysokim stopniu czystości, oczyszczonej przed zapaleniem łuku, lecz potem poddanej działaniu wysokiej temperatury i czynników zewnętrznych. Natomiast obliczenia pól temperaturowych wykonano przy danych cieplno-fizycznych (przewodność cieplna, ciepło właściwe, gęstość) jak dla czystych metali i nie uwzględniono zmian zachodzących w składzie chemicznym powierzchni katody poddanej działaniu łuku.

3. PODSUMOWANIE

Prowadzone w ostatnich latach prace teoretyczno-eksperymetalne nad wyładowaniem elementrarnym w łuku elektrycznym ukierunkowane są na stworzenie możliwie pełnego modelu matematyczno-fizycznego stopy katodowej łuku oraz wyznaczenia parametrów plamek elementarnych. Środki techniczne będące obecnie w dyspozycji badaczy ograniczają zakres prowadzonych badań, a wyniki eksperymentów nie zawsze spełniają pokładane nadzieje, gdyż współzależność i złożoność zjawisk utrudnia interpretację wyników badań, a więc i wyznaczenie parametrów plamek elementarnych. Szeroka i niekiedy daleko idąca rozbieżność w warunkach i parametrach prowadzonych badań wymaga od autorów ostrożnego korzystania z wyników innych badaczy.

Wyznaczenie takich parametrów plamek elementarnych, jak: prąd i gęstość prądu w plamce elementarnej, czas „życia" plamki elementarnej, mechanizm powstawania

i zanikania plamek elementarnych w dalszym ciągu wymagają odpowiedzi, zwłaszcza od strony ilościowej. Aktualny stan badań nie pozwala na jednoznaczną odpowiedź, gdyż publikowane wyniki różnią się nawet o rząd wielkości, a różnice parametrów warunków badań (prąd i napięcie łuku, rodzaj i wartość ciśnienia otaczającego gazu, czas palenia się łuku, materiał katody) są również znaczne.

Brak pełnych informacji, szczególnie eksperymentalnych, w zakresie składu chemicznego śladów oraz parametrów kształtujących przebieg i zanik wyładowania elementarnego uniemożliwiają precyzyjne wyznaczenie parametrów plamek elementarnych. Mikroanaliza śladów łuku, mająca na celu wyznaczenie parametrów plamek elementarnych na podstawie jej śladów jest tylko fragmentem całości zagadnienia. Stanowi element uzupełniający lub potwierdzający wnioski wyciągnięte z innych badań.

W śladach łuku ruchomego na katodzie obserwuje się powierzchniowe zniszczenia o nieregularnych kształtach, powierzchniowe nadtopienia i wgłębienia w postaci rowków, wynikające między innymi z ruchu stopy łuku. W śladach łuku nieruchomego oprócz śladów charakterystycznych dla łuku ruchomego występują przede wszystkim kratery, które przyjęto traktować jako ślady plamek elementarnych. Kratery mają zróżnicowane średnice, wzajemnie nakładają się, na krawędziach wielu kraterów występują obwałowania w postaci stopionego metalu. W śladach widoczne są struktury krystaliczne i krople stopionego metalu.

Istotną rolę w przebiegu nagrzewania katody plamkami elementarnymi oraz tworzeniu się ich śladów odgrywa temperatura na powierzchni katody, tzw. temperatura podstawowa [33, 34].

W łuku nieruchomym palącym się przez dłuższy czas, istnieją warunki zapewniające wytworzenie się temperatury podstawowej równej lub zbliżonej do temperatury topnienia materiału katody. W łuku ruchomym temperatura podstawowa będzie zwykle niższa od temperatury topnienia metalu (na wielkość tej temperatury ma wpływ wartość prądu łuku, prędkość stopy łuku, rodzaj materiału katody). Impulsy energii docierające do powierzchni katody o różnej temperaturze podstawowej mogą więc pozostawić ślady o zróżnicowanych kształtach. Świadczy to o tym, że impulsy energii nie zawsze doprowadzają do wybuchowego parowania materiału katody albo materiał katody nie jest w stanie miejscowo przegrzać się do takiej temperatury, przy której może nastąpić eksplozja materiału.

Mikroskopowa analiza śladów łuku z powodu różnorodności kształtu i rozmiarów zniszczeń na powierzchni katody napotyka również na trudności w jednoznacznym określeniu śladów plamek cząstkowych i elementarnych, stąd też powstają rozbieżności w ocenie gęstości prądu. Stosowanie powiększeń śladów rzędu 10000 razy i więcej powoduje niejasności w interpretacji wyników. Obserwacje tego samego krateru w powiększeniu 1000 i np. 100 000 razy zmieniają jakościowo wyniki analizy. Przy powiększeniach 100 000 razy obiektem obserwacji przestaje być widziany uprzednio krater lecz wszelkie mikrostrukturalne deformacje dna krateru nawet w kształcie mikrowgłębień. Jeżeli teraz wszelkie mikrowgłębienia potraktowane zostaną jako ślady wyładowań elementarnych, a prądy przypadające na jedno takie wyładowanie przyjmiemy nawet mniejsze od 1 A, to gęstości prądu w wyładowaniu elementarnym winny być znacznie większe, nawet rzędu 10^{12}, 10^{13} A/m^2 przy czasach trwania mniejszych od mikrosekundy [7, 12, 13, 22, 27].

W dalszym ciągu brak jest przekonujących danych co do wartości prądu jakie można przyjąć, aby wyznaczyć gęstość prądu w plamkach elementarnych, których ślady reprezentują poszczególne kratery. Przedstawione w tablicy 2 rzędy wielkości gęstości prądu w plamkach elementarnych w świetle przytoczonych uwag, jako wartości szacunkowe, wydają się być uzasadnione.

Statystyczne ujęcie częstości występowania kraterów o danej średnicy, a następnie precyzowanie parametrów plamek elementarnych wymaga prowadzenia dalszych kompleksowych, pracochłonnych i kosztownych prac, aby umożliwić porównywanie wyników badań prowadzonych w różnych ośrodkach. Próby wyjaśnienia przebiegu procesów termodynamicznych w wyładowaniach elementarnych za pomocą znanych praw fizyki klasycznej mają jeszcze charakter hipotetyczny.

Problemu ścisłego określenia wartości parametrów wyładowania elementarnego nie można więc uznać za zakończony.

BIBLIOGRAFIA

1. G. D. Amft, *Über den Lichtbogenwandung im Bereich geringer Geschwindigkeiten*, Elektrische Kontakte, 1970.
2. G. D. Amft, *Zur Lichtbogenwanderung auf heterogenen Elektrodenmaterialien*, Sec. Int. Symp. Switching Arc Phenomena, Łódź, 1973.
3. B. Bolanowski, *Gęstość prądu w plamce katodowej oraz pole temperaturowe w katodzie łuku elektrycznego*, ZNPŁ Elektryka, z. 25, 1967.
4. B. Bolanowski, A. Sobieszczuk, *Ein neues Modell des Kathodenfusses des elektrischen Lichtbogens*, Sec. Int. Symp. Switching Arc Phenomena, Łódź, 1973.
5. B. Bolanowski i in., *Mikroanalyse der Kathodenfleckspur*, Sec. Int. Symp. Switching Arc Phenomena, Łódź. 1973.
6. B. Bolanowski, M. Marusik, *Erscheinungshäfigkeit von Kratern Spuren eines Kathoden-Elementarflecks*, Third Int. Symp. Switching Arc Phenomena, Łódź 1977.
7. B. Bolanowski, A. Holi, *Elementary discharge process on an oxide-coated Cu Cathode*, Fourth Int. Symp. Switching Arc Phenomena, Łódź, 1981.
8. A. V. Bolotov i in., *The effect of the magnetic field and the electrodes temperature upon the electric arc motion*, Third Int. Symp. Switching Arc Phenomena, Łódź, 1977, Materiały pokonferencyjne.
9. G. Burkhard, *Über den Zusammenhang zwischen Abbrandverhalten und Kathodenmechanismus von Silber-Verbundwerkstoffen*, Int. Symp. Siela, Plovdiv, 1971.
10. G. Burkhard, *Der Einfluss von Struktureigenschaften galvanischer Niedeschlage auf das Lautverhalten von Lichtbogen*, Sec. Int. Symp. Switching Arc Phenomena, Łódź, 1973, Materiały pokonferencyjne.
11. G. Burkhard, *Die Vorgänge an der Kathode eines Lichtbogens zwischen Metallelektroden und ihre theoretische und experimentalle Deutung*, Dissert, SC, Berlin, 1971.
12. M. G. Drouet, S. Gruber, *Dynamic measurements and a simplified model of cathode emission in a moving arc*, Proc. Int. Conf. Phen. Ion. Gas., Eindhoven, 1975.
13. M. G. Drouet, S. Gruber, *Dynamic measurements of cathode emission in a moving arc*, IEE, 1976.
14. E. Gebauer, *Verhalten des Lichtbogens und des Elektrodenabbrandes in einem Druckbereich von 10^{-6} Torr bis Atmosphärendruck*, Int. Symp. Switching Arc Phenomena, Łódź, 1970.
15. E. Gebauer, *Über die Vorgänge an den Elektroden einer Lichtbogenentladung in einem Druckbereich von 10^{-6} Torr bis 760 Torr*, Elektrie, 1974, z. 9.
16. K. D. Froome, *The Rate of Growth of Current and Bechavior of the Cathode Spot in Transient Arc Discharges*, Proc. Phys. Soc. 1948, t. 60.

17. G. N. Fursey i in., *Discharges and Electrical Insulation in Vaccuum*, Proc. of the Vth. Int. Symp., Poznań, 1972.
18. A. E. Guile, *Model of emitting sites and erosion on nonretractory arc cathodes with relatively thick oxidee film*, Proc. IEE, v. 121, nr 12, 1974.
19. A. E. Guile, A. H. Hitchcock, *Effect of oxide thickness on the microstructure of arc cathode spots on copper*. Third Int. Symp. Switching Are Phenomena, Łódź, 1977.
20. A. E. Guile, A. H. Hitchcock, *Oxide films on arc cathodes on their emission and erosion*, J. Phys. D. Appl. Phys., 1975.
21. A. E. Guile, A. H. Hitchcock, *Arc-cathodes craters on cooper at high currents and with reduced gas pressures*, Proc. IEE, 1978.
22. A. E. Guile, *Arc cathode root microstructure on various metals*, Fourth Int. Symp. Switching Arc Phenomena, Łódź, 1981.
23. I. G. Kesajew, *Katodnyje procesy elektriczeskoj dugi*, Nauka, Moskwa, 1968.
24. R. Kordas, *Wpływ domieszek CuO i CdO w kompozytach Ag-CuO i Ag-CdO na proces ruchu łuku po powierzchni*, Praca doktorska, Wrocław, 1982.
25. J. Kutzner, *Charakterystyka fizyczna i ocena podstawowych parametrów łuku elektrycznego w próżni*, Praca doktorska, Łódź-Poznań, 1967.
26. H. J. Mau, *Stand der Erkenntnisse über die Beanspruchung elektrischer Kontakte in der Starkstromtechnik*, V Int. Tagung über elektrische Kontakte, München, 1970.
27. M. Marusik, *Dynamika wyładowania elementarnego łuku krótkiego palącego się przy ciśnieniu atmosferycznym*, Praca doktorska, Łódź, 1981.
28. V. Rachowski, *Fizyczeskije osnowy elektriczeskogo toka w vacuumie*, Nauka, Moskwa, 1970.
29. V. Rakhowski, *Experimental study of dynamics of cathode spots development*, IEEE Trans. on Plasma. Sc., V. PS-4, nr 2, 1976.
30. J. M. Sommerville, C. T. Graiger, *Elektrode Phenomena in Transient Arcs*, J. Appl. Phys., 1956, t. 7, nr 11.
31. C. Sanger, P. Secker, *Arc-cathode current density measurements*, J. Phys. D. Appl. Phys., 1971.
32. K. H. Schröder, *Die Stabilisierung von Lichtbogenfusspunkten durch heterogene Werkstoffe und ihre Einfluss auf den Materialverschleis*, 4-th Int. Conf. on Electric Conteat Phenomena, Swense'á, 1968.
33. A. Sobieszczuk, *Nagrzewanie katody łuku elektrycznego z uwzględnieniem impulsowego modelu stopy katodowej*, Praca doktorska, Łódź, 1972.
34. A. Sobieszczuk, *Erwärmung heterogener Kathoden durch Elementarflecks*, Third Int. Symp. Switching Arc Phenomena, Łódź, 1977.
35. A. Sobieszczuk, *Impulsentladungen im Kathodenfuss des Lichtbogens*, Fourth Int. Symp. Switching Arc Phenomena, Łódź, 1981.
36. K. Wesołowski, *Metaloznawstwo, t. 1, Badanie metali*, PWT, Warszawa, 1959.
37. K. Wesołowski, *Metaloznawstwo, t. 3, Metale nieżelazne i ich stopy*, WNT, Warszawa 1966.
38. Praca zbiorowa, *Zagadnienia podstawowe spektralnej analizy atomowej*, WNT, Warszawa 1972.

A. SOBIESZCZUK

SOME PROBLEMS OF MICROSCOPIC ANALYSIS OF ARC TRACES ON THE CATHODE

Summary

For many years microscopic analysis of arc traces on the cathode has been used to evaluate current density in the cathode spot and analyze the phenomena which occur in the cathode of an electric arc. On the basis of the microscopic analysis of arc traces on the cathode performed by author himself and discussed in literature, an attempt has been made to estimate parameters of the elementary spots, particularly those which fall in the range of average current density. In his analysis of arc traces on the cathode the author points out problems and doubts connected with precise interpretation of the traces and addresses certain

phenomena related to the apperance and disappearance of the elementary spot which have not been studied before.

A. SOBIESZCZUK

PROBLÈMES CHOISIS DE L'EXAMEN AU MICROSCOPE DES TRACES DE L'ARC SUR LA CATHODE

Résumé

Depuis plusieurs années l'examen au microscope des traces de l'arc sur la cathode servait à déterminer la densité du courant dans la tache cathodique et à l'analyse des phènomènes se produisant dans la cathode de l'arc électrique. L'analyse de ses propres examens au microscope ainsi que de ceux trouvés dans les publications a permis à l'auteur d'estimer les paramètres des taches élémentaires surtout en ce qui concerne la densité moyenne du courant dans ces taches. On trouve important de souligner que leur interprétation univoque est difficile et souvent hypothétique. L'attention de l'auteur fut retenue par certains phénomènes se produisant au moment de formation et de disparition des taches dont l'examen fut souvent négligé usqu'à présent.

A. SOBIESZCZUK

GEWÄHLTE PROBLEME DER MIKROSKOPISCHEN UNTERSUCHUNG VON BOGENSPUREN AN DER KATHODE

Zusammenfassung

Die mikroskopische Analyse von Bogenspuren an der Kathode wurde seit vielen Jahren zur Bestimmung der Stromdichte im Kathodenfleck und zur Analyse der in der Kathode des Lichtbogens entstehenden Erscheinungen genutzt. Der Verfasser versucht in Anlehnung an eigene mikroskopische Untersuchungen von Bogenspuren an der Kathode und auf Grund der aus der Literatur hervorgehenden Daten die Parameter der elementaren Flecke einzuschätzen, insbesondere im Bereich der mittleren Stromdichte im elementaren Fleck. Bei mikroskopischer Analyse von Bogenspuren an der Kathode wurde besondere Aufmerksamkeit den Schwierigkeiten und Zweifeln einer eindeutigen Interpretation geschenkt sowie den vielen, bisher ungenügend berücksichtigten Erscheinungen, die hierbei entstehen und verschwinden.

А. СОБЕЩУК

ИЗБРАННЫЕ ВОПРОСЫ МИКРОСКОПИЧЕСКОГО АНАЛИЗА КАТОДНЫХ СЛЕДОВ ЭЛЕКТРИЧЕСКОЙ ДУГИ

Резюме

Микроскопический анализ катодных следов электрической дуги является основным методом определения плотности тока в катодном пятне а также методом анализа процессов происходящих в катоде электрической дуги. Опираясь на результатах собственных исследований а также литературных данных, автор дает оценку параметров элементарных пятен, особенно средней плотности тока в элементарном пятне а также ряда других явлений происходящих во время их возникновения и исчезания. В работе указывается на основные трудности в однозначной интерпретации дуговых следов на катоде.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 453—468

621.394.662

Optymalizacja energetyczna układów napędu elektrycznego o regulowanej prędkości kątowej

ANDRZEJ TRZYNADLOWSKI (WROCŁAW)

Instytut Układów Elektromaszynowych Politechniki Wrocławskiej

Otrzymano 10.6.1981

Przedstawiono algorytmy sterowania i syntezy układów napędowych prądu stałego i przemiennego o minimalnym rozproszeniu energii przy zmianach prędkości kątowej. Do syntezy energo-optymalnych układów napędowych zastosowano tak zwaną metodę parametrów węzłowych, stanowiącą formę matematycznego programowania nieliniowego. Rozważania teoretyczne zilustrowano przykładami obliczeniowymi.

1. WSTĘP

Postęp techniczny w dziedzinie środków sterowania układów napędu elektrycznego, w szczególności wzrastające zastosowanie mikroprocesów i mikrokomputerów, umożliwia optymalizację sterowania tych układów nie tylko z punktu widzenia wymagań dynamicznych, lecz również poprawy wskaźników energetycznych, coraz bardziej istotnych w warunkach rosnącej wagi problemu oszczędności energii.

W artykule niniejszym rozważono zagadnienie praktycznej realizacji podstawowych układów napędowych prądu stałego i przemiennego o minimalnym rozproszeniu energii przy regulacji prędkości kątowej. Rozpatrzono układ napędowy z obcowzbudnym silnikiem prądu stałego ze sterowaniem od strony twornika oraz układ z silnikiem asynchronicznym ze sterowaniem częstotliwościowym, wskazując na analogie elektromechaniczne między tymi układami. Do syntezy energo-optymalnego układu napędowego o zadanych warunkach działania zastosowano tak zwaną metodę parametrów węzłowych, stanowiącą formę matematycznego programowania nieliniowego.

2. STEROWANIE ENERGO-OPTYMALNE SILNIKA PRĄDU STAŁEGO

Rozpatruje się silnik obcowzbudny o ustalonym, znamionowym wzbudzeniu i sterowaniu napięciem twornika. Mianem sterowania energo-optymalnego określa się takie sterowanie prędkości kątowej napędu, przy którym zmiana tej prędkości o określoną wartość

związana jest z minimalną wartością energii E_R rozproszonej w uzwojeniu twornika silnika.

Oznaczając przez U, I, R odpowiednio napięcie, prąd i rezystancję twornika, zaś przez T czas regulacji prędkości, można powyższe kryterium optymalizacji energetycznej zapisać w postaci zależności

$$E_R = R\int_0^T I^2(t)dt. \tag{1}$$

Zakładając, że moment oporowy M_o urządzenia napędzanego pozostaje stały podczas zmiany prędkości kątowej ω_o tego urządzenia i pomijając straty mechaniczne, można napisać następujące równania napędu:

$$I(t) = \frac{N}{k}[M_o + J_w\dot{\omega}_o(t)], \tag{2}$$

$$U(t) = \frac{k}{N}\omega_o(t) + RI(t), \tag{3}$$

gdzie

$$N = \frac{\omega_o}{\omega} \tag{4}$$

oznacza przełożenie przekładni pomiędzy silnikiem a urządzeniem napędzanym, czyli stosunek prędkości kątowej ω_o do prędkości kątowej ω silnika,

$$J_w = J_o + \frac{J}{N^2} \tag{5}$$

oznacza wypadkowy moment bezwładności na wale urządzenia napędzanego, pochodzący od momentu bezwładności J wirnika silnika i momentu bezwładności J_o urządzenia napędzanego, k — stała elektryczna silnika.

Oznaczając przez $\Delta\omega_o$ przyrost prędkości kątowej $\omega_o(t)$, dokonujący się w okresie czasu od $t = 0$ do $t = T$, można sformułować następujące zadanie wariacyjne: zminimalizować funkcjonał E_R, określony zależnościami (1) i (2), przy warunku brzegowym $\omega_o(T) = \omega_o(0) + \Delta\omega_o$. Łatwo stwierdzić, że rozwiązanie $\omega_o^*(t)$ tego zadania ma postać

$$\omega_o^*(t) = \frac{\Delta\omega_o}{T}t + \omega_o(0). \tag{6}$$

Odpowiadająca zależności (6) energia rozproszona w uzwojeniu twornika silnika równa się

$$E_R^* = \left[\frac{N}{k}\left(M_o + \frac{J_w\Delta\omega_o}{T}\right)\right]^2 RT. \tag{7}$$

Sterowanie według wzoru (6) nazywane będzie dalej sterowaniem suboptymalnym (SS). Określenie to wynika z możliwości dalszej minimalizacji E_R^* po czasie T. Optymalna

wartość czasu T minimalizująca funkcję (7) równa się

$$T^* = \frac{J_w|\Delta\omega_o|}{M_o}, \tag{8}$$

zaś odpowiadająca jej energia rozproszona wynosi

$$E_R^{**} = \left[\frac{N}{k}(1+\operatorname{sign}\Delta\omega_o)\right]^2 RJ_w M_o|\Delta\omega_o|. \tag{9}$$

Sterowanie ze stałą czasową T^* nazywane będzie dalej sterowaniem optymalnym (SO). Łatwo zauważyć, że polega ono na utrzymywaniu wartości bezwzględnej pochodnej prędkości kątowej $\omega_o(t)$ na poziomie równym wartości bezwzględnej opóźnienia kątowego przy hamowaniu napędu wybiegiem swobodnym.

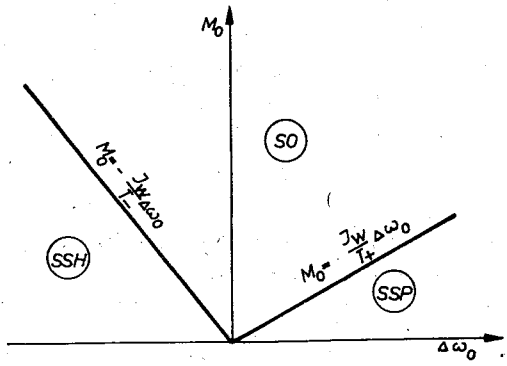
Zakres stosowania sterowania optymalnego ograniczony jest w praktyce warunkiem, by czas zmiany prędkości (w szczególności czas rozruchu lub hamowania) nie przekraczał pewnej ustalonej wartości granicznej. W myśl zależności (8) warunkowi $T^* \leqslant T$ odpowiada nierówność

$$\frac{M_o}{|\Delta\omega_o|} \geqslant \frac{J_w}{T}, \tag{10}$$

wyznaczająca granicę sterowań SS i SO w przestrzeni zmiennych $|\Delta\omega_o|$, M_o. Wobec mniejszego obciążenia silnika przy hamowaniu niż przy przyspieszaniu można dla sterowania suboptymalnego ustalić wartość $T = T_-$ przy $\Delta\omega < 0$ mniejszą od wartości $T = T_+$ przy $\Delta\omega_o > 0$. Wtedy granice sterowań optymalnego (SO), suboptymalnego przy przyspieszaniu (SSP) i suboptymalnego przy hamowaniu (SSH) wyznaczają półproste o równaniach:

$$\frac{M_o}{\Delta\omega_o} = \begin{cases} \dfrac{J_w}{T_+} & \text{dla} \quad \Delta\omega_o > 0, \\ -\dfrac{J_w}{T_-} & \text{dla} \quad \Delta\omega_o < 0, \end{cases} \tag{11}$$

co zilustrowano rysunkiem 1, pokazującym obszary SO, SSP i SSH na płaszczyźnie zmiennych $\Delta\omega_o$, M_o.



Rys. 1. Obszary sterowania na płaszczyźnie zmiennych $\Delta\omega_o$, M_o

3. SYNTEZA ENERGO-OPTYMALNEGO UKŁADU NAPĘDOWEGO Z SILNIKIEM PRĄDU STAŁEGO

Synteza energo-optymalnego układu napędowego polega na doborze silnika, przekładni oraz układów zasilania i sterowania tak, by dla zadanych zakresów prędkości kątowej oraz momentu oporowego obciążenia można było realizować sterowanie optymalne lub suboptymalne według omówionych zasad. Przedstawiony zostanie sposób sprowadzenia tego zadania do problemu programowania nieliniowego, za pomocą tak zwanej metody parametrów węzłowych.

Prawidłowa praca silnika napędowego wymaga spełnienia następujących warunków:

a) w stanach przejściowych ($\omega_o(t) = \text{var}$) wartości prędkości kątowej silnika ω, prądu twornika I i napięcia twornika U nie mogą przekraczać właściwych temu silnikowi wartości dopuszczalnych ω_{dop}, I_{dop} i U_{dop}; również czasy trwania stanów przejściowych, to jest T_+ i T_- nie mogą przekraczać zadanych wartości dopuszczalnych T_{+dop} i T_{-dop},

b) w stanach ustalonych ($\omega_o(t) = \text{const}$) wartości ω, I i U nie mogą przekraczać wartości znamionowych ω_n, I_n i U_n prędkości, prądu i napięcia silnika.

Zakładając $\omega_{dop} = \omega_n$, warunek ograniczenia prędkości kątowej silnika można na podstawie zależności (4) napisać w postaci

$$\frac{\omega_{o\,max}}{N} \leqslant \omega_n, \tag{12}$$

gdzie $\omega_{o\,max}$ — górny zakres prędkości kątowej urządzenia napędzanego.

Nierówność (12) obowiązuje dla każdego stanu pracy i rodzaju sterowania napędu.

Rozważając przykładowo warunek ograniczenia prądu twornika przy zmianie prędkości ω_o według formuły *SSP*, można na podstawie zależności (2) i (6) napisać

$$\frac{N}{k}\left(M_o + \frac{J_w \Delta\omega_o}{T_+}\right) \leqslant I_{dop}. \tag{13}$$

Lewa strona tej nierówności osiąga maksimum przy $\Delta\omega_o = \omega_{o\,max}$ i $M_o = M_{o\,max}$, gdzie $M_{o\,max}$ oznacza największą wartość momentu oporowego obciążenia. Wobec tego warunek (13) będzie zawsze spełniony jeśli

$$\frac{N}{k}\left(M_{o\,max} + \frac{J_w \omega_{o\,max}}{T_+}\right) \leqslant I_{dop}. \tag{14}$$

Nierówność (14) obowiązuje dla przypadku, gdy punkt o współrzędnych $\omega_{o\,max}$, $M_{o\,max}$ leży w obszarze *SSP*, to znaczy gdy

$$M_{o\,max} \leqslant \frac{J_w \omega_{o\,max}}{T_+}. \tag{15}$$

Jeśli punkt ($\omega_{o\,max}$, $M_{o\,max}$) leżałby w obszarze *SO*, to największa wartość prądu twornika wystąpiłaby przy sterowaniu optymalnym, zaś spełnienie warunku ograniczenia prądu gwarantowałaby nierówność

$$2\frac{N}{k} M_{o\,max} \leqslant I_{dop}. \tag{16}$$

Wreszcie, przy sterowaniu *SSH* największa wartość prądu występuje dla $\Delta\omega_o = -\omega_{o\,max}$ i $M_o = 0$, wobec czego konieczne jest spełnienie nierówności

$$\frac{NJ_w\omega_{o\,max}}{kT_-} \leqslant I_{dop}. \tag{17}$$

Nierówności (12), (14), (16) i (17) zwane są dalej nierównościami ograniczającymi. Analogiczne jak dla ograniczenia prądu rozważania przeprowadzono w odniesieniu do ograniczenia napięcia. Komplet nierówności ograniczających zestawiono w tyblicy 1.

W skład nierówności ograniczających wchodzą parametry silnika (np. R), przekładni (N), urządzenia napędzanego (np. $M_{o\,max}$) oraz warunków pracy napędu (np. T_{+dop}).

Tablica 1

Nierówności ograniczające energo-optymalnego układu napędowego z silnikiem prądu stałego

Lp.	Ograniczenie	Sterowanie	Nierówność ograniczająca
1	Prędkości kątowej silnika	dowolne	$\frac{\omega_{o\,max}}{N} \leqslant \omega_n$
2	Prądu twornika silnika	*SO*	$2\frac{N}{k} M_{o\,max} \leqslant I_{dop}$
3		*SSP*	$\frac{N}{k}\left(M_{o\,max}+\frac{J_w\omega_{o\,max}}{T_+}\right) \leqslant I_{dop}$
4		*SSH*	$\frac{NJ_w\omega_{o\,max}}{kT_-} \leqslant I_{dop}$
5		stan ustalony	$\frac{N}{k} M_{o\,max} \leqslant I_n$
6	Napięcia twornika silnika	*SO*	$\frac{k}{N}\omega_{o\,max}+2\frac{N}{k} RM_{o\,max} \leqslant U_{dop}$
7		*SSP*	$\frac{k}{N}\omega_{o\,max}+\frac{N}{k} R\left(M_{o\,max}+\frac{J_w\omega_{o\,max}}{T_+}\right) \leqslant U_{dop}$
8		*SSH*	ograniczenie zawsze spełnione
9		stan ustalony	$\frac{k}{N}\omega_{o\,max} \leqslant U_n$
10	Czasu przyspieszania	*SO*, *SSP*	$T_+ \leqslant T_{+dop}$
11	Czasu hamowania	*SO*, *SSH*	$T_- \leqslant T_{-dop}$

Uwagi

1. Warunek obszaru obowiązywania nierówności *SO*: $M_{o\,max}T_+ > J_w\omega_{o\,max}$
2. Warunek obszaru obowiązywania nierówności *SSP*: $M_{o\,max}T_+ \leqslant J_w\omega_{o\,max}$
3. Wypadkowy moment bezwładności: $J_w = J_o + \frac{J}{N^2}$

Część tych parametrów jest zadana z góry jako dane wyjściowe do projektowania napędu. Podstawiając do nierówności ograniczających parametry konkretnego silnika, dobranego wstępnie na podstawie na przykład jego mocy znamionowej, redukuje się ilość parametrów o nie ustalonych wartościach do kilku, na przykład N, T_+ i T_-. Parametry te będą dalej nazywane węzłowymi, gdyż ustalenie ich wartości pozwala na jednoznaczne określenie projektowanego układu. Nierówności ograniczające wyznaczają obszar dopuszczalny w przestrzeni parametrów węzłowych, w obrębie którego należy zlokalizować rozwiązanie optymalne, które w przypadku wspomnianych parametrów N, T_+ i T_- i w aspekcie zależności (7) i (9) odpowiada najmniejszej możliwej wartości N i największym możliwym wartościom T_+ i T_-. Ponieważ parametry T_+ i T_- nie występują razem w żadnej nierówności ograniczającej, więc powyższe zadanie programowania matematycznego można rozwiązać grafoanalitycznie na płaszczyznach T_+, N i T_-, N. Z uwagi na nieliniowy charakter ograniczeń (należy pamiętać, że $J_w = J_w(N)$) jest to zadanie programowania nieliniowego, łatwe wszakże do rozwiązania dzięki wspomnianej procedurze grafoanalitycznej.

Powierzchnie obszarów dopuszczalnych w układach współrzędnych (T_+, N) i (T_-, N) związane są z wielkością zastosowanego silnika. Oczywiste jest, że w odniesieniu do silnika „za małego" obszary dopuszczalne nie istnieją, zaś przewymiarowanie silnika wiąże się ze wzrostem obszarów dopuszczalnych. Dlatego przedstawione zadanie należy rozwiązywać alternatywnie dla kilku możliwych do zastosowania silników i wybrać ten, dla którego obszary dopuszczalne są najmniejsze. Przy idealnym dopasowaniu silnika do urządzenia napędzanego obszary dopuszczalne są zredukowane do pojedynczych punktów. Tak więc omówiona metoda parametrów węzłowych pozwala na optymalizację napędu zarówno z punktu widzenia eksploatacyjnego (minimalizacja energii), jak i inwestycyjnego (minimalizacja silnika napędowego).

4. STEROWANIE ENERGO-OPTYMALNE SILNIKA PRĄDU PRZEMIENNEGO

Układ napędowy z silnikiem asynchronicznym ze sterowaniem częstotliwościowym przy stałej wartości strumienia magnetycznego wykazuje formalne podobieństwo do omówionego układu z silnikiem prądu stałego. Stosunek prędkości poślizgowej ω_s do momentu M silnika można w praktyce, dla niewielkich zmian poślizgu s, uważać za stały i przedstawiać jako rezystancję mechaniczną R_s. Oznaczając przez E_s energię poślizgową, rozproszoną w tworniku silnika, można zatem napisać

$$E_s = R_s \int_0^T M^2(t)dt. \tag{18}$$

Na zasadzie omówionej analogii do układu z silnikiem prądu stałego sterowanie energo-optymalne silnika asynchronicznego można bezpośrednio opisać zależnościami (6) i (8), zaś podstawiając R_s w miejsce R i 1 zamiast k do wzorów (7) i (9) otrzymuje się

$$E_s^* = \left[N\left(M_o + \frac{J_w \Delta\omega_o}{T}\right)\right]^2 R_s T \tag{19}$$

oraz

$$E_s^{**} = [N(1+\operatorname{sign}\Delta\omega_o)]^2 R_s J_w M_o |\Delta\omega_o| . \tag{20}$$

5. SYNTEZA ENERGO-OPTYMALNEGO UKŁADU NAPĘDOWEGO Z SILNIKIEM PRĄDU PRZEMIENNEGO

Obok ograniczeń prędkości kątowej ω silnika oraz czasów T_+ i T_- regulacji prędkości, identycznych jak w przypadku napędu prądu stałego, przy syntezie układu z silnikiem asynchronicznym należy wziąć pod uwagę ograniczenia:

a) w stanach przejściowych — ograniczenia prądu stojana I_1, prądu wirnika I_2' oraz napięcia zasilającego U do wartości dopuszczalnych I_{1dop}, I_{2dop}' i U_{dop}; ponadto ograniczenie momentu obciążenia do wartości nie przekraczającej momentu rozruchowego M_r — przy przyspieszaniu oraz momentu krytycznego (maksymalnego) M_{kr} — przy hamowaniu,

b) w stanach ustalonych — ograniczenia prądów stojana i wirnika do wartości znamionowych I_{1n} i I_{2n}'.

Wykorzystując znany schemat zastępczy silnika asynchronicznego w postaci czwórnika typu T, uzyskuje się następujące zależności dotyczące wartości skutecznych prądów stojana i wirnika oraz prądu magnesującego I_μ, a także momentu silnika, momentu rozruchowego i momentu krytycznego:

$$I_1 = \frac{U}{\omega_1 L_o}\sqrt{\frac{\alpha\omega_1^2 s^2+1}{(\beta_1\omega_1^2+\beta_o)s^2+\gamma s+1}}, \tag{21}$$

$$I_2' = \frac{UL_\mu}{R_2' L_o}\frac{s}{\sqrt{(\beta_1\omega_1^2+\beta_o)s^2+\gamma_s+1}}, \tag{22}$$

$$I_\mu = \frac{U}{\omega_1 L_o}\sqrt{\frac{\partial\omega_1^2 s^2+1}{(\beta_1\omega_1^2+\beta_o)s^2+\gamma_s+1}}, \tag{23}$$

$$M = \frac{U^2}{\omega_1}\frac{\mu s}{(\beta_1\omega_1^2+\beta_o)s^2+\gamma_s+1}, \tag{24}$$

$$M_r = \frac{U^2}{\omega_1}\frac{\mu}{\beta_1\omega_1^2+\beta_o+\gamma+1},$$

$$M_{kr} = \pm\frac{U^2}{\omega_1}\frac{\mu}{2\sqrt{\beta_1\omega_1^2+\beta_o}\pm\gamma}, \tag{25}$$

gdzie:

$$\alpha = \left(\frac{L_2'+L_\mu}{R_2'}\right)^2, \tag{26}$$

$$\beta_1 = \frac{1}{R_2'}\left(L_2'+\frac{L_1 L_\mu'}{L_o}\right)^2, \tag{27}$$

$$\beta_o = \left(\frac{R_1(L_2' + L_\mu)}{R_2' L_o} \right)^2, \tag{28}$$

$$\gamma = 2 \frac{R_1}{R_2'} \left(\frac{L_\mu}{L_o} \right)^2, \tag{29}$$

$$\delta = \left(\frac{L_2'}{R_2'} \right)^2, \tag{30}$$

$$\mu = 3 \frac{\eta_m}{R_2'} \left(\frac{L_\mu}{L_o} \right)^2, \tag{31}$$

przy czym

$$L_o = L_1 + L_\mu, \tag{32}$$

zaś η_m oznacza współczynnik sprawności mechanicznej silnika, czyli stosunek momentu na wale do momentu elektromagnetycznego w szczelinie.

W odniesieniu do sterowania częstotliwościowego z zachowaniem stałej wartości głównego strumienia magnetycznego można napisać

$$I_\mu = I_{\mu n} = \frac{U_n}{\omega_{1n} L_o} \sqrt{\frac{\delta' s^2 + 1}{\beta' s^2 + \gamma_s + 1}} = \text{const}, \tag{33}$$

gdzie: $I_{\mu n}$ oznacza wartość znamionową prądu magnesującego, ω_{1n} — znamionową prędkość synchroniczną silnika, zaś $\delta' = \delta \omega_{1n}^2$, a $\beta' = \beta_1 \omega_{1n}^2 + \beta_o$.

Na podstawie zależności (33) prąd I_1 przy omawianym sterowaniu może być przedstawiony jako

$$I_1 = \frac{I_1}{I_\mu} I_{\mu n} = \frac{U_n}{\omega_{1n} L_o} \sqrt{\frac{\alpha \omega_s^2 + 1}{\delta \omega_s^2 + 1}} \sqrt{\frac{\delta' s^2 + 1}{\beta' s^2 + \gamma_s + 1}}. \tag{34}$$

Biorąc pod uwagę, że $\delta' \ll \beta'$ można przyjąć (co potwierdzają liczne przykłady praktyczne), że wartość prawego pierwiastka prawej strony zależności (34) dla dowolnych wartości poślizgu s osiąga maksimum bliskie jedności. Zatem ograniczenie wartości skutecznej prądu stojana można przedstawić w postaci

$$\frac{U_n}{\omega_{1n} L_o} \sqrt{\frac{\alpha \omega_s^2 + 1}{\delta \omega_s^2 + 1}} \leqslant I_{1dop}. \tag{35}$$

Zakładając, że rezystancja mechaniczna R_s silnika jest stała i równa stosunkowi znamionowej prędkości kątowej poślizgu ω_{sn} do momentu znamionowego M_n silnika, czyli

$$R_s = \frac{\omega_s}{M} = \frac{\omega_{sn}}{M_n} = \text{const}, \tag{36}$$

i wykorzystując równanie równowagi momentów układu przy sterowaniu energo-optymalnym

$$M = N(M_o + J_w \dot{\omega}_o) = N\left(M_o + \frac{J_w \Delta \omega_o}{T} \right), \tag{37}$$

można przedstawić prędkość kątową poślizgu jako

$$\omega_s = R_s N\left(M_o + \frac{J_w \Delta\omega_o}{T}\right). \tag{38}$$

Po podstawieniu zależności (38) do nierówności (35) otrzymuje się

$$\frac{U_n}{\omega_{1n} L_o}\sqrt{\frac{\alpha}{\delta} - \frac{\frac{\alpha}{\delta}-1}{\delta\left[R_s N\left(M_o + \frac{J_w \Delta\omega_o}{T}\right)\right]^2 + 1}} \leqslant I_{1dop}. \tag{39}$$

Nierówność powyższą można uprościć podstawiając wyrażenia (26) i (30) oraz biorąc pod uwagę, że $L_2' \ll L_\mu$ i $L_\mu + L_2' \approx L_o$. Po przekształceniu otrzymuje się

$$\frac{R_s N U_n}{R_2' \omega_{1n}}\left(M_o + \frac{J_w \Delta\omega_o}{T}\right) \leqslant I_{1dop}. \tag{40}$$

Lewa strona nierówności (40) osiąga maksimum przy $M_o = M_{o\,max}$ i $\Delta\omega_o = \omega_{o\,max}$ (to jest $T = T_+$); zatem ostatecznie nierówność ograniczająca dotycząca prądu stojana przy sterowaniu suboptymalnym przy przyspieszaniu (*SSP*) przyjmuje postać

$$\frac{R_s N U_n}{R_2' \omega_{1n}}\left(M_{o\,max} + \frac{J_w \omega_{o\,max}}{T_+}\right) \leqslant I_{1\,dop}. \tag{41}$$

W przypadku sterowania optymalnego (*SO*), w myśl zależności (8) otrzymuje się

$$2\,\frac{R_s N U_n M_{o\,max}}{R_2' \omega_{1n}} \leqslant I_{1\,dop}, \tag{42}$$

zaś w przypadku sterowania suboptymalnego przy hamowaniu (*SSH*)

$$\frac{R_s N U_n J_w \omega_{o\,max}}{R_2' \omega_{1n} T_-} \leqslant I_{1\,dop}. \tag{43}$$

Wyprowadzone w podobny sposób pozostałe nierówności ograniczające zestawiono w tablicy 2.

6. PRZYKŁADY OBLICZENIOWE

Celem zilustrowania przedstawionych rozważań przykładami obliczeniowymi przeprowadzono syntezę energo-optymalnego układu napędowego prądu stałego o następujących danych wyjściowych: $\omega_{o\,max} = 200$ rad/s, $M_{o\,max} = 100$ N · m, $J_o = 3$ kg · m², $T_{+dop} = 5$ s, $T_{-dop} = 2{,}5$ s. Rozpatrzono możliwość zastosowania silnika prądu stałego typu PXb64b produkcji krajowej, o mocy znamionowej 22 kW. Posiada on następujące parametry: $\omega_n = 314$ rad/s, $I_n = 111$ A, $U_n = 220$ V, $R = 0{,}182\ \Omega$, $k = 0{,}636$ N · m/A, $J = 0{,}11$ kg · m². Przyjęto $I_{dop} = 2{,}5$, $I_n = 278$ A, $U_{dop} = 1{,}1$ *V*, $U_n = 242$ V. Wykres parametrów węzłowych N, T_+, T_- pokazano na rysunku 2. Linie odpowiadające poszczególnym nierównościom ograniczającym oznaczono numeracją według tablicy 1. Linie ciągłe dotyczą parametrów N i T_+, zaś przerywane — N i T_-. Obszary dopuszczalne

Tablica 2

Nierówności ograniczające energo-optymalnego układu napędowego z silnikiem prądu przemiennego

Lp.	Ograniczenie	Sterowanie	Nierówność ograniczająca
1	Prędkości kątowej	dowolne	$\frac{\omega_{o\,max}}{N} \leqslant \omega_n$
2	Prądu stojana silnika	*SO*	$2\frac{R_s N U_n M_{o\,max}}{R'_2 \omega_{1n}} \leqslant I_{1dop}$
3		*SSP*	$\frac{R_s N U_n}{R'_2 \omega_{1n}}\left(M_{o\,max}+\frac{J_w \omega_{o\,max}}{T_+}\right) \leqslant I_{1dop}$
4		*SSH*	$\frac{R_s N U_n J_w \omega_{o\,max}}{R'_2 \omega_{1n} T_-} \leqslant I_{1dop}$
5		stan ustalony	$\frac{R_s N U_n M_{o\,max}}{R'_2 \omega_{1n}} \leqslant I_{1n}$
6	Prądu wirnika silnika	*SO*	$2NM_{o\,max}\sqrt{\frac{R_s}{3R'_2}} \leqslant I'_{2dop}$
7		*SSP*	$N\sqrt{\frac{R_s}{3R'_2}}\left(M_{o\,max}+\frac{J_w \omega_{o\,max}}{T_+}\right) \leqslant I'_{2dop}$
8		*SSH*	$\frac{NJ_w \omega_{o\,max}}{T_-}\sqrt{\frac{R_s}{3R'_2}} \leqslant I'_{2dop}$
9		stan ustalony	$NM_{o\,max}\sqrt{\frac{R_s}{3R'_2}} \leqslant I'_{2n}$
10	Napięcia zasilającego	dowolne	$\frac{\omega_{o\,max}}{N\omega_{1n}} U_n \sqrt{\frac{A\left(\frac{\omega_{o\,max}}{N}\right) B\left(\frac{\omega_{o\,max}}{N}\right)}{C\left(\frac{\omega_{o\,max}}{N}\right) D\left(\frac{\omega_{o\,max}}{N}\right)}} \leqslant U_{dop}$
11		*SO*	$M_{o\,max} \leqslant \frac{\mu\omega_{1\,min}}{2N}\left(\frac{U_n}{\omega_{1n}}\right)^2 \frac{\delta'+1}{(\beta'+\gamma+1)(\delta'\omega^2_{1\,min}+1)}$
12	Momentu obciążenia	*SSP*	$M_{o\,max}+\frac{J_w \omega_{o\,max}}{T_+} \leqslant \frac{\mu\omega_{1\,min}}{N}\left(\frac{U_n}{\omega_{1n}}\right)\times \times\frac{\delta'+1}{(\beta'+\gamma+1)(\delta'\omega^2_{1\,min}+1)}$

Tabela 2 (c.d.)

Lp.	Ograniczenie	Sterowanie	Nierówność ograniczająca
13		*SSH*	$\frac{J_w\omega_{o\,max}}{T_-} \leqslant \frac{\mu\omega_{1\,min}}{N}\left(\frac{U_n}{\omega_{1n}}\right)^2 \frac{\beta_1\omega_{1\,min}^2+\beta_o+\delta'}{\left(\frac{\beta'}{\sqrt{\beta_1\omega_{1\,min}^2+\beta_o}}-\gamma+\sqrt{\beta_1\omega_{1\,min}^2+\beta_o}\right)} \times \times \frac{1}{(\beta_1+\delta)\omega_{1\,min}^2+\beta_o}$
14	Czasu przyspieszenia	*SO*, *SSP*	$T_+ \leqslant T_{+dop}$
15	Czasu hamowania	*SO*, *SSH*	$T_- \leqslant T_{-dop}$

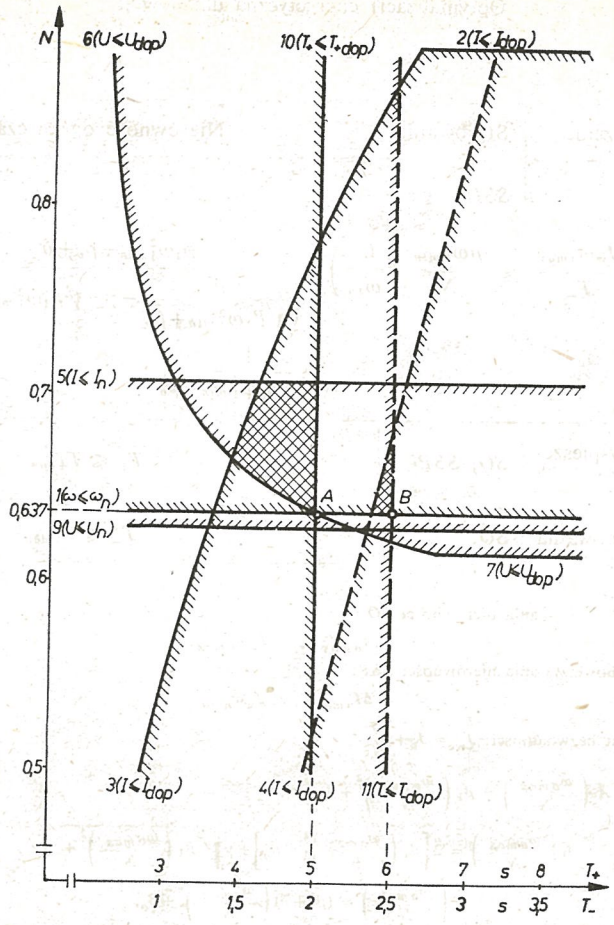
Uwagi

1. Warunek obszaru obowiązywania nierówności *SO*:
$$M_{o\,max}T_+ > J_w\omega_{o\,max}$$
2. Warunek obszaru obowiązywania nierówności *SSP*:
$$M_{o\,max}T_+ \leqslant J_w\omega_{o\,max}.$$
3. Wypadkowy moment bezwładności: $J_w = J_o + \frac{J}{N^2}$
4. W nierówności 10: $A\left(\frac{\omega_{o\,max}}{N}\right) = \beta_1\left(\frac{\omega_{o\,max}}{N}\right)^2 + \beta_o + \delta'$,
$$B\left(\frac{\omega_{o\,max}}{N}\right) = 2\left[\beta_1\left(\frac{\omega_{o\,max}}{N}\right)^2 + \beta_o\right] + \gamma\sqrt{\beta_1\left(\frac{\omega_{o\,max}}{N}\right)^2 + \beta_o},$$
$$C\left(\frac{\omega_{o\,max}}{N}\right) = (\beta_1+\delta)\left(\frac{\omega_{o\,max}}{N}\right)^2 + \beta_o,$$
$$D\left(\frac{\omega_{o\,max}}{N}\right) = \beta_1\left(\frac{\omega_{o\,max}}{N}\right)^2 + \beta_o + \beta' + \gamma\sqrt{\beta_1\left(\frac{\omega_{o\,max}}{N}\right)^2 + \beta_o}$$
5. W nierównościach (11)÷(13) $\omega_{1\,min}$ oznacza dolną granicę regulacji synchronicznej prędkości kątowej ω_1

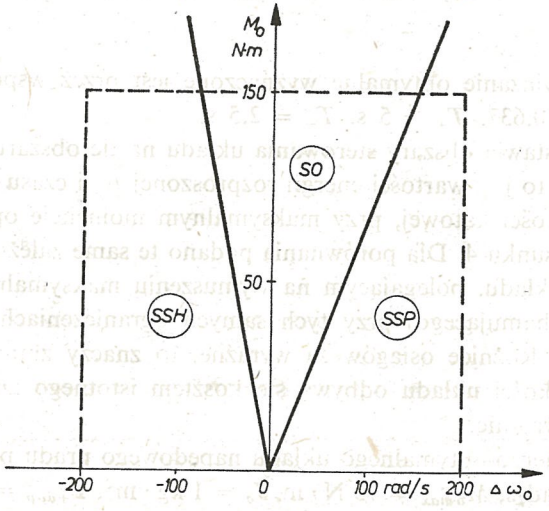
zakreskowano. Rozwiązanie optymalne wyznaczone jest przez współrzędne punktów *A* i *B*, to znaczy $N = 0{,}637$, $T_+ = 5$ s, $T_- = 2{,}5$ s.

Rysunek 3 przedstawia obszary sterowania układu na tle obszaru jego pracy. Podstawowe osiągi układu, to jest wartości energii rozproszonej E_R i czasu regulacji T w funkcji przyrostu $\Delta\omega_o$ prędkości kątowej, przy maksymalnym momencie oporowym obciążenia, przedstawiono na rysunku 4. Dla porównania podano te same zależności przy sterowaniu czaso-optymalnym układu, polegającym na wymuszeniu maksymalnego momentu napędowego (względnie hamującego) przy tych samych ograniczeniach co przy sterowaniu energo-optymalnym. Różnice osiągów są wyraźne, to znaczy zmniejszenie strat energii przy zmianach prędkości układu odbywa się kosztem istotnego zmniejszenia szybkości jego działania i odwrotnie.

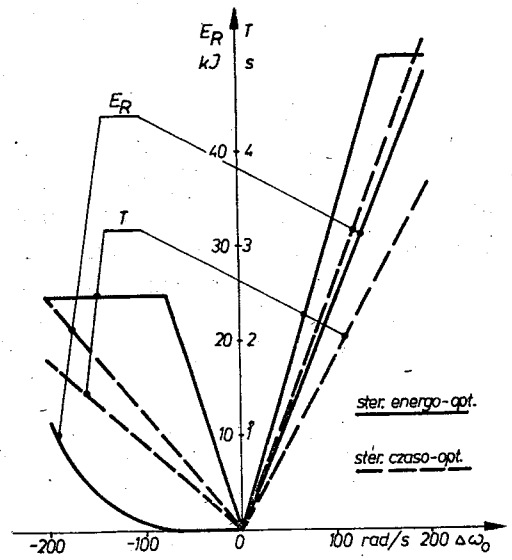
W przykładzie energo-optymalnego układu napędowego prądu przemiennego przyjęto dane: $\omega_{o\,max} = 200$ rad/s, $M_{o\,max} = 75$ N · m, $J_o = 1$ kg · m^2, $T_{+dop} = 3{,}2$ s, $T_{-dop} = 1{,}4$ s. Do obliczeń wykorzystano parametry silnika asynchronicznego pierścieniowego typu



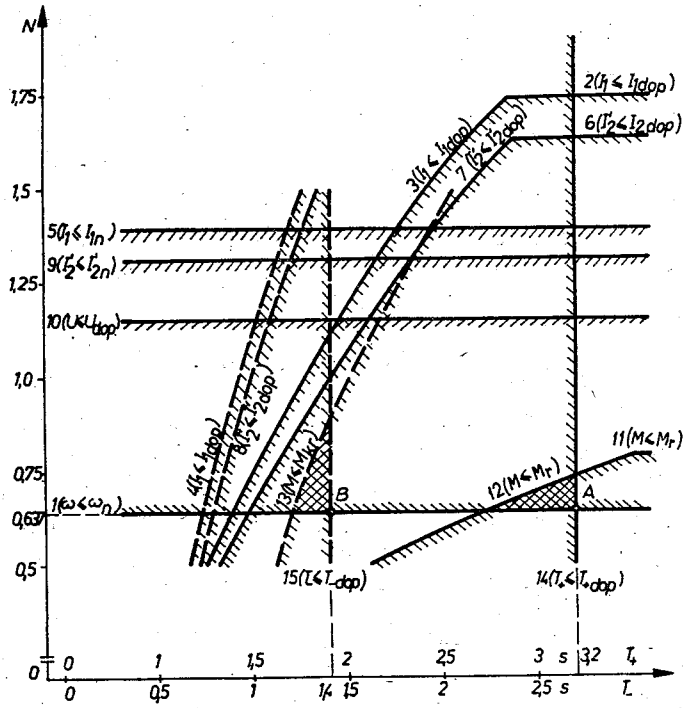
Rys. 2. Wykres parametrów węzłowych przykładowego układu napędowego z silnikiem prądu stałego



Rys. 3. Obszary sterowania przykładowego układu napędowego z silnikiem prądu stałego

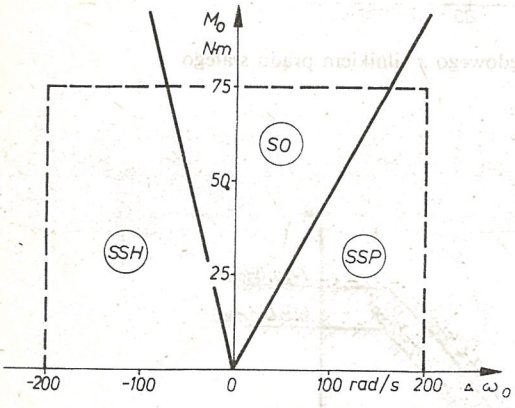


Rys. 4. Osiągi przykładowego układu napędowego z silnikiem prądu stałego

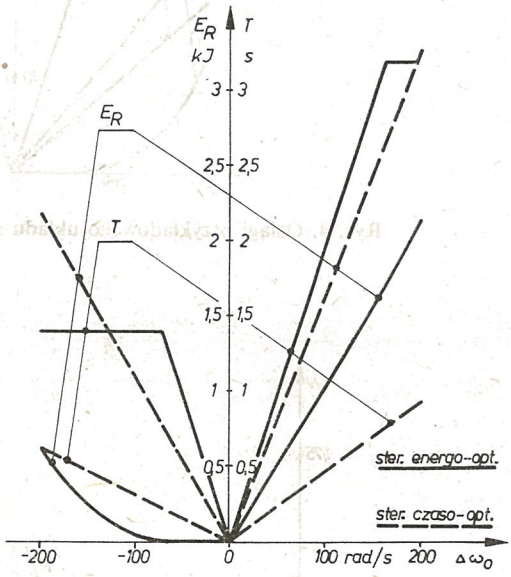


Rys. 5. Wykres parametrów węzłowych przykładowego układu napędowego z silnikiem prądu przemiennego

SUDf 180-L4 (bez dodatkowej rezystancji wirnika). Parametry te równają się: $U_n = 220$ V, $U_{dop} = \sqrt{2}\,U_n = 318$ V, $I_{1n} = 29{,}2$ A, $I'_{2n} = 25{,}3$ A, $\omega_{1n} = 157$ rad/s, $\omega_{1dop} = 2\omega_{1n} = 314$ rad/s, $s_n = 0{,}037$, $M_n = 99$ N · m, $R'_2 = 0{,}295\ \Omega$, $L_\mu = 0{,}1585$ H, $L_o = 0{,}1613$ H, $\alpha = 0{,}31$ s^2, $\beta_1 = 8{,}39 \cdot 10^{-4}$ s^2, $\beta_o = 0{,}344$ s^{-1}, $\gamma = 1{,}113$, $\delta = 3{,}8 \cdot 10^{-4}$ s^2, $\beta' = 21{,}05$, $\delta' = 9{,}38$, $\mu = 9{,}33$ s, $J = 0{,}196$ kg · m^2. Przyjęto $I_{1dop} = 2{,}5\ I_{1n} = 73$ A, $I'_{2dop} = 2{,}5\ I'_{2n} = 63$ A, $\omega_{1min} = 0{,}1\ \omega_{1n} = 15{,}7$ rad/s. Wykres parametrów węzłowych przedstawiono na rysunku 5. Rozwiązanie optymalne stanowią wartości $N = 0{,}637$, $T_+ = 3{,}2$ s, $T_- = 1{,}4$ s. Obszary sterowania oraz osiągi układu przy sterowaniu energo-optymalnym i częstotliwościowym sterowaniu czaso-optymalnym przedstawiono odpowiednio na rysunkach 6 i 7.



Rys. 6. Obszary sterowania przykładowego układu napędowego z silnikiem prądu przemiennego



Rys. 7. Osiągi przykładowego układu napędowego z silnikiem prądu przemiennego

7. UWAGI KOŃCOWE

Przedstawione rozważania wskazują na możliwość redukcji strat w układach napędu elektrycznego drogą sterowania energo-optymalnego w funkcji momentu oporowego obciążenia. Współczesne układy regulacji cyfrowej z zastosowaniem mikroprocesorów zapewniają praktyczne możliwości takiego sterowania [1, 2, 3]. Mikroprocesor działający w czasie rzeczywistym wyliczałby optymalną trajektorię prędkości kątowej napędu według zależności (6) i (8), zaś układ regulacji nadążnej realizowałby zadany przebieg. Układ taki wymagałby ciągłego pomiaru momentu obciążenia, co wszakże jest już realizowane w wielu układach praktycznych [4].

Tendencja wzrostowa stosunku kosztów energii do kosztów oprzyrządowania uzasadnia wprowadzenie rozwiązań energo-optymalnych do praktyki przemysłowej, zwłaszcza w zakresie układów dużej mocy oraz urządzeń o ograniczonym zasilaniu, np. w robotyce,

technice wojskowej, kosmicznej itp. Zasady sterowania energo-optymalnego obowiązują przy tym zarówno proste układy napędowe z pojedynczym silnikiem, jak i układy złożone, gdzie każdy silnik z osobna winien być sterowany energo-optymalnie, stosownie do przypadającego nań obciążenia.

