

10. В. А. Повстень: *Применение теории синтеза электрических цепей для расчета параметров роторных стержней произвольной формы*. И. В. У. З. Электромеханика 1980, но. 10, с. 1022—1026
11. H. Pressler: Kettenleiter-Netzmodell für das elektromagnetische Feld in den Nuten elektrischer Maschinen. AfE, 1974, Bd. 56, S. 212—216
12. P. Rolicz: *Zastosowanie metody Budowa-Galerkina do obliczania impedancji przewodów*. Archiwum Elektrotechniki 1975, t. 24, s. 89—102
13. R. Sikora, W. Lipiński: *Dwuwymiarowe wypieranie prądu w żłobkach maszyn elektrycznych*. Archiwum Elektrotechniki. 1971, s. 467—478

M. DĄBROWSKI, J. KOŁOWROTKIEWICZ, A. MAŁECKI

NUMERICAL COMPUTING OF THE IMPEDANCE OF THE SQUIRREL-CAGE ROTOR BAR OF AN ELECTRIC MACHINE

Summary

The skin-effect in deep bars of arbitrary shapes of a squirrel-cage rotor, is considered. A differential equation for the general case of one-dimensional eddy-current density distribution is given. It is proven that, in the general case, an analytical solution of this equation does not exist. An algorithm for the numerical analysis of the current density distribution and also for the computation of the ratios of a.c. to d.c. resistance as well as for a.c. to d.c. inductance for different bar configurations is developed. On this basis, a computer program is developed and tested. As an example, the parameters for bars of several shapes are calculated.

M. DĄBROWSKI, J. KOŁOWROTKIEWICZ, A. MAŁECKI

CALCUL NUMÉRIQUE DE L'IMPEDANCE DE LA BARRE DE CAGE DE LA MACHINE ÉLECTRIQUE

Résumé

Cet ouvrage concerne l'influence de l'effet pelliculaire du courant sur la résistance et la réactance de la barre en forme quelconque, montée dans l'encoche d'une machine électrique. On y indique l'équation différentielle, exprimant la distribution de la densité de courant dans une barre, dont l'effet pelliculaire est unidirectionnel. On montre qu'en cas général, cette équation ne peut pas être résolue par un nombre fini d'intégrations. Dans cet ouvrage on présente l'algorithme pour calculer numériquement la distribution de la densité de courant dans une barre et aussi pour déterminer sa résistance, sa réactance et ses facteurs k_R et k_L , qui représentent l'influence de l'effet pelliculaire sur ces paramètres. En se basant sur cet algorithme on a élaboré le programme de calcul et puis on l'a vérifié en désignant les paramètres des barres de différents profils.

M. DĄBROWSKI, J. KOŁOWROTKIEWICZ, A. MAŁECKI

NUMERISCHE IMPEDANZBERECHNUNG DES KÄFIGWICKLUNGSTABES EINER ELEKTRISCHEN MASCHINE

Zusammenfassung

In der vorliegenden Bearbeitung wurde die Einwirkung der Stromverdrängung auf den Widerstand und die Reaktanz des Nutenstabes beliebiger Form untersucht. Die allgemeine Differentialgleichung der Stromdichte wurde unter Voraussetzung der eindimensionalen Stromverdrängung dargestellt. Es wurde nachgewiesen, daß eine analytische Lösung der Differentialgleichung für eine beliebig Stabform nicht

möglich ist. Der numerische Berechnungsalgorithmus wurde angegeben zur Bestimmung von: Stromdichteverteilung im Stab, Widerstand und Reaktanz, sowie Koeffizienten k_R und k_L . Das Berechnungsprogramm wurde überprüft, und die Parameter der Stäbe verschiedener Formen wurden berechnet.

М. ДОМБРОВСКИ, Е. КОЛОВРОТКЕВИЧ, А. МАЛЭЦКИ

ЧИСЛЕННЫЙ МЕТОД ОПРЕДЕЛЕНИЯ СОПРОТИВЛЕНИЯ СТЕЖНЯ КЛЕТКИ ЭЛЕКТРИЧЕСКОЙ МАШИНЫ

Р е з ю м е

Рассматривается влияние эффекта вытеснения тока на активное и реактивное сопротивление стержня произвольной конфигурации, расположенного в пазах электрической машины. Приведено дифференциальное уравнение тока в стержне с учётом одностороннего вытеснения. Доказано, что в общем случае это уравнение является неразрешимым аналитически. Представлен алгоритм цифрового вычисления плотности тока в стержне, его активного и реактивного сопротивления, а также коэффициентов k_R и k_L учитывающих влияние вытеснения тока на эти параметры. На основе этого алгоритма разработана вычислительная программа для ЭВМ и определены параметры стержней различных форм.

Some aspects of nonlinear functions simulation

STANISŁAW OSOWSKI, ANDRZEJ CICHOCKI, STEFAN FILIPOWICZ

*Instytut Elektrotechniki Teoretycznej i Miernictwa Elektrycznego,
Politechnika Warszawska*

Received 1985.03.20

Authorized 1985.11.14

The article concerns the method of simulation of the nonlinear transfer characteristics. On the basis of piecewise-linear approximation and Kang and Chua description of this approximation, the review of the hardware methods of simulation are given. Some new propositions of modelling based on application of piecewise-polynomial approximation are suggested. The results of experimental verifications of chosen circuits are inserted and compared with theoretical predictions.

1. INTRODUCTION

The synthesis of nonlinear resistive networks using linear elements and some chosen nonlinear units is a very important subject in applied circuit theory. The generation of nonlinear function finds many applications in simulation of the physical phenomena on the analog and hybrid computers, it is extensively used in measurement technique for the correction of nonlinear characteristics of the instruments as well as in automatic regulation technique for proper shaping of dynamical characteristics of the system.

Usually the synthesis task falls into two categories: synthesis of a driving point or transfer characteristic plot, although each of them can be replaced by the other, using very simple linear transformation circuits shown in Fig. 1a b [15]. Up to now the work

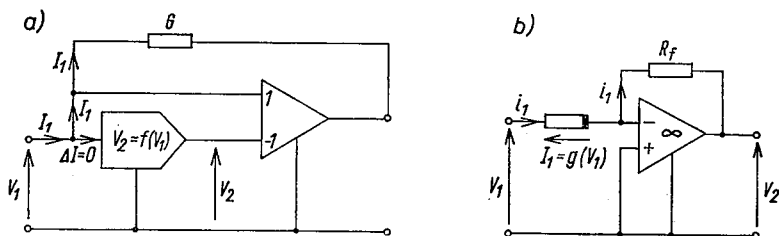


Fig. 1 Transformation of a) Voltage transfer characteristic $V_2 = f(V_1)$ into driving point characteristic $i_1 = Gf(V_1)$ b) driving point characteristic $i_1 = g(V_1)$ into voltage transfer characteristic $V_2 = -Rg(V_1)$

has been oriented toward the approximation of the characteristic plot and strictly associated with this synthesis technique. So two basic approaches arise, one using piecewise linear approximation [1, 2, 3, 4, 5, 10] in connection with the basic blocks used in practical realization of it, like convex and concave type characteristics, absolute value blocks etc., and the other involving rational function approximation, usually connected with analog four quadrant multipliers or linearly variable conductive or resistive elements [11, 12, 13, 14]. Many different design techniques as well as active circuit structures have been proposed for both approaches. Neither of them has found a general application with superiority to the other. For example, rational function approximation is very well suited for characteristic plots of very slow changes in slope, but jump discontinuities cannot be allowed. Moreover, using analog multiplier as the working horse, the technique becomes rather expensive in practical hardware implementation.

On the other hand piecewise linear technique allows any kind of characteristic plot, it is very simple and inexpensive in practical application. However to get the desired accuracy, the number of breakpoints of the characteristic must be very large which complicates the final circuit. Moreover, applying parallel connections of the basic blocks to realize each of the slopes, the sensitivity of the whole realization becomes critical.

One reasonable solution to this problem seems to be the combination of both methods. The most suitable approximation technique becomes in this case the piecewise polynomial approximation, introduced to the circuit theory by the authors of paper [9]. These functions allow to get any degree of smoothness and jump discontinuities while using in practical implementation small number of expensive analog multipliers.

This paper will discuss some chosen aspects of realization of nonlinear functions using piecewise linear and piecewise polynomial approximation of given characteristic plot, suggesting some new solution to older and new problems.

2. PIECEWISE LINEAR APPROXIMATION TECHNIQUE

Piecewise linear approximation approach to the synthesis of nonlinear characteristic plot has been developed in two main directions: indirect approach based on a transformation of an arbitrary type of nonlinear nonmonotonic characteristic to monotonic one, the latter being easily realised by using simple building blocks; and direct approach, based on the description of the given characteristic, using simple functions of an easy hardware implementation.

2.1. TRANSFORMATION METHODS

One of the first techniques applied for transformation of nonlinear characteristic was rotation or reflection, both in conjunction with scaling either in voltage or in current [1, 6]. Denoting current and voltage of the monotonic characteristic by i_m and V_m , respectively, while actually realised by i and V we could describe these type of transformation by

$$\begin{bmatrix} i \\ V \end{bmatrix} = KA_{\text{rot}} \begin{bmatrix} i_m \\ V_m \end{bmatrix} \quad (1)$$

and

$$\begin{bmatrix} i \\ V \end{bmatrix} = K A_{\text{refl}} \begin{bmatrix} i_m \\ V_m \end{bmatrix}, \quad (2)$$

where A_{rot} denotes the chain matrix describing the rotator of the angle of rotation equal to Θ and voltage-to-current coefficient equal to R

$$A_{\text{rot}} = \begin{bmatrix} \cos \Theta & -R \sin \Theta \\ \frac{1}{R} \sin \Theta & \cos \Theta \end{bmatrix}. \quad (3)$$

A_{refl} denotes the chain matrix describing the reflector of the reflection angle equal to Θ and voltage-to-current coefficient equal to R

$$A_{\text{refl}} = \begin{bmatrix} \cos 2\Theta & R \sin 2\Theta \\ \frac{1}{R} \sin 2\Theta & -\cos 2\Theta \end{bmatrix}, \quad (4)$$

while K describes the scalar either voltage or current, of

$$K_v = \text{diag}[K_v, 1] \quad (5)$$

$$K_i = \text{diag}[1, K_v]. \quad (6)$$

Practical implementation of equations (1) and (2) corresponds to the cascade connection of the scalar and rotator or reflector, loaded by nonlinear element of the monotonic characteristic $i_x = f_m(V_x)$. The driving point characteristic seen from the input port terminals corresponds to the realised nonlinear function $i = f_i(V)$ or $V = f_v(i)$.

Thanks to the easy realization of the mentioned 2-port networks, by using active elements like op-amps or controlled sources, this method of synthesis is very effective and was extensively used to generate different types of nonlinear characteristics [1, 6]. However, this approach is mainly oriented to the realization of the driving point characteristics. Its direct application to the transfer characteristics is not possible, unless a transformation circuit from Fig. 1 is applied. In the case of voltage transfer characteristics a new approach was developed, based on the linear transformation [5, 15]. Assume that nonmonotonic voltage-to-voltage characteristics are considered, which are controlled either by input voltage V_1 , that is, $V_2 = f(V_1)$ or by output voltage, that is $V_1 = \varphi(V_2)$. Let $V_y = f_m(V_x)$ and $V_x = \varphi_m(V_y)$ denote the monotonic characteristics. It was proved in [5] that the linear transformation

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \alpha_1 & \alpha_2 \end{bmatrix} \begin{bmatrix} V_x \\ V_y \end{bmatrix} \quad (7)$$

converts the monotonic characteristic $V_y = f_m(V_x)$ into the nonmonotonic one $V_2 = f(V_1)$, controlled by the input voltage V_1 . In a similar way the transformation of the type

$$\begin{bmatrix} V_x \\ V_y \end{bmatrix} = \begin{bmatrix} \beta_1 & \beta_2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad (8)$$

converts the monotonic characteristic $V_x = \varphi_m(V_y)$ into the nonmonotonic one $V_1 = \varphi(V_2)$ controlled by the output voltage V_2 . By choosing proper values of coefficients α_1, α_2 ,

or β_1, β_2 , every nonmonotonic characteristic can be converted into monotonic one with the same breakpoints and the slopes being changed according to the relations.:

$$S_k^{(m)} = \frac{S_k^{(n)} - \alpha_1}{\alpha_2}, \quad (9a)$$

where $S_{(k)}^n = \frac{dV_2}{dV_1}$ denotes the slope of the nonmonotonic characteristic between $(k-1)$ -th and k -th breakpoints, while $S_{(k)}^m = \frac{dV_y}{dV_x}$ denotes the slope of the monotonic one. Similarly

in the case of transformation (8) the nonmonotonic slope $D_k^n = \frac{dV_1}{dV_2}$ of the characteristic $V_1 = \varphi(V_2)$ between $(k-1)$ -th and k -th breakpoints is transformed to a monotonic slope $D_{(k)}^m = \frac{dV_x}{dV_y}$, according to the relation

$$D_{(k)}^m = \beta_1 D_{(k)}^n + \beta_2 \quad (9b)$$

without any change of the breakpoints.

The circuits corresponding to both types of transformations are presented in Fig. 2.

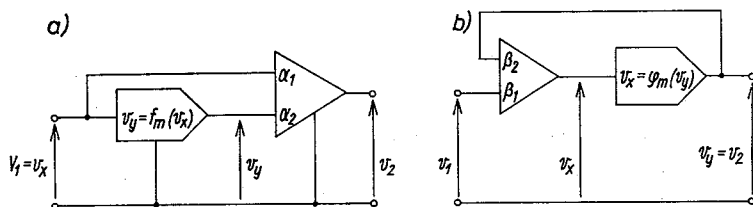


Fig. 2 Circuit realizations of a) Transformation (7) b) Transformation (8)

Both circuits were tested at simulation of different nonlinear, nonmonotonic characteristics of both N and S types. The authors found them especially useful for simulation of different complicated shapes of hysteresis loops with hysteresis of input or output voltage. Fig. 3 presents one of the typical application of the transformation (7).

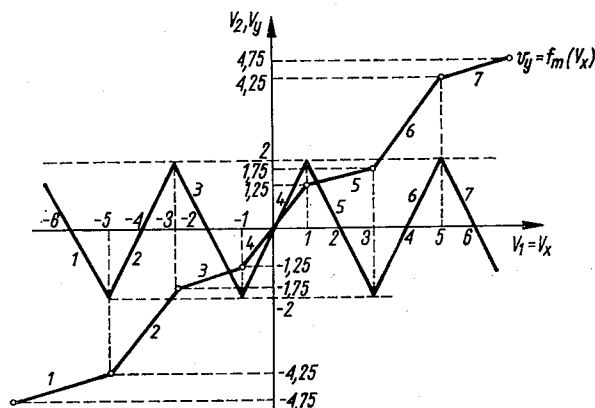


Fig. 3 Theoretical nonmonotonic (a) and monotonic (b) curves under simulation

The curve (a) in this figure shows a nonlinear nonmonotonic characteristic of the frequency multiplying converter. Applying linear transformation (7) of $\alpha_1 = -3$ and $\alpha_2 = 4$ the monotonic characteristic $V_y = f_m(V_x)$ was obtained (curve b of Fig. 3).

The circuit realising this nonmonotonic characteristic on the basis of given monotonic $V_y = f_m(V_x)$ is presented in Fig. 4a [15]. It was built in laboratory with monotonic function

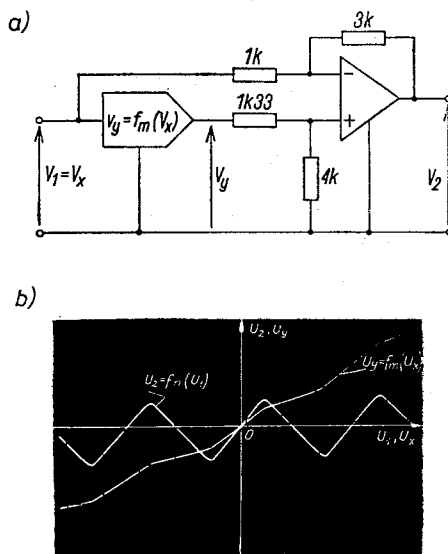


Fig. 4 a) Circuit simulation of nonmonotonic characteristic using monotonic one b) Oscillograms of both characteristics

being simulated by using diodes and resistors. The obtained oscillograms of both monotonic and nonmonotonic functions are given in Fig. 4b. The experimental results confirm the usefulness of the proposed method of simulation.

2.2 DIRECT METHOD OF REALIZATION

Direct realization method of piecewise-linear function makes use of decomposition of given function into sum of simple, basic standard functions being easily implemented in practice. Such standard functions include absolute value modules $|X - X_i|$, concave type characteristics $f_{\text{con}} = \frac{1}{2} [|X - X_i| + (X - X_i)]$, jump functions

$$1(X - X_i) = \begin{cases} 0 & \cup \text{sign}(X - X_i) = -1 \\ 1 & \cup \text{sign}(X - X_i) = 1 \end{cases}$$

or signum function

$$\text{sign} X = \begin{cases} -1 & \cup X < 0 \\ 1 & \cup X > 0 \end{cases}$$

Assuming for simplicity that nonlinear characteristic under consideration begins at $X = a$ it can be proved that any nonmonotonic characteristic $f(X)$ can be decomposed

into the sum of concave type functions of either positive or negative slopes and jump functions,

$$f(X) = f(a) + \sum_{i=1}^n \alpha_i f_{\text{conc}}(X - X_i) + \sum_{i=1}^n \beta_i 1(X - X_i) \quad (9)$$

where jump functions appear only for noncontinuous characteristics. In such a case

$$\beta_i = f(X_i^+) - f(X_i^-) \quad (10)$$

The values of coefficients α_i are described by the differences of slopes S_i of two adjacent sections, that is,

$$\alpha_i = S_{i+1} - S_i \quad (11)$$

In a similar way the arbitrary nonlinear characteristic is described by using other basic standard functions, for example absolute blocks. According to Chua and Kang [10] this description may be given in the form

$$f(X) = a_0 + a_1 X + \sum_{i=1}^n \alpha_i |X - X_i| + \sum_{i=1}^n \beta_i \text{sign}(X - X_i) \quad (12)$$

where

$$a_1 = \frac{1}{2} [S_0 + S_n]; \quad \alpha_i = \frac{1}{2} [S_i - S_{i-1}]; \quad \beta_i = \frac{1}{2} [f(X_i^+) - f(X_i^-)]$$

$$a_0 = f(0) - \sum_{i=1}^n \alpha_i |X_i| - \sum_{i=1}^n \beta_i \text{sign}(X_i)$$

and S_i denotes the slope of the section between i th and $(i+1)$ st breakpoints, while S_0 and S_n are the slopes of the sections placed on the left and right limit of the piecewise linear characteristic, respectively. It is worth noting that if function $f(X)$ is a parametric function of the parameter p , that is, $f(X, p)$, all coefficients $\alpha_i, \beta_i, a_0, a_1$ as well as breakpoints X_i are also functions of parameter p . In this case by changing parameter p a whole family of characteristics can be generated [4, 10]. The above given relations have been generalized to a multidimensional function, of an arbitrary number of variables. Chua and Kang developed a new representation for N -dimensional piecewise linear functions that are affine over a convex polyhedral region bounded by linear partitions. Introducing vector notations for input variables $X_i, X = [X_1 X_2, \dots, X_N]^T$, coefficients $a_1 = [a_{11}, a_{12}, \dots, a_{1N}]$ and $\alpha_i = [\alpha_{i1}, \alpha_{i2}, \dots, \alpha_{iN}]$ the general expression describing any continuous single multidimensional piecewise linear function is given as follows

$$f(X) = a_0 + a_1 x + \sum_{i=1}^L g_i |\alpha_i x - \beta_i| \quad (13)$$

where a_0, g_i, β_i are real numbers and L the number of breakline borders between linear regions.

The values of α_{ik} and β_i are given explicitly by the linear equation of the borders between the linear regions while the values of g and vector a_1 could be obtained after solving a set of linear equations (14) following from the gradient of the function $f(X)$ in each region, where the gradient of $f(X)$ is given in the form

$$\nabla f(x) = a_1 + \sum_{i=1}^L g_i \alpha_i \text{sign}(\alpha_i x - \beta_i) \quad (14)$$

Finally after computing a_1 and g_i the remaining constant a_0 is obtained as follows

$$a_0 = f(0) - \sum_{i=1}^L g_i |\beta_i| \quad (15)$$

The usefulness of description (13) was proved in practice for simulation of three-variable function $f(x, y, z)$ by the authors in [4]. The circuit realizing this equation is shown in Fig. 5, where both polarities of variables X and switches were needed to adjust

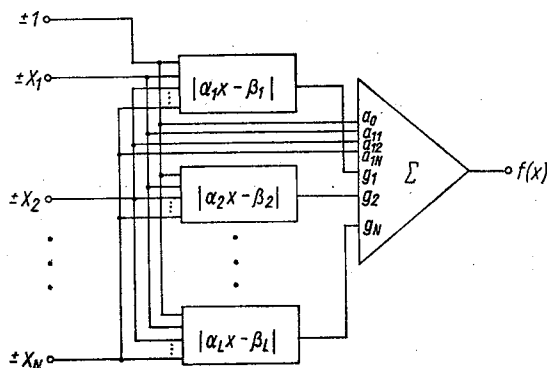


Fig. 5 The circuit representation of the equation (13)

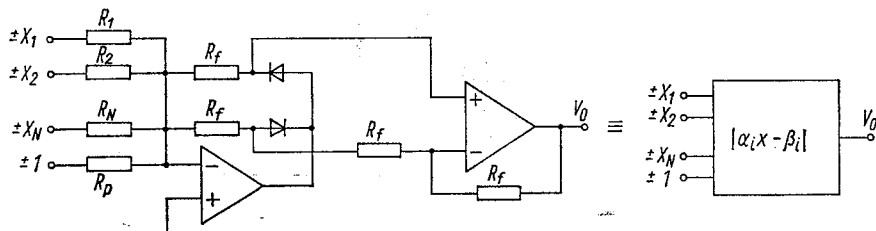


Fig. 6 The realization of vector block absolute value module using two op-smps

proper sign of each of the variables in the vector block absolute value modules $|\alpha_i X - \beta_i|$. The simplest realization of the vector block absolute value module by using only two operational amplifiers is presented in Fig. 6, where the values of α_i are set by the ratios of R_f and R_i , $\alpha_i = \frac{R_f}{R_i}$ and β_i are described by $\beta_i = \frac{R_f}{R_p}$.

The expression (13) was further generalized to be used for the nonlinear functions of an implicit form $F(Z, X) = 0$, where Z represents output variable of the circuit. Defining function

$$f(Z, X) = a_0 + a_1 x + \sum_{i=1}^L g_i |\alpha_i x + b_i Z - \beta_i| = Z \quad (16)$$

we can state that

$$F(Z, X) = f(Z, X) - Z = 0. \quad (17)$$

The implicit relation $F(Z, X) = 0$ was solved practically by the use of a nonlinear function $F(Z, X)$ in a high gain ($A_1 \rightarrow \infty$) negative feedback circuit, presented in Fig. 7.

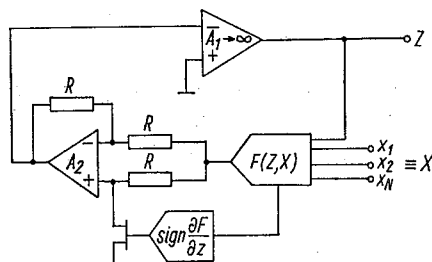


Fig. 7 The circuit realizing implicit relation $F(Z, X) = 0$

The additional circuit on op-amp A_2 is the inverting (if $\text{sign } \frac{\partial F}{\partial z} > 0$) or noninverting (if $\text{sign } \frac{\partial F}{\partial z} < 0$) unity gain block added to the circuit to provide the net incremental

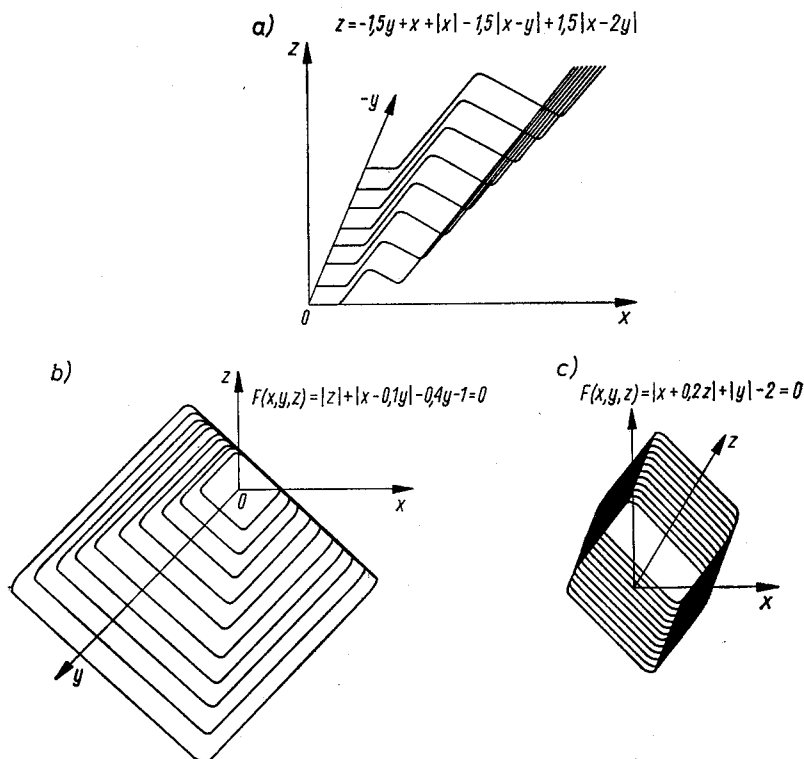


Fig. 8 The plots of nonlinear functions (19), (20) and (21)

negative feedback, needed to ensure stability of the implicit computation loop. Experimental verifications of the relations presented above for three dimensional explicit and implicit functions performed in the laboratory of the Institute of Theory of Electrical Engineering and Electrical Measurements confirmed the coincidence of the obtained results with theory. Plots of some simulated functions

$$Z = -1.5Y + X + |X| - 1.5|X - Y| + 1.5|X - 2Y| \quad (19)$$

$$F(Z, X) = |Z| + |X - 0.1Y| - 0.4Y - 1 = 0 \quad (20)$$

and

$$F(Z, X) = |X + 0.2Z| + |Y| - 2 = 0 \quad (21)$$

obtained in the laboratory experiments are presented in Fig. 8a, b and c, respectively. It should be noted at this point that piecewise linear approximation is the simplest approximation of nonlinear function. The approximation error can be made as small as we wish by making ΔX_k sufficiently small for all k . However, this practice increases the number of basic blocks needed in hardware implementation, complicates the final electronic circuit and raises the costs of its manufacture. For these reasons it seems that application of piecewise, polynomial approximation, instead of piecewise linear, can produce approximation to $f(X)$ with significantly smaller number of breakpoints by keeping the same accuracy of approximation. In the next chapter we will consider this type of approximation limiting ourselves to cubic polynomials only.

3. PIECEWISE POLYNOMIAL APPROXIMATION TECHNIQUE

Let us introduce first some closed form formulas for representing a plot of a given function $f(X)$ which belongs to either cubic Hermite or cubic Spline family of piecewise polynomials. Variable X can represent either voltage or current of the circuit.

Let $[a, b]$ be an interval of the real variable X and let Δ denotes the discrete set of $(N+1)$ points in $[a, b]$ defined by the partition

$$a = X_1 < X_2 < X_3 < \dots < X_{N+1} = b \quad (22)$$

We wish to construct a piecewise-cubic polynomial function $g_3(X)$ which interpolates $f(X)$ at these points. That means that on each interval $[X_i, X_{i+1}]$ we construct $g_3(X)$ as a certain cubic polynomial $P_i(X)$ ($i = 1, 2, \dots, N$), where

$$P_i(X) = C_{0i} + C_{1i}(X - X_i) + C_{2i}(X - X_i)^2 + C_{3i}(X - X_i)^3 \quad (23)$$

Since we want $g_3(X_i) = f(X_i)$ for $i = 1, 2, \dots, N+1$, we must have

$$P_i(X_i) = f(X_i) \quad (24)$$

$$P_i(X_{i+1}) = f(X_{i+1}) \quad (25)$$

and

$$P_{i-1}(X_i) = P_i(X_i) \quad (26)$$

for $i = 2, 3, \dots, N$ so that $g_3(X)$ is guaranted to be continuous on $[a, b]$. In piecewise-cubic Hermite interpolation two additional conditions for the interpolating function are

so that also

$$P'_i(X_i) = f'(X_i) \quad (27)$$

and

$$P'_i(X_{i+1}) = f'(X_{i+1}). \quad (28)$$

Using abbreviations $f_i = f(X_i)$, $S_i = f'(X_i)$ for $i = 1, 2, \dots, N+1$ we get then [8]

$$C_{0,i} = f_i \quad (29)$$

$$C_{1,i} = S_i \quad (30)$$

$$C_{2,i} = \frac{f[X_i, X_{i+1}] - S_i}{\Delta X_i} - C_{3,i} \Delta X_i \quad (31)$$

$$C_{3,i} = \frac{S_{i+1} + S_i - 2f[X_i, X_{i+1}]}{(\Delta X_i)^2}, \quad (32)$$

where

$$\Delta X_i = X_{i+1} - X_i$$

$$f[X_i, X_{i+1}] = \frac{f(X_{i+1}) - f(X_i)}{\Delta X_i}.$$

Piecewise-cubic interpolation requires knowledge of $f'(X)$. If it is impossible in practice, we can use some reasonable approximation to $f'(X_i)$. In cubic Bessel interpolation one uses

$$S_i = \frac{\Delta X_{i-1} f[X_i, X_{i+1}] + \Delta X_i f[X_{i-1}, X_i]}{\Delta X_{i-1} - \Delta X_i} \quad (33)$$

instead of $S_i = f'(X_i)$ but proceeds otherwise as before determining the coefficients C_{ji} for the cubic pieces by (29)—(32).

In spline interpolation one uses the knowledge, that $g_3(X)$ is twice continuously differentiable, which is equivalent to

$$P''_{i-1}(X_i) = P''_i(X_i) \quad (34)$$

or with (23)

$$2C_{2,i-1} + 6C_{3,i-1}\Delta X_{i-1} = 2C_{2,i} \quad (35)$$

for $i = 2, 3, \dots, N$. Applying (31) and (32) to (35) we express C_{3i} and C_{2i} in terms of f and unknown S_j , getting finally the following system of $(N-1)$ equations with $(N+1)$ unknown variables $S_1, S_2, \dots, S_{N+1}$, $\Delta X_i S_{i-1} + 2(\Delta X_{i-1} + \Delta X_i) S_i + \Delta X_{i-1} S_{i+1} = 3f[X_{i-1}, X_i] \Delta X_i + 3f[X_i, X_{i+1}] \Delta X_{i-1}$. Choosing S_1 and S_{N+1} somehow (for example by applying (33)) and applying Gauss elimination we get all needed values of S , and then from (29)—(32) the coefficients of polynomial C_{0i} , C_{1i} , C_{2i} and C_{3i} .

The circuit simulation of the piecewise polynomial describing nonlinear characteristic is based on the simulation of the relation (23) in each subinterval of $[a, b]$.

To minimize the number of components needed, we present $P_i(X)$ in the form.

$$P_i(X) = C_{0i} + C_{1i}(X - X_i) + (X - X_i)C_{2i}(X - X_i) + C_{2i}(X - X_i)^2 \frac{C_{3i}}{C_{2i}} (X - X_i) \quad (37)$$

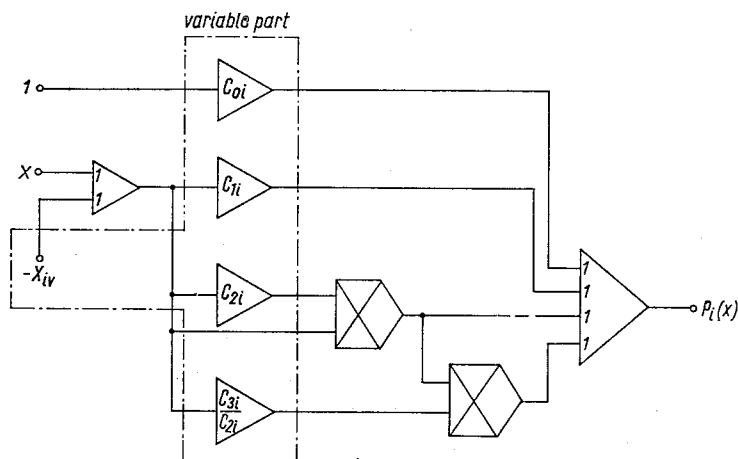


Fig. 9 The general circuit for simulation piecewise-polynomial function $f(X)$

The circuit simulating this relation is presented in Fig. 9. This circuit is composed of two main parts: the fixed elements changing neither their structure nor their values, and the variable part, which includes changing values of X_i and contains amplifiers C_i of the gains fixed in one interval, say (X_i, X_{i+1}) but being changed by the change of this interval.

The value X_{iv} as well as all C_i , should be changed automatically with change of X . To satisfy this requirement we have to build comparator circuit comparing actual value of X with some given X_i . The simplest circuit fulfilling this function built by using op-amp is shown in Fig. 10a.

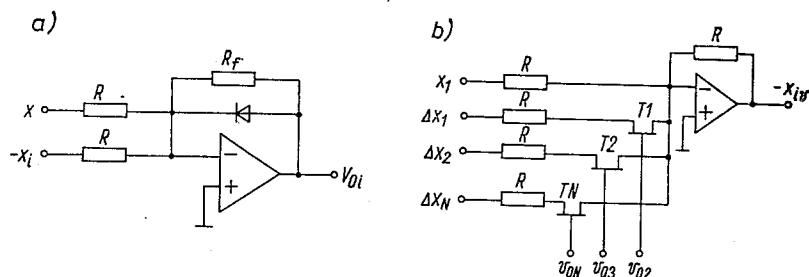


Fig. 10 a) Simple comparator circuit b) The generator of staircase function $-X_{iv}$

For $X < X_i$ diode D is conducting and the voltage $X_{oi} = 0$; for $X > X_i$ the output voltage is equal $V_{oi} = -\frac{R_f}{R}(X - X_i)$. By choosing sufficiently large ratio $\frac{R_f}{R}$, voltage V_{oi} is equal to saturation voltage of the amplifier immediately after $X = X_i$. This voltage, V_{oi} , generated by the comparator is used as the controlling voltage of the transistor switches in the circuit generating proper value of X_{iv} (Fig. 10b). For $X < X_1 + \Delta X_1 = X_2$ all voltages V_{oi} ($i = 2, 3, \dots, N$) in circuit of Fig. 10a are zero and all transistor switches are off. The output signal $-X_{iv}$ of the summing circuit is $-X_1$, that is, the lower limit of X . When $X \geq X_2 = X_1 + \Delta X_1$, the output signal $V_{oi} = V_{02}$ of Fig. 10a changes its value from

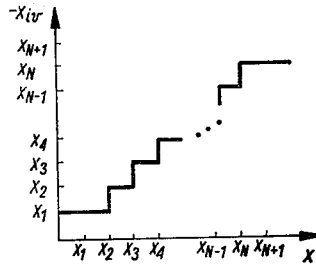
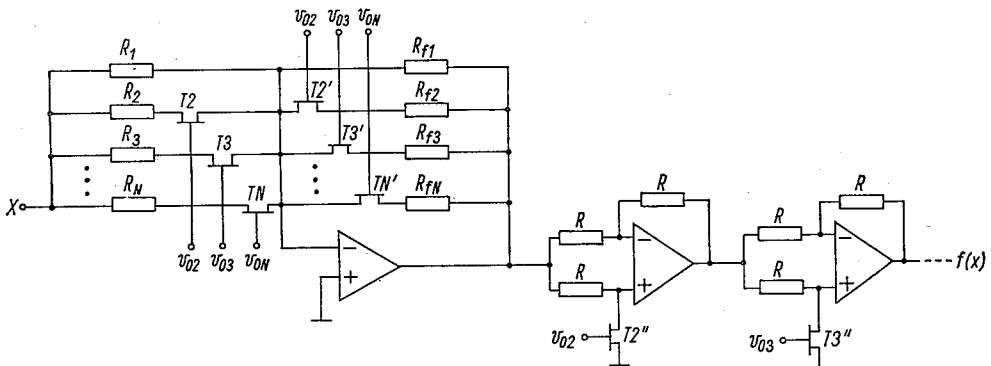


Fig. 11 The staircase function generated by circuit of Fig. 10b

0 to saturation voltage of op-amp, which turns on the transistor $T1$, increasing the value of $-X_{iv}$ to $-(X_1 + \Delta X_1)$. In this way the circuit of Fig. 10b generates the staircase signal $-X_{iv}$ (Fig. 11) needed in circuit of Fig. 9. To produce the gains of amplifier C_i dependent on the range of (X_i, X_{i+1}) we should be able to change both magnitude and sign of the

Fig. 12 The amplifier circuit of variable gain C_i

gain, automatically with X . Fig. 12 illustrates the general diagram of an amplifier circuit with the gain adjusted by the changing value of X .

The transistor switches T_i or T'_i , controlled by the output V_{0i} of Fig. 10a, connect adjusted resistances parallel to either R_{fi} or R_i ($i = 1, 2, \dots, N$) to simulate proper value of gain in each interval (X_i, X_{i+1}) . In this way the gain of the amplifier A_1 in the interval X_i, X_{i+1} is described by

$$k_i = - \left[G_1 + \sum_{k=2}^i G_k \right] / \left[G_{f1} + \sum_{k=2}^i G_{fk} \right] \quad (38)$$

Where $\sum_{k=2}^i G_k$ is zero if $i < 2$. To change the sign of the gain (if needed) the special circuits on op-amps $A_2 \div A_N$ may be added. At the lack of signal V_{0i} , the amplification factor of the section is $+1$, at given signal $V_{0i} \neq 0$, transistor is on, and the gain is changed to -1 .

4. CONCLUSIONS

Two different approaches to the simulation of nonlinear transfer characteristics $f(X)$ have been discussed in the paper. Very well known piecewise linear approximation presents simplicity and not expensive hardware for realization of the function. However the error introduced by the approximation is proportional to ΔX_k^2 , where ΔX_k denotes the subinterval $[X_k, X_{k+1}]$ of X . It is known [8] that for all $X \in [a, b]$ the maximum error is equal

$$\varepsilon_{\max} = \frac{1}{8} \max_k [\Delta X_k]^2 \max_{X_k < \xi < X_{k+1}} |f''(\xi)| < \max_{\xi \in [a, b]} \frac{1}{8} [f''(\xi) (\max_k \Delta X_k)^2] \quad (39)$$

provided that given nonlinear function $f(X)$ is twice differentiable on $[a, b]$. Note that to decrease the error of approximation we have to increase the number of breakpoints $X_2, \dots, \dots, X_N$. The increasing number of breakpoints complicates the final circuit and makes the circuit more sensitive to element variations. The authors see the cure of this in application of piecewise polynomial approximation instead of piecewise-linear. The maximum error in cubic polynomial approximation is proportional to $(\Delta X)^4$ and is much smaller than in piecewise linear. According to [8] in piecewise cubic Hermite interpolation this error is given by

$$\varepsilon_{\max} = \frac{1}{384} \max_k [(\Delta X)^4 \max_{X_k < \xi < X_{k+1}} f^{(4)}(\xi)] < \frac{1}{384} \max_{a < \xi < b} [f^{(4)}(\xi) \max_k \Delta X_k^4], \quad (40)$$

while in cubic spline interpolation by

$$\varepsilon_{\max} = \frac{5}{384} \max_{\xi \in [a, b]} [f^{(4)}(\xi) 5 (\max_k \Delta X_k)^4]. \quad (41)$$

The circuit implementation of piecewise polynomial approximation is only a bit more complicated, using transistor switches instead of diodes. The sensitive points of this realisation are comparators needed to recognise the interval of X . The circuit is cheap since it needs only two analog multipliers, for practical implementation. Although there only piecewise-cubic approximation was discussed, any other kind of piecewise polynomial approximation may be applied using the same approach. Discontinuities of the characteristics present no problem in this method.

REFERENCES

1. L. O. Chua, *Introduction to Nonlinear Network Theory*. Mc Graw Hill, 1969
2. L. O. Chua, S. Wong: *Synthesis of piecewise-linear networks*. IEE J. ECS, 1978, July, pp. 102—108
3. L. O. Chua, S. M. Kang: *A global representation of multidimensional piecewise linear functions with linear partitions*. IEEE Trans. CAS, 1978, November, pp. 938—940
4. A. Cichocki, S. Filipowicz, S. Osowski: *The multidimensional piecewise linear function simulation*. Proc. AMSE Conf. Modelling & Simulation, Paris 1982
5. A. Cichocki, S. Filipowicz, S. Osowski: *Unified structure for the 2-port network realization*. Electronics Letters, 1977, vol. 13, pp. 328—329
6. A. Cichocki, S. Filipowicz, S. Osowski: *Analog simulation of the nonlinear transfer characteristics—a novel approach*. IEE Journal on ECS, 1978, vol. 2., pp. 26—28

7. D. Sheingold: *Nonlinear Circuit Handbook, Analog Devices*, Norwood, 1976
8. S.D. Conte, C. de Boor: *Elementary Numerical Analysis*. Mc Graw Hill, 1981
9. A. Rueda, J. Huertas, A. Rodriguez-Vazquez: *Basic circuit structures for the synthesis of piecewise polynomial one-port resistors*. IEE J. ECS, 1985
10. L.O. Chua, S.M. Kang: *Sectionwise piecewise linear functions, canonical representation, properties and applications*. Proc. IEEE, vol. 65, 1977, June, pp. 915—929
11. N.R. Malik, G.L. Jackson, Y.S. Kim: *Theory and applications of resistor and linear controlled resistor networks*. IEEE Trans. CAS, 1976, vol. 23, pp. 222—229
12. J. Huertas, J. Acha, A. Gago: *Design of general voltage or current controlled resistive elements*. IEEE Trans. CAS, 1980, vol. 27, pp. 92—104
13. S. Osowski: *Nonlinear resistive network synthesis*. IEE J. ECS, 1983, vol. 130, pp. 15—19
14. A. Cichocki: *Synteza układów nieliniowych przy użyciu wzmacniaczy operacyjnych i elementów sterowanych*. Zesz. Nauk. Elektryka, 1982 r
15. A. Cichocki, S. Filipowicz, A. Osowski: *Zastosowanie transformacji liniowych w syntezie charakterystyk nieliniowych*. Archiwum Elektrotechniki, 1978, t XXVII, pp. 891—904

S. OSOWSKI, A. CICHOCKI, S. FILIPOWICZ

WYBRANE ZAGADNIENIA SYMULACJI CHARAKTERYSTYK NIELINIOWYCH

Streszczenie

Praca stanowi przegląd metod symulacji nieliniowych charakterystyk rezystancyjnych. W oparciu o metodę odcinkowo liniowej aproksymacji charakterystyk, ze szczególnym uwzględnieniem formuły Kanga i Chua, przedstawiono pewne nowe rozwiązania zagadnienia syntezy układów nieliniowych o cechach aplikacyjnych. Podano przykłady liczbowe ilustrujące rozważania teoretyczne. W części końcowej pracy przedstawiono wyniki badań laboratoryjnych wybranych układów nieliniowych. Stwierdzono zgodność uzyskanych wyników eksperymentalnych z teoretycznymi.

S. OSOWSKI, A. CICHOCKI, S. FILIPOWICZ

LES PROBLÈMES CHOISIS DE LA SYMULATION DES CARACTÉRISTIQUES NON-LINÉAIRES

Résumé

L'article concerne des méthodes de la simulation des caractéristiques transitoires. En se basant sur la méthode d'approximation linéaire par segments des caractéristiques non-linéaires et en appliquant la formule de Chua et Kang on établit une nouvelle solution du problème de synthèse des systèmes non-linéaires du point de vue de l'utilisation pratique. En outre on présente des exemples concrets de calcul pour illustrer les considérations théoriques. Dans la partie finale de l'article on donne des résultats des essais pratiques opérés sur des systèmes particuliers. On constate la concordance entre les résultats expérimentaux et théoriques.

S. OSOWSKI, A. CICHOCKI, S. FILIPOWICZ

AUSGEWÄHLTE PROBLEME DER SIMULATION VON NICHTLINEAREN KENNLINIEN

Zusammenfassung

Der Beitrag ist ein Überblick der Simulation methoden von nichtlinearen Übergangskennlinien. Gründend an die Aproximation mit strecken-lineare Methode, unter speziellen Besehtung von Chua und Kang Formel, einige neue Lösungen der Synthese von nichtlinearen Systems zur unmittelbaren Applikation

wurden vorgestellt. Die Zahlenbeispiele illustrieren die theoretische Überlegungen. Am Ende des Beitrags die experimentelle Ergebnisse der Untersuchung von einigen Schaltungen wurden vorgestellt. Eine gute Übereinstimmung von erzielten Resultaten und theoretischen Überlegungen wurde fortgestellt.

С. ОСОВСКИ, А. ЦИХОВСКИ, С. ФИЛИПОВИЧ

ИЗБРАННЫЕ ВОПРОСЫ СИМУЛЯЦИИ НЕЛИНЕЙНЫХ СИСТЕМ

Р е з ю м е

Приведен обзор методов симуляции нелинейных переходных характеристик. Применяя метод кусочно-линейной аппроксимации характеристик, с учетом формул Хулы и Канга, представлены некоторые новые решения вопросов синтеза нелинейных систем, характеризующиеся практической применимостью. Поданы числовые примеры иллюстрирующие результаты лабораторных исследований избранных систем. Констатируется хорошее совпадение экспериментальных результатов с теоретическими.

Straty mocy i moment napędowy silnika indukcyjnego o symetrii kulistej

JAN PURCZYŃSKI, LESZEK KASZYCKI

Wydział Elektryczny, Politechnika Szczecińska

Otrzymano 1986.06.20

Autoryzowano do druku 1987.07.30

W pracy wyznaczono straty mocy oraz moment napędowy silnika indukcyjnego o symetrii kulistej. Do analizy wykorzystano funkcję prądową oraz potencjał skalarny pola magnetycznego. Wyprowadzone zależności umożliwiają prowadzenie analizy dla dowolnego rozmieszczenia uzwojeń stojana.

W wyniku analizy numerycznej określono wpływ poślizgu, grubości warstwy przewodzącej, wielkości szczeliny powietrznej oraz kąta rozchyłu uzwojeń stojana na wartości strat mocy i momentu napędowego.

1. WSTĘP

W ostatnich latach pojawiło się szereg nowoczesnych konstrukcji przetworników elektromechanicznych o symetrii kulistej, umożliwiających ruch o dwóch kątowych stopniach swobody. Głównymi dziedzinami zastosowań tych urządzeń są: nawigacja wykorzystująca urządzenia inercyjne [3, 9] (żyroskopy z kulistym wirnikiem) oraz robotyka [2, 5, 9] (silniki kuliste o ruchu złożonym). W przypadku układów żyroskopowych (żyroskop kriogeniczny, elektrostatyczny, kulożyroskop z ferromagnetycznym wirnikiem) wykorzystuje się wirujące pole elektromagnetyczne wytwarzane przez uzwojenia dwu i trójfazowe do napędu wirnika w początkowej fazie pracy urządzenia. W silnikach o ruchu złożonym uzwojenia wytwarzają wzajemnie prostopadłe pola wirujące umożliwiające ruch twornika o dwóch kątowych stopniach swobody [8, 9].

W związku z tym istnieje zainteresowanie problematyką analizy wirującego pola magnetycznego oraz momentu obrotowego i strat mocy żyroskopów i silników elektrycznych w symetrii kulistej, czego wyrazem są prace [5, 7, 8, 10].

W niniejszej pracy dokonano analizy pola elektromagnetycznego, a następnie wyznaczono straty mocy i moment napędowy dla konkretnego rozwiązania konstrukcyjnego silnika. Silnik kulisty składa się z kulistego twornika oraz sferycznego stojana. Twornik wykonany ze stali jest wydrążony wewnątrz oraz pokryty od zewnątrz cienką warstwą miedzi. Na powierzchni stojana rozmieszczone są dwa rodzaje uzwojeń podstawowe i pomocnicze. Uzwojenie podstawowe powodujące ruch obrotowy ułożone jest biegunowo,

natomiast, uzwojenie pomocnicze służące do odchyłania osi obrotu rozmieszczone jest równoleżnikowo. Odnosnie uzwojenia podstawowego czyni się następujące założenia upraszczające:

1. uzwojenie wykonane jest w postaci okładu prądowego (odcinka warstwy)
2. uzwojenie rozciąga się od bieguna do bieguna (w rzeczywistości uzwojenia jak i stojan są ograniczone w kierunku horyzontalnym dla umożliwienia przenoszenia napędu).

Odnosnie stojana, przyjmuje się przenikalność magnetyczną $\mu = \infty$ oraz konduktywność $\gamma = 0$ (stojan wykonany jest z blach elektrotechnicznych). W przypadku twornika zakłada się $\mu = \infty$ uwzględniając natomiast konduktywność warstwy miedzi na jego powierzchni.

W pracy ograniczono się do analizy pola elektromagnetycznego wytworzonego przez uzwojenie podstawowe. Prezentowana analiza stosuje się do żyroskopów ze sferycznym wirnikiem, przykładowo do kulożyroskopu Sperry.

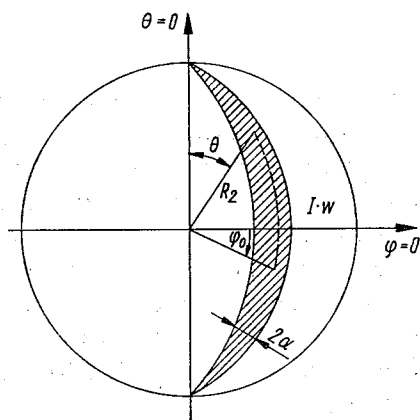
2. ANALIZA POLA ELEKTROMAGNETYCZNEGO

Elementarna sekcja uzwojenia podstawowego przedstawiona jest na rys. 1. Gęstość liniowa prądu płynącego w sekcji uzwojenia dana jest wzorem

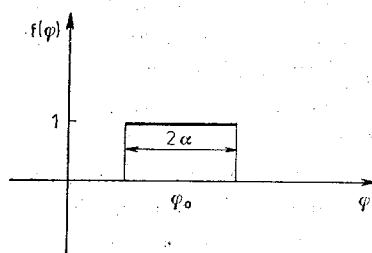
$$K = \frac{I_m \cdot w}{l} = \frac{I_m \cdot w}{2\alpha R_2 \sin\Theta}, \quad (1)$$

gdzie: l — zmienna szerokość uzwojenia,

$I_m \cdot w$ — przepływ jednej sekcji.



Rys. 1. Elementarna sekcja uzwojenia stojana



Rys. 2. Gęstość prądu elementarnej sekcji uzwojenia jako funkcja współrzędnej φ

Traktując gęstość prądu jako funkcję współrzędnych Θ i φ otrzymuje się:

$$K(\Theta, \varphi) = \frac{K_0}{\sin\Theta} \cdot f(\varphi), \quad (2)$$

gdzie $K_0 = \frac{I_m \cdot w}{2\alpha R_2}$.

Przebieg funkcji $f(\varphi)$ obrazuje rys. 2

Funkcję $K(\theta, \varphi)$ rozwija się w szereg harmonik sferycznych [1, 6] o postaci:

$$K(\theta, \varphi) = \sum_{n=1}^{\infty} \sum_{m=1}^n P_n^m(\cos\theta) (\alpha_{mn} \cdot \cos m\varphi + \beta_{mn} \cdot \sin m\varphi) \quad (3)$$

W wyniku obliczeń otrzymuje się:

$$K(\theta, \varphi) = \frac{K_0}{\pi \sin\theta} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{2n+1}{m} C_{mn} \sin(m\alpha) \cdot \frac{(n-m)!}{(n+m)!} \cdot P_n^m(\cos\theta) \cdot \cos m(\varphi - \varphi_0). \quad (4)$$

Współczynnik C_{mn} w wyrażeniu (4) wyraża się wzorem [6, 10] $C_{mn} = \int_0^\pi P_n^m(\cos\theta) \cdot \sin \theta d\theta$ i jego wartość wynosi (dodatek A):

$$C_{mn} = 2 \cdot (m)!! \sum_{l=0}^{\left[\frac{n-m}{2}\right]} \frac{(-1)^l \cdot (2n-2l-1)!!}{(n-2l+1)!!(n-m-2l)!!(2l)!!} \cdot m, n \text{ parzyste}$$

$$C_{mn} = \pi(m)!! \sum_{l=0}^{\left[\frac{m-n}{2}\right]} (-1)^{l+1} \cdot \frac{(2n-2l-1)!!}{(n-2l+1)!!(n-m-2l)!!(2l)!!} \cdot m, n \text{ nieparzyste}$$

$$C_{mn} = 0; \quad m - \text{parzyste}, n - \text{nieparzyste}$$

$$C_{mn} = 0; \quad m - \text{nieparzyste}, n - \text{parzyste}$$

Przy czym $0!! = 1$, $k!! = 1.3.5, \dots k - \text{nieparzyste}$
 $2.4.6, \dots k - \text{parzyste}$

Korzystając z wzorów Eulera zależność (4) można przedstawić następująco:

$$K(\theta, \varphi) = \frac{K_0}{2\pi \sin\theta} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{(2n+1)}{m} \cdot C_{mn} \cdot \sin(m\alpha) \cdot \frac{(n-m)!}{(n+m)!} P_n^m(\cos\theta) \cdot [e^{+jm(\varphi-\varphi_0)} + e^{-jm(\varphi-\varphi_0)}]. \quad (5)$$

Wzór (5) przedstawia rozwinięcie w szereg harmonik sferycznych elementarnej sekcji uzwojenia zasilanej prądem jednofazowym. Uzwojenie stojana składa się z parzystej liczby takich elementarnych sekcji, zasilanych wielofazowo. Stosując zasadę superpozycji, na podstawie wzoru (5) można znaleźć wyrażenie na zespoloną amplitudę gęstości prądu płynącego w uzwojeniu podstawowym. Na podstawie (5) wypadkową gęstość prądu można przedstawić zależnością:

$$K(\theta, \varphi) = \frac{K_0}{2 \sin\theta} \cdot \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{2n+1}{m} \cdot \sin(m\alpha) \cdot \frac{(n-m)!}{(n+m)!} \cdot C_{mn} \cdot P_n^m(\cos\theta) \cdot [K_m e^{jm(\varphi-\varphi_0)} + L_m \cdot e^{-jm(\varphi-\varphi_0)}], \quad (6)$$

gdzie K_m, L_m są w ogólności zespolonymi współczynnikami zależnymi od liczby, sposobu rozmieszczenia i zasilania elementarnych sekcji uzwojeń.

Wzór (6) określa zespoloną wartość amplitudy gęstości prądu. Natomiast zależność czasowo-przestrzenną gęstości prądu wyznacza się ze wzoru

$$K(\Theta, \varphi, t) = \text{Im}[K(\Theta, \varphi) \cdot e^{j\omega t}].$$

Wykorzystując ostatnią zależność oraz wzór (6), otrzymuje się następującą postać czasowo-przestrzenną gęstości prądu:

$$K(\Theta, \varphi, t) = \frac{K_0}{2\sin\Theta} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{2n+1}{m} \cdot \sin(m\alpha) \cdot \frac{(n-m)!}{(n+m)!} \cdot P_n^m(\cos\Theta) \cdot C_{nm} \cdot \{K_m e^{j[\omega t + m(\varphi - \varphi_0)]} + L_m e^{j[\omega t - m(\varphi - \varphi_0)]}\}. \quad (6a)$$

Oznacza to, że uzwojenie wytwarza pole wirujące, przy czym współczynnik L_m związany jest z harmonicznymi wirującymi zgodnie z harmoniczną podstawową, natomiast współczynnik K_m związany jest z harmonicznymi wirującymi przeciwnie. W dalszej części pracy zostanie podany sposób wyznaczania współczynników K_m i L_m , na przykładzie konkretnego uzwojenia.

Znajomość rozwinięcia w szereg harmonik gęstości prądu płynącego przez uzwojenie stojana (wzór (6)) umożliwia zastosowanie metody rozdzielania zmiennych do analizy pola magnetycznego od wzbudzenia oraz pola prądów indukowanych na powierzchni twornika w układzie współrzędnych kulistych. Z uwagi na to, że pole elektromagnetyczne wytworzone przez uzwojenie podstawowe jest polem trójwymiarowym, zdecydowano się na użycie do jego analizy potencjału skalarnego pola magnetycznego U określonego równaniem

$$H = -\text{grad } U = -\frac{\partial U}{\partial r} \cdot \vec{1}_r - \frac{1}{r\sin\Theta} \frac{\partial U}{\partial \varphi} \cdot \vec{1}_\varphi + \frac{1}{r} \frac{\partial U}{\partial \Theta} \cdot \vec{1}_\Theta. \quad (7)$$

Potencjał U spełnia równanie Laplace'a [6].

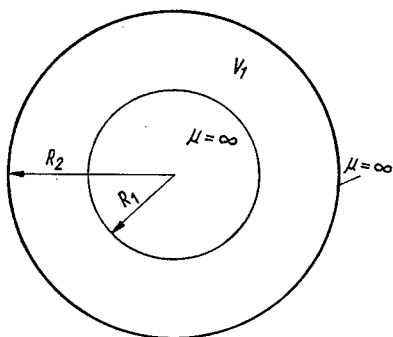
$$U = \frac{\partial^2 U}{\partial r^2} + \frac{2}{r} \frac{\partial U}{\partial r} + \frac{1}{r^2} \frac{\partial^2 U}{\partial \Theta^2} + \frac{\text{ctg}\Theta}{r^2} \frac{\partial U}{\partial \Theta} + \frac{1}{r^2 \sin\Theta} \frac{\partial^2 U}{\partial \varphi^2} = 0, \quad (8)$$

którego rozwiązanie posiada ogólną postać [1, 6]

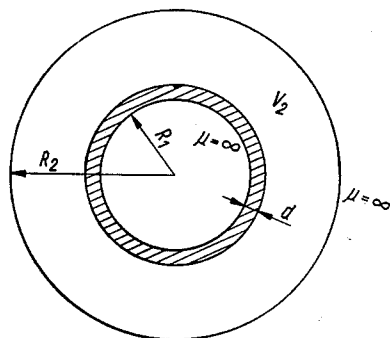
$$U = \sum_{n=1}^{\infty} \sum_{m=1}^n \left(a_n r^n + \frac{b_n}{r^{n+1}} \right) \cdot P_n^m(\cos\Theta) \cdot (\alpha_m \cos m\varphi + \beta_m \sin m\varphi). \quad (9)$$

Na wstępie rozpatrywane jest pole od wzbudzenia określone potencjałem U_1 . Na rys. 3 przedstawiono przekrój równikowy modelu silnika służący do wyznaczania pola magnetycznego od wzbudzenia $\vec{H}_1$ opisanego potencjałem U_1 . Jest on zgodny z przyjętymi na wstępie założeniami upraszczającymi. Uwzględniając postać wzoru (6) (wymuszenia) i wzór (9) przewiduje się następującą postać potencjału U_1 :

$$U_1 = \sum_{n=1}^{\infty} \sum_{m=1}^n \left(r^n + \frac{a_n}{r^{n+1}} \right) \cdot b_{nm} (e^{jm(\varphi - \varphi_0)} + c_m e^{-jm(\varphi - \varphi_0)}) \cdot P_n^m(\cos\Theta) \quad (10)$$



Rys. 3. Model obliczeniowy silnika służący do wyznaczania pola magnetycznego od wzbudzenia



Rys. 4. Model obliczeniowy silnika służący do wyznaczania pola magnetycznego prądów zaindukowanych na powierzchni twornika

Z warunków brzegowych

$$H_{1\varphi} \Big|_{r=R_1} = - \frac{1}{r \sin \Theta} \frac{\partial U_1}{\partial \varphi} \Big|_{r=R_1} = 0 \quad (11)$$

$$H_{1\varphi} \Big|_{r=R_2} = K(\Theta, \varphi)$$

wyznacza się stałe a_n , b_{nm} , c_m .

Ze wzorów (2), (6), (10) i (11) otrzymuje się

$$U_1 = \frac{I_m \cdot w}{4\pi \cdot \alpha R_2} \sum_{n=1}^{\infty} \sum_{m=1}^n \left(r^n - \frac{R_1^{2n+1}}{r^{n+1}} \right) \frac{(2n+1) C_{mn} \sin m\alpha R_1^{(1-n)} (n-m)!}{(1 - \mathcal{L}^{2n+1}) m^2 (n+m)!} \cdot$$

$$j[K_m e^{jm(\varphi - \varphi_0)} - L_m e^{-jm(\varphi - \varphi_0)}] \cdot P_n^m(\cos \Theta), \quad (12)$$

gdzie $\mathcal{L} = \frac{R_1}{R_2}$.

Ferromagnetyczny twornik pokryty jest cienką warstwą przewodzącą (np. miedź), w której indukują się prądy wirowe. Prądy te wytwarzają pole magnetyczne, którego potencjał oznaczono przez U_2 . Całkowity potencjał magnetyczny wynosi $U = U_1 + U_2$. W celu wyznaczenia potencjału U_2 przyjmuje się model obliczeniowy silnika pokazany na rys. 4. Na podstawie wzorów (6) i (9) przewiduje się następującą postać potencjału U_2 .

$$U_2 = \sum_{n=1}^{\infty} \sum_{m=1}^n \left(r^n + \frac{b_n}{r^{n+1}} \right) \cdot d_{mn} \cdot (e^{jm(\varphi - \varphi_0)} + e_m e^{-jm(\varphi - \varphi_0)}) P_n^m(\cos \Theta). \quad (13)$$

Poszukiwany potencjał U_2 spełnia warunek brzegowy:

$$- \frac{1}{R_2 \sin \Theta} \frac{\partial U_2}{\partial \varphi} \Big|_{r=R_2} = 0. \quad (14)$$

Uwzględniając (14) wyznacza się stałą b_n i wzór (13) przyjmuje postać

$$U_2 = \sum_{n=1}^{\infty} \sum_{m=1}^n \left(r^n - \frac{R_2^{2n+1}}{r^{n+1}} \right) \cdot d_{mn} [e^{jm(\varphi - \varphi_0)} + e_m \cdot e^{-jm(\varphi - \varphi_0)}] \cdot P_n^m(\cos \Theta). \quad (15)$$

Rozpływ prądów wirowych w powłoce miedzianej twornika opisuje funkcja prądowa I , zdefiniowana równaniem [4]

$$\mathbf{K} = -\vec{n} \times \text{grad} I, \quad (16)$$

gdzie $\mathbf{K}$ jest wektorem liniowej gęstości prądów wirowych indukowanych na powierzchni twornika, $\vec{n}$ — wersor normalny do powierzchni twornika.

Funkcja I spełnia równanie [4]

$$I = \Delta U |r = R_1 = U_2 |r = R_1. \quad (17)$$

Stąd i ze wzoru (15) otrzymuje się

$$I = \sum_{n=1}^{\infty} \sum_{m=1}^n \left(R_1^n - \frac{R_2^{2n+1}}{R_1^{n+1}} \right) \cdot d_{mn} \cdot P_n^m(\cos \Theta) \cdot [e^{jm(\varphi - \varphi_0)} + e_m \cdot e^{-jm(\varphi - \varphi_0)}] \quad (18)$$

Współczynniki d_{mn} i e_m zostaną wyznaczone w dalszej części pracy.

Dla obracającego się z prędkością $\vec{v}$ twornika w polu magnetycznym o indukcji $\mathbf{B}$, gęstość powierzchniowa prądów indukowanych na jego powierzchni wyraża się wzorem

$$\vec{K} = \sigma \cdot [\vec{E} + \vec{v} \times \vec{B}], \quad (19)$$

gdzie: $\sigma = \gamma \cdot d$, d — grubość warstwy przewodzącej.

Korzystając z zależności

$$\text{rot}_r E = - \frac{\partial B_r}{\partial t}$$

równanie (19) przyjmuje postać:

$$\text{rot}_r \vec{K} = \sigma \cdot \left[- \frac{\partial B_r}{\partial t} + \text{rot}_r (\vec{v} \times \vec{B}) \right]. \quad (20)$$

Wypadkowe pole magnetyczne $\vec{H}$ jest superpozycją pola od wzbudzenia $\vec{H}_1$ i pola prądów indukowanych $\vec{H}_2$:

$$\vec{H} = \vec{H}_1 + \vec{H}_2, \quad (21)$$

gdzie:

$$\begin{aligned} \vec{H}_1 &= - \text{grad} V_1 \\ \vec{H}_2 &= - \text{grad} V_2. \end{aligned}$$

Wykorzystując zależność (20) i (21) równanie (16) w układzie współrzędnych kulistych przyjmuje następującą postać:

$$\begin{aligned} \frac{\partial^2 I}{\partial \Theta^2} + \text{ctg} \Theta \cdot \frac{\partial I}{\partial \Theta} + \frac{1}{\sin^2 \Theta} \cdot \frac{\partial^2 I}{\partial \varphi^2} &= \sigma \mu R_1^2 \left[\frac{\partial}{\partial t} \left(H_{1r} - \frac{\partial U_2}{\partial r} \right) + \right. \\ &\left. + \omega_w \cdot \frac{\partial}{\partial \varphi} \left(H_{1r} - \frac{\partial U_2}{\partial r} \right) \right] \Big|_{r=R_1}, \end{aligned} \quad (22)$$

gdzie ω_w jest prędkością kątową wirowania twornika.