Zaproponowana metoda parametrów węzłowych jest wygodnym i elastycznym narzędziem syntezy układu napędowego, zwłaszcza jej pierwszego etapu, polegającego na skonkretyzowaniu elementów elektromechanicznych napędu. Zastosowanie współczesnych technik grafiki komputerowej wybitnie upraszcza zagadnienie wyznaczenia obszarów dopuszczalnych na płaszczyźnie parametrów węzłowych. Uzyskana w ten zalgorytmizowany sposób wstępna wersja układu napędowego stanowi dogodny punkt wyjściowy do dalszej, bardziej precyzyjnej analizy na gruncie współczesnych metod opisu układów dynamicznych.

Wykorzystanie prostych modeli matematycznych elementów układu napędowego, w szczególności równań stanu quasi-ustalonego silników, uzasadnione jest rozważaniem strony ilościowej, a nie jakościowej, procesu przetwarzania energii. Oczywistym jest fakt, potwierdzony zależnością (9), że znaczące oszczędności energii drogą energo-optymalnego sterowania napędu przy zmianach prędkości mogą być uzyskane tylko w układach o dużym momencie bezwładności, w których mechaniczna stała czasu napędu znacznie przewyższa stałą elektryczną. Energia rozproszona w elektrycznych stanach przejściowych stanowi więc drobny ułamek strat energii odpowiadających rozważanym stanom quasi-ustalonym.

BIBLIOGRAFIA

1. A. Kusko, *Application of microprocessors to AC and DC electric motor drive systems*, IEEE Conference Record, IAS-77, New York 1977.
2. T. Konishi i in., *An application technology of microprocessor to adjustable speed motor drives*, IEEE Conference Record, IAS-78, New York 1978.
3. S.J. Lukus, *Microprocessor control of a DC drive*, IEEE Conference Record, IAS-79, New York 1979.
4. A.B. Plunkett, *Direct torque and flux regulation in a PWM inverter-induction motor drive*, IEEE Trans. Ind. Appl., Vol. IA-13, nr 2, Mar./Apr. 1977.

A. TRZYNADLOWSKI

ENERGY OPTIMIZATION OF ELECTRICAL DRIVE SYSTEMS WITH CONTROLLED ANGULAR VELOCITY

Summary

The paper presents control and synthesis algorithms of *DC* and *AC* drives characterized by minimum energy dissipation while changing angular velocity. The, so called, nodal parameters method is applied for synthesis of energy optimal drive systems. The method relies on the mathematical theory of non-linear programming. Theoretical considerations are illustrated by computational examples.

A. TRZYNADLOWSKI

OPTIMALISATION ÉNERGÉTIQUE DES SYSTÈMES DE COMMANDE ÉLECTRIQUE À VITESSE ANGULAIRE RÉGLABLE

Résumé

On a présenté des algorithmes de commande et de synthèse des systèmes de commande à courant continu et alternatif dont la dissipation de l'énergie est minimale au cours des changements de vitesse angulaire. On a appliqué à la synthèse des systèmes de commande énergétiques optimaux la méthode dite des paramètres de noeuds, qui constitute une forme de programmation mathématique non-linéaire. Les considérations théoriques ont été illustrées par des exemples de calculs.

A. TRZYNADLOWSKI

ENERGETISCHE OPTIMIERUNG DER ELEKTROANTRIEBSYSTEME MIT EINSTELLBARER WINKELGESCHWINDIGKEIT

Zusammenfassung

Dargestellt werden die Algorithmen zur Steuerung und Synthese von Antriebssystemen des Gleich- und Wechselstroms mit minimaler Energiestreuung bei veränderlicher Winkelgeschwindigkeit. Für die Synthese der im Hinblick auf die Energie optimalen Antriebssysteme wird das sog. Verfahren der Knotenparameter angewendet, das eine Form der nichtlinearen mathematischen Programmierung darstellt. Die theoretischen Überlegungen werden anhand von Rechenbeispielen veranschaulicht.

А. ТШИНАДЛЁВСКИ

ЭНЕРГЕТИЧЕСКАЯ ОПТИМИЗАЦИЯ СИСТЕМ ЭЛЕКТРОПРИВОДА С РЕГУЛИРУЕМОЙ УГЛОВОЙ СКОРОСТЬЮ

Резюме

Представлены алгоритмы управления и синтеза приводных систем постоянного и переменного токов с минимальным рассеянием энергии при изменениях угловой скорости. Для синтеза энергооптимальных приводных систем применён так называемый метод узловых параметров, представляющий собой форму математического нелинейного программирования. Теоретические рассуждения проиллюстрированы расчётными примерами.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 469—479

621.313.17

Wpływ komutacji w układzie prostownikowym na moc czynną i graniczną wartość prędkości samochodowej prądnicy prądu przemiennego

JERZY MORAWSKI (KIELCE)

Instytut Automatyki Politechniki Świętokrzyskiej

Otrzymano 15.2.1984

Przedstawiono analityczną metodę wyznaczania mocy czynnej oraz dopuszczalnej prędkości kątowej samochodowej prądnicy prądu przemiennego w różnych warunkach pracy, uwzględniając zjawisko komutacji zaworów w prostowniku maszyny.

1. WPROWADZENIE

Samochodowe prądnice prądu przemiennego, zwane przez producentów alternatorami, stanowią podstawowy element elektrycznego wyposażenia pojazdów samochodowych. Ich produkcja zwiększa się każdego roku. Rosną też stawiane im wymagania, zwłaszcza, że obserwuje się tendencję użytkowników i producentów pojazdów do zwiększania liczby odbiorników energii elektrycznej poprawiających komfort i bezpieczeństwo jazdy. Doskonalenie konstrukcji tych maszyn winno pociągać za sobą potrzebę dokładniejszego określania ich parametrów oraz własności eksploatacyjnych już na etapie projektowania, a nie na podstawie badań prototypu.

Ulepszanie metod projektowania alternatorów polega obecnie na udoskonalaniu analizy rozkładu pola elektromagnetycznego w obwodach magnetycznych prądnicy i wyznaczaniu na tej podstawie optymalnych kształtów stojana i wirnika. W tej dziedzinie notuje się duże osiągnięcia, czego przykładem są prace [5, 7].

Doświadczenia projektantów i producentów alternatorów dowodzą jednak, że zaprojektowanie obwodu magnetycznego i obwodów elektrycznych, nawet w sposób optymalny z punktu widzenia wykorzystania obwodu magnetycznego i kształtu przebiegów prądów i napięć w uzwojeniach maszyny, nie gwarantuje jeszcze poprawności pracy prądnicy i często uzyskiwane parametry eksploatacyjne różnią się znacznie od założonych. Bezpośrednią tego przyczyną jest fakt, że stosowane metody nie uwzględniają jeszcze, lub uwzględniają w minimalnym stopniu [3, 4, 5], specyfiki współpracy pomiędzy podstawowymi elementami składowymi alternatora — synchroniczną prądnicą prądu przemiennego i prostownikiem.

Projekt samochodowej prądnicy prądu przemiennego wykonuje się najczęściej dla ściśle określonych warunków pracy prądnicy, takich jak prędkość kątowa, prąd znamionowy, napięcie znamionowe. W trakcie normalnego użytkowania w układzie elektrycznym samochodu prądnica pracuje jednak w warunkach często znacznie odbiegających od znamionowych. Dotyczy to zwłaszcza zmian prędkości kątowych, które w nowoczesnych konstrukcjach osiągają $\omega_{min} : \omega_{max} = 1:20$.

Duża reaktancja rozproszenia uzwojeń fazowych stojana alternatora, zmieniająca swoją wartość w szerokim zakresie, proporcjonalnie do zmian prędkości kątowej, sprawia, że zjawisko komutacji pomiędzy zaworami układu prostownikowego może zachodzić w trzech etapach [6]. Ostatni z nich jest stanem awaryjnym pracy prądnicy, w czasie którego współpracujący z nią akumulator ulega szybkiemu rozładowaniu. Na podstawie stosowanych dotychczas metod projektowych i sposobów doboru alternatora nie da się wyznaczyć prędkości prądnicy, przy której następuje przejście układu prostownikowego w trzeci etap komutacji. Podawane przez producentów w katalogach wartości graniczne prędkości wynikają przeważnie z mechanicznej wytrzymałości konstrukcji prądnicy.

Zjawisko komutacji sprawia dodatkowo, że wraz ze zmianą parametrów komutacji (kąta komutacji γ i kąta opóźnienia komutacji α_o) zmienia się kąt fazowy pomiędzy pierwszą harmoniczną napięcia indukowanego i prądem fazowym prądnicy. Wynika z tego, że przy zaprojektowaniu maszyny na określoną moc pozorną, można będzie uzyskać z niej moc czynną o różnej wartości. Znajomość tej mocy w różnych warunkach pracy prądnicy jest bardzo istotna, chodzi bowiem o to, by w trakcie eksploatacji była ona zdolna do oddania mocy czynnej znamionowej w jak najszerszym zakresie prędkości.

Omawiane zjawiska występują szczególnie wyraźnie wraz ze wzrostem mocy maszyny. Zgodnie z przewidywanym rozwojem przemysłu motoryzacyjnego, w najbliższych latach pojazdy samochodowe będą wyposażone w prądnicę o mocy około 1 kW pracującą z prędkościami do 2000 rad $\cdot$ s^{-1}. Istnieje więc uzasadniona potrzeba weryfikacji obliczeń.

Niniejsza praca przedstawia analityczny sposób wyznaczenia mocy czynnej oraz granicznej prędkości kątowej prądnicy. Przedstawione wzory są również przydatne przy obliczeniach sprawdzających w trakcie doboru alternatora do konkretnego układu.

2. ZAKRES ZMIAN MOCY CZYNNEJ ALTERNATORA

Podstawą do obliczenia głównych wymiarów geometrycznych obwodów magnetycznych alternatora jest znajomość mocy pozornej (zwanej również wewnętrzną mocą prądu przemiennego [4]) niezbędnej do wytworzenia znamionowej wielkości mocy czynnej zużywanej przez odbiorniki energii elektrycznej przy znamionowej prędkości kątowej prądnicy. W oparciu o jej wartość oraz po przyjęciu pewnych stałych parametrów materiałowych i konstrukcyjnych, wyznacza się według jednej z metod, najczęściej [4], średnicę twornika D i długość pakietu L.

Moc pozorną S oblicza się według wzoru:

$$S = 3\left(\frac{P_N}{U_N} + I_{wN}\right) k_I \left(\frac{U_\% + 100}{100}\right) \frac{0{,}75}{\sqrt{3}} (U_N + 2sU_D), \qquad (1)$$

gdzie:

P_N — moc znamionowa prądu stałego alternatora,

U_N — napięcie znamionowe alternatora,

k_I — współczynnik wyprostowania prądu, $k_I = \frac{I_f}{I_N}$,

I_N — znamionowa średnia wartość prądu wyprostowanego,

I_f — skuteczna wartość prądu fazowego alternatora przy znamionowym obciążeniu i znamionowej prędkości,

$U_{\%}$ — procentowy spadek napięcia na impedancji fazowej,

s — liczba diod w gałęzi prostownika,

U_D — spadek napięcia na diodzie prostownika,

I_{wN} — znamionowy prąd wzbudzenia.

We wzorze (1) współczynniki k_I i $U_{\%}$ przyjmuje się na podstawie badań porównawczych różnych typów alternatorów w granicach $k_I = 0{,}81 \div 0{,}75$, $U_{\%} = 50 \div 100$, przy czym nie uwzględnia się żadnej korelacji pomiędzy nimi.

W taki sposób liczony obwód magnetyczny prądnicy o wymiarach D, L, zapewnić ma oddawanie przez alternator mocy P_N również dla wszystkich prędkości kątowych większych od znamionowej. Jest to uogólnienie błędne. W takt zmian prędkości kątowej alternatora zmieniają się bowiem parametry komutacyjne (kąt komutacji i kąt opóźnienia komutacji) układu prostownikowego. Ulega zatem zmianom kąt fazowy pomiędzy pierwszą harmoniczną napięcia indukowanego w uzwojeniach fazowych i pierwszą harmoniczną prądu fazowego [6]. Ponieważ zaprojektowany obwód elektromagtnetyczny alternatora może przenieść stałą, obliczoną wcześniej, wartość mocy pozornej S, znaczy to, że wraz ze zmianą kąta fazowego ulega zmianie wartość mocy czynnej, jaką można z tego obwodu uzyskać. Maleje ona ze wzrostem prędkości kątowej prądnicy i teoretycznie dla wszystkich prędkości większych od znamionowej otrzymana moc czynna będzie mniejsza od znamionowej. Różnica pomiędzy nimi musi być wtedy pobrana z równolegle pracującego z prądnicą akumulatora, co prowadzi do jego szybkiego rozładowania.

Przyjmując założenie, że układ prostownikowy alternatora w całym zakresie zmian obciążeń i prędkości kątowych nie wchodzi w trzeci etap komutacji, skuteczne wartości napięcia indukowanego w uzwojeniu fazowym E_f i prądu fazowego I_f wyrazić można zależnościami [6]:

$$E_f = \frac{\pi U_{al}}{3\sqrt{6}F}, \qquad I_f = I_{al}\sqrt{\frac{2}{3} - \frac{\gamma}{3\pi}}. \tag{2}$$

Moc pozorna S jest wtedy równa

$$S = \frac{\pi U_{al} I_{al}}{\sqrt{6}F}\sqrt{\frac{2}{3} - \frac{\gamma}{3\pi}} = \frac{\pi R_o I_{al}^2}{\sqrt{6}F}\sqrt{\frac{2}{3} - \frac{\gamma}{3\pi}}, \tag{3}$$

$$S = P_{al}\frac{\sqrt{\frac{2}{3} - \frac{\gamma}{3\pi}}}{\sqrt{6}F}\pi,$$

gdzie:

U_{al} — średnia wartość napięcia wyprostowanego alternatora,

I_{al} — średnia wartość prądu wyprostowanego alternatora,

P_{al} — moc czynna oddawana przez alternator, $P_{al} = I_{al} U_{al}$,

R_o — rezystancja zastępcza obciążenia, $R_o = \dfrac{U_{al}}{I_{al}}$,

γ — kąt komutacji zaworów prostownika,

F — funkcja bezwymiarowa, odzwierciedlająca wpływ komutacji na przebiegi elektryczne.

Funkcję F wyraża wzór

$$
\begin{aligned}
F = {} & \frac{R_o}{\sqrt{\left[R_o + \frac{2}{3}(R_f + R_d)\right]^2 + \left(\frac{3}{2} L_f \omega\right)^2}} \left\{ \frac{\sin\left(\pi \frac{2}{3} - \psi_k\right)}{\operatorname{ctg} \varphi_k} (e^{-\gamma \operatorname{ctg} \varphi_k} - 1) + \right. \\
& - 2 \sin \frac{\gamma}{2} \sin\left(\frac{\gamma}{2} + \frac{2}{3}\pi + \psi_k\right) - \sin \frac{\gamma}{2} \sin\left(\frac{\gamma}{2} - \psi_k\right) + \frac{\sin \psi_k}{\operatorname{ctg} \varphi_k} (1 - e^{-\gamma \operatorname{ctg} \varphi_k}) + \\
& + \left[\sin\left(\gamma + \frac{2\pi}{3} - \psi_k\right) - \sin\left(\frac{2}{3}\pi - \psi_k\right) e^{-\gamma \operatorname{ctg} \varphi_k} + \frac{1}{2} \sin(\gamma - \psi_k) + \right. \\
& \left. + \frac{1}{2} \sin \psi_k e^{-\gamma \operatorname{ctg} \varphi_k} \right] \left[\frac{e^{-\left(\pi \frac{1}{3} - \gamma\right) \operatorname{ctg} \varphi}}{\operatorname{ctg} \varphi_k} (1 - e^{-\gamma \operatorname{ctg} \varphi_k}) + \right. \\
& \left. \left. + \frac{1}{\operatorname{ctg} \varphi} \left(1 - e^{-\left(\frac{\pi}{3} - \gamma\right) \operatorname{ctg} \varphi}\right) \right] \frac{1}{1 - e^{-\frac{\pi}{3} \operatorname{ctg} \varphi + \gamma (\operatorname{ctg} \varphi - \operatorname{ctg} \varphi_k)}} \right\} + \\
& + \frac{R_o}{\sqrt{[R_o + 2(R_f + R_d)]^2 + 4\omega^2 L_f^2}} \left\{ \frac{\sin\left(\gamma + \frac{\pi}{3} \mid \psi\right)}{\operatorname{ctg} \varphi} \left(e^{-\left(\frac{\pi}{3} - \gamma\right) \operatorname{ctg} \varphi} - 1 \right) + \right. \\
& - 2 \sin\left(\frac{\gamma}{2} - \psi + \frac{\pi}{2}\right) \sin\left(\frac{\gamma}{2} - \frac{\pi}{6}\right) + \left[\sin\left(\frac{2}{3}\pi - \psi\right) + \right. \\
& \left. - \sin\left(\gamma + \frac{\pi}{3} - \psi\right) e^{-\left(\frac{\pi}{3} - \gamma\right) \operatorname{ctg} \varphi} \right] \left[\frac{e^{-\gamma \operatorname{ctg} \varphi_k}}{\operatorname{ctg} \varphi} \left(1 - e^{-\left(\frac{\pi}{3} - \gamma\right) \operatorname{ctg} \varphi}\right) + \right. \\
& \left. \left. + \frac{1}{\operatorname{ctg} \varphi_k} (1 - e^{-\gamma \operatorname{ctg} \varphi_k}) \right] \frac{1}{1 - e^{-\frac{\pi}{3} \operatorname{ctg} \varphi + \gamma (\operatorname{ctg} \varphi - \operatorname{ctg} \varphi_k)}} \right\}, \qquad (4)
\end{aligned}
$$

gdzie:

$$
\begin{aligned}
\varphi &= \arccos \frac{R_o + 2(R_f + R_d)}{\sqrt{[R_o + 2(R_f + R_d)]^2 + 4\omega^2 L_f^2}}, \\
\varphi_f &= \arccos \frac{R_f + R_d}{\sqrt{(R_f + R_d)^2 + L_f^2 \omega^2}},
\end{aligned} \qquad (5)
$$

$$\varphi_k = \arccos \frac{R_o + \frac{3}{2}(R_f + R_d)}{\sqrt{\left[R_o + \frac{3}{2}(R_f + R_d)\right]^2 + \left(\frac{3}{2} L_f \omega\right)^2}}, \tag{5}$$

$$\psi_k = \varphi_k - \alpha_o,$$

$$\psi = \varphi - \alpha_o,$$

α_o — kąt opóźnienia komutacji,

R_f — rezystancja uzwojenia fazowego alternatora,

R_d — rezystancja diody prostownika,

L_f — indukcyjność rozproszenia fazy alternatora.

Przyjmując stałą wartość mocy pozornej S oddawanej przez zaprojektowaną maszynę oraz posługując się obliczonymi w trakcie prac projektowych parametrami R_f, L_f, R_d, można w oparciu o wzór (3) wyznaczyć moc czynną możliwą do uzyskania z prądnicy, przy różnych prędkościach kątowych i różnych obciążeniach. Wygodnie posługiwać się stosunkiem:

$$\frac{P_{al}}{S} = k_p,$$

zwanym często współczynnikiem wyprostowania mocy [4]. Współczynnik ten przyjmowany był dotychczas jako stały. Wyznaczyć go można w następujący sposób.

1. Dla zadanej wielkości obciążenia prądnicy R_o, prędkości kątowej ω_p, parametrów R_f, L_f, R_d wyznacza się parametry komutacyjne γ, α_o. Dokładny sposób wyznaczenia tych parametrów podano w [6], dlatego nie będzie tu przytaczany.

2. Posługując się parametrami R_f, L_f, R_d, R_o, ω_p, γ, α_o obliczyć należy F a następnie współczynnik k_p z przekształconego wzoru (3)

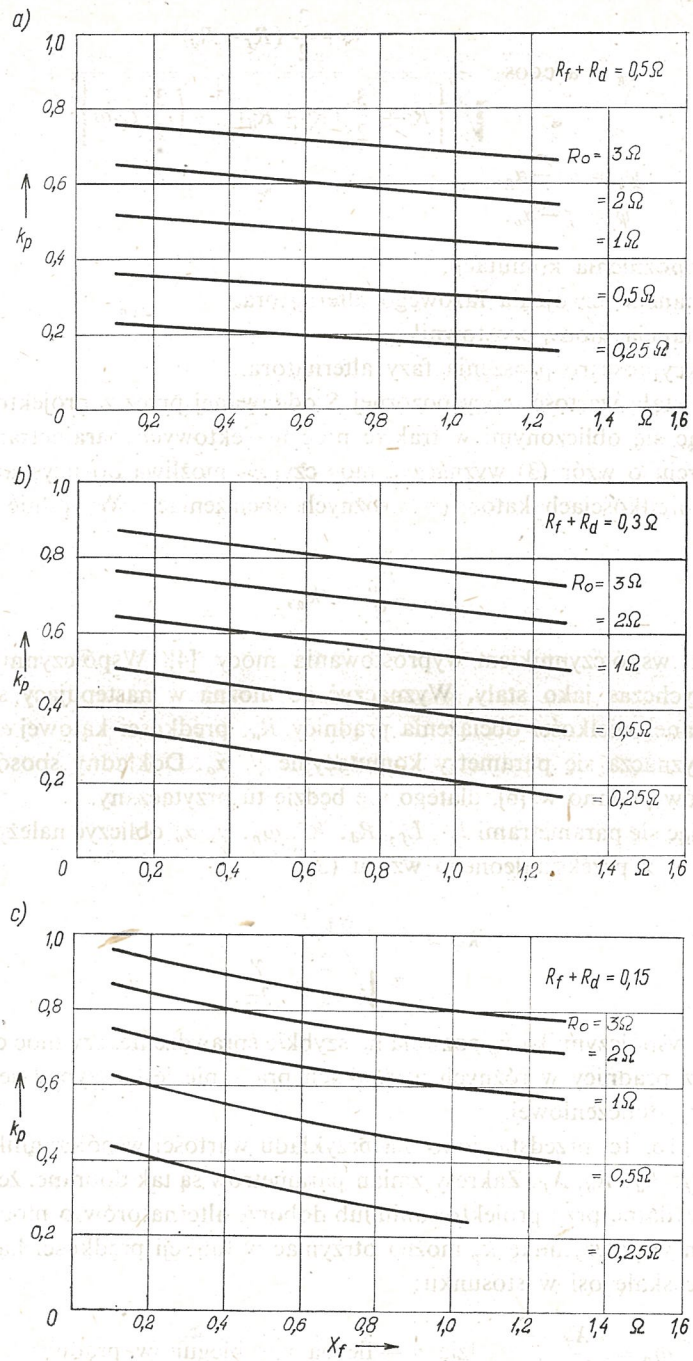
$$k_p = \frac{6\sqrt{F}}{\pi \sqrt{\frac{2}{3} - \frac{\gamma}{3\pi}}}. \tag{6}$$

Znajomość współczynnika k_p pozwala na szybkie sprawdzenie, czy moc czynna możliwa do uzyskania z prądnicy w różnych warunkach pracy nie jest przypadkiem mniejsza od założonej mocy obliczeniowej.

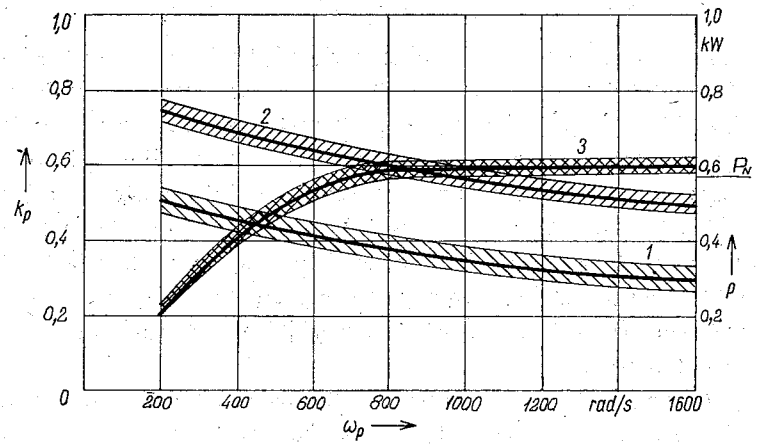
Na rys. 1a, 1b, 1c, przedstawiono dla przykładu wartości współczynnika k_p w funkcji parametrów R_f, R_d, R_o, X_f. Zakresy zmian parametrów są tak dobrane, że przedstawione wykresy są przydatne przy projektowaniu lub doborze alternatorów o mocy 0,3÷1,0 kW. Przebieg zmian współczynnika k_p można otrzymać w funkcji prędkości kątowej prądnicy ω_p, zmieniając skalę osi w stosunku:

$$\omega_p = \frac{X_f}{L_f p}, \quad \text{gdzie } p \text{ — liczba par biegunów prądnicy.}$$

Rys. 2 przedstawia wykresy zmian współczynnika k_p (krzywa *1*), mocy czynnej możliwej do otrzymania z prądnicy (krzywa *2*) oraz mocy czynnej uzyskanej w trakcie badań 9 prądnic (krzywa *3*) A12M produkcji krajowej (P_n = 588 W, U_n = 14 V, I_n = 42 A,



Rys. 1. Wartości współczynnika k_p w funkcji parametrów R_o, $R_f + R_d$, X_f



Rys. 2. Wyniki badań prądnic A12M.

1 — wartości współczynnika k_p wyznaczone na podstawie pomiaru parametrów prądnic, *2* — wartości mocy czynnej możliwej do uzyskania z prądnicy A12M przy obciążeniu rezystancją $R_o = U_n/I_n$, wyznaczone w oparciu o wzór (3) i pomiar parametrów prądnic, *3* — maksymalne wartości mocy czynnej uzyskane z badanych prądnic

$n_n = 5000\ \text{min}^{-1}$). Krzywe *1*, *2* są obliczone teoretycznie jako średnie na podstawie pomiarów parametrów badanych prądnic oraz po przyjęciu założenia, że maszyny są obciążone znamionową rezystancją

$$R_o = \frac{U_n}{I_n}.$$

Krzywa *3* jest zdjęta doświadczalnie, jako średnia, w układzie, w którym badane prądnice obciążane były regulowaną rezystancją, do granicy, przy której, dla konkretnej wartości prędkości kątowej, uzyskano maksymalną wartość mocy czynnej na ich wyjściu.

Obwody wzbudzenia prądnic zasilono napięciem znamionowym.

Obszarami zakreskowanymi oznaczono rozrzut pomiędzy odpowiednimi krzywymi, wynikający z różnicy parametrów badanych prądnic.

Z rysunku widać, że przy prędkościach kątowych prądnicy większych od $(1100 \div 1200)\ \text{rad} \cdot \text{s}^{-1}$ teoretycznie możliwa do otrzymania moc czynna jest mniejsza od podawanej przez producenta wartości mocy znamionowej. Z badania laboratoryjnego prądnic wynika jednak, że przy prędkościach $(1100 \div 1200)\ \text{rad} \cdot \text{s}^{-1}$ lub większych można uzyskać moce równe lub nawet większe od znamionowej. Jest to spowodowane tym, że w trakcie badań laboratoryjnych prądnice pracowały przy granicznych obciążeniach obwodu elektromagnetycznego a nie przy obciążeniach znamionowych, jakie zakłada się przy projektowaniu alternatora lub przy wyznaczaniu krzywej *2*. Wynika z tego, że przy prędkościach większych od $(1100 \div 1200)\ \text{rad} \cdot \text{s}^{-1}$ moc znamionową można otrzymać jedynie kosztem przeciążenia maszyn.

W zakresie niskich prędkości kątowych alternatora obwód magnetyczny nie jest w pełni wykorzystany i uzyskiwana moc czynna jest dużo mniejsza od możliwości obwodu elektromagnetycznego.

3. ZAKRES ZMIAN PRĘDKOŚCI KĄTOWEJ ALTERNATORA

Podawana przez producentów, w katalogach, maksymalna wartość prędkości kątowej prądnicy wynika z granicznej wytrzymałości konstrukcji wirnika (zwłaszcza obwodu wzbudzenia i łożysk) [1, 2, 3]. Nie bierze się pod uwagę, że istnieje możliwość wystąpienia, w trakcie pracy alternatora, zwarcia trójfazowego w prostowniku, spowodowanego komutacją złożoną zaworów półprzewodnikowych. Zjawisko to może wystąpić szczególnie ostro w nowoczesnych konstrukcjach o większej mocy, przystosowanych do pracy z prędkościami kątowymi do 2000 rad $\cdot$ s^{-1}. Celowe wydaje się zatem, by dla każdej projektowanej konstrukcji alternatora określenie granicznej prędkości kątowej ustalane było po sprawdzeniu zarówno wytrzymałości mechanicznej, jak również możliwości bezawaryjnej pracy układu elektrycznego.

Prędkość graniczną, przy której istnieje możliwość wystąpienia pomiędzy zaworami prostownika komutacji złożonej (znalezienie się układu w III etapie komutacji) wyznaczyć można [6] z równania opisującego warunki komutacji na granicy II i III etapu komutacji przy kącie komutacji i kącie opóźnienia komutacji równych:

$$\gamma = \frac{\pi}{3}, \qquad \alpha_o = \frac{\pi}{6}.$$

Równanie to przedstawia się następująco

$$\frac{\cos\varphi_k}{R_o + \frac{3}{2}(R_f + R_d)} \left\{ -\frac{\left[R_o + \frac{3}{2}(R_f + R_d)\right]\cos\varphi_f}{(R_f + R_d)\cos\varphi_k} \left[\cos\varphi_f - \sin\left(\frac{\pi}{6} - \varphi_f\right) e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k}\right] + \right.$$

$$+ \sin\left(\frac{7}{6}\pi - \varphi_k\right) - \sin\left(\frac{5}{6}\pi - \varphi_k\right) e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k} + \frac{1}{2}\left[\cos\varphi_k + \right.$$

$$\left. \left. - \sin\left(\frac{\pi}{6} - \varphi_k\right) e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k}\right]\right\} + \frac{1}{2}\left[e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k} + e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_f}\right] \times$$

$$\times \frac{\cos\varphi_k}{\left(1 - e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k}\right)\left[R_o + \frac{3}{2}(R_f + R_d)\right]} \cdot \left\{\sin\left(\frac{7\pi}{6} - \varphi_k\right) + \right.$$

$$\left. - \sin\left(\frac{5}{6}\pi - \varphi_k\right) e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k} + \frac{1}{2}\left[\cos\varphi_k - \sin\left(\frac{\pi}{6} - \varphi_k\right) e^{-\frac{\pi}{3}\operatorname{ctg}\varphi_k}\right]\right\} = 0, \tag{7}$$

gdzie:

$$\varphi_f = \arccos \frac{R_f + R_d}{\sqrt{(R_f + R_d)^2 + p^2\omega_p^2 L_f^2}},$$

$$\varphi_k = \arccos \frac{R_o + \frac{3}{2}(R_f + R_d)}{\sqrt{\left[R_o + \frac{3}{2}(R_f + R_d)\right]^2 + \frac{9}{4}p^2\omega_p^2 L_f^2}},$$

ω_p — prędkość prądnicy, $p\omega_p L_f = X_f$.

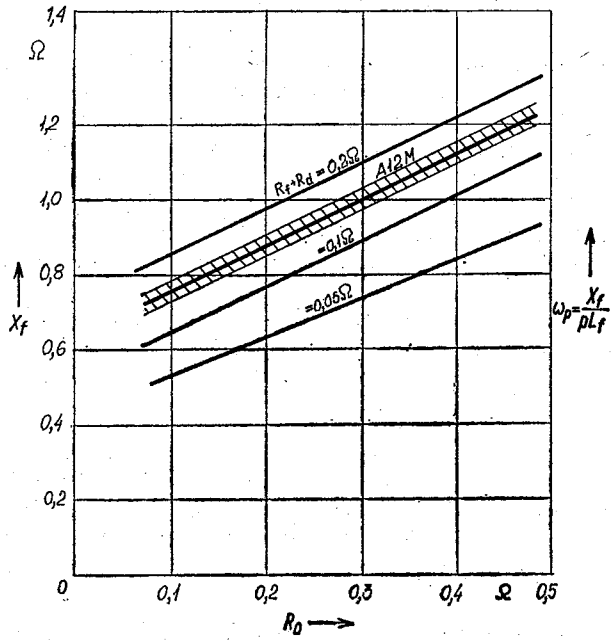
Dla znanych parametrów R_o, R_f, R_d, L_f z równania (7) wyznacza się prędkość ω_p. Jest to prędkość, przy której komutacja zaworów prostownika odbywa się na granicy II i III etapu. Maksymalna prędkość kątowa maszyny w trakcie jej eksploatacji musi być od niej mniejsza.

Na rysunku 3 przedstawione są przebiegi ilustrujące omawiany problem, które równocześnie mogą być nomogramami pomocniczymi przy sprawdzaniu zaprojektowanej konstrukcji lub doborze prądnicy do konkretnego układu. Są to zależności reaktancji fazowej X_f lub po zastosowaniu przeliczenia

$$\omega_p = \frac{X_f}{pL_f},$$

prędkości kątowej prądnicy w funkcji obciążenia R_o dla stałych parametrów $R_f + R_d$. Na tym samym rysunku znajduje się również krzywa, obliczona teoretycznie, jako średnia, dla alternatora A12M (na podstawie pomiaru parametrów 9 prądnic A12M). Obszarem zakreskowanym zaznaczono rozrzut pomiędzy krzywymi spowodowany różnicami parametrów badanych prądnic.

Z rysunku 3 wynika, że alternatory A12M o parametrach $R_f + R_d = (0{,}15 \div 0{,}18)\ \Omega$, $L_f = (0{,}85 \div 0{,}90)\ 10^{-4}$ H, przy obciążeniu znamionowym ($R_o \cong 0{,}33\ \Omega$) z punktu widzenia prawidłowej pracy prostownika mogą być bezpiecznie eksploatowane do prędkości $(1900 \div 2000)$ rad · s^{-1}. Ponieważ ich dopuszczalna prędkość podawana przez producenta sięga 1500 rad · s^{-1} nie zachodzi niebezpieczeństwo wejścia prostownika w III etap komutacji.



Rys. 3. Wartości reaktancji X_f lub prędkości kątowej ω_p, przy których zachodzi komutacja złożona w prostowniku alternatora

Z rysunku 3 wynika również, że konstrukcje o mocach rzędu 1,0 kW, charakteryzujące się parametrami $R_f + R_d = (0{,}05 \div 0{,}10)\ \Omega$, $L_f = (0{,}7 \div 1{,}0)\ 10^{-4}$ H, $R_o = (0{,}15 \div 0{,}20)\ \Omega$ już przy prędkościach $(1100 \div 1200)\ \text{rad} \cdot \text{s}^{-1}$ mogą być niezdolne do prawidłowej pracy na skutek komutacji złożonej w prostowniku. Dla tych konstrukcji istnieje więc uzasadniona potrzeba sprawdzania zakresu zmian prędkości kątowej.

4. WNIOSKI KOŃCOWE

Zjawisko komutacji występujące w prostowniku samochodowej prądnicy prądu przemiennego wpływa zarówno na wielkość mocy czynnej możliwej do uzyskania z prądnicy jak również ogranicza zakres prędkości kątowej, przy której może być ona stosowana.

Zakres mocy czynnej możliwej do uzyskania z zaprojektowanej lub dobranej prądnicy wyznaczyć można z wzoru (3) i na tej podstawie odpowiednio skorygować projekt lub parametry pracy prądnicy. Nie wystąpi wtedy zjawisko przeciążania maszyny przy pracy z prędkościami większymi od znamionowej.

Określenie prędkości granicznej prądnicy nie może być ustalone jedynie w oparciu o mechaniczną wytrzymałość konstrukcji. W trakcie eksploatacji maszyny może bowiem dochodzić do komutacji złożonej w prostowniku, czego efektem jest szybkie rozładowanie akumulatora. Prędkość, przy której ten rodzaj komutacji może w prostowniku wystąpić, wyznaczyć można z równania (7).

BIBLIOGRAFIA

1. V. V. Apsit, *Sinchronnyje masziny s koktieobraznymi polusami*, Łotewska Akademia Nauk, 1959.
2. *Baukasten Drehstromgeneratoren*, Robert Bosch, Stuttgart 1979.
3. *Drehstrom-Generatoren, Technische Unterrichtung*, Robert Bosch, Stuttgart 1982.
4. A. Kupeev, *Metodika wybora osnownych razmierow awtomobilnych generatorow peremennogo toka*, Trudy N.I.I. Awtopribora, 1970.
5. R. Matla, A. Sęk, *Numeryczne projektowanie samochodowych maszyn elektrycznych metodą potencjału skalarnego*, Archiwum Elektrotechniki, 3/1979.
6. J. Morawski, *Analityczna metoda wyznaczania charakterystyk alternatora*, Z. N. Politechniki Św. Elektryka, 12, 1983.
7. A. Sęk, *Polmag — system programów do wyznaczania przestrzennego pola magnetycznego w alternatorze samochodowym*, Sympozjum Elektrotechniki Pojazdów Samochodowych, Ustronie 1983.

J. MORAWSKI

THE INFLUENCE OF COMMUTATION IN THE RECTIFIER SYSTEM ON THE ACTIVE POWER AND THE GRAPHIC VALUE OF THE VELOCITY OF THE AUTOMOTIVE ALTERNATOR

Summary

An analytic method of calculating active power and limit velocity of the automotive alternator is offered in the paper. The method has been developed as a result of the analysis of commutation in the rectifier system of the automotive alternator.

J. MORAWSKI

INFLUENCE DE LA COMMUTATION DANS UN MONTAGE REDRESSEUR SUR LA PUISSANCE ACTIVE ET SUR LA VALEUR GRAPHIQUE DE LA VITESSE D'UN ALTERNATEUR DE VOITURE

Résumé

Dans cet article on a présenté la méthode analytique de détermination de la puissance active et de la vitesse angulaire admissible d'un alternateur de voiture pour les conditions de travail diverses. Cette méthode tient compte de l'effet de commutation dans le redresseur d'une machine.

J. MORAWSKI

EINFLUß DER KOMMUTATION IM GLEICHRICHTERSYSTEM AUF DIE NUTZKLEISTUNG UND DEN GRAPHISCHEN DREHZAHLWERT EINES WECHSELSTROM-KLAUENPOL--GENERATORS

Zusammenfassung

Es wird eine analytische Methode der Bestimmung von Nutzleistung sowie der zulässigen Drehzahl eines Klauenpol-Generators bei verschiedenen Arbeitsumständen dargestellt. Diese Methode berücksichtigt auch das Kommutationsphänomen im Gleichrichter.

Е. МОРАВСКИ

ВЛИЯНИЕ КОММУТАЦИИ В СХЕМЕ ВЫПРЯМИТЕЛЕЙ НА АКТИВНУЮ МОЩНОСТЬ И ПРЕДЕЛЬНУЮ СКОРОСТЬ ВРАЩЕНИЯ АВТОМОБИЛЬНОГО АЛЬТЕРНАТОРА

Резюме

Представлен аналитический метод определения активной мощности и предельной угловой скорости автомобильного генератора переменного тока в различных условиях роботы учитывающий влияние явления коммутации в выпрямителе машины.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 481—497

621.315.592

Synteza układów kombinacyjnych metodą dekompozycji funkcji boolowskich

TADEUSZ ŁUBA (WARSZAWA)

Instytut Telekomunikacji Politechniki Warszawskiej

Otrzymano 10.1.1982

W artykule omówiono syntezę układów kombinacyjnych metodą dekompozycji funkcji boolowskich. Wprowadzając do opisu funkcji boolowskiej podział dwublokowy sformułowano warunek wystarczający dekompozycji umożliwiający zastosowanie rachunku podziałów. Warunek ten rozszerzono następnie na przypadek układów funkcji boolowskich. Przedstawiono przykłady zastosowania rachunku podziałów do dekompozycji funkcji i układów funkcji boolowskich na bloki realizowalne z typowych elementów średniej i wielkiej skali integracji.

1. WPROWADZENIE

Jedną z możliwości upraszczania realizacji układów kombinacyjnych konstruowanych z modułów średniej i wielkiej skali integracji jest dekompozycja funkcji boolowskich. Metoda ta, stosowana niezbyt powszechnie w praktyce projektowania na elementach małej skali integracji ze względu na konkurencyjność innych metod optymalizacji układu (np. metody Quine'a-McCluskey'a), stanowić może w dziedzinie projektowania na elementach *MSI* i *LSI* skuteczne narzędzie. Narzędzie to z powodzeniem może być stosowane w projektowaniu układów kombinacyjnych na multiplekserach, dekoderach, jak też pamięciach *ROM* i matrycach *PLA*.

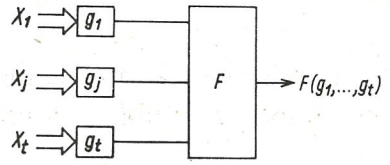
Stosowanie metod dekompozycji jest szczególnie istotne w syntezie funkcji boolowskich silnie nieokreślonych, stanowiących jak wiadomo częsty model rzeczywistych układów kombinacyjnych. Proces minimalizacji funkcji boolowskich ulega bowiem znacznej komplikacji w przypadku silnie nieokreślonych funkcji wielu zmiennych [1]. Stosowanie tablic Karnaugha jest wówczas niemożliwe ze względu na liczbę zmiennych, natomiast stosowanie metody Quine'a-McCluskey'a nie pozwala wykorzystać silnej nieokreśloności funkcji, gdyż kombinacje nieokreślone uważa się tam wstępnie za określone.

W przypadku ogólnym dekompozycję funkcji boolowskiej można opisać wzorem

$$f(x_1, \ldots, x_n) = F(g_1, \ldots, g_t); \quad t < n,$$

gdzie $g_j (1 \leqslant j \leqslant t)$ są funkcjami boolowskimi określonymi na podzbiorach (niekoniecznie rozłącznych) zbioru $\{x_1, \ldots, x_n\}$.

Dekompozycji tej odpowiada schemat blokowy jak na rys. 1. Realizując blok F na multiplekserze (dekoderze itp.), uzyskujemy możliwość zmniejszenia liczby wejść adreso-



Rys. 1. Schemat blokowy dekompozycji funkcji boolowskiej; $X_j \subseteq \{x_1, \ldots, x_n\}$

wych multipleksera z n na t. Oczywiście zysk ten jest okupiony kosztem wprowadzenia do realizacji dodatkowych układów g_j. Często jednak uzyskuje się dekompozycje, w których układy g_j są proste i mogą być zrealizowane na typowych elementach małej skali integracji. W najgorszym przypadku niektóre z układów g_j można zrealizować na multiplekserze o $|X_j|$ zmiennych adresowych. Dobierając odpowiednio funkcje g_j uzyskuje się rozwiązania, w których dla każdego $j \in \{1, \ldots, n\}$ zachodzi $|X_j| < n$. Mając na uwadze fakt, że produkowane moduły średniej (również wielkiej) skali integracji mają ograniczoną liczbę wejść adresowych, uzyskuje się w ten sposób znaczne rozszerzenie zakresu ich stosowalności.

2. DEKOMPOZYCJA FUNKCJI BOOLOWSKICH

Niech f będzie funkcją boolowską n zmiennych daną odwzorowaniem $f: D^n \to \{0, 1\}$, gdzie $D^n \subseteq \{0, 1\}^n$. Odwzorowanie f wyznacza na zbiorze wektorów D^n podział $P = (f^{-1}(0), f^{-1}(1))$. Przyjmując do opisu wektorów z D^n zapis według naturalnego kodu dwójkowego

$$\varepsilon(\mathbf{x}) = \sum_{i=1}^{n} x_i 2^{n-i}; \quad \mathbf{x} = (x_1, \ldots, x_n),$$

można stwierdzić, że podziałowi P odpowiada wzajemnie jednoznacznie podział $\tilde{P} = (\tilde{f}^{-1}(0), \tilde{f}^{-1}(1))$, gdzie odwzorowanie $\tilde{f}: \varepsilon(D^n) \to \{0, 1\}$, przy czym $f = \varepsilon \circ \tilde{f}$ [1]. Niech dla $f: D^n \to \{0, 1\}$ g będzie funkcją boolowską zmiennych $x_{i_1}, \ldots, x_{i_p}$ daną wyrażeniem boolowskim $g(x_{i_1}, \ldots, x_{i_p})$, gdzie $p < n$ oraz $i_j \in \{1, \ldots, n\}$. W zbiorze wektorów z D^n wyróżniamy dwa podzbiory G^0 i G^1. Wektor $\mathbf{x} = (x_1, \ldots, x_n)$ zaliczać będziemy do G^0, jeśli $g(x_{i_1}, \ldots, x_{i_p}) = 0$, natomiast do G^1, jeśli $g(x_{i_1}, \ldots, x_{i_p}) = 1$. Oczywiście zbiory wektorów prawdziwych i fałszywych funkcji g są rozłączne, a więc również zbiory G^0 i G^1 są rozłączne. Zachodzi też

$$G^0 \cup G^1 = D^n.$$

[1] Symbol o oznacza złożenie odwzorowań.

Dowolna funkcja g wyznacza więc na zbiorze D^n podział $P(g) = (G^o, G^1)$, a na zbiorze $\varepsilon(D^n)$ podział $\tilde{P}(g) = (\tilde{f}^{-1}(0), \tilde{f}^{-1}(1))$. Niech na przykład dla funkcji $f: D^4 \to \{0, 1\}$, dla której zbiór mintermów odpowiadających wektorom z D^4 jest

$$\varepsilon(D^4) = \{0, 2, 3, 4, 5, 12, 13, 15\},$$

funkcja $g_1(x_2, x_4) = x_2 \oplus x_4$ i funkcja $g_2(x_3) = x_3$. Wyznaczymy $P(g_1)$, $P(g_2)$.

Wektory dla g_1 są wektorami o współrzędnych x_2, x_4, przy czym wektory fałszywe to (0, 0) i (1, 1), a prawdziwe to (0, 1) i (1, 0) i w takim razie $G_1^o = (0, 0, 0, 0)$; (0, 0, 1, 0), (0, 1, 0, 1), (1, 1, 0, 1), (1, 1, 1, 1) oraz $G_1^1 = (0, 0, 1, 1)$, (0, 1, 0, 0,), (1, 1, 0, 0). Natomiast wektory z G_2^1 zapisane w postaci minterów tworzą zbiór {0, 4, 5, 12, 13}, a wektory z G_2^1 — {2, 3, 15}, czyli $\tilde{P}(g_1) = (0, 2, 5, 13, 15; 3, 4, 12)$ i $P(g_2) = (0, 4, 5, 12, 13; 2, 3, 15)$.

Niech dla $f: D^n \to \{0, 1\}$, $g_1, \ldots, g_t$ będą funkcjami boolowskimi określonymi na t podzbiorach (niekoniecznie rozłącznych) zbioru zmiennych $\{x_1, \ldots, x_n\}$. Przyporządkujemy każdemu blokowi B_j iloczynu podziałów $P(g_1) \ldots P(g_i) \ldots P(g_t)$ wektor zmiennych binarnych $(g_1, \ldots, g_t)$, w taki sposób, że:

$$g_i = 0, \quad \text{gdy } B_j \subseteq G_i^o,$$

$$g_i = 1, \quad \text{gdy } B_j \subseteq G_i^1.$$

Oznaczmy zbiór uzyskanych wektorów przez D^t. Zbiór podziałów $\mathbf{P} = \{P(g_1), \ldots, P(g_t)\}$ wyznacza więc na zbiorze D^n odwzorowanie $\mathbf{P}: D^n \to D^t (D^t \subseteq X^t)$, które definiujemy następująco

$$\mathbf{P}(x_1, \ldots, x_n) = (g_1, \ldots, g_t),$$

gdzie

$$g_i = 0, \quad \text{gdy } (x_1, \ldots, x_n) \in G_i^o,$$

natomiast

$$g_i = 1, \quad \text{gdy } (x_1, \ldots, x_n) \in G_i^1.$$

Nietrudno zauważyć, że odwzorowanie $\mathbf{P}: D^n \to D^t$ jest iniekcją. Określając z kolei na zbiorze D^t odwzorowanie $F: D^t \to \{0, 1, *\}$ w sposób następujący:

$$F(g_1, \ldots, g_t) = 0, \quad \text{gdy } \mathbf{P}^{-1}(g_1, \ldots, g_t) \subseteq f^{-1}(0);$$

$$F(g_1, \ldots, g_t) = 1, \quad \text{gdy } \mathbf{P}^{-1}(g_1, \ldots, g_t) \subseteq f^{-1}(1) \text{ oraz}$$

$$F(g_1, \ldots, g_t) = *, \quad \text{gdy nie zachodzi żaden z powyższych przypadków,}$$

stwierdzić można, że F jest odwzorowaniem $D^t \to \{0, 1\}$ wtedy i tylko wtedy, gdy

$$P(g_1) \ldots P(g_t) \leqslant P(f). \tag{1}$$

Tym samym przy spełnieniu warunku (1) na zbiorze D^t określona jest funkcja boolowska zmiennych $g_1, \ldots, g_t$. Inaczej mówiąc funkcję $f: D^n \to \{0, 1\}$, która w istocie jest funkcją n zmiennych, można jednocześnie opisać odwzorowaniem (funkcją) $F: D^t \to \{0, 1\}$, której zbiór argumentów jest $g_1, \ldots, g_t$, przy czym istotne jest to, że $t < n$. Prawdziwe jest więc następujące twierdzenie.

Twierdzenie 1. Funkcję $f = f(x_1, \ldots, x_n): D^n \to \{0, 1\}$ można przedstawić jako superpozycję

$$f = F(g_1, g_2, \ldots, g_t)$$

wtedy i tylko wtedy, gdy

$$\tilde{P}(g_1) \cdot \tilde{P}(g_2) \ldots \tilde{P}(g_t) \leqslant P(f),$$

przy czym $F(g_1, \ldots, g_t) = f(a_1, \ldots, a_n)$, gdzie $(a_1, \ldots, a_n)$ jest wektorem z D^n takim, że

$$(a_1, \ldots, a_n) \in \mathbf{P}^{-1}(g_1, \ldots, g_t).$$

Z praktycznego punktu widzenia istotne są przede wszystkim takie dekompozycje, w których t jest minimalne oraz funkcje g są funkcjami możliwie najmniejszej liczby argumentów. Dlatego najpierw rozważamy dekompozycje opisane przez

$$f(x_1, \ldots, x_n) = F(x_{i_1}, \ldots, x_{i_t}),$$

gdzie $i_j \in \{1, \ldots, n\}$.

Definicja 1. Dekompozycję funkcji $f = f(x_1, \ldots, x_n)$ taką, że

$$f = F(x_{i_1}, \ldots, x_{i_t})$$

nazywać będziemy minimalno-argumentową realizacją funkcji f wtedy i tylko wtedy, gdy nie istnieje zbiór zmiennych $\{x_{j_1}, \ldots, x_{j_{t-1}}\} \subseteq \{x_1, \ldots, x_n\}$ taki, że $f = F'(x_{j_1}, \ldots, x_{j_{t-1}})$.

Niech $a \in f^{-1}(0)$, $b \in f^{-1}(1)$. Utwórzmy dla wektorów a, b wyrażenie boolowskie $W(a, b)$

$$W(a, b) = \bigvee_i x_i,$$

gdzie sumowanie jest po wszystkich i takich, że $a_i \oplus b_i = 1$. Inaczej mówiąc x_i jest składnikiem $W(a, b)$ tylko wtedy, gdy wektory a oraz b mają składowe $a_i, b_i: a_i \neq b_i$.

Definicja 2. Reprezentacją argumentową funkcji f (oznaczoną dalej przez RA_f) nazywać będziemy wyrażenie boolowskie RA_f

$$RA_f = \bigwedge_{a,\, b} W(a, b),$$

gdzie iloczyn logiczny $\wedge$ brany jest po wszystkich parach $a, b: a \in f^{-1}(0)$, $b \in f^{-1}(1)$.

Twierdzenie 2. Dla każdego rozwiązania równania $RA_f = 1$

$$RA_f(x_{i_1}, \ldots, x_{i_t}) = 1$$

istnieje funkcja F o argumentach $x_{i_1}, \ldots, x_{i_t}$, dla której

$$f(x_1, \ldots, x_n) = F(x_{i_1}, \ldots, x_{i_t}).$$

Dowód. Jeżeli $RA_f(x_{i_1}, \ldots, x_{i_t}) = 1$, to zgodnie z def. 2, dla każdej pary wektorów $a, b: a \in f^{-1}(0)$, $b \in f^{-1}(1)$, wektory $(a_{i_1}, \ldots, a_{i_t})$ oraz $(b_{i_1}, \ldots, b_{i_t})$ różnią się na co najmniej jednej składowej. Tym samym zbiory wektorów A, B:

$$A = \{(x_{i_1}, \ldots, x_{i_t}): (x_1, \ldots, x_n) \in f^{-1}(0)\},$$

$$B = \{(x_{i_1}, \ldots, x_{i_t}): (x_1, \ldots, x_n) \in f^{-1}(1)\},$$

są rozłączne. Przyporządkowując więc każdemu wektorowi z A wartość 0, a każdemu wektorowi z B wartość 1, uzyskujemy funkcję F

$$\forall (x_1, \ldots, x_n) \in D^n \quad f(x_1, \ldots, x_n) = F(x_{i_1}, \ldots, x_{i_t}).$$

Jeżeli zbiór zmiennych $\{x_{i_1}, \ldots, x_{i_t}\}$ spełniających równanie $RA_f(x_{i_1}, \ldots, x_{i_t}) = 1$

jest zbiorem o możliwie najmniejszej liczności, to funkcja F uzyskana w sposób przedstawiony w dowodzie twierdzenia 1 jest minimalno-argumentową realizacją funkcji f.

Z konstrukcji wyrażenia RA_f wynika, że jest to iloczyn czynników o postaci $x_{i_1}+\ldots+x_{i_k}(i_j \in \{1, \ldots, n\})$. Liczba czynników wyrażenia RA_f jest równa $|f^{-1}(0)| \cdot |f^{-1}(1)|$, i jakkolwiek rozwiązanie równania $RA_f = 1$ jest znanym problemem pokrycia, to w ogólnym przypadku zapis równania o dużej liczbie czynników jest niewygodny w obliczeniach. Proces ten można znacznie uprościć wprowadzając pojęcie zmiennej niezbędnej.

D e f i n i c j a 3. Zmienna x_i jest zmienną niezbędną dla realizacji minimalno-argumentowej funkcji $f(x_1, \ldots, x_n)$, jeśli istnieją wektory $a \in f^{-1}(0)$, $b \in f^{-1}(1)$ takie, że $a_i \neq b_i$ oraz $a_j = b_j$ dla każdego $j \neq i$.

Oznaczmy przez K odwzorowanie $D^n \to \{1, \ldots, k\}$ (D^n — dziedzina funkcji $f(x_1, \ldots, x_n)$ oraz $k = |D^n|$), skonstruowane w taki sposób, że wektorom z $f^{-1}(0)$ przyporządkowane są liczby naturalne $1, \ldots, k'$, a wektorom z $f^{-1}(1)$ — liczby $k'+1, \ldots, k$. Tym samym dowolną funkcję f można opisać dwublokowym podziałem $P(f)$

$$P(f) = (B_f^0, B_f^1) = (1, \ldots, k'; k'+1, \ldots, k).$$

Niech $\mathscr{P} = P(x_{i_1}^{\sigma}) \ldots P(x_{i_r})$ będzie iloczynem podziałów indukowanych zmiennymi niezbędnymi funkcji $f(x_1, \ldots, x_n)$. Oczywiście podział $\mathscr{P} = (B_1, \ldots, B_j, \ldots, B_m)$ jest podziałem na zbiorze $\{1, \ldots, k\}$, gdzie przez B_j oznaczamy ustalony blok tego podziału. Utwórzmy wyrażenie boolowskie

$$W'(a, b) = \bigvee_i x_i$$

w sposób następujący: jeśli dla $p, q: p \in B_f^0$, $q \in B_f^1$ istnieje blok B_j podziału $\mathscr{P}$: $p, q \in B_j$ oraz wektory $a \in K^{-1}(p)$, $b \in K^{-1}(q)$ mają składowe $a_i, b_i: a_i \neq b_i$, to zmienna x_i jest składnikiem wyrażenia $W'(a, b)$.

Iloczyn logiczny RA_f'

$$RA_f' = \bigwedge_{a, b} W'(a, b)$$

uzyskany dla wszystkich $a \in f^{-1}(0)$, $b \in f^{-1}(1)$ takich, że dla $p = K(a)$, $q = K(b)$ istnieje $B_j \in \mathscr{P}$ zawierający parę p, q, będzie w takim razie taką reprezentacją argumentową funkcji f, która nie zawiera zmiennych niezbędnych. Stąd wniosek, że jeżeli zbiór zmiennych $\{x_{i_{r+1}}, \ldots, x_{i_t}\}$ jest minimalnym rozwiązaniem równania $RA_f' = 1$, to zbiór zmiennych $\{x_{i_1}, \ldots, x_{i_r}, x_{i_{r+1}}, \ldots, x_{i_t}\}$, gdzie $x_{i_1}, \ldots, x_{i_r}$ są zmiennymi niezbędnymi, jest jednocześnie minimalnym rozwiązaniem równania $RA_f = 1$, a tym samym wyznacza minimalno-argumentową realizację funkcji f.

Przykład 1. Obliczymy wszystkie realizacje minimalno-argumentowe funkcji $f = f(x_1, \ldots, x_{10})$, której tablica prawdy podaną jest w tabl. 1. Ponumerujmy wektory prawdziwe i fałszywe tej funkcji liczbami naturalnymi ze zbioru $\mathscr{K} = \{1, \ldots, 25\}$. Mamy wtedy (przy oznaczeniach $P(x_i) = P_i$):

$$P_1 = (1, 3, 6, 8, 9, 10, 11, 13, 14, 15, 19, 21, 24, 25;\\ 2, 4, 5, 7, 12, 16, 17, 18, 20, 22, 23)$$

$$P_2 = (1, 2, 4, 11, 15, 19, 21, 23;\\ 3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 16, 17, 18, 20, 22, 24, 25)$$

Tablica 1

Tablica prawdy funkcji f z przykładu 1

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	f
1	0	0	1	0	1	1	1	0	1	0	0
2	1	0	1	0	0	1	0	1	0	0	0
3	0	1	0	0	0	1	1	1	1	0	0
4	1	0	1	1	1	0	1	0	1	1	0
5	1	1	0	0	0	1	0	0	1	1	0
6	0	1	0	0	0	1	0	1	1	0	0
7	1	1	1	0	1	0	0	1	1	0	0
8	0	1	0	0	1	1	0	0	0	0	0
9	0	1	0	1	0	0	0	0	1	0	0
10	0	1	1	1	1	1	1	0	1	1	1
11	0	0	0	0	0	1	0	1	0	0	1
12	1	1	0	1	1	1	0	0	1	1	1
13	0	1	0	0	1	0	0	0	0	0	1
14	0	1	0	0	0	1	1	1	1	1	1
15	0	0	1	0	0	0	0	1	1	0	1
16	1	1	1	1	0	1	0	0	0	1	1
17	1	1	1	1	1	0	1	0	0	1	1
18	1	1	1	1	1	1	1	1	1	1	1
19	0	0	1	0	0	0	0	0	0	0	1
20	1	1	0	1	1	0	0	1	1	1	1
21	0	0	1	0	0	0	1	1	1	1	1
22	1	1	1	1	1	0	0	0	1	0	1
23	1	0	1	0	1	1	1	1	0	1	1
24	0	1	1	0	0	0	0	1	1	0	1
25	0	1	0	0	1	1	1	0	0	0	1

$P_3 = (3, 5, 6, 8, 9, 11, 12, 13, 14, 20, 25;$
$1, 2, 4, 7, 10, 15, 16, 17, 18, 19, 21, 22, 23, 24)$
$P_4 = (1, 2, 3, 5, 6, 7, 8, 11, 13, 14, 15, 19, 21, 23, 24, 25;$
$4, 9, 10, 12, 16, 17, 18, 20, 22)$
$P_5 = (2, 3, 5, 6, 9, 11, 14, 15, 16, 19, 21, 24;$
$1, 4, 7, 8, 10, 12, 13, 17, 18, 20, 22, 23, 25)$
$P_6 = (4, 7, 9, 13, 15, 17, 19, 20, 21, 22, 24;$
$1, 2, 3, 5, 6, 8, 10, 11, 12, 14, 16, 18, 23, 25)$
$P_7 = (2, 5, 6, 7, 8, 9, 11, 12, 13, 15, 16, 19, 20, 22, 24;$
$1, 3, 4, 10, 14, 17, 18, 21, 23, 25)$
$P_8 = (1, 4, 5, 8, 9, 10, 12, 13, 16, 17, 19, 22, 25;$
$2, 3, 6, 7, 11, 14, 15, 18, 20, 21, 23, 24)$
$P_9 = (2, 8, 11, 13, 16, 17, 19, 23, 25;$
$1, 3, 4, 5, 6, 7, 9, 10, 12, 14, 15, 18, 20, 21, 22, 24)$
$P_{10} = (1, 2, 3, 6, 7, 8, 9, 11, 13, 15, 19, 22, 24, 25;$
$4, 5, 10, 12, 14, 16, 17, 18, 20, 21, 23).$

Zmiennymi niezbędnymi tej funkcji są x_6, x_7, x_{10}. Dlatego najpierw wyznaczymy iloczyn $\mathscr{P} = P_6 \cdot P_7 \cdot P_{10}$. Bloki iloczynu $\mathscr{P} = (B_1, \ldots, B_7)$ są:

$B_1 = \{7, 9, 13, 15, 19, 22, 24\}$ $B_5 = \{4, 17, 21\}$
$B_2 = \{2, 6, 8, 11\}$ $B_6 = \{20\}$
$B_3 = \{5, 12, 16\}$ $B_7 = \{10, 14, 18, 23\}.$
$B_4 = \{1, 3, 25\}$

W dalszych obliczeniach pomijamy bloki B_6 i B_7, gdyż $B_6 \subseteq B_f^1$ oraz $B_7 \subseteq B_f^1$. Natomiast dla każdej pary p, q: $p \in B_j$, $q \in B_j (j \in \{1, \ldots, 5\})$ oraz $p \in B_f^0$, $q \in B_f^1$ tworzymy czynnik iloczynu logicznego RA_f', który jest sumą logiczną zmiennych $x_{i_1}, \ldots, x_{i_k}$. Zmienne x_{i_j} wchodzące w skład takiej sumy wygodnie jest wyznaczać, korzystając z wypisanych poprzednio podziałów $P(x_i)$. Na przykład dla pary $p = 7$, $q = 13$ mamy, że (7, 13) należy do jednego bloku podziałów P_1, P_3, P_8, P_9, a więc odpowiadający tej parze czynnik RA_f' będzie

$$C_1 = x_1 + x_3 + x_8 + x_9.$$

W rezultacie uzyskujemy $RA_f' = C_1 \ldots C_{19}$, gdzie wyrażenia C zapisane w postaci zbiorów indeksów występujących przy zmiennych x_i wchodzących w ich skład będą:

$C_1 = 1, 3, 8, 9$ $C_{11} = 1, 3$
$C_2 = 1, 2, 5$ $C_{12} = 2, 9$
$C_3 = 1, 2, 5, 8, 9$ $C_{13} = 2, 5, 8$
$C_4 = 4, 8$ $C_{14} = 4, 5$
$C_5 = 1, 5$ $C_{15} = 3, 4, 9$
$C_6 = 4, 5, 9$ $C_{16} = 2, 3, 9$
$C_7 = 2, 3, 4, 8$ $C_{17} = 5, 8, 9$
$C_8 = 2, 3, 4, 9$ $C_{18} = 2, 9$
$C_9 = 1, 3, 5$ $C_{19} = 1, 4, 5, 8.$
$C_{10} = 3, 4, 8$

Redukując w zbiorach C każde C_j, dla którego istnieje C_i: $C_j \supseteq C_i$ otrzymujemy równanie $RA'_f = 1$ o postaci

$$(4+8)(1+5)(1+3)(2+9)(2+5+8)(4+5)(3+4+9)(5+8+9) = 1\ ^{2)},$$

z którego wyznaczamy wszystkie minimalne rozwiązania:

$$R_1 = 1,4,5,9 \qquad R_6 = 1,2,4,9$$
$$R_2 = 1,5,8,9 \qquad R_7 = 2,3,4,5$$
$$R_3 = 3,4,5,9 \qquad R_8 = 2,3,5,8$$
$$R_4 = 3,5,8,9 \qquad R_9 = 1,2,4,5$$
$$R_5 = 1,4,8,9 \qquad R_{10} = 1,2,4,8.$$

Funkcję f z przykładu 1 można więc zrealizować w układzie kombinacyjnym F o wejściach $x_{i_1}, \ldots, x_{i_7}$, gdzie $x_{i_1} = x_6$, $x_{i_2} = x_7$, $x_{i_3} = x_{10}$ są zmiennymi niezbędnymi występującymi w każdej minimalno-argumentowej realizacji tej funkcji oraz $x_{i_4}, \ldots, x_{i_7}$ są zmiennymi, których indeksy i_j przyjmują wartości jednego ze zbiorów $R_1, \ldots, R_{10}$.

W realizacjach minimalno-argumentowych uzyskuje się dekompozycję $f = F(x_{i_1}, \ldots, x_{i_t})$ wymagającą zastosowania t wejściowego układu kombinacyjnego do realizacji F. Dalsze zmniejszenie liczby wejść do bloku F jest możliwe wyłącznie przy zastosowaniu funkcji g, których argumenty stanowiłyby co najmniej dwuelementowe podzbiory zbioru $\{x_{i_1}, \ldots, x_{i_t}\}$.

Niech π będzie iloczynem podziałów $P(x_{i_1}) \ldots P(x_{i_r})$, którego bloki oznaczamy $B_1, \ldots, B_i, \ldots, B_m$. Z bloków podziału π tworzymy zbiór D_w

$$D_w = \bigcup_i B_i,$$

gdzie sumowanie jest po wszystkich i takich, że $B_i \not\subseteq B_f^0$ oraz $B_i \not\subseteq B_f^1$. Oznaczmy:

π_w — podział na D_w, którego blokami są te bloki podziału π, które nie są zawarte ani w B_f^0 ani w B_f^1,

π_{wf} — dwublokowy podział na $D_w = (B_w^0, B_w^1)$: $B_w^0 \subseteq B_f^0$ oraz $B_w^1 \subseteq B_f^1$,

π_g — podział na D_w: $\pi_g = P(x_{i_{r+1}}) \ldots P(x_{i_t}) \cdot P(x_{j_1}) \ldots P(x_{j_p})$.

Twierdzenie 3. Minimalno-argumentowa realizacja funkcji f: $f = F(x_{i_1}, \ldots, x_{i_t})$ ma dekompozycję F'

$$F = F'(x_{i_1}, \ldots, x_{i_r}, \quad g(x_{i_{r+1}}, \ldots, x_{i_t}, x_{j_1}, \ldots, x_{j_p})) \tag{2}$$

wtedy i tylko wtedy, gdy istnieje dwublokowy podział $\pi'_g \geqslant \pi_g$: $\pi'_g \cdot \pi_w \leqslant \pi_{wf}$.

Z twierdzenia 3 oraz warunku wystarczającego dekompozycji podanego w twierdzeniu 1 wynika, że dla założonego podzbioru argumentów wziętego ze zbioru zmiennych występujących w minimalno-argumentowej realizacji funkcji f można zawsze wyznaczyć zbiór zmiennych $x_{j_1}, \ldots, x_{j_p}$, które łącznie ze zmiennymi $x_{i_{r+1}}, \ldots, x_{i_t}$ tworzyć będą zbiór argumentów funkcji g

$$P(x_{i_1}) \ldots P(x_{i_r}) \cdot P(g) \leqslant P(f).$$

W szczególności może się okazać, że zbiór zmiennych $\{x_{j_1}, \ldots, x_{j_p}\}$ jest zbiorem pustym

[2)] Symbol „+" oznacza tu sumę logiczną.

i wtedy dekompozycja opisana wzorem (2) jest dekompozycją prostą rozłączną

$$F = F'(x_{i_1}, \ldots, x_{i_r}, g(x_{i_{r+1}}, \ldots, x_{i_t})).$$

Sposób wyboru argumentów $x_{j_1}, \ldots, x_{j_p}$ występujących w dekompozycji opisanej wzorem (2) wyjaśnia następujący przykład.

Przykład 2. Jedną z realizacji minimalno-argumentowych funkcji f z przykładu 1 jest $f = F(x_1, x_2, x_4, x_5, x_6, x_7, x_{10})$. Obliczymy dekompozycję $F = F'(x_4, x_6, x_7, x_{10}, g(x_1, x_2, x_5, x_{j_1}, \ldots, x_{j_p}))$, wyznaczając jednocześnie zmienne $x_{j_1}, \ldots, x_{j_p}$. W tym celu obliczamy najpierw iloczyn podziałów

$$P(x_4) \cdot P(x_6) \cdot P(x_7) \cdot P(x_{10})$$

z odpowiednich bloków tego iloczynu tworzymy podział π_w, zbiór D_w oraz podział π_{wf}:

$$\pi_w = (7, 13, 15, 19, 24;\ 2, 6, 8, 11;\ 9, 22;\ 4, 17;\ 1, 3, 25),$$

$$D_w = \{1, 2, 3, 4, 6, 7, 8, 9, 11, 13, 15, 17, 19, 22, 24, 25\},$$

$$\pi_{wf} = (1, 2, 3, 4, 6, 7, 8, 9;\ 11, 13, 15, 17, 19, 22, 24, 25).$$

Na zbiorze D_w tworzymy podział $P_g = P(x_1) \cdot P(x_2) \cdot P(x_5)$:

$$P_g = (11, 15, 19;\ 1;\ 3, 6, 9, 24;\ 8, 13,25;\ 2;\ 4;\ 7, 17, 22).$$

Z bloków podziału P_g należy utworzyć (jeśli to możliwe) dwublokowy podział π'_g: $\pi'_g \cdot \pi_w \leqslant \pi_{wf}$. Przeznaczając blok (7, 17, 22) do pierwszego bloku π'_g, a pozostałe bloki (z P_g) do drugiego (w π'_g) uzyskujemy

$$\pi'_g = ((7, 17, 22);\ (1)\,(2)\,(8, 13, 25)(11, 15, 19)(3, 6, 9, 24)(4)).$$

Niestety $\pi'_g \cdot \pi_w \not\leqslant \pi_{wf}$, gdyż elementy 2, 6, 8 występują w drugim bloku π'_g łącznie z elementem 11 oraz elementy 1, 3 łącznie z 25. Należy więc do zmiennych x_1, x_2, x_5 dołączyć takie zmienne, które oddzielą elementy 8 od 25, 13 od 25, 11 od 15 oraz 11 od 19. Zmienne $x_{j_1}, \ldots, x_{j_p}$ wyznaczymy więc z równania

$$x_7(x_6+x_7)(x_3+x_6+x_9)(x_3+x_6+x_8) = 1.$$

Jednym z minimalnych rozwiązań tego równania jest zbiór $\{x_6, x_7\}$, a więc π_g będzie teraz $\pi_g = P(x_1) \cdot P(x_2) \cdot P(x_5) \cdot P(x_6) \cdot P(x_7)$, a stąd

$$\pi_g = (1; 2; 3; 4; 6; 7, 22; 8; 9, 24; 13; 15, 19; 17; 25; 11)$$

i jak łatwo sprawdzić istnieje $\pi'_g \geqslant \pi_g$: $\pi'_g \cdot \pi_w \leqslant \pi_{wf}$. Podział π'_g bezpośrednio określa funkcję g. Wystarczy w tym celu elementom zaliczonym do pierwszego bloku tego podziału przyporządkować wartość logiczną 0, a elementom przeznaczonym do drugiego bloku, wartość 1, co jednocześnie określa tablicę prawdy funkcji g. Należy przy tym podkreślić, że niezależnie od konstrukcji podziału π'_g (ale takiej, że $\pi'_g \geqslant \pi_g$ oraz $\pi'_g \cdot \pi_w \leqslant \pi_{wf}$) funkcja g jest funkcją zależną (co najwyżej) od zmiennych ze zbioru $\{x_1, x_2, x_5, x_6, x_7\}$.

Na przykład dla

$$\pi'_{g_1} = (7, 11, 17, 22, 25;\ 1, 2, 3, 4, 6, 8, 9, 13, 15, 19, 24)$$

będzie

$$\pi'_{g_1} \cdot \pi_w = (7;\ 13, 15, 19, 24;\ 2, 6, 8;\ 11;\ 22;\ 9;\ 17;\ 4;\ 25;\ 1, 3)$$

$$\text{i } \pi'_g \cdot \pi_w \leqslant \pi_{wf}, \text{ ale również dla}$$

$$\pi'_{g_2} = (1, 2, 3, 4, 6, 7, 8, 22;\ 9, 11, 13, 15, 17, 19, 24, 25)$$

będzie $\pi'_{g_2} \cdot \pi_w \leqslant \pi_{wf}$. Tablice prawdy funkcji g'_1, g'_2 przedstawione są w tabl. 2, a odpowiadające im tablice Karnaugha w tabl. 3a oraz 3b. Z tablic 3a i 3b obliczamy, że

$$g'_1 = x_1\bar{x}_6 + x_2\bar{x}_5 + \bar{x}_2x_7 + x_1x_6 + x_2x_6\bar{x}_7;$$

$$g'_2 = \bar{x}_1\bar{x}_6 + x_2x_5x_7 + \bar{x}_1\bar{x}_2\bar{x}_7.$$

Tablica 2

Tablica prawdy funkcji g_1, g_2

	x_1	x_2	x_5	x_6	x_7	g_1	g_2
1	0	0	1	1	1	1	0
2	1	0	0	1	0	1	0
3	0	1	0	1	1	1	0
4	1	0	1	0	1	1	0
6	0	1	0	1	0	1	0
7	1	1	1	0	0	0	0
8	0	1	1	1	0	1	0
9	0	1	0	0	0	1	1
11	0	0	0	1	0	0	1
13	0	1	1	0	0	1	1
15	0	0	0	0	0	1	1
17	1	1	1	0	1	0	1
19	0	0	0	0	0	1	1
22	1	1	1	0	0	0	0
24	0	1	0	0	0	1	1
25	0	1	1	1	1	0	1

Tablica 3

Tablice Karnaugha funkcji: a) g_1, b) g_2

a)

x_1	x_2	x_5	x_6x_7 00	01	11	10
0	0	0	1	—	—	0
0	0	1	—	—	1	—
0	1	1	1	—	0	1
0	1	0	1	—	1	1
1	1	0	—	—	—	—
1	1	1	0	0	—	—
1	0	1	—	1	—	—
1	0	0	—	—	—	1

b)

x_1	x_2	x_5	x_6x_7 00	01	11	10
0	0	0	1	—	—	1
0	0	1	—	—	0	—
0	1	1	1	—	1	0
0	1	0	1	—	0	0
1	1	0	—	—	—	—
1	1	1	0	1	—	—
1	0	1	—	0	—	—
1	0	0	—	—	—	0

Oczywiście dla realizacji na elementach małej skali integracji lepsza jest funkcja g_2, ale dla elementów średniej skali integracji (np. multiplekserów) obie postaci tych funkcji są w zasadzie równoważne, gdyż każdą z tych funkcji można zrealizować na multiplekserze o czterech wejściach adresowych. Jedyną różnicą w tych realizacjach jest konieczność zastosowania inwertora (dla funkcji g'_1).

Dalszym etapem syntezy funkcji f jest wyznaczenie odwzorowania F takiego, aby dla każdego $\mathbf{x} \in D^n$ było

$$f = F(x_4, x_6, x_7, x_{10}, g_2).$$

W tym celu należy obliczyć iloczyn podziałów

$$P_F = P(x_4) \cdot P(x_6) \cdot P(x_7) \cdot P(x_{10}) \cdot P(g_2),$$

gdzie $P(g_2)$ jest podziałem indukowanym funkcją g_2 na zbiorze K, a następnie każdemu blokowi B_i podziału P_F przyporządkować należy wektor binarny o składowych $(x_4, x_6, x_7, x_{10}, x_g)$ taki, że

$$B_i \to (x_4, x_6, x_7, x_{10}, x_g),$$

gdzie $x_j = 0$, jeśli B_i jest podzbiorem pierwszego bloku podziału $P(x_j)$, natomiast $x_j = 1$, jeśli B_i jest podzbiorem drugiego bloku podziału $P(x_j)$; $j \in \{4, 6, 7, 10, g\}$.

Odwzorowanie F przyporządkowuje wektorowi uzyskanemu z bloku B_i wartość logiczną 0, jeśli $B_i \subseteq B_f^0$, nastomiast 1, jeśli $B_i \subseteq B_f^1$. Odpowiednie wyniki obliczeń są następujące:

$$P_F = (B_1, \ldots, B_{16}),$$

gdzie $B_1 = 1, 3$, $B_2 = 2, 6, 8$, $B_3 = 4$, $B_4 = 5$, $B_5 = 7$, $B_8 = 9$, $B_7 = 10, 18$, $B_7 = 11$, $B_9 = 12, 16$, $B_{10} = 13, 15, 19, 24$, $B_{11} = 14, 23$, $B_{12} = 17$, $B_{18} = 20$, $B_{14} = 21$, $B_{15} = 22$, $B_{16} = 25$.

I tak na przykład blok B_2 jest podzbiorem odpowiednio pierwszego bloku podziału $P(x_4)$, drugiego bloku $P(x_6)$, pierwszego bloku $P(x_7)$, pierwszego bloku $P(x_{10})$ oraz pierwszego bloku podziału $P(g_2)$, a więc odpowiadający mu wektor $(x_4, x_6, x_7, x_{10}, x_g)$ będzie (0, 1, 0, 0, 0). Ponadto $B_2 \subseteq B_f^0$, czyli odwzorowanie F przyporządkowuje wektorowi (0, 1, 0, 0, 0) wartość 0. Ostateczna tablica prawdy funkcji F jest podana w tabl. 4. Zapisując F w tablicy Karnaugha (tabl. 5) uzyskujemy wyrażenie boolowskie reprezentujące

Tablica 4

Tablica prawdy funkcji F z przykładu 2

x_4	x_6	x_7	x_{10}	g_2'	F
0	1	1	0	0	0
0	1	0	0	0	0
1	0	1	1	0	0
0	1	0	1	0	0
0	0	0	0	0	0
1	0	0	0	1	0
1	1	1	1	1	1
0	1	0	0	1	1
1	1	0	1	0	1
0	0	0	0	1	1
0	1	1	1	0	1
1	0	1	1	1	1
1	0	0	1	0	1
0	0	1	1	1	1
1	0	0	0	0	1
0	1	1	0	1	1

Tablica 5

Tablica Karnaugha funkcji F

x_4	x_6	x_7	$x_{10}g_2$ 00	01	11	10
0	0	0	0	1	—	—
0	0	1	—	—	1	—
0	1	1	0	1	—	1
0	1	0	0	1	—	0
1	1	0	—	—	—	1
1	1	1	—	—	1	—
1	0	1	—	—	1	0
1	0	0	1	0	—	1

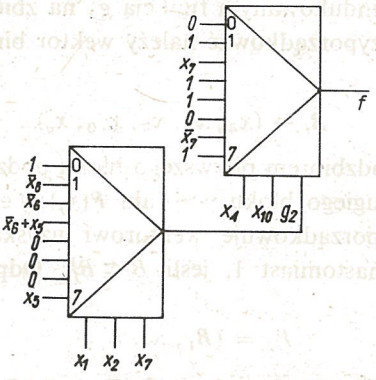
funkcję F

$$F = \bar{x}_4 g_2 + x_{10} g_2 + x_4 \bar{x}_{10} \bar{g}_2 + \bar{x}_4 x_7 x_{10} + x_4 \bar{x}_7 x_{10}.$$

W rezultacie otrzymujemy dla funkcji $f(x_1, \ldots, x_{10})$ z tabl. 1 dekompozycję

$$f = F(x_4, x_7, x_{10}, g_2(x_1, x_2, x_5, x_6, x_7)),$$

której odpowiada realizacja na multiplekserach taka, jak na rys. 2.

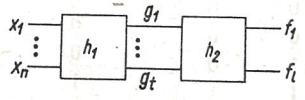


Rys. 2. Realizacja funkcji f z przykładu 2

3. DEKOMPOZYCJA UKŁADÓW FUNKCJI BOOLOWSKICH

Układem funkcji boolowskich nazywamy odwzorowanie $h: D^n \to \{0, 1\}^l$, gdzie $D^n \subseteq \{0, 1\}^n$. Realizacją układu funkcji boolowskich jest wielowyjściowy układ kombinacyjny. Układ taki można traktować jako zbiór l funkcji $f_i: D^n \to \{0, 1\}$ $(1 \leqslant i \leqslant l)$ i do realizacji każdej z tych funkcji stosować poprzednio podane metody syntezy. W wielu jednak przypadkach, a szczególnie w realizacjach na dekoderach, pamięciach, układach *PLA*, wygodniej jest odwzorowanie h realizować w układzie o schemacie blokowym jak na rys. 3. Dekompozycja taka umożliwia zmniejszenie liczby wejść (np. adresowych) w bloku h_2. Jest to istotne w tych wszystkich przypadkach, gdy blok *MSI* (np. dekoder) lub blok *LSI* (np. pamięć) jest elementem, w którym liczba wejść adresowych jest mniejsza niż n.

Obecnie sformułujemy warunki na to, aby odwzorowanie $h: D^n \to \{0, 1\}^l$ było realizowalne w układzie o schemacie blokowym, jak na rys. 3.



Rys. 3. Schemat blokowy dekompozycji układu funkcji boolowskich

Oznaczmy (dla uproszczenia zapisów) wektory z D^n liczbami naturalnymi $K = \{1, \ldots, |D^n|\}$. Niech $P(h)$ będzie podziałem na zbiorze wektorów z D^n skonstruowanym w następujący sposób

$$(s, t) \in B_P \Leftrightarrow h(s) = h(t).$$

Inaczej mówiąc wektory s i t należą do jednego bloku B podziału $P(h)$ tylko wtedy, gdy odwzorowanie h przyporządkowuje tym wektorom taki sam wektor z $\{0, 1\}^l$. Na przy-

Tablica 6

Tablica prawdy układu h funkcji boolowskich

	x_1	x_2	x_3	x_4	x_5	y_1	y_2	y_3
1	0	0	0	0	0	0	0	0
2	0	0	0	1	1	0	1	0
3	0	0	0	1	0	1	0	0
4	0	1	1	0	0	0	1	1
5	0	1	1	0	1	0	0	1
6	0	1	1	1	0	0	1	0
7	0	1	0	0	0	0	0	1
8	1	1	0	0	0	0	0	1
9	1	1	0	1	0	0	0	0
10	1	1	1	0	0	1	0	0
11	1	1	1	1	1	0	1	1
12	1	1	1	1	0	0	1	0
13	1	0	0	0	1	0	0	1
14	1	0	0	1	1	0	0	0
15	1	0	0	1	0	1	0	0

kład dla odwzorowania h określonego jak w tabl. 6, podział $P(h)$ (zapisany w postaci podziału na zbiorze $\{1, 2, \ldots, 15\}$) będzie

$$P(h) = (1, 9, 14;\ 5, 7, 8, 13;\ 2, 6, 12;\ 4, 11;\ 3, 10, 15).$$

Warunek istnienia dekompozycji o schemacie blokowym jak na rys. 3 formułuje następujące twierdzenie.

T w i e r d z e n i e 4. Oznaczmy symbolem $\beta(P)$ liczbę bloków podziału P. Odwzorowanie $h: D^n \to \{0, 1\}^l$ można zrealizować w układzie o schemacie blokowym jak na rys. 3, gdzie $\lceil \log_2 \beta(P) \rceil \leqslant t \leqslant n$[3], wtedy i tylko wtedy, gdy

$$P(g_1) \ldots P(g_t) \leqslant P(h) = P, \tag{3}$$

gdzie $P(g_i)$ jest podziałem indukowanym na zbiorze D^n funkcją g_i określoną na podzbiorze X_i (niekoniecznie właściwym) zbioru $\{x_1, \ldots, x_n\}$.

W selekcji funkcji g_i spełniających warunek (3) z twierdzenia 1 pomocnym jest pojęcie r - przydatności zbioru podziałów względem podziału P. Oznaczmy przez $\varepsilon(\tau|\sigma)$ liczbę elementów w największym bloku ilorazu podziałów τ i σ [4]. Niech $e(\tau|\sigma) = \lceil \log_2 \varepsilon(\tau|\sigma) \rceil$.

Zbiór $\{P_1, \ldots, P_k\}$ jest r-przydatny względem P, gdzie

$$r = k + e(P_1 \ldots P_k | P \cdot P_1 \ldots P_k).$$

Obliczymy r dla $P_1 = (1, 2, 3, 4, 5, 6, 7;\ 8, 9, 10, 11, 12, 13, 14, 15)$
oraz $P = (1, 9, 14;\ 5, 7, 8, 13;\ 2, 6, 12;\ 4, 11;\ 3, 10, 15)$.
Mamy tu:

$$P \cdot P_1 = (1; 9, 14;\ 5, 7;\ 8, 13;\ 2, 6;\ 12;\ 4;\ 11;\ 3;\ 1, 15),$$

$$P_1 | P \cdot P_1 = \big((1)\ (2, 6)\ (3)\ (4)\ (5, 7);\ (8, 13)\ (9, 14)\ (10, 15)\ (11)\ (12)\big),$$

$$\varepsilon(P_1 | P \cdot P_1) = 5,\ e(P_1 | P \cdot P_1) = 3, \text{ czyli } r = 4.$$

[3] Symbol $[a]$ oznacza najmniejszą liczbę całkowitą większą lub równą a.

[4] Iloraz podziałów $\tau, \sigma: \tau \geqslant \sigma$, to podział, którego elementami są bloki podziału σ.

Jeśli $\{P_1, \ldots, P_k\}$ jest r-przydatny ($k < r$) względem P, to istnieje zbiór podziałów $\{P_{k+1}, \ldots, P_r\}$, dla których

$$P_1 \ldots P_k \cdot P_{k+1} \ldots P_r \leqslant P,$$

natomiast nie istnieje $\{P'_{k+1}, \ldots, P'_{r-1}\}$ taki, że

$$P_1 \ldots P_k \cdot P'_{k+1} \ldots P'_{r-1} \leqslant P.$$

Inaczej mówiąc, warunkiem koniecznym i wystarczającym na to, aby

$$P_1 \ldots P_k \cdot P_{k+1} \ldots P_m \leqslant P,$$

jest m-przydatność zbioru $P_1, \ldots, P_k$.

Można również wykazać, że jeśli $\mathscr{P} = \{P_1, \ldots, P_k\}$ jest m-przydatny względem P, to każdy podzbiór z $\mathscr{P}$ jest m'-przydatny, gdzie $m' \leqslant m$.

Obliczymy r-przydatność podziałów $P(x_i)$ dla układu h funkcji z tablicy 6; $i \in \{1, \ldots, 5\}$. Mamy tu

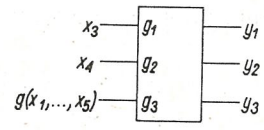
$$P(x_1) = \big((1)\,(2, 6)\,(3)\,(4)\,(5, 7);\ (8, 13)\,(9, 14)\,(10, 15)\,(11)\,(12)\big),$$

czyli $P(x_1)$ jest 4-przydatny względem $P(h)$.

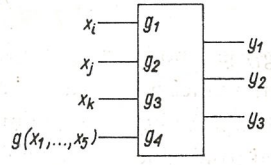
Dokonując analogicznych obliczeń dla pozostałych podziałów uzyskujemy, że $P(x_3)$, $P(x_4)$ są 3-przydatne, natomiast $P(x_2)$, $P(x_5)$ 4-przydatne. Jeśli więc istnieje dekompozycja układu h taka jak na rys. 4, to byłoby to możliwe tylko wtedy, gdyby zbiór $\{P(x_3), P(x_4)\}$ był 3-przydatny. Ale

$$P(x_3) \cdot P(x_4) = \big((1)\,(7, 8, 13);\ (9, 14)\,(2)\,(3, 15);\ (4)\,(5)\,(10);\ (11)\,(6, 12)\big),$$

czyli $\{P(x_3), P(x_4)\}$ jest 4-przydatny. Sprawdzimy więc, czy istnieje dekompozycja o schemacie blokowym jak na rys. 5. W tym celu obliczymy r dla każdego zbioru $\{P(x_i), P(x_j)\}$; $i, j \in \{1, \ldots, 5\}$. W wyniku uzyskujemy, że 4-przydatnymi parami są tylko $\{P_1, P_3\}$[5],



Rys. 4. Schemat blokowy dekompozycji układu h przy założeniu, że $\{P(x_3), P(x_4)\}$ jest 3-przydatny



Rys. 5. Schemat blokowy dekompozycji układu h w przypadku, gdy zbiór $\{P(x_i), P(x_j), P(x_k)\}$ jest 4-przydatny

$\{P_1, P_4\}$, $\{P_2, P_3\}$, $\{P_2, P_4\}$, $\{P_3, P_4\}$, $\{P_3, P_5\}$ oraz $\{P_4, P_5\}$. W związku z tym 4-przydatnymi trójkami mogą być tylko: a) $\{P_1, P_3, P_4\}$; b) $\{P_2, P_3, P_4\}$; c) $\{P_3, P_4, P_5\}$. Następnie stwierdzamy, że wyłącznie trójka a) oraz c) są 4-przydatne. Dla trójek tych odpowiednie iloczyny podziałów są:

$$P(x_1) \cdot P(x_3) \cdot P(x_4) = ((1)\,(7);\ (8, 13);\ (2)\,(3);\ (9, 14)\,(15);\ (4)\,(5);\ (10);\ (6);\ (11)\,(12));$$

$$P(x_3) \cdot P(x_4) \cdot P(x_5) = ((1)\,(7,\ 8);\ (13);\ (9)\,(3, 15);\ (14)\,(2);\ (4)\,(10);\ (5);\ (6, 12);\ (11)).$$

[5] Stosujemy tu oznaczenie $P(x_i) = P_i$.

Dla trójki a) funkcja g_1, przy której

$$P(x_1) \cdot P(x_3) \cdot P(x_4) \cdot P(g_1) \leqslant P(h),$$

może być np.: $g_1^{-1}(0), = \{1,3,4,9,11,14\}$; $g_1^{-1}(1) = \{2,5,7,12,15\}$. Wtedy $g_1 = \bar{x}_1 x_5 + x_2 \bar{x}_3 \bar{x}_4 + x_1 x_3 \bar{x}_5 + x_1 \bar{x}_2 \bar{x}_5$ oraz $P(g_1) = (1,3,4,6,9,11,13,14;\ 2,5,7,8,10,12,15)$.

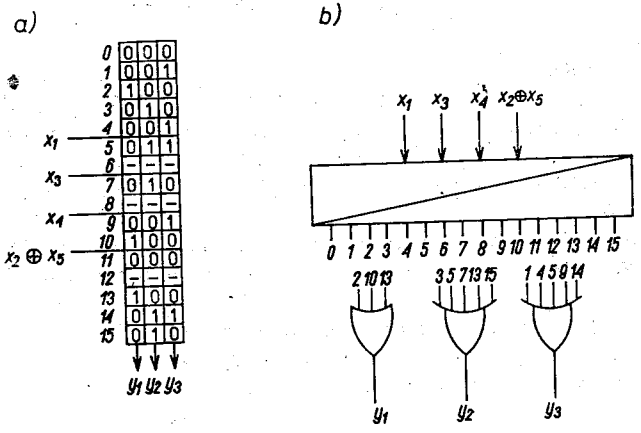
Prostsze rozwiązanie uzyskamy w następujący sposób. $P(g)$ ma być podziałem, w którym elementy 1 i 7, 2 i 3, 9, 14 i 15, 4 i 5, 11 i 12 nie mogą należeć do jednego (tego samego bloku). Obliczymy więc najmniejszy podzbiór X zbioru $x_1, \ldots, x_5$, który umożliwi konstrukcję 2-blokowego podziału $P(g) \geqslant \pi_g$ „rozdzielającego" jednocześnie wymienione wyżej elementy, gdzie π_g jest podziałem indukowanym przez zmienne z X. W tym celu tworzymy równanie $RA_g = 1$. Czynnikami wyrażenia RA_g będą sumy logiczne tych składowych x_i, dla których wektory 1 i 7, 2 i 3 itd. mają różne wartości. Stąd $RA_g = x_2 \cdot x_5$, a więc do rozdziału tych wektorów wystarczą zmienne z $X = \{x_2, x_5\}$. Zmienne te indukują na zbiorze $\{1,2,3,4,5,7,9,11,12,14,15\}$ podział π_g

$$\pi_g = (1,3,15;\ 2,14;\ 4,7,9,12;\ 5,11).$$

Przyjmując $P(g) = (1,3,5,11,15;\ 2,4,7,9,12,14)$ uzyskujemy wymagany rozdział wektorów.

Z powyższego wynika również, że do spełnienia warunku

$$P(x_1) \cdot P(x_3) \cdot P(x_4) \cdot P(g) \leqslant P(h) \tag{4}$$



Rys. 6. Realizacja układu h: a) na pamięci stałej *ROM*, b) na dekoderze

należy przyjąć funkcję $g = x_2 \oplus x_5$. Skoro do spełnienia warunku (4) wystarczy funkcja g istotnie zależna od 2 zmiennych i skoro poprzednio wykazaliśmy, że w zbiorze $\{P(x_1), P(x_2), \ldots, P(x_5)\}$ nie istnieje 4-przydatna czwórka, to obliczenia odpowiednich g dla trójki $P(x_3), P(x_4), P(x_5)$ na pewno nie dadzą funkcji g prostszej od już uzyskanej. Realizacja dekompozycji układu h odpowiadająca podziałom $P(x_1), P(x_3), P(x_4), P(g)$ jest przedstawiona na rys. 6, przy czym na rys. 6a — na pamięci stałej *ROM*, a na rys. 6b — na ekoderze.

4. UWAGI KOŃCOWE

Dotychczasowe dekompozycyjne metody syntezy układów kombinacyjnych za punkt wyjścia do obliczeń przyjmują przede wszystkim minimalne lub zminimalizowane wyrażenie boolowskie dekomponowanej funkcji. Metody te ulegają więc znacznej komplikacji, gdy funkcja boolowska jest zadana w postaci tablicy prawdy. Obliczenie minimalnej sumo-iloczynowej postaci funkcji jest bowiem dla funkcji wielu zmiennych procesem skomplikowanym. Stosowane w praktyce funkcje wielu zmiennych to najczęściej funkcje silnie nieokreślone. Należy więc poszukać takich metod dekompozycji, w których złożoność obliczeniowa byłaby bardziej uzależniona od liczby wektorów prawdziwych i fałszywych funkcji, a mniej od liczby jej zmiennych. Wydaje się, że warunek dekompozycji sformułowany w niniejszym artykule spełnia takie założenia. Przyjęcie bowiem w dekompozycji $f(x_1, \ldots, x_n) = F(g_1, \ldots, g_t)$ założenia, że $g_1 = x_{i_1}, \ldots, g_r = x_{i_r}$ $(r < t)$ powoduje podział wektorów z dziedziny f (zbiór D^n) na dziedziny funkcji $g_{r+1}, \ldots, g_t$. Dziedziny te stanowią rozłączne podzbiory zbioru D^n, są więc mniej liczne.

Sformułowany warunek dekompozycji rozszerzono następnie na przypadek układów funkcji boolowskich, co może znaleźć praktyczne zastosowanie w realizacjach funkcji wzbudzeń automatów, które z natury rzeczy wygodniej jest traktować (szczególnie w realizacjach na dekoderach i pamięciach) jako układ, a nie jako zbiór niezależnych funkcji boolowskich.

BIBLIOGRAFIA

1. W. Dudek, *Minimalizacja wieloargumentowych funkcji logicznych słabo określonych z wiązaniami wieloargumentowymi*, Archiwum Automatyki i Telemechaniki, PWN, Warszawa 1979, z. 1.
2. H. A. Sholl, *Direct Transition Memory and its Application in Computer Design*, IEEE Trans. on Computers, No 10, 1974.
3. W. Traczyk, *Multipleksery w układach kombinacyjnych*, Archiwum Automatyki i Telemechaniki, PWN, Warszawa 1979, z. 3.
4. M. A. Harrison, *Wstęp do teorii sieci przełączających i teorii automatów*, PWN, Warszawa 1975.
5. H. C. Torng, *Switching Circuits — Theory and Logic Design*, Addison-Wesley Publishing Company, 1972.
6. W. Majewski, A. Albicki, *Algebraiczna teoria automatów*, WNT, Warszawa 1980.
7. T. Łuba, *Synteza układów kombinacyjnych metodą dekompozycji funkcji boolowskich*, Instytut Telekomunikacji PW, referaty, z. 108, 1983.

T. ŁUBA

SYNTHESIS OF COMBINATORIAL SYSTEMS THROUGH DECOMPOSITION OF BOOLEAN FUNCTIONS

Summary

A synthesis of combinatorial systems through decomposition of Bollean functions is discussed. Introduction of two-block partition to the description of Boolean functions has made it possible to formulate the sufficient condition for application of partition algebra to problems of decomposition. This condition has been subsequently extended to systems of Boolean functions. It is argued that partition algebra can

be applied to break down Boolean functions and systems of Boolean functions into a block made of typical elements of medium and large scale of integration.

T. ŁUBA

SYNTHÈSE DES SYSTÈMES DE COMBINAISON PAR LA MÉTHODE DE DÉCOMPOSITION DES FONCTIONS DE BOOLOW

Résumé

L'article présente une nouvelle méthode de décomposition des fonctions et des systèmes de Boolow. La caractéristique essentielle de cette méthode est l'application du calcul de partage, ce qui permet la vérification de la condition suffisante d'existence de la décomposition sans qu'il soit nécesaire de calculer la forme minimale de la fonction. La présente méthode peut être appliquée à la synthèse des systèmes de combinaison sur éléments moyens et grands d'intégration.

T. ŁUBA

SYNTHESE DER KOMBINATIONSSYSTEME MITTELS DEKOMPOSITION DER BOOLOWSCHEN FUNKTIONEN

Zusammenfassung

Es wurde in diesem Artikel eine neue Methode für die Dekomposition der Boolowschen Funktionen und der Funktionssysteme dargestellt. Die Ausnutzung der Teilungsrechnung ist eine wesentliche Eigenschaft dieser Methode, die Anwendung bei Synthese der Kombinationssysteme mit Elementen mittleren und großen Intergrationsmaßstabes finden kann.

Т. ЛУБА

СИНТЕЗ КОМБИНАЦИОННЫХ СХЕМ ПРИ ПОМОЩИ МЕТОДА ДЕКОМПОЗИЦИИ БУЛЕВЫХ ФУНКЦИИ

Резюме

Показан новый метод декомпозиции булевых функции и систем этих функции. Существенной чертой этого метода является применение алгебры разбеиний. Это даёт возможность проверки достаточного условия существования декомпозиции, причём не нужно считать минимальную реализацию И — Или. Представленный метод можно применить в синтезе комбинационных схем на типичных элементах средней и большой степени интегрирования.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 499—526

A Natural Approach in Pattern Recognition

ZBIGNIEW M. WÓJCIK (WARSZAWA)

Przemysłowy Instytut Elektroniki

The paper received 17.8.1983

The paper is concerned with an approach to pattern recognition in which objects are describable with sentences and phrases. The description is then transformed into a set of graphs by means of grammar rules of a natural language. Parts of speech are used as labels of arcs and of nodes in the graphs. In particular, nodes of the graphs represent zero-argument relations (i.e. features) of a pattern and the nodes are marked with nouns or with pronouns; arcs tying together two nodes represent double-argument relations on the features and are marked with verbs or with combinations verb-preposition; arcs tying to a single node merely specify a feature and represent single-argument relations on features and are labeled with adjectives, numerals, pronouns, etc; directed arcs represent functions and are marked by means of verbs or combinations verb-preposition.

Operators and recognition algorithms capable of transforming patterns into the prepared graphs are to be worked out. Every operator corresponds to a set of names labelling the graph nodes. The operators take into account a gray level distribution in a preprocessing and numbers of objects, background fragments and connected peripheries within a circular neighbourhood in the process of object tracking.

Results of recognition fulfil the basic condition on representing images if they are independent of rotations of picture around the operator field centre and if all picture points are processed.

An object transformed into a graph gets the name of the standard (template) if arrangement of names of relations in the graph of the object is identical to that of the standard, i.e. if the two graphs are of the same qualitative structure. In some cases similar quantitative structure is required for the two graphs, eg. similar total change in slopes, uniformity in the slope changes, and lengths of contours, too.

1. SENTENCE DESCRIPTION OF OBJECTS AND THEIR GRAPH REPRESENTATION

A recognition process is presented here methodologically starting from a description of an object to be recognized. The description can be accomplished quite naturally using sentences and phrases. The natural description is then transformed into its graph representation by means of synonymous grammar rules. The graph representation of patterns is very important for three reasons: a) at the stage of devising algorithms and operators the description enables one to know which features of a pattern are subject to recognition and which relations on them will be searched for; b) recognition results take the form

of graphs at the stage of object tracking and coding; c) the graph data base facilitates reduction in object data and identification of patterns.

Let us consider a sentence l_t describing a triangle as an example:

l_t = „the first vertex adjoins a side, this side adjoins the next vertex, this vertex adjoins the following side, this side adjoins the next vertex, this vertex adjoins the following side, this side adjoins the first vertex".

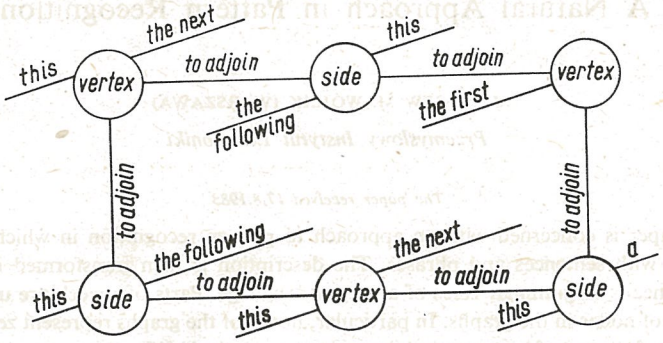


Fig. 1. Simple graph of the triangle definition

Grammar rules of English make it possible to transform, even automatically, any pattern defined in English into a synonymous graph or set of graphs being semantics of the definition [9—15] (fig. 1). Predicates of clauses are usually transformed into directed arcs of the semantics. Since the two clauses „side adjoins vertex" and „vertex adjoins side" are true and are equivalent, the function „to adjoin" is a symmetrical one and therefore it is transformed into a non-directed arc. All graphs of static images are mostly non-directed. Zero-argument relations (features), as „side" and „vertex", are argument of the double-argument relations, such as „to adjoin", and therefore they are transformed into nodes originating these arcs. Some single-argument relations as „this", „the next", „the first", „the following" are auxiliary and they define which feature adjoins another. For the basic condition on representing images to be satisfied (see Appendix) one can use only words which represent a pattern. Therefore, all auxiliary words and arcs should be discarded from any graph. A triangle graph shown in fig. 2b is obtained as a result of this reduction.

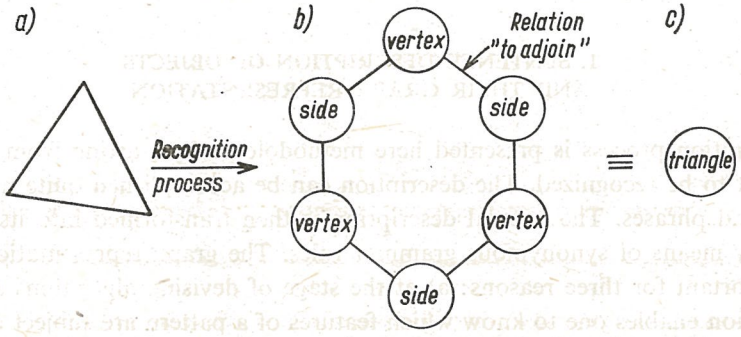


Fig. 2. Triangle (a), its graph (b), and its simple node representing all triangles

General rules of transforming any sentence or phrase description into a graph representation are the following [9, 13, 15]:

a) the subject of any independent clause is transformed into a node of a graph. The subject is the label of the node;
b) the predicate of the clause is transformed into a directed arc having its origin at the subject node. The predicate labels the arc;
c) the object, if it exists, is transformed into another node at which the directed arc ends. The object labels the node;
d) the adjuncts, if they exist, (i.e. the attributes of subject and of object) are transformed into arcs tied to subject node or to object node. The adjuncts label the arcs;
e) the resulting graph, so called simple graph (e.g. shown in fig. 1) is transformed into so called simple node (e.g. into a node labeled with the word „triangle") (see fig. 2c);
f) the attributes of predicate (int. al. adverbials) including mood, tense and voice, are transformed into arcs connected to the simple node. The attributes label the node;
g) any subordinate clause is transformed into a proper node or into an arc in the same manner as the corresponding part of sentence.

2. THE EDGE-PRESERVING TECHNIQUE IN IMAGE PROCESSING

Image pre-processing is described in terms of grey-level distribution within a circular operator field. Elementary operations such as smoothing, enhancement, noise suppression, derivation of textures can be performed without blurring edges of objects if the edge-preserving technique is incorporated. The edge-preserving technique devised by Z. Wójcik [7, 8] relies upon examination of homogeneity in grey levels of pixels (i.e. picture elements) of a rotating neighbourhood (fig. 3) with respect to grey level of the centre pixel of a circular operator field [7, 8]. One can assume the reciprocal of the variance $r_n(X, Y)$ as the measure of the homogeneity of the nth neighbourhood:

$$r_n(X, Y) = \frac{1}{I+1} \sum_{i=1}^{I} (v_{in}(X_i, Y_i) - v(X, Y))^2. \tag{1}$$

The variance $r_n(X, Y)$ is not purely statistical as I differences in grey levels of I pixles of nth neighbourhood with respect to the operator centre grey level $v(X, Y)$ is calculated and not with respect to the mean grey level of the nth neighbourhood.

A maximal variance (i.e. minimal admissible homogeneity) r (e.g. equal to 0.5) is assumed for a picture. The first neighbourhood is selected whose variance $r_f(X, Y)$ is smaller than the maximal admissible variance r. The mean value $v_f(X, Y)$ in grey levels of pixels of this fth neighbourhood is calculated:

$$v_f(X, Y) = \frac{1}{I} \sum_{i=1}^{I} v_{if}(X_i, Y_i). \tag{2}$$

The fth neighbourhood of the minimal variance is selected if no variance in the operator field is smaller than the threshold r. The set of the mean values of all pixels forms the

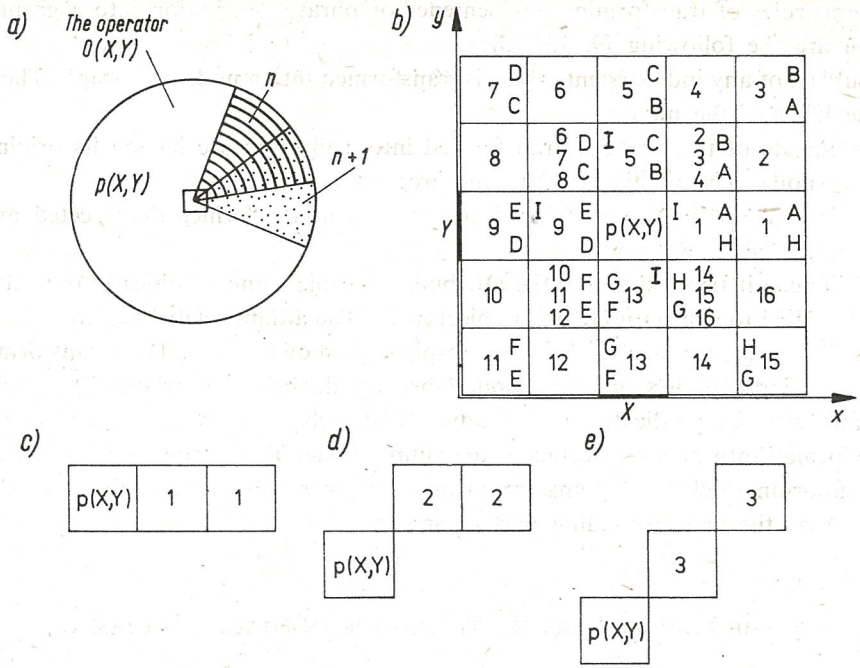


Fig. 3. An operator field with two positions of a rotating neighbourhood n and $n+1$ (a), and the 25-pixel operator field (b) with 16 positions of a two-pixel rotating neighbourhood labeled through digits 1 to 16; capital letters A through H label pixels of the 4-pixel neighbourhood used for noise suppression; figs. (c), (d), (e) are exemples of typical two-pixel neighbourhoods

smoothed picture within so called active area B of the image:

$$B = \{p(X, Y): q \leqslant X \leqslant M-q,\ q \leqslant Y \leqslant N-q\} \tag{3}$$

where M, N are numbers of image pixels along picture coordinates axes Ox and Oy, q is the operator radius.

Any increase in the value of the threshold r speeds up the smoothing but the quality of the process decreases.

When dealing with coloured pictures it is very useful to enhance contrast of objects towards their background, e.g. in the case of remote-sensing data. In the enhancement-smoothing method, the covariance of grey levels between every two or more spectrum channels is calculated for every nth neighbourhood. The covariance $c_{nkl}(X, Y)$ between the nth neighbourhood grey levels $v_{nli}(X_i, Y_i)$ in the channel l and the grey levels $v_{nki}(X_i, Y_i)$ in the channel k in the same neighbourhood with respect to the centre grey levels $v_k(X, Y)$ and $v_l(X, Y)$ of the two channels is:

$$c_{nkl}(X, Y) = \frac{1}{I+1} \sum_{i=1}^{I} \left(v_{nki}(X_i, Y_i) v_{nli}(X_i, Y_i) - v_k(X, Y) v_l(X, Y)\right)^2. \tag{4}$$

The neighbourhood of the maximal covariance $c_{kl}(X, Y)$ from among the covariances $c_{1kl}(X, Y)$, $c_{2kl}(X, Y)$, ..., $c_{nkl}(X, Y)$, ..., $c_{Pkl}(X, Y)$ is selected. The correlation coefficient

$\varrho_{kl}(X, Y)$ is then calculated as objects of the same colour are usually represented by pixels of similar value in the correlation coefficient between at least two channels k and l:

$$\varrho_{kl}(X, Y) = \frac{c_{kl}(X, Y)}{\sqrt{r_k(X, Y)\, r_l(X, Y)}} \tag{5}$$

where variances $r_k(X, Y)$ and $r_l(X, Y)$ refer to the neighbourhood of the maximal covariance $c_{kl}(X, Y)$.

The centre of the operator acquires the average grey level $v_{kl}(X, Y)$ of pixels of the two channels of the neighbourhood of the maximal covariance multiplied by the correlation coefficient:

$$v_{kl}(X, Y) = \frac{\varrho_{kl}(X, Y)}{2I} \sum_{i=1}^{I} (v_{ki}(X_i, Y_i) + v_{li}(X_l, Y_l)). \tag{6}$$

The enhanced and smoothed image is the set of all pixels in the raster B, each pixel of the grey level $v_{kl}(X, Y)$ (compare expression (3)).

The edge-preserving technique is applicable to noise removing, too, and to derivation of textures which will be discussed in the next items. The smoothing suppresses noise to some extent, but high-level disturbances are to be discarded after the thresholding process.

3. DERIVATION OF OBJECTS OF SIMILAR GREY LEVEL

Thresholding follows pre-processing of image. Pixel $p(X, Y)$ represents an object and is labeled with digit 1 if the pre-processed grey-level $v'(X, Y)$ is greater than threshold $Q(j, X, Y)$:

$$(j, X, Y) = \begin{cases} 1, & \text{if } v'(X, Y) \geqslant Q(j, X, Y), \\ 0, & \text{if } v'(X, Y) < Q(j, X, Y), \end{cases} \tag{7}$$

where: (j, X, Y) denotes digital either 1 or 0 after the thresholding (0 represents the background); $v'(X, Y)$ is the grey level either $v_f(X, Y)$, or $v_{kl}(X, Y)$, or $v(X, Y)$; $Q(j, X, Y)$ is the threshold:

$$Q(j, X, Y) = k_m \frac{w_o v_o(X, Y) + v(j)}{w_o + 1}, \tag{8}$$

where: $v_o(X, Y)$ denotes the mean grey level of an operator field, e.g. in the case of the square operator shown in fig. 3b with the radius $q = 2$ the $v_o(X, Y)$ is:

$$v_o(X, Y) = \frac{1}{(2q+1)^2} \sum_{X_i=X-q}^{X+q} \sum_{Y_i=Y-q}^{Y+q} v(X_i, Y_i); \tag{9}$$

w_o is the weight (importance) of the mean $v_o(X, Y)$: any increase in the weight w_o implies that the mean $v_o(X, Y)$ (i.e. a local grey level) affects the threshold $Q(j, X, Y)$ more intensively. The w_o may be assumed to be equal to zero in practice and then:

$$Q(j, X, Y) = Q(j) = k_m v(j), \tag{10}$$

$v(j)$ is the value of abscissa of the beginning of a consocutive local jth minimum of the histogram of grey levels of all image pixels.

L e m m a 1. The coefficient k_m satisfies the inequalities:

$$k_1 > 1, \quad \text{if } v_o(X, Y) \leqslant v(j), \tag{11a}$$

$$k_2 \leqslant 1, \quad \text{if } v_o(X, Y) > v(j). \tag{11b}$$

Proof (by contradiction). If the inverse to (11a) is assumed (i.e. $k_1 \leqslant 1$ for $v_o(X, Y) \leqslant v(j)$, e.g. for the background of objects: the background is defined approximately as a subset of pixels for which $v_o(X, Y \leqslant v(j))$, small disturbances and shading generating greater $v'(X, Y)$ readily result in fulfilling the inequality $v'(X, Y) \geqslant Q(j, X, Y)$ for some pixels representing a darker background (compare expressions (7), (8) and (10)). The incorrect value $(j, X, Y) = 1$ is then obtained for some background pixels and therefore the inequality (11a) is true. Inequality (11b) holds true analogously, which was to be proved.

L e m m a 2. For every jth grey level histogram minimum abscissa $v(j)$:

a) the maximal value of the coefficient k_1 in equation (8) for an assumed value k_2 is constrained by:

$$\forall(j)\left(k_1 \frac{w_o v_o(X, Y)+v(j)}{w_o+1} < k_2 \frac{w_o v_o(X, Y)+v(j+1)}{w_o+1}\right) \tag{12a}$$

and in equation (10) by:

$$\forall(j)\left(k_1 v(j) < k_2 v(j+1)\right) \tag{12b}$$

where $\forall$ is the universal quantifier, and $v(j) < v(j+1)$;

b) the minimal value of the coefficient k_2 in equation (8) for an assumed value of k_1 is constrained by:

$$\forall(j)\left(k_2 \frac{w_o v_o(X, Y)+v(j)}{w_o+1} > k_1 \frac{w_o v_o(X, Y)+v(j-1)}{w_o+1}\right) \tag{13a}$$

and in equation (10) by:

$$\forall(j)\left(k_2 v(j) > k_1 v(j-1)\right). \tag{13b}$$

Proof. Both left and right sides of inequalities (12a, b) and (13a, b) are thresholds: $Q(j, X, Y) = Q(k_m, j, X, Y)$. Pixels of grey levels exceeding the thresholds represent objects and all other pixels represent their background. If the inverse to inequalities (12a, b) is assumed, both $(j, X, Y) = 1$ and $(j+1, X, Y) = 1$ for two different grey level layers are obtainable. This is a contradiction because brighter objects implied by $v(j)$ would not be separated from darker objects implied by $v(j+1)$. Darker objects should be detected by (7) for $v(j+1)$ and brighter for $v(j)$ but the inverse to inequalities (12a, b) contradicts this. Therefore, (12a, b) are true. Inequalities (13a, b) hold true analogously, which was to be proved.

In the case when the values $v(j)$ are arranged in an ascending order, the brightest objects are detected first (i.e. for $j = 1$) and the darkest last.

Objects of grey levels $v(j)$ through $v(j+1)$ are represented by the set $F(j, j+1)$:

$$F(j, j+1) = F(j) \setminus F(j+1), \tag{14}$$

where $F(j+1)$ and $F(j)$ are sets of image pixels of digital 1:

$$F(j+1) = \{p(X, Y) \in B\colon\ (j+1, X, Y) = 1\}, \tag{15a}$$

$$F(j) = \{p(X, Y) \in B\colon\ (j, X, Y) = 1\} \tag{15b}$$

and $\setminus$ denotes the difference of sets.

The same set $F(j, j+1)$ is achievable when using the exclusive-or operation $\oplus$, because any image $F(j)$ covers the image $F(j+1)$:

$$F(j, j+1) = \{p(X, Y) \in B\colon\ ((j, X, Y) \oplus (j+1, X, Y)) = 1\}. \tag{16}$$

Equations (7) through (16) define regions of similar grey level. Frontiers of the similarity are determined by the coefficients k_1 and k_2, and by statistical parameters of the image, especially by the histogram and by the mean local grey level. In order to compensate stronger disturbances, the coefficients k_1 and k_2 should more differ from 1. One can say also of some self-adapting property of the threshold $Q(j, X, Y)$ to unknown a priori shading [14] varying with the mean value $v_o(X, Y)$.

4. REMOVAL OF DISTURBANCES IN BINARY IMAGES

Disturbances in any binary image under investigation can be discarded by using the dge-preserving technique [7, 8]. Several (e.g. eight) positions are assumed for a neigh-e bourhood rotating within an operator, e.g. within the 25-pixel operator (fig. 3b). Capital letters $A, B, \ldots H$ inside the pixels in fig. 3b indicate membership of the pixels for the eight positions of the rotating four-pixel neighbourhood. The neighbourhood in its nth position is homogeneous with the central pixel $p(X, Y)$ of the operator if all pixels of the neighbourhood are of the same digital value as the central pixel $p(X, Y)$. Additionally, homogeneity of the symmetrical four-pixel neighbourhood (fig. 3b) labelled with capitals I with the central pixel $p(X, Y)$ is inspected. The digital value of the central pixel $p(X, Y)$ remains unchanged if there exists at least one neighbourhood homogeneous with the $p(X, Y)$. If there is no neighbourhood homogeneous with the $p(X, Y)$, the binary value of the central pixel is complemented.