Rozwiązując równanie (22) wyznacza się nieokreślone współczynniki d_{mn} i e_m występujące

we wzorach (15) i (18). Następnie wyznacza się z zależności (18) prądową funkcję I . Do rozwiązania równania (22) wykorzystuje się zależności (12), (15), (18) oraz fakt, że funkcja

$$Y = [e^{jm(\varphi-\varphi_0)} + e^{-jm(\varphi-\varphi_0)}] \cdot P_n^m(\cos\Theta)$$

spełnia równanie różniczkowe [11]

$$\frac{\partial^2 Y}{\partial \Theta^2} + \operatorname{ctg} \Theta \cdot \frac{\partial Y}{\partial \Theta} + \frac{1}{\sin^2 \Theta} \cdot \frac{\partial^2 Y}{\partial \varphi^2} + (n+1) \cdot n = 0. \quad (23)$$

W efekcie przeprowadzonych obliczeń otrzymano następującą postać funkcji I

$$I = -\frac{I_m \cdot w}{4 \cdot \pi \cdot \alpha} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{(2n+1)^2 \cdot C_{mn} \cdot \sin m\alpha \cdot (n-m)!}{m^2 \cdot (n+m)!} \times \\ \times \left[\frac{b_s \cdot K_m \cdot e^{jm(\varphi-\varphi_0)}}{jb_s(n \cdot \mathcal{L}^{2n+1} + n+1) + n(n+1)(1 - \mathcal{L}^{2n+1})} - \right. \\ \left. - \frac{b_{-s} L_m e^{-jm(\varphi-\varphi_0)}}{jb_{-s}(n \mathcal{L}^{2n+1} + n+1) + n(n+1) \cdot (1 - \mathcal{L}^{2n+1})} \right] \cdot P_n^m(\cos\Theta),$$

gdzie:

$$b_s = \omega \mu \sigma R_1 [1 + m(1-s)]$$

$$b_{-s} = \omega \mu \sigma R_1 [1 - m(1-s)]$$

$$s = \frac{\omega - \omega_w}{\omega}.$$

Wzór (24) opisuje rozptyw prądów wirowych w powłoce miedzianej obracającego się twornika. Zależność (24) pozwala wyznaczyć ze wzoru (16) składowe wektora gęstości powierzchniowej prądów indukowanych.

3. WYZNACZANIE STRAT MOCY I MOMENTU NAPĘDOWEGO

W wyniku założenia zblachowanego stojana, zasadnicze straty mocy wystąpią w tworniku, tzn. w warstwie przewodzącej. Wielkość tych strat określa wzór

$$P = \frac{1}{2\sigma} \iint_s |K|^2 ds, \quad (25)$$

s — powierzchnia twornika.

Uwzględniając zależność (16) oraz wzór Greena, ostatnia zależność przyjmuje postać

$$P = -\frac{1}{2\sigma} \iint_s I^* \cdot \Delta I ds \quad (26)$$

gdzie I określa wzór (8).

W trakcie obliczeń strat mocy korzysta się z zależności (18), (23), (24) oraz ortogonalności funkcji $e^{jm\varphi}$ oraz $P_n^m(\cos\Theta)$

Własność ortogonalności opisana jest wzorami [1]

$$\int_0^{2\pi} e^{-jm\varphi} \cdot e^{jk\varphi} d\varphi = \begin{cases} 0 & k \neq m \\ 2\pi & k = m \end{cases} \quad (27a)$$

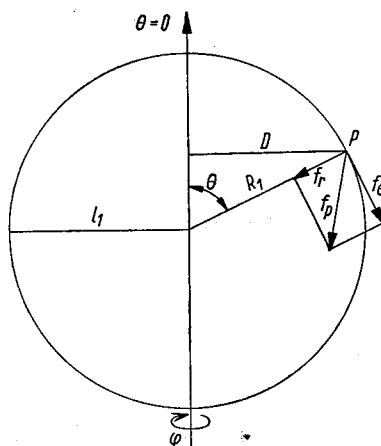
$$\int_0^{2\pi} P_n^m(\cos\Theta) \cdot P_L^m(\cos\Theta) \sin\Theta d\Theta = \begin{cases} 0 & \text{dla } l \neq n \\ \frac{2}{2n+1} \frac{(n+m)!}{(n-m)!} & \text{dla } l = n. \end{cases} \quad (27b)$$

W wyniku przeprowadzonych obliczeń otrzymano następującą zależność na straty mocy:

$$P = \frac{I_m \cdot w}{8\pi\alpha^2 \cdot \gamma \cdot d} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{n(n+1)(2n+1)^3 \sin^2(m\alpha) C m n^2 \mathcal{L}^{2n}(n-m)!}{m^4(n+m)!} \times$$

$$\times \left\{ \frac{b_s^2 \cdot |K_m|^2}{[n(n+1)(1-\mathcal{L}^{2n+1})]^2 + b_s^2(n\mathcal{L}^{2n+1} + n+1)^2} + \right.$$

$$\left. + \frac{b_{-s}^2 |L_m|^2}{[n(n+1)(1-\mathcal{L}^{2n+1})]^2 + b_{-s}^2(n\mathcal{L}^{2n+1} + n+1)^2} \right\} \quad (28)$$



Rys. 5. Rozkład sił działających na twornik silnika kulistego

Na rysunku 5 przedstawiono rozkład sił działających na obracający się twornik silnika kulistego. Wyznacza się moment napędowy powodujący obrót twornika względem osi $\Theta = 0$

Gęstość objętościowa siły działająca w punkcie P na powierzchni twornika wyraża się wzorem

$$\vec{f} = \vec{K} \times \vec{B} \quad (29)$$

Za obrót twornika względem osi $\Theta = 0$ odpowiedzialna jest składowa f_φ , która wyraża się zależnością

$$\vec{f}_\varphi = -K_\Theta \cdot B_{1r} \cdot \vec{1}_\varphi \quad (30)$$

gdzie: B_{1r} oznacza składową normalną wektora indukcji magnetycznej wytworzoną przez uzwojenie podstawowe, a K_θ składową wektora gęstości powierzchniowej prądu indukowanego na powierzchni twornika.

Moment elementarny m_0 działający na punkt P wyraża się wzorem

$$\vec{m}_0 = D \cdot \vec{f}_\varphi$$

Uwzględniając relację: $D = R_1 \sin\theta$ (rys. 5) otrzymuje się wzór określający całkowity moment M_0

$$M_0 = -\frac{1}{2} R_1^3 \cdot \operatorname{Re} \left\{ \int_0^\pi \int_0^{2\pi} B_{1r} \cdot K_\theta^* \sin^2\theta d\varphi d\theta \right\}. \quad (31)$$

Stosując wzory

$$B_{1r} = \mu_0 \frac{\partial U_1}{\partial r}$$

$$K_\theta = -\frac{1}{r \sin\theta} \frac{\partial I}{\partial \varphi}.$$

Na podstawie zależności (12), (24) i (31) wyznacza się wzór określający moment M_0 działający na twornik

$$M_0 = \frac{\mu_0 (I_m \cdot w)^2 R_1}{8\pi\alpha^2} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{n(n+1) \cdot (2n+1)^3 \cdot \mathcal{L}^{2n} \cdot \sin^2(\alpha m) \cdot C_{mn}^2 \cdot (n-m)!}{m^3 \cdot (n-m)!} \times$$

$$\times \left\{ \frac{b_s |K_m|^2}{[n(n+1)(1 - \mathcal{L}^{2n+1})]^2 + b_s^2 [n\mathcal{L}^{2n+1} + n+1]^2} - \right. \quad (32)$$

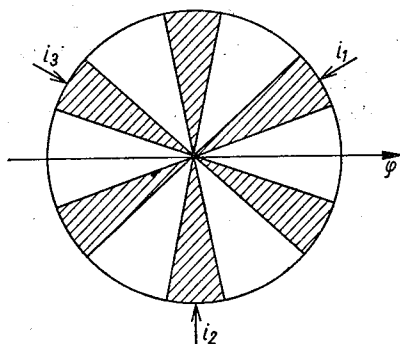
$$\left. - \frac{b_{-s} |L_m|^2}{[n(n+1)(1 - \mathcal{L}^{2n+1})]^2 + b_{-s}^2 [n\mathcal{L}^{2n+1} + n+1]^2} \right\}.$$

Wzory (28) i (32) określają straty mocy i moment w przyjętym modelu silnika kulstego w zależności od wymiarów geometrycznych układu, parametrów materiałowych i częstotliwości zasilania. Należy podkreślić, że zależności (28) i (32) zostały wyprowadzone przy poczynieniu dość ogólnych założeń odnośnie rozmieszczenia i zasilania elementarnych sekcji uzwojeń stojana. W celu zastosowania wzorów (28) i (32), do konkretnej realizacji uzwojenia stojana, należy zastosować zasadę superpozycji do wzoru (5) i wyznaczyć wstępnie współczynniki K_m i L_m . Zostanie to pokazane na przykładzie obliczeniowym w p. 4.

4. PRZYKŁAD OBLICZENIOWY

Jako przykład obliczeniowy przyjmuje się silnik, którego uzwojenie składa się z sześciu sekcji elementarnych przesuniętych o kąt $\frac{\pi}{3}$ i zasilanych prądem trójfazowym.

Na rys. 6 pokazano widok sekcji od strony bieguna z zaznaczeniem poszczególnych faz i kierunków prądu.



Rys. 6. Model uzwojenia podstawowego

Zgodnie z zasadą superpozycji, ze wzoru (5) otrzymuje się następującą postać zespolonej amplitudy gęstości prądu (rys. 6)

$$K(\Theta, \varphi) = \frac{K_0}{2\pi \sin \Theta} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{(2n+1)}{m} \frac{\sin m\alpha (n-m)!}{(n+m)!} \cdot P_n^m(\cos \Theta) \cdot$$

$$\cdot \left\{ e^{jm(\varphi - \varphi_0)} + e^{-jm(\varphi - \varphi_0)} - e^{jm(\varphi - \varphi_0 + \pi)} - e^{-jm(\varphi - \varphi_0 + \pi)} + \right.$$

$$+ e^{-j\frac{2}{3}\pi} \left[-e^{jm(\varphi - \varphi_0 + \frac{\pi}{3})} - e^{-jm(\varphi - \varphi_0 + \frac{\pi}{3})} + e^{jm(\varphi - \varphi_0 + \frac{4}{3}\pi)} + e^{-jm(\varphi - \varphi_0 + \frac{4}{3}\pi)} \right] +$$

$$\left. + e^{-j\frac{4}{3}\pi} \left[e^{jm(\varphi - \varphi_0 + \frac{2}{3}\pi)} + e^{-jm(\varphi - \varphi_0 + \frac{2}{3}\pi)} - e^{jm(\varphi - \varphi_0 + \frac{5}{3}\pi)} - e^{-jm(\varphi - \varphi_0 + \frac{5}{3}\pi)} \right] \right\}.$$

Po wykonaniu obliczeń otrzymuje się

$$K(\Theta, \varphi) = \frac{K_0}{2\pi \sin \Theta} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{(2n+1)}{m} \frac{\sin m\alpha \cdot (n-m)!}{(n+m)!} \cdot P_n^m(\cos \Theta) \cdot$$

$$\cdot \left\{ e^{jm(\varphi - \varphi_0)} \cdot (1 - e^{jm\pi}) \left[1 - e^{j\left(\frac{m+2}{3}\pi\right)} + e^{-j\left(\frac{2m-4}{3}\pi\right)} \right] + \right.$$

$$\left. + e^{-jm(\varphi - \varphi_0)} (1 - e^{jm\pi}) \left[1 - e^{-j\left(\frac{m+2}{3}\pi\right)} + e^{-j\left(\frac{2m+4}{3}\pi\right)} \right] \right\}. \quad (33)$$

Na podstawie zależności (5) i (33) współczynniki K_m i L_m wyrażają się wzorami.

$$K_m = (1 - e^{jm\pi}) \cdot \left[1 - e^{j\left(\frac{m+2}{3}\pi\right)} + e^{j\left(\frac{2m-4}{3}\pi\right)} \right],$$

$$L_m = (1 - e^{jm\pi}) \cdot \left[1 - e^{-j\left(\frac{m+2}{3}\pi\right)} + e^{-j\left(\frac{2m-4}{3}\pi\right)} \right]. \quad (34)$$

Analizując wzory (34) stwierdza się, że współczynniki K_m i L_m przyjmują wartości

$$K_m = \begin{cases} 6, & m = 5, 11 \\ 0, & \text{pozostaje } m \end{cases} \quad L_m = \begin{cases} 6, & m = 1, 7, 13 \\ 0, & \text{pozostaje } m \end{cases}$$

W rozwinięciu gęstości prądu w szereg harmonik wystąpią zatem wyrazy o nieparzystych indeksach m . Uwzględniając, że współczynnik $c_{mn} = 0$ dla $n = 2p$ i $m = 2k + 1$ otrzymuje się w rozwinięciu w szereg gęstości prądu wyrazy wyłącznie o nieparzystych indeksach n . Zgodnie z uwagą poczynioną przy wzorze (6a), oznacza to, że analizowane uzwojenie wytwarza wirujące pole magnetyczne. Zgodnie z podstawową harmoniczną wiruje pole wytworzone przez 7, 13, ... harmoniczną, a przeciwnie 5 i 11 harmoniczną.

W związku z powyższymi uwagami wzory (28), (32) upraszczają się do postaci:

$$P = \frac{9(I \cdot w)^2}{\pi \alpha^2 \sigma} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{n(n+1) \cdot (2n+1)^3 \cdot \mathcal{L}^{2n} \sin^2(m\alpha) \cdot C_{mn}^2 b_s^2 \cdot (n-m)!}{\{[n(n+1) \cdot (1 - \mathcal{L}^{2n+1})]^2 + b_s^2 (n \cdot \mathcal{L}^{2n+1} + n+1)^2\} m^4 (n+m)!} \quad (35)$$

$$M = \frac{9(I \cdot w)^2 \mu_0 R_1}{\pi \alpha^2} \sum_{n=1}^{\infty} \sum_{m=1}^n \frac{n(n+1)(2n+1)^3 \cdot \mathcal{L}^{2n} \sin^2(m\alpha) \cdot C_{mn}^2 \cdot b_s \cdot (n-m)!}{\{[n(n+1)(\mathcal{L}^{2n+1} - 1)]^2 + b_s^2 (n \cdot \mathcal{L}^{2n+1} + n+1)^2\} m^3 (n+m)!} \cdot ZN \quad (36)$$

gdzie:

$$b_s = \omega \mu \sigma R_1 \cdot [1 + ZN(1-s) \cdot m]$$

$$ZN = \begin{cases} -1 & m = 1, 7, 13 \\ 1 & m = 5, 11, \dots \end{cases}$$

$I = I_m / \sqrt{2}$ — wartość skuteczna wzbudzenia

$n = 1, 3, 5, \dots$

$m = 1, 5, 7, 11, \dots$

Przeprowadzono analizę numeryczną zależności (35) i (36).

Wprowadza się następujące oznaczenia:

$$R_2 = R_1 + d + d_0,$$

gdzie:

R_2 — promień stojana

R_1 — promień kuli ferromagnetycznej

d — grubość warstwy miedzianej

d_0 — grubość szczeliny powietrznej pomiędzy twornikiem a stojanem.

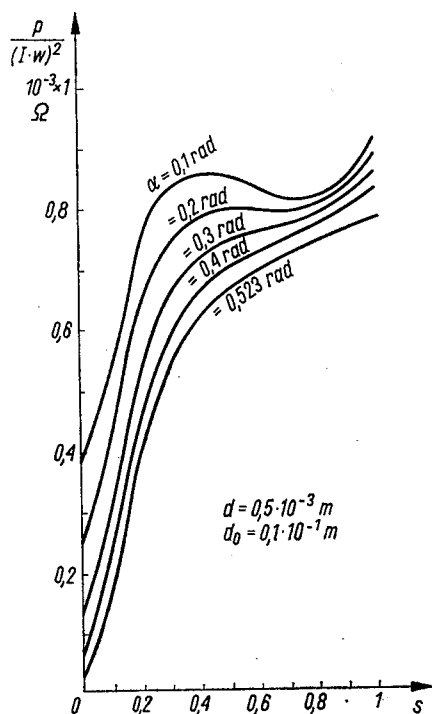
Do obliczeń przyjęto następujące dane liczbowe:

$$R_1 = 0.05 \text{ m}, \gamma = 56 \cdot 10^6 \cdot \Omega \text{m}^{-1}, \omega = 314 \text{ s}^{-1}, \mu_0 = 4\pi \cdot 10^{-7} \text{ H/m}.$$

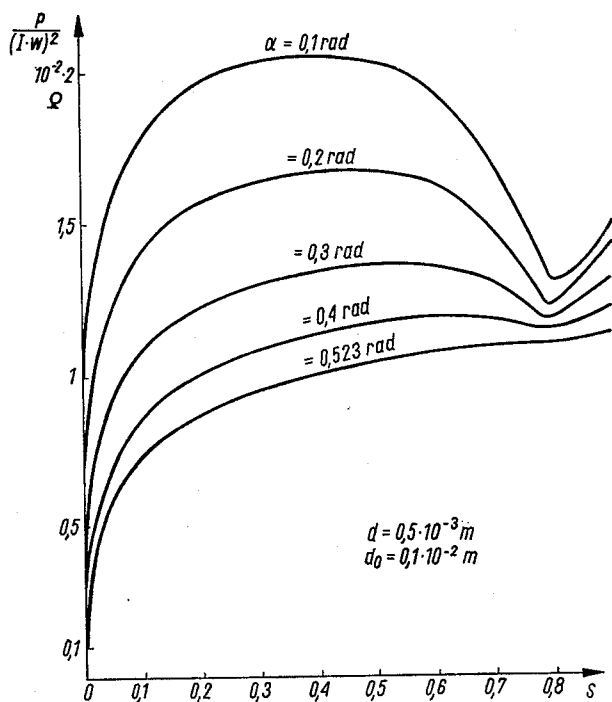
Wyniki analizy strat mocy i momentu napędowego silnika kulistego przedstawiono na rysunkach nr 7÷15.

Na rys. 7 i 8 przedstawiono zależność strat mocy w warstwie przewodzącej wirnika w funkcji poślizgu s dla różnych wartości kąta α rozchyłu uzwojeń stojana. Wzrost kąta rozchyłu uzwojeń powoduje zmniejszenie wartości strat. Obserwuje się (szczególnie dla małych kątów α) obniżenie wartości strat mocy w pobliżu wartości poślizgu $s = 0.8$. Wiąże się to z brakiem oddziaływania piątej harmonicznego — współczynnik b_s we wzorze (35) przyjmuje wartość równą zero.

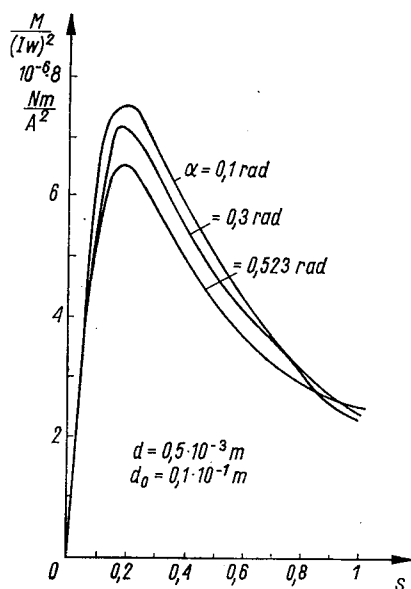
Rys. 9 i 10 przedstawiają zależność momentu napędowego od poślizgu s dla różnych wartości kąta α . Szczególnie na rys. 10, wykonanym dla $d_0 = 1 \cdot 10^{-3} \text{ m}$, widać wyraźnie obniżenie wartości momentu dla poślizgu $s = 0.7$. Jest to spowodowane wpływem wyż-



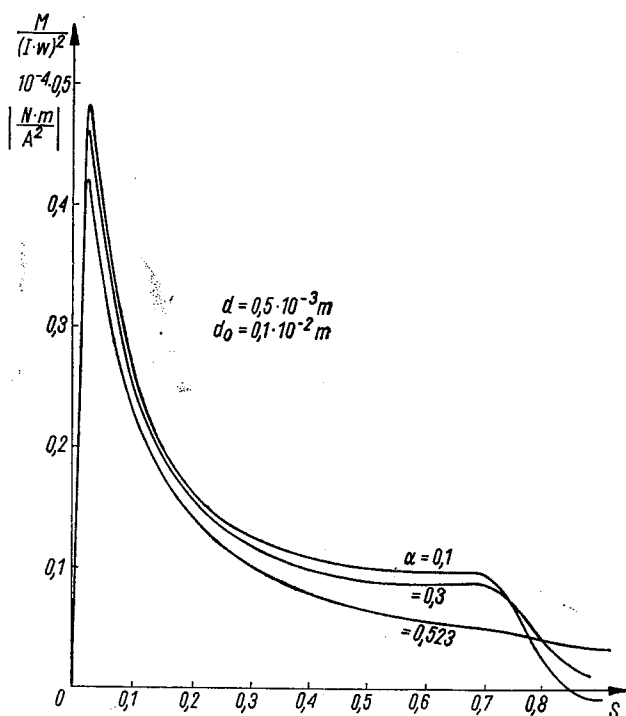
Rys. 7. Straty mocy w funkcji poślizgu (duża wartość grubości szczeliny powietrznej d_0)



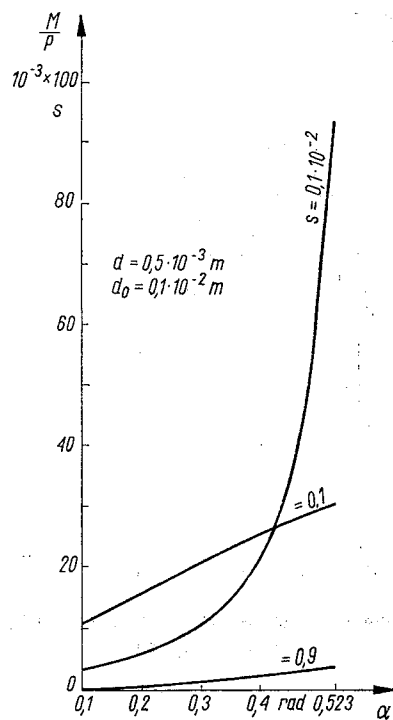
Rys. 8. Straty mocy w funkcji poślizgu (mała wartość grubości szczeliny powietrznej d_0)



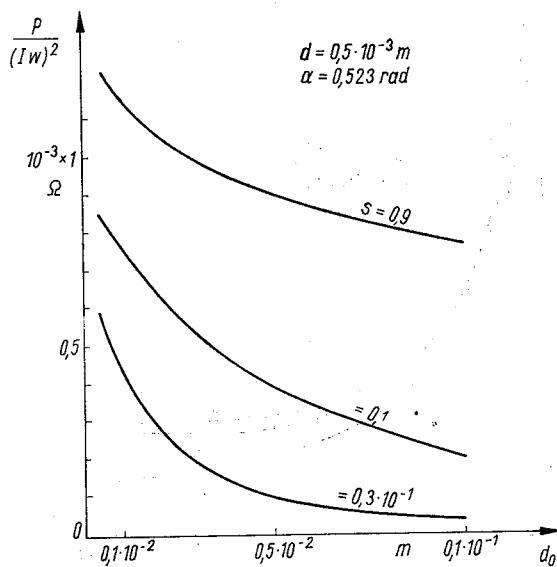
Rys. 9. Moment napędowy silnika kulistego w funkcji poślizgu (duża wartość grubości szczeliny powietrznej d_0)



Rys. 10. Moment napędowy silnika kulistego w funkcji poślizgu (mała wartość grubości szczeliny powietrznej d_0)



Rys. 11. Stosunek wartości momentu napędowego do wartości strat mocy w funkcji kąta rozchyłu uzwojeń α

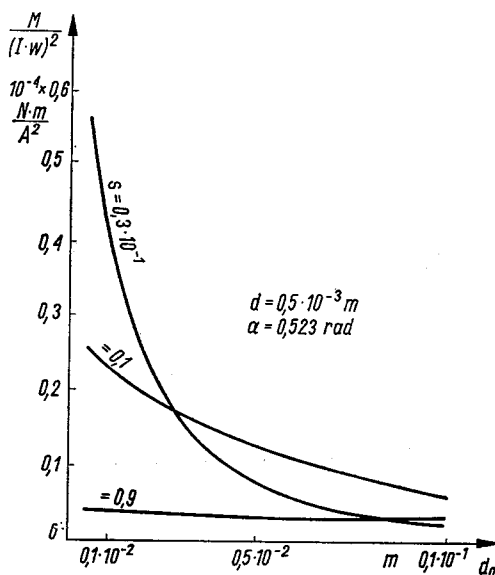


Rys. 12. Wpływ grubości szczeliny powietrznej pomiędzy twornikiem a stojanem na straty mocy silnika kulistego

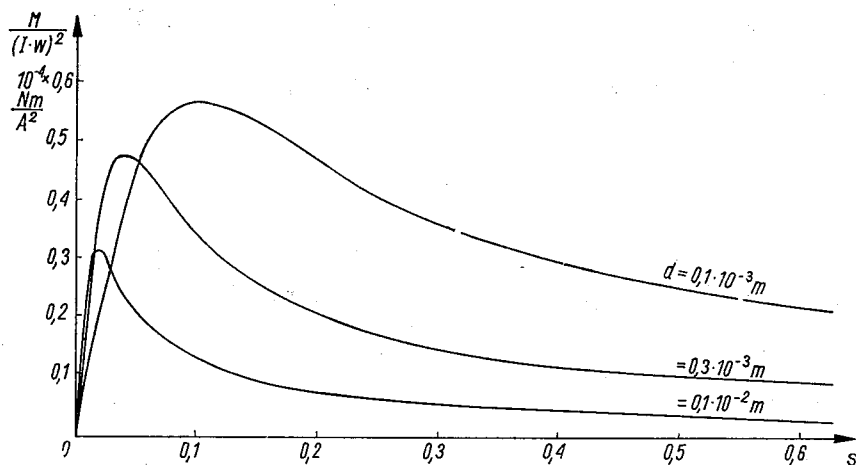
szych harmonicznych (zerowanie się współczynników b_s dla $s = 0.8$ (5 harmoniczna) oraz $s = 0.9$ (11 harmoniczna)). Dla krzywej odpowiadającej $\alpha = 0.523$ (rys. 10) nie obserwuje się tego zjawiska. Wiąże się to z małym udziałem wyższych harmonicznych w wytworzonym polu wirującym. Wzrost kąta α (rys. 9 i rys. 10) powoduje obniżenie wartości maksymalnej momentu natomiast wzrost momentu rozruchowego ($M(s=1)$).

Należy dodać, że rys. 7—15 wykonano przy założeniu stałej wartości przepływu $I \cdot w$ niezależnie od wartości kąta α .

Na rys. 11 przedstawiono zależność stosunku wartości momentu napędowego do strat mocy w funkcji kąta rozchyłu α . Z rysunku wynika, że optymalny kąt wynosi $\alpha_{opt} = 0.523$ rad, co odpowiada całkowitemu wypełnieniu na obwodzie stojana.



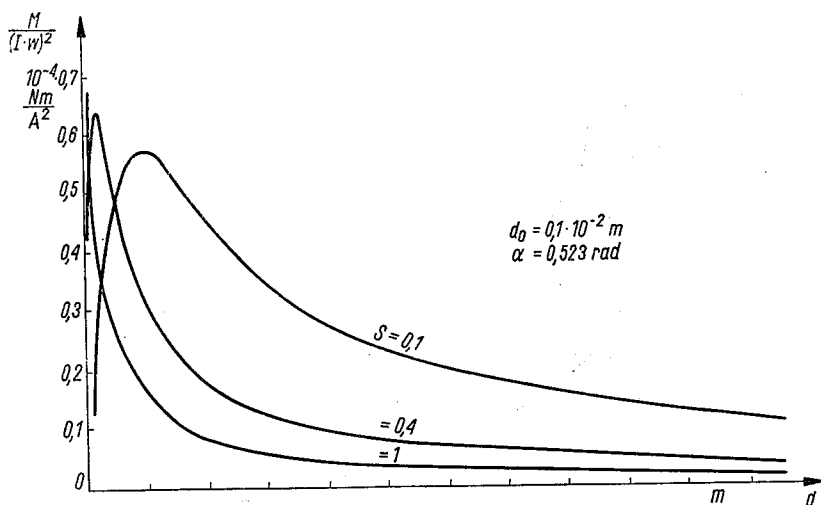
Rys. 13. Wpływ grubości szczeliny powietrznej d_0 na moment napędowy silnika kulistego



Rys. 14. Wpływ grubości warstwy przewodzącej na charakterystykę $M = f(s)$

Rys. 12 i 13 obrazują zależność strat mocy oraz momentu napędowego od wielkości szczeliny powietrznej d_0 . Wzrost d_0 powoduje obniżenie wartości P oraz M .

Na rys. 14 przedstawiono zależność momentu napędowego od poślizgu s dla różnych wartości grubości warstwy przewodzącej d . Obserwuje się wzrost momentu rozruchowego przy malejącej wartości grubości warstwy przewodzącej d (wzrost rezystancji warstwy przewodzącej). Występuje tu pełna analogia do zależności momentu rozruchowego klasycznych silników indukcyjnych od rezystancji obwodu wirnika. Dla uwypuklenia zależności momentu napędowego od grubości warstwy przewodzącej, wykonano rys. 15



Rys. 15. Zależność momentu napędowego silnika kulistego od grubości warstwy przewodzącej d

przedstawiający zależność $M = f(d)$ dla różnych wartości poślizgu s . Dla poszczególnych wartości poślizgu, zależność $M = f(d)$ posiada ekstremum. Rysunek nie ujmuje ekstremum dla poślizgu $s = 1.0$ z powodu małej wartości d odpowiadającej temu ekstremum.

PODSUMOWANIE

Wyprowadzone w pracy zależności (wzory (28) i (32)) umożliwiają wyznaczanie strat mocy oraz momentu napędowego dla różnych wariantów rozmieszczenia uzwojeń stojana.

W wyniku przeprowadzonej analizy numerycznej określono wpływ takich wielkości jak: poślizg, grubość warstwy przewodzącej, wielkość szczeliny powietrznej oraz kąt rozchyłu uzwojeń stojana na wartość strat mocy i momentu napędowego.

Stwierdzono silną zależność wartości momentu napędowego od grubości warstwy przewodzącej (rys. 14 i 15). Umożliwia to kształtowanie charakterystyki $M = f(s)$ przez dobór odpowiedniej grubości warstwy przewodzącej.

DODATEK A.

Obliczanie całki oznaczonej $C_{mn} = \int_0^\pi P_n^m(\cos\theta) \sin\theta d\theta$

$$C_{mn} = \int_0^\pi P_n^m(\cos\theta) \sin\theta d\theta = \int_{-1}^1 P_n^m(x) dx, \quad \text{A.1}$$

gdzie $x = \cos\theta$, $n = 1, 2, 3, \dots$, $m = 1, 2, 3, 4, \dots$

Stowarzyszona funkcja Legendre'a wyraża się wzorem [11, 12]

$$\begin{aligned} P_n^m(x) &= \frac{(-1)^m (2n-1)!! (1-x^2)^{\frac{m}{2}}}{(n-m)!} \cdot \left[x^{n-m} - x^{n-m-2} \cdot \frac{(n-m)(n-m-1)}{2(2n-1)} + \right. \\ &\quad \left. + \frac{(n-m)(n-m-1)(n-m-2)(n-m-3)}{2 \cdot 4 \cdot (2n-1)(2n-3)} \cdot x^{n-m-4} - \dots \right] = \\ &= \frac{(-1)^m (2n-1)!!}{(n-m)!} \sum_{l=0}^{\left[\frac{n-m}{2}\right]} (-1)^l \cdot x^{n-m-2l} \frac{(n-m)!(2n-1-2l)!!}{(n-m-2l)!(2n-1)!!(2l)!!} (1-x^2)^{\frac{m}{2}} \end{aligned} \quad \text{A.2}$$

Ze wzorów A.1 i A.2 wynika

$$C_{mn} = \frac{(-1)^m (2n-1)!!}{(n-m)!} \sum_{l=0}^{\left[\frac{n-m}{2}\right]} (-1)^l \frac{(n-m)!}{(n-m-2l)!} \cdot \frac{(2n-2l-1)!!}{(2n-1)!!(2l)!!} \cdot I_{lmn} \quad \text{A.3}$$

gdzie

$$I_{lmn} = \int_{-1}^1 (1-x^2)^{\frac{m}{2}} \cdot x^{n-m-2l} dx \quad \text{A.4}$$

Przyjmując za n liczbę parzystą oraz za m liczbę nieparzystą otrzymuje się wyrażenie podcałkowe (wzór A.4) będące funkcją nieparzystą zmiennej x . Stąd $I_{Lmn} = 0$.