Disturbances are removable both in any binary image layer $F(j) \cup CF(j)$ and in $F(j, j+1) \cup CF(j, j+1)$, where $\cup$ denotes the topological sum of sets, and $CF(j)$ and $CF(j, j+1)$ are sets of pixels of binary zero:

$$CF(j) = \{p(X, Y) \in B\colon\ (j, X, Y) = 0\}, \tag{17}$$

$$CF(j, j+1) = \{p(X, Y) \in B\colon\ (j+1, X, Y) \oplus (j, X, Y) = 0\}. \tag{18}$$

The process of noise removing from a binary image $W(j, j+1)$ whose binary values 1 represent objects of grey levels $v(j)$ through $v(j+1)$:

$$W(j, j+1) = F(j, j+1) \cup CF(j, j+1) = \{p(X, Y) \in B\colon\ (j+1, X, Y) \oplus (j, X, Y)\} \tag{19}$$

or from a binary image $W(j)$ whose binary 1 represent objects of grey levels exceeding $v(j)$:

$$W(j) = F(j) \cup CF(j) = \{p(X, Y) \in B\colon\ (j, X, Y)\} \tag{20}$$

can be described formally through an operator $d(X, Y)$ related to an operator field, and through an operator $o_n(X, Y)$ related to nth position of a neighbourhood rotating inside the operator. The neighbourhood in its nth position is homogeneous with the operator centre $p(X, Y)$ if the operator $o_n(X, Y)$ acquires the digital 1 and is not homogeneous if $o_n(X, Y) = 0$:

$$o_n(X, Y) = \bigwedge_{i=1}^{I} \sim(X_{in}, Y_{in}) \vee \bigwedge_{i=1}^{I} (X_{in}, Y_{in}) \tag{21}$$

where: indexes j and $j+1$ are omitted for the sake of simplicity; thus (X, Y) denotes binary of pixel $p(X, Y)$ of an image $W(j, j+1)$ or of an image $W(j)$; and (X_{in}, Y_{in}) denotes binary of ith pixel of rotating neighbourhood in its nth position in the operator field; $\wedge$ denotes logical product and $\vee$ denotes logical sum; I is the number of all pixels of the rotating neighbourhood.

AND operation $\wedge$ examines a logical homogeneity of the rotating neighbourhood towards the operator centre binary (X_{on}, Y_{on}) in equation (21).

The operator $d(X, Y)$ acquires binary 1 if there exists at least one position at which the rotating neighbourhood is homogeneous with the operator centre:

$$d(X, Y) = \bigvee_{n=1}^{P} o_n(X, Y) \tag{22}$$

where P is the number of all positions of the rotating neighbourhood.

The binary (X, Y) of the operator centre is complemented if $d(X, Y) = 0$ and remains unchanged if $d(X, Y) = 1$:

$$(X, Y)' = \begin{cases} (X, Y), & \text{if } d(X, Y) = 1, \\ \sim(X, Y), & \text{if } d(X, Y) = 0. \end{cases} \tag{23}$$

The binary image W free from disturbances forms the set of pixels of binary values $(X, Y)'$:

$$W = \{p(X, Y) \in B\colon (X, Y)'\}. \tag{24}$$

5. SEGMENTATION OF BINARY IMAGES

Every binary image W can consist of several objects of the same binary 1. Contours of all objects are detected and copied on so called edge image in the first step of separation of single objects from all the binary image. In the second step each connected contour is tracked. Every pixel belonging to the contour is copied on another image without objects and at the same time the contour pixel is discarded from the edge image when tracking. The tracking of a single contour is completed if there are no more connected contour pixels. Interior of the contour that has been tracked is filled to restore a single object in the third and last step of separation of an object. Segmentation is completed if all contours detected are separated and filled. Segmentation facilitates and even makes possible the process of measurement of parameters of single objects.

Pixel $p(X, Y)$ represents an edge in a binary image if it represents an object (i.e. if it is of binary 1) and if it has at least one background pixel in its nearest four-pixel neigh-

bourhood. Edge pixels are detected by the operator $e(X, Y)$ and are represented by $e(X, Y) = 1$:

$$e(X, Y) = (X, Y)' \wedge \sim \big((X+1, Y)' \wedge (X-1, Y)' \wedge (X, Y+1)' \wedge (X, Y-1)'\big). \quad (25)$$

The set E of all pixels of binary values $e(X, Y)$ forms the edge image:

$$E = \{p(X, Y) \in B: e(X, Y)\}. \quad (26)$$

The contours in image E are without breaks because they are detected as edges of objects in binary image, and a single break in an object implies two objects and thus two connected contours.

One can clean an extra image to locate pixels representing any connected edge while tracking it. The tracking process starts when the centre of the nine-pixel operator comes across a pixel representing edge, and then:

i) The central pixel representing an edge is either copied on another extra image and is descarded from the edge image or acquires an extra digital value (e.g. equal to 2) in the edge image (this extra digit cannot be treated as representing edge when tracking);
ii) The eight pixel neighbourhood of the operator is searched for the presence of edge pixels. The tracking continues from (iv) if there are no edge pixels in the neighbourhood;
iii) The first edge pixel of the neighbourhood becomes the new centre of the operator field and coordinates of all other edge pixels of the neighbourhood are recorded. The tracking continues from (i);
iv) Centre of the operator field acquires consecutive coordinates recorded in (iii) and tracking continues from (i). The tracking of the single object is completed if there are no more coordinates recorded in (iii).

Filling of the contour tracked is the last step of separation of a single object. Minimum and maximum coordinates X_{min}, X_{max}, Y_{min} and Y_{max} of edge gathered while tracking are very useful for filling edge. Coordinates X, Y of any pixel of the contour tracked satisfy the inequalities:

$$X_{min} \leqslant X \leqslant X_{max}, \quad (27a)$$

$$Y_{min} \leqslant Y \leqslant Y_{max}. \quad (27b)$$

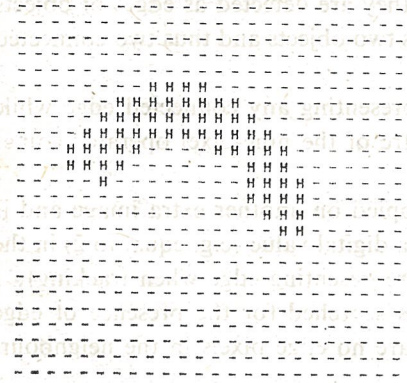
The edge tracked is filled with a pixel if there is the connectivity of the pixel with the edge through an object of the binary image W. The connectivity is investigated only for coordinates X, Y specified by (27a, b).

6. GRAPH CODING AND SHAPE RECOGNITION OF OBJECTS IN BINARY IMAGE

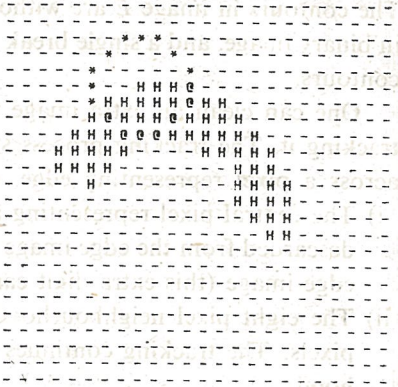
A linear object filled can be subjected to graph coding. While coding, the object is tracked by a circular operator field. It is assumed that the diameter of the operator field is grater than any width of a linear object but smaller than an extension (e.g. length) of a linear object. The tracking of a single object begins when an object pixel passes the operator field centre.

A single line segment is implied if the number I_p of intersections of an object with

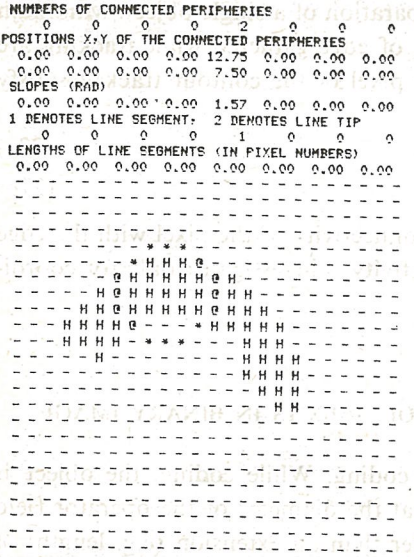
NUMBERS OF CONNECTED PERIPHERIES
0 0 0 0 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
SLOPES (RAD)
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 0 0 0 0 0 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00



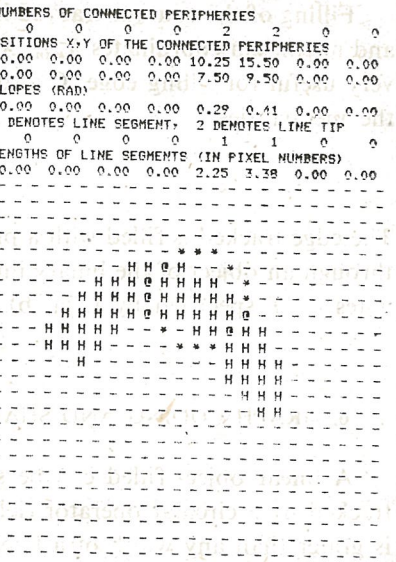
NUMBERS OF CONNECTED PERIPHERIES
0 0 0 0 1 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 0.00 9.86 0.00 0.00 0.00
0.00 0.00 0.00 0.00 8.00 0.00 0.00 0.00
SLOPES (RAD)
0.00 0.00 0.00 0.00 1.04 0.00 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 0 0 0 1 0 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 0.00 0.71 0.00 0.00 0.00



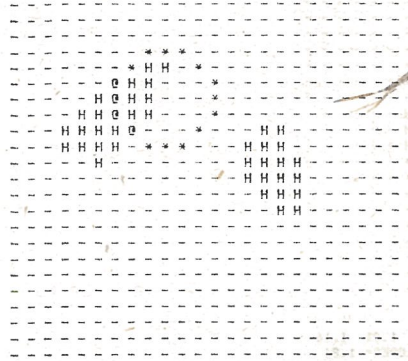
NUMBERS OF CONNECTED PERIPHERIES
0 0 0 0 2 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 0.00 12.75 0.00 0.00 0.00
0.00 0.00 0.00 0.00 7.50 0.00 0.00 0.00
SLOPES (RAD)
0.00 0.00 0.00 0.00 1.57 0.00 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 0 0 0 1 0 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00



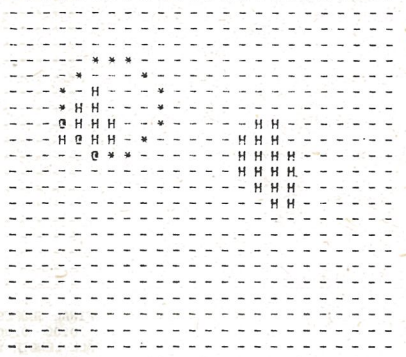
NUMBERS OF CONNECTED PERIPHERIES
0 0 0 0 2 2 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 0.00 10.25 15.50 0.00 0.00
0.00 0.00 0.00 0.00 7.50 9.50 0.00 0.00
SLOPES (RAD)
0.00 0.00 0.00 0.00 0.29 0.41 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 0 0 0 1 1 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 0.00 2.25 3.38 0.00 0.00



NUMBERS OF CONNECTED PERIPHERIES
0 0 0 1 2 2 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 8.00 10.25 15.50 0.00 0.00
0.00 0.00 0.00 8.20 7.50 9.50 0.00 0.00
SLOPES (RAD)
0.00 0.00 0.00 -0.30 0.29 0.41 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 0 0 1 1 1 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 2.36 2.25 3.38 0.00 0.00

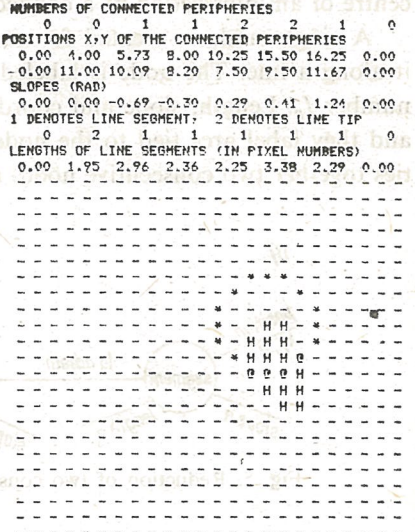


NUMBERS OF CONNECTED PERIPHERIES
0 0 1 1 2 2 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 5.73 8.00 10.25 15.50 0.00 0.00
0.00 0.00 10.09 8.20 7.50 9.50 0.00 0.00
SLOPES (RAD)
0.00 0.00 -0.69 -0.30 0.29 0.41 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 0 1 1 1 1 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 2.96 2.36 2.25 3.38 0.00 0.00



NUMBERS OF CONNECTED PERIPHERIES
0 0 1 1 2 2 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 4.00 5.73 8.00 10.25 15.50 0.00 0.00
0.00 11.00 10.09 8.20 7.50 9.50 0.00 0.00
SLOPES (RAD)
0.00 0.00 -0.69 -0.30 0.29 0.41 0.00 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 2 1 1 1 1 0 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 1.95 2.96 2.36 2.25 3.38 0.00 0.00

NUMBERS OF CONNECTED PERIPHERIES
0 0 1 1 2 2 1 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 4.00 5.73 8.00 10.25 15.50 16.25 0.00
0.00 11.00 10.09 8.20 7.50 9.50 11.67 0.00
SLOPES (RAD)
0.00 0.00 -0.69 -0.30 0.29 0.41 1.24 0.00
1 DENOTES LINE SEGMENT; 2 DENOTES LINE TIP
0 2 1 1 1 1 1 0
LENGTHS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 1.95 2.96 2.36 2.25 3.38 2.29 0.00



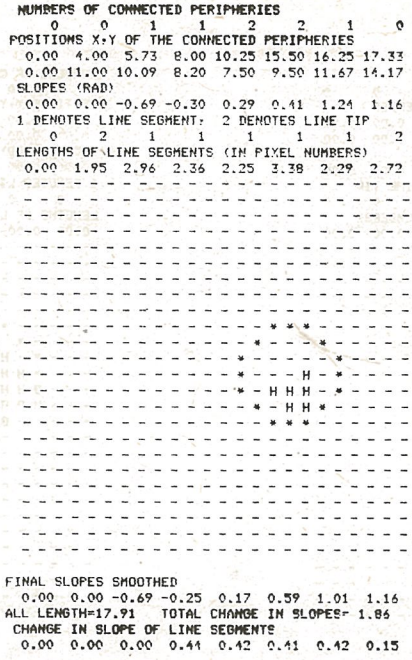


Fig. 4. A half-ring (a) and consecutive stages of tracking it by a circular operator field

the periphery of the operator field equals 2. A line tip is recognized if $I_p = 1$. A bifurcation is stated if $I_p = 3$. A junction is implied if $I_p = 4$.

Subsequently, the operator field is translocated to the next position determined by centre of an intersection of a connected periphery of the operator with a line (fig. 4).

A single node is created for each position of the operator field while translocating it along a line. The node is labeled with a name of a line segment recognized upon the number I_p. Length, slope and coordinates are one-argument relations of the line segment and they label arcs tied to the node [10—14]. An arc labeled with the verb „to adjoin" ties together two consecutive nodes representing an object (fig. 5a).

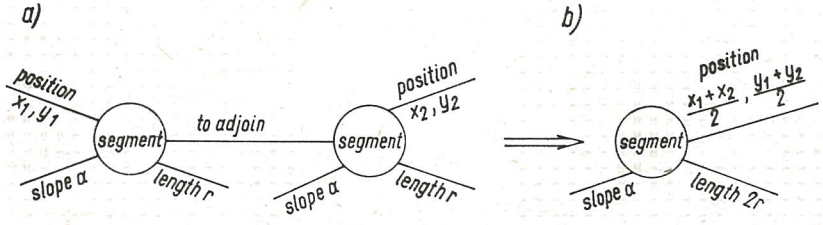


Fig. 5. Reduction of two consecutive nodes (a) of similar slope to one node (b)

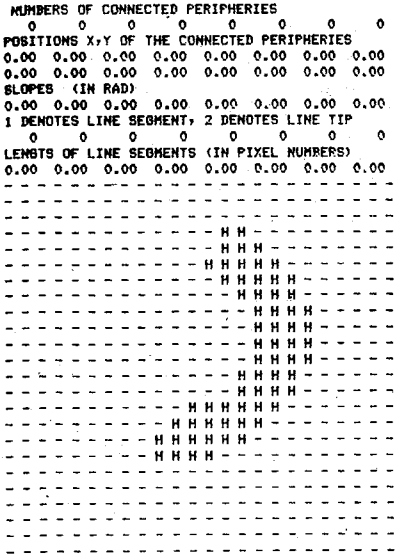
Every two consecutive nodes are replaced by one node in the case when their direction attributes are of similar value (fig. 5). Reduction in the graph data base is carried out in this way.

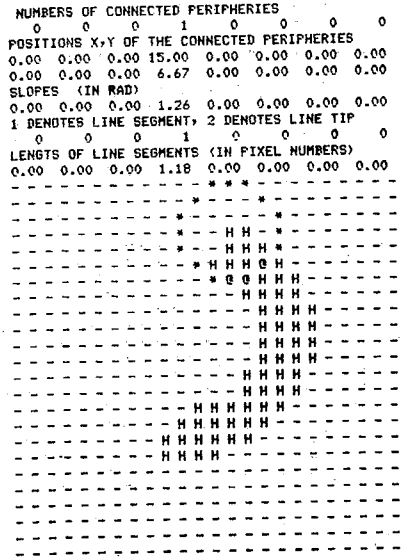
The recognized object (i.e. converted to a graph) gets the name of the standard graph (template) if the graph of the object is identical qualitatively (i.e. with respect to arrangement of names of features and relations in the graph) to the graph of the standard.

Figure 4 presents a half-ring while tracking it by a circular operator. Positions of the connected periphery, slopes and lengths of line segments are on display for each stage of the tracking. They are determined upon the centres of connected peripheral pixels of the operator field representing object. To facilitate computations all area of the operator is cleared after all parameters of the line segment are determined. Some steps of the tracking at the outset (figs. 4b, c, d) are to embed the operator field on the object properly. The continous change in the line slope in one direction merely (see bottom of fig 4i) and the total change in the slope are the basis of a half-ring identification.

Figs. 6 and 7 present the steps of tracking of the same object as in fig. 4a but situated in two different positions.

Contour of a non-linear object separated through tracking by the nine-pixel operator can be coded through the repeated tracking by a greater operator field. The diameter of the operator field should be smaller than any dimension (e.g. width) of a non-linear object. Further, the greater is the diameter of the operator, the faster is the process of contour coding. The number I_p is smaller than three when tracking nad coding the edge of any single object. The number $I_p > 2$ (e.g. implying a bifurcation or junction of two contours) can be stated in an image obtained as the topological sum of two various edge images received for two various abscissae $v(j)$ and $v(j+1)$ of histogram minima. Decomposition of two dimensional images into three-dimensional bodies is possible through the analysis of the contour connections; grey levels of adjacent parts of the bodies (e.g. walls) providing additional information [2].





NUMBERS OF CONNECTED PERIPHERIES

0 0 0 2 0 0 0 0

POSITIONS X,Y OF THE CONNECTED PERIPHERIES

0.00 0.00 0.00 17.25 0.00 0.00 0.00 0.00

0.00 0.00 0.00 8.50 0.00 0.00 0.00 0.00

SLOPES (IN RAD)

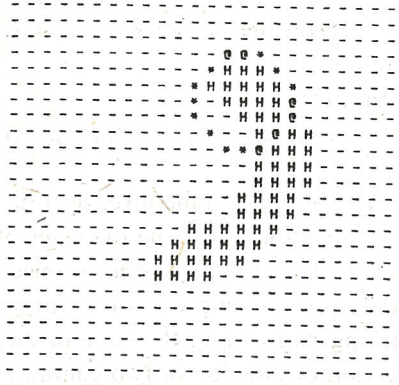
0.00 0.00 0.00 1.57 0.00 0.00 0.00 0.00

1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP

0 0 0 1 0 0 0 0

LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)

0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00



NUMBERS OF CONNECTED PERIPHERIES

0 0 2 2 0 0 0 0

POSITIONS X,Y OF THE CONNECTED PERIPHERIES

0.00 0.00 16.00 17.50 0.00 0.00 0.00 0.00

0.00 0.00 6.33 11.75 0.00 0.00 0.00 0.00

SLOPES (IN RAD)

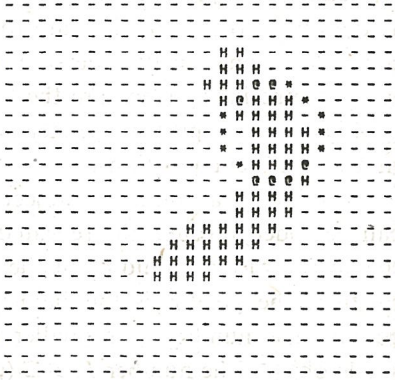
0.00 0.00 1.21 1.43 0.00 0.00 0.00 0.00

1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP

0 0 1 1 0 0 0 0

LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)

0.00 0.00 3.38 2.27 0.00 0.00 0.00 0.00



NUMBERS OF CONNECTED PERIPHERIES

0 0 2 2 1 0 0 0

POSITIONS X,Y OF THE CONNECTED PERIPHERIES

0.00 0.00 16.00 17.50 17.00 0.00 0.00 0.00

0.00 0.00 6.33 11.75 13.44 0.00 0.00 0.00

SLOPES (IN RAD)

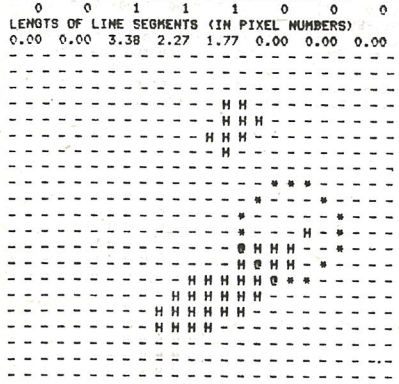
0.00 0.00 1.21 1.43 -1.28 0.00 0.00 0.00

1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP

0 0 1 1 1 0 0 0

LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)

0.00 0.00 3.38 2.27 1.77 0.00 0.00 0.00



NUMBERS OF CONNECTED PERIPHERIES

0 0 2 2 1 1 0 0

POSITIONS X,Y OF THE CONNECTED PERIPHERIES

0.00 0.00 16.00 17.50 17.00 14.78 0.00 0.00

0.00 0.00 6.33 11.75 13.44 15.44 0.00 0.00

SLOPES (IN RAD)

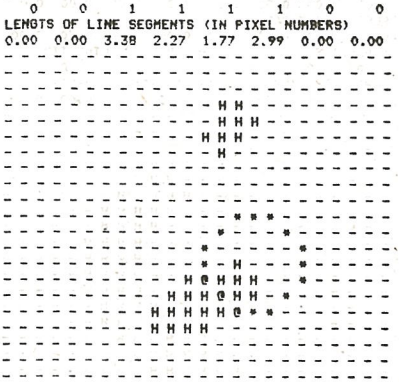
0.00 0.00 1.21 1.43 -1.28 -0.73 0.00 0.00

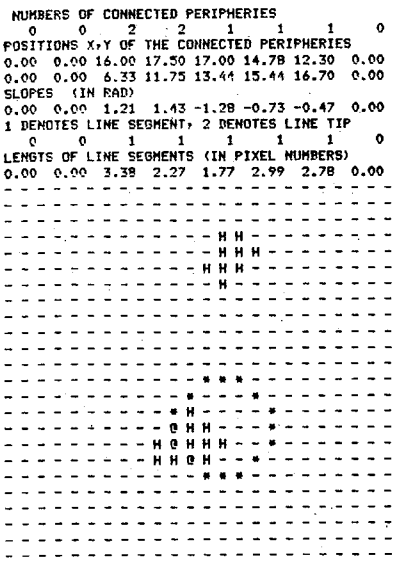
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP

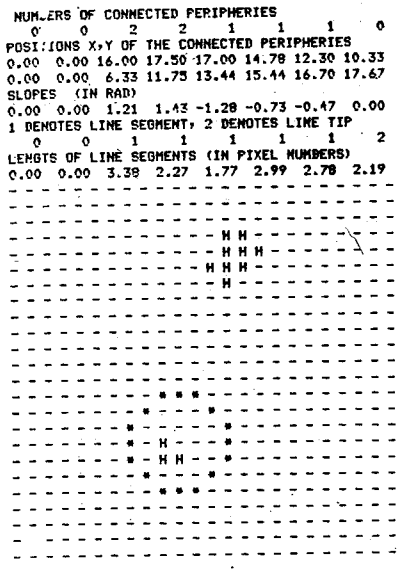
0 0 1 1 1 1 0 0

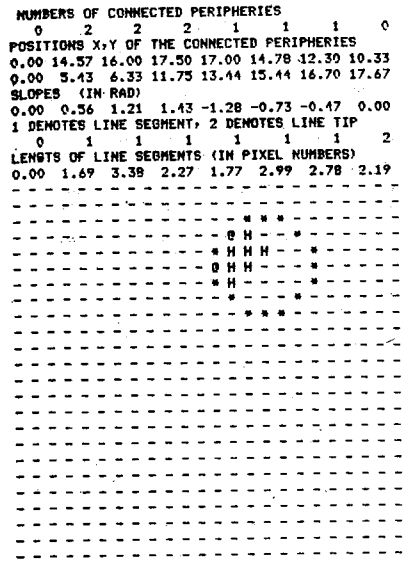
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)

0.00 0.00 3.38 2.27 1.77 2.99 0.00 0.00









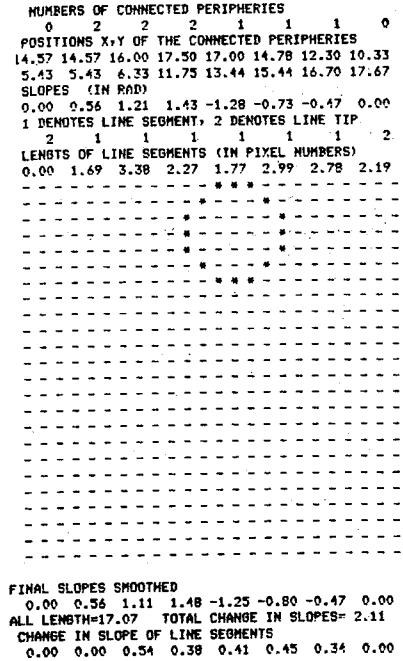
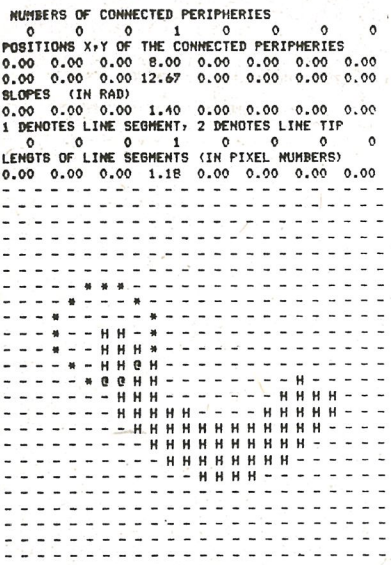
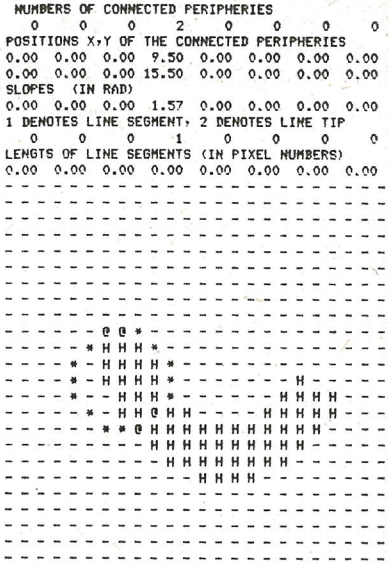
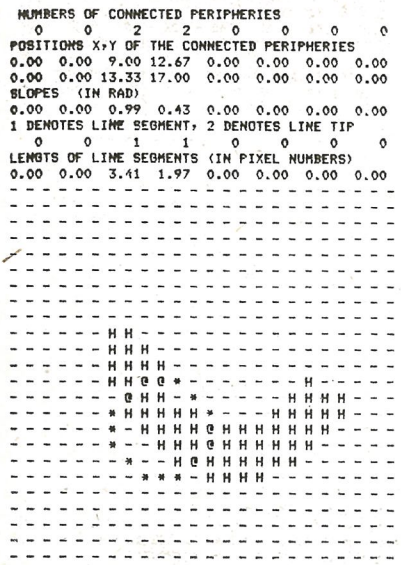
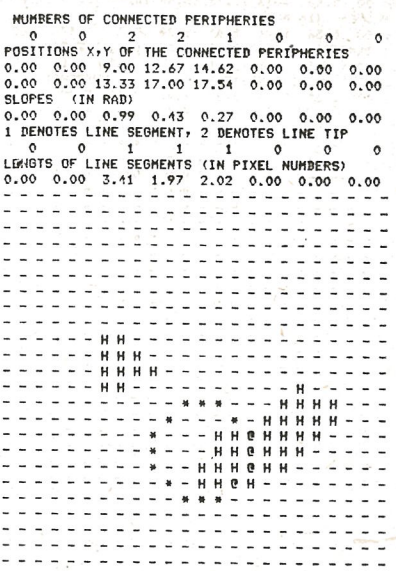


Fig. 6. The half-ring (a) in another position and consecutive stages of tracking it by a circular operator field

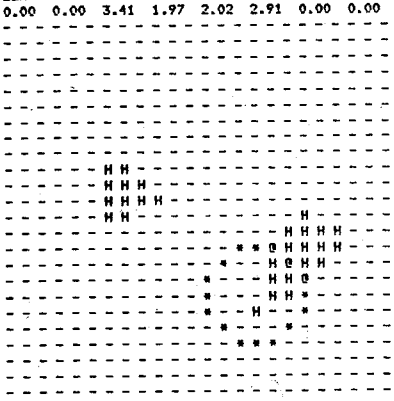

NUMBERS OF CONNECTED PERIPHERIES
0 0 0 1 0 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 8.00 0.00 0.00 0.00 0.00
0.00 0.00 0.00 12.67 0.00 0.00 0.00 0.00
SLOPES (IN RAD)
0.00 0.00 0.00 1.40 0.00 0.00 0.00 0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
0 0 0 1 0 0 0 0
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 1.18 0.00 0.00 0.00 0.00


NUMBERS OF CONNECTED PERIPHERIES
0 0 0 2 0 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 0.00 9.50 0.00 0.00 0.00 0.00
0.00 0.00 0.00 15.50 0.00 0.00 0.00 0.00
SLOPES (IN RAD)
0.00 0.00 0.00 1.57 0.00 0.00 0.00 0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
0 0 0 1 0 0 0 0
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 0.00 0.00 0.00 0.00 0.00 0.00

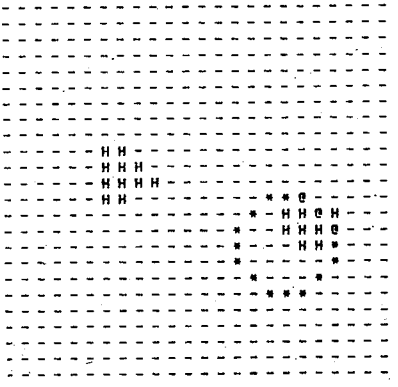

NUMBERS OF CONNECTED PERIPHERIES
0 0 2 2 0 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 9.00 12.67 0.00 0.00 0.00 0.00
0.00 0.00 13.33 17.00 0.00 0.00 0.00 0.00
SLOPES (IN RAD)
0.00 0.00 0.99 0.43 0.00 0.00 0.00 0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
0 0 1 1 0 0 0 0
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 3.41 1.97 0.00 0.00 0.00 0.00


NUMBERS OF CONNECTED PERIPHERIES
0 0 2 2 1 0 0 0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00 0.00 9.00 12.67 14.62 0.00 0.00 0.00
0.00 0.00 13.33 17.00 17.54 0.00 0.00 0.00
SLOPES (IN RAD)
0.00 0.00 0.99 0.43 0.27 0.00 0.00 0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
0 0 1 1 1 0 0 0
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00 0.00 3.41 1.97 2.02 0.00 0.00 0.00

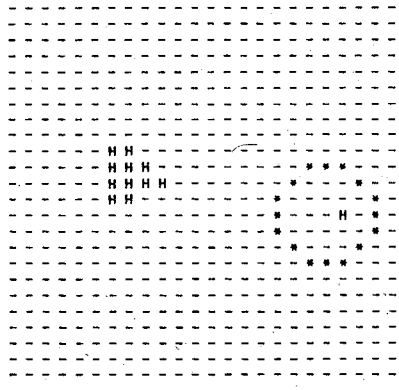
```
NUMBERS OF CONNECTED PERIPHERIES
   0     0     2     2     1     1     0     0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00  0.00  9.00 12.67 14.62 17.50  0.00  0.00
0.00  0.00 13.33 17.00 17.54 17.13  0.00  0.00
SLOPES  (IN RAD)
0.00  0.00  0.99  0.43  0.27 -0.14  0.00  0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
   0     0     1     1     1     1     0     0
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00  0.00  3.41  1.97  2.02  2.91  0.00  0.00
```



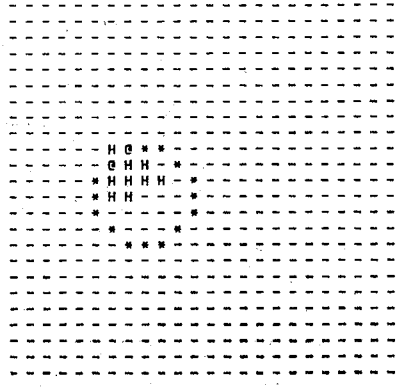
```
NUMBERS OF CONNECTED PERIPHERIES
   0     0     2     2     1     1     1     0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00  0.00  9.00 12.67 14.62 17.50 19.50  0.00
0.00  0.00 13.33 17.00 17.54 17.13 14.88  0.00
SLOPES  (IN RAD)
0.00  0.00  0.99  0.43  0.27 -0.14 -0.84  0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
   0     0     1     1     1     1     1     0
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00  0.00  3.41  1.97  2.02  2.91  3.01  0.00
```



```
NUMBERS OF CONNECTED PERIPHERIES
   0     0     2     2     1     1     1     1     0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00  0.00  9.00 12.67 14.62 17.50 19.50 21.00
0.00  0.00 13.33 17.00 17.54 17.13 14.88 14.00
SLOPES  (IN RAD)
0.00  0.00  0.99  0.43  0.27 -0.14 -0.84  0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
   0     0     1     1     1     1     1     2
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00  0.00  3.41  1.97  2.02  2.91  3.01  1.74
```



```
NUMBERS OF CONNECTED PERIPHERIES
   0     1     2     2     1     1     1     0
POSITIONS X,Y OF THE CONNECTED PERIPHERIES
0.00  8.00  9.00 12.67 14.62 17.50 19.50 21.00
0.00 11.88 13.33 17.00 17.54 17.13 14.88 14.00
SLOPES  (IN RAD)
0.00  0.97  0.99  0.43  0.27 -0.14 -0.84  0.00
1 DENOTES LINE SEGMENT, 2 DENOTES LINE TIP
   0     1     1     1     1     1     1     2
LENGTS OF LINE SEGMENTS (IN PIXEL NUMBERS)
0.00  1.77  3.41  1.97  2.02  2.91  3.01  1.74
```



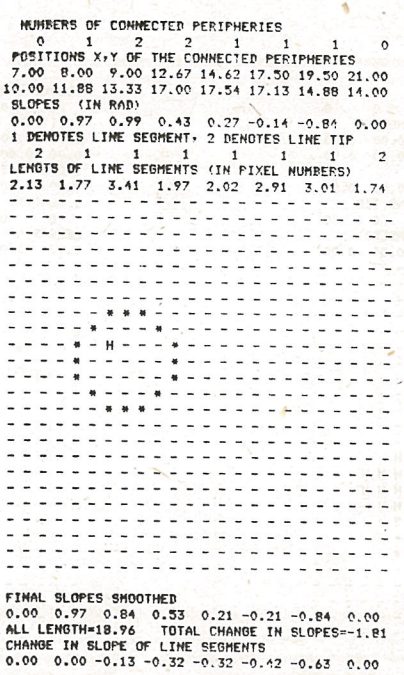


Fig. 7. The half-ring (a) in another position and consecutive stages of tracking it

7. RECOGNITION OF TEXTURES AND THE SECOND-ORDER PARAMETERS

Some statistical parameters of graphic images used in quantitative metallography are measured in each position n of a rotating neighbourhood. Pixel $p(X, Y)$ represents a texture $T(u_p, w_p)$ if there is such position n of a neighbourhood in an operator field or a circular neighbourhood of the same area as that rotating that value s_p of the parametr p measured within the neighbourhood belongs to the interval (u_p, w_p) assumed a priori, and the pixel $p(X, Y)$ does not represent the texture in the other cases [9, 13]:

$$\left(\exists\left(n \in 0(X, Y)\right)\left(s_p \in (u_p, w_p)\right)\right) \Rightarrow p(X, Y) \in T(u_p, w_p), \tag{28a}$$

$$\left(\forall\left(n \in 0(X, Y)\right)\left(s_p \notin (u_p, w_p)\right)\right) \Rightarrow p(X, Y) \notin T(u_p, w_p). \tag{28b}$$

The texture $T(u_p, w_p)$ consists of binary values $t_p(X, Y) = 1$:

$$t_p(X, Y) = \begin{cases} 1, & \text{if } \exists\left(n \in 0(X, Y)\right)\ s_p \in (u_p, w_p) \\ 0, & \text{if } \forall\left(n \in 0(X, Y)\right)\ s_p \notin (u_p, w_p) \end{cases} \tag{29}$$

and

$$T(u_p, w_p) = \{p(X, Y) \in B:\ t_p(X, Y) = 1\}. \tag{30}$$

All local parameters measured within an operator field $0(X, Y)$ at stages preceding the thresholding are first-order parameters (e.g. the variance (1), mean value (2), the covariance (4), the correlation coefficient (5)), and all parameters following the thresholding are second-order parameters (e.g. the textural parameters used in the quantitative metallo-

graphy). The third-order parameters are measured following a second thresholding (e.g. textures of textures).

A second-order image is generated through substitution of second-order parameters for grey levels of image pixels. Incorporating a rotating neighbourhood one can assume the neighbourhood f of the maximal value $s_{pf}(X, Y)$ in a parameter p from among of all values $\{s_{p1}(X, Y), s_{p2}(X, Y), \ldots s_{pP}(X, Y)\}$ measured in neighbourhoods $\{1, 2, \ldots P\}$:

$$v'(X, Y) = s_{pf}(X, Y). \tag{31}$$

Thus, the second-order image S_p is:

$$S_p = \{p(X, Y): v'(X, Y) = s_{pf}(X, Y)\}. \tag{32}$$

Any second-order image can be subjected to smoothing, enhancement and thresholding resulting in textures. Repeating the procedure of image processing of the first-order stage one can guarantee better quality of second-order images and textures. Removal of disturbances from any texture is possible as from binary images. A topological product of two textures results in a new texture.

The maximum radius q of any neighbourhood should be smaller than any size of the smallest extension of a texture under investigation. The minimum radius q should be great enough so as to make it possible to find such a neighbourhood for each pixel $p(X, Y)$ belonging to a texture in question that a value s_p or $s_{pf}(X, Y)$ measured inside the neighbourhood still belongs to an interval assumed.

The well known statistical parameters of optic images used in the quantitative metallography are [6]: size of area A_A, length of perimeters of objects L_A and the number N_A of objects, all referred to the area unit of the image measured. The A_A, L_A and N_A are the basic image parameters used for statistical evaluations in quantitative metallography. Many other parameters can be derived from these three ones.

A texture extracted with respect to an interval of the parameter A_A forms regions of approximate areas referred to an area unit of the image under investigation. Analogously, one can extract regions of an approximate number of objects referred to an area unit.

An approximation L of a real length of perimeters of objects is [13]:

$$L = ax + by \tag{33}$$

where:

$$x = \operatorname{card}\left\{p(X, Y) \in 0(X, Y): \bigvee_{n=1}^{4} K_n(X, Y) = 1\right\}, \tag{34}$$

$$y = \operatorname{card}\left\{p(X, Y) \in 0(X, Y): \bigvee_{n=1}^{4} I_n(X, Y) = 1\right\}, \tag{35}$$

card $\{\ldots\}$ is the number of all elements of the set $\{\ldots\}$ the logical functions $K_n(X, Y)$ and $I_n(X, Y)$ are shown in fig. 8 and the coefficients a, b are calculated by means of the least squares method, as shown below.

To assess the coefficients a, b some measurements $L_1, L_2, \ldots L_m, \ldots L_D$ of a real and well known length of perimeter L' are performed, each measurement m being:

$$L_m = ax_m + by_m. \tag{36}$$

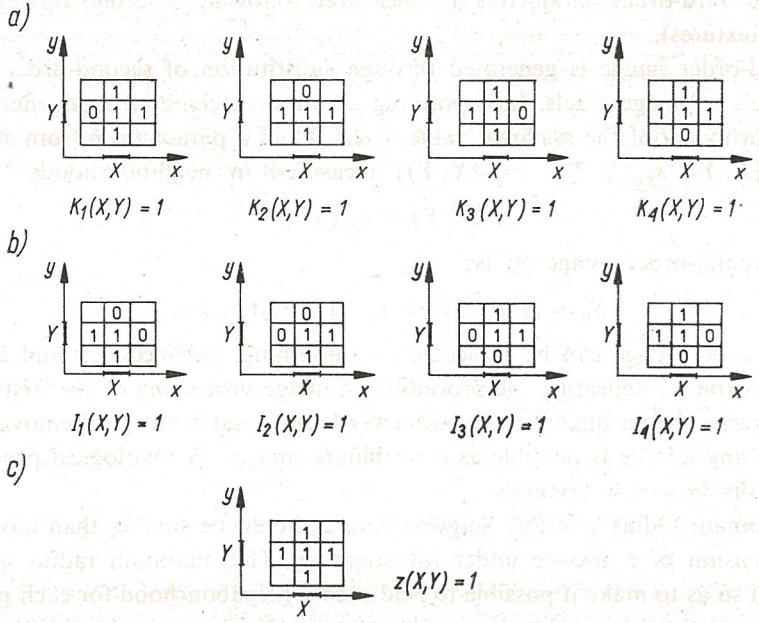


Fig. 8. Graphical interpretations of value 1 of the logical functions: a) $K_n(X, Y)$, b) $I_n(X, Y)$ and c) $z(X,Y)$

The initial and sufficient condition for the minimum error E_D:

$$E_D = \sum_{m=1}^{D} (L' - L_m)^2 \tag{37}$$

is the following:

$$\frac{\partial}{\partial a} E_D = \frac{\partial}{\partial b} E_D = 0. \tag{38}$$

Solution of the two equations (38) gives the two unknowns a, b.

The length L_A of perimeters covered by an operator field $0(X, Y)$ (see equations (33)—(35)) per area unit is:

$$L_A = \frac{L}{A_0} \tag{39}$$

where A_0 is all area of the operator field $0(X, Y)$ under investigation.

The same considerations refer to any part of the operator field $0(X, Y)$, e.g. to any position of a rotating neighbourhood.

An approximation A of real object area covered by an operator field $0(X, Y)$ is:

$$A = c_1 x + c_2 y + c_3 z \tag{40}$$

where:

$$z = \text{card}\{p(X, Y) \in 0(X, Y): z(X, Y) = 1\} \tag{41}$$

and the logical function $z(X, Y)$ is shown in fig. 8c.

The area A_A of objects covered by an operator field $0(X, Y)$ per area unit is:

$$A_A = \frac{A}{A_0}. \tag{42}$$

The third basic parameter is the number I of objects. It can be determined by the propagation method [5], with the aid of segmentation presented in point 5 of the paper, or from the following equation [9, 13]:

$$I = I_{max} - I_{min} \tag{43}$$

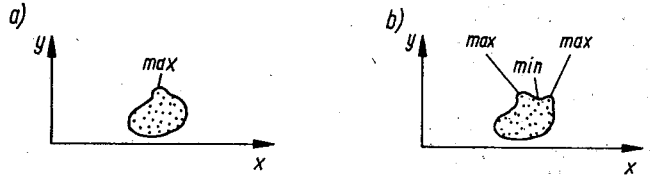


Fig. 9. An illustration of equation (41) for one local maximum (a) and after adding a single maximum more (b)

where I_{max} and I_{min} are numbers of local maxima and local minima of objects covered by an operator field and investigated in an image coordinate system. A very simple case of equation (43) is presented in fig. 9. Increasing the I_{max} by 1 for any single object results in increasing the I_{min} by 1. Therefore, equation (43) still holds true. Addition of any new single object results in the appearance of some new local maxima and minima, however, since the difference between them is 1 for any single object, the global difference between local maxima and local minima will increase only by 1 after the appearance of a single object. Any displacement or rotation of an image on a raster will result in appearing and disappearing of local minima and maxima; however, the difference between them will have remaind unchanged [9, 13]. Thus, the number I of objects fulfils the basic condition of representing images. The algorithm for counting the number I is presented in fig. 10. The number of objects N_A per unit area is:

$$N_A = \frac{I}{A_0}. \tag{44}$$

8. CONCLUSIONS

Results of all algorithms presented in this paper satisfy the basic condition on representing images because all operators are symmetrical with respect to their centres and all active image area is subjected to processing by the operators.

M. Nowakowska [4] has introduced graphs for the purposes of representation of visual scenes. Subsequently, various types of graphs have been developed in pattern recognition (see Rosenfeld A. and Kak C. [5]). Substantial innovations in one type of these graphs, int. al. substitution of simple nodes for simple graphs, discarding of auxiliary arcs which do not fulfil the basic condition of representing images, introduction of arcs representing one-argument relations of objects in images, revealment of synonymous

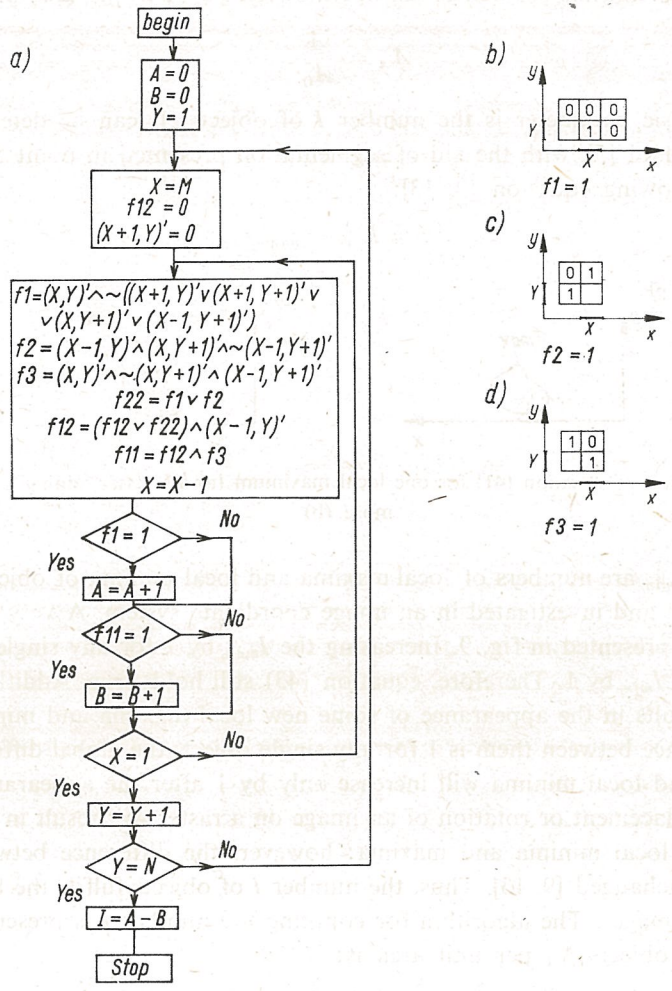


Fig. 10. Algorithm for counting the number I of objects (a) and illustrations of logical 1 of the functions $f1$ (b), $f2$ (c), and $f3$ (d) used in it

correspondences of graphs with sentences and phrases, have been worked out by Z. Wójcik [9, 13] to prove relevancy in using natural languages for the description of objects in pattern recognition.

The idea of checking distances between objects in images with the aid of the number of objects within an operator field has been introduced by P. E. Danielsson and B. Kruse [1]. Equipped with the analysis of dependances between the numbers of objects within an operator field and types of linear objects crossing the operator field, with the method of graph coding of every recognized line segment (including determination of non-argument relations of any line segment), with the method of reduction of graphs representing linear objects, the idea of calculating the number of objects covered by an operator field has been spread to a method of tracking and recognition of linear objects by Z. Wójcik

[10—16]. Types of line segments can be determined upon the number of objects covered by all the operator field and not only by its periphery. When a single line (e.g. contour) passes through an operator field, the number I_o of objects corresponding to background parts equals 2 and the number I_w of objects corresponding to line is equal to 1 (fig. 11a). A tip of a single line is recognized within an operator field if $I_w = 1$ and $I_o = 1$ (fig. 11b). Single bifurcation is implied if $I_w = 1$ and $I_o = 3$ (fig. 11c). A single junction is stated if $I_w = 1$ and $I_o = 4$ (fig. 11d). When the two non-crossing lines pass through the operator field, $I_w = 2$ and $I_o = 3$ (fig. 11e). A single break in a line within the operator field decreases the number I_o by 1 and increases I_w by 1.

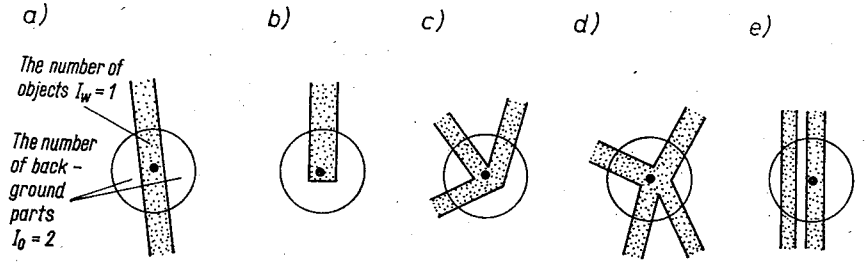


Fig. 11. Several cases of lines within an operator field: a) a single line; b) a tip; c) a bifurcation; d) a junction; e) two non-crossing lines

The diameter of the operator field should be fitted to the width of any line segment subjected to the tracking. The diameter can be determined using the distance-checking algorithms [1, 10].

The edge-preserving technique has been conceived by Z. Wójcik [7, 8]. Subsequently, the idea of investigating variance in grey levels of a rotating neighbourhood with regard to the average grey level of all the neighbourhood was utilized in [3] which is a variation upon the method by Z. Wójcik, and therefore, the smoothing process by [3] had to be iterated many times. In the original procedure [7, 8] the homogeneity in grey levels of a rotating neighbourhood with respect to grey level of the central pixel of the operator field is calculated.

Also, the method of thresholding incorporating the floating threshold with the coefficients k_1 and k_2 suppressing noise, and the formulation of the second-order image and second-order thresholding (textures) has been worked out by Z. Wójcik. Together with the basic condition on representing images which is helpful at all stages of pattern recognition and which has been formulated and proved by Z. Wójcik, the contents of the paper is a genuine approach to pattern recognition.

Generally, recognition algorithms generate symbols of an object language in a formal system of representation of events (Z. Wójcik [15]) in which events (including images) are considered as a metalanguage (fig. 12). Recognition algorithms derive graphs from images. Sentences describing these images are achieveable from these graphs directly through substitution of labels (parts of speech) of arcs and nodes of linear fragments of these graphs for variables (parts of sentence) in corresponding syntax formulae [9—15]

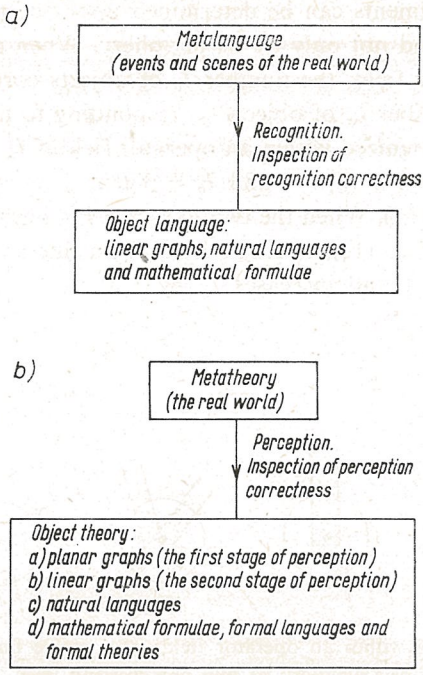


Fig. 12. Formal system of representation of the real world a) being subject of consideration of the paper; b) general formal system or the real world representation

e.g. in:

attributes of subject (adjuncts) — subject — predicate — — attributes of object — object — adverbials and obligatory attributes of predicate (including mood, tense) (45)

Thus, both graphs and sentences representing images are elements of the subject language and are generable from the metalanguage.

In the context of the above formal model, the metatheory is all the real world, and the theory is: the set of all the graphs (e.g. planar and linear), all natural languages describing events, and mathematical equations expressible through these sentences. Decidability of symbols of subject language is the general property of the formal system consisting of a metalanguage and a subject language. Each decidable expression of a formal theory is complete, consistent (non-discrepant) and reducible to axioms [17]. The object language formulated by Z. Wójcik [15] is decidable (conclusive) if:

a) the object language is complete, ie. if there exists a rule allowing to decide whether any symbol (not only the symbol of the object language) is element of the object language. The basic condition on representing images is the rule: a symbol is the member of the object language if the symbol represents an image and if the symbol is invariable with respect to the group of possible displacements of a recognition system within the field of observation of the image;

b) and if any symbol of the object language is consistent (non-discrepant), i.e. if the symbol is true, i.e. if the symbol satisfies the basic condition on representing images;

c) and if any symbol of the object language is reducible to axioms or if it is an axiom, i.e. if the symbol satisfies the basic condition on representing images.

Completness of the object language in the model of representation of images eliminates all symbols that do not stand for images, for instance, one can not uncritically describe a triangle using useless clauses such as: „It is not a person. It does not fly. It is not com-

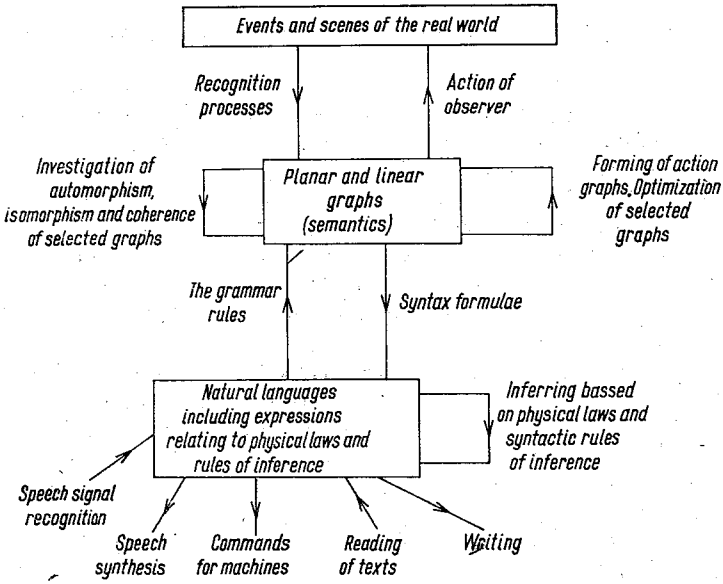


Fig. 13. A functional model of real world representation

fortable", etc. Definition of consistency and significance of axiom are here similar to that in formal theories.

A general model of real world representation discussed in the paper is shown in fig. 13. It is suggested in the model that another sentences describing images are generable from true sentences by using syntactic rules of inference and physical laws definitions.

APPENDIX

THE BASIC CONDITION ON REPRESENTING IMAGES

Any recognition system can be in any position x from among an infinite set $X = \{1, 2, \ldots x, \ldots\}$ of possible positions. No position x can be distinguished before the recognition is accomplished. Thus, the recognition process must be executed for any position x being member of the set X. Let us assume that for each position x one can obtain a symbol $w(x)$ representing the same fragment s of the image as a result of recognition process h. Since the set X is infinite, the infinite set $W(x) = \{w(1), w(2), \ldots w(x), \ldots\}$ of symbols is obtained. This infinite set $W(x)$ can be stored only within an infinite memory of the infinite time of construction. This is a contradiction, and therefore, each symbol

which represents any fragment of the real world is independent of possible displacements (and thus positions) of recognition system in the field of observation (for full formal proof see references [12, 13, 15]).

REFERENCES

1. P. E. Danielsson, B. Kruse, *Distance checking algorithms*, Comp. Graph. Image Proc., 4, 1979, 349—376. pp.
2. A. Guzman, *Decomposition of a visual scene into three dimmensional bodies*, in: *Automatic interpretation and classification of images*, ed. by A. Grasseli, New York, Acad. Press, 1969, pp. 243—276.
3. M. Nagao, T. Matsuyama, *Edge preserving smoothing*, Proc. Forth. Int. Conf. Pattern Recognition, Kyoto, Japan, 1978, pp. 518—520.
4. M. Nowakowska, *Some psychological problems of psychometry*, and, *Quantitative approach to the dynamics of perception*, General Systems, XII, 1967.
5. A. Rosenfeld, C. A. Kak, *Digital Picture Processing*, Acad. Press, 1966.
6. J. Ryś, *Introduction to Metallography*, Śląsk (publisher), Katowice (in Polish), 1972.
7. Z. M. Wójcik, *Automatic detection of semiconductor mask defects*, Microelectronics and Reliability, Perg. Press, v. 15, 1976.
8. Z. M. Wójcik, *A system for an automatic detection of defects of semiconductor masks and PC boards*, Electron Technology, 4, Warsaw, 1977 (published by the Institute of Electron Technology).
9. Z. M. Wójcik, *A model of pattern recognition*, and, *A decidability of the natural language*, Int. Symp. System-Modelling-Control, Zakopane, Poland, 1979 (org. by Polish Cybernetic Society).
10. Z. M. Wójcik, *Algorithms of recognition of lines and characters*, Rozprawy Elektrotechniczne, 1, 1980, Warsaw (in Polish).
11. Z. M. Wójcik, *A natural method of pattern recognition*, Proc. Int. Conf. Systems Eng., Coventry Polytechnic, England, 1980.
12. Z. M. Wójcik, *Praxiological studies of conterporary tools of man by means of the basic condition of representing events*, Prakseologia, 3 (published by the Polish Academy of Sciences), Poland, 1980
13. Z. M. Wójcik, *A model of semantics of the natural language*, and, *A fundamental condition for events representation processes and its application in pattern recognition*, in Progress in Cybernetics and Systems Research, v. 11, Hemisphere, Washington, 1982.
14. Z. M. Wójcik, *An approach to recognition of contours and line-shaped objects*, CVGIP, Acad. Press, 1984, 25, pp. 184—204.
15. Z. M. Wójcik, *Conclusiveness of natural language and recognition of images*, Cybernetics and Systems: An International Journal, Hemisphere, Washington 1983, 14, pp. 1—31.
16. Z. M. Wójcik, *Method of contour recognition*, Digital Systems for Industrial Automation, v. 2, no. 1, 1983, Crane, Russak and Company Inc.
17. R. C. Lyndon, *Notes on Logic*, D. Van Nostrand Reinhold Company Inc., 1966.