Przyjmując za n liczbę nieparzystą oraz za m liczbę parzystą otrzymuje się wyrażenie podcałkowe we wzorze A.4 będące funkcją nieparzystą zmiennej x . Stąd $I_{Lmn} = 0$.

Dla nieparzystej wartości $n = 2k-1$ oraz $m = 2p-1$ otrzymuje się

$$I = 2 \int_0^1 (1-x^2)^{p-\frac{1}{2}} \cdot x^{2k-2p-2l} dx. \quad \text{A.5}$$

Stosując wzory [11]

$$\int_0^1 (1-x^2)^{r-1} x^{q-1} dx = \frac{1}{2} B\left(r, \frac{q}{2}\right) \quad \text{A.6}$$

gdzie

$B(x, y)$ — całka Eulera II-go rodzaju oraz

$$B(x+y) = \frac{\Gamma(x) \cdot \Gamma(y)}{\Gamma(x+y)} \quad \text{A.7}$$

Korzystając z własności funkcji [11]

$$\begin{aligned} \Gamma'(x+1) &= x\Gamma'(x) \\ \Gamma\left(n + \frac{1}{2}\right) &= \frac{\sqrt{\pi}}{2^n} (2n-1)!! \end{aligned} \quad \text{A.8}$$

otrzymuje się wartość całki wzór A.5

$$I = \frac{\pi(2p-1)!!(2k-2p-2l-1)!!}{2^{k-1}(k-l)!}$$

Wracając do pierwotnych oznaczeń m i n , na podstawie wzoru A.3 otrzymuje się:

$$C_{mn} = \pi(m)!! \sum_{l=0}^{\left[\frac{n-m}{2}\right]} (-1)^{l+1} \frac{(2n-2l-1)!!}{(n-2l+1)!!(n-m-2l)!!(2l)!!} \quad \text{A.9}$$

gdzie $m = 1, 3, 5, \dots, n$; $n = 1, 3, 5, \dots$

Przyjmując parzystą wartości $n = 2k$ oraz $m = 2p$ i wykonując analogiczne obliczenia jak poprzednio otrzymuje się:

$$I_{Lmn} = 2^{\frac{m+2}{2}} \left(\frac{m}{2}\right)! \frac{(n-m-2l-1)!!}{(n-2l+1)!!}$$

oraz

$$C_{mn} = 2(m)!! \sum_{l=0}^{\left[\frac{m-n}{2}\right]} (-1)^l \frac{(2n-2l-1)!!}{(n-2l+1)!!(n-m-2l)!!(2l)!!} \quad \text{A.10}$$

gdzie $m = 2, 4, 6, \dots, n$; $n = 2, 4, 6, \dots$

WYKAZ OZNACZEŃ

- $\vec{B}$ — wektor indukcji magnetycznej
- $\vec{E}$ — wektor natężenia pola elektrycznego
- $\vec{H}$ — wektor natężenia pola magnetycznego
- I — prądowa funkcja
- $I_m \cdot w$ — przepływ (wartość maksymalna)
- $\vec{J}$ — wektor gęstości prądu
- $\vec{K}$ — wektor gęstości powierzchniowej prądu
- M — moment napędowy (wartość średnia)
- P — straty mocy
- U — potencjał skalarny pola magnetycznego
- r, Θ, φ — współrzędne układu sferycznego
- s — poślizg
- α — kąt
- γ — konduktywność
- μ — przenikalność magnetyczna

- σ — konduktywność powierzchniowa
 ω — pulsacja

Uwaga: W pracy operuje się amplitudą zespoloną wektorów:

$\vec{B}$, $\vec{E}$, $\vec{H}$, $\vec{J}$, $\vec{K}$ oraz skalarów: I , U .

BIBLIOGRAFIA

1. M. Abramowitz, I. A. Stegun: *Handbook of mathematical functions*. New York, Dover 1970
2. W. W. Chruscew: *Elektriceskije mikromasyny awtomaticheskich ustrojstw*. Leningrad: Energiya, 1976
3. J. A. Gorejstejn, J. A. Sumlacz: *Inercialnyjne nawigacijonnoje sistimey*. Moskwa: Masynostrojenije, 1970
4. M. Gramz, J. Purczyński, R. Sikora: *Calculation of eddy currents induced in thin curved sheets*, Conference on the Computation of Magnetic Field COMPUMAG 78, Grenoble Sept. 1978
5. G. Kamiński, M. Kant: *Elektryczny silnik o ruchu złożonym z wirnikiem kulistym*. Przegląd Elektrotechniczny, 1981, nr 8
6. G. A. Korn, T. M. Korn: *Mathematical handbook for scientists and engineers*. New YORK, Toronto, London, 1961
7. E. A. Mendrela, E. Gierczak: *Własności elektromechaniczne silnika indukcyjnego o wirniku kulistym*, Arch. Elektrot. t. 30, 1984, z. 3
8. E. A. Mendrela, G. Kamiński: *Silniki elektryczne o trzech stopniach swobody*. Rozprawy Elektrotechniczne, 1985, t. 31, z. 3—4
9. A. W. Milach, W. A. Barabanow, E. W. Dwoinach: *Triechstiepiennyje elektriceskije masyny*, Kijów 1979
10. J. Purczyński: *Pola magnetyczne silnika elektrycznego o trzech stopniach swobody*. Archiwum Elektrotechniki, 1974, t. 23, z. 2
11. J. M. Ryżyk, J. S. Gradsztejn: *Tablice calek, sum, szeregów i iloczynów*. Warszawa: PWN, 1964
12. *Tablicy prisojedinionnych funkcji Lieżandra*. Moskwa: Bibl. Matemat. Tablic, 1962, Wypusk 4

J. PURCZYŃSKI, L. KASZYCKI

POWER LOSSES AND TORQUE OF INDUCTION MOTOR OF SPHERICAL SYMMETRY

Summary

The power losses and torque of an induction motor of spherical symmetry are determined. The analysis is carried out using the current function and the scalar potential of the magnetic field. Derived formulae allow to consider an arbitrary space distribution of stator windings. As a result of a numerical analysis the influence of the values of: slip, thickness of conductive layer and air gap as well as the flare angle of windings on the power loss value and torque, is determined.

J. PURCZYŃSKI, L. KASZYCKI

DISSIPATIONS DE PUISSANCE ET LE MOMENT MOTEUR DU MOTEUR D'INDUCTION A SYMETRIE SPHERIQUE

Résumé

Dans le travail on a déterminé les dissipations de puissance et le moment moteur du moteur sphérique. Dans l'analyse on a utilisé la fonction de courant et le potentiel scalaire. On a établi des formules, permet-

tant d'analyser une disposition quelconque des enroulements du stator. A partir de l'analyse numérique on a déterminé l'influence du glissement, de la grosseur de la fente d'aire et de l'angle entre des enroulements du stator sur la valeur des dissipations de puissance et du moment moteur.

J. PURCZYŃSKI, L. KASZYCKI

VERLUSTLEISTUNG UND ANTRIEBSDREHMOMENT EINER KUGELSYMMETRISCHEN INDUKTIONSMASCHINE

Zusammenfassung

In der Bearbeitung wurden Wirbelstromverluste und das Antriebsdrehmoment eines kugelförmigen Motors bestimmt. Für die Analyse wurden Stromfunktion und Skalarpotential des magnetischen Feldes angewandt. Die abgeleiteten Formeln ermöglichen die Analyse im Falle einer beliebigen Verteilung der Ständerwicklung. Als Ergebnis der numerischen Analyse wurden die Einflüsse des Schlups, der Stärke der leitenden Schicht, des Luftspalts und des Differenzwinkels der Wicklung auf die Werte der Verluste und des Drehmomentes bestimmt.

Я. ПУРЧИНСКИ, Л. КАШИЦКИ

ПОТЕРИ МОЩНОСТИ И ПРИВОДНОЙ МОМЕНТ АСИНХРОННОГО ШАРОВОГО ЭЛЕКТРОДВИГАТЕЛЯ

Резюме

Определены потери мощности и приводной момент шарового электродвигателя. Для анализа использовано токовую функцию, а также скалярный потенциал магнитного поля. Выведённые формулы дают возможность проведения анализа для любого размещения обмоток статора. В результате проведённого цифрового анализа определено влияние скольжения, мощности токоведущего слоя, величины воздушного зазора, а также угла отклонения обмоток статора на значение потерь мощности и приводного момента.

Przetworniki wiropądowe do nieniszczących badań rur nieferromagnetycznych

ANNA LEWIŃSKA-ROMICKA

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

Otrzymano 1987.06.10

Autoryzowano do druku 1987.11.11

W artykule opisano wiropądowe przelotowe przetworniki transformatorowe o quasi-równomiernym rozkładzie pola magnetycznego i sposób doboru warunków pracy przetworników, przeznaczonych do wykrywania nieciągłości szwu rur. Przedstawiono wyniki badań.

OBIEKTY BADANE

Obiektem badanym są rury wykonane z materiałów nieferromagnetycznych: miedzi i mosiądzu.

Rozróżnia się następujące postacie rur:

- 1) wyciskane,
- 2) wyciskane — kalibrowane,
- 3) ciągnione,
- 4) zwijane z taśmy, a następnie spawane.

Wymagania odnośnie rur, wymienionych w pkt. 1 ÷ 3, są przedmiotem normy PN-77/H-74586 (Miedź i stopy miedzi. Rury, ogólne wymagania i badania), a także norm PN-79/H-87053 (Miedź stopowa), PN-77/H-87025 (Mosiądze), PN-77/H-87050 (Brązy) i PN-77/H-82120 (Miedź — skład chemiczny). Brak jest natomiast normy dotyczącej jakości rur spawanych. Stosuje się jedynie normy DIN 1754-1755 i DIN 17671 (Rury miedziane i ze stopów miedzi).

Wymagania odnośnie rur

Zgodnie z normą PN-77/H-74586 powierzchnia zewnętrzna i wewnętrzna rur powinna być czysta i gładka, odpowiednio do technologii wytwarzania. Na powierzchni rur dopuszcza się:

- a) drobne pęcherze (tylko na powierzchni rur wyciskanych i kalibrowanych), pojedyncze rysy, odciski, ślady usuwania drobnych wad oraz ślady pierścieniowe i spiralne, jeżeli

ich głębokość nie przekracza dla rur wyciskanych pola tolerancji, dla rur ciągnionych — połowy pola tolerancji grubości ścianki,

b) barwy nalotowe oraz plamy po obróbce cieplnej i trawieniu.

Makrostruktura rur wyciskanych nie powinna wykazywać obcych wtrąceń, porów, pęknięć i tzw. wciągów. Dopuszczalne są drobne wady w postaci punktów w liczbie nie większej niż 2 na 1 cm² próbki i o powierzchni nie większej niż 0,5 mm² każda.

Rury z miedzi i mosiądzu w stanie rekryształizacji, o grubości ścianki do 3 mm, spłaszczane do zetknięcia się przeciwległych ścianek, nie powinny wykazywać pęknięć, rys i rozwarstwień widocznych nieuzbrojonym okiem.

Rury z mosiądzu M63 w stanie półtwardym, poddane próbie spłaszczenia do 1/3 pierwotnej średnicy nie powinny wykazywać pęknięć ani rozwarstwień.

Jakość rur spawanych (zwijanych z taśmy) zależy głównie od jakości szwu. Rury nieferromagnetyczne spawane są tzw. metodą TIG (elektroda wolframowa, osłona argonowa). Szew przebiega wzdłuż długości rur. Wadą rur spawanych jest brak przetopu (nie zgrzana spoina), dziury powstałe z wypaleń elektrodą oraz przetop jedynie na części grubości ścianki rury. Typowymi zastosowaniami rur, wykonanych z miedzi i mosiądzu, są wymienniki ciepła, osprzęt spawalniczy, instalacje, otuliny części grzejnych, armatura, osprzęt oświetleniowy. Rury spawane stanowią zarówno produkt finalny, jak i półprodukt do obróbki plastycznej na zimno. W procesie ciągnięcia otrzymywane są z tych rur rury o mniejszych średnicach, również kapilarne. Wady szwu rur spawanych, jak np. częściowy przetop są przyczyną powstawania nieszczelności rur otrzymywanych w procesie dalszego ciągnięcia.

PROBLEMY KONTROLI RUR ZA POMOCĄ PRZETWORNIKÓW WIROPRAĐOWYCH PRZEŁOTOWYCH

W przypadku badań rur interesujące są następujące parametry:

- przewodność właściwa γ , związana ze strukturą i składem chemicznym materiału,
- przenikalność magnetyczna względna μ_r (dla materiałów ferromagnetycznych),
- średnica zewnętrzna rury D_z ,
- średnica wewnętrzna otworu rury D_w ,
- grubość ścianki G ,
- mimosrodowość ścianki,
- występowanie nieciągłości materiału i szwu rur.

Przy badaniu rur mamy do czynienia z większą liczbą zmiennych w porównaniu z badaniem elementów pełnych, a więc całokształt zależności fizycznych jest tu bardziej złożony. W związku z tym wygodnie jest rozpatrzyć osobno przypadek rury cienkościennej i rury grubościenniej [8]. Dla oceny możliwości kontroli rur cienkościennych, dla których grubość ścianki $G < 0,5\% D_z$, istotne jest wzajemne położenie następujących kierunków na płaszczyźnie zmiennej zespolonej, na której przedstawiane są sygnały otrzymywane z przetworników:

- kierunku „ f/f_g ”
- kierunku „ D_z ”.

Jako f oznaczono tu częstotliwość prądu wzbudzającego przetwornik, a jako f_g tzw. czę-

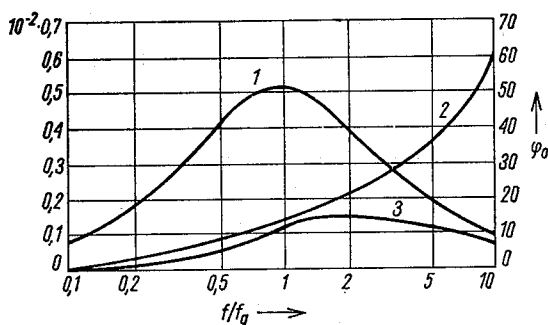
stotliwość graniczną [wzór (1)]. Częstotliwość graniczna f_g zależna jest od przewodności właściwej γ , przenikalności magnetycznej μ_r materiału rury oraz średnicy wewnętrznej i grubości jej ścianki.

W kierunku f/f_g występują sygnały przetwornika, wywołane przez zmiany parametrów materiału tj. przewodności γ , a więc m.in. wywołane przez nieciągłości i przenikalności magnetycznej μ_r oraz zmiany wymiarów, tj. grubości ścianki G i średnicy wewnętrznej D_w (zmiana iloczynu GD_w), gdy średnica zewnętrzna jest stała.

W kierunku D_z występują sygnały przetworników, wywołane przez zmiany średnicy zewnętrznej (zmiany grubości ścianki rury), gdy jej średnica wewnętrzna D_w jest stała.

Istnieje możliwość odróżniania efektów występujących w tych obu kierunkach [8]. A więc możliwe jest np. wykrywanie nieciągłości rur, wykonanych z materiałów nieferromagnetycznych, gdy średnice wewnętrzne i zewnętrzne rury są stałe lub odróżnienie zmian średnicy wewnętrznej od zmian średnicy zewnętrznej rur, gdy przewodność materiału rury jest stała lub też odróżnienie nieciągłości na powierzchni zewnętrznej i wewnętrznej. Problemem podstawowym przy realizacji kontroli jakości rur metodą prądów wirowych jest odpowiedni dobór częstotliwości pracy. Optymalne odróżnienie sygnałów, występujących w obu kierunkach, otrzymuje się przy przyjęciu stosunku częstotliwości pracy do częstotliwości granicznej $f/f_g = 0,8 \div 8$ [8].

Największy wpływ zmian parametrów γ , D_w i G rur występuje dla wartości stosunku częstotliwości pracy do częstotliwości granicznej $f/f_g = 1$ (rys. 1) [8].



Rys. 1. Przyrost przenikalności skutecznej rury cienkościennej z materiału nieferromagnetycznego, spowodowany 1-procentową zmianą parametrów γ , μ_r oraz $(D_w G)$ w zależności od stosunku f/f_g
1 — moduł przyrostu, 2 — kąt Φ między kierunkiem f/f_g , a kierunkiem D_z , 3 — składowa przyrostu prostopadła do kierunku D_z [8]

W przypadku kontroli rur grubościennych, wykonanych z materiałów nieferromagnetycznych, dla których $G \geq 0,5\% D_z$, można odróżnić zmiany przewodności od zmian grubości ścianki i średnicy zewnętrznej [8].

Dla rur o grubości ścianki od $0,5\% D_z$ do $20\% D_z$ optymalna wartość stosunku częstotliwości pracy do częstotliwości granicznej f/f_g dla odróżnienia tych zmian w przypadku zastosowania przetwornika zewnętrznego wynosi od 200 do 8. Gdy $G < 10\% D_z$ i $f/f_g < 15$ sygnały przetwornika wiropądowego wywołane przez wady na powierzchni wewnętrznej i na powierzchni zewnętrznej mają pod kątem fazowym ten sam kierunek. A więc można

sformułować wniosek, że wady te mogą być wykrywane jednakowo skutecznie zarówno przy zastosowaniu przetworników zewnętrznych, jak i wewnętrznych.

RURA CIENKOŚCIENNA W PRZETWORNIKU PRZELOTOWYM

Dla rury cienkościennej częstotliwość graniczną można obliczyć według następującej zależności [8]:

$$f_g = \frac{5066}{\gamma \mu_r D_w G}, \quad [\text{Hz}] \quad (1)$$

gdzie:

γ — przewodność właściwa materiału rury, S/m,

μ_r — przenikalność magnetyczna względna,

D_w — średnica wewnętrzna rury, cm,

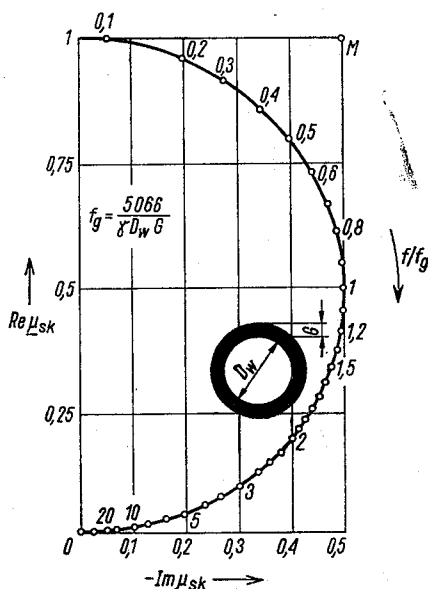
G — grubość ścianki rury, cm.

Dla rur z materiałów nieferromagnetycznych $\mu_r = 1$.

W przypadku kontroli rur z materiałów austenitycznych badane rury podmagnesowuje się dla ustabilizowania wartości przenikalności magnetycznej.

W celu umożliwienia analizy właściwości przetworników wiroprowadowych wprowadzono pojęcie tzw. przenikalności skutecznej μ_{sk} , zależnej od częstotliwości pracy f przetwornika i częstotliwości granicznej f_g [8]

$$\text{Re} \mu_{sk} = \frac{1}{1 + (f/f_g)^2} \quad \text{Im} \mu_{sk} = \frac{f}{1 + (f/f_g)^2} \quad (2)$$



Rys. 2. Przenikalność skuteczna cienkościennej rury z materiału nieferromagnetycznego w funkcji stosunku f/f_g [8]

Przenikalność skuteczną przedstawia na płaszczyźnie zmiennej zespolonej półokrąg oparty na osi rzędnych średnicy równej 1 (rys. 2) [8]. Napięcie wyjściowe, względnie impedancja przetworników wiroprowodowych, zawierających obiekt badany, przedstawiane są w postaci tzw. unormowanej, tj. w odniesieniu do odpowiednich wielkości, charakteryzujących przetwornik pusty. Unormowana impedancja (napięcie wyjściowe) przetwornika zawierającego rurę opisana jest następującymi zależnościami:

$$\frac{\omega L}{\omega L_0} = \frac{ImE}{E_0} = 1 - \eta + \eta Re\mu_{sk}$$

$$\frac{R}{\omega L_0} = \frac{ReE}{E_0} = \eta Im\mu_{sk}$$
(3)

gdzie:

ω — pulsacja prądu zasilającego przetwornik,

L_0 — indukcyjność przetwornika pustego,

L — indukcyjność przetwornika zawierającego obiekt badany,

E_0 — napięcie indukowane w przetworniku pustym,

E — napięcie indukowane w przetworniku zawierającym obiekt badany,

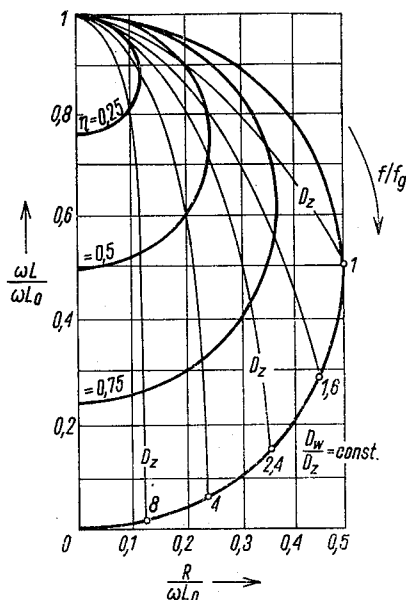
$\eta = \left(\frac{D_z}{D_s}\right)^2$ — tzw. współczynnik wypełnienia przetwornika przez obiekt badany,

D_z — średnica zewnętrzna rury,

D_s — średnica wewnętrzna przetwornika,

R — rezystancja przetwornika zawierającego obiekt.

Jeżeli założyć, że stosunek średnicy zewnętrznej D_z rury do średnicy wewnętrznej D_w przetwornika D_z/D_w jest stały, to częstotliwość graniczną można będzie wyrazić nastę-



Rys. 3. Unormowana impedancja przetwornika przelotowego zawierającego cienkościenną rurę z materiału nieferromagnetycznego, gdy $D_w/D_z = \text{const}$ [8]

dla próbki pełnej i zbiegające się w punkcie $\mu_{sk} = 1$, obrazują, jak zmienia się przenikalność skuteczna, gdy próbka pełna zmienia się w rurę cienkościenną. Krzywe ciągłe odwzorowują *kierunek G* tj. kierunek zmian grubości ścianki, gdy $D_z = \text{const}$. Jak widać krzywe przecinają się pod coraz większym kątem w miarę wzrostu f/f_g , co oznacza, że można odróżnić zmiany grubości ścianki od zmiany przewodności i średnicy zewnętrznej. Optymalna wartość f/f_g dla odróżnienia tych zmian, gdy grubość ścianki zawarta jest

Tabela 1

Optymalne częstotliwości pracy

Grubość ścianki % D_z	$(f/f_g)_{opt}$	Optymalny zakres f/f_g
20	8,35	6—12
15	9,5	7—14
10	12,5	9—18
5	22,5	15—30
2,5	42	30—60
1	100	70—140
0,5	200	140—280

w granicach $(0,01 \div 0,4) D_z$ podana jest w tabeli 1 [8]. Wartość stosunku częstotliwości pracy do częstotliwości granicznej można obliczyć wg wzoru (5) [8]:

$$f/f_g = \frac{1}{\left(1 - \frac{2G}{D_z}\right) \frac{G}{D_z}} \quad (5)$$

We wzorze (5) jako f_g podstawiono częstotliwość graniczną obliczoną ze wzoru dla próbki pełnej [8], tj.

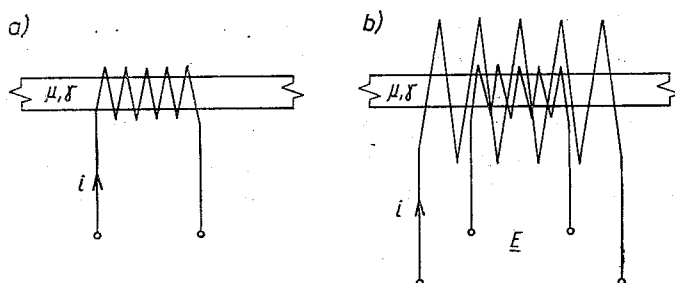
$$f_g = \frac{5066}{\gamma \mu_r D_z^2} \quad (6)$$

Przedstawione wyżej rozważania przeprowadzono dla przypadków uproszczonych, nie uwzględniających kształtu i parametrów przetwornika.

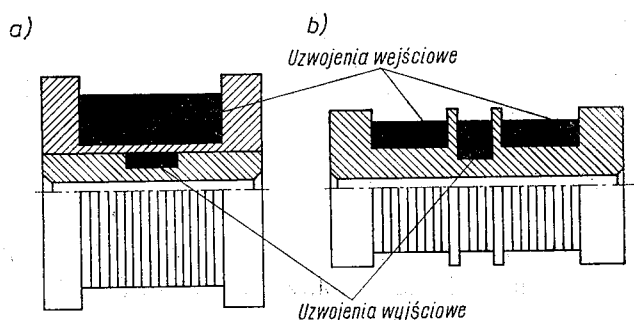
PRZETWORNIKI POMIAROWE

W metodzie prądów wirowych do badania drutów, prętów i rur stosowane są przetworniki przelotowe o różnorodnej budowie. W zależności od wielkości, stanowiącej sygnał wyjściowy, będący miarą wpływu obiektu badanego, przetworniki wiropądowe można podzielić na dwie podstawowe grupy: parametryczne (rys. 5a) i transformatorowe (rys. 5b). W przypadku przetworników transformatorowych uzwojenie wyjściowe umieszczone jest możliwie blisko obiektu badanego w celu zapewnienia silnego sprzężenia z obiektem. W zależności od charakteru rozkładu składowej osiowej pola magnetycznego (w granicach uzwojenia wyjściowego) przetworniki wiropądowe można podzielić na dwie podstawowe grupy:

1. przetworniki o jednorodnym rozkładzie pola magnetycznego (rys. 6a)
2. przetworniki o niejednorodnym rozkładzie pola magnetycznego (rys. 6b).



Rys. 5. Schematy przetworników wiroprowadowych przelotowych a) przetwornik parametryczny, b) przetwornik transformatorowy



Rys. 6. Konfiguracje przetworników pomiarowych przelotowych a) o jednorodnym rozkładzie pola magnetycznego, b) o niejednorodnym rozkładzie pola magnetycznego

Dotychczas są stosowane (w znanych opracowaniach) przetworniki o polu jednorodnym. Rozkład składowej osiowej pola magnetycznego dla przetwornika jednowarstwowego w postaci cewki długiej, o znikomej grubości uzwojenia można wyznaczyć korzystając z następującej zależności [22]:

$$H = \frac{zI}{4L} \left[\frac{L+x}{\sqrt{r_0^2 + (L+x)^2}} + \frac{L-x}{\sqrt{r_0^2 + (L-x)^2}} \right], \quad (7)$$

przy czym:

- H — pole magnetyczne,
- z — liczba zwojów przetwornika,
- I — natężenie prądu wzbudzającego przetwornik,
- $2L$ — długość przetwornika,
- r_0 — średnica przetwornika,
- x — odległość punktu, w którym obliczamy natężenie pola, względem środka uzwojenia przetwornika.

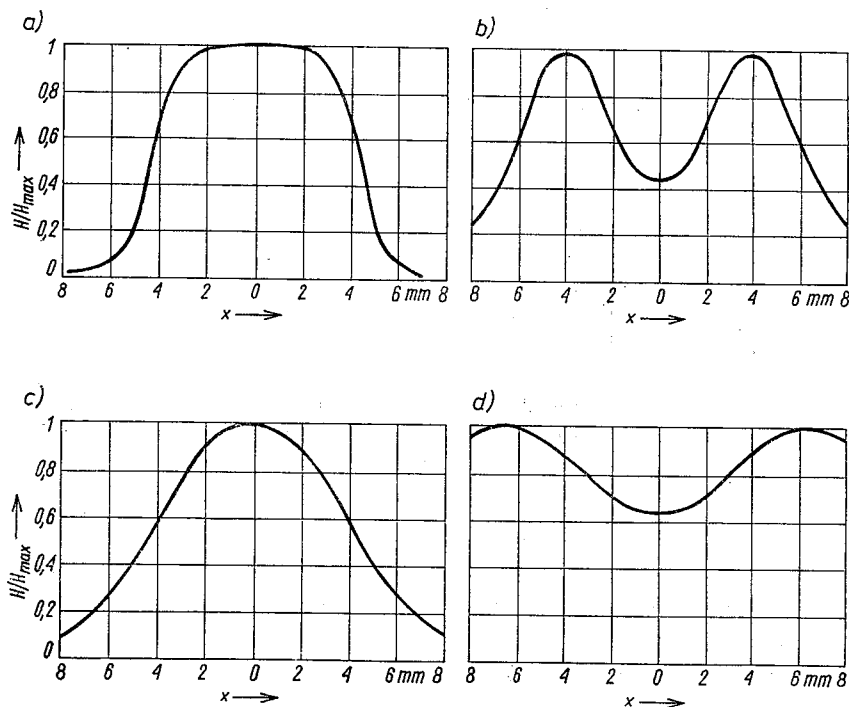
Rozkład składowej osiowej pola magnetycznego dla przetwornika wielowarstwowego można wyznaczyć korzystając z następującej zależności [22]

$$H = \frac{zI}{4L(r_2 - r_1)} \left[(L+x) \ln \frac{r_2 + \sqrt{r_2^2 + (L+x)^2}}{r_1 + \sqrt{r_1^2 + (L+x)^2}} + (L-x) \ln \frac{r_2 + \sqrt{r_2^2 + (L-x)^2}}{r_1 + \sqrt{r_1^2 + (L-x)^2}} \right] \quad (8)$$

przy czym:

r_2 i r_1 — promień zewnętrzny i wewnętrzny przetwornika.

Rozkład składowej osiowej pola magnetycznego przetworników, pokazanych na rys. 6, przedstawiono na rys. 7.



Rys. 7. Rozkład składowej osiowej pola magnetycznego wirowoprądowych przetworników transformatorowych

a) przetwornika o jednorodnym rozkładzie pola magnetycznego (rys. 6a), przetwornika o niejednorodnym rozkładzie pola magnetycznego (rys. 6b), x — odległość wzdłuż osi obiektu badanego względem środka uzwojenia wyjściowego

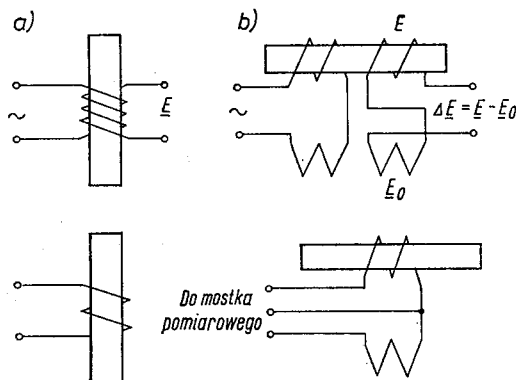
Poszczególne rodzaje przetworników mogą być stosowane pojedynczo, bądź parami — w połączeniach zależnych od celu i wymaganej czułości kontroli oraz od tego, wpływ jakich parametrów obiektu chcemy uwypuklić lub stłumić. Zasadniczo budowane są trzy układy przetworników do stosowania:

- 1) w metodzie bezwzględnej,
- 2) w metodzie porównawczej,
- 3) w metodzie samoporównawczej.

METODA BEZWZGLĘDNA

Metoda bezwzględna polega na pomiarze bezwzględnej wartości (tzn. całkowitej wartości) bądź impedancji przetwornika parametrycznego lub też napięcia indukowanego w uzwojeniu pomiarowym przetwornika transformatorowego. Metoda bezwzględna

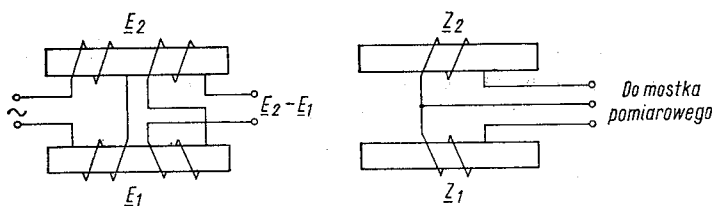
nadaje się jedynie do wykrywania znacznych różnic właściwości obiektów badanych. Układy przetworników, stosowanych w metodzie bezwzględnej oraz w metodzie bezwzględnej z cewką kompensującą (zawsze pustą) pokazano na rys. 8.



Rys. 8. Schematy ogólnie przetwornika w układzie bezwzględnym (a) i w układzie bezwzględnym z przetwornikiem kompensującym (b)

METODA PORÓWNAWCZA

Metoda porównawcza jest stosowana, gdy celem badania jest stwierdzenie różnic właściwości dwóch elementów. W metodzie porównawczej wykorzystywane są dwa identyczne przetworniki połączone w sposób pokazany na rys. 9.