Z. M. WÓJCIK

NATURALNE PODEJŚCIE DO ROZPOZNAWANIA OBRAZÓW

Streszczenie

Praca przedstawia podejście do rozpoznawania obrazów, w którym obiekty są opisywalne za pomocą zdań i faz. Opis jest następnie przekształcany w zbiór grafów za pomocą reguł gramatycznych języka naturalnego.

Łuki i węzły tych grafów są oznaczane częściami mowy. W szczególności, węzły reprezentują relacje zero-argumentowe (tj. cechy) obrazu i są oznaczane rzeczownikami, łuki łączące dwa węzły

reprezentują relacje dwu-argumentowe na cechach, łuki przyłączone do pojedynczego węzła specyfikują cechę i reprezentują relacje jedno-argumentowe, a łuki skierowane reprezentują funkcje,

Operatory i aldorytmy rozpoznawania umożliwiające przekształcanie obrazów w przygotowane grafy są następnie do opracowania. Każdy operator odpowiada zbiorowi nazw oznaczającyh węzły grafu. Operatory biorą pod uwagę rozkład stopni szarości we wstępnych proceach przetwarzania oraz licsby obiektów, części tła i spójnych brzegów operatora kolistego w procesie śledzenia objektu, Wyniki rozpoznawania spełniają podstawowy warunek reprezentowania obrazów, jeśli są one niezależne od obrotów obrazu wokół centrum pola operatora i jeśli wszystkie punkty obrazu są przedmiotem przetwarzania obrazów. Obiekt przekształcony w graf otrzymuje nazwę wzorca, jeżeli rozkład nazw relacji w grafie objektu jast identyczny z rozkładem nazw we wzorcu.

Z. WÓJCIK

APPROCHE NATURELLE DE LA RECONNAISSANCE DES FORMES

Résumé

L'étude concerne l'approche de la reconnaissance des formes selon laquelle les objets sont décrits à l'aide des phrases de la langue naturelle. Cette description est ensuite transformée en un certain ensemble de graphes moyennant des règles de grammaire de la langue naturelle. Les parties du discours sont utilisées à des graphes en tant qu'étiquette des noeuds et des arcs. En particulier, les noeuds de ces graphes représentent les relations zéroarguments (c'est à dire les traits) de l'image et sont étiquetés par les substantifs ou par les pronoms; les arcs attachant deux noeuds représentent les relations biargumentes, dont les arguments sont des traits et les arcs sont étiquetés par des conjunctions ou des verbes, ou par des combinaisons verbe-prépositions; chaque arc, se rattachant à un seul noeud, spécifie un trait et représente une relation monoargumente, dont l'argument est ce trait et cet arc est étiqueté par un adjectif numéral ou un pronom, ou un adverbe etc., chaque arc sagittal représente une fonction et est étiqueté par un verbe ou une combinaison d'un verbe-préposition.

A partir des graphes sont élaborés des opérateurs et des algorithmes de reconnaissances capables de transformer des formes en graphes préparés. Chaque opérateur correspond à un ensemble de noms étiquetant des noeuds de graphe. En cas du traitement préliminaire des images, les opérateurs prennent en considération la répartition du niveau du grisé à l'intérieur d'un opérateur, et en cas de dépistage des objects on prend en considération le nombre des objets et le nombre des parties du fond à l'intérieur d'un opérateur et le nombre des périphéries jointes d'un opérateur. Les résultats de la reconnaissance sont conformes à la condition principale de représentation des images s'ils sont indépendents des rotations de l'image autour du centre d'un opérateur et si chaque point de l'image est traité.

Un objet transformé en graphe rečoit le nom d'un graphe standard (de gabarit) si un arrangement des noms des relations dans le graphe d'un object reconnu est le même que dans le graphe standard, c'est à dire si ces deux graphes possèdent la même structure qualitative. En certaines situations une structure quantitative est demandée pour ces deux graphes, par ex. un simulaire changement total de l'inclinaison d'une ligne, une uniformité de ces changements et aussi la longueur d'un contour.

Z. M. WÓJCIK

NATÜRLICHE ANNÄHERUNG AN DIE BILDERWAHRNEHMUNG

Zusammenfassung

Der Artikel befaßt sich mit der Annäherung an die Wahrnehmung von Bildern, deren Objekte mittels Sätze und Phrasen beschrieben werden. Die Beschreibung wird dann in eine Graphensammlung mittels grammatischer Regeln der natürlichen Sprache umgestaltet. Bogen und Knoten dieser Graphen werden

mit Sprachteilen bezeichnet. Im besonderen stellen Knoten die Relationen der Nullargumente (d.h. Eigenschaften) des Bildes dar, die zwei Knoten verbindende Bogen sind Zweiargumentrelationen hinsichtlich der Eigenschaften, die an einen einzelnen Knoten angeknüpften Bogen bezeichnen eine Eigenschaft und sind Einargumentrelationen, die gerichteten Bogen stellen Funktionen dar.

Die Wahrnehmungsoperatoren und -algorithmen, die die Umgestaltung der Bilder in die vorbereiteten Graphen ermöglichen, müssen dann bearbeitet werden. Jeder Operator entspricht einer die Graphenknoten bezeichnenden Namensammlung. Die Operatoren berücksichtigen die Verteilung der Graugrade bei Vorumformungsprozessen sowie die Objektzahl, Teile des Bildgrundes und der kohärenten Ränder des kreisförmigen Operators im Prozeß der Beobachtung des Objektes.

Die Ergebnisse der Wahrnehmung erfüllen die Bedingung der Darstellung des Bildes, wenn sie von den Rotationen des Bildes um das Zentrum des Operatorfeldes unabhängig sind, und wenn sämtliche Bildpunkte Gegenstand der Bildumformung sind. Das in einen Graphen umgeformte Objekt erhält die Bezeichnung eines Mustergraphen, wen die Verteilung der Relationsbezeichnungen der Objekte im Graphen identisch mit der Verteilung der Bezeichnungen im Mustergraphen ist.

З. М. ВУЙЦИК

НАТУРАЛЬНЫЙ ПОДХОД В РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Р е з ю м е

Обсужден подход к распознаванию изображений согласно которому объекты можно описать при помощи предложений натурального языка. После того это описание трансформируется в некоторое множество графов при помощи грамматических правил натурального языка. Части речи используются как этикеты узлов и дуг графов. В частности узлы этих графов представляющие соотношения нолевого аргумента изображения и дуги можно представить при помощи существительных и прилагательных. Дуги, которые связывают два узла представляют отношения двойного аргумента. Дуги представляют собой глаголы, наречия и комбинации глагол-наречие. Каждая дуга связанная с одним узлом определяет его черту и представляет соотношение одно-аргументовое, которого аргумент является чертой и которого дуга определена прилагательным или местоимением или нумерическим прилагательным. Каждая направлённая дуга представляет функцию и описана глаголом или комбинацией глагол-наречие. Разработаны операторы и алгоритмы идентификации способные трансформировать изображения в приготовленные графы. Любому оператору соответствует множество названий определяющих узлы графов. В случае первоначального исследования изображений операторы учитывают распределение уровня серого цвета и число объектов, а тоже число периметров связаных с оператором. Результаты этого изображения исполняют основное условие представления изображений если они независимы от вращений изображения вокруг оператора центра поля и если все пункты изображения разработаны.

Объект трансформированный в граф получает название стандарда если распределение названий реляции в графе объекта идентичное стандарду. Это значит, если два графа имеют ту же качественную структуру. Иногда одна структура нужна для двух графов, например, такое-же изменение наклонения или длина контуров.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 527—537

621.3.072.9
621.394.14

Synchronizacja słów sygnału kodowego

JAN S. KUNICKI (WARSZAWA)

Instytut Telekomunikacji Politechniki Warszawskiej

Otrzymano 12.5.1983

W artykule przedstawiono teoretyczne aspekty zagadnienia synchronizacji słów kodowych dla systemów cyfrowych pracujących z kodem transmisyjnym. Zaprezentowano na przykładach kodów *AMI-II* i *AMI-III* modele analityczne procesu kontroli transmisji oparte o opis kodu łańcuchem Markowa. Przedyskutowano zagadnienie wiarygodności decyzji przy wykrywaniu błędu podziału ciągu kodowego na słowa oraz omówiono parametry kodu ważne z punktu widzenia projektowania algorytmów synchronizacji. Rozważono zastosowanie dwu typów algorytmów statystycznych (*TSP* i *STI*) oraz dokonano porównania ich wydajności na podstawie praktycznego przykładu obliczeniowego wykonanego dla kodu *AMI-III* i kanału transmisyjnego o przeciętnym poziomie zakłóceń.

1. WSTĘP

Sygnały przesyłane w większości telekomunikacyjnych łączy *PCM* o średniej i dużej przelotności są sygnałami kodowymi. Oznacza to, że pierwotny, binarny sygnał informacyjny jest przekształcany, przed skierowaniem go do transmisji, przy użyciu kodu transmisyjnego, w celu nadania sygnałowi szeregu własności niezbędnych do jego przesłania poprzez kanał telekomunikacyjny [1].

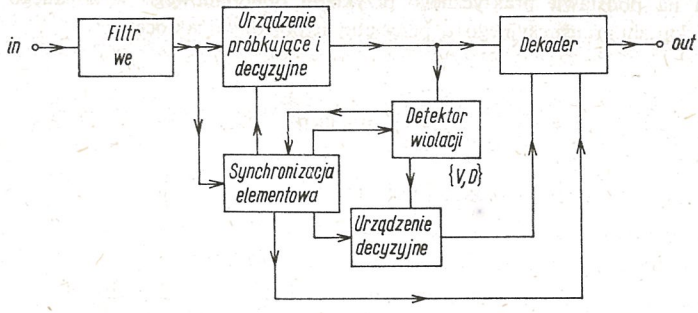
Stosowanie w tych systemach alfabetowych kodów transmisyjnych o skomplikowanej strukturze [10] stwarza, oprócz innych problemów [3, 5], zagadnienie synchronizacji słów sygnału kodowego [7], którego pominięcie, czy też zaniedbanie prowadzi do przykrych i trudnych do przewidzenia konsekwencji pojawiających się czasami dopiero w eksploatacji publicznej [8]. Jest to spowodowane faktem, że odbierany i interpretowany ciąg kodowy składa się z następujących po sobie kolejno słów kodowych o ustalonej i jednakowej długości, ale żadna dodatkowa informacja wskazująca początki słów nie jest do odbiornika przekazywana, gdyż jest to system samosynchronizujący się. Jednocześnie należy pamiętać, że jeśli słowo kodowe ma k symboli, to tylko jeden spośród k możliwych sposobów podziału ciągu kodowego na słowa jest poprawny, reszta z nich jest błędna. Z drugiej strony istnieje szereg różnorodnych przyczyn mogących spowodować utratę synchronizacji słów; są to między innymi: przerwy w transmisji, zmiany zasilania centralowego (w łączach komutowanych [8]), paczki błędów transmisji [9] i inne. Zagadnieniem

synchronizacji słów kodowych w samosynchronizujących się systemach impulsowych (głównie światłowodowych) zajmowało się wielu autorów [4, 6], lecz uzyskiwane wyniki odnosiły się wyłącznie do poszczególnych kodów transmisyjnych nie tworząc jednolitego systematycznego przejścia do zastosowań.

Celem tej pracy jest przedyskutowanie teoretycznych podstaw dla algorytmów pozwalających na możliwie szybkie podejmowanie odpowiednio wiarygodnych decyzji o wystąpieniu błędu podziału lub jego braku. Wykorzystane zostaną modele opisujące wyjściowy sygnał kodera, używające metod skończonych łańcuchów Markowa [1, 2, 11] oraz metody statystycznej teorii decyzji [12÷15]. Omówione zostaną zagadnienia: kontroli poprawności transmisji łącznie z podaniem modelu analitycznego, detekcji błędu podziału oraz wydajności proponowanych algorytmów synchronizacyjnych. Prezentowane rozważania ilustrowane będą przykładami zastosowań dla kodów światłowodowych *AMI-II* i *AMI-III* [16, 17, 1].

2. KONTROLA POPRAWNOŚCI TRANSMISJI

Rozważmy jaką informacją dysponuje układ odbiornika kontrolujący poprawność transmisji w samosynchronizującym się systemie impulsowym (patrz rys. 1). Umieszczony jest on za układami próbkującym i decyzyjnym, a więc obserwuje odebrany ciąg symboli



Rys. 1. Schemat blokowy odbiornika sygnału *PCM*

kodowych, wiedząc à priori, jakiego kodu użyto do ich zakodowania. Tak więc kontrola poprawności transmisji, wobec dowolności przekazywanej informacji, opiera się wyłącznie na ciągłym sprawdzaniu, czy odbierany ciąg kodowy spełnia wybrane reguły kodu transmisyjnego, który został użyty do jego wytworzenia. Jest to proces z natury rzeczy sekwencyjny, gdyż sekwencyjny jest również proces przekazywania informacji. Analizowane są dłuższe lub krótsze fragmenty odbieranego ciągu, czasami nawet zachodzące częściowo na siebie, wszystko zależnie od weryfikowanych nieustannie reguł charakterystycznych dla użytego kodu transmisyjnego. Każdy analizowany fragment ciągu kodowego opatrywany jest etykietą: D jeśli weryfikowane reguły są spełnione lub V w przeciwnym przypadku (D — dopuszczalny, V — wiolacja). Często podaje się również jaka reguła była weryfikowana, mówiąc np. wiolacja sumy cyfrowej słowa itp. W tej pracy uwaga zostanie skupiona na kolejnych słowach kodowych, jako kolejno analizowanych fragmentach

ciągu kodowego, jakkolwiek analogiczne rozważania mogą być łatwo przeprowadzone dla przypadku, gdy analizowane fragmenty obejmują kilka słów i następują bądź to kolejno po sobie, bądź też zachodzą na siebie. Z drugiej strony rozważanie kolejnych słów kodowych jest regułą w większości praktycznie stosowanych systemów z kodami transmisyjnymi.

Przyczyny powstawania wiolacji mogą być różne, ogólnie jednak należy odróżniać wiolacje będące wynikiem błędów transmisji od wiolacji wywołanych błędem podziału ciągu kodowego na słowa. Z praktyki wiadome jest, że przy wystąpieniu błędu fazy ilość pojawiających się wiolacji znacznie wzrasta i fakt ten stanowił podstawę dla większości dotychczas stosowanych algorytmów synchronizacji słów kodowych [4, 6, 7, 8]. Istotnym utrudnieniem w projektowaniu tych algorytmów jest fakt, że wiolacje wywołane błędami transmisyjnymi występują niezależnie od pojawiających się błędów podziału i co więcej nie są rozłożone równomiernie, a raczej mają tendencję do grupowania się [9]. Taka natura wiolacji wywołanych błędami liniowymi oraz konieczność podejmowania decyzji o poprawności podziału nawet w przypadkach, gdy średni wzrost liczby wiolacji wywołany błędem podziału jest niewielki, powodują konieczność szczegółowego zbadania powstającego problemu od strony teoretycznej i opracowania odpowiednich modeli pozwalających na uzyskanie praktycznych wniosków. W rozważaniach zawartych w następnych punktach główną rolę gra podejście wynikające ze statystycznej teorii decyzji, pozwalającej na utworzenie modeli analitycznych opisujących rozważane zjawiska.

3. MODEL PROCESU KONTROLI TRANSMISJI

W układzie z rys. 1 proces kontroli transmisji realizowany jest przez detektor wiolacji. Polega to na ciągłej obserwacji wyselekcjonowanych parametrów odebranego ciągu kodowego, analizowaniu ich wartości przez porównanie z regułami zastosowanego kodu transmisyjnego i podejmowaniu decyzji o wystąpieniu wiolacji (symbol V) lub fragmentu dopuszczalnego (symbol D). Tak więc układ synchronizacji słów odbiera informację z detektora wiolacji w postaci ciągu symboli D lub V zależnie od wcześniej podjętych decyzji. Własności nadchodzącej sekwencji porównywane są ze statystycznymi parametrami dwu możliwych przypadków (błąd podziału i jego brak) i następnie podejmowana jest decyzja, który z tych przypadków ma miejsce. Tak działa w skrócie każdy układ kontroli transmisji. Jego najbardziej skomplikowanym fragmentem jest detektor wiolacji, gdyż jego działanie powinno postępować za pracą kodera, a jednocześnie częściowo przewidywać kolejne jego stany.

Aby dokładnie zbadać pracę detektora wiolacji, warto zacząć od przypadku transmisji bezbłędnej, wtedy bowiem schemat pracy odpowiada dokładnie modelowi kodera.

W rozważanych tu kodach symbol 1 kodowany jest na przemian jako 11 lub 00, a symbol 0 jako 01 lub 10 stosownie do reguły kodu. W kodzie *AMI-II* [16] symbol 0 kodowany jest jako 01 po 11 i po 01, a jako 10 po 00 i po 10, w kodzie *AMI-III* [17] symbol 0 kodowany jest jako 10 po 11 i 01, a jako 01 po 00 i po 10. Reguły kodowania kodów *AMI-II* i *AMI-III* podane zostały w tablicach 1 i 2. W tablicach tych użyta została konwencja stosowana w literaturze (np. [1]).

Tablica 1

Kod *AMI-II*

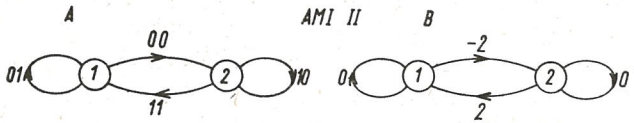
F	G_1	G_2
0	01,1	10,2
1	00,2	11,1

Tablica 2

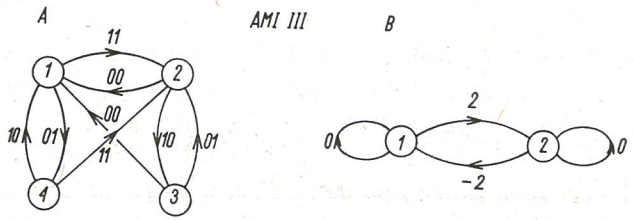
Kod *Ami-III*

F	G_1	G_2	G_3	G_4
0	01,4	10,3	01,2	10,1
1	11,2	00,1	00,1	11,2

Na rys. 2 i 3 pokazano grafy procesu kontroli transmisji dla wymienionych tu kodów. Są one identyczne z grafami opisującymi sygnały wyjściowe koderów [1, 2, 10], gdyż przyjęto, że transmisja jest bezbłędna. Warto zauważyć, że jakkolwiek dla kodu *AMI-II* grafy kontroli słów i sum cyfrowych słów mają identyczną strukturę i wartości przejść, czyli różnice polegają tylko na zmianie nazw, to dla kodu *AMI-III* obserwowanie sumy cyfrowych słów wprowadza istotne różnice w stosunku do obserwacji samych słów. Stany grafu z rys. 3A odpowiadają numerom alfabetów z tabl. 2, a stany grafu z rys. 3B reprezentują możliwe do otrzymania wartości kumulowanej sumy cyfrowej liczonej na końcu każdego słowa. Uproszczenie procesu kontroli przejawia się w zmniejszeniu liczby stanów



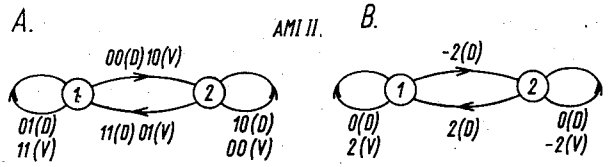
Rys. 2. Grafy procesu kontroli transmisji dla kodu *AMI-II* i przypadku bezbłędnej transmisji oraz poprawnej synchronizacji słów
Obserwowano: *A* — słowa, *B* — sumy cyfrowe słów



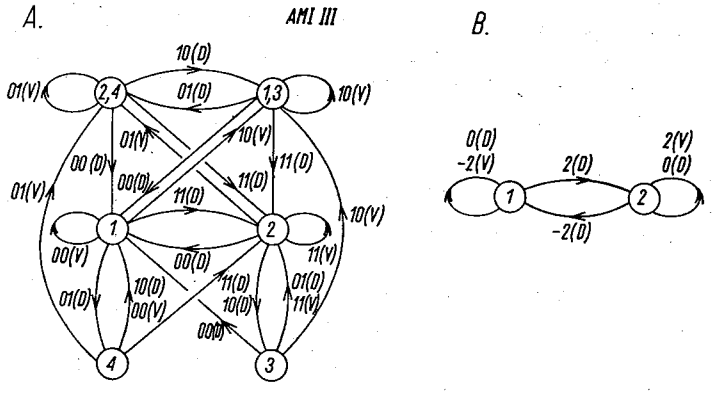
Rys. 3. Grafy procesu kontroli transmisji dla kodu *AMI-III* i przypadku bezbłędnej transmisji oraz poprawnej synchronizacji słów
Obserwowano: *A* — słowa, *B* — sumy cyfrowe słów

grafu reprezentującego ten proces. Jest to zgodne z prawidłowością, że funkcyjne przekształcenie procesu losowego może dać w wyniku procesu o strukturze izomorficznej lub prostszej, nigdy bardziej skomplikowanej.

Kolejnym krokiem przy modelowaniu procesu kontroli transmisji jest przedstawienie modelu uwzględniającego błędy (transmisji i podziału). W modelu takim, oprócz gałęzi reprezentujących wiolacje, pojawić się mogą nowe stany odzwierciedlające sytuacje, kiedy wskutek zaistniałych błędów zakłóceniu uległ proces wyznaczania alfabetów dla kolejnych interpretowanych fragmentów odbieranego ciągu kodowego.



Rys. 4. Grafy procesu kontroli transmisji dla kodu *AMI-II* uwzględniające możliwe błędy
Obserwowano: *A* — słowa, *B* — sumy cyfrowe słów



Rys. 5. Grafy procesu kontroli transmisji dla kodu *AMI-III* uwzględniające możliwe błędy
Obserwowano: *A* — słowa, *B* — sumy cyfrowe słów

Schematy z rys. 4 i 5 są ściśle związane ze schematami prezentowanymi dla transmisji bezbłędnej. Zmiany jakie można zauważyć na rys. 5A wynikają z analizy rys. 3A i tabl. 2. Warto zwrócić uwagę na to, iż po zaobserwowaniu wiolacji następuje przejście do stanu zależnego tylko od tej wiolacji np. po 10 (*V*) przejście następuje zawsze do stanu {1, 3}, a po 00 (*V*) do {1}. Wynika to z następującego rozumowania: otóż po wykryciu fragmentu (w rozważanym przypadku słowa), będącego wiolacją, zostaje ona zasygnalizowana przez nadanie symbolu *V*, natomiast fragment jest rozpatrywany dalej jako początek ciągu, którego kolejne elementy pozwolą na ustalenie właściwego następstwa alfabetów. Tak więc po zaobserwowaniu słowa 00, ocenionego jako wiolacja, traktuje się je jako słowo poprawne i na tej podstawie wskazuje alfabet 1, stan {1}, jako następny w grafie. Taka sama sytuacja występuje, gdy pojawi się słowo 10, ocenione jako wiolacja z tym tylko, że z tabl. 2 wynika, że do analizowania następnego słowa należy użyć albo alfabetu 1, albo 3.

Dlatego w grafie z rys. 5A występuje stan reprezentujący tę sytuację, zaznaczony jako {1, 3}. Porównując schematy z rys. 5A i 5B łatwo spostrzec, że jakkolwiek w schemacie 5A występują wiolacje postaci 01 i 10, to na rys. 5B nie są one uwzględnione (brak wiolacji 0 (V)). Jest tak dlatego, że schemat 5B reprezentuje sytuację, gdy obserwowana jest funkcja nadchodzącego ciągu kodowego, jeden z jego parametrów — suma cyfrowa słów, a ponieważ przyjęto uproszczenie przy obserwacji, część wiolacji pozostaje nierozpoznana. Jest to cena jaką płaci się za takie uproszczenie. Wszystkie przedstawione tutaj postaci grafów można otrzymać drogą analityczną, modelując sygnał kodowy łańcuchem Markowa {11}. Dotyczy to również schematów uwzględniających błędy. Jak okaże się to w następnym punkcie, ważnymi parametrami kodu są prawdopodobieństwa wiolacji: przy wystąpieniu błędu podziału P_{VF} i przy jego braku $P_{V\bar{F}}$. Parametry te można uzyskać przekształcając i analizując pokazane tu grafy metodami prezentowanymi w [2, 11]. Stanowi to jednak odrębny problem i nie będzie tu rozważane.

4. DETEKCJA BŁĘDU FAZY

W dalszym ciągu głównym celem pozostaje szybkie podjęcie odpowiednio wiarygodnej decyzji o tym, czy błąd podziału wystąpił, czy go nie było. Rozwijając należy stwierdzić, za statystyczną teorią decyzji, iż takie postępowanie polega na porównaniu wartości parametrów à priori z aktualnymi i decydowaniu w zależności od różnic. Dopuszczalne różnice określane są przez założone z góry dopuszczalne prawdopodobieństwa błędnych decyzji, określające wiarygodność procedury. Błędne decyzje to fałszywy alarm (F) (decyzja — błąd podziału, choć go nie ma) i strata (S) (decyzja — błędu nie ma, choć naprawdę jest). Tak więc parametry wiarygodnościowe procedury decyzyjnej to P_F i P_S oznaczające odpowiednie dopuszczalne prawdopodobieństwa błędnych decyzji. Są to parametry wyjściowe wyznaczone przez przewidywane zastosowania procedury.

W omawianym przypadku proces decyzyjny oparty jest na analizowaniu ciągu wyjściowych symboli detektora wiolacji, a więc na porównywaniu ich aktualnej częstości z wartościami à priori. Wartości te, tj. P_{VF} i $P_{V\bar{F}}$, czyli prawdopodobieństwa wiolacji przy błędzie podziału i prawdopodobieństwo wiolacji przy braku błędu, są dla ustalonego poziomu błędów liniowych wyznaczone jednoznacznie przez użyty kod transmisyjny. Ich wartości liczbowe mogą być wyznaczone przez analizowanie modeli z p. 3.

Operując terminami statystycznej teorii decyzji, należy dokonać testowania hipotez:

$$\begin{aligned} H_0 &= F, \\ H_1 &= \bar{F}, \end{aligned} \tag{1}$$

gdzie: H_0 — hipoteza główna, H_1 — hipoteza alternatywna, w momencie pojawienia się pierwszej wiolacji na wyjściu detektora wiolacji.

Zakładając, że symbole V i D reprezentują ciąg niezależnych zmiennych losowych Bernoulliego z parametrami P_{VF} lub $P_{V\bar{F}}$ zależnie od prawdziwości odpowiednich hipotez, stosować można dwa typy testów statystycznych: 1) test o stałej wielkości próby (TSP) i 2) sekwencyjny test ilorazowy (STI). Oba te testy przedstawione zostaną tylko z użytkowego punktu widzenia, gdyż ich aspekty teoretyczne są szeroko znane, np. [12, 13, 14]

Dla testu stałej próby droga postępowania jest następująca: należy określić jaka liczba symboli podlegać będzie sprawdzeniu, tj. rozmiar testu R oraz kiedy podejmować się będzie decyzję, tj. próg krytyczny K. Jest to równoznaczne z podaniem planu badania $K||R$. Rozwiązanie uzyskuje się poprzez analizę układu nierówności:

$$\begin{aligned} P(X_R \geqslant K|H_1) &< P_F, \\ P(X_R < K|H_0) &< P_S, \end{aligned} \tag{2}$$

gdzie X_R jest zmienną losową wskazującą liczbę symboli V wśród R kolejno analizowanych.

Rozkłady warunkowe zmiennej X_R są albo rozkładami dwumianowymi, przy niewielkich K i R, albo zbliżone do normalnego i wtedy rozwiązania uzyskuje się poprzez zastosowanie twierdzenia de Moivre'a-Laplace'a, dla dużych przewidywanych K i R. Wykonanie obliczeń przy dostępnych danych P_F, P_S, P_{VF}, $P_{V\bar{F}}$ nie przedstawia trudności rachunkowych, aczkolwiek w niektórych przypadkach użycie komputera może okazać się nieodzowne.

Sekwencyjny test ilorazowy jest testem, który pozwala uwzględnić informację zawartą w kolejności nadchodzących symboli V i D. Informacja taka jest zawsze tracona, gdy stosowany jest test *TSP*. Rozpatrywany jest iloraz wiarygodności T_m

$$T_m = T_{m-1} \begin{cases} T_V = \dfrac{P_{VF}}{P_{V\bar{F}}}, & \text{gdy odebrany jest symbol } V, \\ T_D = \dfrac{1-P_{VF}}{1-P_{V\bar{F}}}, & \text{gdy odebrany jest symbol } D, \end{cases} \tag{3}$$

$$T_0 = T_V, \quad m = 1, 2, \ldots,$$

odwzorowujący względne szanse prawdziwości hipotezy H_0 względem H_1 dla analizowanego fragmentu ciągu symboli V i D. Porównywany jest on z dwoma progami decyzji A i B

$$A = (1-P_S)/P_F, \quad B = P_S/(1-P_F), \tag{4}$$

mówiącymi o granicznych, możliwych do akceptacji, wartościach obliczanych kolejno ilorazów T_m.

Kryterium podejmowania decyzji sformułowane jest w postaci następujących nierówności:

$$\begin{aligned} T_m &\geqslant A \Rightarrow \text{prawdziwa jest hipoteza } H_0, \\ T_m &\leqslant B \Rightarrow \text{prawdziwa jest hipoteza } H_1, \end{aligned} \tag{5}$$

$T_m \in (A, B) \Rightarrow$ decyzja może być podjęta po przeanalizowaniu kolejnego symbolu.

Tak więc w przypadku *STI* liczba analizowanych elementów nie jest z góry ustalona, a zależy od ich kolejności w rozpatrywanym fragmencie ciągu. Może oczywiście zdarzyć się przypadek, że będzie trzeba przeanalizować wiele symboli zanim decyzja zostanie podjęta, ale teoria [12÷15] mówi, że prawdopodobieństwo takiej sytuacji jest bardzo małe i dąży do zera wraz z wydłużaniem się fragmentu podlegającego analizie. Praktyka wskazuje, że postępowanie sekwencyjne kończy się zazwyczaj bardzo szybko i tym samym wydajność algorytmów sekwencyjnych jest duża w porównaniu do algorytmów stałopróbowych.

Porównując testy *STI* i *TSP* należy stwierdzić, że *STI* wymaga średnio znacznie mniej elementów do analizy, aby podjąć decyzję na tym samym poziomie wiarygodności. Jest to łatwe do zaakceptowania, jeśli zauważyć, że *STI* wykorzystuje dodatkową, w porównaniu do *TSP*, informację zawartą w kolejności nadchodzących symboli *V* i *D*. To spostrzeżenie potwierdza również teoria. Z drugiej strony przy praktycznych realizacjach układowych obu rozważanych algorytmów należy liczyć się ze zwiększoną komplikacją, a tym samym kosztem układów realizujących *STI*.

5. PRZYKŁADY UŻYTKOWANIA I PORÓWNANIE WYDAJNOŚCI PROPONOWANYCH ALGORYTMÓW

W tym rozdziale zostaną podane przykłady ilustrujące sposób użytkowania omówionych poprzednio algorytmów synchronizacji słów kodowych oraz przedstawione wyniki obliczeń dla możliwych praktycznych zastosowań. Zaprezentowane dane liczbowe dostarczą przesłanek do porównania ilościowego wydajności proponowanych algorytmów.

5.1. Użytkowanie algorytmu stałopróbowego (*TSP*)

Mając dane w postaci liczbowej parametry P_{VF}, $P_{V\bar{F}}$, P_S, P_F, należy skorzystać z zależności podanych w p. 4 i dokonać obliczeń niezbędnych do wyznaczenia planu badania $K||R$, czyli rozmiaru próby R i progu K.

Następnie nadchodzącą informację podaje się na liczniki: pierwszy zliczający wszystkie symbole po kolei i sprawdzający każdorazowo czy drugi, liczący wiolacje nie został przepełniony. Jeśli pierwszy przepełni się wcześniej niż drugi, należy przyjąć, że błąd podziału nie wystąpił. Gdy drugi licznik zostanie przepełniony wcześniej niż pierwszy, badanie można przerwać, gdyż spełnione zostały warunki uprawniające do przyjęcia, że wystąpił błąd podziału. Tak więc w tym przypadku liczba symboli badanych może być mniejsza niż R (oczywiście trzeba sprawdzić co najmniej K symboli).

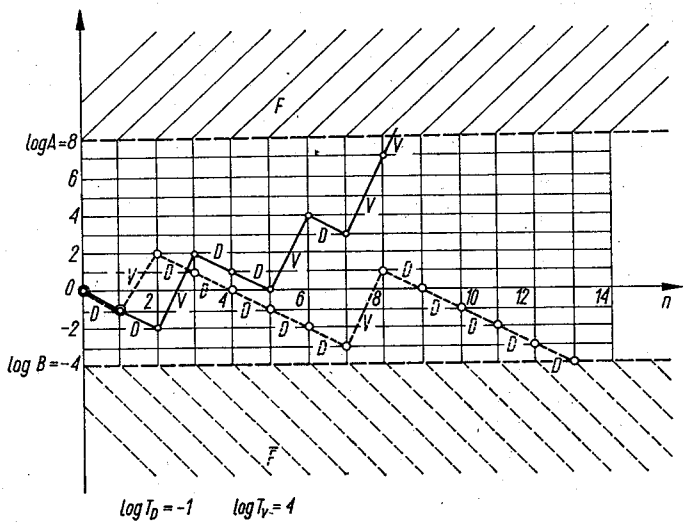
5.2 Użytkowanie algorytmu sekwencyjnego (*STI*)

Działanie algorytmu sekwencyjnego po uprzednim wyznaczeniu jego parametrów A, B T_V, T_D na podstawie znanych P_{VF}, $P_{V\bar{F}}$, P_F, P_S najlepiej jest przedstawić na przykładzie zastosowania kodu *AMI-III*.

Na rysunku 6 wyraźnie widać, w jaki sposób przebiega proces podejmowania decyzji przez odpowiednie urządzenie zależnie od wzoru nadchodzącej do badania sekwencji symboli *V* i *D*.

5.3 Przykład praktycznych obliczeń dla kodu *AMI-III*

W tablicy 3 przedstawiono przykładowe wyniki obliczeń wynikowych parametrów obu typów algorytmów wykonanych dla kodu *AMI-III*, założonego prawdopodobieństwa błędu transmisji P oraz przyjętych à priori P_F i P_S. Obserwowano wiolacje alfabetowe i wiolacje sumy cyfrowej słów.



Rys. 6. Przykład działania algorytmu sekwencyjnego dla kodu *AMI-III*

Tablica 3

Wyniki obliczeń wynikowych parametrów algorytmów *TSP* i *STI* wykonanych dla kodu *AMI-III*

Algorytm	*TSP*	*STI*		
Wiolacje	*KR*	E_F (*m*)	$E_{\bar{F}}$ (*m*)	*E*(*m*)
Alfabetowe	4 || 170	21	69	45
Sumy cyfrowej	6 || 666	132	325	228
Założono: $P = 0{,}001$, $P_S = 1{,}28 \cdot 10^{-12}$, $P_F = 1{,}12 \cdot 10^{-20}$				

Z danych zamieszczonych w tabl. 3 wynika, że algorytm sekwencyjny wymaga mniej symboli do podjęcia decyzji, tak samo wiarygodnej jak w przypadku *TSP*. Warto również zauważyć, że w przypadku gdy obserwowane są tylko wiolacje sumy cyfrowej słów czas podejmowania decyzji wydłuża się. Jest to jasne, gdy brać pod uwagę fakt, że w tym przypadku część wiolacji możliwych do wykrycia przy innym sposobie obserwacji zostaje stracona.

6. ZAKOŃCZENIE

Wskazanie ogólnych metod synchronizowania ciągów kodowych staje się ważnym zagadnieniem w momencie wchodzenia w użycie dużej liczby skomplikowanych kodów transmisyjnych przeznaczonych do stosowania w różnorodnych systemach cyfrowych, zwłaszcza gdy występują słowa o większej niż dwa liczbie symboli. Użycie metod statystycznej teorii decyzji umożliwia precyzyjne określenie wiarygodności podejmowanych

decyzji, a nawet uzyskanie reguł decyzyjnych o założonej wiarygodności. Dzięki stosowaniu sekwencyjnego testu ilorazowego możliwe jest skrócenie czasu podejmowania decyzji, a tym samym poprawienie jakości transmisji. Używając prezentowanych metod można wypowiedzieć się o przydatności poszczególnych sposobów kontroli jakości transmisji (czyli typu obserwowanych wiolacji) do stosowania w systemach synchronizacji słów kodowych. Problem wyznaczania paramtetrów P_{VF} i $P_{V\bar{F}}$ stanowi odrębne zagadnienie wymagające szerszego przedyskutowania.

BIBLIOGRAFIA

1. M. Rydel, *Charakterystyki kodów transmisyjnych dla światłowodowych systemów telekomunijkacyjnych*, Rozprawy Elektrotechniczne, 1980, 26, z. 3.
2. M. Rydel, *Analiza sygnałów kodowych*, Zeszyty Naukowe PW Elektronika, 1979, z. 39.
3. J. Kunicki, *Wpływ kodu transmisyjnego na pracę systemu z modulacją cyfrową*, Referaty IT-PW, 1978, z. 56.
4. V. Castellani, F. Preparata, *Synchronization Properties of an Alphabetic Multilevel Nonlinear Code*, Alta Frequenza, vol. XLVI, n 6, Giugno 1977.
5. M. Rydel, *Problemy korelacyjno-widmowej analizy kodów transmisyjnych*, Referaty IT-PW, z. 32.
6. R. Pietroiusti, *Binary Line Codes For Digital Transmission on Optical Fibres*, CCITT Study Group XVIII Contribution, Question 13/XVIII, Feb. 1979.
7. M. Słomiński, J. Kunicki, *Proces synchronizacji sygnału kodowego*, IX Symp. PTETiS MME, Pokrzywna, 1980.
8. M. Zawadzki, Informacja ustna. IT-PW, 1983.
9. E. Bloch i in., *Model istočnika ošibok w kanalach pieredači čifrowoj informačji*, Izd. Swiaź, Moskwa 1971.
10. J. Kunicki, M. Rydel, *Kody transmisyjne*, Refefraty IT-PW, 1979, z. 79.
11. J. Kunicki, *Modelowanie sygnału kodowego łańcuchem Markowa*, Rozprawy Elektrotechniczne (w druku).
12. M. Fisz, *Rachunek prawdopodobieństwa i statystyka matematyczna*, PWN, 1967.
13. M. H. DeGroot, *Optimal Statistical Decisions*, McGrow-GHill, 1970.
14. E. L. Franks, *Teoria sygnałów*, PWN, 1975.
15. E. L. Lehmann, *Testowanie hipotez statystycznych*, PWN, 1968.
16. Y. Takasaki i in., *Optical pulse formats for fiber optic digital communication*, IEEE Trans. on Comm., vol. COM-24, n. 4, 1976.
17. R. Pietrovic, *A new transmission code for digital optical communication*, Electronic Letters, vol. 14, 1978.

J. KUNICKI

WORD SYNCHRONIZATION FOR CODED SOIGNALS

Summary

The article deals with theoretical aspects of codeword synchronization for numerical systems operating with a transmission code. The author presents analytical models of the process of transmission control based on a description of a code in terms of Markov chains. Two codes, *AMI-II* and *AMI-III*, are used as the examples. Problems of the likelihood of correct decision are discussed with respect to the discovery of errors in partitioning of a code string into words. The parameters of the code that are important in

designing algorithms of synchronization are singled out and analyzed. Two types of statistical algorithms, *TSP* and *STI*, are discussed, and their efficiency is compared by reference to a practical computation problem to be solved with the use of code *AMI-III* and a transmission channel of an average level of interferences.

J. KUNICKI

SYNCHRONISATION DES MOTS DU SIGNAL DE CODE

Résumé

On a discuté le problème de synchronisation des mots du signal de code pour le système *PCM*. On a étudié les plus importants paramètres de détermination des algorithmes suffisanment efficaces et infaillibles. On a présenté deux algorithmes (*TSP* et *STI*) et on a discuté brièvement leurs relations et paramètres caractéristiques. On a donné les exeples des codes de transmission typiques, tels que *AMI-II* et *AMI-III*.

J. KUNICKI

WORTSYNCHRONISIERUNG BEI KODESIGNALEN

Zusammenfassung

Es wurden Probleme der Wortsynchronisierung bei Kodesignalen für *PCM*-Systeme diskutiert. Anhand von Beispielen der *AMI-II* und *AMI-III* Kodes wurden analytische Modelle für die Kontrollprozesse der Transmission geschildert in Anlehnung an die Kodebeschreibung mittels einer Markow-Kette. Es wurden das Problem der Entscheidungglaubwürdigkeit beim Fehlernachweis der Teilung von Kodefolgen zu Worten erwogen, als auch die vom Gesichtspunkt der Projektierung von Synchronisierungsalgorithmen wichtigen Kodeparameter. Dargestellt wurde die Anwendung von zwei Typen statistischer Algorithmen (*TSP* und *STI*). Deren Leistungsfähigkeit wurde auf Grund eines praktischen Berechnungsbeispieles verglichen, das für den *AMI-III*-Kode und einen Transmissionskanal mit einem mittleren Störungsniveau angefertigt worden ist.

Я. КУНИЦКИ

СИНХРОНИЗАЦИЯ СЛОВ КОДОВЫХ СИГНАЛОВ

Резюме

Представлены некоторые проблемы конструирования оптимальных алгоритмов синхронизаций слов кодовых сигналов, при описании сигналов цепями Маркова. Принятые результаты выходят из общей статистической теории решений. Определены и продискутированы общие параметры кодов и алгоритмов синхронизации необходимые при конструировании синхронизационных устройств. Теоретические выводы иллюстрованы примерами кодов *AMI–II* и *AMI–III*.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 539—556

621.372.5

Warunki realizowalności i synteza pewnej klasy czwórników *RC* o strukturze drabinkowej

TERESA JAKUBOWSKA (WROCŁAW), MARIAN S. PIEKARSKI (WROCŁAW)

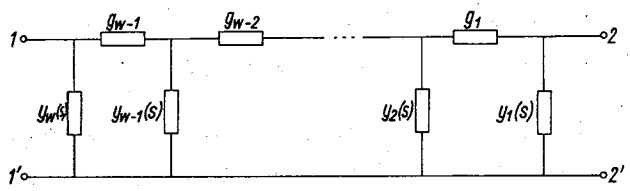
Instytut Telekomunikacji i Akustyki Politechniki Wrocławskiej

Otrzymano 24.11.1983

W pracy zbadano własności czwórników *RC* o strukturze drabinkowej takiej, że w gałęziach szeregowych występują rezystory, a gałęzie równoległe są dowolnymi dwójnikami *RC*. Sformułowano i udowodniono twierdzenie określające konieczne i dostateczne warunki realizowalności macierzy admitancji zwarciowych tych czwórników, które jest podstawą do opracowania algorytmu syntezy i generacji wszystkich możliwych układów równoważnych realizujących zadaną macierz. Podano również algorytm syntezy pozwalający uzyskać realizację danej macierzy admitancji zwarciowych w postaci czwórnika o tzw. minimalnej strukturze zredukowanej.

1. WSTĘP

Rozwój metod syntezy aktywnych układów *RC* spowodował potrzebę rozwiązania problemów syntezy macierzy beztransformatorowych układów pasywnych zbudowanych z elementów *RC*. Ogólnie postawiony problem beztransformatorowej syntezy pasywnych układów *RC* jest trudny i do chwili obecnej nie rozwiązany [1, 2]. W niniejszej pracy ograniczono się do klasy czwórników *RC* o strukturze drabinkowej pokazanej na rys. 1,



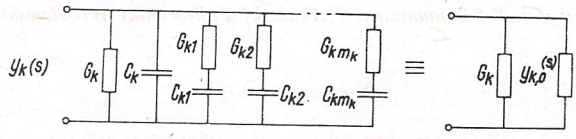
Rys. 1. Czwórniki *RC* o strukturze drabinkowej typu *SzR*

tzn. takiej, że w gałęziach szeregowych występują rezystory, a gałęzie równoległe są dowolnymi dwójnikami *RC*. Rozważana klasa czwórników będzie oznaczana w skrócie literami *SzR* (szeregowe rezystory) [6, 7]. Czwórniki *RC* o strukturze drabinkowej typu *SzR*, realizujące rzeczywiste zera transmitancji zarówno w nieskończoności jak i w skończoności, są uogólnieniem klasy rozważanej w pracy [3].

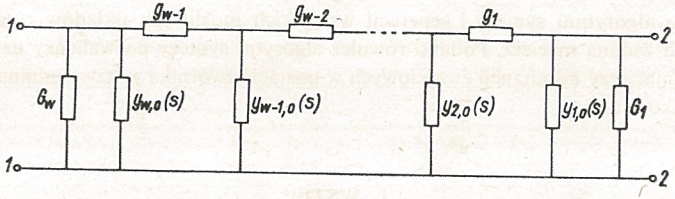
Admitancje $y_k(s)$ gałęzi równoległych czwórnika RC o strukturze drabinkowej typu SzR mają jednoznaczny rozkład o następującej postaci

$$y_k(s) = G_k + C_k s + \sum_{i=1}^{m_k} \frac{\alpha_{ki} s}{s + s_{ki}} = G_k + y_{k,o}(s), \tag{1}$$

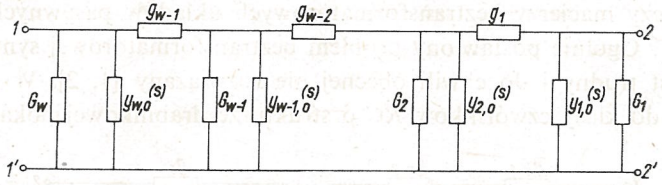
gdzie admitancje $y_{k,o}(s)$ są realizowane w układzie kanonicznym pokazanym na rys. 2. Wtedy czwórnik RC o strukturze drabinkowej typu SzR można przedstawić w postaci pokazanej na rys. 3. Struktura taka, gdy $g_i \neq 0$ $(i = 1, 2, \ldots, w-1)$, będzie nazywana dalej minimalną strukturą drabinkową SzR. W klasie czwórników RC o minimalnej strukturze drabinkowej SzR wyróżnia się takie struktury, w których konduktancje rów-



Rys. 2. Realizacja admitancji $y_{k,o}(s)$ w układzie kanonicznym



Rys. 3. Minimalna struktura drabinkowa typu SzR



Rys. 4. Minimalna realizacja zredukowana

noległe $G_k = 0$ dla $k = 2, 3, \ldots, w-1$ oraz $G_w \geqslant 0$ i $G_1 \geqslant 0$. Wyróżnione struktury będą nazywane minimalnymi strukturami zredukowanymi (rys. 4). Można udowodnić, podobnie jak w pracy [3], że każda realizacja zadanej macierzy admitancji zwarciowych $Y_z(s)$, o minimalnej strukturze drabinkowej typu SzR, daje się przekształcić do równoważnej (ze względu na macierz $Y_z(s)$) minimalnej realizacji zredukowanej, przy czym macierz admitancji węzłowych $Y_w(s)$ (z węzłem odniesień 1′ 2′) czwórnika o minimalnej strukturze drabinkowej i macierz admitancji węzłowych $Y'_w(s)$ równoważnej minimalnej realizacji zredukowanej powiązane są następującą zależnością

$$Y'_w(s) = A^T Y_w(s) A, \tag{2}$$

gdzie A jest dodatnio określoną macierzą diagonalną

$$A = \operatorname{diag}[1, a_2, a_3, \ldots, a_{w-1}, 1].$$

2. WŁASNOŚCI WIELOMIANÓW CHARAKTERYSTYCZNYCH

Niech czwórnik RC będzie opisany za pomocą macierzy admitancji zwarciowych $Y_z(s)$. Macierz $Y_z(s)$ można wyrazić za pomocą wielomianów charakterystycznych w następujący sposób:

$$Y_z(s) = \begin{bmatrix} y_{11}(s) & y_{12}(s) \\ y_{12}(s) & y_{22}(s) \end{bmatrix} = \frac{1}{n_{zz}(s)} \begin{bmatrix} n_{oz}(s) & -n(s) \\ -n(s) & n_{zo}(s) \end{bmatrix}, \tag{3}$$

$$\det Y_z(s) = \frac{n_{oo}(s)}{n_{zz}(s)},$$

gdzie wielomiany zbioru $N = \{n_{oz}(s), n_{zo}(s), n_{oo}(s), n_{zz}(s), n(s)\}$ są względnie pierwsze, tzn. stopień największego wspólnego dzielnika jest równy zeru oraz jest spełniona tożsamość

$$n_{zo}(s)n_{oz}(s) - n_{oo}(s)n_{zz}(s) = n(s)^2. \tag{4}$$

W dalszych rozważaniach będą używane następujące pojęcia:

— biegun kompaktny — biegun $(-s_i)$ macierzy admitancji, w którym warunek residuów, czyli $\operatorname{Re}_{s=-s_i}\left(y_{11}(s)\frac{1}{s}\right)\left(\operatorname{Res}_{s=-s_i} y_{22}(s)\frac{1}{s}\right) - \left(\operatorname{Res}_{s=-s_i} y_{12}(s)\right)^2 \geqslant 0$, jest spełniony ze znakiem równości,

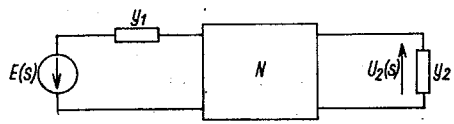
— biegun niekompaktny — biegun, w którym warunek residuów jest spełniony ze znakiem większości,

— biegun własny — biegun funkcji $y_{11}(s)$ i/lub $y_{22}(s)$, który nie jest biegunem funkcji $y_{12}(s)$,

— zera transmitancji — zera transmitancji napięciowej czwórnika (rys. 5)

$$T(s) = \frac{U_2(s)}{E(s)} = \frac{-y_1 y_{21}(s)}{[y_{11}(s)+y_1][y_{22}(s)+y_2] - y_{12}^2(s)}, \tag{5}$$

niezależne od y_1 i y_2.



Rys. 5. Współpraca czwórnika ze źródłem i obciążeniem

Z definicji tych pojęć, własności wielomianów charakterystycznych oraz z własności funkcji admitancji dwójników RC:

$$y_{11}(s) = \frac{n_{oz}(s)}{n_{zz}(s)}, \quad y_{22} = \frac{n_{zo}(s)}{n_{zz}(s)}, \quad \frac{1}{z_{11}(s)} = \frac{n_{oo}(s)}{n_{zo}(s)}, \quad \frac{1}{z_{22}(s)} = \frac{n_{oo}(s)}{n_{oz}(s)}$$

wynika, że:

a) wielomiany n_{oz}, n_{zo}, n_{oo} i n_{zz} nie mogą mieć zer o krotności większej niż 2,

b) wszystkie zera wielomianu $n_{zz}(s)$ $[n_{oo}(s)]$ są biegunami macierzy admitancji zwarciowych $Y_z(s)$ [impedancji rozwarciowych $Z_o(s)$],

c) pojedyncze zera wielomianu $n_{zz}(s)$ $[n_{oo}(s)]$ są kompaktnymi biegunami macierzy $Y_z(s)$ $[Z_o(s)]$, natomiast zera podwójne są biegunami niekompaktnymi macierzy $Y_z(s)$ $[Z_o(s)]$,

d) podwójne zera wielomianu $n_{oz}(s)[n_{zo}(s)]$ są zerami funkcji $y_{11}(s) = \dfrac{n_{oz}(s)}{n_{zz}(s)}$ $\left[y_{22}(s) = \dfrac{n_{zo}(s)}{n_{zz}(s)}\right]$.

W dalszych rozważaniach użyteczne będą następujące lematy.

Lemat 1. Jedynymi zerami wielomianu $n(s)$ są wszystkie skończone zera funkcji $y_{12}(s)$, wszystkie skończone bieguny własne funkcji $y_{11}(s)$ i/lub $y_{22}(s)$ oraz wszystkie niekompaktne bieguny $Y_z(s)$.

Dowód. W celu określenia zer wielomianu $n(s)$ rozważa się funkcję

$$y_{12} = -\frac{n(s)}{n_{zz}(s)}. \tag{6}$$

Z (6) wynika, że każde skończone zero funkcji $y_{12}(s)$ jest zerem wielomianu $n(s)$. Wszystkie pozostałe zera wielomianu $n(s)$ są również zerami wielomianu $n_{zz}(s)$. Z własności c) wielomianów charakterystycznych wynika, że jedynymi wspólnymi zerami wielomianów $n(s)$ i $n_{zz}(s)$ są wszystkie bieguny własne funkcji $y_{11}(s)$ i $y_{22}(s)$ oraz wszystkie niekompaktne bieguny macierzy $Y_z(s)$.

(c.b.d.o.)

Zera wielomianu $n(s)$ są skończonymi zerami transmitancji czwórnika. Zera transmitancji w nieskończoności są zerami funkcji $y_{12}(s)$ w nieskończoności i/lub biegunami własnymi funkcji $y_{11}(s)$ i $y_{22}(s)$ w nieskończoności.

Lemat 2. Skończone wartości funkcji $\dfrac{n_{zo}(s)}{n_{oo}(s)}, \dfrac{n_{zz}(s)}{n_{oz}(s)}, \dfrac{n_{zo}(s)+n_{zz}(s)\,y}{n_{oo}(s)+n_{oz}(s)\,y}$ w zerach transmitancji są sobie równe dla dowolnej wartości y.

Dowód. Z podstawowej tożsamości (4) otrzymuje się

$$\frac{n_{zo}(s)}{n_{oo}(s)} - \frac{n_{zz}(s)}{n_{oz}(s)} = \frac{n(s)^2}{n_{oo}(s)\,n_{oz}(s)}. \tag{7}$$

Wówczas mamy:

1. Dla zer transmitancji w nieskończoności $\operatorname{st} n(s) < \operatorname{st} n_{zz}(s)$ lub $\operatorname{st} n(s) = \operatorname{st} n_{zz}(s) < \max\{\operatorname{st} n_{oz}(s), \operatorname{st} n_{zo}(s)\}$. Z własności funkcji immitancji dwójników *RC* wynika, że $\operatorname{st} n_{oz}(s) \geqslant \operatorname{st} n_{zz}(s)$ oraz $\operatorname{st} n_{oo}(s) \geqslant \max\{\operatorname{st} n_{zo}(s), \operatorname{st} n_{oz}(s)\}$. Dlatego $\operatorname{st} n^2(s) < \operatorname{st} n_{oo}(s) n_{oz}(s)$ i

$$\lim_{s\to\infty} \frac{n(s)^2}{n_{oo}(s)\,n_{oz}(s)} = 0. \tag{8}$$

Z (7) i (8) wynika, że

$$\lim_{s\to\infty} \frac{n_{zo}(s)}{n_{oo}(s)} = \lim_{s\to\infty} \frac{n_{zz}(s)}{n_{oz}(s)}. \tag{9}$$

Korzystając z zależności (4) funkcję impedancji wejściowej czwórnika obciążonego dowolną admitancją y można wyrazić następująco

$$Z_{we}(s) = \frac{n_{zo}(s)+n_{zz}(s)y}{n_{oo}(s)+n_{oz}(s)y} = \frac{n_{zz}(s)}{n_{oz}(s)} \frac{1}{1-\dfrac{n(s)^2}{n_{oz}(s)\,[n_{zo}(s)+n_{zz}(s)y]}}. \tag{10}$$

Z (10) wynika, że gdy istnieje zero transmitancji w nieskończoności, to

$$\lim_{s\to\infty} Z_{we}(s) = \lim_{s\to\infty} \frac{n_{zz}(s)}{n_{oz}(s)}. \tag{11}$$

Z równości (9) i (11) widać, że funkcje $\dfrac{n_{zo}(s)}{n_{oo}(s)}$, $\dfrac{n_{zz}(s)}{n_{oz}(s)}$ i $\dfrac{n_{zo}(s)+n_{zz}(s)y}{n_{oo}(s)+n_{oz}(s)y}$ w zerach transmitancji w nieskończoności mają jednakowe wartości.

2. Dla zer transmitancji w skończoności z (7) wynika, że funkcje $\dfrac{n_{zo}(s)}{n_{oo}(s)}$ i $\dfrac{n_{zz}(s)}{n_{oz}(s)}$ oraz funkcja

$$Z_{we}(s) = \frac{n_{zo}(s)+n_{zz}(s)y}{n_{oo}(s)+n_{oz}(s)y} = \frac{n_{zz}(s)}{n_{oz}(s)} \frac{n_{zo}(s)+n_{zz}(s)y}{n_{zo}(s)+n_{zz}(s)y-\dfrac{n(s)^2}{n_{oz}(s)}} \tag{12}$$

są sobie równe we wszystkich tych zerach wielomianu $n(s)$, które nie są:
a) podwójnymi zerami wielomainu $n_{oo}(s)$ lub $n_{oz}(s)$,
b) pojedynczymi zerami wielomianu $n_{oo}(s)$ i $n_{oz}(s)$.
Można teraz wykazać, że rozważane funkcje dla zer transmitancji wyszczególnionych w przypadkach a) i b) są nieograniczone. Istotnie, w przypadku a) podwójne zera wielomianu $n_{oo}(s)$ są jednocześnie pojedynczymi zerami wielomianów $n_{oz}(s)$ i $n_{zo}(s)$. Wobec tego dla tych zer transmitancji funkcje $\dfrac{n_{zo}(s)}{n_{oo}(s)}$, $\dfrac{n_{zz}(s)}{n_{oz}(s)}$ i $Z_{we}(s)$ mają bieguny. Natomiast zera transmitancji, które są podwójnymi zerami wielomianu $n_{oz}(s)$, są jednocześnie pojedynczymi zerami wielomianów $n_{oo}(s)$ i $n_{zz}(s)$. Wobec tego również dla tych zer transmitancji rozważane funkcje mają bieguny.

W przypadku b) pojedyncze zero $s = -\alpha$ wielomianów $n_{oo}(s)$ i $n_{oz}(s)$ nie może być zerem wielomianu $n_{zo}(s)$, gdyż zero to jest biegunem macierzy $Z_o(s)$. Teraz, z zależności (4) otrzymuje się

$$n_{zo}(s)n'_{oz}(s)(s+\alpha)-n_{zz}(s)\cdot n'_{oo}(s)(s+\alpha) = (s+\alpha)^2 n'(s)^2.$$

Punkt osobliwy $s = -\alpha$ nie może być zerem wielomianu $n_{zz}(s)$, ponieważ $n_{zo}(-\alpha)n'_{oz}(-\alpha) \neq 0$. Wobec tego funkcje:

$$\frac{n_{zo}(s)}{n_{oo}(s)}, \quad \frac{n_{zz}(s)}{n_{oz}(s)} \quad \text{i} \quad Z_{we}(s) \text{ mają bieguny dla } s = -\alpha.$$

Słuszne jest więc twierdzenie, że skończone wartości rozważanych funkcji w zerach transmitancji są sobie równe.

(c.b.d.o.)

3. WARUNKI REALIZOWALNOŚCI

Warunki realizowalności macierzy $Y_z(s)$ czwórnika RC o strukturze drabinkowej typu SzR określa następujące twierdzenie.

Twierdzenie 1. Wymierna macierz $Y_z(s)$ określona zależnością (3), gdzie zbiór wielomianów N ma rzeczywiste współczynniki, jest realizowalna jako macierz admitancji zwarciowych czwórnika RC o strukturze drabinkowej typu SzR wtedy i tylko wtedy, gdy:

1) $n(s) = K$ lub $n(s) = K \prod_{i=1}^{m} (s+s_i)$, gdzie K jest stałą dodatnią, a wszystkie zera wielomianu $n(s)$, będące zerami transmitancji w skończoności, są rzeczywiste i ujemne ($s_i > 0$ dla $i = 1, 2, \ldots, m$);

2) dla zer wielomianu $n(s)$ i $n_{zz}(s) = K_1 \prod_{i=1}^{n} (s+z_i)$ uporządkowanych według rosnących modułów, to znaczy

$$0 < s_1 \leqslant s_2 \ldots \leqslant s_m,$$
$$0 < z_1 \leqslant z_2 \ldots \leqslant z_n,$$

jest spełniony warunek

$$s_i \geqslant z_i \quad \text{dla } i = 1, 2, \ldots, m;$$

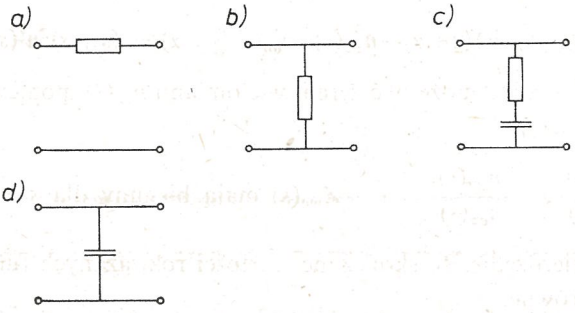
3) $0 < n(0) \leqslant \min\{n_{zo}(0), n_{oz}(0)\}$;

4) funkcje $y_{11}(s) = \dfrac{n_{oz}(s)}{n_{zz}(s)}$, $y_{22}(s) = \dfrac{n_{zo}(s)}{n_{zz}(s)}$, $\dfrac{1}{z_{11}(s)} = \dfrac{n_{oo}(s)}{n_{zo}(s)}$ i $\dfrac{1}{z_{22}(s)} = \dfrac{n_{oo}(s)}{n_{oz}(s)}$ są funkcjami admitancji dwójników RC;

5) $n_{zo}(s)n_{oz}(s) - n_{oo}(s)n_{zz}(s) = n^2(s)$.

Dowód. K o n i e c z n o ś ć. W rozważanej strukturze zera transmitancji są realizowane jako zera impedancji równoległych dwójników RC. Zera impedancji dwójników RC są rzeczywiste i ujemne, wobec czego warunek 1 jest warunkiem koniecznym. Warunek 3 jest warunkiem Fialkowa-Gersta [2] dla trójników realizowanych bez transformatorów, gdy $s = 0$. Warunki 4 i 5 są oczywiste. Konieczność warunku 2 będzie udowodniona poprzez indukcję. Czwórniki RC o strukturze drabinkowej typu SzR są kaskadowym połączeniem prostych czwórników RC przedstawionych na rys. 6.

Niech czwórnik N (rys. 7a) spełniający warunek 2 twierdzenia będzie opisany wielomianami charakterystycznymi $n_{oz}(s)$, $n_{zo}(s)$, $n_{zz}(s)$, $n_{oo}(s)$ i $n(s)$. Aby wykazać, że układy

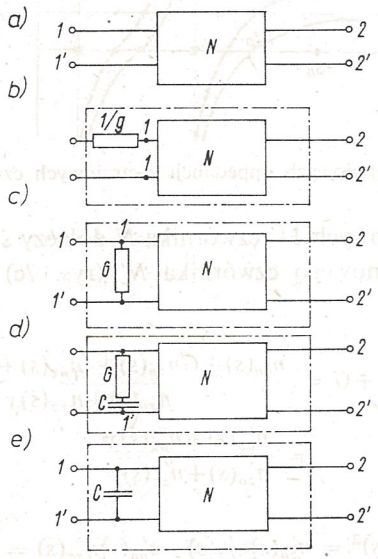


Rys. 6. Elementarne czwórniki RC

z rys. 7b÷e spełniają warunek 2 twierdzenia, zamiast wielomianów charakterystycznych można badać funkcję admitancji wejściowej $y_{we}(s)$ czwórnika N obciążonego dowolną admitancją y

$$y_{we}(s) = \frac{n_{oo}(s)+n_{oz}(s)y}{n_{zo}(s)+n_{zz}(s)y}. \tag{13}$$

Wielomian $n(s)$ nie występujący w zależności (13) można określić z warunku 5 twierdzenia.



Rys. 7. Sposób dołączania elementarnych czwórników *RC*

Z założenia wielomiany $n_{zz}(s)$ i $n(s)$ spełniają warunek 2 twierdzenia.

Dołączając do zacisków wejściowych 11' czwórnika N szeregowy rezystor o rezystancji $\frac{1}{g} > 0$, otrzymuje się czwórnik N' pokazany na rys. 7b. Admitancję wejściową czwórnika N' obciążonego admitancją y i wielomian $n'(s)$ określają zależności:

$$y'_{we}(s) = \frac{1}{\frac{1}{y_{we}(s)}+\frac{1}{g}} = \frac{n_{oo}(s)+n_{oz}(s)y}{n_{zo}(s)+\frac{1}{g}n_{oo}(s)+\left[n_{zz}(s)+\frac{1}{g}n_{oz}(s)\right]y} =$$

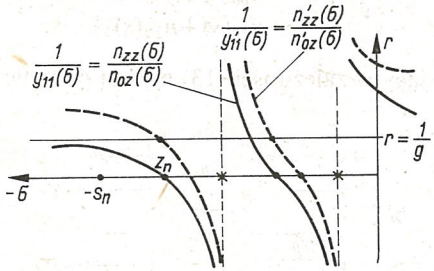
$$= \frac{n'_{oo}(s)+n'_{oz}(s)y}{n'_{zo}(s)+n'_{zz}(s)y}, \tag{14}$$

$$n'(s)^2 = n'_{zo}(s)n'_{oz}(s)-n'_{oo}(s)n'_{zz}(s) =$$

$$= n_{zo}(s)n_{oz}(s)-n_{oo}(s)n_{zz}(s) = n^2(s). \tag{15}$$

Funkcja $\frac{n_{zz}(s)}{n_{oz}(s)}\left[\frac{n'_{zz}(s)}{n'_{oz}(s)}\right]$ jest wejściową impednacją zwarciową czwórnika $N[N']$. Na rys. 8 pokazano typowe przebiegi tych funkcji oraz prostą o równaniu $r = \frac{1}{g}$.

Jak widać z rys. 8, zera wielomianu $n_{zz}(s)$ przesuwają się w kierunku początku układu współrzędnych. Ponieważ wielomian $n'(s)$ nie ulega zmianie, zatem warunek 2 dla czwórnika N' jest w dalszym ciągu spełniony.



Rys. 8. Wykres wejściowych impedancji zwarciowych czwórników N i N'

Jeżeli do zacisków wejściowych 11′ czwórnika N dołączy się czwórnik z rys. 6b, wówczas admitancja wejściowa nowego czwórnika N' (rys. 7c) obciążonego admitancją y wynosi

$$y'_{we}(s) = y_{we}(s)+G = \frac{n_{oo}(s)+Gn_{zo}(s)+[n_{oz}(s)+Gn_{zz}(s)]y}{n_{zo}(s)+n_{zz}(s)y} = \frac{n'_{oo}(s)+n'_{oz}(s)y}{n'_{zo}(s)+n'_{zz}(s)y}. \tag{16}$$

Ponieważ

$$n'(s)^2 = n'_{zo}(s)n'_{oz}(s)-n'_{oo}(s)n'_{zz}(s) = n_{zo}(s)n_{oz}(s)-n_{oo}(s)n_{zz}(s) = n(s)^2, \tag{17}$$

to warunek 2 jest również spełniony dla czwórnika N'.

Dołączając do zacisków wejściowych 11′ czwórnika N układ z rys. 6c otrzymuje się czwórnik N' pokazany na rys. 7d. Dla czwórnika N' obciążonego admitancją y słuszne są zależności:

$$y'_{we}(s) = y_{we}(s)+\frac{sG}{s+\frac{G}{C}} = \frac{n_{oo}(s)\left(s+\frac{G}{C}\right)+sGn_{zo}(s)+\left[n_{oz}(s)\left(s+\frac{G}{C}\right)+sGn_{zz}(s)\right]y}{n_{zo}(s)\left(s+\frac{G}{C}\right)+n_{zz}(s)\left(s+\frac{G}{C}\right)y} = \frac{n'_{oo}(s)+n'_{oz}(s)y}{n'_{zo}(s)+n'_{zz}(s)y}, \tag{18}$$

$$n'(s)^2 = n'_{zo}(s)n'_{oz}(s)-n'_{oo}(s)n'_{zz}(s) = \left(s+\frac{G}{C}\right)n(s)^2. \tag{19}$$

Ponieważ wielomiany $n_{zz}(s)$ i $n(s)$ spełniają warunek 2 twierdzenia, to, jak wynika z powyższych równań, warunek ten jest również spełniony przez wielomiany $n'_{zz}(s)$ i $n'(s)$ czwórnika N'.

Jeżeli do wejścia 11′ czwórnika N dołączy się równoległy kondensator (rys. 6d), wówczas otrzymuje się czwórnik N' (rys. 7e), którego admitancja $y'_{we}(s)$ wynosi

$$y'_{we}(s) = y_{we}(s)+sC = \frac{n_{oo}(s)+sCn_{zo}(s)+[n_{oz}(s)+sCn_{zz}(s)]y}{n_{zo}(s)+n_{zz}(s)y} = \frac{n'_{oo}(s)+n'_{oz}(s)y}{n'_{oz}(s)+n'_{zz}(s)y}. \quad (20)$$

Wielomian $n'(s)$ czwórnika N' określa zależność

$$n'(s)^2 = n'_{zo}(s)n'_{oz}(s)-n'_{oo}(s)n'_{zz}(s) = n_{zo}(s)n_{oz}(s)-n_{oo}(s)n_{zz}(s) = n(s)^2. \quad (21)$$

Ponieważ wielomiany $n'_{zz}(s)$ i $n'(s)$ nie ulegają zmianie, to warunek 2 jest nadal spełniony.

W podobny sposób można wykazać, że dołączając do zacisków wyjściowych 22′ czwórnika N elementarne układy z rys. 6, uzyskuje się czwórniki N' spełniające warunek rozkładu zer wielomianu $n(s)$. Tym samym udowodniono, że warunki podane w twierdzeniu są warunkami koniecznymi realizowalności macierzy $Y_z(s)$ czwórników RC o strukturze drabinkowej typu SzR.

Dostateczność. Z rozważań przeprowadzonych w punkcie 1 wynika, że warunki twierdzenia 1 będą warunkami dostatecznymi realizowalności macierzy $Y_z(s)$, jeżeli macierz $Y_z(s)$ można zrealizować w postaci czwórnika o minimalnej strukturze zredukowanej (rys. 4).

Niech czwórnik N opisany macierzą (3) będzie jednoznacznie określony zbiorem wielomianów charakterystycznych N spełniających warunki twierdzenia

$$N = \{n_{oz}, n_{zo}, n_{oo}, n_{zz}, n(s)\}. \quad (22)$$

Admitancja wejściowa czwórnika N obciążonego dowolną admitancją y określona jest zależnością (13). Z warunku 3 i 4 oraz z własności funkcji admitancji dwójników RC wynika, że:

$$\min R_e y_{11}(j\omega) = R_e y_{11}(0) = \frac{n_{oz}(0)}{n_{zz}(0)} > 0, \quad (23)$$

$$\min R_e y_{22}(j\omega) = R_e y_{22}(0) = \frac{n_{zo}(0)}{n_{zz}(s)} > 0, \quad (24)$$

oraz

$$G_w = \frac{n_{oz}(0)-n(0)}{n_{zz}(0)} \geqslant 0, \quad (25)$$

$$G_1 = \frac{n_{zo}(0)-n(0)}{n_{zz}(0)} \geqslant 0.$$

Możliwe jest więc wyłączenie rezystora o konduktancji G_w równoległego do zacisków wejściowych 11′ czwórnika N. Po wyłączeniu konduktancji G_w otrzymuje się czwórnik resztkowy N'. Admitancja wejściowa czwórnika N' obciążonego admitancją y wynosi

$$y'_{we}(s) = y_{we}(s)-G_w =$$

$$= \frac{n_{oo}(s)-G_w n_{zo}(s)+n_{oz}(s)-G_w n_{zz}(s)y}{n_{zo}(s)+n_{zz}(s)y} = \frac{n'_{oo}(s)+n'_{oz}(s)y}{n'_{zo}(s)+n'_{zz}(s)y} \quad (26)$$

oraz

$$n'(s) = n(s). \quad (27)$$

Można łatwo sprawdzić, że zbiór $N' = \{n'_{oz}, n'_{zo}, n'_{oo}, n'_{zz}, n'(s)\}$ spełnia wszystkie warunki twierdzenia. Niech $\mathscr{Z}$ będzie zbiorem skończonych i nieujemnych wartości funkcji $Z'_{we}(s) = \dfrac{1}{y'_{we}(s)}$ w zerach transmitancji, to znaczy

$$\mathscr{Z} = \{Z'_{we}(-s_i) | 0 \leqslant Z'_{we}(-s_i) < \infty, \quad s_i \in \mathscr{S}\}, \tag{28}$$

gdzie $\mathscr{S}$ jest zbiorem zer transmitancji

$$\mathscr{S} = \{s_1, s_2, \ldots, s_m, \infty\}. \tag{29}$$

Niech wartość rezystancji szeregowej $\dfrac{1}{g}$ będzie określona jako

$$\frac{1}{g} = \min \mathscr{Z}. \tag{30}$$

Warunek 2 twierdzenia gwarantuje, że zbiór $\mathscr{Z}$ nie jest zbiorem pustym. W przypadku, gdy nie występują zera transmitancji w nieskończoności, to st $n(s)$ = st $n_{zz}(s)$ = max {st n_{oz}, st n_{zo}}. Wówczas z warunku 2 otrzymuje się

$$s_n \geqslant z_n. \tag{31}$$

Z wykresu funkcji $\dfrac{1}{y'_{11}(\sigma)} = \dfrac{n'_{zz}(\sigma)}{n'_{oz}(\sigma)}$ (rys. 8) wynika, że przy spełnionym warunku (31) funkcja $\dfrac{n'_{zz}(-s_n)}{n'_{oz}(-s_n)} \geqslant 0$ oraz

$$0 \leqslant \frac{n'_{zz}(-s_n)}{n'_{oz}(-s_n)} < \lim_{s\to\infty} \frac{n'_{zz}(s)}{n'_{oz}(s)}. \tag{32}$$

Ponieważ skończone wartości funkcji $Z'_{we}(s)$ i $\dfrac{1}{y'_{11}(s)}$ w zerach transmitancji są sobie równe (lemat 2), to z (32) otrzymuje się

$$0 \leqslant Z'_{we}(-s_n) < \infty. \tag{33}$$

Wobec tego zbiór $\mathscr{Z}$ zawiera przynajmniej jeden element. W przypadku, gdy istnieją zera transmitancji w nieskończoności, wówczas zbiór $\mathscr{Z}$ jest zawsze niepusty, gdyż

$$0 \leqslant \lim_{s\to\infty} \frac{n'_{zz}(s)}{n'_{oz}(s)} = \lim_{s\to\infty} Z'_{we}(s) < \infty. \tag{34}$$

Ponieważ zbiór $\mathscr{Z}$ jest niepusty, to rezystancja szeregowa przyjmuje wartości:

I. $\dfrac{1}{g} = 0$, II. $\dfrac{1}{g} > 0$.

I. Gdy $\dfrac{1}{g} = 0$, wówczas może się zdarzyć, że:

a) $\dfrac{1}{g} = Z'_{we}(-s_i) = 0$ dla pewnego s_i, lub

b) $\dfrac{1}{g} = \lim\limits_{s\to\infty} Z'_{we}(s) = 0.$

W przypadku a) funkcja $y'_{we}(s) = \dfrac{1}{Z'_{we}(s)}$ ma biegun dla $s = -s_i$. Wtedy wielomiany $n'_{zo}(s)$ i $n'_{zz}(s)$ mają wspólny czynnik $(s+s_i)$. Wobec tego, jak wynika z lematu 1, zero transmitancji dla $s = -s_i$ jest biegunem własnym funkcji $y'_{11}(s)$ lub niekompaktnym biegunem macierzy $Y_z(s)$. W obu sytuacjach możliwe jest wyłączenie równoległej gałęzi o admitancji $\dfrac{\alpha_i \cdot s}{s+s_i}$, gdzie

$$\alpha_i = \operatorname*{Res}_{s=-s_i} \frac{y'_{we}(s)}{s} = \lim_{s\to -s_i} \frac{[n'_{oo}(s)+n'_{oz}(s)y](s+s_i)}{[n'_{zo}(s)+n'_{zz}(s)y]s} =$$

$$= \lim_{s\to -s_i} \frac{n'_{oo}(s)(s+s_i)}{n'_{zo}(s)s} \frac{1+\dfrac{n'_{oz}(s)}{n'_{oo}(s)}y}{1+\dfrac{n'_{zz}(s)}{n'_{zo}(s)}y}. \tag{35}$$

Z warunku 5 twierdzenia otrzymuje się

$$\lim_{s\to -s_i}\left(\frac{n'_{oz}(s)}{n'_{oo}(s)} - \frac{n'_{zz}(s)}{n'_{zo}(s)}\right) = \lim_{s\to -s_i} \frac{n'(s)}{n'_{oo}(s)n'_{zo}(s)} = 0. \tag{36}$$

Wobec tego z zależności (35) i (36) wynika, że

$$\alpha_i = \operatorname{Res} \frac{y_{we}}{s} = \operatorname{Res} \frac{n'_{oo}(s)}{n'_{zo}(s)s} = \operatorname{Res} \frac{1}{z'_{11}(s)s} =$$

$$= \operatorname{Res} \frac{y'_{11}(s)}{s} - \frac{\operatorname*{Res}_{s=-s_i} \dfrac{y'_{12}(s)^2}{s}}{\operatorname*{Res}_{s=-s_i} \dfrac{y_{22}(s)}{s}} \leqslant \operatorname{Res} \frac{y'_{11}(s)}{s}. \tag{37}$$

Po wyłączeniu równoległej gałęzi realizującej zero dla transmitancji $s = -s_i$ otrzymuje się czwórnik resztkowy N''. Admitancja $y''_{we}(s)$ czwórnika N'' wynosi

$$y''_{we}(s) = y'_{we}(s) - \frac{\alpha_i s}{s+s_i} =$$

$$= \frac{n'_{oo}(s) - \dfrac{\alpha_i s}{s+s_i} n'_{zo}(s)}{s+s_i} + \frac{n'_{oz}(s) - \dfrac{\alpha_i s}{s+s_i} n'_{zz}(s)}{s+s_i} y \Big/ \frac{n'_{zo}(s)}{s+s_i} + \frac{n'_{zz}(s)}{s+s_i} y = \frac{n'_{oo}(s)+n''_{oz}(s)y}{n''_{zo}(s)+n''_{zz}(s)y}, \tag{38}$$

a wielomian $n''(s) = \dfrac{n'(s)}{s+s_i}$.

Zbiór wielomianów $N'' = \{n''_{oz}, n''_{zo}, n''_{oo}, n''_{zz}, n''(s)\}$ spełnia warunki 1, 2, 3 i 5 twierdzenia. Korzystając z zależności (37) można również wykazać, że zbiór N'' spełnia także warunek 4. Wobec tego wszystkie warunki twierdzenia są w dalszym ciągu spełnione.

Ponadto z zależności (38), (26) i (25) otrzymuje się:

$$n''_{oz}(0) = \frac{n'_{oz}(0)}{s_i} = \frac{n(0)}{s_i} = \frac{n'(0)}{s_i} = n''(0),$$

$$\frac{n''_{zo}(0)-n''(0)}{n''_{zz}(0)} = \frac{\dfrac{n'_{zo}(0)}{s_i} - \dfrac{n'(0)}{s_i}}{\dfrac{n'_{zz}(0)}{s_i}} = \frac{n_{zo}(0)-n(0)}{n_{zz}(0)} = G_1. \tag{39}$$

W przypadku b) $y'_{we}(s) = \dfrac{1}{Z'_{we}(s)}$ ma biegun w nieskończoności i możliwe jest wyłączenie równoległego kondensatora o pojemności

$$C = \lim_{s\to\infty} \frac{y'_{we}(s)}{s} = \lim_{s\to\infty} \frac{n'_{oz}(s)}{sn'_{zz}(s)} \frac{\dfrac{n'_{oo}(s)}{n'_{oz}(s)} + y}{\dfrac{n'_{zo}(s)}{n'_{zz}(s)} + y}. \tag{40}$$

Gdy zero transmitancji występuje w nieskończoności oraz $\lim\limits_{s\to\infty} Z'_{we}(s) = 0$, to $\operatorname{st} n(s)^2 < \operatorname{st}\{n_{zz}\, n_{oz}\}$. Z warunku 5 twierdzenia otrzymuje się

$$\lim_{s\to\infty} \frac{n'_{zo}(s)}{n'_{zz}(s)} - \lim_{s\to\infty} \frac{n'_{oo}(s)}{n'_{oz}(s)} = \lim_{s\to\infty} \frac{n'(s)^2}{n'_{zz}(s)n'_{oz}(s)} = 0. \tag{41}$$

Z zależności (40) i (41) otrzymuje się

$$C = \operatorname*{Res}_{s\to\infty} \frac{y'_{we}(s)}{s} = \operatorname*{Res}_{s\to\infty} \frac{n'_{oz}(s)}{sn'_{zz}(s)} = \operatorname*{Res}_{s\to\infty} \frac{y'_{11}(s)}{s}. \tag{42}$$

Czwórnik resztkowy N'' będzie opisany funkcją

$$y''_{we}(s) = y'_{we}(s) - sC = \frac{n'_{oo}(s) - sCn'_{zo}(s) + [n'_{oz}(s) - sCn'_{zz}(s)]}{n'_{zo}(s) + n'_{zz}(s)y} =$$

$$= \frac{n''_{oo}(s) + n''_{oz}(s)y}{n''_{zo}(s) + n''_{zz}(s)y} \tag{43}$$

oraz wielomianem $n''(s) = n'(s)$.

Łatwo można wykazać, że zbiór N'' wielomianów charakterystycznych spełnia warunki 1, 2, 3 i 5 twierdzenia. Z lematu 2 wynika, że

$$C = \lim_{s\to\infty} \frac{n_{oz}(s)}{sn_{zz}(s)} = \lim_{s\to\infty} \frac{n_{oo}(s)}{sn_{zo}(s)}. \tag{44}$$

Z zależności (43) i (44) wynika, że również warunek 4 jest w dalszym ciągu spełniony.

Ponadto z (43), (26) i (25) otrzymuje się:

$$n''_{oz}(0) = n'_{oz}(0) = n(0) = n'(0) = n''(0),$$

$$\frac{n''_{zo}(0)-n''(0)}{n''_{zz}(0)} = \frac{n_{zo}(0)-n(0)}{n_{zz}(0)} = G_1. \tag{45}$$

II. Gdy $\frac{1}{g} > 0$, wówczas wyłączenie szeregowego rezystora o rezystancji $\frac{1}{g} > 0$ prowadzi do uzyskania czwórnika resztkowego N'' opisanego funkcją

$$Z''_{we}(s) = Z'_{we}(s) - \frac{1}{g} = \frac{n'_{zo}(s) - \frac{1}{g} n'_{oo}(s) + n'_{zz}(s) - \frac{1}{g} n'_{oz}(s) y}{n'_{oo}(s) + n'_{oz}(s) y} =$$

$$= \frac{n''_{zo}(s) + n'_{zz}(s) y}{n''_{oo}(s) + n''_{oz}(s) y}. \tag{46}$$

Wielomian $n''(s)$ czwórnika resztkowego jest równy wielomianowi $n'(s)$. Zbiór $N'' = \{n''_{oz}, n''_{zo}, n''_{oo}, n''_{zz}, n''\}$ spełnia warunki 1 i 5 twierdzenia. Z (30), (32) i lematu 2 wynika, że $\frac{1}{g} \leqslant Z_{we}(\infty) = \frac{n'_{zz}(\infty)}{n'_{oz}(\infty)} = \frac{n'_{zo}(\infty)}{n'_{oo}(\infty)}$. Stąd i z ogólnych własności funkcji impedancji dwójników RC wynika, że również warunek 4 jest spełniony. Aby wykazać, że zbiór N'' spełnia warunek 3, korzysta się z zależności (46), (26) i (25) i otrzymuje się następujące równości:

$$n''_{oz}(0) = n'_{oz}(0) = n_{oz}(0) - G_w n_{zz}(0) = n(0) = n'(0) = n''(0), \tag{47}$$

$$n''_{zo}(0) = n'_{zo}(0) - \frac{1}{g} n'_{oo}(s) = n'_{oo}(0) \left[\frac{n'_{zo}(0)}{n'_{oo}(0)} - \frac{1}{g} \right]. \tag{48}$$

Z warunków 2 i 4 oraz z (30), (32) i z lematu 2 wynika, że

$$\frac{1}{g} = Z_{we}(-s_i) = \frac{n_{zz}(-s_i)}{n_{oz}(-s_i)} \leqslant \frac{n'_{zz}(\infty)}{n'_{oz}(\infty)}. \tag{49}$$

Funkcja $\frac{1}{y'_{11}(s)} = \frac{n'_{zz}(s)}{n'_{oz}(s)}$ jako funkcja impednancji dwójnika RC spełnia następującą nierówność

$$\frac{n'_{zz}(0)}{n'_{oz}(0)} \geqslant \frac{n'_{zz}(\infty)}{n'_{oz}(\infty)}. \tag{50}$$

Zatem:

$$\frac{1}{g} \leqslant \frac{n'_{zz}(\infty)}{n'_{oz}(\infty)} \leqslant \frac{n'_{zz}(0)}{n'_{oz}(0)}. \tag{51}$$

Z zależności (48), (51) i warunku 5 twierdzenia otrzymuje się

$$n'_{zo}(0) = n'_{oo}(0) \left[\frac{n'_{zo}(0)}{n'_{oo}(0)} - \frac{1}{g} \right] \geqslant n'_{oo}(0) \left[\frac{n'_{zo}(0)}{n'_{oo}(0)} - \frac{n'_{zz}(0)}{n'_{oz}(0)} \right] =$$

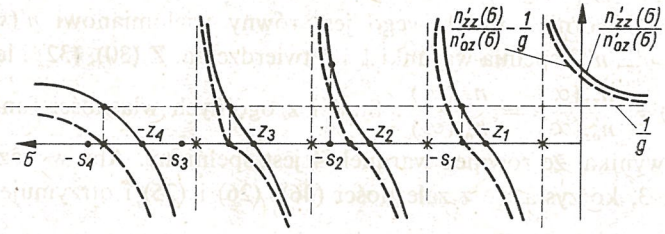
$$= \frac{n'(0)^2}{n'_{oz}(0)} = \frac{n'(0)^2}{n'(0)} = n'(0) = n(0). \tag{52}$$

Z zależności (47) i (52) wynika wprost, że warunek 3 jest w dalszym ciągu spełniony.

W celu wykazania, że wielomiany charakterystyczne czwórnika resztkowego N'' spełniają warunek 2, należy rozważyć wielomian

$$n''_{zz}(s) = n'_{zz}(s) - \frac{1}{g} n'_{oz}(s) = n'_{oz}(s) \left(\frac{n'_{zz}(s)}{n'_{oz}(s)} - \frac{1}{g} \right). \tag{53}$$

Zera wielomianu $n''_{zz}(s)$ są wspólmyni zerami wielomianów $n'_{zz}(s)$ i $n'_{oz}(s)$ oraz zerami funkcji $\frac{n'_{zz}(s)}{n'_{oz}(s)} - \frac{1}{g}$. Jak widać z rys. 9 zera funkcji $\frac{n'_{zz}(s)}{n'_{oz}(s)} - \frac{1}{g}$ przesuwają się względem zer funkcji $\frac{n'_{zz}(s)}{n'_{oz}(s)}$ na lewo od początku układu współrzędnych. Zależność (30) określa wartość rezystancji szeregowej $\frac{1}{g}$ jako minimalną dodatnią wartość $Z'_{we}(-s_i) = \frac{n_{zz}(-s_i)}{n_{oz}(-s_i)}$, gdzie $s_i \in \mathcal{S}$. Wobec tego, jeżeli wielomian $n'_{zz}(s)$, zgodnie z warunkiem 2, ma przynaj-



Rys. 9. Wykres wejściowych impedancji zwarciowych czwórników N' i N''

mniej i pierwiastków nie większych co do modułu od s_i $(i = 1, 2, \ldots, m)$, to również wielomian $n'_{zz}(s) - \frac{1}{g} n'_{oz}(s)$ spełnia ten warunek. Zatem wszystkie warunki twierdzenia są spełnione.

Z zależności (46), (26) i (25) otrzymuje się:

$$n''_{oz}(0) = n'_{oz}(0) = n'(0) = n''(0), \tag{54}$$

$$n''_{zo}(0) = n'_{zo}(0) - \frac{1}{g} n'_{oo} = n_{zo}(0) - \frac{1}{g} n(0) G_1$$

oraz

$$\frac{n''_{zo}(0) - n''(0)}{n''_{zz}(0)} = \frac{n_{zo}(0) - \frac{1}{g} n(0) G_1 - n(0)}{n_{zz}(0) - \frac{1}{g} n(0)} = \frac{G_1 n_{zz}(0) - G_1 \frac{1}{g} n(0)}{n_{zz}(0) - \frac{1}{g} n(0)} = G_1. \tag{55}$$

Po wydzieleniu szeregowego rezystora funkcja $Z''_{we}(s)$ ma zero dla $s = -s_i$ $[s_i \in \mathcal{S}]$. Taka sytuacja odpowiada przypadkowi I i możliwe jest wydzielenie równoległych gałęzi realizujących zera transmitancji. Wobec tego jest możliwe kolejne wydzielanie z wejścia 11′ czwórnika równoległych gałęzi realizujących zera transmitancji i szeregowych rezystorów określonych zależnością (30). Proces ten powoduje sukcesywne obniżanie stopni wielomianów charakterystycznych. W końcowym etapie otrzymuje się zbiór wielomianów charakterystycznych stopnia zerowego

$$N_o = \{n_{oz}, n_{zo}, n_{oo}, n_{zz}, n\}. \tag{56}$$

Zbiór N_o spełnia wszystkie warunki twierdzenia. Ponadto, jak wynika z (39), (45), (54) i (55), wartości $n_{oz}(0)$ i $\frac{n_{zo}(0) - n(0)}{n_{zz}(0)}$ są niezmiennicze, to znaczy po każdym etapie syn-

tezy otrzymuje się

$$n_{oz}(0) = n(0) \tag{57}$$

oraz

$$\frac{n_{zo}(0)-n(0)}{n_{zz}(0)} = G_1. \tag{58}$$

Z warunku 1 twierdzenia i z zależności (57) wynika, że:

$$\begin{aligned} n &= n_o = \text{const}, \\ n_{oz} &= n_o. \end{aligned} \tag{59}$$

Z zależności (58) otrzymuje się

$$n_{zo} = G_1 n_{zz} + n_o. \tag{60}$$

Zbiór n_o spełnia warunek 5, wobec czego

$$n_0^2 = n_{zo}n_{oz} - n_{oo}n_{zz} = n_o(G_1 n_{zz} + n_o) - n_{oo}n_{zz}. \tag{61}$$

Zależność (61) jest słuszna, gdy

$$(n_o G_1 - n_{oo})n_{zz} = 0. \tag{62}$$

Z (62) wynikają dwa przypadki:
a) $n_{zz} = 0$,
b) $n_{zz} \neq 0$, $n_o G_1 - n_{oo} = 0$,
W przypadku a) zbiór N_o ma postać

$$N_o = \{n_o, n_o, n_o G_1, 0, n_o\}. \tag{63}$$

Ponieważ jest możliwe wydzielenie równoległej konduktancji G_1, to w efekcie otrzymuje się zbiór N_o' reprezentujący zaciski wyjściowe 22' czwórnika

$$N_o' = \{n_o, n_o, 0, 0, n_o\}. \tag{64}$$

W przypadku b) zbiór N_o ma postać

$$N_o' = [n_o, G_1 n_{zz} + n_o, n_o G_1, n_{zz}, n_o]. \tag{65}$$

Z porównania zależności (65) z (46) i (26) wynika, że jest możliwe wydzielenie szeregowej rezystancji o wartości $\frac{1}{g} = \frac{n_{zz}}{n_o}$ oraz równoległej konduktancji G_1. Prowadzi to do zbioru N_o' o postaci (64) i proces syntezy jest zakończony.

(c.b.d.o.)

Z dowodu przeprowadzonego powyżej wynika, że zbiór warunków twierdzenia 1 jest jednocześnie zbiorem warunków koniecznych i dostatecznych realizowalności zbioru wielomianów charakterystycznych N w rozpatrywanej strukturze. Na podstawie przeprowadzonego dowodu można stwierdzić, że każdą macierz spełniającą warunki podane w twierdzeniu można zrealizować jako macierz admitancji czwórnika RC o minimalnej strukturze zredukowanej. Dowód dostateczności warunków jest jednocześnie algorytmem syntezy. Algorytm ten polega na rozkładzie funkcji $Y_{we}(s)$ na następujący ułamek łań-

cuchowy

$$y_{we}(s) = \frac{n_{oo}(s)+n_{oz}(s)y}{n_{zo}(s)+n_{zz}(s)y} = G_w + \sum_{i=1}^{r_1} \frac{s\alpha_i}{\frac{s}{p_{w_i}}+1} + \cfrac{1}{\cfrac{1}{g_{w-1}} + \cfrac{1}{\cfrac{1}{Z_{we_{w-1}}(s)} \longrightarrow \sum_{i=1}^{r_2} \cfrac{s\alpha_{w-1,i}}{\frac{s}{p_{w-1,i}}+1} \longrightarrow y_{we_{w-1}} + \cfrac{1}{\cfrac{1}{g_{w-2}} + \cfrac{1}{y_{we_{w-2}}(s) \longrightarrow \vdots + \cfrac{1}{\cfrac{1}{g_1} + \cfrac{1}{Z_{we_1} \longrightarrow \sum_{i=1}^{r_w} \cfrac{s\alpha_{1,i}}{\frac{s}{p_{1_i}}+1} \longrightarrow y_{we_1}(s) + G_1 + y}}}}}}, \tag{66}$$

gdzie:

$$G_w = \frac{n_{oz}(0)-n(0)}{n_{zz}(0)} \geqslant 0, \tag{67}$$

$$G_1 = \frac{n_{zo}(0)-n(0)}{n_{zz}(0)} \geqslant 0 \tag{68}$$

oraz

$$\frac{1}{g_k} = \min \mathscr{Z}_{we_k} = \min \{Z_{we_k}(-s_j) | Z_{we_k}(-s_j) \geqslant 0,\ s_j \in \mathscr{S}\} = z_{we_k}(-p_{ki}), \tag{69}$$

dla $p_{ki} \in \mathscr{S} = \{s_1, s_2, \ldots, s_m, \infty\}$ i $k = w-1, w-2, \ldots, 1$,

$$\alpha_{ki} = \lim_{s \to -p_{ki}} y_{we_k}(s) \frac{\frac{s}{p_{ki}}+1}{s}. \tag{70}$$

Z rozkładu (66) można łatwo określić wartości elementów czwórnika *RC* o minimalnej strukturze zredukowanej.

4. WNIOSKI

Stosując opisany algorytm uzyskuje się minimalne realizacje, w których liczba kondensatorów jest minimalna. Wyróżnienie w klasie czwórników *RC* o strukturze drabinkowej typu *SzR* minimalnych realizacji zredukowanych ma szczególne znaczenie, gdyż algorytm syntezy macierzy $Y_z(s)$ w tych strukturach jest prostszy. Poza tym, minimalne realizacje zredukowane są strukturami wyjściowymi do opracowanego prostego sposobu generacji wszystkich możliwych układów równoważnych realizujących macierz $Y_z(s)$ w klasie czwórników drabinkowych *RC* typu *SzR*. Mając na uwadze transformację

$RC \rightarrow LC$ [2], twierdzenie 1 może być wykorzystane do sformułowania koniecznych i dostatecznych warunków realizowalności odpowiedniej klasy układów drabinkowych LC, która jest szersza niż klasa rozpatrywana przez Fujisawę [5]. Szczegółowy algorytm syntezy i metoda generacji wszystkich możliwych, równoważnych czwórników RC o strukturze drabinkowej typu SzR w oparciu o macierz admitancji zwarciowych będzie omówiona w następnej pracy.

BIBLIOGRAFIA

1. A. D. Fialkow, *Inductance, capacitance networks terminated in resistance*, IEEE Transactions on CT, vol. CAS-26, Nr 8, August 1979.
2. L. Weinberg, *Network Analysis and Synthesis*, Mc Graw-Hill, New York 1962.
3. M. S. Piekarski, *Synteza i realizacje równoważne dwójników i czwórników RC o strukturze drabinkowej prostej*, Archiwum Elektrotechniki, z. 3, 1976.
4. M. Białko i inni, *Filtry aktywne RC*, Warszawa 1979.
5. T. Fujisawa, *Realizability theorem for mid-series or mid-shunt low-pass ladders without Mutual Induction*, IRE Trans. on CT, vol. CT-2, Nr 4, grudzień 1955.
6. T. Jakubowska, *Synteza i realizacje równoważne czwórników drabinkowych RC bez transformatorów*, Praca doktorska, Politechnika Wrocławska, Wrocław 1979.
7. T. Jabubowska, M. S. Piekarski, *Synteza i warunki beztransformatorowej realizacji pewnej klasy czwórników RC o strukturze drabinkowej*, II Krajowa Konferencja „Teoria obwodów i układy elektroniczne", Trzebieszowice 1978.

T. JAKUBOWSKA, M. S. PIEKARSKI

REALIZATION CONDITIONS AND SYNTHESIS OF A CLASS OF *RC* TWO-PORT LADDER NETWORKS

Summary

Properties of a class of *RC* networks are investigated. The class is defined as a ladder structure with resistors in the series branches and arbitrary one-port *RC* networks in the parallel branches. A theorem is formulated and proved which identifies the necessary and sufficient conditions for any matrix to be a short-circuit admittance matrix of *RC* two-ports. The theorem is used to work out an algorithm of synthesis which generates all possible systems equivalent to a given matrix. One more algorithm is discussed in the paper which enables realization of a given short-circuit admittance matrix in the form of a two-port minimal reduced structure.

T. JAKUBOWSKA, M. S. PIEKARSKI

CONDITIONS DE LA RÉALISATION ET SYNTHÈSE DES QUADRIPÔLES *RC* À STRUCTURE DE L'ÉCHELLE

Résumé

Dans l'article les propriétés des quadripôles *RC* à structure de l'échelle ont été examinées. Les quadripôles en question ont des résistances dans les branches en dérivation et les *RC* dipôles dans les branches en parallèle. On a formulé et prouvé le théorème présentant les conditions nécessaires et suffisantes pour réaliser les matrices admittances avec sortie en court-circuit. Ce théorème est la base pour élaborer l'algorithme de la synthèse et de la génération de tous les circuits possibles réalisant la matrice donnée. On a présenté aussi l'algorithme de la synthèse qui permet d'obtenir la réalisation de la matrice admittance avec sortie en courti-circuit en forme de quadripôle à structure réduite au minimum.

T. JAKUBOWSKA, M. S. PIEKARSKI

REALISIERUNGSBEDINGUNGEN UND SYNTHESE BESTIMMTER KLASSE VON *RC*-VIERPOLEN ALS KETTENBRUCHSCHALTUNGEN

Zusammenfassung

Es wurden Eigenschaften von *RC*-Vierpolen als Kettenbruchschaltungen untersucht, deren Reihenzweige aus Widerständen und Parallelzweigen aus beliebigen *RC*-Zweipolen bestehen. Man hat notwendige und hinreichende Realisierungsbedingungen für Admittanzmatrizen solcher Vierpole angeführt, und auf Grund dieser Bedingungen auch Synthesealgorithmen aller äquivalenten Schaltungen. Ein Syntheseverfahren wird ebenfalls angegeben, wodurch ein sog. minimaler reduzierter Vierpol gewonnen werden kann.

Т. ЯКУБОВСКА, М. С. ПЕКАРСКИ

УСЛОВИЯ РЕАЛИЗУЕМОСТИ И СИНТЕЗ НЕКОТОРОГО КЛАССА *RC* ЧЕТЫРЕХПОЛЮСНИКОВ ЛЕСТНИЧНОЙ СТРУКТУРЫ

Резюме

Исследованы свойства резистивно-емкостных четырехполюсников лестничной структуры, у которой рядовые ветви — резисторы а параллельные ветви — произвольные двухполюсники. Даны тоже необходимые и достаточные условия реализуемости матрицы проводимостей короткого замыкания схемы четырёхполюсника исследованной структуры. Приведен алгоритм реализации матрицы проводимостей короткого замыкания в виде минимальной сокращенной схемы.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 557—572

621.373.4

Sterowanie częstotliwością drgań sinusoidalnych generatorów *RC* za pomocą dodatkowego wzmacniacza

RYSZARD WOJTYNA (BYDGOSZCZ)

Instytut Telekomunikacji i Elektrotechniki Akademii Techniczno-Rolniczej w Bydgoszczy

Otrzymano 11.11.1983

W pracy omówione są generatory *RC* drgań sinusoidalnych zawierające dwa wzmacniacze (źródła napięciowe sterowane napięciem). Jeden z nich służy do umiejscowienia pierwiastków równania charakterystycznego na osi $j\omega$ płaszczyzny zmiennej zespolonej s, a drugi służy do przesuwania tych pierwiastków wzdłuż osi $j\omega$ (przestrajanie częstotliwości drgań). Określono wymagania stawiane poszczególnym elementom generatora, podano warunki generacji i przestrajania drgań. Przytoczono dwa przykłady realizacji generatorów tego typu.

1. WSTĘP

Jednym z podstawowych elementów elektronicznych, który znajduje coraz szersze zastosowanie, jest generator przestrajany napięciem, oznaczany w skrócie z angielskiego przez *VCO* (*Voltage Controlled Oscillator*). Generatory takie znajdują zastosowanie między innymi w pętlach sprzężenia fazoczułego (w tłumaczeniu z angielskiego *Phase-Locked Loop*), które wykorzystuje się do przetwarzania częstotliwości, filtracji, modulacji i demodulacji sygnałów, przetwarzania sygnałów analogowych na cyfrowe itd. Generatory *VCO* są również używane w aparaturze kontrolno-pomiarowej. Klasa generatorów *VCO* jest obszerna. Oprócz szeroko stosowanych rozwiązań typu multiwibrator, generatorów *LC*, obserwuje się również coraz większe zainteresowanie generatorami *RC* drgań sinusoidalnych [1], [2], [3], [4], [8]. Generatory *RC* drgań sinusoidalnych pozwalają w łatwy sposób uzyskać małą zawartość harmonicznych generowanego przebiegu.

W niniejszej pracy przedstawiona jest grupa generatorów *RC* drgań sinusoidalnych, które oprócz elementów pasywnych zawierają dwa wzmacniacze typu *ZNSN* (źródło napięciowe sterowane napięciem). Jeden wzmacniacz, oznaczony przez A_1, służy do podtrzymywania drgań w stanie ustalonym, a drugi, oznaczony przez A_2, służy do przestrajania częstotliwości. Odbywa się to przez zmiany wzmocnienia A_2 sterowane za pomocą napięcia. W układach takich, aby przestroić częstotliwość drgań wystarczy zmienić wartość jednego elementu. Najczęściej wykorzystuje się do tego celu jeden rezystor podłączony jednym zaciskiem do masy. W takim przypadku do przestrajania można zastosować tranzystor polowy, którego źródło jest podłączone do masy, a rezystancja dren-źródło

jest sterowana napięciem bramka-źródło. W porównaniu z generatorami przestrajanymi przez jednoczesną zmianę wartości przynajmniej dwu elementów unika się trudności związanych z odpowiednim doborem tych elementów pod kątem korelacji ich charakterystyk napięciowych. W pracach jakie ukazały się dotychczas na temat takich generatorów podane są najczęściej konkretne rozwiązania, warunki generacji i przestrajania. Stosunkowo pobieżnie omawia się ich właściwości i parametry. Odczuwa się ponadto brak pozycji, która obejmowałaby szerszą grupę układów tego typu, w celu ich klasyfikacji i porównania. W pracach [5], [6] podano sposób syntezy omawianych generatorów. Rozważania prowadzone w pracy [6] ograniczają się do układów zawierających wzmacniacze dowolnego typu (źródła napięciowe sterowane napięciem, źródła prądowe sterowane prądem, źródła napięciowe sterowane prądem, źródła prądowe sterowane napięciem) oraz tylko cztery elementy pasywne (dwa kondensatory i dwa rezystory). Ograniczenie rozważań do układów zawierających tylko cztery elementy pasywne często uniemożliwia uzyskanie korzystnych parametrów generatora. W niniejszej pracy rozpatrywane są układy opisane równaniem charakterystycznym drugiego rzędu, w których liczba użytych elementów pasywnych *RC* może być większa niż cztery. Ograniczenie dotyczy natomiast rodzaju użytych wzmacniaczy. Rozpatrywane są układy zawierające dwa *ZNSN*, co ma swoje uzasadnienie, ponieważ *ZNSN* są najczęściej stosowanym typem wzmacniaczy.

W pracy omówiono warunki powstawania drgań oraz zasadę przestrajania częstotliwości. Z przytoczonych rozważań wynikają między innymi wymagania stawiane obu wzmacniaczom oraz pasywnym podukładom generatora. Wymagania dotyczące układów pasywnych podane są w postaci warunków, jakie muszą spełniać transmitancje napięciowe opisujące te układy. Stanowią one podstawę syntezy generatorów omawianego typu. Przedstawione podejście do problemu syntezy jest inne niż proponowane w publikacjach [5], [6].

W pracy dokonano podziału omawianej klasy generatorów na trzy grupy różniące się postacią transmitancji opisującej sprzężenie zwrotne wzmacniacza używanego do podtrzymywania drgań. Przedstawiony podział stanowi podstawę porównania właściwości i możliwości różnych rozwiązań generatorów. Podział ten może być pomocny nie tylko w przypadku omawianej klasy układów, ale również w przypadku układów, gdzie transmitancja opisująca sprzężenie zwrotne wzmacniacza A_1 ma dowolny wymiar (prądowo-prądowa, transimpedancja, transkonduktancja).

Przytoczone rozważania zilustrowano dwoma przykładami.

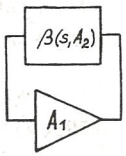
2. WARUNKI GENERACJI I PRZESTRAJANIA. WYMAGANIA STAWIANE POSZCZEGÓLNYM ELEMENTOM GENERATORA

Analizowane układy można przedstawić przy pomocy schematu blokowego pokazanego na rys. 1, gdzie:

A_1 — *ZNSN* służące do podtrzymywania drgań w stanie ustalonym,

$\beta(s, A_2)$ — transmitancja napięciowa, zależna od wzmocnienia wzmacniacza A_2 (*ZNSN*) używanego do przestrajania częstotliwości.

W przedstawionym modelu zakłada się, że pasma przenoszenia obu wzmacniaczy są dostatecznie szerokie w porównaniu z częstotliwością drgań, tzn., że 3 dB częstotliwości graniczne górne obu wzmacniaczy są pomijalnie duże. Jeżeli użyte wzmacniacze mają pomijalnie duże rezystancje wejściowe i pomijalnie małe rezystancje wyjściowe, wówczas w przedstawionym modelu transmitancja $\beta(s, A_2)$ jest określona przez zewnętrzne ele-



Rys. 1. Schemat blokowy generatora

menty pasywne RC oraz przez wartość wzmocnienia A_2. W przypadku przeciwnym transmitancja $\beta(s, A_2)$ jest dodatkowo funkcją rezystancji wejściowych i wyjściowych obu wzmacniaczy.

Załóżmy, że funkcja $\beta(s, A_2)$ jest transmitancją drugiego rzędu o ogólnej postaci

$$\beta(s, A_2) = \frac{s^2 c_2 + s c_1 + c_0}{s^2 n_2 + s n_1 + n_0}, \tag{1}$$

gdzie:

s — zmienna zespolona,

c_i, n_i — współczynniki rzeczywiste.

Dla transmitancji (1) równanie charakterystyczne układu z rys. 1 można przedstawić w postaci

$$s^2 \underbrace{(n_2 - A_1 c_2)}_{b_2} + s \underbrace{(n_1 - A_1 c_1)}_{b_1} + \underbrace{(n_0 - A_1 c_0)}_{b_0} = 0. \tag{2}$$

Warunek generacji oraz pulsacja drgań opisane są wzorami

$$b_1 = n_1 - A_1 c_1 = 0, \tag{3}$$

$$\omega_0^2 = \frac{n_0 - A_1 c_0}{n_2 - A_1 c_2} > 0. \tag{4}$$

Jeżeli spełnione będą warunki (3), (4), wówczas pierwiastki równania (2) będą leżały na osi $j\omega$ płaszczyzny zmiennej zespolonej s.

Przestrajanie częstotliwości drgań polega na takich zmianach współczynników b_0 lub b_2 równania (2), by nie naruszały one warunku $b_1 = 0$. Oznacza to, że współczynnik b_1 równania (2) nie może zależeć od wzmocnienia wzmacniacza A_2 używanego do przestrajania. Wynika z tego, że współczynniki n_1, c_1 transmitancji (1) nie mogą być zależne od wzmocnienia A_2. Wzmocnienie A_2 musi natomiast mieć wpływ na wartość przynajmniej jednego z pozostałych współczynników c_2, c_0, n_2, n_0 tej transmitancji. Ze względu na tłumienie składowych harmonicznych generowanego przebiegu pożądane jest, by współczynnik c_2 był równy zeru, tzn. by $\lim_{\omega \to \infty} |\beta(\omega, A_2)| = 0$. Jeżeli warunek ten nie jest spełniony, wówczas oprócz słabego tłumienia wyższych harmonicznych generowanego

przebiegu istnieje niebezpieczeństwo powstawania pasożytniczych drgań wielkiej częstotliwości. Z tego powodu z dalszych rozważań wyłączamy przypadek przestrajania częstotliwości przez zmiany współczynnika c_2. Biorąc pod uwagę przytoczone rozważania można wyróżnić trzy podstawowe postacie transmitancji $\beta(s, A_2)$, podane wzorami (5), prowadzące do sterowania częstotliwością drgań przez zmiany wzmocnienia A_2:

$$\beta(s, A_2) = \frac{sc_1 + c_0(A_2)}{s^2 n_2 + sn_1 + n_0}, \tag{5a}$$

$$\beta(s, A_2) = \frac{sc_1 + c_0}{s^2 n_2(A_2) + sn_1 + n_0}, \tag{5b}$$

$$\beta(s, A_2) = \frac{sc_1 + c_0}{s^2 n_2 + sn_1 + n_0(A_2)}. \tag{5c}$$

Dla podanych funkcji przestrajanie odbywa się przez zmiany wartości kolejno współczynników $c_0(A_2)$, $n_2(A_2)$, $n_0(A_2)$. Każda z podanych trzech funkcji ma odmienne charakterystyczne właściwości decydujące między innymi o stałości częstotliwości drgań, tłumieniu składowych harmonicznych itd.

W przypadku analizowanych w niniejszej pracy układów funkcje (5) mają wymiar transmitancji napięciowych. Przedstawiony podział może być również pomocny przy klasyfikacji i ocenie układów, w których transmitancje (5) mają wymiar transmitancji prądowych, transimpedancji lub transkonduktancji.

Stałość częstotliwości drgań generatorów omawianego typu jest jednym z podstawowych jego parametrów. Jest to jednak problem na tyle złożony, że wymaga odrębnego potraktowania i będzie przedmiotem oddzielnej publikacji.

Wracając do wzorów (5) zauważmy, że dla wszystkich trzech funkcji warunek generacji (3) prowadzi do zależności

$$A_1 = \frac{n_1}{c_1}. \tag{6}$$

Jest to przypadek graniczny powstawania drgań i w praktyce wzmocnienie wzmacniacza A_1 powinno być nieco większe od wartości występującej po prawej stronie równania (6).

Pulsacje drgań dla podanych trzech typów transmitancji opisane są odpowiednio wzorami

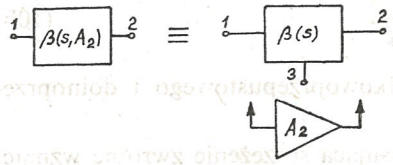
$$\omega_0 = \sqrt{\frac{n_0 - A_1 c_0(A_2)}{n_2}}, \tag{7a}$$

$$\omega_0 = \sqrt{\frac{n_0 - A_1 c_0}{n_2(A_2)}}, \tag{7b}$$

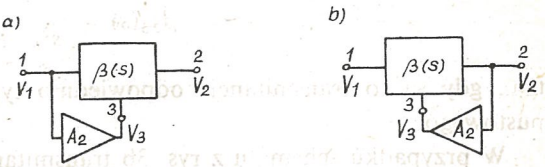
$$\omega_0 = \sqrt{\frac{n_0(A_2) - A_1 c_0}{n_2}}. \tag{7c}$$

Transmitancje (5) stanowią podstawę syntezy nowych rozwiązań generatorów. Rozpatrzmy przypadek, gdy układ sprzężenia zawrotnego wzmacniacza A_1 można przedstawić za pomocą schematu pokazanego na rys. 2, zawierającego trójnik pasywny RC

oraz wzmacniacz A_2 (*ZNSN*) włączony między dwa dowolne zaciski trójnika. Oznaczmy przez *1* zacisk dołączony do wyjścia, a przez *2* zacisk dołączony do wejścia wzmacniacza podtrzymującego drgania. Spośród wszystkich możliwych sposobów dołączenia wzmacniacza A_2 do układu trójnika tylko dwa sposoby pozwalają uzyskać transmitancję $\beta(s, A_2)$ o postaci podanej wzorami (5). Pokazano je na rys. 3.



Rys. 2. Ogólny schemat układu sprzężenia zwrotnego wzmacniacza A_1



Rys. 3. Schematy blokowe układu sprzężenia zwrotnego wzmacniacza A_1

Najpierw wykażemy, że oprócz podanych dwu sposobów żaden inny nie jest możliwy. Oba wzmacniacze A_1, A_2 są typu *ZNSN* i mają bardzo małe rezystancje wyjściowe, co oznacza, że wyjścia obu wzmacniaczy nie mogą być z sobą połączone (jeden zwierałby wyjście drugiego). Z tego powodu w schemacie 3a nie można zamienić miejscami zacisku wejściowego i wyjściowego wzmacniacza A_2. Zamiana miejscami zacisków wejściowego i wyjściowego wzmacniacza A_2 w schemacie z rys. 3b daje transmitancję $\beta(s, A_2)$, która jest iloczynem wzmocnienia A_2 i odpowiedniej transmitancji układu pasywnego od zacisku *1* do *3*, a więc współczynnik c_1 transmitancji $\beta(s, A_2)$ jest liniowo zależny od wzmocnienia A_2. Gdyby wzmcniacz A_2 był włączony między zaciski *1* i *2*, wówczas dla przypadku, gdy wejściem wzmacniacza A_2 jest zacisk *1*, transmitancja $\beta(s, A_2)$ byłaby równa A_2. Dla przypadku przeciwnego, gdy wejściem wzmacniacza A_2 jest zacisk *2*, oba wzmacniacze będą miały połączone wyjścia przy zacisku *1*.

Pokażemy teraz, że schematy przedstawione na rys. 3 pozwalają zrealizować każdy z podanych trzech typów transmitancji (5). Trójnik pasywny musi być opisany transmitancjami drugiego rzędu, by uzyskać równanie charakterystyczne drugiego rzędu. Wprowadźmy następujący opis trójnika

$$\beta_{21}(s) = \left.\frac{V_2}{V_1}\right|_{V_3=0} = \frac{s^2a_2+sa_1+a_0}{s^2e_2+se_1+e_0}, \tag{8a}$$

$$\beta_{23}(s) = \left.\frac{V_2}{V_3}\right|_{V_1=0} = \frac{s^2d_2+sd_1+d_0}{s^2e_2+se_1+e_0}, \tag{8b}$$

gdzie:

a_i, d_i, e_i — współczynniki rzeczywiste dodatnie,

s — zmienna zespolona.

Mianowniki obu funkcji są identyczne, co wynika z właściwości układów pasywnych *RC*.

W przypadku schematu z rys. 3a transmitancja opisująca sprzężenie zwrotne wzmacniacza A_1 opisana jest zależnością

$$\beta(s, A_2) = \frac{V_2}{V_1} = \beta_{21}(s)+A_2\beta_{23}(s) = \frac{s^2(a_2-A_2d_2)+s(a_1-A_2d_1)+a_0-A_2d_0}{s^2e_2+se_1+e_0}. \tag{9}$$

Porównując wzór (9) z funkcjami (5) widać, że możliwa jest realizacja transmitancji typu (5a). Uzyskuje się to, gdy trójnik pasywny opisany jest wzorami

$$\beta_{21}(s) = \frac{sa_1}{s^2e_2+se_1+e_0}, \tag{10a}$$

$$\beta_{23}(s) = \frac{d_0}{s^2e_2+se_1+e_0}, \tag{10b}$$

tzn., gdy są to transmitancje odpowiednio typu środkowoprzepustowego i dolnoprzepustowego.

W przypadku schematu z rys. 3b transmitancja opisująca sprzężenie zwrotne wzmacniacza A_1 określona jest zależnością

$$\beta(s, A_2) = \frac{V_2}{V_1} = \frac{\beta_{21}(s)}{1-A_2\beta_{23}(s)} = \frac{s^2a_2+sa_1+a_0}{s^2(e_2-A_2d_2)+s(e_1-A_2d_1)+e_0-A_2d_0}. \tag{11}$$

Porównując tę zależność z wzorami (5) widać, że możliwa jest realizacja transmitancji typu (5b) i (5c). Transmitancję typu (5c) uzyskuje się, gdy trójnik pasywny opisany jest wzorami (10), tzn. gdy są to transmitancje odpowiednio typu środkowoprzepustowego i dolnoprzepustowego. Dla podanych funkcji opisujących trójnik pasywny współczynnik c_0 we wzorze (5c) jest równy zeru.