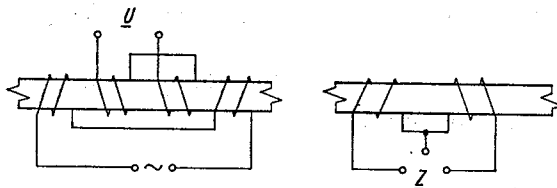


Rys. 9. Schematy ogólnie przetwornika w układzie porównawczym

Do jednego z przetworników wkładany jest element odniesienia (wzorcowy), do drugiego — element badany. Element odniesienia, będąc jednak w sposób ciągły pod wpływem prądów wirowych, nagrzewa się, powodując płynięcie wskazań. Często zamiast przetwornika z elementem odniesienia stosowane jest źródło napięcia kompensującego o regulowanej amplitudzie i fazie [15, 30]. Metoda porównawcza jest metodą bardzo czułą, pozwalającą przy odpowiedniej czułości układu pomiarowego uzyskać pełne wychylenie wskaźnika przyrządu dla różnic właściwości badanego obiektu wynoszących 1% lub mniej.

METODA SAMOPORÓWNAWCZA

Metoda samoporównawcza stosowana jest do wykrywania nieciągłości materiałów w prętach i rurach. Metody bezwzględna i porównawcza nie nadają się do tego celu, ponieważ zmiany impedancji lub napięcia wyjściowego przetwornika spowodowane występowaniem nieciągłości w tych obiektach na ogół są małe w porównaniu z rozrzutami wywołanymi przez różnice struktury, składu chemicznego i wymiarów pomiędzy elementem odniesienia i elementem badanym. W metodzie samoporównawczej stosowany jest



Rys. 10. Schematy ogólne przetwornika w układzie samoporównawczym

analogiczny układ przetwornika, jak w metodzie porównawczej, przy czym obie jego sekcje obejmują ten sam obiekt badany (rys. 10).

Ze względu na małe prawdopodobieństwo wystąpienia większych różnic składu chemicznego, struktury i wymiarów w sąsiednich miejscach tego samego elementu, przyczyną pojawienia się wskazania są przede wszystkim wady o charakterze nieciągłości materiału. Wadą metody samoporównawczej jest brak wykrywalności wad o stałej głębokości, występujących wzdłuż całej długości obiektu badanego. Możliwe jest jednak wykrywanie wad o łagodnie zmieniającej się głębokości. Należy wówczas odpowiednio dobrać odstęp pomiędzy uzwojeniami wyjściowymi przetwornika.

WYBÓR PRZETWORNIKA POMIAROWEGO

Skuteczność kontroli w dużym stopniu zależy od właściwości zastosowanego przetwornika pomiarowego. Przy wyborze rodzaju przetwornika należy wziąć pod uwagę kształt obiektu badanego, natomiast przy wyborze układu, w którym pracuje przetwornik — punktem wyjścia są mierzone właściwości obiektu. I tak na przykład do ciągłej kontroli rur można zastosować przetwornik przelotowy, a do lokalnej kontroli właściwości materiału rury — przetwornik stykowy [8]. Spotykane są również jednoczesne zastosowania przetworników przelotowych zewnętrznych, wewnętrznych i stykowych w przypadku kontroli grubościennych rur o dużej średnicy. Do kontroli struktury materiału można zastosować przetwornik, pracujący w układzie porównawczym, do wykrywania nieciągłości materiału (defektoskopia) — przetwornik pracujący w układzie samoporównawczym.

Cechą charakterystyczną przetworników przelotowych jest to, że umożliwiają one lokalizację miejsca wadliwego tylko na długości rury, nie dają zaś informacji odnośnie położenia wady na obwodzie. Przetworniki przelotowe mają duże zastosowanie do badania stanu zewnętrznych i wewnętrznych powierzchni rur. Rzadziej natomiast stosowane są przetworniki stykowe, wirujące. Przy zapewnieniu ruchu liniowego rur badanych

kontrola odbywa się po spirali. Wydajność kontroli jest mniejsza, ale istnieje wówczas możliwość oceny głębokości wad.

Kontrola prowadzona za pomocą przetworników przelotowych jest mało czuła na wady równomierne, biegnące wzdłuż całej długości rury oraz na wady zaczynające się i kończące z dala od końca rury, które to wady mają łagodnie zmienną głębokość. Jak wskazuje praktyka, poza pęknięciami naprężeniowymi i nie zgrzanymi spoinami w rurach zgrzewanych i spawanych, spotykane wady mają przeważnie początek lub koniec w obrębie badanego wyrobu oraz wykazują różne nieregularności ułatwiające ich wykrycie. W celu wykrycia wad o równomiernej głębokości można zastosować kombinację metody samoporównawczej z metodą porównawczą albo z metodą wirującego przetwornika stykowego.

Przetworniki przelotowe są odpowiednie do wykrywania wad wzdłużnych, natomiast słabo wykrywają pęknięcia poprzeczne.

DEFEKTOSKOPJA RUR NIEFERROMAGNETYCZNYCH SPAWANYCH

Obecnie zostaną przedstawione niektóre wyniki prac dotyczących problemów badań defektów rur nieferromagnetycznych przy zastosowaniu metody prądów wirowych, zarówno w aspekcie badań podstawowych odnoszących się do:

- opracowania przetworników pomiarowych do badania wad rur i
- określenia warunków pracy przetworników,
- jak i w aspekcie zastosowania opracowanych przetworników do budowy defektoskopów, przeznaczonych do kontroli rur w procesie ich wytwarzania. Przytoczone zostaną wyniki badań dotyczących problemów kontroli rur nieferromagnetycznych spawanych. Przewiduje się bowiem opracowanie defektoskopu do kontroli jakości rur wykonanych z miedzi, mosiądzu (a w przyszłości również i brązu) i wdrożenie go do potokowej kontroli podczas procesu ich wytwarzania.

Ze względu na zamierzone uzyskanie dużej szybkości kontroli rur opracowano modele przetworników wiropądowych przelotowych zewnętrznych, tj. takich, przez które będą prowadzone badane rury.

Konfiguracja opracowanych przetworników została podyktowana przez rodzaj obiektu oraz charakter występujących w nim wad.

Opracowano następujące rodzaje przetworników transformatorowych o quasi-równomiernym polu magnetycznym (składowa wzdłuż osi obiektu):

- przetwornik w układzie stosowanym w metodzie porównawczej oraz
- przetwornik w układzie stosowanym w metodzie samoporównawczej i w metodzie bezwzględnej.

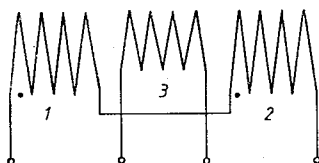
Przetwornik w układzie stosowanym w metodzie porównawczej wykrywa wszystkie rodzaje wad szwu. Natomiast przetwornik w układzie stosowanym w metodzie samoporównawczej, służy do wykrywania wad rur, takich jak dziury i częściowy przetop. Przetwornik taki wykrywa początek i koniec wady obiektu. Przetwornik w układzie stosowanym w metodzie bezwzględnej służy do wykrywania odcinków rur z brakiem przetopu. Dla realizacji przetwornika w układzie bezwzględnym możliwe jest wykorzystanie jednej sekcji przetwornika w układzie samoporównawczym, względnie uzupełnienie przetwornika

w układzie samoporównawczym, przetwornikiem w układzie bezwzględny. Przyjęto drugie rozwiązanie.

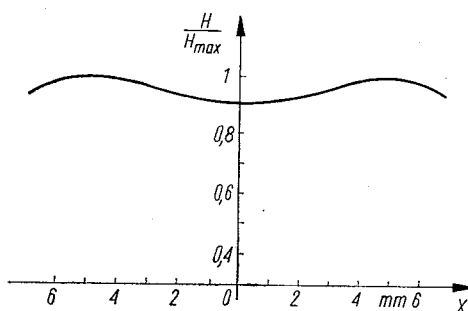
Przetworniki transformatorowe wybrano ze względu na łatwość analizy ich właściwości metrologicznych i eksploatacyjnych, choć możnaby również opracować przetworniki impedancyjne do kontroli jakości szwu rur.

PRZETWORNIK W UKŁADZIE PORÓWNAWCZYM

Do nieniszczących badań rur miedzianych spawanych opracowano przetwornik przetłotowy transformatorowy zewnętrzny o schemacie elektrycznym przedstawionym na rys. 11.



Rys. 11. Schemat opracowanego przetwornika w układzie porównawczym 1, 2 — uzwojenia wzbudzające, 3 — uzwojenie wyjściowe

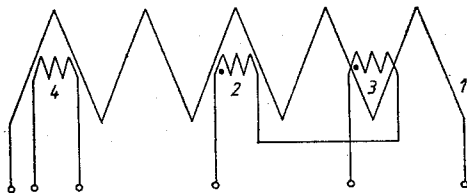


Rys. 12. Rozkład składowej osiowej pola magnetycznego przetwornika w układzie porównawczym

Dwie identyczne sekcje uzwojeń wzbudzających 1 i 2 przetwornika, połączonych szeregowo o strumieniach magnetycznych skierowanych zgodnie, umieszczono po przeciwnych stronach uzwojenia wyjściowego 3. Sygnały wyjściowe realizowanego przetwornika w układzie porównawczym porównywane są z sygnałami odpowiedniego źródła napięcia odniesienia, o regulowanej fazie i amplitudzie napięcia wyjściowego. Opracowany przetwornik charakteryzuje się quasi-równomiernym rozkładem pola magnetycznego wzdłuż osi obiektu. Rozkład pola magnetycznego modelu przetwornika obliczono korzystając z zależności (8) [22]. Rozkład pola magnetycznego przetwornika, przeznaczonego do kontroli rur o średnicy zewnętrznej 8 mm, przedstawiono na rys. 12. Rozkład ten unormowano względem wartości pola dla punktów odpowiadających środkom uzwojeń wzbudzających. W przyszłości przewiduje się opracowanie przetworników o nierównomiernym rozkładzie pola magnetycznego wzdłuż osi rur.

PRZETWORNIK W UKŁADZIE SAMOPORÓWNAWCZYM I BEZWZGLĘDNYM

Opracowany przetwornik w układzie samoporównawczym i bezwzględnym (rys. 13), zawiera trzy uzwojenia wyjściowe, z których dwa symetryczne połączone są szeregowo przeciwsośnie oraz jedno wspólne uzwojenie wzbudające. Uzwojenia wyjściowe przet-

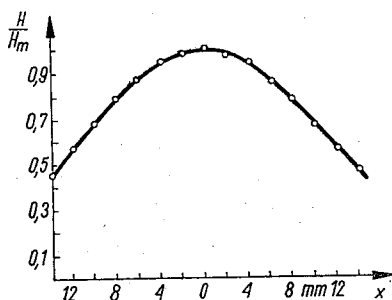


Rys. 13. Schemat opracowanego przetwornika w układzie samoporównawczym i bezwzględnym 1 — uzwojenie wzbudające, 2, 3 — uzwojenie wyjściowe przetwornika samoporównawczego, 4 — uzwojenie wyjściowe przetwornika bezwzględnego

wornika znajdują się możliwie blisko obiektu badanego. Zapewniono w ten sposób możliwie dużą wartość współczynnika wypełnienia przetwornika przez obiekt.

W przypadku konstrukcji przetwornika tzw. wewnętrznego, wsuwanego do wnętrza rur, konfiguracja przetwornika powinna być odwrotna. Ostatnio przetworniki wewnętrzne zyskują coraz większe zastosowanie [5, 16]. Są to zarówno przetworniki transformatorowe, jak i parametryczne, realizowane w układzie samoporównawczym. Przetworniki te stosowane są m.in. do badania stanu rur, wykonanych ze stali austenicznych, zamontowanych w wytwornicach pary w elektrowniach atomowych, a także do kontroli zainstalowanych rurociągów.

Opracowany przetwornik w układzie samoporównawczym charakteryzuje się quasi-równomiernym rozkładem pola magnetycznego wzdłuż osi obiektu. Rozkład pola magne-



Rys. 14. Rozkład składowej osiowej pola magnetycznego przetwornika w układzie samoporównawczym i bezwzględnym

tycznego przetwornika, obliczony w oparciu o zależność (8) [22], przedstawiono na rys. 14. Rozkład pola magnetycznego przetwornika unormowano względem wartości pola dla punktu odpowiadającego środkowi uzwojenia wzbudającego.

Tabela 2

Częstotliwość graniczna dla rur z materiałów nieferromagnetycznych wg kryterium dla rur cienkościennych

Materiał	D_z [mm] G [mm] $\gamma [\times 10^6 \text{S/m}]$	f_g [Hz]													
		6,5	6,5	8,0	10,0	10,0	10,0	12,0	14,0	16,0	16,0	16,0	18,0	18,0	18,0
		0,5	0,6	0,5	0,5	0,6	0,5	0,5	0,5	0,5	0,6	0,6	0,6	1,0	1,2
		—	—	—	—	—	—	—	—	—	—	—	—	—	—
Miedź	59	3 122	2 700	2 453	1 908	1 626	1 561	1 321	1 145	967	613	852	537	459	
Aluminium	35	5 263	4 551	4 136	3 217	2 741	2 632	2 227	1 930	1 630	1 034	1 436	905	773	
PA4	27	6 823	5 900	5 361	4 170	3 554	3 412	2 887	2 502	2 112	1 340	1 861	1 173	1 002	
Mosiądz	15	12 281	10 621	9 650	7 505	6 397	6 141	5 196	4 503	3 803	2 412	3 351	2 111	1 804	
Braz fosforowy	9	20 469	17 700	16 083	12 509	10 661	10 234	8 660	7 505	6 339	4 021	5 584	3 518	3 007	
Braz cynowy	3,6	51 172	44 252	40 206	31 272	26 652	25 586	21 650	18 763	15 847	10 052	13 961	8 795	7 517	

CZĘSTOTLIWOŚĆ PRACY PRZETWORNIKÓW

Dla określenia częstotliwości pracy przetworników przeprowadzono rozważania uwzględniające następujące przesłanki:

- 1) materiały, z których wykonywane są rury badane,
- 2) zakres średnic zewnętrznych badanych rur,
- 3) grubość ścianek badanych rur.

Uwzględniono również następujące ograniczenia. Można, oczywiście, dobrać optymalną częstotliwość pomiarową dla przypadku kontroli rur o poszczególnych średnicach zewnętrznych, określonych grubościach, wykonanych z różnych materiałów. Jednak w związku z przewidywanym zastosowaniem opracowanych przetworników w defektoskopie do kontroli rur, wykonanych z materiałów nieferromagnetycznych, celowy jest wybór jednej lub co najwyżej dwóch częstotliwości pracy i określenie rodzajów rur, jakie będą badane przy poszczególnych częstotliwościach. Zagadnienie to nie jest krytyczne, gdyż jak to wynika z badań modelowych [8] (patrz rys. 1) składowa zmiany napięcia wyjściowego przetwornika, wywołana np. przez nieciągłość, prostopadła do kierunku zmian średnicy zewnętrznej, przy wartościach stosunku częstotliwości pracy do częstotliwości granicznej w zakresie około 0,8—8 zmienia się stosunkowo niewiele.

W tabeli 2 podano wartość częstotliwości granicznej dla rur różnych średnic i różnych grubości ścianek, wykonanych z różnych materiałów nieferromagnetycznych. Obliczenia wykonano według zależności (1) — dla rur cienkościennych. Jak wynika z tej tabeli, przyjmując wartość stosunku częstotliwości pracy do częstotliwości granicznej rur równą 1, częstotliwości pracy przetworników do kontroli rur nieferromagnetycznych powinny zawierać się w zakresie około od 500 Hz do 50000 Hz. Częstotliwości te, dla wyszczegół-

Tabela 3

Częstotliwość graniczna dla rur miedzianych i mosiężnych (spawanych)
wytwarzanych w Polsce — wg zależności dla rur cienkościennych

Materiał	D_z [mm] D_z [mm]	f_g [Hz]			
		0,5	0,6	1,0	1,2
Miedź $\gamma = 59 \cdot 10^6$ S/m	6,5	3122	2700	—	—
	10,0	1908	1626	—	—
	16,0	1145	967	613	—
	18,0	—	852	537	459
Mosiądz $\gamma = 15 \cdot 10^6$ S/m	6,5	12281	10621	—	—
	10,0	7505	6397	—	—
	16,0	4503	3803	2412	—
	18,0	—	3351	2111	1804

Tabela 4

Parametry kontroli rur miedzianych i mosiężnych (spawanych), wytwarzanych w Polsce — wg zależności dla rur grubościennych

Materiał	G	mm	0,5			0,6			1,0			1,2		
			$\frac{G}{D_z}$	$(f/f_0)_{opt}$	$f(f/f_0)_{opt}$	$\frac{G}{D_z}$	$(f/f_0)_{opt}$	$f(f/f_0)_{opt}$	$\frac{G}{D_z}$	$(f/f_0)_{opt}$	$f(f/f_0)_{opt}$	$\frac{G}{D_z}$	$(f/f_0)_{opt}$	$f(f/f_0)_{opt}$
Miedź $\gamma = 59 \cdot 10^6$ S/m	mm	Hz	%	—	Hz	%	—	Hz	%	—	Hz	%	—	Hz
	6,5	203	7,7	15,4	3 126	9,2	13,2	2 680	—	—	—	—	—	—
	10,0	86	5	22,2	1 909	6	18,9	1 625	—	—	—	—	—	—
	16,0	34	3,1	34,1	1 160	3,8	28,8	979	6,3	18,3	622	—	—	—
	18,0	27	—	—	—	3,3	32,1	867	5,6	20,3	548	6,7	17,3	467
Mosiądz $\gamma = 15 \cdot 10^6$ S/m	6,5	799	7,7	15,4	12 305	9,2	13,2	10 547	—	—	—	—	—	—
	10,0	338	5	22,2	7 504	6	18,9	6 388	—	—	—	—	—	—
	16,0	132	3,1	34,1	4 501	3,8	28,8	3 802	6,3	18,3	2 416	—	—	—
	18,0	104	—	—	—	3,3	32,1	3 338	5,6	20,3	2 111	6,7	17,3	1 799

* obliczone według zależności (6) — jak dla próbki pełnej

nionego w tabeli 2 asortymentu rur miedzianych i mosiężnych, wynosiłyby w przybliżeniu od 500 Hz do 12000 Hz, zaś częstotliwości pracy dla rur, wykonanych z brązów, wynosiłyby w przybliżeniu od 3000 Hz do 50000 Hz.

Częstotliwości pracy, odpowiednie do kontroli rur wykonanych z brązów, byłyby kilkakrotnie (zależnie od składu poszczególnych brązów) większe niż częstotliwości pracy, odpowiednie do kontroli rur miedzianych i mosiężnych dla danej średnicy rur i grubości ścianki. Problemy kontroli rur, wykonanych z brązów, zostaną podjęte po uruchomieniu w naszym kraju produkcji tych rur (spawanych); tj. w szczególności — po określeniu gatunków brązów i typoszeręgu średnic i grubości ścianek rur.

Przytoczone niżej wyniki badań odnoszą się do przypadku kontroli rur spawanych, produkowanych obecnie w Polsce. Są to rury miedziane i mosiężne wg głównego asortymentu podanego w tabeli 3. W tabeli podano wartości częstotliwości granicznej dla tych rur. Obliczenia wykonano wg zależności (1) — dla rur cienkościennych. Jeżeli przyjąć wartość stosunku częstotliwości pracy do częstotliwości granicznej równą 1, to częstotliwość pracy przetworników do kontroli rur miedzianych i mosiężnych (spawanych) zawierałaby się w zakresie od około 500 Hz do około 12000 Hz, jak to wynika z tabeli.

W tabeli 4 zestawiono parametry kontroli wiroprowodowej wg zależności podanych w pracy [8], jak dla rur grubościennych. W tabeli tej podano grubość ścianek w procentach średnicy zewnętrznej rur G/D_z , optymalne wartości stosunku częstotliwości pracy przetwornika do częstotliwości granicznej $(f/f_g)_{opt}$, oraz optymalne wartości częstotliwości pracy $f[(f/f_g)_{opt}]$. Dane dotyczą rur miedzianych i mosiężnych (spawanych), wytwarzanych w naszym kraju. Przytoczone w tabeli 4 wartości częstotliwości granicznej f_g obliczono wg zależności (6) — jak dla próbek pełnych — zgodnie z wyżej przytoczonym sposobem analizy właściwości przetworników.

Jak widać z porównania danych zestawionych w tab. 3 i tab. 4 otrzymano całkowitą zbieżność wartości optymalnej częstotliwości pracy przetworników wg zależności dla rur grubościennych i częstotliwości pracy, przy przyjęciu stosunku częstotliwości pracy do częstotliwości granicznej f/f_g równej 1, obliczonej wg zależności dla rur cienkościennych.

WŁAŚCIWOŚCI METROLOGICZNE I EKSPLOATACYJNE PRZETWORNIKÓW

W celu określenia właściwości metrologicznych i eksploatacyjnych przetworników, skonstruowanych z przeznaczeniem do kontroli rur, opracowano metodykę analizy przetworników. W wyniku analizy właściwości przetworników można otrzymać, przedstawione na płaszczyźnie zmiennej zespolonej, unormowane składowe napięcia wyjściowego przetworników.

METODYKA BADAŃ

Właściwości metrologiczne przetwornika mogą być określone na podstawie:

- zmian wartości bezwzględnej napięcia wyjściowego przetwornika ΔX_R , wywołanej istnieniem wady w obiekcie,
- zmian wartości składowej napięcia wyjściowego przetwornika ΔX_{R1} , prostopadłej do

wartości bezwzględnej zmian napięcia wyjściowego przetwornika ΔX_S , wywołanej zmianą średnicy zewnętrznej rury,

- przebiegu kąta fazowego Φ pomiędzy wielkością ΔX_R , spowodowanej istnieniem wady w rurze i wielkością ΔX_S , wywołaną zmianą średnicy zewnętrznej rury, przy czym

$$\Delta X_{R1} = \Delta X_R \sin \Phi. \quad (9)$$

Analiza powyższych wielkości pozwala na zbadanie relacji warunki pracy przetwornika — czułość wykrywania wad obiektu (niezależnie od zmian średnicy zewnętrznej rur). Opracowany sposób analizy umożliwia znalezienie takiej częstotliwości pracy przetwornika do kontroli rur, dla której z dużą czułością będą wykrywane wady typu nieciągłości, a jednocześnie stworzone zostaną przesłanki składowej (układem elektronicznym) eliminacji wpływu zmian rozmiarów rur.

Wymienione wielkości można określić na podstawie pomiarów dwóch składowych napięcia wyjściowego przetwornika:

1. Składowej U_R (dla przetwornika zawierającego obiekt badany) o fazie zgodnej z fazą prądu I_w zasilającego przetwornik,
2. Składowej U_I (dla przetwornika zawierającego obiekt badany), przesuniętej względem prądu I_w o kąt fazowy równy $\pi/2$.

Odpowiednie składowe napięcia przetwornika pustego oznaczono odpowiednio jako U_{R0} i U_{I0} . Unormowane składowe napięcia wyjściowego przetwornika obliczono z następujących zależności

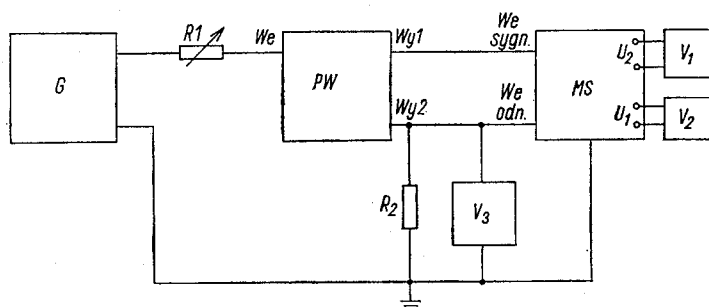
$$ReX = \frac{U_R - U_{R0}}{U_{I0}}, \quad ImX = \frac{U_I}{U_{I0}}. \quad (10)$$

Pomiary i obliczenia wykonano w zakresie częstotliwości od 1 kHz do 10 kHz przy stałej wartości prądu zasilającego przetwornika dla poszczególnych częstotliwości prądu zasilającego. Składowe napięcia wyjściowego przetwornika mierzono przy użyciu miernika składowych (Transfer function analyzer), typ 272 produkcji Unipan, którego maksymalna częstotliwość pomiarowa wynosi 10 kHz. Nieprzydatnym do tego celu okazał się miernik składowych (Resolved Component Indicator), model VP 250.2 produkcji firmy Solartron, umożliwiający wprowadzić pomiary w zakresie częstotliwości do 100 kHz ale nie mający możliwości oddzielnej regulacji czułości obu kanałów. Miernik składowych, typ VP 250.2 firmy Solartron, przydatny jest do pomiarów składowych napięcia wyjściowego przetworników, zawierających obiekty wykonane z materiałów ferromagnetycznych, gdyż w tym przypadku składowe te są tego samego rzędu wielkości [19]. Natomiast składowe napięcia wyjściowego przetwornika zawierającego obiekt (rurę, pręt) wykonany z materiału nieferromagnetycznego, różnią się co najmniej o rząd wielkości (patrz np. rys. 17). Schemat układu pomiarowego przedstawiono na rys. 15.

W y k a z p r z y r z ą d ó w:

- G — generator wewnętrzny miernika składowych,
 PW — przetwornik wirowprądowy,
 MS — miernik składowych, typ 272 prod. Unipanu,
 R₁ — rezystor dekadowy, typ MDR-93-5b prod. Inco,

R_2 — rezystor typ MŁT, 10Ω ,
 V_1, V_2 — woltomierze cyfrowe typ V560 prod. Meratroniku,
 V_3 — woltomierz cyfrowy V541, prod. Meratroniku.

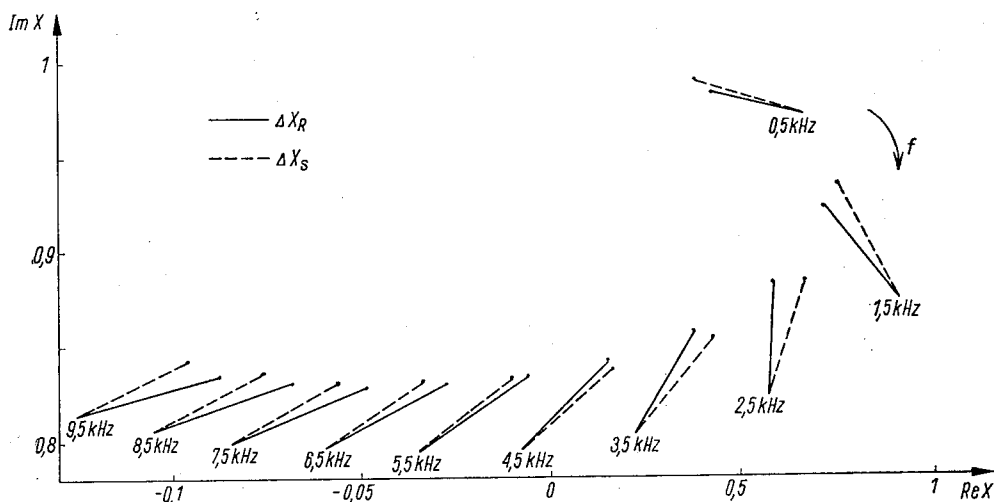


Rys. 15. Układ do pomiaru składowych napięcia wyjściowego przetworników wiroprowadowych

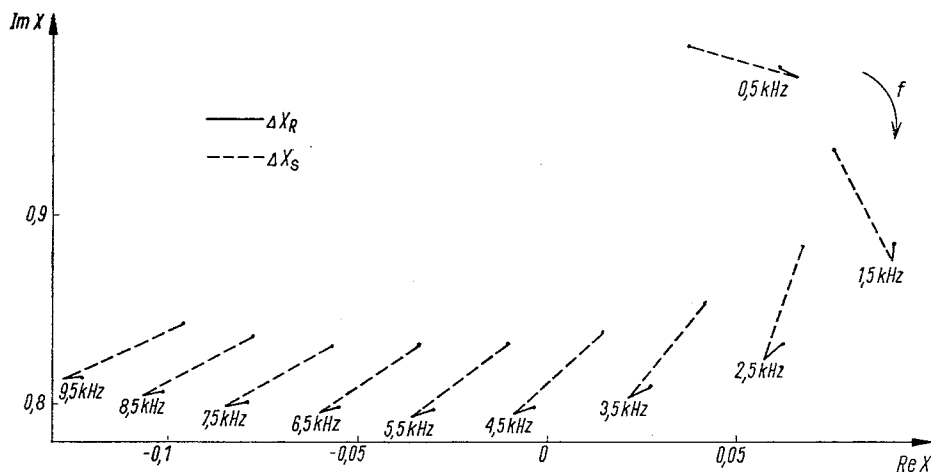
WYNIKI BADAŃ

Na rys. 16—33 przedstawiono wyniki badań przetworników zrealizowanych w układzie stosowanym w metodzie bezwzględnej. Wykorzystano w tym celu uzwojenia wyjściowe 3 przetworników o schemacie pokazanym na rys. 11. Przetworniki opracowano z przeznaczeniem do kontroli rur średnic w zakresie od 6,5 do 20 mm. W rurach wykonano piłką tarczową wady sztuczne w postaci rowków wzdłużnych o różnych głębokościach; również przelotowe. Miało to na celu zmodelowanie wad naturalnych (częściowy przetop, dziury, braki przetopu). Na określonych, pozbawionych wad, odcinkach rur zmniejszono ich średnicę zewnętrzną.

Współrzędne punktów opisanych wartością częstotliwości prądu zasilającego przetwornik na wykresach na płaszczyźnie zmiennej zespolonej (rys. 16, 17, 21, 24, 25, 29,



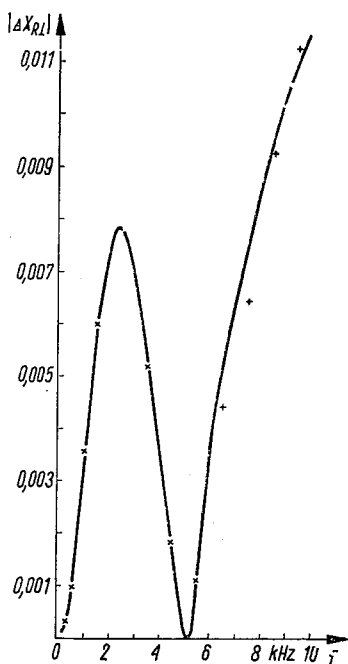
Rys. 16. Zespolone napięcie przetwornika zawierającego rurę miedzianą średnicy 8 mm i grubości ścianki 1,1 mm z wadą sztuczną przelotową i zmniejszeniem średnicy do 7,2 mm



Rys. 17. Zespolone napięcie przetwornika zawierającego rurę miedzianą średnicy 8 mm i grubości ścianki 1,1 mm z wadą sztuczną nieprzelotową i zmniejszeniem średnicy do 7,2 mm

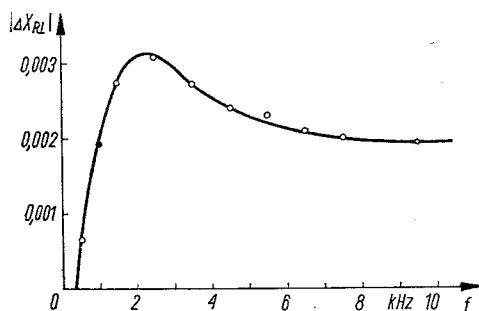
30) przedstawiają odpowiednio $Re X$ i $Im X$ wg zależności (10) dla przetwornika zawierającego odcinek rury o średnicy nominalnej, bez wady. Wielkość ΔX_R przedstawia zmianę wartości X , wywołaną istnieniem wady w rurze, natomiast wielkość ΔX_S odpowiada zmianie wartości X , wynikającej ze zmniejszenia średnicy odcinka rury.

Na rys. 18, 19, 22, 26, 27, 31, 32 przedstawiono rzut modułu unormowanej składowej



Rys. 18. Rzut modułu składowej unormowanego napięcia przetwornika, wywołanej wadą na kierunku prostopadły do kierunku średnicy zewnętrznej. Rura miedziana średnicy 8 mm i grubości ścianki 1,1 mm.