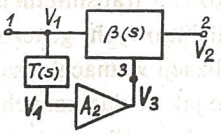
Transmitancję typu (5b) uzyskuje się, gdy trójnik pasywny opisany jest wzorem (10a) i wzorem (12)

$$\beta_{23}(s) = \frac{s^2d_2}{s^2e_2+se_1+e_0}, \tag{12}$$

tzn. gdy są to transmitancje odpowiednio typu środkowoprzepustowego i górnoprzepustowego. Dla podanych funkcji współczynnik c_0 we wzorze (5b) jest również równy zeru. Realizacja trójnika pasywnego opisanego wzorami (10a), (10b) lub (10a), (12) nie nastręcza większych trudności.

Omawiane układy są generatorami drgań sinusoidalnych. Aby uzyskać małą zawartość harmonicznych generowanego przebiegu, układ sprzężenia zwrotnego wzmacniacza A_1 powinien być liniowy, tzn. wzmacniacz A_2 występujący w tym układzie powinien pracować w zakresie liniowym. Narzuca to określone wymagania dotyczące amplitudy sygnału wzmacnianego przez wzmacniacz A_2, a pośrednio również wymagania dotyczące dopuszczalnej wartości wzmocnienia tego wzmacniacza. Widać więc, że małe zniekształcenia nieliniowe można uzyskać tylko w ograniczonym przedziale zmian wzmocnienia A_2, zależnym od amplitudy sygnału na wejściu tego wzmacniacza, a więc w ograniczonym przedziale przestrajania częstotliwości drgań. Zwróćmy uwagę, że pod tym względem układ sprzężenia zwrotnego z rys. 3b jest korzystniejszy od układu z rys. 3a. W przypadku schematu z rys. 3b na wejściu wzmacniacza A_2 występuje taki sam sygnał jak na wejściu wzmacniacza A_1, natomiast w przypadku schematu z rys. 3a jest to sygnał wyjściowy wzmacniacza A_1, a więc sygnał o większej amplitudzie. Z tego powodu celowe jest wprowadzenie modyfikacji schematu z rys. 3a, która pozwoliłaby zmniejszyć amplitudę sygnału na wejściu wzmacniacza A_2. Można to uzyskać łącząc kaskadowo ze wzmacniaczem A_2

dodatkowy blok pasywny opisany transmitancją $T(s)$, co pokazano na rys. 4. W takim przypadku, aby uzyskać transmitancję sprzężenia zwrotnego $\beta(s, A_2)$ drugiego rzędu, wystarczy by transmitancje opisujące pasywne podukłady generatora były pierwszego



Rys. 4. Zmodyfikowany schemat blokowy układu sprzężenia zwrotnego wzmacniacza A_1

rzędu o ogólnej postaci

$$\beta_{21}(s) = \left.\frac{V_2}{V_1}\right|_{V_3=0} = \frac{sp_1+p_0}{se_1+e_0}, \tag{13a}$$

$$\beta_{23}(s) = \left.\frac{V_2}{V_3}\right|_{V_1=0} = \frac{sh_1+h_0}{se_1+e_0}, \tag{13b}$$

$$T(s) = \frac{V_4}{V_1} = \frac{sg_1+g_0}{sl_1+l_0}, \tag{13c}$$

gdzie:

s — zmienna zespolona,

p_i, h_i, g_i, e_i, l_i — współczynniki rzeczywiste dodatnie.

Postępując w sposób analogiczny jak w przypadku schematu z rys. 3, można wykazać, że transmitancja sprzężenia zwrotnego będzie miała postać podaną wzorem (5a), gdy bloki pasywne na rys. 4 opisane będą funkcjami

$$\beta_{21}(s) = \frac{p_0}{se_1+e_0}, \tag{14a}$$

$$\beta_{23}(s) = \frac{h_0}{se_1+e_0}, \tag{14b}$$

$$T(s) = \frac{g_0}{sl_1+l_0}. \tag{14c}$$

Są to transmitancje pierwszego rzędu typu dolnoprzepustowego. Realizacja układów pasywnych opisanych wzorami (14) nie nastręcza większych trudności.

Zwróćmy uwagę, że przy odpowiednim doborze wartości elementów pasywnych moduł transmitancji $T(s)$ w zakresie generowanych częstotliwości może być znacznie mniejszy od jedności. W ten sposób można uzyskać bardzo małą wartość amplitudy sygnału zmiennego na wejściu wzmacniacza A_2. W rezultacie wprowadzonej modyfikacji możliwe jest znaczne rozszerzenie zakresu zmian wzmocnienia A_2 w obrębie liniowej pracy tego układu. Zauważmy, że układ sprzężenia zwrotnego pokazany na rys. 4 musi zawierać przynajmniej pięć elementów pasywnych (trzy elementy w bloku $\beta(s)$ i dwa elementy w bloku $T(s)$. W tym miejscu widać, że przedstawiony sposób syntezy, nie ograniczający liczby elementów pasywnych do czterech, pozwala znaleźć rozwiązania generatorów o korzystniejszych właściwościach.

Podobnie jak schemat 3a, również schemat 3b można zmodyfikować wprowadzając między zacisk 2 trójnika i wejście wzmacniacza A_2 dodatkowy blok pasywny $T(s)$. Jeżeli transmitancje bloków pasywnych w tak zmodyfikowanym schemacie opisane będą wzorami o postaci (14), wówczas całkowita transmitancja $\beta(s, A_2)$ będzie typu (5c).

Mówiąc o zniekształceniach nieliniowych generowanego sygnału należy dodać, że zależą one również od sposobu realizacji wzmacniacza A_1 używanego do podtrzymywania drgań w stanie ustalonym. Podobnie jak w klasycznych generatorach *RC* można je znacznie zmniejszyć obejmując wzmacniacz A_1 pętlą automatycznej regulacji wzmocnienia.

3. OGRANICZENIA ZAKRESU PRZESTRAJANIA GENERATORA

Przeprowadzona w poprzednim punkcie analiza zakłada nieskończenie szerokie pasma przenoszenia obu wzmacniaczy. Z analizy tej wynika, że omawiane układy mogą pracować jako generatory przestrajane w szerokim zakresie, gdy współczynniki c_1, n_1 transmitancji (1) są niezależne od wzmocnienia A_2, co uzyskać można przy odpowiednich strukturach części pasywnej generatora. Jeżeli w takich układach rezystancje wejściowe wzmacniaczy są pomijalnie duże, a wyjściowe pomijalnie małe, wówczas współczynniki c_1, n_1 są określone przez pasywne elementy generatora i nie zależą od parametrów wzmacniaczy. Jeżeli natomiast rezystancje wejściowe i wyjściowe wzmacniaczy są porównywalne z parametrami pasywnych podukładów generatora, wówczas współczynniki c_1, n_1 zależą w sposób istotny od tych rezystancji (w generatorach dużych częstotliwości dotyczy to również pojemności wejściowych i wyjściowych wzmacniaczy). W takim przypadku rezystancje (pojemności) wejściowe i wyjściowe wzmacniaczy powinny być nieczułe na zmiany wzmocnienia tych wzmacniaczy. Dotyczy to głównie wzmacniacza A_2 używanego do sterowania częstotliwością drgań. Jeżeli wymaganie to nie jest spełnione, wówczas zmiany wzmocnienia A_2 powodują zmiany jego rezystancji (pojemności) wejściowych lub wyjściowych, co prowadzi do zmian wymaganej wartości wzmocnienia A_1 (wzór 6). Konsekwencją tego są zmiany amplitudy drgań wywołane przestrajaniem częstotliwości, a przy szerokim zakresie przestrajania może wystąpić zrywanie drgań lub generacja przebiegu o dużych zniekształceniach nieliniowych. Zależy to od sposobu rozwiązania wzmacniacza A_1.

W tym miejscu warto dodać, że nawet gdy rezystancje (pojemności) wejściowe i wyjściowe wzmacniaczy są nieczułe na zmiany wzmocnienia, to są one często stosunkowo wrażliwe na zmiany takich czynników, jak temperatura, napięcia zasilania itd. Z tego powodu wskazane jest, by impedancje wejściowe wzmacniaczy były pomijalnie duże, a wyjściowe pomijalnie małe w porównaniu z parametrami pasywnych podukładów generatora.

Pokażemy teraz, że istotnym czynnikiem ograniczającym zakres przestrajania generatora jest szerokość pasma przenoszenia użytych wzmacniaczy. Rozpatrzymy to na przykładzie generatora opisanego transmitancją $\beta(s, A_2)$ o postaci (5a), gdzie trójnik pasywny opisany jest funkcjami (10). Wstawiając wzory (10) do równania (9) otrzymuje się

$$\beta(s, A_2) = \frac{sa_1 - A_2 d_0}{s^2 e_2 + s e_1 + e_0}. \tag{15}$$

Dla podanej transmitancji równanie charakterystyczne układu można przedstawić w postaci

$$\underbrace{s^2 e_2}_{b_2} + s\underbrace{(e_1 - A_1 a_1)}_{b_1} + \underbrace{e_0 - A_1 A_2 d_0}_{b_0} = 0, \tag{16}$$

z której widać, że współczynnik b_1 nie zależy od A_2.

Rozpatrzmy teraz przypadek, gdy oba wzmacniacze mają ograniczone pasmo i można je opisać w przybliżeniu wzorami

$$A_1 = \frac{A_{10}}{1 + \dfrac{s}{\omega_{g1}}} \cong A_{10} - s\frac{A_{10}}{\omega_{g1}}, \tag{17a}$$

$$A_2 = \frac{A_{20}}{1 + \dfrac{s}{\omega_{g2}}} \cong A_{20} - s\frac{A_{20}}{\omega_{g2}}, \tag{17b}$$

gdzie: ω_{g1}, ω_{g2} — 3 dB pulsacje graniczne górne wzmacniaczy.

Wstawiając zależności (17) do wzoru (16) i wykonując proste przekształcenia otrzymuje się równanie charakterystyczne

$$s^2 b_2 + s b_1 + d_0 = 0, \tag{18}$$

gdzie:

$$b_2 = e_2 - a_1 \frac{A_{10}}{\omega_{g1}} - d_0 \frac{A_{10} A_{20}}{\omega_{g1}\omega_{g2}}, \tag{19a}$$

$$b_1 = e_1 - a_1 A_{10} + d_0 \left(A_{20} \frac{A_{10}}{\omega_{g1}} + A_{10} \frac{A_{20}}{\omega_{g2}} \right), \tag{19b}$$

$$b_0 = e_0 - d_0 A_{10} A_{20}. \tag{19c}$$

Ze wzoru (19b) widać, że współczynnik b_1 równania charakterystycznego jest funkcją wzmocnienia A_{20}. Zależność ta jest tym silniejsza, im mniejsze są wartości pulsacji granicznych ω_{g1}, ω_{g2}. Wynikająca z warunku $b_1 = 0$ wartość wzmocnienia A_{10}, wymagana dla podtrzymywania drgań w stanie ustalonym, będzie się zmieniała ze zmianami wzmocnienia A_{20}. Prowadzi to do zmian amplitudy drgań wywołanych przestrajaniem częstotliwości, a przy szerokim zakresie przestrajania może nastąpić zerwanie drgań lub generacja sygnału o dużych zniekształceniach nieliniowych. Jak z tego wynika, szerokość pasma przenoszenia wzmacniaczy jest istotnym czynnikiem ograniczającym zakres przestrajania generatora.

Wzmacniacz A_1 pracuje zwykle przy małych wzmocnieniach (rzędu kilku), natomiast moduł wzmocnienia A_2 przy szerokim zakresie przestrajania zmienia się w bardzo szerokim przedziale. Widać to ze wzorów (7), w których współczynniki sterujące częstotliwością drgań są zwykle liniowo zależne od A_2 (równanie 16). Moduł wzmocnienia A_2 nie może przyjmować zbyt małych wartości, ponieważ dla małych wartości wzmocnienia A_2 uzyskuje się złą stałość częstotliwości drgań oraz słabe tłumienie składowych harmonicznych. Szersze naświetlenie tego problemu będzie przedstawione w publikacji poświęconej stałości częstotliwości drgań. Minimalne wartości modułu wzmocnienia A_2 powinny być rzędu 10^0. Żeby uzyskać szeroki zakres przestrajania (1 dekada lub więcej), maksymalna wartość

modułu wzmocnienia musi być rzędu $10^2 \div 10^3$. Tak więc pole wzmocnienia wzmacniacza A_2, przy szerokim zakresie przestrajania, powinno być o około dwa rzędy wielkości większe niż pole wzmocnienia wzmacniacza A_1.

Z przytoczonych rozważań wynika, że głównym czynnikiem ograniczającym zakres przestrajania generatora jest ograniczony przedział wartości, jakie może przyjmować moduł wzmocnienia A_2. Jak już wspomniano, ograniczenie to może wynikać ze zbyt małej wartości pola wzmocnienia wzmacniacza A_2. Inną przyczyną ograniczającą maksymalną wartość modułu wzmocnienia A_2 może być wymaganie dotyczące liniowej pracy tego wzmacniacza, na co zwrócono uwagę w poprzednim rozdziale. Z licznych badań eksperymentalnych autora oraz innych publikacji na ten temat wynika, że dobre parametry generowanego sygnału (dobra stałość częstotliwości i amplitudy, małe zniekształcenia harmoniczne) można uzyskać, gdy stosunek maksymalnej do minimalnej częstotliwości drgań, określających przedział przestrajania, wynosi co najwyżej kilkadziesiąt.

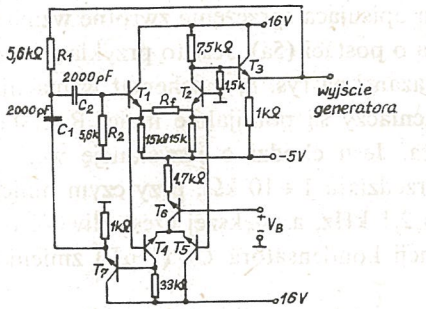
Rząd wielkości generowanych częstotliwości zależy od wartości elementów pasywnych RC. Dla częstotliwości małych ograniczeniem są dopuszczalne maksymalne wartości rezystancji i pojemności elementów pasywnych generatora, które z różnych powodów nie mogą być dowolnie duże. Ograniczenie to można złagodzić w przypadku generatorów opisanych transmitancją $\beta(s, A_2)$ typu (5b), dla których pulsację drgań opisuje wzór (7b). Ze wzoru (7b) widać, że dla bardzo dużych wartości współczynnika $n_2(A_2)$ uzyskuje się małe częstotliwości drgań. Przy ustalonych wartościach elementów pasywnych RC wartość współczynnika $n_2(A_2)$ można znacznie zwiększyć (o kilka rzędów wielkości) przez odpowiedni dobór wartości wzmocnienia A_2. W ten sposób można zmniejszyć rząd wielkości generowanych częstotliwości. Pozwala to uzyskać drgania sinusoidalne o częstotliwości rzędu 10^0 Hz. Układy opisane transmitancją $\beta(s, A_2)$ typu (5a) lub (5c) pozwalają natomiast zwiększyć rząd generowanych częstotliwości. Widać to ze wzorów (7a) i (7c), gdzie dla dużych wartości bezwzględnych współczynników $c_0(A_2)$, $n_0(A_2)$ uzyskuje się duże wartości częstotliwości drgań. W przypadku wzoru (7a) współczynnik $c_0(A_2)$ powinien być ujemny (dla $A_1 > 0$), a w przypadku wzoru (7c) współczynnik $n_0(A_2)$ dodatni. W takich układach można uzyskać drgania o częstotliwościach rzędu 10^6 Hz.

4. PRZYKŁADY GENERATORÓW STEROWANYCH JEDNYM ELEMENTEM

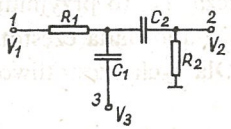
Na rys. 5 pokazano układ generatora, w którym transmitancja napięciowa opisująca sprzężenie zwrotne wzmacniacza podtrzymującego drgania opisana jest funkcją o postaci (5b). Jest to przykład zaczerpnięty z pracy [6]. Autor tej pracy nie pokazuje w sposób jawny jaka jest postać transmitancji $\beta(s, A_2)$. Podany układ odpowiada schematowi blokowemu z rys. 3b. Wzmacniacz A_1 zrealizowano na tranzystorach T_1, T_2, T_3, gdzie wejściem wzmacniacza jest baza tranzystora T_1, a wyjściem emiter tranzystora T_3. Jest to $ZNSN$ nie odwracające fazy wzmacnianego sygnału. Wartość rezystora R_f dobierana jest eksperymentalnie tak, by uzyskać wymagane wzmocnienie wzmacniacza A_1. Zależy ona głównie od parametrów h_{21e} tranzystorów T_1, T_2 i przyjmuje wartości rzędu kilku kΩ. Wzmacniacz A_2 zrealizowano na tranzystorach T_4, T_5, T_6, T_7. Amplituda sygnału zmiennego na bazie tranzystora T_5 jest o około dwa rzędy wielkości mniejsza niż na

bazie tranzystora T_4. Można więc w przybliżeniu przyjąć, że wejściem wzmacniacza A_2 jest baza tranzystora T_4. Wzmocnienie tranzystora T_1, definiowane jako stosunek napięcia na jego emiterze do napięcia na jego bazie, jest w przybliżeniu równe jeden, a więc na wejściach obu wzmacniaczy występuje w przybliżeniu taki sam sygnał zmienny. Wyjściem wzmacniacza A_2 jest emiter tranzystora T_7. Wzmanciacz ten jest *ZNSN* odwracającym fazę wzmacnianego sygnału. Wzmocnienie wzmacniacza A_2 jest regulowane przez zmiany prądu kolektora tranzystora T_6, sterowane napięciem V_B, co powoduje zmiany wzmocnienia pary różnicowej T_4, T_5. Rezystancje wejściowe obu wzmacniaczy są na tyle duże, a wyjściowe na tyle małe, że można pominąć ich wpływ na postać transmitancji $\beta(s, A_2)$.

Blok pasywny generatora składa się z elementów R_1, C_1, R_2, C_2 i pokazano go oddzielnie na rys. 6, oznaczając zaciski zewnętrzne zgodnie z oznaczeniami trójnika na rys. 3b.



Rys. 5. Schemat generatora. Tranzystory są typu 2N3569



Rys. 6. Schemat części pasywnej układu sprzężenia zwrotnego wzmacniacza A_1

Blok ten opisany jest transmitancjami:

$$\beta_{21}(s) = \frac{sC_2R_2}{s^2C_1R_1C_2R_2+s(C_2R_2+C_2R_1+C_1R_1)+1}, \tag{20a}$$

$$\beta_{23}(s) = \frac{s^2C_1R_1C_2R_2}{s^2C_1R_1C_2R_2+s(C_2R_2+C_2R_1+C_1R_1)+1}. \tag{20b}$$

Transmitancja $\beta(s, A_2)$, opisująca sprzężenie zwrotne wzmacniacza A_1, podana jest wzorem

$$\beta(s, A_2) = \frac{\beta_{21}(s)}{1-A_2\beta_{23}(s)} = \frac{sC_2R_2}{s^2C_1R_1C_2R_2(1-A_2)+s(C_2R_2+C_2R_1+C_1R_1)+1}. \tag{21}$$

Wymagane wzmocnienie wzmacniacza A_1 opisuje wzór (6), co w tym przypadku prowadzi do zależności [1]

$$A_1 = 1+\frac{R_1}{R_2}\left(1+\frac{C_1}{C_2}\right). \tag{22}$$

[1] W pracy [6] do odpowiednika wzoru (22) wkradła się pomyłka.

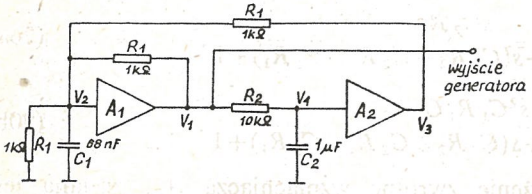
Pulsacja drgań określona jest zależnością

$$\omega_0 = \frac{1}{\sqrt{R_1 C_1 R_2 C_2}} \frac{1}{\sqrt{1-A_2}}, \tag{23}$$

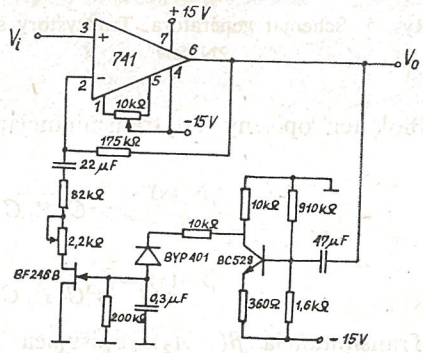
co odpowiada zależności (7b), gdy $c_0 = 0$.

Podany w pracy [6] zakres przestrajania częstotliwości wynosi 2,38 ÷ 17,97 kHz przy zmianach napięcia sterującego V_B w przedziale $-0{,}55 \div -4{,}7$ V. Dla tego przedziału podano wartość tłumienia drugiej i trzeciej harmonicznej generowanego sygnału, gdzie sygnał wyjściowy generatora zbierany jest z emitera tranzystora T_3. Z podanych wartości wynika, że współczynnik zawartości harmonicznych jest większy niż 1%. Stosunkowo duża wartość współczynnika zawartości harmonicznych wynika z faktu, że wzmacniacz A_1 nie jest objęty pętlą ARW, a ponadto sygnał na wejściu wzmacniacza A_2 ma stosunkowo dużą amplitudę.

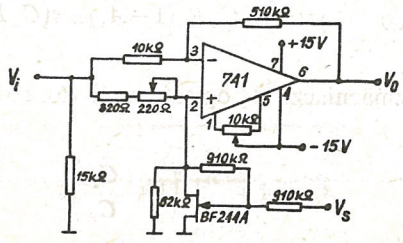
Na rys. 7 pokazano przykład generatora zrealizowanego na wzmacniaczach operacyjnych μA 741, w którym transmitancja napięciowa opisująca sprzężenie zwrotne wzmacniacza podtrzymującego drgania opisana jest funkcją o postaci (5a). Jest to przykład zaczerpnięty z pracy [7]. Schemat wzmacniacza A_1 pokazano na rys. 8, a schemat wzmacniacza A_2 na rys. 9. Rezystancje wyjściowe obu wzmacniaczy są pomijalnie małe. Rezystancja wejściowa wzmacniacza A_1 jest pomijalnie duża. Jeśli chodzi o rezystancję wejściową wzmacniacza A_2, to przyjmuje ona wartości z przedziału $1 \div 10$ kΩ, przy czym mniejszej rezystancji odpowiada częstotliwość drgań około 2,1 kHz, a większej częstotliwość około 210 Hz. Dla tych częstotliwości moduł impedancji kondensatora C_2 (1 μF) zmienia się



Rys. 7. Schemat generatora



Rys. 8. Schemat wzmacniacza A_1



Rys. 9. Schemat wzmacniacza A_2

w przedziale 75÷750 Ω, a więc jest około 13 razy mniejszy od rezystancji wejściowej wzmacniacza A_2. W dalszych rozważaniach potraktujemy rezystancję wejściową wzmacniacza A_2 jako pomijalnie dużą wobec impedancji kondensatora C_2. Dokładniejszą analizę, uwzględniającą wpływ tej rezystancji, można znaleźć w pracy [7].

Układ sprzężenia zwrotnego wzmacniacza A_1 odpowiada schematowi z rys. 4. Trójnik pasywny opisany jest funkcjami

$$\beta_{21}(s) = \beta_{23}(s) = \frac{1}{sR_1C_1+3}, \tag{24}$$

a transmitancja $T(s)$ ma postać

$$T(s) \cong \frac{1}{sR_2C_2+1}. \tag{25}$$

Transmitancja opisująca sprzężenie zwrotne wzmacniacza A_1 opisana jest wzorem

$$\beta(s, A_2) = \beta_{21}(s)+T(s)A_2\beta_{23}(s) = \frac{sR_2C_2+1+A_2}{s^2R_1C_1R_2C_2+s(R_1C_1+3R_2C_2)+3}. \tag{26}$$

Wymagane wzmocnienie wzmacniacza A_1, odpowiadające zależności (6), opisane jest wzorem

$$A_1 = 3+\frac{R_1C_1}{R_2C_2}, \tag{27}$$

a pulsacja drgań opisana jest wzorem

$$\omega_0 = \frac{1}{\sqrt{R_1C_1R_2C_2}}\sqrt{3-A_1(1+A_2)}, \tag{28}$$

co odpowiada zależności (7a).

Elementy generatora są tak dobrane, że spełniona jest nierówność

$$R_1C_1 \ll R_2C_2. \tag{29}$$

W rezultacie otrzymuje się

$$A_1 \cong 3, \tag{30}$$

$$\omega_0 \cong \frac{1}{\sqrt{R_1C_1R_2C_2}}\sqrt{-A_1A_2}. \tag{31}$$

Ze wzorów tych widać, że wzmacniacz A_2 musi być układem odwracającym fazę wzmacnianego sygnału ($A_2 < 0$).

Spełnienie nierówności (29) zapewnia silne tłumienie generowanego sygnału przez blok pasywny opisany transmitancją $T(s)$, tzn. amplituda sygnału na wejściu wzmacniacza A_2 jest znacznie mniejsza niż na wyjściu wzmacniacza A_1. Wzmacniacz A_1, pokazany na rys. 8, objęty jest pętlą ARW. W rezultacie uzyskano w szerokim zakresie przestrajania (1 dekada) współczynnik zawartości harmonicznych rzędu $10^{-1}\%$.

Wzmocnienie wzmacniacza A_2 jest regulowane przez zmiany rezystancji dren-źródło tranzystora polowego BF 244A, sterowane za pomocą napięcia V_s. Zakres generowanych częstotliwości w podanym przykładzie wynosi 210 Hz÷2,15 kHz przy zmianach napięcia V_s w przedziale 0÷−2,5 V. Wyjściem generatora jest wyjście wzmacniacza A_1. Amplituda

drgań wynosi około 770 mV, a jej zmiany w całym zakresie przestrajania nie przekraczają 2%.

Czynnikiem ograniczającym zakres przestrajania częstotliwości jest tu pole wzmocnienia wzmacniacza A_2 (około 1 MHz). W tym miejscu warto dodać, że wzmacniacz A_2 został celowo tak zaprojektowany, by jego rezystancja wejściowa zmieniała się przy przestrajaniu częstotliwości drgań. Zmiany te wpływają bowiem na wartość współczynnika b_1 równania charakterystycznego w ten sposób, że kompensują wpływ ograniczonego pasma przenoszenia wzmacniacza A_2 na wymaganą wartość wzmocnienia wzmacniacza A_1. Z analizy przedstawionej w pracy [7] wynika, że podany przedział zmian wartości rezystancji wejściowej wzmacniacza A_2 ($1 \div 10$ kΩ) czyni tę kompensację najbardziej skuteczną.

5. PODSUMOWANIE

W pracy omówiono warunki, jakie muszą spełniać elementy generatora, by możliwa była generacja drgań sinusoidalnych przestrajanych przez zmiany wzmocnienia wzmacniacza A_2. Warunki dotyczące elementów pasywnych podano w postaci wymagań stawianych transmitancjom opisującym pasywne podukłady generatora. Sformułowane warunki stanowią podstawę syntezy generatorów omawianego typu.

Dokonano podział omawianej klasy generatorów na trzy grupy. Kryterium podziału jest postać transmitancji opisującej sprzężenie zwrotne wzmacniacza podtrzymującego drgania w stanie ustalonym. Każda grupa posiada odmienne charakterystyczne właściwości. Dokonany podział pozwala na przeprowadzenie klasyfikacji różnych rozwiązań generatorów, porównanie ich właściwości i możliwości.

W pracy nie omówiono problemu stałości częstotliwości drgań generatorów tego typu. Jest to zagadnienie na tyle obszerne, że wymaga odrębnego potraktowania i będzie przedmiotem oddzielnej publikacji.

Przeprowadzone w pracy rozważania bazują na liniowym opisie układu generatora. Zwrócono jednak uwagę na najważniejsze czynniki decydujące o zniekształceniach nieliniowych generowanego przebiegu.

W pracy omówiono ograniczenia zakresu przestrajania generatora. Stosunek maksymalnej do minimalnej częstotliwości drgań, określających przedział przestrajania, może wynosić co najwyżej kilkadziesiąt. Rząd wielkości generowanych częstotliwości zależy nie tylko od wartości elementów pasywnych RC, ale również od typu transmitancji opisującej sprzężenie zwrotne wzmacniacza podtrzymującego drgania w stanie ustalonym. Transmitancja $\beta(s, A_2)$ typu (5b) pozwala zmniejszyć rząd wielkości generowanych częstotliwości, co jest korzystne w przypadku częstotliwości małych. W układach opisanych tego typu transmitancją można generować drgania o częstotliwościach rzędu 10^0 Hz. Transmitancja $\beta(s, A_2)$ typu (5a) lub (5c) pozwala zwiększyć rząd wielkości generowanych częstotliwości, co jest korzystne w przypadku wielkich częstotliwości. W układach opisanych tego typu transmitancjami można generować drgania o częstotliwościach rzędu 10^6 Hz. Podane w pracy przykłady generatorów dotyczą częstotliwości średnich (kHz).

Generatory omawianego typu można przestrajać przez zmiany wartości jednego elementu, co pozwala na łatwe sterowanie częstotliwością drgań za pomocą napięcia.

BIBLIOGRAFIA

1. M. Hribsec, R. W. Newcomb, *VCO Controlled by One Variable Resistor*, IEEE Transactions on Circuits and Systems, 1976, CAS-23, s. 166—169.
2. R. Nandi, *Another Very Low Frequency Sinusoid Oscillator*, Int. J. Electron., 1977, 43, s. 197—200.
3. R. Senani, *Novel Sinusoidal Oscillator Employing Grounded Capacitors*, Electronics Letters, January 1980, Vol. 16, Nr 2, s. 62—63.
4. M. Sundaramurthy i in., *A Simple Voltage Controlled Oscillator with Grounded Capacitors*, Proceedings of the IEEE, Vol. 65, November 1977, s. 1612—1614.
5. V. S. Ganguly, *RC-Oscillator Synthesis: Sub-Audio Frequency Generation with a Single Control*, Proceedings of the IEEE, Vol. 66, No. 4, April 1978, s. 516—518.
6. Y. Sun, *Generation of Sinusoidal Voltage (Current) — Controlled Oscillators for Integrated Circuits*, IEEE Transactions on Circuit Theory, Vol. 19, No. 2, March 1972, 137—141.
7. R. Wojtyna, *Analiza generatorów RC drgań sinusoidalnych z dwupętlowym selektywnym sprzężeniem zwrotnym*, Praca doktorska, Poznań 1982.
8. R. Wojtyna, L. Weiss, *Przestrajany napięciem generator RC drgań sinusoidalnych wysokiej częstotliwości*, VI Krajowa Konferencja TO i UE, Kozubnik 1983.

R. WOJTYNA

FREQUENCY CONTROL IN *RC* SINUSOIDAL OSCILLATORS BY AN ADDITIONAL AMPLIFIER

Summary

Sinusoidal *RC* oscillators containing two amplifiers (voltage controlled voltage sources) are considered. One amplifier is used to confine the roots of the characteristic equation on the $j\omega$ axis of the complex plane, and the other is used to move these roots along the $j\omega$ axis (control of oscillation frequency). Requirements to be fulfilled by the oscillator elements are specified. Conditions for oscillation and frequency control are given. Two examples of oscillators are presented.

R. WOJTYNA

COMMANDE DE LA FRÉQUENCE DES OSCILLATIONS SINUSOÏDALES DE L'OSCILLATEUR *RC* À L'AIDE DE L'AMPLIFICATUER ADDITIONNEL

Résumé

On a décrit, dans cet article, l'oscillateur *RC* des oscillation sinusoïdales, comprenant deux amplificateurs (sources de tension commandées par tension). L'un d'eux sert à placer les racines de l'équation caractéristique sur l'axe $j\omega$ du plan compleze (s — variable) — pendant que l'autre sert à les déplacer le long de l'axe $j\omega$ (accord de la fréquence de l'oscillation). On a précisé les exigences auxquelles doivent répondre les éléments de l'oscillateur, on a indiqué les conditions de l'oscillation et de l'accord de la fréquence des oscillations. On a cité deux exemples de la réalisation de l'oscillateur de ce type.

R. WOJTYNA

FREQUENZSTEUERUNG DER SINUSFÖRMIGEN SCHWINGUNGEN IN *RC*-GENERATOREN MIT HILFE EINES ZUSÄTZLICHEN VERSTÄRKERS

Zusammenfassung

Im Beitrag wurden zwei Verstärker (spannungsgesteuerte Spannungsquellen) enthaltende *RC*-Generatoren für sinusförmige Schwingungen erwogen. Einer von ihnen dient zur Lokalisierung der Wurzeln der charakteristischen Gleichungen auf der $j\omega$-Achse der Ebene für die komplexe Variable s, der andere dient zur Verschiebung dieser Wurzeln längs der $j\omega$-Achse (Verstimmung der Schwingungsfrequenz).

Anforderungen den einzelnen Elementen des Generators gegenüber wurden bestimmt. Es wurden auch Bedingungen für Generierung und Verstimmung der Schwingungen angegeben. Zwei Realisierungsbeispiele für Generatoren dieser Art wurden gebracht.

Р. ВОЙТЫНА

ПЕРЕСТРАИВАНИЕ ЧАСТОТЫ СИНУСОИДАЛЬНЫХ КОЛЕБАНИЙ В ГЕНЕРАТОРАХ ПРИ ПОМОЩИ ДОБАВОЧНОГО УСИЛИТЕЛЯ *RC*

Р е з ю м е

Представлены генераторы *RC* синусоидальных колебаний содержащие два усилителя (источники напряжения управляемые напряжением). Один используется для смещения корней характеристического уравнения на ось $j\omega$ комплексной пласкости s, а второй используется для перемещения этих корней вдоль оси $j\omega$ (перестраивание частоты колебаний). Определены требования, которые должны выполнять элементы генератора. Представлены условия для генерации и перестраивания колебаний. Приведены два примера реализации генераторов этого типа.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31 z. 2, ss 573—591

621.375.026

Wykorzystanie dowolnych, dwustanowych przetworników napięcia stałego do budowy wzmacniaczy mocy klasy *D* i metoda ich analizy

STANISŁAW KUTA (KRAKÓW)

Instytut Elektroniki AGH

Otrzymano 15.2.1984

W pracy przedstawiono zasadę działania i teorię wzmacniacza klasy *D* małej częstotliwości, zbudowanego w oparciu o przetwornik napięcia stałego Middlebrooka i Čuka. Wzmacniacz ten posiada szereg zalet w porównaniu z konwencjonalnym układem klasy *D*. Podano przykłady stopni końcowych wzmacniaczy klasy *D* zrealizowanych w oprarciu o inne dwustanowe przetworniki napięcia stałego. Opisano tzw. metodę uśrednionych równań stanu analizy dwustanowych przetworników napięcia stałego, pozwalającą na wyznaczenie kanonicznego modelu zastępczego każdego przetwornika. Model ten może być wykorzystany do analizy charakterystyk częstotliwościowych każdego wzmacniacza klasy *D* małej częstotliwości.

1. WSTĘP

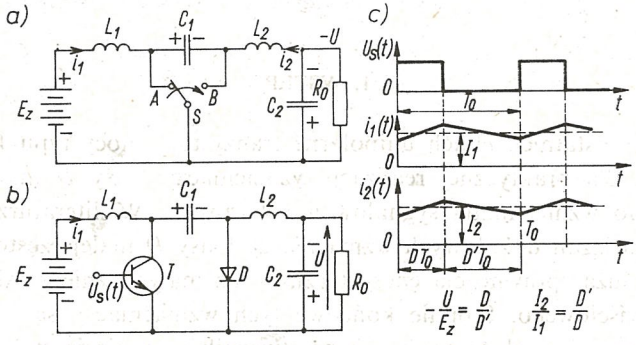
Opracowane w ostatnich latach unipolarne tranzystory mocy typu *VMOS* otworzyły nowe możliwości dla praktycznej realizacji wzmacniaczy klasy *D* małej częstotliwości, przeznaczonych do wzmacniania sygnałów akustycznych. W literaturze [7, 8, 9, 13, 14] opisano wiele rozwiązań układowych wzmacniaczy klasy *D* małej częstotliwości, charakteryzujących się dużą sprawnością energetyczną oraz małymi zniekształceniami nieliniowymi sygnału wyjściowego. Stopnie końcowe tych wzmacniaczy są układami przeciwsobnymi, utworzonymi ze złożenia dwóch przetworników napięcia stałego z szeregowym tranzystorem i szeregową indukcyjnością, sterowanych w przeciwfazie. Stopnie te obciążają zasilacze impulsowo, co może być źródłem dużych zakłóceń radiowych. Stosowanie w każdej szynie napięcia zasilającego filtru wejściowego [10] oraz ekranowanie układu pozwala tylko na częściowe wyeliminowanie wpływu tych zakłóceń. Do wad tych stopni końcowych należy również konieczność stosowania dwóch oddzielnych źródeł napięć zasilających (nie dotyczy jedynie układów przeciwsobnych). Ponadto układy sterowania tranzystorów stopnia końcowego są dość złożone, nawet w przypadku stosowania tranzystorów komplementarnych, ponieważ emitery tych tranzystorów (lub źródła tranzystorów *VMOS*) nie są uziemione. O efektywności wzmacniaczy klasy *D* w dużym stopniu decyduje właściwy dobór elementów ze względu na szybkość ich działania, jak również dobór odpowiednich

układów sterowania tranzystorów, w celu zminimalizowania strat mocy związanych z procesami przełączania, bowiem ten składnik strat może być dominujący w ogólnym bilansie. Nieuziemienie wszystkich emiterów (lub źródeł) tranzystorów stopnia końcowego stanowi dodatkowe utrudnienie przy realizacji układów sterujących, zapewniających krótkie czasy przełączania tych tranzystorów. Wzmacniacz klasy D zrealizowany w oparciu o impulsowy przetwornik napięcia stałego Middlebrooka i Čuka jest rozwiązaniem układowym, w którym wymienione we wstępie wady zostały wyeliminowane [6, 9].

Na przykładzie tworzenia wzmacniacza klasy D małej częstotliwości z przetwornikiem Middlebrooka i Čuka zostanie pokazane, że każdy impulsowy, dwustanowy przetwornik napięcia stałego może stanowić podstawę dla realizacji stopni końcowych tych wzmacniaczy. Na podstawie literatury [5, 11, 12] przedstawiono metodę uśrednionych równań stanu, prowadzącą do kanonicznego modelu zastępczego dowolnego, dwustanowego przetwornika napięcia stałego. Stosując tę metodę przedstawiono teorię wzmacniacza klasy D, wykorzystującego przetwornik Middlebrooka i Čuka. Metoda ta może być wykorzystana również do analizy i optymalizacji charakterystyk stopni końcowych wzmacniaczy klasy D małej częstotliwości, realizowanych w oparciu o inne dwustanowe przetworniki napięcia stałego.

2. IMPULSOWY PRZETWORNIK NAPIĘCIA STAŁEGO MIDDLEBROOKA I ČUKA

Na rysunku 1 przedstawiono podstawowy schemat ideowy przetwornika, jego praktyczną realizację oraz idealizowane przebiegi czasowe w układzie [4].



Rys. 1. a) Podstawowy schemat ideowy impulsowego przetwornika napięcia stałego Middlebrooka i Čuka b) praktyczna realizacja przetwornika, c) idealizowane przebiegi czasowe w układzie

W przedziale czasu $D'T_o = (1-D)\,T_o$, tranzystor T znajduje się w stanie odcięcia, dioda D zostaje spolaryzowana w kierunku przewodzenia (na schemacie ideowym przełącznik S znajduje się w położeniu B), a pojemność C_1 ładowana jest prądem i_1 ze źródła napięcia zasilającego E_z. W przedziale czasu DT_o (rys. 1c), gdy tranzystor T zostaje wysterowany do stanu nasycenia, dioda D zostaje spolaryzowana w kierunku zaporowym, a kondensator C_1 rozładowuje się prądem i_2, przekazując energię do obciążenia. Zarówno prąd po stronie pierwotnej i_1, jak i po stronie wtórnej i_2 są przebiegami ciągłymi, zawierającymi składowe stałe I_1, I_2 oraz składowe zmienne, utworzone w przybliżeniu z liniowo

narastających i opadających funkcji. Amplitudy składowych zmiennych prądów i_1, i_2 maleją wraz ze zwiększaniem indukcyjności L_1 i L_2 i dla zadanego okresu T_o mogą być pomijalnie małe w porównaniu z wartościami składowych stałych. Pojemności C_1 i C_2 powinny być na tyle duże, aby otrzymać małe zmiany napięć na tych pojemnościach, zarówno w taktach ich ładowania, jak i rozładowania. Z warunku, że średnia wartość napięcia na każdej indukcyjności wynosi zero, otrzymujemy:

$$\begin{aligned} E_z D T_o &= (U_{c1} - E_z) D' T_o, \\ U D' T_o &= (U_{c1} - U) D T_o \end{aligned} \tag{1}$$

(oznaczenia zgodne z rys. 1).

Z równań (1) otrzymujemy

$$\frac{U}{E_z} = -\frac{D}{D'} = -\frac{D}{1-D}. \tag{2}$$

Stosunek prądów I_1, I_2 możemy wyznaczyć z warunku, że napięcie na kondensatorze C_1 zmienia się w czasie jego ładowania o taką samą wartość, jak w czasie jego rozładowania:

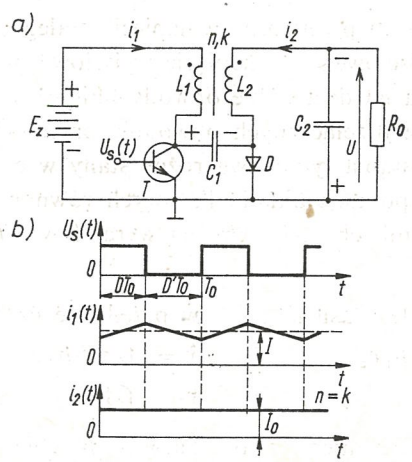
$$\begin{aligned} C_1 \Delta U_c &= I_1 D T_o, \\ C_1 \Delta U_c &= I_2 D' T_o; \end{aligned} \tag{3}$$

stąd

$$\frac{I_2}{I_1} = \frac{D'}{D} = \frac{1-D}{D}. \tag{4}$$

Główną zaletą opisanego układu przetwornika napięcia stałego jest to, że przy prostej strukturze układowej zapewnia on ciągłe prądy zarówno po jego pierwotnej jak i wtórnej stronie, przez co nie wymaga stosowania na jego wejściu dodatkowego filtru.

Niewielka modyfikacja opisanego układu, polegająca na sprzężeniu ze sobą indukcyjności L_1 i L_2 prowadzi do dalszego polepszenia jego właściwości. Układ taki wraz z idealizowanymi przebiegami czasowymi przedstawiono na rysunku 2. Dla średnich wartości napięć i prądów w układzie z rysunku 2 obowiązują w dalszym ciągu zależności (2), (4).



Rys. 2. a) Schemat ideowy ulepszonego przetwornika, b) idealizowane przebiegi czasowe w układzie

Jak wykazano w pracach [2, 3], zależnie od parametrów zastosowanego transformatora:

$$n = \frac{L_1}{L_2}, \quad k = \frac{L_M}{\sqrt{L_1 L_2}}, \tag{5}$$

gdzie:
L_1 — indukcyjność uzwojenia pierwotnego,
L_2 — indukcyjność uzwojenia wtórnego,
L_M — indukcyjność wzajemna,
n — przekładnia transformatora,
k — współczynnik indukcyjności wzajemnej,
możemy otrzymać dwa interesujące przypadki.

Gdy $n = 1 \neq k$, otrzymujemy tzw. warunek zrównoważonej redukcji, przy którym amplitudy składowych zmiennych prądów i_1, i_2 maleją $(1+k)$ razy, w porównaniu do wartości otrzymywanych w układzie z rysunku 1 (tzn., gdy $k = 0$). Praktycznie oznacza to prawie dwukrotne zmniejszenie amplitud składowych zmiennych prądów: pierwotnego i wtórnego. W przypadku tzw. niezrównoważonej redukcji, gdy n jest bliskie wartości k, amplituda składowej zmiennej prądu pierwotnego i_1 pozostaje bez zmiany w porównaniu z układem z rysunku 1, zaś amplituda składowej zmiennej prądu i_2 zredukowana zostaje o ponad rząd wartości. W przypadku tzw. warunku dopasowania, gdy $n = k$, po stronie wtórnej otrzymujemy tylko składową stałą I_2, jak zaznaczono na rysunku 2b. Wyeliminowanie składowej zmiennej po stronie wtórnej przetwornika, w przypadku dopasowania $(n = k)$, otrzymujemy w przybliżeniu niezależnie od wartości okresu przełączania T_o, współczynnika wypełnienia prostokątnego przebiegu sterującego D, czy też parametrów obwodu R_L, C_2, L_1, L_2. Pozwala to na wyeliminowanie pojemności C_2 i uzyskanie bardzo dobrych parametrów dynamicznych przetwornika.

3. KANONICZNY MODEL ZASTĘPCZY DOWOLNYCH, DWUSTANOWYCH PRZETWORNIKÓW NAPIĘCIA STAŁEGO

W dowolnym impulsowym przetworniku napięcia stałego, zachodzący w nim proces przetwarzania napięcia wejściowego na napięcie wyjściowe polega na cyklicznym przełączaniu struktury regulatora na dwa różne obwody liniowe, zawierające elementy bierne. Przy założeniu, że każdy z przełączanych obwodów przetwornika znajduje się w stanie przewodzenia ciągłego, wystąpią tylko dwa różne stany w całym układzie. Zatem przetwornik można opisać za pomocą układu liniowych równań różniczkowych pierwszego rzędu o stałych współczynnikach, które można wyrazić w postaci wektorowo-macierzowej,

w przedziale czasu dT_o:	w przedziale czasu $d'T_o$:
$\dot{x} = A_1 x + b_1 e_z$	$\dot{x} = A_2 x + b_2 e_z$
$y_1 = C_1^T x,$	$y_2 = C_2^T x,$

(6)

gdzie: d — współczynnik wypełnienia prostokątnego przebiegu sterującego o okresie T_o $(d' = 1-d)$,

x — kolumnowy wektor zmiennych stanu (prądy w indukcyjnościach i napięcia na pojemnościach),

e_z — kolumnowy wektor zmiennych wejściowych (w naszym przypadku wielkością wejściową jest napięcie E_z),

A_1, A_2, b_1, b_2 — macierze obwodów,

y_1, y_2 — poszukiwane wielkości wyjściowe,

C_1^T, C_2^T — macierze określające związek między wielkością wyjściową a wektorem stanu.

Równania (6), opisujące dwa oddzielne obwody liniowe, wynikające z kolejnych faz przełączania przetwornika w ciągu okresu T_o, można zastąpić jednym równaniem, które z pewnym przybliżeniem opisze zachowanie się wypadkowego obwodu, w czasie trwania całego okresu T_o. Taki uśredniony układ opisany za pomocą zmiennych stanu możemy otrzymać w wyniku sumowania, uśrednionych za cały okres T_o, równań (6), gdy częstotliwość przełączania F_o jest dużo większa niż częstotliwość graniczna dolnoprzepustowego filtru wyjściowego ($F_o \gg f_g$):

$$\begin{aligned} \dot{x} &= (dA_1 + d'A_2)x + (db_1 + d'b_2)e_z, \\ y &= dy_1 + d'y_2 = (dC_1^T + d'C_2^T)x. \end{aligned} \tag{7}$$

Metoda uśrednionych równań stanu wynika z liniowej aproksymacji wyrażeń:

$$\begin{aligned} e^{dA_1T_o} &\cong I + dA_1T_o, \\ e^{d'A_2T_o} &\cong I + d'A_2T_o, \end{aligned} \tag{8}$$

co odpowiada uwzględnieniu pierwszego wyrazu w szeregu Baker-Campbell-Hausdorffa [1]

$$AT_o = (dA_1 + d'A_2)T_o + dd'(A_1A_2 - A_2A_1)T_o^2 + \cdots, \tag{9}$$

gdzie:

$$e^{AT_o} = e^{d'A_2T_o}e^{dA_1T_o},$$

I — macierz jednostkowa, tego samego rzędu co macierz A.

Podstawiając:

$$\begin{aligned} A &= dA_1 + d'A_2, \\ b &= db_1 + d'b_2, \\ C^T &= dC_1^T + d'C_2^T, \end{aligned} \tag{10}$$

z równań (7, 10) otrzymujemy:

$$\begin{aligned} \dot{x} &= Ax + be_z, \\ y &= C^Tx. \end{aligned} \tag{11}$$

Załóżmy, że w każdym okresie impulsowania zmienia się współczynnik wypełnienia prostokątnej fali sterującej oraz wektor zmiennych wejściowych:

$$\begin{aligned} d(t) &= D + \hat{d}, \\ d'(t) &= D' - \hat{d}, \\ e_z(t) &= E_z + \hat{e}_z, \end{aligned} \tag{12}$$

gdzie: D, D', E_z są wartościami stałymi, na które są nałożone składowe zmienne. Wtedy

zaburzone zostają wektor stanu oraz wielkości wyjściowe:

$$x = X + \hat{x},$$
$$y = Y + \hat{y}. \tag{13}$$

Z równań (11)÷(13) otrzymujemy:

$$\dot{\hat{x}} = AX + bE_z + A\hat{x} + b\hat{e}_z + [(A_1 - A_2)X + (b_1 - b_2)E_z]\hat{d} + [(A_1 - A_2)\hat{x} + (b_1 - b_2)\hat{e}_z]\hat{d},$$
$$Y + \hat{y} = C^T X + C^T\hat{x} + (C_1^T - C_2^T)X\hat{d} + (C_1^T - C_2^T)\hat{x}\hat{d}. \tag{14}$$

Układy równań (14) są nieliniowe, wskutek występujących w nich iloczynów $\hat{x}\hat{d}$, $\hat{e}_z\hat{d}$.

Zakładając, że odchylenia składowych zmiennych od wartości stałych są niewielkie:

$$\frac{\hat{e}_z}{E_z} \ll 1, \quad \frac{\hat{d}}{D} \ll 1, \quad \frac{\hat{x}}{X} \ll 1, \tag{15}$$

otrzymujemy aproksymowany, małosygnałowy model uśredniony, opisany za pomocą zmiennych stanu.

Rozdzielając równania (14) dla składowych stałych i zmiennych oraz uwzględniając zależności (15), otrzymujemy:

dla składowych stałych

$$X = -A^{-1}bE_z,$$
$$Y = C^T X = -C^T A^{-1} bE_z, \tag{16}$$

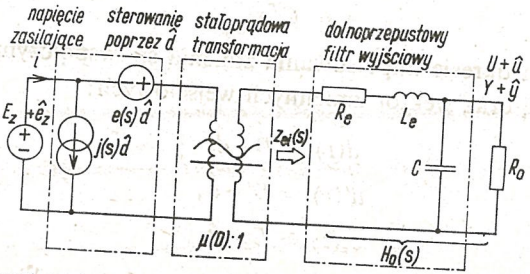
dla składowych zmiennych

$$\dot{\hat{x}} = A\hat{x} + b\hat{e}_z + [(A_1 - A_2)X + (b_1 - b_2)E_z]\hat{d},$$
$$\hat{y} = C^T\hat{x} + (C_1^T - C_2^T)X\hat{d}. \tag{17}$$

Równania (16), (17) reprezentują małosygnałowe modele w zakresie małych częstotliwości dowolnego, dwustanowego przetwornika napięcia stałego, pracującego w stanie przewodzenia ciągłego. Stosując przekształcenie Laplace'a równań (17), otrzymujemy:

$$\hat{x}(s) = (sI - A)^{-1}b\hat{e}_z(s) + (sI - A)^{-1}[(A_1 - A_2)X + (b_1 - b_2)E_z]\hat{d}(s),$$
$$y(s) = C^T\hat{x}(s) + (C_1^T - C_2^T)X\hat{d}(s). \tag{18}$$

Na rys. 3 przedstawiono kanoniczny model zastępczy przetwornika napięcia stałego dowolnego typu, opisujący trzy jego podstawowe bloki funkcyjne: blok sterowania poprzez



Rys. 3. Kanoniczny schemat zastępczy dwustanowego przetwornika napięcia stałego dowolnego typu

zmianę współczynnika wypełnienia fali prostokątnej, blok stałoprądowej transformacji napięcia wejściowego na wyjściowe, dolnoprzepustowy filtr wyjściowy.

Dla układu z rys. 3 możemy napisać:

$$y(s) = \frac{1}{\mu(D)} H_e(s) e_z(s) + e(s) + \frac{1}{\mu(D)} H_e(s) \hat{d}(s),$$
$$i(s) = \frac{1}{[\mu(D)]^2 Z_{ei}(s)} e_z(s) + \left[j(s) + \frac{e(s)}{[\mu(D)]^2 Z_{ei}(s)} \right] \hat{d}(s), \tag{19}$$

gdzie:

$\hat{y}(s)$ — składowa zmienna napięcia wyjściowego,

$i(s)$ — składowa zmienna prądu wejściowego,

$Z_{ei}(s)$ — operatorowa impedancja na zaciskach wejściowych filtru,

$H_e(s)$ — operatorowa funkcja przenoszenia filtru wyjściowego, obciążonego rezystancją R_o.

Równania (19) mogą być zapisane w postaci ogólnej, przy pomocy odpowiednich funkcji przenoszenia:

$$\hat{y}(s) = G_{uz} \hat{e}_z(s) + G_{ud} \hat{d}(s),$$
$$i(s) = G_{iz} \hat{e}_z(s) + G_{id} \hat{d}(s). \tag{20}$$

Występujące w równaniach (20) operatorowe funkcje przenoszenia są wielkościami znanymi z układu równań (18).

Tak więc, z równań (18), (20) otrzymujemy:

$$e(s) = \frac{G_{uz}(s)}{G_{uz}(s)},$$
$$j(s) = G_{id}(s) - e(s) G_{iz}(s), \tag{21}$$
$$H_e(s) = \mu G_{uz}(s).$$

Parametr $\frac{1}{\mu(D)}$ oznacza idealne, stałoprądowe wzmocnienie układu (bezstratne indukcyjności i pojemności). Parametr ten możemy wyznaczyć z równania (18)

$$\frac{Y}{E_z} = -C^T A^{-1} b = \frac{1}{\mu(D)}, \tag{22}$$

gdy w macierzy układu przyjmiemy zerowe wartości rezystancji pasożytniczych.

3.1. Kanoniczny model zastępczy przetwornika Middlebrooka i Čuka

Stosując zależności (9) ÷ (22) możemy wyznaczyć kanoniczny schemat zastępczy układu z rysunku 2. Schemat ten przedstawiono na rysunku 4, na którym przyjęto oznaczenia:

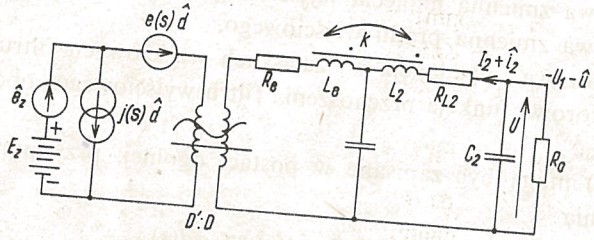
R_{L1} — rezystancja uzwojenia pierwotnego (L_1),

R_{L2} — rezystancja uzwojenia wtórnego (L_2),

$$R_e = \left(\frac{D}{D'}\right)^2 R_{L1}, \quad L_e = \left(\frac{D}{D'}\right)^2 L_1, \quad C_e = \frac{C_1}{D^2}, \tag{23}$$

$$e(s) = \frac{U}{D^2} \frac{\left(1+\frac{R_{L2}}{R_o}\right)\left(sR_eC_eD' + s^2\frac{(L_1-L_M)C_1}{D} + 1\right) + s\left(\frac{DL_M}{D'R_o} - \frac{L_e}{R_o}\right) - \frac{R_e}{R_o}}{1 - \frac{L_MC_1}{DD'}s^2}, \tag{24}$$

$$j(s) = \frac{U}{(D')^2R_o} \frac{1 - sC_eR_oD'\left(1+\frac{R_{L2}}{R_o}\right)}{1 - \frac{L_MC_1}{DD'}s^2}. \tag{25}$$



Rys. 4. Kanoniczny schemat zastępczy układu z rys. 2

Uwzględniając rezystancje uzwojeń transformatora, na podstawie schematu z rysunku 4, możemy wyznaczyć związki pomiędzy składowymi stałymi napięć i prądów:

$$\frac{U}{E_z} = -\frac{D}{D'} \frac{1}{1+\alpha_1\left(\frac{D}{D'}\right)^2 + \alpha_2}, \tag{26}$$

$$\frac{I_2}{I_1} = \frac{D'}{D}, \tag{27}$$

gdzie:

$$\alpha_1 = \frac{R_{L1}}{R_o}, \quad \alpha_2 = \frac{R_{L2}}{R_o}. \tag{28}$$

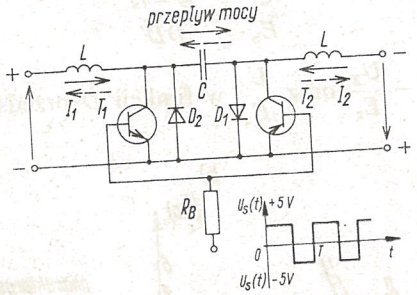
W palszej części pracy, do analizy charakterystyk częstotliwościowych wzmacniacza przyjęto nieco uproszczony model zastępczy z rysunku 4, w którym $R_{L1} = R_{L2} = 0$. Dla uproszczonego schematu zastępczego jego parametry określone są zależnościami:

$$e(s) = \frac{U}{D^2} \frac{1 - \frac{D(DL_1+D'L_M)s}{(D')^2R_o} + \frac{(L_1-L_M)C_1s^2}{D'}}{1 - \frac{s^2L_MC_1}{DD'}}, \tag{29}$$

$$j(s) = \frac{U}{(D')^2R_o} \frac{1 - \frac{sC_1R_oD'}{D^2}}{1 - \frac{s^2L_MC_1}{DD'}}. \tag{30}$$

4. BUDOWA WZMACNIACZA KLASY *D* MAŁEJ CZĘSTOTLIWOŚCI, ZREALIZOWANEGO W OPARCIU O PRZETWORNIK MIDDLEBROOKA I ĆUKA

Opisany w rozdz. 1 przetwornik napięcia stałego charakteryzuje się jednokierunkowym przepływem prądu (i mocy) od źródła zasilającego po obciążenia. Przez dodanie komplementarnych kluczy diodowo-tranzystorowych (T_2, D_2), jak na rysunku 5, otrzymujemy

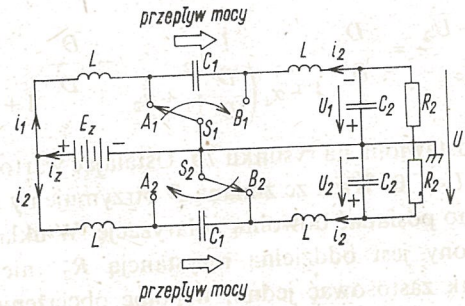


Rys. 5. Symetryczny regulator impulsowy

w pełni symetryczny układ, w którym prądy mogą płynąć w obydwu kierunkach, zależnie od współczynnika wypełnienia fali prostokątnej, sterującej układem.