Wada przelotowa

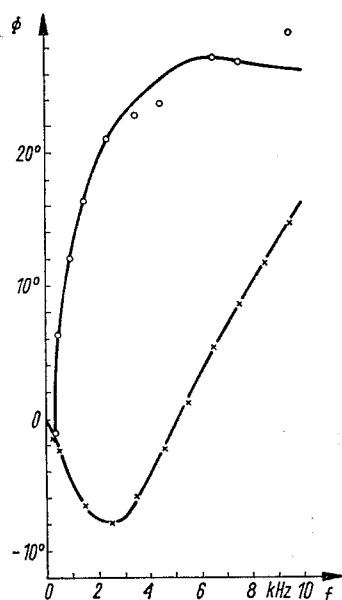


Rys. 19. Rzut modułu składowej unormowanego napięcia przetwornika wywołanej wadą, na kierunek prostopadły do kierunku średnicy zewnętrznej. Rura miedziana średnicy 8 mm i grubości ścianki 1,1 mm. Wada nieprzelotowa

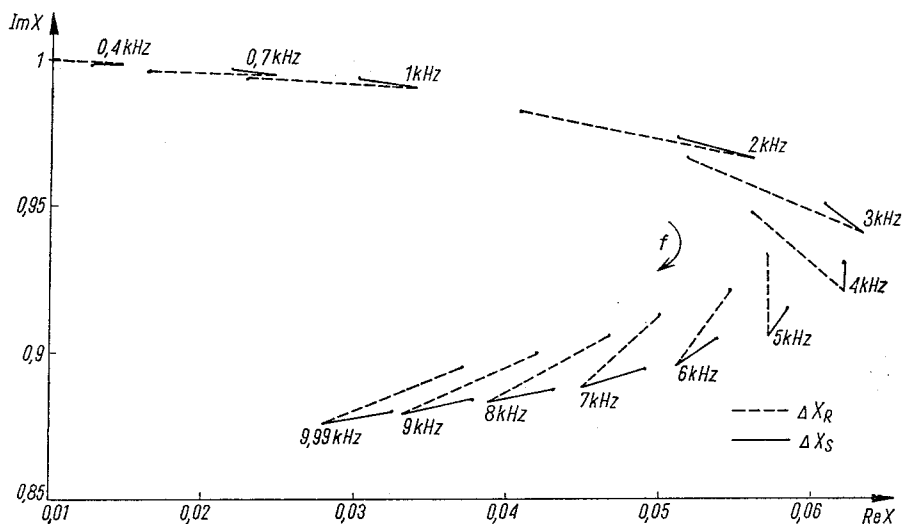
napięcia przetwornika, pochodzącej od wady na kierunek prostopadły do kierunku D_z (pod względem fazowym kierunek, w którym uzewnętrzniają się na płaszczyźnie zmiennej zespolonej zmiany pochodzące od zmian średnicy zewnętrznej rur).

Na rys. 20, 23, 28, 33 pokazano przebieg kąta fazowego pomiędzy unormowanymi składowymi napięć przetwornika, wywołanymi przez wady i zmiany średnicy rur.

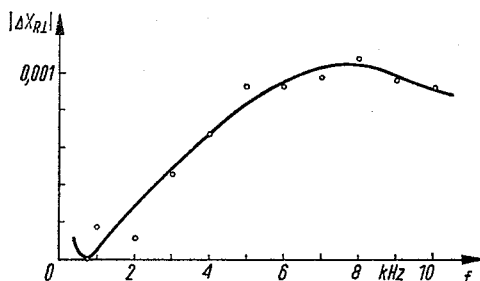
W celu dokonania oceny wykrywalności wad szwu przez opracowany przetwornik w układzie stosowanym w metodzie porównawczej (rys. 11) przeprowadzono pomiary napięcia wyjściowego przetwornika dla rur z wadami naturalnymi (brak przetopu, dziury



Rys. 20. Kąt fazowy pomiędzy unormowanymi: napięciem przetwornika, pochodzącym od wady i napięciem wywołanym przez zmianę średnicy. Rura miedziana średnicy 8 mm i grubości ścianki 1,1 mm. Wady przelotowa i nieprzelotowa. Zmniejszenie średnicy do 7,2 mm



Rys. 21. Zespolone napięcie przetwornika zawierającego rurę mosiężną średnicy 12 mm i grubości ścianki 0,9 mm, z wadą nieprzelotową i zmniejszeniem średnicy do 11,5 mm

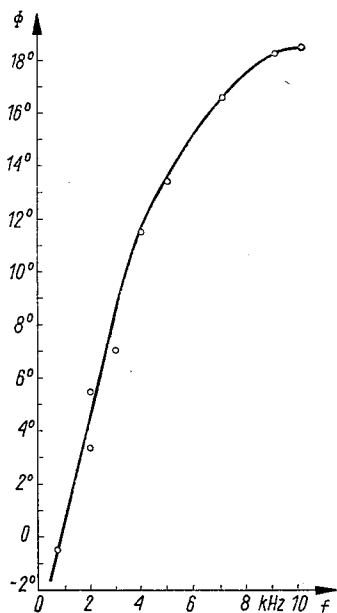


Rys. 22. Rzut modułu składowej unormowanego napięcia przetwornika, wywołanej wadą, na kierunku prostopadły do kierunku średnicy zewnętrznej. Rura mosiężna średnicy 12 mm i grubości ścianki 0,9 mm, z wadą nieprzelotową i zmniejszeniem średnicy do 11,5 mm

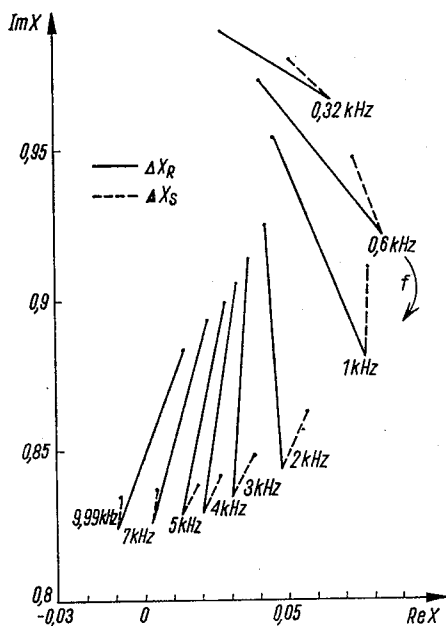
pochodzące od wypaleń elektrodą, częściowy przetop) oraz z wadami sztucznymi (frezowane kanały nieprzelotowe oraz przelotowe wzdłużne, a także zmiany średnicy zewnętrznej). Pomiary przeprowadzono w układzie detekcji amplitudowej, przedstawionym na rys. 34.

Na rys. 35, 36 pokazano przykładowe wyniki badań przetwornika zawierającego rury miedziane o średnicy 8 mm; na rys. 35 — dla rury z wadami sztucznymi; na rys. 36 — z wadami naturalnymi w postaci dziur, pochodzących z wypaleń elektrodą. Rury spawane o średnicy 8 mm są również produkowane w naszym kraju. Badania przeprowadzono w zakresie częstotliwości prądu zasilającego przetwornik od 1 kHz do 50 kHz.

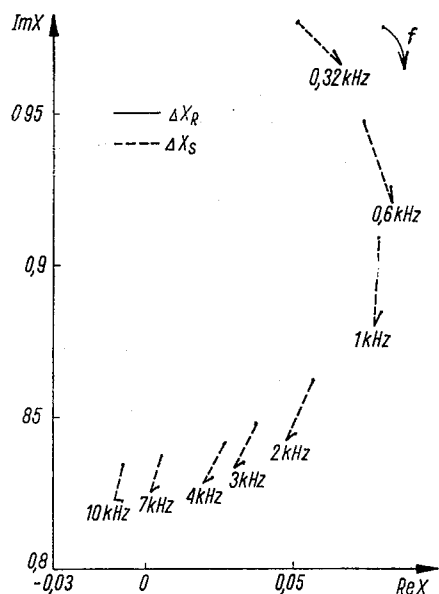
Na rys. 37 pokazano przykładowe wyniki badań siedmiu rur miedzianych spawanych z wadami naturalnymi w postaci dziur o różnej długości, spowodowanych wypaleniem elektrodą. Średnica rur wynosiła 8 mm. Częstotliwość prądu zasilającego przetwornik była równa 5 kHz.



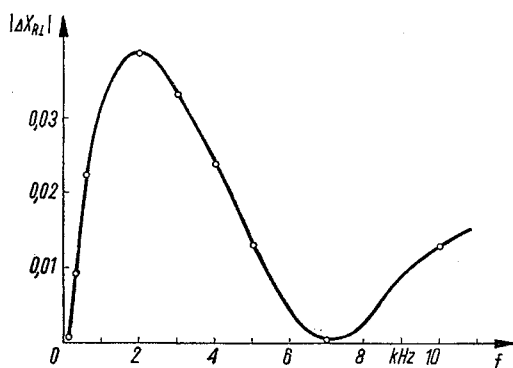
Rys. 23. Kąt fazowy pomiędzy unormowanymi: napięciem przetwornika, pochodzącym od wady i napięciem wywołanym przez zmiany średnicy. Rura mosiężna średnicy 12 mm i grubości ścianki 0,9 mm z wadą nieprzelotową i zmniejszeniem średnicy do 11,5 mm



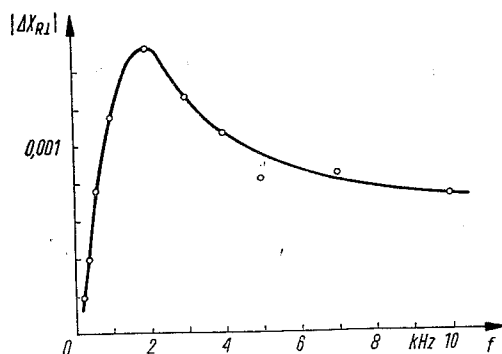
Rys. 24. Zespółone napięcie przetwornika zawierającego rurę miedzianą średnicy 16 mm i grubości ścianki 1,0 mm z wadą przelotową i zmniejszeniem średnicy do 15,4 mm



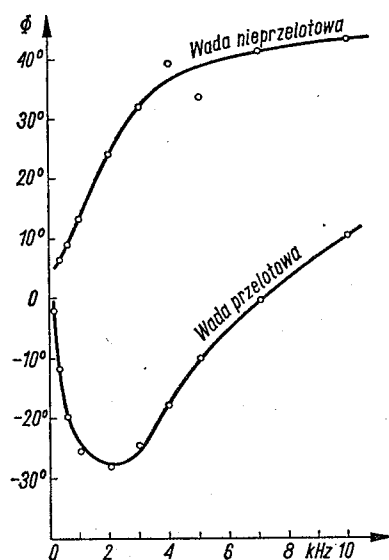
Rys. 25. Zespolone napięcie przetwornika zawierającego rurę miedzianą średnicy 16 mm i grubości ścianki 1,0 mm z wadą nieprzetłową i zmniejszeniem średnicy do 15,4 mm



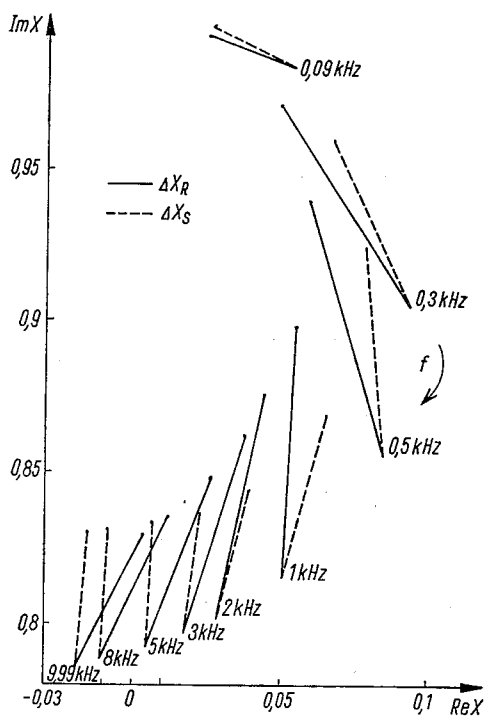
Rys. 26. Rzut modułu składowej unormowanego napięcia przetwornika, wywołanej wadą, na kierunku prostopadły do kierunku średnicy zewnętrznej. Rura miedziana średnicy 16 mm i grubości ścianki 1,0 mm, z wadą przetłową i zmniejszeniem średnicy do 15,4 mm



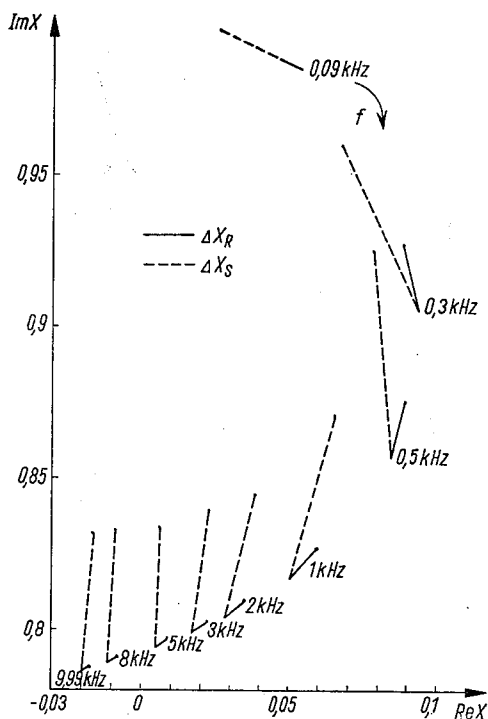
Rys. 27. Rzut modułu składowej unormowanego napięcia przetwornika, wywołanej wadą, na kierunek prostopadły do kierunku średnicy zewnętrznej. Rura miedziana średnicy 16 mm i grubości ścianki 1,0 mm z wadą nieprzelotową i zmniejszeniem średnicy do 15,4 mm



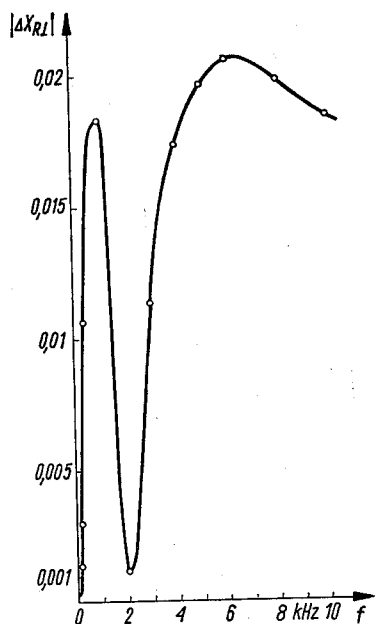
Rys. 28. Kąt fazowy pomiędzy unormowanymi: napięciem przetwornika, pochodzącym od wady i napięciem, wywołanym przez zmiany średnicy. Rura miedziana średnicy 16 mm i grubości ścianki 1,0 mm, z wadami przelotową i nieprzelotową i zmniejszeniem średnicy do 15,4 mm



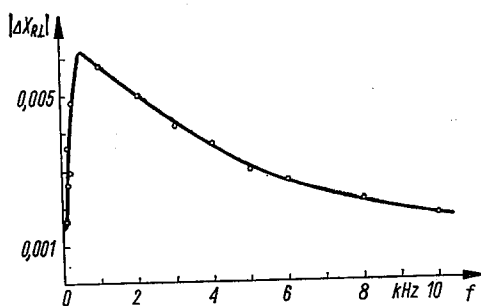
Rys. 29. Zespółone napięcie przetwornika zawierającego rurę miedzianą średnicy 20 mm i grubości ścianki 2, 0 mm z wadą przelotową i zmniejszeniem średnicy do 18,1 mm



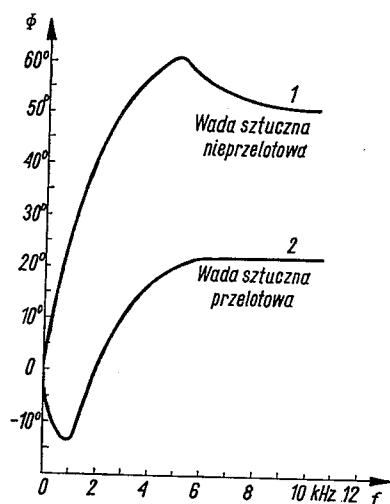
Rys. 30. Zespółone napięcie przetwornika zawierającego rurę miedzianą średnicy 20 mm i grubości ścianki 2,0 mm z wadą nieprzelotową i zmniejszeniem średnicy do 18,1 mm



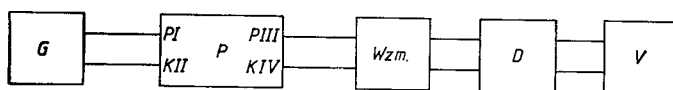
Rys. 31. Rzut modułu składowej unormowanego napięcia przetwornika, wywołanej wadą, na kierunek prostopadły do kierunku średnicy zewnętrznej. Rura miedziana średnicy 20 mm i grubości ścianki 2,0 mm, z wadą przelotową i zmniejszeniem średnicy do 18,1 mm



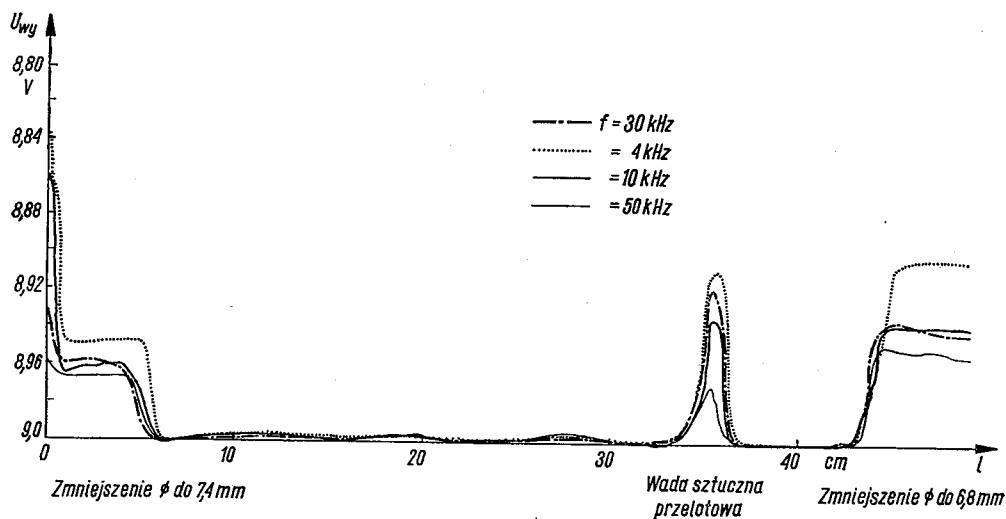
Rys. 32. Rzut modułu składowej unormowanego napięcia przetwornika, wywołanej wadą, na kierunek prostopadły do kierunku średnicy zewnętrznej. Rura miedziana średnicy 20 mm i grubości ścianki 2,0 mm z wadą nieprzelotową i zmniejszeniem średnicy do 18,1 mm



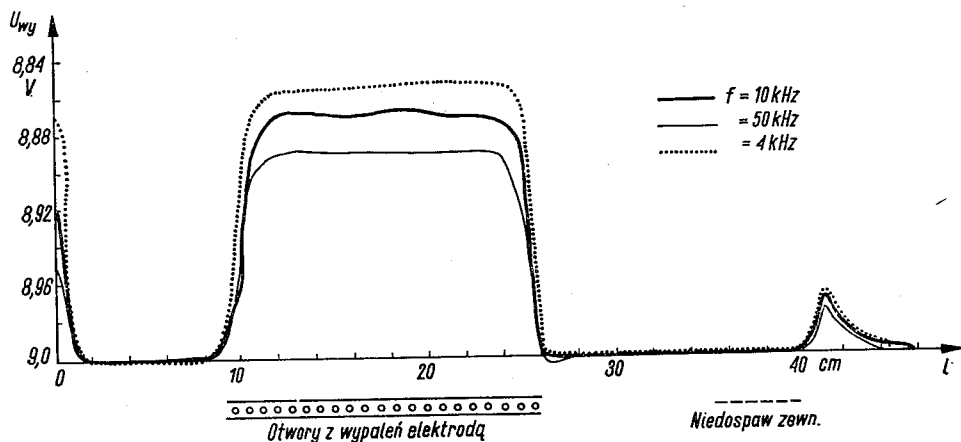
Rys. 33. Kąt fazowy pomiędzy unormowanymi: napięciem przetwornika, pochodzącym od wady i napięciem wywołanym przez zmiany średnicy. Rura miedziana średnicy 20 mm i grubości ścianki 2,0 mm z wadami przelotową i nieprzelotową i zmniejszeniem średnicy do 18,1 mm



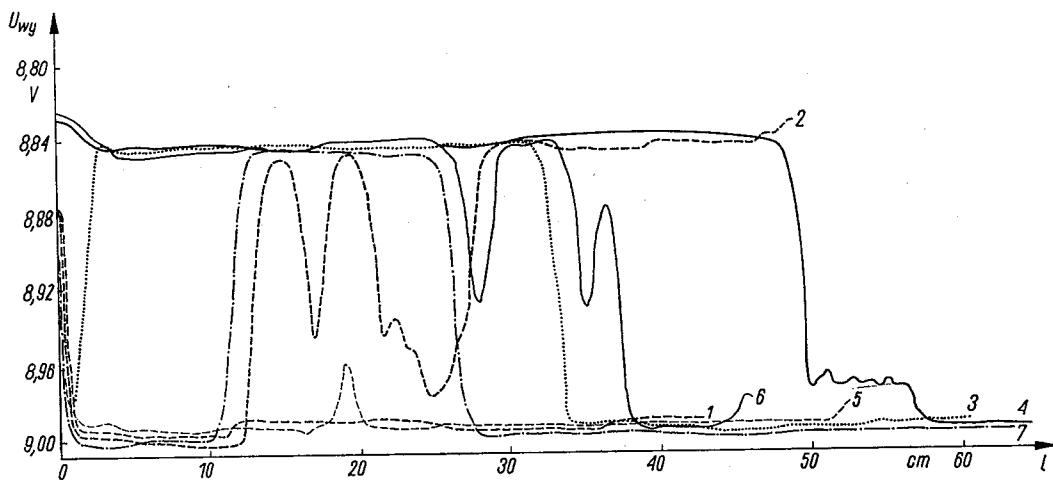
Rys. 34. Schemat blokowy układu pomiarowego, G — generator funkcyjny typ F51A produkcji IEC Corporation, V — woltomierz cyfrowy typ V541, produkcji Meratroniku, Wzm — wzmacniacz, D — detektor szczytowy, P — przetwornik w układzie porównawczym



Rys. 35. Wyniki badania przetwornika, zawierającego rurę miedzianą ciągnioną średnicy 8 mm z wadami sztucznymi w postaci kanałów wzdłużnych i z ubytkami średnicy



Rys. 36. Wyniki badań przetwornika zawierającego rurę miedzianą spawaną średnicy 8 mm z dziurami i częściowym przetopem



Rys. 37. Wyniki badania przetwornika zawierającego rury miedziane spawane średnicy 8 mm; z dziurami. Częstotliwość prądu zasilającego przetwornik wynosiła 5 kHz

Opis właściwości metrologicznych i eksploatacyjnych przetwornika w układzie samoporównawczym i bezwzględny będą przedmiotem drugiej części artykułu.

GŁĘBOKOŚĆ WNIKANIA PRĄDÓW WIROWYCH

Głębokość wnikania prądów wirowych można obliczyć z następującej zależności [8]

$$\delta = \frac{503}{\sqrt{f\gamma\mu_r}} \quad [\text{mm}] \quad (11)$$

przy czym:

f — częstotliwość pracy przetwornika, Hz,

γ — przewodność właściwa materiału, Sm,

μ_r — przenikalność magnetyczna względna.

Dla głębokości wnikania, obliczonej według zależności (11) natężenie prądów wirowych w próbce badanej maleje e -krotnie w stosunku do natężenia na powierzchni próbki.

Tabela 5

Głębokość wnikania prądów wirowych dla miedzi i mosiądzu

Materiał	f	kHz	1	2	3	4	5	10	20	50
Miedź	δ	mm	2,07	1,46	1,19	1,04	0,93	0,66	0,46	0,29
Mosiądz	δ	mm	4,11	2,90	2,37	2,05	1,84	1,30	0,92	0,58

W tabeli 5 podano głębokość wnikania prądów wirowych dla miedzi ($\gamma = 59 \cdot 10^6$ S/m) i mosiądzu ($\gamma = 15 \cdot 10^6$ S/m). Jak widać z tabeli w zakresie częstotliwości do około 10 kHz, przy przyjęciu tak określonej głębokości wnikania prądów wirowych, badany jest cały przekrój rur miedzianych i mosiężnych.

ZAKOŃCZENIE

Z przeprowadzonych badań wynika, co następuje.

- Otrzymano bardzo dobrą wykrywalność wad szwu (dziury, braki przetopu) badanych rur miedzianych o średnicy 8 mm za pomocą opracowanego przetwornika, w układzie stosowanym w metodzie porównawczej. Odpowiednio mniejsze napięcie wyjściowe przetwornika otrzymane dla miejsc rury z częściowym przetopem. I tak np. przy określonym wzmocnieniu układu detekcji sygnałów, dla danej częstotliwości prądu zasilającego przetwornik, czułość przetwornika dla rur z częściowym przetopem jest około 5-krotnie mniejsza niż dla rur z dziurami i brakiem przetopu dla rur z częściowym przetopem otrzymano różnicę napięć ok. 30 mV, dla dziur — około 150 mV w stosunku do napięć przetwornika dla rury bez wad, jak to wynika z rys. 36. Możliwe jest więc rozróżnienie tych wad.
- Jak wynika z obliczeń przedstawionych w tab. 3 i 4 dla rur miedzianych i mosiężnych (spawanych), wytwarzanych obecnie w Polsce, otrzymano całkowitą zbieżność optymalnej częstotliwości pracy przetwornika wg zależności dla rur grubościennych ($G \geq 0,5\% D_z$) i częstotliwości pracy wg zależności, jak dla rur cienkościennych (przy założeniu wartości stosunku częstotliwości pracy do częstotliwości równej jedności). Zgodnie z przytoczonymi obliczeniami, częstotliwości pracy przetworników, przeznaczonych do kontroli rur miedzianych i mosiężnych, produkowanego obecnie asortymentu, zawierają się w zakresie od około 500 Hz do około 12000 Hz. Dla powyższych częstotliwości badany jest praktycznie cały przekrój rur (przy przyjęciu jako kryterium głębokości wnikania prądów wirowych).

3. Dla rur miedzianych o średnicy 8 mm moduł unormowanej składowej napięcia wyjściowego przetwornika, pochodzącej od wad ma maksimum dla częstotliwości prądu zasilającego przetwornik w zakresie 1—5 kHz.
4. W zależności od wielkości ubytku średnicy zewnętrznej rur otrzymano (co do modułu) sygnały wyjściowe przetwornika porównywalne lub nawet większe od sygnałów dla odcinków rur z wadami przelotowymi. Wobec tego konieczne jest zastosowanie układu fazoczułej analizy sygnałów otrzymywanych z przetwornika.
5. Kąt fazowy pomiędzy składową napięcia wyjściowego przetwornika, pochodzącą od wad i składową, pochodzącą od zmiany średnicy, jak to wynika z badań na modelach rur [8], przy wzroście częstotliwości pracy przetwornika rośnie i dąży do około 60° dla wartości stosunku pracy do częstotliwości granicznej $f/f_g = 8$. Dla badanych rur miedzianych i mosiężnych dla częstotliwości prądu zasilającego przetwornik $f = 10$ kHz, dla przypadku wad nieprzelotowych ten kąt fazowy osiąga dopiero wartość około 45° . Dla rur z wadami przelotowymi odpowiedni kąt fazowy jest zawsze mniejszy.

Przy zastosowaniu detekcji fazoczułej sygnałów, uzyskiwanych z przetwornika, możliwe jest uniezależnienie wskazań defektoskopu od zmian średnicy zewnętrznej rur. W tym przypadku należałoby przyjąć odpowiednio większe częstotliwości pracy przetwornika, dla których powyższy kąt fazowy jest większy. Należy się przy tym liczyć ze zmniejszeniem wartości napięcia wyjściowego przetwornika, wywołanego przez wady. Z kolei przyjęcie częstotliwości pracy defektoskopu odpowiadających maksimum składowej unormowanego napięcia przetwornika, prostopadłej do składowej wywołanej zmianą średnicy utrudniłoby praktyczną realizację fazowej analizy sygnałów uzyskiwanych z przetwornika dla rur z wadami przelotowymi (dla częstotliwości pracy do ok. 5 kHz kąt fazowy pomiędzy składową napięcia wyjściowego, pochodzącą od wad i składową, pochodzącą od zmiany średnicy jest ujemny).

6. Na podstawie przeprowadzonych badań oraz rozważań należy przyjąć częstotliwość pracy defektoskopu, przeznaczonego do kontroli rur miedzianych i mosiężnych spawanych równą 10 kHz.
7. Niezbadany odcinek rur wynosi około 20 mm z każdej strony. Ten efekt brzegowy można zmniejszyć poprzez odpowiednie ekranowanie przetwornika.

BIBLIOGRAFIA

1. Boulanger, J. Dufayet, A. Samoel, B. Spriet, A. Stossel: *Automatic non-destructive testing system for thin tubes*. Materials Evaluation, 1974, 32, Nr 1, str. 18—24
2. B. Brown: *The eddy current slide rule*. Materials Evaluation, 1968, 26, Nr 6, str. 120—123
3. T. Busse, N. Beyer: *Pulsed eddy current inspection of thinwalled reactor fuel tubing*. Materials Evaluation, 1970, 28, Nr 10, str. 228—236
4. M. Curati, S. Gabbicini, F. Giovanelli: *Eddy current examination of the tubing of power plant feedwater preheaters*. 3 rd European Conference on Nondestructive Testing, Firenze 1984, Conference Proceedings, t. 2, str. 89—89/10
5. C. De Michelis, C. Sala: *Eddy current internal probe for in service inspection of heat exchanger carbon steel tubes*. 3 rd European Conference on Nondestructive Testing, Firenze 1984, Conference Proceedings, t. 2, str. 67

6. W. G. Gierasimow: *Czuwstwielnost' pnachodnykh datchikow pri kontrole parametrov mnogo-slojnykh cilindricheskikh izdelij*. Defektoskopiya, 1969, Nr 3, str. 82—92
7. J. Hellet, G. Van Druenen, V. Cecco: *An eddy current probe for separating defects from resistivity variations in zirconium alloy tubes*. Materials Evaluation, 1984, 42, Nr 10, str. 1276—1280
8. H. Heptner, H. Stroppe: *Magnetyczne i indukcyjne badania metali*. Katowice, Wyd. Śląsk; 1972
9. R. Hochschild: *Electromagnetic testing of tubing: standardization of specifications and test methods*, Nondestructive Testing, 1958, 16, Nr 6, str. 495—499
10. B. Holler: *Separation of parameters in the eddy current testing of tubes using a multifrequency technique*. Materialprüfung, 1974, 16, Nr 12, str. 391—392
11. T. Judd: *Orbitest for round tubes*. Materials Evaluation, 1970, 28, Nr 1, str. 8—12
12. A. Julier: *Recent developments in the eddy current examination of air conditioner tubing*. Materials Evaluation, 1985, 42, Nr 13, str. 1655—1659
13. A. Lewińska-Romicka: *Devices for eddy-currents testing of cracks in steel wires and bars*. 3rd European Conference on Nondestructive Testing, Firenze, 1984, Conference Proceedings, t. 4, str. 100—109
14. A. Lewińska-Romicka: *Eddy current wire crack detectors*. XXXI Annual Gathering JUREMA, Zagreb — Plitvicka Jezera, 1986, Jurema Proceedings, t. 3, str. 53—56
15. A. Lewińska-Romicka: *Nieniszczące wykrywanie defektów drutów i prętów niemagnetycznych i magnetycznych metodą prądów wirowych*. Aparatura Naukowa i Dydaktyczna, 1984, Nr 3 (37), str. 11—16
16. A. Lewińska-Romicka: *Problemy nieniszczących badań rur niemagnetycznych metodą prądów wirowych*. Rozprawy Elektrotechniczne, 1986, 32, Nr 3, str. 801—812
17. A. Lewińska-Romicka: *Przetwornik wiroprowodowy do kontroli rur spawanych z materiałów nieferromagnetycznych*. Sympozjum Metrologia'86 Metrologia wielkości mechanicznych i kontrola jakości produkcji, Warszawa, 1986, Referaty i komunikaty, t. II, str. 334—343
18. A. Lewińska-Romicka: *The eddy current inspection of nonmagnetic and magnetic wires*. 4th International Symposium on Technical Diagnostics, Kupari-Dubrownik, 1986, Proceedings, str. 555—558
19. A. Lewińska-Romicka, S. Piskorski: *Badania prętów stalowych defektoskopami DW 21÷DW 23*. XV Krajowa Konferencja Badań Nieniszczących, Rynia, 1986, Materiały Konferencji, str. 277—282
20. I. Nikitin, *Kontrol tolszcziny stienki trub s pomoszczu prachodnykh priobrazowatelej*. Defektoskopiya, 1978, Nr 3, str. 53—62
21. A. T. Nikołajenko: *Kontrol trub w adnarodnom popierecznom elektromagnitnom pole*. Defektoskopiya 1969, Nr 3, str. 67—74
22. M. Nałęcz, J. Jaworski: *Pomiary magnetyczne*. Warszawa—Łódź: PWN, 1965
23. D. Roach: *Defects and damage in cast heater tubes*. Materials Evaluation, 1972, 30, Nr 11, str. 229—235
24. J. Rudlin: *Developments in the eddy current inspection of heat exchanger tubing*. 3 rd European Conference on Nondestructive Testing, Firenze, 1984, Conference Proceedings t. 4, str. 79—89
25. J. Ryden, Jr: *Nondestructive testing of small — diameter stainless steel fuel clad tubing*. Materials Evaluation, 1970, 28, Nr 3, str. 67—71
26. W. Solnica, A. Lewińska-Romicka, K. Dydyński, M. Owsiany: *Układ cyfrowej rejestracji wad do defektoskopów*. Patent PRL Nr 103 326, 1908.03.18
27. W. Solnica, A. Lewińska-Romicka, K. Dydyński, M. Owsiany: *Przetwornik pomiarowy do badania elementów ferromagnetycznych*. Patent PRL Nr 113 703, 1982.10.25
28. W. Solnica, A. Lewińska-Romicka, M. Owsiany: *Układ do wykrywania wad drutów z materiałów przewodzących prąd elektryczny*. Patent PRL Nr 118 366, 1984.03.22
29. W. Solnica, A. Lewińska-Romicka, S. Piskorski: *Sposób wzorcowania detektora fazy w defektoskopach wiroprowodowych*. Patent PRL Nr 124 916, 1986.03.10
30. W. Solnica, A. Lewińska-Romicka, K. Dydyński, S. Piskorski, M. Owsiany: *Defektoskop do badania elementów ferromagnetycznych*. Patent PRL Nr 113 172, 1983.01.26

31. W. Solnica, A. Lewińska-Romicka, S. Piskorski: *Urządzenie do nieniszczących badań jakości i znakowania wad materiałowych w drutach niemagnetycznych*. Rozprawy Elektrotechniczne, 1980, 26, Nr 3, str. 609—615
32. J. Spanner: *An eddy current technique for measuring the tube — to — tube annulus in a nuclear test reactor*. Nondestructive Testing, 1962, 20, Nr 6, str. 334—397
33. W. Stumm: *New developments in the eddy current testing of hot wire and hot tube*. Materials Evaluation, 1971, 29, Nr 7, str. 141—147
34. W. J. Szaternikow, W. A. Denisow: *Wysokotemperaturnyje wichirotokowyje datcziki*. 1968, Nr 3, str. 31—34
35. W. I. Wołkow: *Izuczenije wazmożnosti apriedielenija głubiny powierzchnostnych trieszczin w paroprowodnych trubach iz niemagnitnoj stali*. Defektoskopija, 1968, Nr 2, str. 1—5
36. W. I. Wołkow: *Wichirotokowyj pribor dla apriedielenija głubiny trieszczin w swarnych sajedinienjach paroprowodnych trub z austenitnoj stali*. Defektoskopija, 1968, Nr 3, str. 25—31

A. LEWIŃSKA-ROMICKA

EDDY-CURRENT TRANSDUCERS FOR NON DESTRUCTIVE TESTING OF WELDED NONFERROUS TUBES

Summary

Eddy current through-coil transformer transducers with a quasi-uniform magnetic field distribution and their service conditions when the latter are applied for the detection of tube discontinuity, are presented. Test results are quoted.

A. LEWIŃSKA-ROMICKA

TRANSDUCTEURS A COURANTS DE FOUCAULT POUR LE CONTROLE NON-DESTRUCTIF DES TUYAUX NON FERROMAGNÉTIQUES SOUDÉS L'PARTIE

Résumé

Dans l'article on a décrit les transducteurs transformatiques encerclants à quasi-uniforme répartition du champ magnétique, ainsi que le mode du choix des conditions du travail des transducteurs, destinés à la détection des discontinuités des soudures des tuyaux. On a présenté les résultats des essais.

A. LEWIŃSKA-ROMICKA

WIRBELSTROMUMSETZER FÜR ZERSTÖRUNGSFREIE PRÜFUNGEN SCHMELZSCHWEIßBARER NICHTFERROMAGNETISCHER RÖHRE

Zusammenfassung

Im Artikel wurden Wirbelstrom-Durchlauf-Umspannumsetzer beschrieben mit quasi-gleichmäßiger Verteilung des Magnetfeldes sowie eine Methode für die Zusammensetzung der Betriebsverhältnisse für Umsetzer, die für die Detektion der Unkontinuerlichkeit von Rohnähten bestimmt sind. Prüfungsergebnisse wurden angeführt.

А. ЛЕВИНЬСКА-РОМИЦКА

**ВИХРЕТОКОВЫЕ ПЕРЕХОДНЫЕ ТРАНСФОРМАТОРНЫЕ ПРЕОБРАЗОВАТЕЛИ
НЕРАЗРУШАЮЩИХ ИСПЫТАНИЙ ДЛЯ НЕФЕРРОМАГНИТНЫХ СВАРЕННЫХ ТРУБ****Р е з ю м е**

Описаны вихрековые переходные трансформаторные преобразователи с квазиравномерным распределением магнитного поля и способ подбора условий работы этих преобразователей предназначенных для обнаруживания повреждений шва труб. Приведены результаты испытаний.

Propagacja fal elektromagnetycznych wzdłuż niejednorodnej kolumny plazmy gazowej

RYSZARD PAROSA, EDWARD RESZKE

Instytut Telekomunikacji i Akustyki, Politechnika Wrocławska

Otrzymano 1986.09.16

Autoryzowano do druku 1987.08.05

W pracy zaproponowano nową procedurę numeryczną pozwalającą wyliczać stałą propagacji (stałą falową) dla fali elektromagnetycznej quasi TEM propagowanej wzdłuż linii współosiowej z kolumną niejednorodnej plazmy stanowiącą przewód środkowy linii. Przedstawiono wyniki obliczeń dla plazmy azotowej oraz wyniki badań eksperymentalnych wykonanych dla zweryfikowania przydatności pokazanego modelu teoretycznego.

1. WPROWADZENIE

Plazma gazowa jest od wielu już lat przedmiotem stałego zainteresowania nauki i techniki. Wynika to zarówno z jej roli w przyrodzie jak i możliwości wykorzystania w praktyce jej unikalnych właściwości. Spośród różnych metod generacji plazmy gazowej coraz ważniejszą rolę odgrywają plazmatrony mikrofalowe, w których plazma gazowa wytwarzana jest wewnątrz różnorodnych mikrofalowych struktur wyładowczych: w linii współosiowej, wewnątrz falowodów i rezonatorów mikrofalowych oraz w specjalnej konstrukcji liniach opóźniających.

W badaniach plazmotronów mikrofalowych ważną rolę odgrywają prace z zakresu modelowania zjawisk charakterystycznych dla takich wyładowań, między innymi modele procesów wymiany energii pomiędzy polem elektromagnetycznym, plazmą i ściankami rezonatora. Stąd też wynika potrzeba opracowania procedur numerycznych umożliwiających wyliczenie parametrów elektrycznych mikrofalowych struktur częściowo wypełnionych plazmą. Wobec niejednorodnych przestrzennych rozkładów parametrów elektrycznych plazmy w rzeczywistych wyładowaniach, analiza taka jest zwykle bardzo złożona.

W niniejszej pracy zaproponowano oryginalną procedurę numeryczną pozwalającą wyliczać stałą propagacji dla fali elektromagnetycznej propagowanej w linii współosiowej z kolumną niejednorodnej plazmy stanowiącą przewód środkowy linii. Przedstawiono także wyniki wykonanych obliczeń dla plazmy azotowej oraz wyniki badań eksperymentalnych wykonanych dla zweryfikowania przydatności pokazanego modelu teoretycznego.

Warto tu podkreślić, że w oparciu o modele znane z literatury wyliczyć można współczynnik propagacji i rozkładu pola elektromagnetycznego jedynie dla przypadku wybranych radialnych rozkładów przenikalności elektrycznych plazmy [1] oraz dla przypadku wielowarstwowych skokowych zmian przenikalności elektrycznej plazmy [2, 3].

Zaproponowany model przydatny jest do opisu procesu absorpcji energii elektromagnetycznej w plazmotronach współosiowych, w których plazma generowana jest na przedłużeniu przewodu środkowego otwartej linii współosiowej (por. rys. 1). Wewnątrz współosiowej struktury wyładowczej pokazanej w uproszczeniu na rys. 1 wyróżnić można odrębne obszary:

- I — obszar w pobliżu końcówki przewodu środkowego, gdzie można przewidywać występowanie znacznego zakłócenia przestrzennego rozkładu pola elektromagnetycznego w linii,
- II — obszar, w którym znajduje się praktycznie cała użyteczna część kolumny plazmy i który w rzeczywistości stanowi linię współosiową z plazmą stanowiącą stratny niejednorodny przewód środkowy.

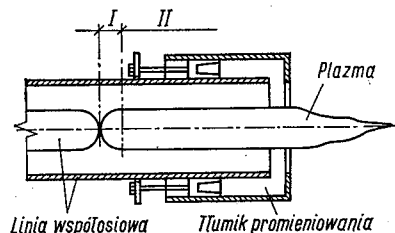
Z punktu widzenia bilansu energii w plazmie, obszar II odgrywa najważniejszą rolę bowiem w tym obszarze plazma pochłania największą część energii elektromagnetycznej doprowadzonej do głowicy wyładowczej.

2. OPIS MODELU TEORETYCZNEGO

Ocena gęstości mocy mikrofal pochłanianych przez plazmę wymaga wyliczenia współczynnika tłumienia dla fali quasi TEM propagowanej wzdłuż kolumny plazmy.

Rozważając proces propagacji monochromatycznej fali elektromagnetycznej wzdłuż linii współosiowej z plazmą przyjęto układ współrzędnych jak na rys. 1 oraz następujące założenia:

- 1 — Zespolona przenikalność elektryczna plazmy ϵ_{pl} jest symetryczną funkcją promienia „ r ”, nie zależy natomiast od współrzędnej „ z ” (wzdłuż kolumny plazmy).



Rys. 1 Konstrukcja plazmotronu współosiowego

- 2 — Przewód zewnętrzny toru współosiowego (ekran) jest idealnym przewodnikiem, $\sigma_m = \infty$.
- 3 — Wymiary toru (średnicę ekranu) dobrano tak, aby w warunkach, gdy przewodność plazmy jest bardzo duża, w torze mógł być propagowany jedynie rodzaj fali TEM.

Przy tych założeniach składowa pola elektrycznego E_z opisana jest równaniem falowym:

$$\frac{\partial^2 E_z}{\partial r^2} + \frac{1}{r} \frac{\partial E_z}{\partial r} + (\omega^2 \mu_0 \varepsilon_{pl}(r) - h^2) E_z = 0 \quad (1)$$

gdzie: $h = \beta - j\alpha$, β — współczynnik przesunięcia fazy, α — współczynnik tłumienia, ε_{pl} — zespolona przenikalność elektr. plazmy,

$$\varepsilon_{pl} = 1 - \frac{\omega_p^2}{\omega^2 + \nu^2} - j \frac{\nu}{\omega} \frac{\omega_p^2}{\omega^2 + \nu^2}; \quad \omega_p = \left(\frac{N_e e^2}{\varepsilon_0 m_e} \right)^{1/2}$$

N_e — koncentracja elektronów w plazmie, ν — częstość zderzeń w plazmie.

Warunki brzegowe mają następującą postać:

$$E_z(R) = 0; \partial E_z(r=0)/\partial r = 0. \quad (2)$$

Wygodnie jest tu wprowadzić znormalizowaną wartość natężenia pola elektrycznego w osi linii $E(r=0) = E_0$

Proponowana metoda numeryczna może opierać się na prostym schemacie różnicowym:

$$\frac{\partial E_z}{\partial r} \simeq \frac{E_{i+1} - E_{i-1}}{2\Delta R}; \quad \frac{\partial^2 E_z}{\partial r^2} \simeq \frac{E_{i+1} - 2E_i + E_{i-1}}{\Delta R^2} \quad (3)$$

gdzie $\Delta R = \frac{R}{n}$, i n — jest ilością warstw, na jakie podzielono promień ekranu R .

Wstawiając (3) do równania falowego (1) otrzymuje się równanie różnicowe:

$$\frac{E_{i+1} - 2E_i + E_{i-1}}{\Delta R^2} - \frac{1}{r_i} \frac{E_{i+1} - E_{i-1}}{2\Delta R} + (\omega^2 \mu_0 \varepsilon_{pl}(r_i) - h^2) \cdot E_i = 0 \quad (4)$$

$i = 1, 2, 3, \dots, n-1$.

W przyjętym schemacie różnicowym warunki brzegowe (2) przyjmują postać:

$$E_1 = \frac{1}{4} E_0 (4 - \gamma_0^2 \Delta R^2); \quad E_z(R) = E_n = 0 \quad (5)$$

gdzie:

$$\gamma_i^2 = \omega^2 \mu_0 \varepsilon_{pl}(r_i) - h^2. \quad (6)$$

Występująca w równaniu (6) stała propagacji h stanowi początkowe przybliżenie poszukiwanej dokładnej wartości stałej propagacji h^* .

Wstawiając do równania różnicowego (4) $\gamma_i = \gamma_i^{(0)}$ (gdzie $\gamma_i^{(0)}$ — przybliżenie początkowe γ_i^*) i po niewielkich przekształceniach otrzymuje się wzór rekurencyjny pozwalający wyliczyć kolejne wartości E_{i+1} :

$$E_{i+1} = \frac{(2i)^{-1} - 1}{(2i)^{-1} + 1} E_{i+1} + \frac{2 - \gamma_i^{(0)^2} \cdot \Delta R^2}{(2i)^{-1} + 1} E_i. \quad (7)$$

Ponieważ wstawiona do równania (7) wartość $\gamma^{(0)}$ różni się od dokładnej wartości γ_i^* , to obliczona ze wzoru rekurencyjnego wartość natężenia pola elektrycznego przy ekranie E_n nie będzie równa zeru — a zatem nie będzie spełniony warunek brzegowy.

Dokładne wartości natężenia pola elektrycznego określone są przez analogiczne równanie rekurencyjne:

$$E_{i+1}^* = \frac{(2i)^{-1}-1}{(2i)^{-1}+1} E_{i+1}^* + \frac{2-\gamma_i^{*2} \cdot \Delta R^2}{(2i)^{-1}+1} E_i^* \quad (8)$$

gdzie:

$$\gamma_i^{*2} = \omega^2 \mu_0 \varepsilon_{pl}(r_i) - h^{*2}. \quad (9)$$

Odejmując stronami równania (7) i (8) oraz dzieląc wynik przez $\Delta h^{(0)2} = h^{*2} - h^{(0)2}$ otrzymuje się:

$$D_{i+1} = \frac{(2i)^{-1}-1}{(2i)^{-1}+1} \cdot D_{i-1} + \frac{2-\gamma_i^{(2)2} \Delta R^2}{(2i)^{-1}+1} \cdot D_i - \frac{\Delta R^2}{(2i)^{-1}+1} E_i^* \quad (10)$$

gdzie:

$$D_i = \frac{E_i^* - E_i}{\Delta h^{(0)2}} \quad \text{oraz} \quad \Delta h^{(0)2} = h^{*2} - h^{(0)2}.$$

Warunki brzegowe dla D_i mają postać:

$$D_0 = \frac{E_0^* - E_0}{\Delta h^{(0)2}} = 0, \quad \text{bowiem} \quad E_0^* = E_0, \quad (11a)$$

$$D_1 = -\frac{\Delta R^2}{2} \cdot E_0, \quad (11b)$$

$$D_n = \frac{E_n^* - E_n}{\Delta h^{(0)2}} = -\frac{E_n}{\Delta h^{(0)2}}, \quad \text{bowiem} \quad E_n^* = 0. \quad (11c)$$

Korzystając z wzorów (11a), (11b) oraz wzoru rekurencyjnego (10) oraz (7) można wyliczyć kolejne wartości E_i oraz D_i dla $i = 1, 2, 3, \dots, n$. Następnie na podstawie warunku brzegowego (11 c) wylicza się poprawkę $\Delta h^{(0)2}$:

$$\Delta h^{(0)2} = -\frac{E_n}{D_n}. \quad (12)$$

Po obliczeniu poprawki $\Delta h^{(0)2}$ każdorazowo obliczamy nową, poprawioną wartość stałej propagacji:

$$(h^{m+1})^2 = (h^{(m)})^2 + (\Delta h^{(m)})^2, \quad m = 0, 1, 2, \dots \quad (13)$$

Procedurę wyliczania poprawionej wartości współczynnika propagacji powtarza się wielokrotnie, aż do uzyskania wystarczająco małych wartości poprawek $\Delta h^{(m)}$.

Obliczenia wskazały, że opisana procedura jest monotonicznie zbieżna, tzn:

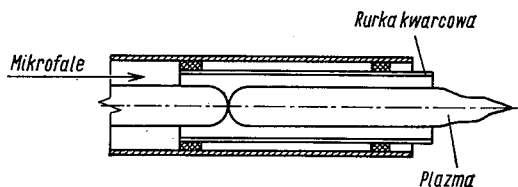
$$\lim_{k \rightarrow \infty} h^{(k)} \rightarrow h^* \quad \text{i} \quad |\Delta h^{(m)}| \gtrsim |\Delta h^{(m+1)}|$$

jeśli tylko spełniony jest warunek:

$$|\Delta h^{(0)}| \lesssim |\gamma_i^2|_{\max}$$

3. PROCEDURA OBLICZEŃ DLA PRZYPADKU LINII WSPÓŁOSIOWEJ WYPEŁNIONEJ DIELEKTRYKIEM ZE SKOKOWO ZMIENIAJĄCYMI SIĘ PARAMETRAMI ELEKTRYCZNYMI

Rozważany dalej będzie przypadek, gdy radialny rozkład przenikalności elektrycznej jest funkcją nieciągłą. Rozkład taki może modelować np. linię współosiową z plazmą generowaną wewnątrz rurki kwarcowej (por. rys. 2).



Rys. 2 Linia współosiowa z plazmą generowaną wewnątrz rurki kwarcowej

Warunek brzegowy w miejscu nieciągłości funkcji $\varepsilon = \varepsilon(r)$ może być wyprowadzony w oparciu o równania Maxwella i warunek ciągłości impedancji. Jeśli przyjąć oznaczenia:

$$\varepsilon(r = r_{d-}) = \varepsilon_{pl}(r_d); \varepsilon(r = r_{d+}) = \varepsilon_d,$$

to warunki brzegowe dla $r = r_d$ mogą być zapisane w następującej postaci:

$$\frac{\partial E_z(r = r_{d-})}{\partial r} \frac{\omega^2 \mu_0 \varepsilon_{pl}(r_d)}{\omega^2 \mu_0 \varepsilon_{pl}(r_d) - h^2} = \frac{\partial E_z(r = r_{d+})}{\partial r} \frac{\omega^2 \mu_0 \varepsilon_d}{\omega^2 \mu_0 \varepsilon_d - h^2} \quad (14)$$

$$E_z(r = r_{d-}) = E_z(r = r_{d+}). \quad (15)$$

Stąd łatwo wyprowadzić zależności rekurencyjne w postaci:

$$E_{k+1} = \frac{\varepsilon_{pl}}{\varepsilon_d} \frac{\omega^2 \mu_0 \varepsilon_d - h^2}{\omega^2 \mu_0 \varepsilon_{pl} - h^2} (E_k - E_{k-1}) + E_k, \quad (16)$$

$$E_{k+1}^* = \frac{\varepsilon_{pl}}{\varepsilon_d} \frac{\omega^2 \mu_0 \varepsilon_d - h^{*2}}{\omega^2 \mu_0 \varepsilon_{pl} - h^{*2}} (E_k^* - E_{k-1}^*) + E_k^*. \quad (17)$$

Po odjęciu stronami równań (16) oraz (17) i po dalszych przekształceniach wyprowadzić można zależność pozwalającą wyliczyć kolejną wartość współczynnika D_{k+1} w obszarze dielektryka, tzn. dla $r = r_d + \Delta r$:

$$D_{k+1} = (D_k - D_{k-1}) \frac{\gamma_{pl}^2 \varepsilon_{pl}}{\gamma_{pl}^2 \varepsilon_d} + (E_k - E_{k-1}) \left(\frac{\gamma_d^2}{\gamma_{pl}^4} - \frac{1}{\gamma_{pl}^2} \right) \frac{\varepsilon_{pl}}{\varepsilon_d} + D_k \quad (18)$$

gdzie: $\gamma_{pl}^2 = \omega^2 \mu_0 \varepsilon_{pl}(r_d) - h^2$; $\gamma_d^2 = \omega^2 \mu_0 \varepsilon_d - h^2$

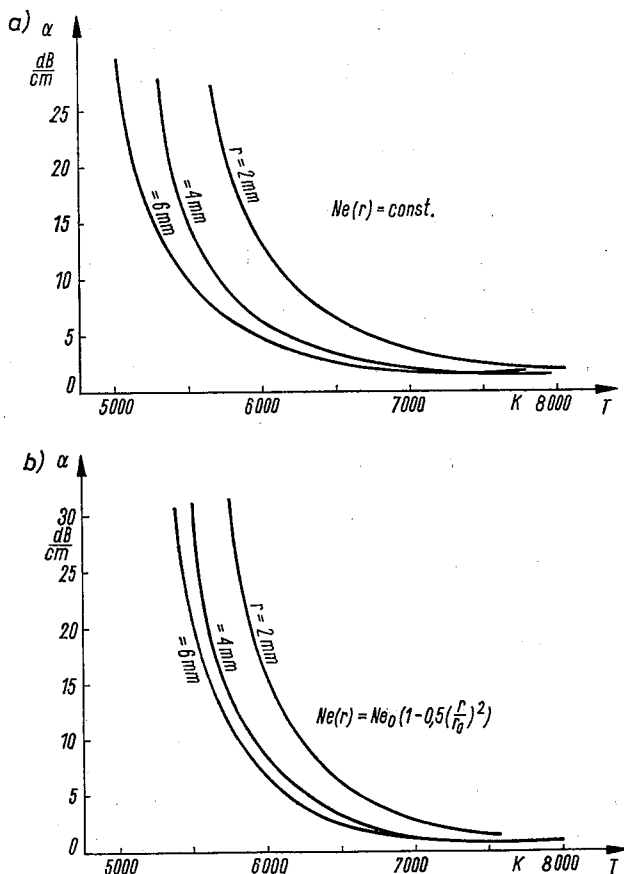
Kolejne wartości E_i oraz D_i dla $i \geq k+2$ oblicza się już według normalnej procedury, tzn. z wzorów (7) i (10). Warto tu jeszcze zaznaczyć, że ponownie należy stosować podane tu wzory (16) oraz (18) dla $r = r_p$, t.j. na granicy ośrodka kwarcowego i powietrza.

4. WYNIKI OBLICZEŃ DLA NIEJEDNORODNEJ PLAZMY AZOTOWEJ

Wykorzystując model przedstawiony w rozdz. 2 opracowano program komputerowy i wyliczono zależność zespolonej stałej propagacji od temperatury plazmy azotowej. Obliczenia wykonano dla jednorodnej kolumny plazmy i dla plazmy o radialnym rozkładzie koncentracji w postaci parabolicznej funkcji promienia „ r ”:

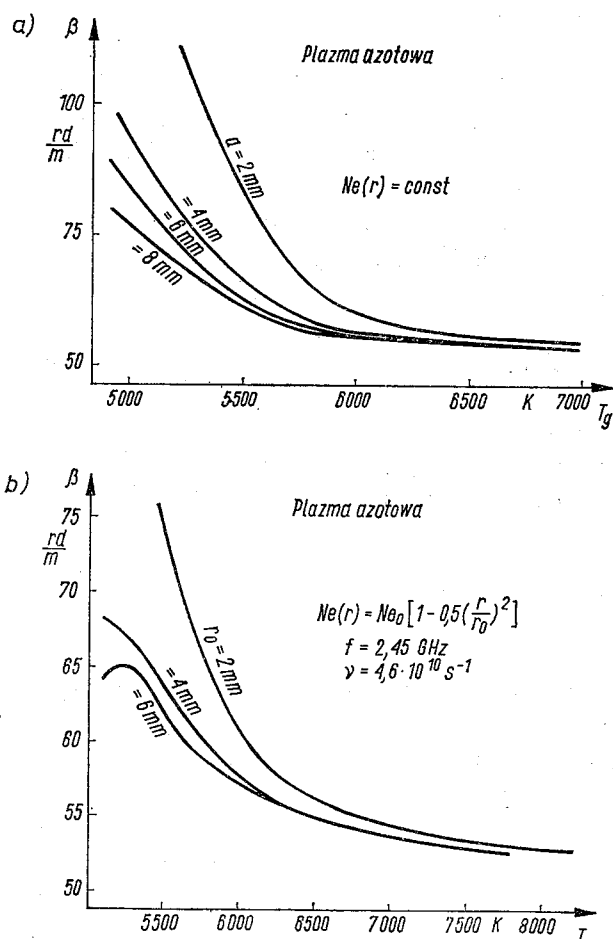
$$N_e(r) = N_{e0} \cdot \left[1 - \left(\frac{r}{r_0} \right)^2 \right]$$

Wyniki tych obliczeń zestawiono na rys. 3 i 4.



Rys. 3 Zależność tłumienności jednostkowej α dla linii współosiowej z plazmą od temperatury plazmy; a) dla jednorodnej kolumny plazmy, b) dla niejednorodnej plazmy

Pokazane na rys. 3 i 4 zależności stałej propagacji od temperatury plazmy pozwalają na dokonanie oceny wpływu niejednorodnego rozkładu parametrów plazmy na parametry falowe współosiowej struktury z plazmą. Zgodnie z oczekiwaniami tłumienność α i przesuwność β w torze współosiowym z plazmą są znacznie większe, gdy plazma jest niejed-



Rys. 4 Zależność przesuwności jednostkowej β dla linii współosiowej z plazmą od temperatury plazmy;
 a) dla jednorodnej kolumny plazmy, b) dla niejednorodnej plazmy

norodna. Warto tu podkreślić, że podobnie różnice uzyskano zakładając w obliczeniach inne radialne rozkłady koncentracji elektronów w plazmie.

Porównanie rezultatów obliczeń wykonanych przy założeniu jednorodności ośrodka plazmowego oraz dla plazmy niejednorodnej wskazuje, że jeśli promień kolumny plazmy jest niewielki (dla plazmy azotowej $r_0 \leq 4 \text{ mm}$), to wykresy $\alpha = \alpha(N_e)$ oraz $\beta = \beta(N_e)$ są zbliżone z zadowalającą dokładnością (około 30%), jeśli w przypadku niejednorodnej plazmy przyjąć średnią koncentrację elektronów w kolumnie plazmy. Można zatem dla „smukłych” plazm przyjąć, że parametry falowe toru α oraz β są funkcją średniej koncentracji elektronów w plazmie i korzystać z prostego rozwiązania równania dyspersyjnego dla linii współosiowej z jednorodnym stratnym przewodem środkowym (por. np. [4]). Jeśli natomiast promień kolumny plazmy jest większy, niż 4 mm (dla azotu), to konieczna jest dokładniejsza analiza, bowiem uzyskane tłumienności jednostkowe α są w takim przypadku wyraźnie większe, niż tłumienności wyliczone przy założeniu średniej kon-

centracji elektronów w plazmie. Ten zaskakujący na pierwszy rzut oka rezultat można jednak stosunkowo prosto wyjaśnić, biorąc pod uwagę omówiony wcześniej efekt naskórkowy w plazmie. Wyniki obliczeń pokazane w [5] wskazują bowiem, że w plazmie azotowej efekt naskórkowy jest wyraźny, jeśli kolumna plazmy ma promień mniejszy niż $r_0 \lesssim 4\text{--}5$ mm (przy $T \lesssim 6000$ K). Można zatem sądzić, że w przypadku kolumny plazmy o promieniu $r_0 > 4$ mm pole elektromagnetyczne penetruje głównie peryferyjne obszary plazmy, natomiast w osiowej części kolumny plazmy, gdzie temperatura i koncentracja elektronów są największe, pole elektryczne jest najmniejsze. A zatem pole elektromagnetyczne jakby „odczuwało” obecność plazmy jedynie o mniejszych temperaturach i koncentracjach elektronów.

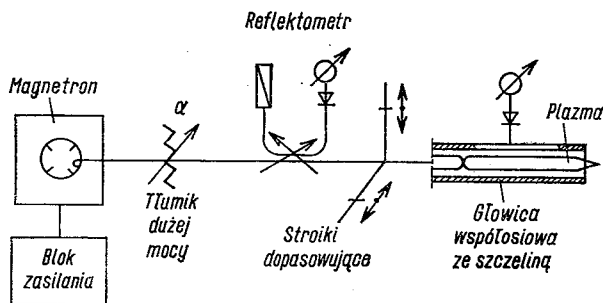
Podsumowując wyniki obliczeń, należy podkreślić, że omówiony tu wpływ efektu naskórkowego na parametry falowe linii współosiowej z plazmą występuje w podobny sposób w przypadku dowolnych gazów plazmotwórczych. Jeśli efekt naskórkowy jest duży, to dla oceny warunków propagacji fali quasi TEM konieczne jest uwzględnienie w analizie radialnego rozkładu koncentracji elektronów w plazmie.

5. WERYFIKACJA EKSPERYMENTALNA MODELU LINII WSPÓŁOSIOWEJ Z PLAZMĄ

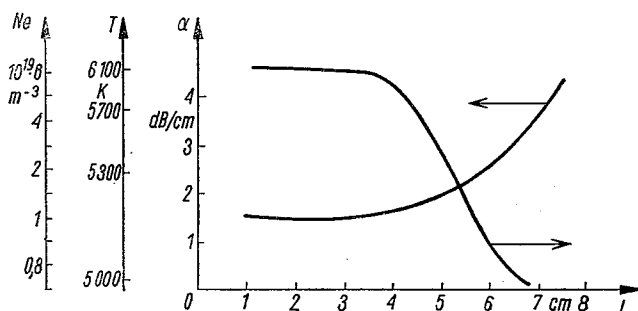
Opisany model linii współosiowej z plazmą, jakkolwiek już złożony, bo uwzględniający niejednorodność ośrodka plazmowego, to opisuje rzeczywiste zjawiska w znacznym przybliżeniu. Przykładowo, w analizie nie uwzględniono wyższych rodzajów pola elektromagnetycznego, które mogą być wzbudzone i propagowane w takiej linii. Mimo, że w mikrofalowych plazmotronach współosiowych konstrukcja głowicy wyładowczej wskazuje raczej na dominację fali TEM, to obecność innych rodzajów pola jest możliwa, szczególnie gdy temperatury plazmy są wysokie ($|\epsilon_w| > 1$) oraz gdy średnica kolumny plazmy jest duża.

Celem weryfikacji przedstawionych rozważań teoretycznych wykonano pomiary tłumienności jednostkowej α w linii współosiowej z plazmą (w plazmotronie współosiowym) oraz wykonano pomiary spektralne takiej plazmy. Celem eksperymentu było porównanie temperatury plazmy wyliczonej w oparciu o przedstawiony model (przyjmując w obliczeniach zmierzoną wartość α) z temperaturą wyznaczoną na drodze pomiarów spektroskopowych.

Schemat blokowy stanowiska eksperymentalnego pokazano na rys. 5. W skład tego stanowiska wchodzi uniwersalny zestaw do generacji plazmy mikrofalowej opisany szczegółowo w [6]. Źródłem energii mikrofalowej był magnetron fali ciągłej o mocy wyjściowej regulowanej w zakresie od 200 W do 2500 W, o częstotliwości $f = 2.45 \text{ GHz} \pm 0.12 \text{ GHz}$. Energia mikrofalowa doprowadzana była do głowicy współosiowej torem wyposażonym w układ stroików umożliwiających dopasowanie energetyczne głowicy wyładowczej z plazmą do generatora mikrofalowego. W przewodzie zewnętrznym głowicy współosiowej wycięto szczelinę i zainstalowano ruchomą karetkę z dwiema niewielkimi sondami mikrofalowymi i detektorami. Sondy te umieszczone wewnątrz szczeliny i przesuwane wzdłuż linii współosiowej z plazmą pozwoliły zmierzyć rozkład natężenia pola elektrycznego wzdłuż linii i stąd wyznaczyć rozkład tłumienności $\alpha(z)$. Z porównania zmierzonych wartości α z wy-



Rys. 5 Schemat blokowy stanowiska eksperymentalnego



Rys. 6 Wyznaczone średnie koncentracje elektronów w plazmie wzdłuż kolumny plazmy

kresami na rys. 3 wyznaczono średnie koncentracje elektronów w plazmie wzdłuż kolumny plazmy — por. rys. 6.