Zastosowanie komplementarnych tranzystorów T_1, T_2 z uziemionymi emiterami pozwala na stosowanie najprostszego z możliwych układów sterujących te tranzystory. Oprócz tego załączenie jednego z tych tranzystorów jest możliwe dopiero wtedy, gdy z bazy drugiego tranzystora zostanie usunięty nadmiarowy ładunek nasycenia i przejdzie on w stan odcięcia. Zatem w układzie występuje automatyczne zabezpieczenie przed równoczesnym przewodzeniem obydwu tranzystorów w czasie ich przełączania.

Zakładając, że po stronie wtórnej przetwornika zostanie dołączone obciążenie, to jednak w układzie tym nie ma możliwości zmiany polaryzacji napięcia wyjściowego.



Rys. 6. Schemat ideowy układu przeciwsobnego

W celu uzyskania takiej możliwości należy zastosować układ przeciwsobny, złożony z dwóch identycznych układów z rysunku 5, sterowanych w przeciwnej fazie. Otrzymany układ ideowy przedstawiono na rys. 6.

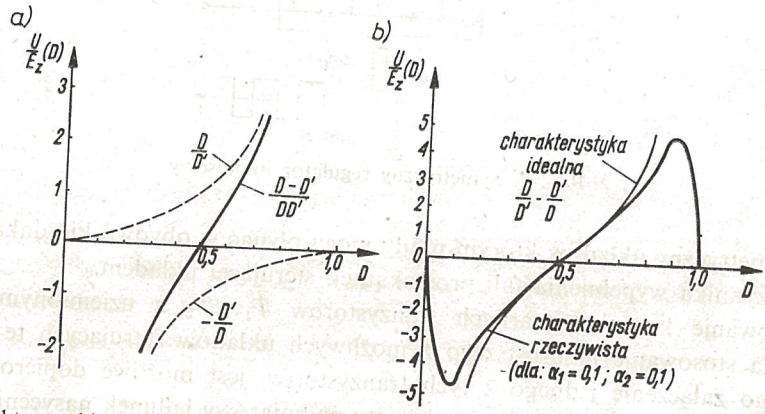
Przy spełnieniu warunków przewodzenia ciągłego oraz przy bezstratnych indukcyjnościach L, napięcia U_1 i U_2 wynoszą:

$$U_1 = \frac{D}{D'} E_z, \quad U_2 = \frac{D'}{D} E_z. \tag{31}$$

Wyznaczając różnicę tych napięć: $U = U_1 - U_2$, otrzymujemy

$$\frac{U}{E_z} = \frac{D - D'}{DD'}. \tag{32}$$

Przebieg zależności $\frac{U_1}{E_z}$, $-\frac{U_2}{E_z}$ oraz $\frac{U}{E_z}$ w funkcji D przedstawiono na rysunku 7a.



Rys. 7. Charakterystyki statyczne wzmacniacza: a) z idealnym transformatorem, b) z uwzględnieniem rezystancji uwzojeń transformatora

Większą liniowość charakterystyki statycznej wzmacniacza, w jej środkowej części, otrzymujemy w przypadku, gdy rezystancje uzwojeń transformatora nie są równe zeru.

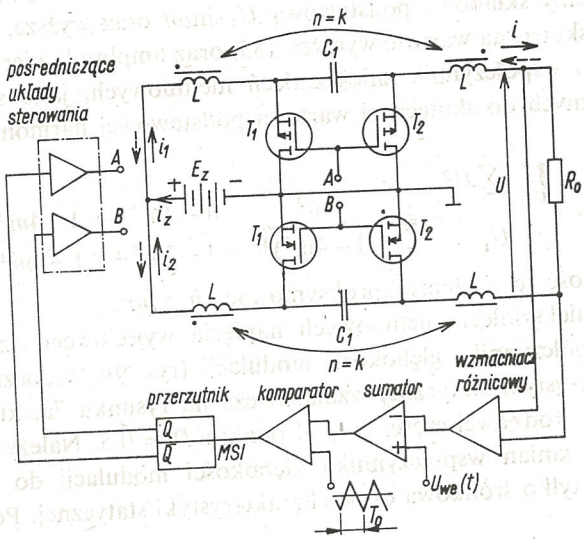
Wykorzystując zależności (26) ÷ (28) możemy wyznaczyć napięcie różnicowe wzmacniacza

$$\frac{U}{E_z}(D) = \frac{U_1 - U_2}{E_z} = \frac{D}{D'} \frac{1}{1 + \alpha_1 \left(\frac{D}{D'}\right)^2 + \alpha_2} - \frac{D'}{D} \frac{1}{1 + \alpha_1 \left(\frac{D'}{D}\right)^2 + \alpha_2}. \tag{33}$$

Zależność ta została przedstawiona na rysunku 7b. Ustalając wartość spoczynkową $D = 0{,}5$ otrzymujemy na wyjściu $U = 0$. Wraz ze zmianą D otrzymujemy zmianę napięcia wyjściowego, przy czym może ono posiadać dowolną polaryzację. W układzie na rysunku 6 każdy z przetworników obciążony jest oddzielną rezystancją R_2, nie oddziałując wzajemnie na siebie. Możemy jednak zastosować jedno, wspólne obciążenie „pływające" (odizolowane od masy układu), przez które płynie prąd różnicowy, jak to przedstawiono na rysunku 8. Ponadto na rysunku 8 przełączniki S_1 i S_2 zostały zastąpione kluczami tranzystorowymi.

Zastosowano komplementarne tranzystory mocy *VMOS* typu *HEXFET*, charakteryzujące się małą rezystancją w stanie załączenia i dużą szybkością przełączania. Należy

podkreślić, że przy inwersyjnej polaryzacji tranzystora prąd drenu jest wypadkową prądu przewodzącego kanału (przy wysterowaniu tranzystora) i prądu pasożytniczej diody równoległej w strukturze tranzystora *VMOS*. Na rysunku 8 indukcyjności zostały sprzężone w warunkach dopasowania. Dzięki dopasowaniu prąd wyjściowy nie zawiera składowej



Rys. 8. Schemat ideowy wzmacniacza klasy *D* małej częstotliwości zrealizowanego w oparciu o przetwornik Middlebrooka i Čuka

zmiennej o częstotliwości fali nośnej (sygnał sterujący jest zmodulowany w szerokości impulsów), co pozwala na wyeliminowanie z układu pojemności C_2. Prąd i_z pobierany ze źródła zasilającego E_z stanowi różnicę prądów i_1 oraz i_2, których składowe zmienne utworzone są w przybliżeniu z odcinków liniowo narastających i opadających funkcji o nachyleniu odwrotnie proporcjonalnym do wartości indukcyjności L. Zatem, przy współczynniku wypełnienia $D = 0{,}5$, prąd $i_z = i_1 - i_2$ nie zawiera również składowej zmiennej o częstotliwości nośnej.

5. ANALIZA ZNIEKSZTAŁCEŃ NIELINIOWYCH I OPTYMALIZACJA CHARAKTERYSTYKI STATYCZNEJ WZMACNIACZA

Przy liniowej zależności szerokości impulsu od napięcia wejściowego modulatora, równanie (32) opisuje charakterystykę statyczną wzmacniacza, tj. zależność napięcia wyjściowego od napięcia wejściowego wzmacniacza. Ustalając początkową wartość współczynnika wypełnienia fali prostokątnej $D = 0{,}5$ oraz podając na wejście modulatora sinusoidalny przebieg modulujący $U_{we} = U_m \sin \omega t$, otrzymujemy:

$$d(t) = 0{,}5 + m \sin \omega t, \qquad (34)$$
$$d'(t) = 0{,}5 - m \sin \omega t,$$

gdzie $m \leqslant 0{,}5$ — współczynnik głębokości modulacji.

Z równań (32), (34) otrzymujemy

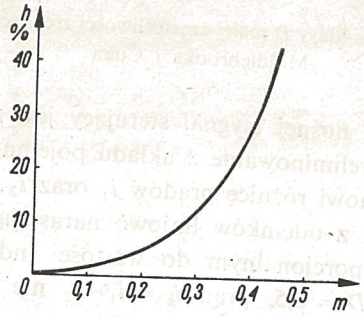
$$U(U_{we}) = \frac{2m\sin\omega t}{0{,}25 - m^2\sin^2\omega t} E_z. \tag{35}$$

Zależność (35) jest periodyczną i nieparzystą funkcją czasu. Po rozwinięciu jej w szereg Fouriera otrzymujemy składową podstawową $U_1 \sin\omega t$ oraz wyższe, nieparzyste harmoniczne. Obliczając skuteczną wartość wyrażeń (35) oraz amplitudę pierwszej harmonicznej, możemy wyznaczyć współczynnik zniekształceń nieliniowych, jako stosunek skutecznej wartości harmonicznych do skutecznej wartości podstawowej harmonicznej

$$h = \frac{\sqrt{\sum_{i=2}^{\infty} U_i^2}}{U_1} = \frac{m^2}{(1-4m^2)^{3/4}} \frac{1-4m^2+\sqrt{1-4m^2}}{1-2m^2+\sqrt{1-4m^2}}. \tag{36}$$

Dla $m < 0{,}2$ zależność tę możemy aproksymować: $h \cong m^2$.

Współczynnik zniekształceń nieliniowych napięcia wyjściowego szybko rośnie wraz ze zwiększaniem współczynnika głębokości modulacji (rys. 9). Widoczne to jest również z przebiegu charakterystyki statycznej wzmacniacza na rysunku 7a, którą można uznać za liniową jedynie w środkowej części, wokół punktu $D = 0{,}5$. Należy zatem ograniczyć dopuszczalny zakres zmian współczynnika głębokości modulacji do małych wartości, tak aby wykorzystać tylko środkową część charakterystyki statycznej. Pomimo tego ogra-



Rys. 9. Charakterystyka zniekształceń nieliniowych napięcia wyjściowego

niczenia w rozważanym kładzie, możemy uzyskać taką samą moc wyjściową jak w klasycznych wzmacniaczach klasy D małej częstotliwości. Jest to możliwe ponieważ nachylenie charakterystyki statycznej (rys. 7), wyrażające wzmocnienie napięciowe stopnia końcowego jest większe od jedności (w punkcie charakterystyki statycznej $D = 0{,}5$ jej nachylenie wynosi $k = 8$). Przyjmując jako dopuszczalną wartość współczynnika głębokości modulacji $m = 0{,}12$, amplituda napięcia wyjściowego $U = 1{,}02\ E_z$ jest nawet nieco większa od napięcia zasilającego. Dla porównania, w klasycznym stopniu końcowym wzmacniacza klasy D, przy maksymalnej głębokości modulacji $m = 0{,}5$, amplituda napięcia wyjściowego może być co najwyżej równa wartości połówkowego napięcia zasilającego. Przy $m < 0{,}1$ współczynnik zniekształceń nieliniowych napięcia wyjściowego $h < 1\%$. Zastosowanie ujemnego sprzężenia zwrotnego pozwoli na dalsze zmniejszenie zniekształceń nieliniowych. Większą liniowość charakterystyki statycznej wzmacniacza

; to w szerszym zakresie zmian współczynnika głębokości modulacji niż wyżej otrzymujemy w przypadku, gdy rezystancje uzwojeń transformatora nie są równe zeru (rys. 7b). Optymalizacja charakterystyki statycznej polega na znalezieniu wartości parametrów α_1 i α_2, przy których zależność (33) jest najbardziej zbliżona do funkcji liniowej wokół spoczynkowego punktu pracy $D = 0{,}5$. Podstawiając do równania (33): $D = 0{,}5+m$, $D' = 0{,}5-m$ otrzymujemy

$$\frac{U}{E_z}(m) = \frac{8m}{1-4m^2}\,\frac{1-\dfrac{\alpha_1}{1+\alpha_2}}{K(m)}, \tag{37}$$

gdzie

$$K(m) = 1+\alpha_2+\frac{\alpha_1}{1+\alpha_2}+2\alpha_1\frac{1+24m^2+16m^4}{(1-4m^2)^2}. \tag{38}$$

Rozwijając mianownik równania (37): $(1-4\ m^2)\,K(m)$ w szereg Maclaurina i uwzględniając dwa pierwsze wyrazy tego szeregu otrzymujemy wyrażenie niezależne od m, gdy:

$$14\alpha_1-\left(1+\alpha_2+\frac{\alpha_1}{1+\alpha_2}\right)=0. \tag{39}$$

Po rozwiązaniu równania (39) otrzymujemy

$$\alpha_1 = (7\pm4\sqrt{3})(1+\alpha_2). \tag{40}$$

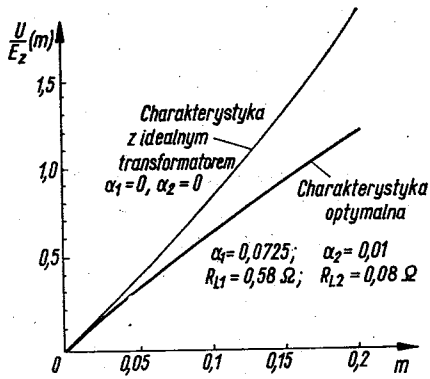
Dodatni znak w rozwiązaniu (40) prowadzi do zależności

$$\alpha_1 = (7+4\sqrt{3})(1+\alpha_2) = 13{,}93(1+\alpha_2) > 13{,}93.$$

Rozwiązanie to jest praktycznie nieużyteczne, bowiem wymaga, aby rezystancja uzwojenia pierwotnego R_{L1} była prawie 14 razy większa od rezystancji obciążenia.

Drugie rozwiązanie jest możliwe fizycznie i określa warunki dla otrzymania optymalnej charakterystyki statycznej wzmacniacza

$$\alpha_1 = (7-4\sqrt{3})(1+\alpha_2) = 0{,}0718(1+\alpha_2). \tag{41}$$



Rys. 10. Optymalna charakterystyka statyczna wzmacniacza

Normalnie $\alpha_2 = \frac{R_{L2}}{R_o} \ll 1$, zatem

$$\alpha_{1opt} \cong 0{,}0718.$$

Przykładowo dla rezystancji obciążenia $R_o = 8$ [Ω], optymalna wartość rezystancji uzwojenia pierwotnego wynosi $R_{L1opt} = 0{,}58$ [Ω]. Przy praktycznej realizacji wzmacniacza najłatwiej jest wykonać symetryczny transformator o indukcyjnościach $L_1 = L_2 = L$ i $R_{L1} = R_{L2} = R_L$, zaś w obrót pierwotny transformatora dołączyć dodatkowo niewielką rezystancję szeregową, w celu spełnienia warunku: $R_{L1opt} = R_d + R_L$.

Na rysunku 10 przedstawiono przebieg optymalnej charakterystyki statycznej wzmacniacza. Jest ona prawie idealnie liniowa w zakresie $m < 0{,}2$.

6. CHARAKTERYSTYKI CZĘSTOTLIWOŚCIOWE WZMACNIACZA

W celu wyznaczenia charakterystyk częstotliwościowych wzmacniacza wykorzystamy kanoniczny schemat zastępczy przetwornika (rys. 4), przy założeniu, że rezystancje uzwojeń transformatora są równe zeru. Pominięcie tych rezystancji przy analizie przebiegów zmiennych nie wnosi dużych błędów, ponieważ $R_{L1} + R_{L2} \ll R_o$. Zakładając, że wzmacniacz pracuje w warunkach dopasowania ($n = k$, $L_M = L_1$) oraz pomijając pojemność wyjściową C_2 jako zbędną przy dopasowaniu, na podstawie schematu zastępczego możemy wyznaczyć operatorową funkcję przenoszenia

$$G(D) = \frac{\hat{u}(s)}{\hat{d}(s)}, \tag{41}$$

gdzie:

$\hat{u}(s)$ — transformata składowej zmiennej napięcia wyjściowego,

$\hat{d}(s)$ — transformata składowej zmiennej przebiegu modulującego (współczynnika wypełnienia zmodulowanej fali prostokątnej).

Wykorzystując ponadto zależność (2), otrzymujemy

$$G(D) = \frac{\frac{E_z}{(D')^2}\left(1 - \frac{DL_1}{(D')^2 R_o}s\right)\left(1 - \frac{L_1 C_1}{D^2}s^2\right)}{\left(1 - \frac{L_1 C_1}{DD'}s^2\right)} \times$$

$$\times \frac{1}{1 + \frac{L_1 D^2 + 2L_1(D')^2 + L_2(D')^2}{R_o(D')^2}s + \frac{L_1 C_1}{(D')^2}s^2 + \frac{L_1 C_1 [L_2 D^2 - L_1(D')^2]}{R_o(DD')^2}s^3}. \tag{42}$$

Rozważany wzmacniacz jest układem przeciwsobnym, w którym dwa przetworniki sterowane w przeciwfazie pracują różnicowo na wspólne obciążenie. Wraz ze wzrostem współczynnika wypełnienia przebiegu sterującego górny przetwornik maleje współczynnik wypełnienia przebiegu sterującego dolny przetwornik (i na odwrót). Jeżeli funkcja przenoszenia górnego przetwornika wynosi $G(D)$, to dla dolnego funkcja ta wynosi $-G(D')$ (a nie $+G(D')$. Z tego względu różnicowa funkcja przenoszenia wzmacniacza wynosi

$$G_r(D) = G(D) + G(D'), \tag{43}$$

gdzie: $G(D)$, $G(D')$ określone są zależnością (42).

Przy $D = D' = 0{,}5$ oraz $m < 0{,}1$ lub $m < 0{,}2$ w przypadku doboru optymalnej wartości R_{L1}, z równań (42), (43) otrzymujemy

$$G_r|_{D=0,5} = 8E_z \frac{1-2\frac{L_1}{R_o}s}{1+s\frac{3L_1+L_2}{R_o}+4s^2L_1C_1+4s^3\frac{L_1C_1L_e}{R_o}}, \tag{44}$$

gdzie: $L_e = L_2 - L_1$ — indukcyjność rozproszenia.

Przy $s = 0$ otrzymujemy stałoprądowe wzmocnienie układu $k = 8$. Zależność (44) może być z dobrym przybliżeniem aproksymowana do postaci

$$G_r = 8E_z \frac{1-2\frac{L_1}{R_o}s}{\left(1+s\frac{4L_1}{R_o}+s^2 4L_1C_1\right)\left(1+s\frac{L_e}{R_o}\right)}, \tag{45}$$

gdy $C_1R_o \gg \frac{L_e}{R_o}$.

W praktyce warunek ten jest zawsze spełniony.

Wprowadzając oznaczenia:

$$\begin{aligned} f_c &= \frac{1}{4\pi\sqrt{L_1C_1}}, & \omega_c &= 2\pi f_c, \\ f_z &= \frac{R_o}{4\pi L_1}, & \omega_z &= 2\pi f_z, \\ f_p &= \frac{R_o}{2\pi L_e}, & \omega_p &= 2\pi f_p, \\ Q &= \frac{R_o}{2}\sqrt{\frac{C_1}{L_1}}, \end{aligned} \tag{46}$$

zależność (45) możemy przedstawić

$$G_r = 8E_z \frac{1-\frac{s}{\omega_z}}{\left[1+\frac{s}{Q\omega_c}+\left(\frac{s}{\omega_c}\right)^2\right]\left(1+\frac{s}{\omega_p}\right)}. \tag{47}$$

Funkcja przenoszenia posiada rzeczywiste zero, leżące w prawej półpłaszczyźnie zmiennej zespolonej, co mogłoby stwarzać potencjonalne niebezpieczeństwo niestabilności wzmacniacza z ujemnym sprzężeniem zwrotnym. Niebezpieczeństwo to nie występuje w rzeczywistym układzie, w którym rezystancje uzwojeń transformatora nie są równe zeru. Szczegółowa analiza operatorowej funkcji przenoszenia $G_r = \frac{\hat{u}(s)}{\hat{d}(s)}$, wyznaczonej na podstawie rysunku 4 pozwala na stwierdzenie, że rezystancje R_{L1} i R_{L2} w niewielkim stopniu wpływają

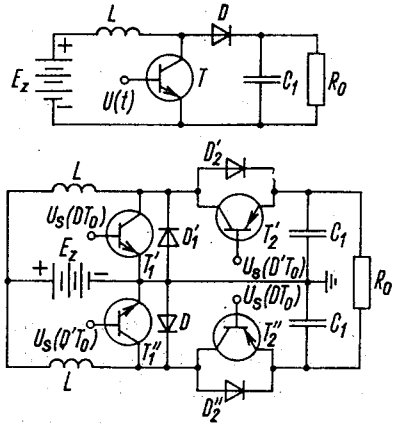
na rozkład biegunów tej funkcji, ponieważ $R_{L1} \ll R_o$, $R_{L2} \ll R_o$. Natomiast w liczniku funkcji przenoszenia otrzymujemy wyrażenie

$$1 - 2s\left(\frac{L_1}{R_o} - R_{L1} C_1\right).$$

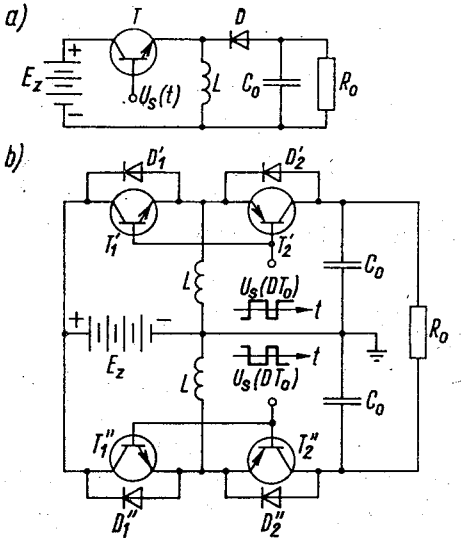
Przy spełnieniu warunku: $\frac{L_1}{R_o} - R_{L1} C_1 < 0$, funkcja przenoszenia wzmacniacza z nieidealnym transformatorem posiada rzeczywiste zero, leżące w lewej półpłaszczyźnie zmiennej zespolonej. Powyższy warunek jest łatwy do spełnienia nawet dla bardzo małych wartości R_{L1}. Ponieważ zero tej funkcji leży w zakresie niewielkich częstotliwości, więc istnieje dodatkowa możliwość redukcji tego zera z jednym biegunem, leżącym również w zakresie małych częstotliwości. Przy takiej redukcji otrzymujemy układ drugiego rzędu, chociaż fizycznie schemat wzmacniacza zawiera 6 elementów magazynujących energię.

7. INNE ROZWIĄZANIA UKŁADOWE WZMACNIACZY KLASY *D* MAŁEJ CZĘSTOTLIWOŚCI

Na rysunku 11b przedstawiono schemat ideowy stopnia końcowego wzmacniacza klasy *D*, którego zasada działania oparta jest na przetworniku napięcia stałego z szeregową indukcyjnością i równoległym tranzystorem (rys. 11a). W podobny sposób możemy utworzyć schemat ideowy stopnia końcowego wzmacniacza klasy *D*, wykorzystując przetwornik napięcia stałego z szeregowym tranzystorem i równoległą indukcyjnością. Schemat ten przedstawiono na rysunku 12. Podstawową wadą otrzymanych układów są trudności



Rys. 11. a) Schemat ideowy przetwornika z szeregową indukcyjnością i równoległym tranzystorem (*SLRT*), b) schemat ideowy stopnia końcowego wzmacniacza klasy *D* z przetwornikiem *SLRT*



Rys. 12 a) Schemat ideowy przetwornika z szeregowym tranzystorem i rownoległą indukcyjnością (*STRL*), b) schemat ideowy stopnia końcowego wzmacniacza klasy *D* z przetwornikiem *STRL*

związane ze sterowaniem tranzystorów, których emitery (lub źródła) nie są uziemione. Stosując metodę uśrednionych równań stanu możemy wyznaczyć charakterystyki nowo utworzonych wzmacniaczy klasy *D*. Jakkolwiek wzmacniacze, których schematy ideowe przedstawiono na rysunkach 11 i 12, nie posiadają żadnych zalet w porównaniu do klasycznych rozwiązań wzmacniaczy klasy *D* małej częstotliwości, to jednak stanowią dobrą ilustrację w jaki sposób mogą być realizowane nowe, nieznane jeszcze rozwiązania układowe wzmacniaczy, wykorzystując do tego celu dwustanowy przetwornik napięcia stałego.

8. WYNIKI BADAŃ EKSPERYMENTALNYCH I WNIOSKI KOŃCOWE

Do badań eksperymentalnych zbudowano wzmacniacz klasy *D* z przetwornikiem Middlebrooka i Čuka na podstawie schematu ideowego, przedstawionego na rysunku 8. W stopniu końcowym wzmacniacza zastosowano komplementarne tranzystory *VMOS* o nazwie firmowej *HEXFET* typu IRF130 (z kanałem *n*) i IRF9130 (z kanałem *p*). Zastosowano dwubrzegową modulację szerokości impulsów, przy częstotliwości nośnej $F_o = 150$ [kHz]. Indukcyjności, sprzężone w warunkach bliskich dopasowania, zrealizowano na modelowych transformatorach z rdzeniami ferrytowymi. Wartości zastosowanych elementów układu były następujące: $C_1 = 220$ [μF], $L = 120$ [μH], $R_{L1} = R_{L2} = 0{,}1$ [Ω]. W celu uzyskania optymalnej charakterystyki statycznej, w szereg z uzwojeniami pierwotnymi transformatorów włączono dodatkowe rezystancje o wartości $R_d = 0{,}48$ [Ω]. Przy napięciu zasilającym $E_z = 25$ [V] i rezystancji obciążenia $R_o = 8$ [Ω] uzyskano moc wyjściową 50 [W] i całkowitą sprawność energetyczną $\eta = 90\%$. Przy zastosowaniu ujemnego sprzężenia zwrotnego uzyskano pasmo przepustowe wzmacniacza $0 \div 20$ [kHz] i zniekształcenia nieliniowe sygnału wyjściowego $h_n < 0{,}5\%$. Współczynnik zniekształceń nieliniowych mierzono w ten sposób, że sygnał wyjściowy podano na wejście miernika zniekształceń nieliniowych za pośrednictwem filtru pasmowego o płaskiej charakterystyce amplitudowej w zakresie częstotliwości 20 Hz$\div$20 [kHz] i tłumienności 60 [dB/dek] poza tym pasmem.

Na podstawie przeprowadzonych w niniejszej pracy rozważań można stwierdzić, że wzmacniacz klasy *D* zrealizowany w oparciu o przetwornik Middlebrooka i Čuka posiada szereg zalet, takich jak:

a) łatwość sterowania tranzystorów stopnia końcowego, ponieważ emitery lub źródła tych tranzystorów są uziemione,
b) generuje małe zakłócenia radiowe, ponieważ prąd pobierany z zasilacza ma charakter ciągły,
c) wymaga stosowania jednego źródła zasilającego.

Należy również stwierdzić, że wzmacniacze klasy *D* małej częstotliwości stanowią szeroką klasę układów, bowiem mogą być realizowane w oparciu o dowolny, dwustanowy przetwornik napięcia stałego. Analiza tych nowych, nawet nieznanych jeszcze rozwiązań układowych wzmacniaczy może być przeprowadzona według zaprezentowanej metody uśrednionych równań stanu, wykorzystując kanoniczny model zastępczy przetwornika.

BIBLIOGRAFIA

1. H. F. Baker, *On the Integration of Linear Differential Equations*, Proc. London Math. Soc., 34, 347—360, 1902, 35, 333—374, 1903, second series, 2, 293—296. 1904.
2. S. Čuk, *Switching DC-to-DC Converter with Zero Input or Output Current Ripple*, Industry Applications Society, IEEE-IAS-1978, Annual Meeting 1978, Toronto, Ontario, October 1—5, 1978, pp. 1131—1146.
3. S. Čuk, R. D. Middlebrook, *A New Optimum Topology Switching DC-to-DC Converter*, IEEE Power Electronics Specialists Conference, 1977, Palo Alto, California, June 14—16, 1977, pp. 160—179.
4. S. Čuk, R. D. Middlebrook, *Coupled-Inductor and Other-Extensions of a New Opitmum Topology Switching DC-to-DC Converter*, IAS Annual Meeting, IEEE Industry Applications Society 1977, Los Angeles, California, October 2—6, 1977, No 77 CH1246-8-IA, pp. 1110—1126.
5. S. Čuk, R. D. Middlebrook, *A General Unified Approach to Modelling Switching DC-to-DC Converters in Discontinuous Conduction Mode*, IEEE Power Electronics Specialists Conference, 1977, Palo Alto, California, June 14—16, 1977, No 77CH1213-8AES.
6. S. Čuk, R. W. Erickson, *A Conceptually New High-Frequency Switched-Mode Power Amplifier Technique Eliminates Current Ripple*, Proc. Power 5, 5th National Solid-State Power Conversion Conf., San Francisco, California, 4—6 May, 1978.
7. J. A. Dutra, *Digital amplifiers for analog power*, IEEE Trans. on Consumer Electronics, vol. CE-24, No 3, August 1978.
8. A. D. Evans, D. Hoffman, *Higher power ratings extend V-MOS FET's dominion*, Electronics, June 22, 1978, pp. 105—112.
9. St. Kuta, *Ograniczenia wzmacniaczy mocy z impulsową pracą elementów aktywnych*, Zeszyty Naukowe AGH, Nr 914, Automatyka z. 30, 1982, ss. 1—130.
10. R. D. Middlebrook, *Input Filter Considerations in Design and Application of Switching Regulators*, IAS Annual Meeting. IEEE Industry Applications Society, 1976, Chicago, October 11—14, 1976, No 76CH1122-1-1A.
11. R. D. Middlebrook, S. Čuk, *Modelling and Analysis Methods for DC-to-DC Switching Converters*, IEEE 1977, Int'l Sem. Power Converter Conf., 2-8-31 March, 1977, 77CH1183-31A.
12. R. D. Middlebrook, S. Čuk, *A General Unified Approach to Modelling Switching Converter Power Stages*, IEEE Power Electronics Specialists Conf., 1976, Cleveland, Ohio, June 8—10, 1976, No 76CH1084-3AES.
13. *Amplificateur audio 100 W classe D a V-MOS*, ELECTRON and APL. IND. (France), No 271, pp. 54—55.
14. *PWM Audio amplifiers*, ELEKTOR (Conterburg), vol. 4, No 12 (Dec. 1978), pp. 12-04/-7.

S. KUTA

D-CLASS LOW FREQUENCY AMPLIFIERS WITH TWO-STATE *DC/DC* SWITCHING CONVERTERS

Summary

The principle of operation and a theory of the low frequency *D*-class amplifier built with two-stage *DC/DC* Middlebrook-Čuk-type converter are presented in the paper. The amplifier described has some advantages over the conventional *D*-class circuit. Examples of power stages of *D*-class amplifiers, built with different *DC/DC* converters, are given. The state-space averaging method of analyzing two-stage *DC/DC* converters is offered. The method allows to obtain a canonical model of the *DC/DC* switching converter regardless of its specific configuration, and can be used in the analysis of the dynamic frequency response of a low frequency push-pull *D*-class amplifier.

S. KUTA

APPLICATION DES CONVERTISSEURS DE TENSION CONSTANTE BISTABLE AUX AMPLIFICATEURS DE PUISSANCE CLASSE *D* ET LA MÉTHODE DE LES ANALYSER

Résumé

L'article présente l'analyse théorique de l'amplificateur classe *D*, basse fréquence se servant du convertisseur de tension constante de Middlebrock et Čuk. On a présenté aussi les étages de sortie de l'amplificateur construit à base des autres convertisseurs de tension constante bistable. On a décrit la méthode des équations moyennes d'état, permettant de trouver le modèle canonique équivalent de chaque convertisseur. Ce modèle peut être utilisé à l'analyse fréquentielle des amplificateurs classe *D* basse fréquence.

S. KUTA

NF-VERSTÄRKER DER *D*-KLASSE MIT EINSATZ DER *DC*/*DC*-WANDLER

Zusammenfassung

In der Bearbeitung wurden die Wirkprinzipien sowie die Theorie eines *NF*-Verstärkers der *D*-Klasse mit Einsatz der Middlebrook-Čuk *DC*/*DC*-Wandler dargestellt. Ein solcher Verstärker ist vorteilhafter im Vergleich mit konventionellen Schaltkreisen der *D*-Klasse. Es wurden auch Beispiele für die Endstufen dieser Verstärker mit verschiedenen *DC*/*DC*-Wandlern gebracht. Zur Analyse der *DC*/*DC*-Wandler wurde eine Methode der Mittelwert-Zustandsgleichungen angewandt. Diese Methode ermöglicht die kanonische Ersatzschaltung zu bestimmen, und dieserhalb eignet sie sich zur Frequenzanalyse eines beliebigen *NF*-Verstärkers der *D*-Klasse.

С. КУТА

НИЗКОЧАСТОТНЫЕ УСИЛИТЕЛИ КЛАССА *D* ВЫПОЛНЕННЫЕ НА ОСНОВЕ ПРОИЗВОЛЬНЫХ ПРЕОБРАЗОВАТЕЛЕЙ ПОСТОЯННОГО НАПРЯЖЕНИЯ

Резюме

Представлен принцип работы и теория усилителя класса *D* звуковой частоты, построеного на основе преобразователя постоянного напряжения Миддлеброока и Цука. Указаны преимущества этого усилителя по сравнению с конвенциональным усилителем класса *D*. Представлены примеры выходных каскадов низкочастотных усилителей класса *D*, построенных на основе других преобразователей постоянного напряжения. Описан метод усредненных уравнений состояния, который позволяет определить каноническую эквивалентную модель произвольного преобразователя постоянного напряжения. На основе этой модели можно анализировать частотные характеристики любого низкочастотного усилителя класса *D*.

ROZPRAWY ELEKTROTECHNICZNE 1985, 31, z. 2, ss. 593—610

621.395.74

Metodyka tworzenia oprogramowania protokołów komunikacyjnych dla sieci teleinformatycznych

ANATOL BADACH (GDAŃSK)

Instytut Telekomunikacji Politechniki Gdańskiej

Otrzymano 21.4.1983

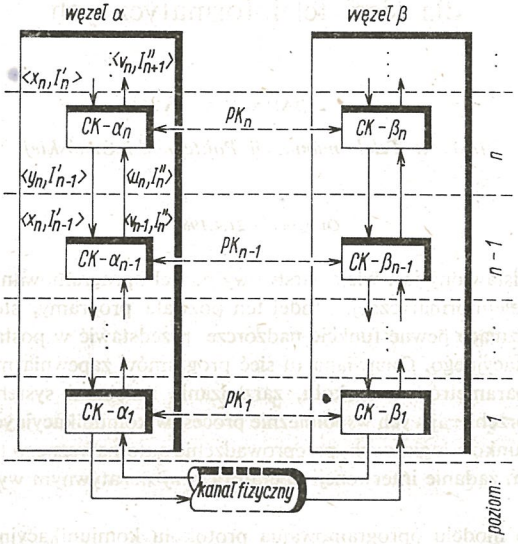
W pracy przedstawiony jest wielowarstwowy model oprogramowania protokołu komunikacyjnego sieci teleinformatycznej. Model ten pozwala programy, sterujące przebiegiem komunikacji i realizujące pewne funkcje nadzorcze, przedstawić w postaci sieci programów protokołu komunikacyjnego. Omawiana tu sieć programów zapewnia między innymi: kontrolę wszystkich parametrów protokołu, zarządzanie zasobami systemu komputerowego i ich przydział dla przebiegających współbieżnie procesów komunikacyjnych, kontrolę zdarzeń i spełniania się warunków logicznych, przeprowadzenie automatycznego testowania w sytuacjach niepożądanych, żądanie interwencji operatora przy negatywnym wyniku automatycznego testowania.

Przedstawienie modelu oprogramowania protokołu komiunikacyjnego w postaci sieci programów powinno spełniać podobną funkcję, co tzw. sieć czynności przy realizacji złożonych przedsięwzięć. Dążono przy tym, by wychodząc z formalnego opisu protokołu w miarę sprecyzować jakie funkcje, wynikające z jego formalnego opisu, mają realizować poszczególne programy. Przedstawiona w niniejszej pracy metodyka powinna się przyczynić do bardziej systematcznego podejścia do tworzenia oprogramowania protokołów komunikacyjnych.

1. WPROWADZENIE

W związku z burzliwym rozwojem sieci teleinformatycznych w latach siedemdziesiątych i stosowania w nich różnorodnego sprzętu telekomunikacyjnego i informatycznego zaistniała konieczność ujednolicenia opisu sterowania siecią niezależnie od jej topologii i sprzętu, na bazie którego są one budowane. Podjęcie próby kompleksowego opisu sterowania siecią teleinformatyczną doprowadziło do przyjęcia jednolitego modelu sieci mającego hierarchiczną strukturę warstwową, w którym dokonana została dekompozycja funkcji komunikacyjnych [13, 20, 22]. Model sieci teleinformatycznej powstały w wyniku takiej dekompozycji zwany jest *architekturą logiczną sieci teleinformatycznej.* Ilustrację idei architektury przedstawiono na rys. 1, gdzie dwa węzły sieci teleinformatycznej zostały przedstawione jako zespół warstw o strukturze hierarchicznej. W każdej z warstw zostały wyszczególnione tzw. człony komunikacyjne [3, 4, 7] realizujące ściśle określone funkcje związane ze sterowaniem. Dwa człony, np. $CK-\alpha_n$ oraz $CK-\beta_n$ komunikują się wza-

jemnie, co sprowadza się do wymiany ściśle określonych porcji informacji I'_n, I''_n, zwanych segmentami informacji poziomu n-tego. Komunikacja pomiędzy CK-α_n i $CK-\beta_n$ odbywa się za pośrednictwem podobnych członów $CK-\alpha_{n-1}$ i $CK-\beta_{n-1}$ warstwy funkcjonalnej poziomu $n-1$. Do sterowania przebiegiem procesu komunikowania się wymagane jest także przesyłanie pewnych rozkazów i generowanie na nie odpowiedzi. Ogólnie rozkazy i odpowiedzi oznaczone będą symbolami u, v, x, y z odpowiednimi indeksami i nazywane *informacją sterująco-kontrolną*.



Rys. 1. Ilustracja idei architektury logicznej sieci teleinformatycznej

Zachodzi konieczność określenia zasad komunikowania się członów $CK-\alpha_n$ i $CK-\beta_n$ należących do jednej warstwy, jak również zasad korzystania członów komunikacyjnych warstwy poziomu n-tego z usług komunikacyjnych warstwy poziomu $(n-1)$-szego. Sformułowanie tych zasad prowadzi do pojęcia *protokołu komunikacyjnego* (PK_n) warstwy poziomu n-tego. Ogólnie protokół komunikacyjny PK_n należy traktować jako język, którego semantyka określona jest znaczeniem definiowanych informacji sterująco-kontrolnych, a syntaktyka sposobem usytuowania informacji sterująco-kontrolnych podczas przebiegu komunikacji pomiędzy parą członów komunikacyjnych warstwy poziomu n-tego.

Jeżeli przez $R-\alpha_n$ i $R-\beta_n$ oznaczymy odpowiednio model $CK-\alpha_n$ i $CK-\beta_n$, to modelem protokołu komunikacyjnego warstwy poziomu n-tego jest zbiór

$$\mathbf{PK}_n \stackrel{\text{df}}{=} \{R-\alpha_n, R-\beta_n\}.$$

Wobec tego opracowanie modelu oprogramowania protokołu sprowadza się więc do opracowania modelu oprogramowania odpowiedniej pary członów komunikacyjnych.

W szczególności, jeżeli modele $R-\alpha_n$ i $R-\beta_n$ są identyczne, możemy mówić, że PK_n jest *protokołem symetrycznym*, np. *HDLC* w trybie *ARM* i przy konfiguracji zrównoważonej, a w przypadku przeciwnym — asymetrycznym, np. *HDLC* w trybie *NRM* [3, 13].

Węzeł sieci teleinformatycznej zbudowany na bazie systemu komputerowego jest złożonym systemem czasu rzeczywistego, w którym przebiega współbieżnie wiele różnorodnych procesów. Zachodzi przy tym dodatkowo konieczność uwzględnienia szeregu pewnych *funkcji nadzorczych* w oprogramowaniu protokołów komunikacyjnych. Wyróżniane będą w pracy następujące funkcje nadzorcze:

— zarządzanie zasobami systemu komputerowego i ich przydział dla przebiegających współbieżnie procesów komunikacyjnych,
— kontrola zdarzeń i spełniania się warunków logicznych (np. kontrola różnych odcinków czasowych, kontrola czasu czekania w kolejkach),
— przeprowadzanie automatycznego testowania,
— żądanie interwencji operatora w przapadkach pojawienia się negatywnego wyniku testowania automatycznego.

Określenie zbioru zarówno programów sterujących przesyłaniem informacji zgodnie z założonym protokołem, jak również programów realizujących wymienione funkcje nadzorcze oraz zasad współdziałania tych wszystkich programów jest przedmiotem niniejszej pracy.

Do opracowania modelu oprogramowania protokołu komunikacyjnego niezbędne jest przyjęcie określonego modelu samego protokołu. W literaturze znany jest szereg takich modeli zbudowanych między innymi na bazie: grafów [21], aparatu gramatyk formalnych [14], teorii automatów skończonych [10, 12, 14], modelu maszyny sekwencyjnej [14, 18], modelu komunikowania się pary interlokutorów [1, 15], sieci Petriego z uwzględnieniem czasu [11, 16, 17]. W wymienionych modelach, których syntetyczne ujęcie zawierają prace [20, 23], nie są uwzględniane aspekty tworzenia oprogramowania protokołów, w szczególności wymienione uprzednio funkcje nadzorcze. W żadnym ze znanych modeli nie uwzględnia się możliwości powstawania zniekształceń przesyłanej informacji sterująco-kontrolnej.

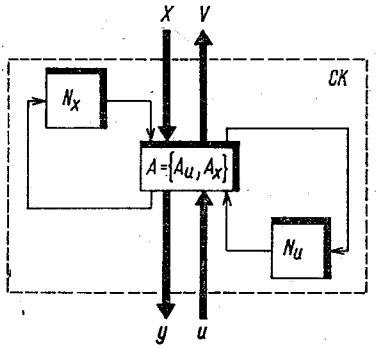
Przedstawiony w pracy wielowarstwowy model oprogramowania protokołu komunikacyjnego bazuje na zasadach tworzenia oprogramowania reguł sterowania informacji cyfrowych zawartych w pracach [5—9] i pozwala on zasady współdziałania programów (sterujących przebiegiem komunikacji i programów realizujących różnego typu funkcje nadzorcze) przedstawić w postaci pewnej sieci, zwanej *siecią programów protokołu komunikacyjnego* [7]. Sieć taka spełniać może analogiczną funkcję, jaką spełnia tzw. sieć czynności przy realizacji złożonych przedsięwzięć. Przedstawiona tu sieć programów zapewnia możliwość kontroli wszystkich parametrów protokołu oraz realizację samoczynnej diagnostyki. Należy na wstępie zaznaczyć, że pojęcie programu rozumiane jest tu bardzo ogólnie, tzn. p r o g r a m stanowi dowolny ciąg rozkazów sterujących wykonywaniem ściśle określonych czynności.

2. WARSTWOWY MODEL OPROGRAMOWANIA PROTOKOŁU

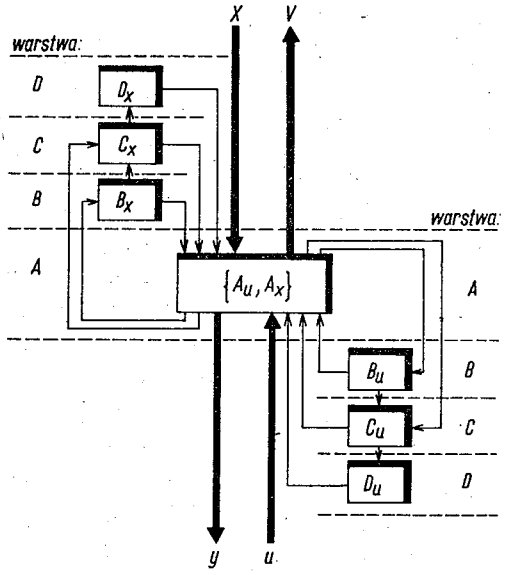
Ze względu na to, że w członie komunikacyjnym interpretowane są jedynie informacje sterująco-kontrolne, przeto traktowany będzie taki człon dalej jako pewien moduł, na którego wejściach pojawiają się asynchronicznie w czasie informacje sterująco-kontrolne

wejścioweoznaczane przez u i x, zaś człon ten generuje wyjściowe informacje sterująco-kontrolne oznaczane przez v i y. Informacje sterująco-kontrolne wejściowe będą interpretowane jako z g ł o s z e n i a żądające programowej obsługi. Z kolei v i y będą przyjmowane jako „wyniki" wykonywania odpowiednich programów obsługi zgłoszeń u i x.

Ogólną strukturę modelu oprogramowania członu komunikacyjnego przedstawiono na rys. 2. Wyszczególniono tu *regułę programowej obsługi zgłoszeń*, oznaczoną symbolem



Rys. 2. Ogólna struktura modelu oprogramowania członu komunikacyjnego



Rys. 3. Warstwowa struktura modelu oprogramowania członu komunikacyjnego

$\underline{A}$, oraz dwie reguły nadzorcze N_u i N_x. Regułę nadzorczą N_u tworzy zbiór odpowiednio ze sobą sprzężonych programów realizujących funkcje nadzorcze w procesie programowej obsługi zgłoszeń u. Analogicznie interpretujemy N_x.

Przyjmujemy dalej, że programowa obsługa zgłoszeń u i x realizowana jest na identycznych zasadach. Z tego względu w regule $\underline{A}$ wyróżniamy dwie reguły wewnętrzne A_u oraz A_x, z których A_u jest regułą programowej obsługi zgłoszeń u, zaś A_x — regułą programo-

wej obsługi zgłoszeń x. Zakładamy, że odpowiednia koordynacja reguł A_u i A_x jest zadaniem podsystemu koordynacji procesów systemu operacyjnego.

Pomiędzy wymienionymi funkcjami nadzorczymi w procesie programowego sterowania komunikacją według ustalonego protokołu istnieje ściśle określona hierarchia, którą przedstawia na rys. 3 warstwowa struktura modelu oprogramowania członu komunikacyjnego. W modelu tym wyróżniamy następujące warstwy oprogramowania:

A — *warstwa funkcjonalna*, którą tworzą wszystkie programy realizujące reguły A_u i A_x,

B — *warstwa kontroli zasobów i zdarzeń* (warstwa KZ i Z), którą tworzą wszystkie programy realizujące, oznaczone symbolami B_u, i B_x reguły kontroli zasobów i zdarzeń,

C — *warstwa automatycznego testowania*, którą tworzą wszystkie programy realizujące reguły automatycznego testowania C_u i C_x,

D — *warstwa interwencji operatora*, którą tworzą wszystkie programy realizujące reguły D_u i D_x.

Warstwy B, C i D nazywane będą dalej *warstwami nadzorczymi*.

Warstwę KZ i Z tworzą pewne programy kontroli zasobów systemu komputerowego oraz programy kontrolujące występowania zdarzeń będących rezultatem spełnienia się ściśle określonych warunków logicznych. W szczególności upłynięcie określonego odcinka czasu (*time out*) interpretowane jest jako spełnienie się pewnego warunku logicznego, np. licznik odciąga stan zerowy, a więc występuje pewne zdarzenie. Na występienie takiego zdarzenia mógą oczekiwać pewne procesy, w tym przypadku najczęściej powodujące retransmisję. Do warstwy KZ i Z oczywiście mają następować odwołania z warstwy funkcjonalnej, w takich sytuacjach, gdy proces przebiegający w warstwie funkcjonalnej żąda przydziału określonego zasobu, lub też proces nie może dalej przebiegać i musi oczekiwać na spełnienie się pewnych warunków, tj. zajście określonych zdarzeń. Jest oczywiste, że czas oczekiwania zarówno na zasoby jak i na warunki nie może być nieskończenie długi. Musi być on dla każdego procesu odpowiednio ograniczony. W sytuacji, gdy czas oczekiwania upływa, a np. zasoby nie są przydzielone, mamy sytuację niepożądaną, która powinna być zarejestrowana i w odpowiedni sposób ewentualnie sprawdzona. Z tego względu zakładamy, że w sytuacji, gdy pewne zgłoszenie oczekuje na zasoby lub zdarzenia oraz w zadanym okresie czasu zasoby te mogą być przydzielone lub oczekiwane zdarzenie nie występuje, następuje z warstwy KZ i Z odwołanie się do warstwy automatycznego testowania.

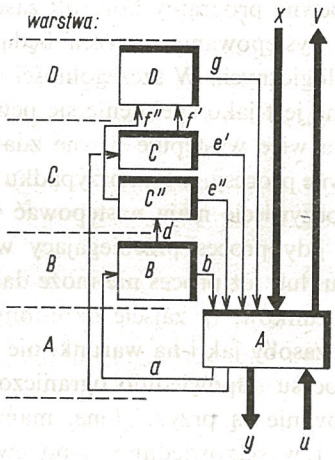
Warstwę automatycznego testowania tworzą programy, które w sposób automatyczny w zależności od aktualnego stanu dokonują testowania odpowiednich elementów (np. kanałów). Do warstwy tej może nastąpić odwołanie z omówionej uprzednio warstwy KZ i Z oraz z warstwy funkcjonalnej.

W procesie komunikacji mogą wystąpić takie sytuacje niepożądane, w których wymagana jest ostatecznie interwencja operatora. W strukturze przedstawionej na rys. 3 założono, że interwencja operatora wymagana jest w takiej sytuacji, w której wynik testowania automatycznego okazał się negatywny. Dalej przyjmuje się, że jeżeli z warstwy testowania następuje odwołanie się do warstwy interwencji operatora, to dla operatora przekazywana jest „historia" tej niepożądanej trwałej sytuacji. W szczególności może to być w odpowiedniej formie przekazany dla operatora stan (faza), w którym dana sytuacja niepożądana wystąpiła. Warstwę interwencji operatora tworzą więc wszystkie takie programy, które

w odpowiednio zakodowanej formie przedstawiają dla operatora uszkodzenia z ewnetualną dla niego wskazówką dotyczącą postępowania. Podejście to w pełni przypomina uruchamianie programów obliczeniowych (np. w FORTRANie). Uruchamiając program obliczeniowy na ogół na początku mamy wydruk różnych błędów z podaniem ich numerów, z kolei, sięgając odpowiedniego opisu danego systemu komputerowego, odczytujemy znaczenie tych błędów. W podanej tu strukturze oprogramowania nadzorczego istnieje możliwość podania operatorowi takiej „wskazówki", że nastąpiła określona trwała sytuacja niepożądana i ze wskazaniem czego ona może dotyczyć.

Ponieważ programowa obsługa zgłoszeń u i x realizowana jest na identycznych zasadach, przeto omawiać będziemy dalej programową obsługę zgłoszeń jednego typu pomijając przy tym indeks u lub x.

Współdziałanie reguły A z pozostałymi regułami z warstw nadzorczych ilustruje rys. 4, gdzie przyjęto $C = \{C', C''\}$. W przedstawionym tu schemacie poszczególne wejścia



Rys. 4. Ilustracja współdziałania poszczególnych warstw modelu oprogramowania

i wyjścia interpretujemy następująco:

a — odwołanie się do programu przydziału zasobów lub kontroli określonego zdarzenia,

b — zasób został przydzielony lub wystąpiło określone zdarzenie (został spełniony określony warunek),

c — odwołanie się do programu automatycznego testowania,

d — wymagane zasoby nie są przydzielone lub zdarzenie, na które oczekiwano, nie wystąpiło; wymagane jest testowanie,

e', e'' — wyniki automatycznego testowania są pozytywne,

f', f'' — wyniki automatycznego testowania są negatywne,

g — po zlikwidowaniu (trwałej) sytuacji niepożądanej ustawienie przez operatora odpowiednio określonego stanu spoczynkowego reguły A.

Realizację każdej z reguł nadzorczych oraz reguły A przedstawiać będziemy jako proces zmiany ściśle określonych stanów. Każda z tych reguł ma ściśle określony zbiór stanów

dopuszczalnych. Ogólnie współpracę wszystkich reguł podczas programowego sterowania przebiegiem komunikacji opisywać będziemy przez zmianę stanów ze zbioru

$$\mathbf{S}^* \stackrel{\mathrm{df}}{=} \mathbf{Q} \cup \mathbf{R} \cup \mathbf{S}' \cup \mathbf{S}'' \cup \mathbf{W}, \tag{2.1}$$

gdzie zbiór:

$\mathbf{Q}$ — jest podzbiorem stanów reguły A,

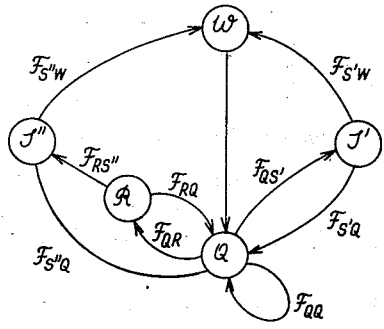
$\mathbf{R}$ — jest podzbiorem stanów reguły B,

$\mathbf{S}' \cup \mathbf{S}''$ — jest podzbiorem stanów reguły $C = \{C', C''\}$, przy czym $\mathbf{S}'$ dotyczy reguły C', zaś $\mathbf{S}''$ dotyczy C'',

$\mathbf{W}$ — jest podzbiorem stanów reguły D.

Traktując poszczególne podzbiory zbioru $\mathbf{S}^*$ jako pewne *stany zagregowane*, współdziałanie reguły A z oprogramowaniem nadzorczym według rys. 3 daje się opisać przy pomocy grafu przedstawionego na rys. 5, którego łuki reprezentują poszczególne relacje pomiędzy tak interpretowanymi stanami zagregowanymi.

Dalej przyjmować będziemy, że dla danego stanu ze zbioru $\mathbf{S}^*$ przyporządkowany jest jeden program, zgodnie z którym w danym stanie przebiega sterowanie komunikacją.



Rys. 5. Relacje pomiędzy poszczególnymi stanami zagregowanymi modelu oprogramowania

Tak przedstawione na rys. 5 relacje są w istocie relacjami pomiędzy programami spełniającymi różne funkcje. Z tego założenia oraz z funkcji realizowanych przezposzczególne warstwy nadzorcze wynika, że:

— relacja $\mathbf{F}_{QQ}$ określa współdziałanie programów w sytuacji, gdy nie są wymagane zasoby, nie oczekuje się na żadne zdarzenia oraz nie występują sytuacje niepożądane,

— relacja $\mathbf{F}_{QR}$ określa zasoby odwołania się z warstwy funkcjonalnej do warstwy KZ i Z,

— relacja $\mathbf{F}_{RQ}$ określa zasady „powrotu" z warstwy KZ i Z do warstwy funkcjonalnej,

— relacje $\mathbf{F}_{RS''}$, $\mathbf{F}_{QS'}$ określają sposób przekazywania sterowania do warstwy automatycznego testowania,

— relacja $\mathbf{F}_{S'Q}$ i $\mathbf{F}_{S''Q}$ określają sposób przekazania sterowania do warstwy funkcjonalnej przy pozytywnym wyniku automatycznego testowania,

— relacje $\mathbf{F}_{S'W}$, $\mathbf{F}_{S''W}$ opisują przy negatywnym wyniku automatycznego testowania, żądane interwencje operatora w zależności od stanu odpowiednio reguły A i B.

— relacja $\mathbf{F}_{WQ}$ określa stan spoczynkowy w warstwie funkcjonalnej ustawiany przez operatora.

Wyszczególnione relacje są podstawą formalnego opisu współdziałania poszczególnych warstw modelu oprogramowania protokołu.

3. FORMALNY OPIS WSPÓŁDZIAŁANIA WARSTW MODELU PROGRAMOWANIA PROTOKOŁU

Określimy teraz poszczególne relacje z grafu przedstawionego na rys. 5 tak, by można było je wykorzystać do przedstawienia zasad współdziałania programów przy sterowaniu przebiegiem komunikacji w postaci pewnej sieci, zwanej siecią programów protokołu komunikacyjnego [7]. W tym celu sprecyzujemy dokładniej przyjmowane dalej założenia. Zakładamy mianowicie, że:

Z1. Każdy stan zagregowany $\mathbf{Q}$, $\mathbf{R}$, $\mathbf{S}'$, $\mathbf{S}''$, $\mathbf{W}$ jest złożony z pewnej ściśle określonej liczby *stanów szczegółowych.*

Z2. Każdy stan szczegółowy ma przyporządkowany dokładnie jeden program, którego adres w pamięci zależy jednoznacznie poprzez pewne przyporządkowanie od danego stanu. Niech poszczególnym stanom $q \in \mathbf{Q}$ przyporządkowane będą adresy programów ze zbioru H_0 poprzez przyporządkowanie

$$H_0: \mathbf{Q} \to \mathbf{H}_0. \tag{3.1}$$

Do zbioru H_0 należą adresy $h_0(q)$ takich programów P_q, które w poszczególnych stanach $q \in \mathbf{Q}$ rozpoczynają programową obsługę zgłoszenia; programy te będą nazywane programami *wstępnej obsługi zgłoszeń.*

Z3. Każdy stan szczegółowy $s \in \mathbf{S}^*$ ma przyporządkowane na stałe określone rejestry, w których zapisywana jest informacja sterująco-kontrolna stanu s. Dalej konkretną postać tej informacji oznaczać będziemy przez $\xi(s)$, zaś zbiór jej możliwych postaci przez $\Xi(s)$.

Z4. Wszystkie sytuacje, jakie w stanie $q \in \mathbf{Q}$ po nadejściu zgłoszenia $\sigma \in \{u, x\}$ mogą wystąpić, są jednoznacznie ponumerowane przez funkcję klasyfikacyjną:

$$K(q): \Sigma x \Xi(q) \to K(q), \quad q \in \mathbf{Q}, \tag{3.2}$$

gdzie Σ jest zbiorem możliwych postaci σ.

Z5. Możliwe sytuacje w stanie $q \in \mathbf{Q}$ są podzielone na sytuacje pożądane, które są zgodne z protokołem komunikacyjnym i sytuacje niepożądane, czyli niezgodne z przyjętym protokołem. Niech

$$\mathbf{K}(q) = \mathbf{K}'(q) \cup \mathbf{K}''(q), \quad \mathbf{K}'(q) \cup \mathbf{K}''(q) = \emptyset, \tag{3.3}$$

gdzie $\mathbf{K}'(q)$ jest zbiorem numerów k'_q sytuacji pożądanych w stanie $q \in \mathbf{Q}$, zaś $\mathbf{K}''(q)$ zbiorem numerów k''_q sytuacji niepożądanych.

Z6. Niech $\mathbf{Q}(q) \subseteq \mathbf{Q}$ będzie podzbiorem stanów q bezpośrednio osiągalnych ze stanu $q \in \mathbf{Q}$, tj. takich wszystkich stanów q, że możliwe jest bezpośrednie przejście ze stanu q do stanu q', oznaczone symbolicznie jako $q \to q'$. Zbiorem $\mathbf{K}'(q)$ numerów sytuacji pożądanych w stanie q jest podzielony na następujący sposób:

$$\mathbf{K}'(q) = \bigcup_{q' \in \mathbf{Q}(q)} \mathbf{K}'(q, q'), \tag{3.4}$$

gdzie $K'(q, q') \cap K'(q, q'') = \emptyset$ dla $q \neq q''$.