Następnie zakładając istnienie w plazmie stanu lokalnej równowagi termicznej (LRT) i w oparciu o wzór Saha, wyliczono rozkład temperatury plazmy wzdłuż kolumny plazmy:

$$\frac{N_i N_e}{N_0} = \frac{g_i g_e}{g_0} \cdot \left(\frac{2\pi m_e k T}{h^2} \right)^{3/2} \cdot \exp \left(-\frac{w_j}{k T} \right)$$

gdzie:

- N_e, N_i, N_0 — odpowiednio: koncentracja elektronów, jonów i cząstek obojętnych,
- g_e, g_i, g_0 — odpowiednio: wagi statystyczne elektronu, jonu i cząstki obojętnej,
- m_e — masa elektronu,
- h — stała Plancka,
- W_j — energia jonizacji.

Średnicę kolumny plazmy oceniano metodą fotograficzną przyjmując granicę plazmy w miejscu, w którym intensywność świecenia plazmy malała około 10-krotnie w porównaniu z intensywnością świecenia centralnej części plazmy. Pomiary wykonano wielokrotnie uzyskując powtarzalność wyników (przy ustalonych warunkach zasilania plazmotronu) w granicach $\pm 7\%$.

W pomiarach spektroskopowych wykorzystano znaną metodę spektroskopii cząsteczkowej polegającą na rejestracji względnych intensywności linii pasm rotacyjnych cząsteczki azotu. Metoda ta pozwala wyznaczyć tzw. temperaturę rotacyjną plazmy, którą w warunkach

kach gdy rozkład prędkości molekuł jest makswelewski oraz rozkład obsadzeń poziomów rotacyjnych i oscylacyjnych zgodny jest z prawem Boltzmana, równa jest temperaturze gazu. W przeprowadzonych eksperymentach zmierzono intensywności względne linii rotacyjnych cząsteczki N_2^+ ($\lambda = 3914.4 \text{ \AA}$) systemu $B^2 \sum_u^+ - x^2 \sum_g^+$ (przejście 0—0). Mierzono intensywności linii rotacyjnych dla $40 \leq J \leq 94$ gałęzi R wybierając z tego zbioru jedynie linie dobrze separowane. Pomiary wykonano za pomocą spektrometru PGS-2 w III rzędzie z filtrami wybierającymi rejestrując badane widmo rotacyjne na kliszach Agfa-Gevaert 34B50 i 26D50 oraz rejestratorem sprzężonym mechanicznie z siatką dyfrakcyjną spektrometru (pomiaru fotopowielaczem).

Przy pomiarach głowicę współosiową zasilano mocą mikrofalową $P_u = 1.2 \text{ kW}$ i przy wymuszonym przepływie azotu $v = 1850 \text{ 1/h}$. Zmierzono średnią temperaturę rotacyjną plazmy w odległości około 1 cm od końcówki przewodu środkowego oraz w odległości około 3 cm od końcówki tego przewodu. Przy pomiarze temperatury rotacyjnej bliżej końcówki przewodu środkowego, uzyskano $T_{\text{rot}} = 5600 + 500 \text{ K}$, natomiast w drugim punkcie pomiarowym zmierzono $T_{\text{rot}} = 5450 \pm 500 \text{ K}$.

Uzyskane na drodze pomiarów spektroskopowych temperatury plazmy azotowej są z zadowalającą dokładnością zbieżne z wynikami pokazanymi na wykresie na rys. 6. Warto tu jeszcze podkreślić, że podobną zgodność stwierdzono porównując wyniki pomiarów spektroskopowych znane z pracy [7] z wynikami obliczeń opartych na zaproponowanym modelu linii współosiowej z plazmą.

PODSUMOWANIE

Opisana w pracy stosunkowo prosta procedura numeryczna znalazła już swoje zastosowania zarówno jako element rozwiniętego modelowania zjawisk zachodzących w skali makroskopowej w plazmotronach mikrofalowych (por. praca [8]) jak również stanowi podstawę nowej metody diagnostyki plazmy. Z przytoczonych rozważań teoretycznych wynika bowiem, że dla dowolnej plazmy wyznaczyć można jednoznaczną zależność stałej propagacji w torze współosiowym z plazmą od makroskopowych parametrów elektrycznych plazmy. Wystarczy zatem wykonać pomiary stałej propagacji w takim torze, aby wykorzystując zaproponowaną procedurę obliczeniową wyznaczyć makroskopowe parametry badanej plazmy.

Warto tu podkreślić, że analogiczną procedurę numeryczną opracowano dla rezonatorów mikrofalowych częściowo wypełnionych niejednorodną plazmą. Procedura ta umożliwia wyznaczenie częstotliwości rezonansowej rezonatora, dobroci i rozkładu pola elektromagnetycznego wewnątrz rezonatora [9].

BIBLIOGRAFIA

1. D. N. P i n d e r: *Propagation of fast electromagnetic waves along a non-uniform annular plasma column*. Int. Journal of Electronics, 1974, vol. 36, nr 2, s. 223
2. G. M u r: *Field analysis and complex resonance frequency of the quasi- TE_{011} -mode in an inhomogeneously filled resonator with losses.*, Appl. Sci. Res., April 1974, vol. 29, s. 137

3. J. Rogers, J. Asmussen.: *Standing waves along a microwave surface wave plasma*. IEEE Trans. on Plasma Science, March 1982, vol. PS-10, nr 1, s. 11
4. R. Parosa, E. Reszke: *Propagacja fal elektromagnetycznych wzdłuż kolumny wysokociśnieniowej plazmy*. Proc. of II Symp. on Plasma Physics, Gdańsk 1978, vol. 1, s. 45
5. R. Parosa: *Analiza stanu stacjonarnego w mikrofalowym wyładowaniu plazmowym*. Archiwum Elektrotechniki, 1980, vol. 29, nr 3, s. 725
6. R. Parosa, E. Reszke: *Problemy mikrofalowej generacji plazmy gazowej przy ciśnieniu atmosferycznym*. Zespołowa praca doktorska. Inst. Telekomunikacji i Akustyki Politechniki Wrocławskiej, Wrocław 1976
7. V. Baltin i inni.: *Stacionarnyj SVC-razriad v azocie pri atmosfernom davlenii*. Teplofiz. Vysokich Temp., vol. 9, 1971, s. 1105
8. R. Parosa, E. Reszke: *Simplified Modelling of Atmospheric Pressure Microwave Discharges*. XVII Int. Conf. on Phenomena in Ionized Gases. Budapest 1985, s. 846
9. R. Parosa: *Analiza rezonatorów cylindrycznych z niejednorodnym wypełnieniem*. VI Krajowa Konf. Mikrofalowa MIKON-83 Zakopane 1983, tom 1, s. 288

R. PAROSA, E. RESZKE

PROPAGATION OF ELECTROMAGNETIC WAVES ALONG A NONHOMOGENEOUS PLASMA COLUMN

Summary

A new numerical procedure has been proposed that enables calculations of the propagation constant for quasi TEM wave propagated along coaxial line with a plasma — formed inner conductor. Results are presented for nonhomogeneous nitrogen plasma and a good measure of agreement between the calculated and experimental data is pointed out in conclusions.

R. PAROSA, E. RESZKE

PROPAGATION DES ONDES ÉLECTROMAGNÉTIQUES LE LONG DE LA COLONNE DU PLASMA HÉTÉROGÈNE A GAZ

Résumé

Dans l'étude on propose une nouvelle procédure numérique, permettant de calculer la constante de la propagation (constante de l'onde) pour l'onde électromagnétique quasi TEM propagée le long de la ligne coaxiale à la colonne du plasma hétérogène, constituant le conduit central de la ligne. On présente les résultats des calculs pour le plasma à azote, ainsi que les résultats des essais expérimentaux, faits en vue de la vérification de l'utilité pratique du modèle théorique présenté.

R. PAROSA, E. RESZKE

AUSBREITUNG DER ELEKTROMAGNETISCHEN WELLE LÄNGS EINER INHOMOGENEN GASPLASMASÄULE

Zusammenfassung

Ein neues Verfahren für die numerische Berechnung des Ausbreitungskoeffizienten einer elektromagnetischen quasi-TELL-Welle längs einer inhomogenen Gasplasmasäule wurde vorgeschlagen. Auch wurden Berechnungsergebnisse sowie Meßresultate für die Stickstoffplasma angeführt.

Р. ПАРОСА, Э. РЕШКЕ

РАСПРОСТРАНЕНИЕ ЭЛЕКТРОМАГНИТНОЙ ВОЛНЫ ВДОЛЬ ПЛАЗМЕННОГО
СТЕРЖНЯ

Р е з ю м е

Предложены машинные методы расчета постоянной распространения электромагнитной волны типа квази ТЕМ, распространяющейся вдоль коаксиальной линии, в которой в качестве центрального проводника используется неоднородная плазма. Представлены расчеты для азотной плазмы и результаты экспериментальных исследований, подтверждающие правильность выбранной теоретической модели.

Właściwości elektryczne wielostykowych modeli złącza metal - metal

HENRYK GIERASIMOWICZ

Instytut Podstaw Elektroniki, Politechnika Warszawska

Otrzymano 1987.07.24

Autoryzowano do druku 1987.11.16

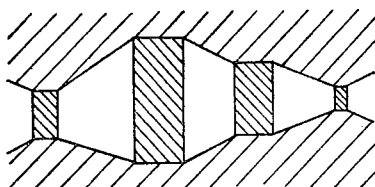
W pracy porównano właściwości elektryczne wielostykowych modeli złącza metal-metal złożonych z mikrostryków o różnych wymiarach. Analizując czynniki mogące prowadzić do degradacji złącza stwierdzono, że dominującą rolę odgrywają czynniki fizyko-chemiczne, natomiast jest mało prawdopodobne aby nastąpiło zniszczenie złącza w następstwie przepływającego przez niego prądu. Procesowi degradacji złącza towarzyszy zmiana względnej intensywności szumów typu $1/f$, co może być wykorzystane do eliminacji złączy potencjalnie zawodnych.

Uszkodzenia elementów i podzespołów elektronicznych są w znacznej mierze spowodowane wadliwie wykonanymi złączami metal-metal i procesami prowadzącymi do ich degradacji [1, 2, 3]. U podstaw degradacji złączy leżą takie procesy fizyko-chemiczne jak korozja, dyfuzja, tworzenie się związków międzymetalicznych, itp. Procesy te oraz chropowatość łączonych powierzchni [4] mogą być przyczyną niejednorodności złącza, przejawiającej się między innymi nieciągłością powierzchni jego poprzecznego przekroju. Jeżeli złącze jest niejednorodne, to jego modele mogą mieć postać modeli wielostykowych, będących równoległym połączeniem wielu elementarnych mikrostryków o różnych polach przekroju poprzecznego. Wielostykowe modele złącza metal-półprzewodnik badał między innymi Vandamme i jego współpracownicy [5, 6].

W niniejszej pracy, realizowanej w ramach CPBP 02.14, porównano właściwości elektryczne modeli wielostykowych złożonych z mikrostryków o różnych wymiarach oraz oceniono wpływ natężenia prądu przepływającego przez mikrostryki na szybkość procesów destrukcyjnych prowadzących do zniszczenia złącza.

1. WPŁYW WYMIARÓW MIKROSTYKÓW NA WŁAŚCIWOŚCI ELEKTRYCZNE WIELOSTYKOWEGO MODELU ZŁĄCZA

Niech dany będzie model złącza wielostykowego (rys. 1) zasilanego prądowo prądem o natężeniu I , składający się z mikrostryków mających postać walców krzywoliniowych



Rys. 1. Wielostykowy model złącza metal-metal

o długości d_i i przekroju poprzecznym S_i ($i = 1, 2, \dots, n$), przy czym liczba mikrostryków o przekroju poprzecznym S_i wynosi l_i .

Jeżeli przewodność właściwa materiału mikrostryku jest równa σ , to przewodność mikrostryku o długości d_i i przekroju poprzecznym S_i wynosi

$$G_i = \frac{\sigma S_i}{d_i}. \quad (1)$$

Względną intensywność C_i szumów typu $1/f$ w takim mikrostryku, zdefiniowaną jako

$$C_i = \frac{(\overline{\Delta U_i})^2}{U_i^2} \cdot \frac{f}{\Delta f} = \frac{(\overline{\Delta G_i})^2}{G_i^2} \cdot \frac{f}{\Delta f}, \quad (2)$$

gdzie:

$(\overline{\Delta U_i})^2$ — wartość średniokwadratowa napięcia szumów w i -tym mikrostryku,

$(\overline{\Delta G_i})^2$ — wartość średniokwadratowa fluktuacji przewodności i -tego mikrostryku,

Δf — rozpatrywane pasmo częstotliwości,

można je wyznaczyć stosując procedurę podaną w [7]. Traktując szумы typu $1/f$ jako zjawisko objętościowe i zakładając brak korelacji między fluktuacjami warstw utworzonych przez sąsiednie powierzchnie ekwipotencjalne otrzymuje się

$$C_i = \frac{\beta}{n_c} \cdot \frac{1}{d_i S_i}, \quad (3)$$

gdzie:

β — stała wynosząca dla metali $2 \cdot 10^{-3}$,

n_c — koncentracja swobodnych nośników w mikrostryku (dla aluminium $n_c = 8,4 \cdot 10^{28} \text{ m}^{-3}$).

Wobec powyższego, przewodność G złącza wielostykowego będzie określona wyrażeniem

$$G = \sum_{i=1}^n l_i G_i = \sigma \sum_{i=1}^n \frac{S_i}{d_i} l_i \quad (4)$$

a początkowe wartości natężenia prądu I_k i gęstości prądu j_k w mikrostryku o przekroju poprzecznym S_k i przewodności G_k wyniosą

$$I_k = \frac{IG_k}{G}, \quad (5)$$

$$j_k = \frac{I_k}{S_k} = \frac{IG_k}{S_k G} \quad (6)$$

Względna intensywność C szumów typu $1/f$ dla złącza wielostykowego będzie opisana zależnością

$$C = \frac{(\overline{\Delta G})^2}{G^2} \cdot \frac{f}{\Delta f}, \quad (7)$$

gdzie:

$(\overline{\Delta G})^2$ — wartość średniokwadratowa fluktuacji przewodności złącza wielostykowego,
 G — przewodność złącza wielostykowego.

Wykorzystując zależności podane w [8], dotyczące wyznaczania fluktuacji dla układu rezystorów, można wartość średniokwadratową fluktuacji przewodności złącza związaną z szumami typu $1/f$ wyrazić następująco:

$$(\overline{\Delta G})^2 = \sum_{i=1}^n l_i (\overline{\Delta G_i})^2 \quad (8)$$

co po wykonaniu niezbędnych przekształceń prowadzi do zależności

$$C = \frac{\sum_{i=1}^n l_i G_i^2 C_i}{G^2} \quad (9)$$

Długość d_i mikrostryków tworzących złącze może być stała, niezależna od przekroju poprzecznego mikrostryków lub może być związana z przekrojem poprzecznym S_i określoną zależnością funkcyjną. Aby przeanalizować wpływ wymiarów mikrostryków na parametry elektryczne modelu, rozpatrzono modele wielostykowe, dla których zachodzą podane poniżej zależności między długością d_i a przekrojem poprzecznym S_i mikrostryku.

model 1 — $d_i \sim S_i$

model 2 — $d_i \sim \sqrt{S_i}$

model 3 — $d_i = d_0 = \text{const}$

model 4 — $d_i \sim 1/\sqrt{S_i}$

model 5 — $d_i \sim 1/S_i$

(10)

Zależności uzyskane dla poszczególnych modeli złączy wielostykowych zestawiono w tabeli 1.

Aby móc porównać ze sobą właściwości elektryczne modeli złączy wielostykowych złożonych z mikrostryków o różnych wymiarach, wykonano obliczenia dla złączy złożonych z $n = 100$ mikrostryków aluminiowych (przewodność właściwa $\sigma = 3,53 \cdot 10^7$ 1/ Ωm). Założono, że

— prąd przepływający przez złącze $I = 1 \mu\text{A}$,

— suma powierzchni przekrojów poprzecznych wszystkich mikrostryków tworzących złącze wynosi $S = 10^{-12}$ m²,

— liczba l_i mikrostryków o danym przekroju poprzecznym S_i wynosi $l_0 = 1$,

— przekroje poprzeczne kolejnych mikrostryków są określone zależnością $S_i = S_1 + (i-1)\Delta S$, gdzie $\Delta S > 0$,

— parametr $D = (S_n - S_1)/(S_n + S_1) = 0,5$, co oznacza że $\Delta S = 2S_1/(n-1)$,

— współczynniki proporcjonalności w zależnościach (10) są tak dobrane, że długości mikrostryków o średnim przekroju poprzecznym równym 10^{-14} m² wynoszą $d_0 = 10^{-6}$ m.

Tabela 1

Zestawienie zależności określających właściwości elektryczne wielostykowych modeli złącza

model 1	model 2	model 3	model 4	model 5
$d_i \sim S_i$	$d_i \sim \sqrt{S_i}$	$d_i = d_0 = const.$	$d_i \sim l/\sqrt{S_i}$	$d_i \sim l/S_i$
$G \sim \sum_{i=1}^n l_i$	$G \sim \sum_{i=1}^n l_i \sqrt{S_i}$	$G \sim \sum_{i=1}^n l_i S_i = S$	$G \sim \sum_{i=1}^n l_i S_i^{3/2}$	$G \sim \sum_{i=1}^n l_i S_i^2$
$j_k \sim \frac{1}{S_k G}$	$j_k \sim \frac{1}{\sqrt{S_k} G}$	$j_k \sim \frac{1}{G}$	$j_k \sim \frac{\sqrt{S_k}}{G}$	$j_k \sim \frac{S_k}{G}$
$C \sim \frac{1}{G^2} \sum_{i=1}^n \frac{l_i}{S_i^2}$	$C \sim \frac{1}{G^2} \sum_{i=1}^n \frac{l_i}{\sqrt{S_i}}$	$C \sim \frac{S}{G^2}$	$C \sim \frac{1}{G^2} \sum_{i=1}^n l_i S_i^{5/2}$	$C \sim \frac{1}{G^2} \sum_{i=1}^n l_i S_i^4$

Tabela 2

Porównanie własności elektrycznych wielostykowych modeli złącza

model	1	2	3	4	5
$G [S]$	35,3	26,2	35,3	14,4	10,7
j_{max} (A/mm ²)	2,00	1,91	1,00	3,00	4,95
°C	$7,94 \cdot 10^{-14}$	$5,84 \cdot 10^{-14}$	$2,38 \cdot 10^{-14}$	$3,24 \cdot 10^{-14}$	$2,41 \cdot 10^{-14}$

Wyniki obliczeń zestawiono w tabeli 2, gdzie przez j_{max} oznaczono maksymalną gęstość prądu w danym złączu (w jednym z mikrostków złącza).

Z danych zawartych w tablicach 1 i 2 wynikają następujące wnioski:

- gęstość prądu jest taka sama we wszystkich mikrostkach tworzących złącze tylko wtedy, gdy wszystkie mikrostryki mają tę samą długość,
- najlepszymi właściwościami elektrycznymi charakteryzuje się złącze złożone z mikrostków o takich samych długościach; przewodność tego złącza jest największa a maksymalna gęstość prądu i względna intensywność szumów typu $1/f$ — są najmniejsze.

2. WPLYW NATEŻENIA PRĄDU NA TEMPERATURĘ ZŁĄCZA

Z grona potencjalnych czynników mogących prowadzić do degradacji złącza wielostykowego wyróżnić można:

- a) kolejne przepalanie się mikrostków poczynając od mikrostryku o najmniejszym polu przekroju poprzecznego,

b) zmniejszanie się przekroju poprzecznego mikrostryków tworzących złącze pod wpływem czynników fizyko-chemicznych.

Aby zbadać, który z powyższych czynników może mieć dominujące znaczenie, wyznaczono przyrost temperatury w mikrostryku wywołany przepływającym prądem. Rozkład temperatury w mikrostryku można otrzymać rozwiązując dla niego równanie przewodnictwa.

Niech dany będzie mikrostryk mający postać jednorodnego walca o długości d i promieniu podstawy r . Równanie przewodnictwa ciepła przez mikrostryk można wówczas zapisać następująco [9]

$$\frac{\partial [\Delta T(x, t)]}{\partial t} = \frac{K}{c_0 \varrho} \frac{\partial^2 [\Delta T(x, t)]}{\partial x^2} + \frac{1}{c_0 \varrho} F(x, t) \quad (11)$$

gdzie:

$\Delta T(x, t)$ — przyrost temperatury w chwili t , w przekroju poprzecznym mikrostryku odległym o x od jednej z jego podstaw,

K — współczynnik przewodnictwa cieplnego materiału mikrostryku,

c_0 — ciepło właściwe materiału mikrostryku,

ϱ — gęstość materiału,

$F(x, t)$ — gęstość źródeł cieplnych w chwili t , w przekroju poprzecznym mikrostryku odległym o x od jednej z jego podstaw.

Jeżeli przez mikrostryk przepływa prąd stały o natężeniu I , to gęstość źródeł cieplnych $F(x, t) = I^2 R_0$, gdzie R_0 jest rezystancją jednostki objętości mikrostryku. W rozpatrywanym przypadku $F(x, t)$ jest niezależne od x i t .

Podstawiając

$$a^2 = \frac{K}{c_0 \varrho} \quad \text{ i } \quad f(x, t) = \frac{F(x, t)}{c_0 \varrho} \quad (12)$$

otrzymuje się równanie w postaci

$$\frac{\partial [\Delta T(x, t)]}{\partial t} = a^2 \frac{\partial^2 [\Delta T(x, t)]}{\partial x^2} + f(x, t) \quad (13)$$

Przy założeniu, że istnieje doskonałe odprowadzanie ciepła od obu podstaw mikrostryku, które w związku z tym posiadają stałą temperaturę ($\Delta T = 0$), przewodnictwo ciepła przez mikrostryk będzie opisane równaniem (13) z warunkiem początkowym

$$\Delta T(x, 0) = 0 \quad (14)$$

i warunkami brzegowymi

$$\Delta T(0, t) = 0 \quad \Delta T(d, t) = 0 \quad (15)$$

Poszukując rozwiązania równania (13) w postaci szeregu Fouriera

$$\Delta T(x, t) = \sum_{n=1}^{\infty} T_n(t) \sin \frac{\pi n x}{d} \quad (16)$$

otrzymuje się następujące wyrażenie opisujące rozkład przyrostu temperatury w mikrostryku

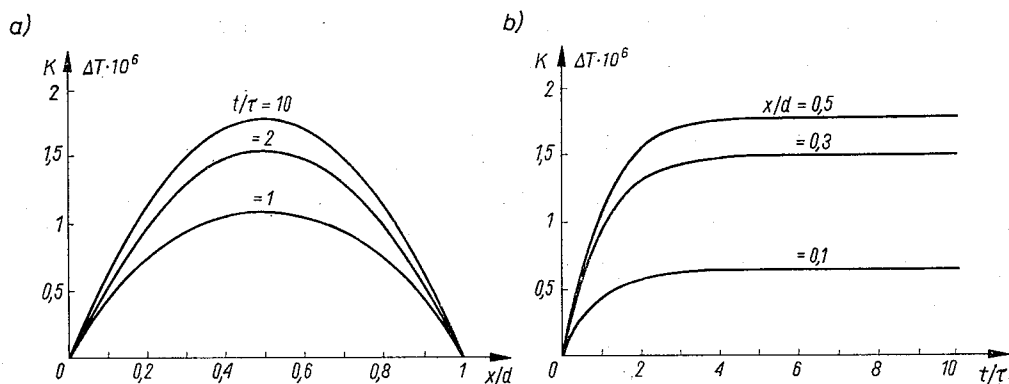
$$\Delta T(x, t) = \frac{4I^2 d^2}{\pi^5 \sigma K r^4} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^3} \cdot \left[1 - \exp\left(-\frac{(2n-1)^2 t}{\tau}\right) \right] \sin \frac{\pi x (2n-1)}{d} \quad (17)$$

gdzie

$$\tau = \frac{\rho c_0 d^2}{\pi^2 K} \quad (18)$$

Otrzymany szereg jest jednostajnie zbieżny dla każdego t i $x \in \langle 0, d \rangle$.

Na rys. 2 przedstawiono przykładowe rozkłady przyrostu ΔT temperatury w jednorodnym mikrostryku aluminiowym o parametrach: $r = 10^{-8}$ m, $d = 10^{-7}$ m, $\sigma = 3,53 \cdot 10^{-7}$ 1/ Ω m, $K = 213$ J/msK, $c_0 = 896$ J/kgK, $I = 1 \mu$ A, $\rho = 2720$ kg/m³.



Rys. 2. Rozkład przyrostu temperatury w mikrostryku walcowym

Ze sporządzonych obliczeń i wykresów wynika, że temperatura w mikrostryku ustala się bardzo szybko, gdyż już po czasie $t = 5\tau$, gdzie $\tau = 1,16 \cdot 10^{-11}$ s, można uważać, że został osiągnięty stan ustalony. Maksymalny przyrost temperatury $\Delta T_{\max}$ występuje dla przekroju poprzecznego mikrostryku, leżącego w jego płaszczyźnie symetrii ($x = d/2$). W stanie ustalonym $\Delta T_{\max}$ jest opisane wyrażeniem:

$$\Delta T_{\max} = \frac{4I^2 d^2}{\pi^5 \sigma K r^4} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)^3} \quad (19)$$

Otrzymane przyrosty temperatury w mikrostryku są bardzo małe mimo, że przyjęta gęstość prądu w mikrostryku znacznie przewyższa gęstości prądu występujące w elementach elektronicznych. W rzeczywistości należy liczyć się ze znacznie większymi od wyznaczonych wartościami przyrostów temperatury w mikrostrykach z uwagi na to, że nie jest dokładnie spełnione założenie dotyczące niezmiennych temperatur obu końców mikrostryku. Mimo to raczej należy wykluczyć hipotezę o możliwości zniszczenia złącza wielo-

stykowego w wyniku kolejnego przepalania się mikrostryków, poczynając od mikrostryku o najmniejszym polu przekroju poprzecznego, jako, że jest bardzo mało prawdopodobne aby ich temperatura osiągnęła wartość temperatury topnienia materiału mikrostryku.

3. DEGRADACJA ZŁĄCZA SPOWODOWANA CZYNNIKAMI FIZYKO-CHEMICZNYMI

Procesy fizyko-chemiczne, takie jak wytrącanie się związków międzymetalicznych, korozja itp., prowadzą do degradacji właściwości złącza, przejawiającej się zmniejszaniem czynnego przekroju poprzecznego mikrostryku w funkcji czasu. Szybkość tych procesów zależy od temperatury T zgodnie z prawem Arrheniusa [2]

$$v = v_{\infty} \exp\left(\frac{-E}{RT}\right) \quad (20)$$

gdzie:

v — szybkość procesu w temperaturze T [K],

v_{∞} — współczynnik proporcjonalności,

E — energia aktywacji,

R — uniwersalna stała gazowa.

Jeżeli w temperaturze T_0 szybkość procesu wynosi v_0 , to szybkość v_1 procesu w temperaturze $T_0 + \Delta T$ będzie określona wyrażeniem

$$v_1 = v_0 \exp\left(\frac{E}{RT_0} \cdot \frac{\Delta T}{T_0 + \Delta T}\right) = v_0 \exp\left(A \frac{\Delta T}{T_0 + \Delta T}\right), \quad (21)$$

gdzie $A = E/RT_0$.

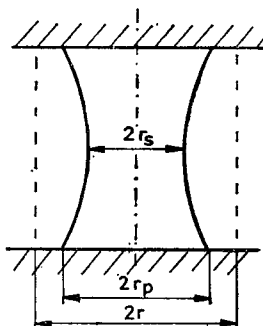
Prawo Arrheniusa można wykorzystać do oceny wpływu natężenia prądu przepływającego przez mikrostryk na czas jego życia.

Niech dla jednorodnego mikrostryku mającego postać walca o długości d i promieniu podstawy r , szybkość procesów prowadzących do degradacji mikrostryku w temperaturze T_0 wynosi v_0 , przy czym v_0 jest rozumiane jako część promienia mikrostryku ulegająca degradacji w jednostce czasu. Wówczas czas życia t_{z0} w temperaturze T_0 jest określony wyrażeniem

$$t_{z0} = \frac{r}{v_0}. \quad (22)$$

Przepływający przez mikrostryk prąd o natężeniu I , wywołuje zmianę rozkładu temperatury w mikrostryku, czego następstwem jest zróżnicowanie szybkości procesów fizyko-chemicznych dla różnych przekrojów poprzecznych mikrostryku. Dla końców mikrostryku utrzymywanych w stałej temperaturze T_0 , promień mikrostryku r_p będzie zmniejszał się ze stałą szybkością v_0 , zaś z największą szybkością będzie malał promień r_s dla $x = d/2$. W wyniku tego, po czasie Δt mikrostryk będzie miał postać pokazaną na rys. 3.

Niech maksymalny przyrost temperatury w mikrostryku, określony wyrażeniem (19), bezpośrednio po przepuszczeniu przez niego prądu I wynosi ΔT . Jeżeli rozpatrzy się odstęp czasu Δt znacznie większy od stałej czasowej τ i jednocześnie znacznie mniejszy

Rys. 3. Przekrój osiowy mikrostryku po czasie $k\Delta t$

od czasu życia mikrostryku w temperaturze T_0 , to można wówczas uważać, że rozkład temperatury w mikrostryku jest ustalony, a procesy prowadzące do jego degradacji w przekroju odległym o $x = d/2$ od końców mikrostryku zachodzą w tym przedziale czasu ze stałą prędkością v_1 opisaną zależnością (21). Wówczas po czasie Δt

$$r_p = r - v_0 \Delta t = r \left(1 - \frac{\Delta t}{t_{z0}} \right) \quad (23)$$

$$r_s = r - v_1 \Delta t = r - \frac{r \Delta t}{t_{z0}} \exp \left(\frac{A \Delta T}{T_0 + \Delta T} \right). \quad (24)$$

Zmiana kształtu przekroju osiowego mikrostryku (rys. 3), przy nieziennej wartości przepływającego przez niego prądu, spowoduje ponowną zmianę rozkładu temperatury w mikrostryku. Analityczne wyznaczenie tego rozkładu dla mikrostryku o postaci pokazanej na rys. 3 byłoby bardzo kłopotliwe, dlatego też do dalszych rozważań zastąpiono ten mikrostryk mikrostrykiem walcowym długości d i promieniu r_x określonym wyrażeniem

$$r_x = \frac{\sqrt{2r_p r_s}}{\sqrt{r_p^2 + r_s^2}}. \quad (25)$$

Taki wybór promienia r_x zapewnia dominujący wpływ przewężenia (promienia r_s) na parametry elektryczne mikrostryku oraz sprawia, że gdy $r_s = 0$ to i $r_x = 0$.

Wzrost gęstości prądu wywołany zmniejszeniem przekroju poprzecznego mikrostryku spowoduje teraz dalszy przyrost temperatury mikrostryku a tym samym wzrost szybkości procesów prowadzących do jego degradacji i dalsze zmniejszanie się promienia r_s . Proces taki będzie przebiegał ze wzrastającą szybkością aż do całkowitego uszkodzenia mikrostryku, tzn. aż do momentu, w którym promień r_s stanie się równy zeru. W ramach realizowanej pracy przeprowadzono symulację numeryczną procesów prowadzących do uszkodzenia mikrostryku, korzystając z podanych poniżej zależności.

Jeżeli w chwili początkowej $t = 0$ temperatura $T = T_0 = 293$ K oraz promienie początkowe spełniają zależność $r_{s0} = r_{p0} = r_{x0} = r$, to po czasie $t = k \Delta t$, gdzie $k \in \mathbb{N}$ a $\Delta t \ll t_{z0}$ i $\Delta t \gg \tau$, słuszną są wyrażenia

$$r_{p,k} = r \left(1 - \frac{k \Delta t}{t_{z0}} \right), \quad (26)$$

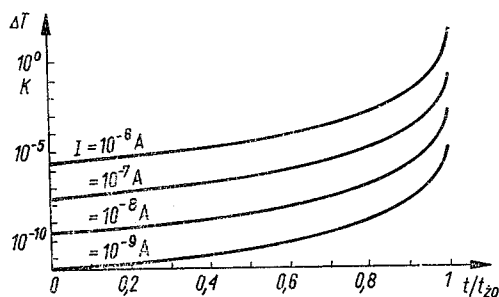
$$r_{s,k} = r_{s,k-1} - \Delta t v_0 \exp\left(A \frac{\Delta T_k}{T_0 + \Delta T_k}\right), \quad (27)$$

$$r_{x,k} = \frac{\sqrt{2} r_{p,k} \cdot r_{s,k}}{\sqrt{r_{p,k}^2 + r_{s,k}^2}}, \quad (28)$$

$$\Delta T_k = \frac{A_0 I^2}{r_{x,k-1}^4}, \quad (29)$$

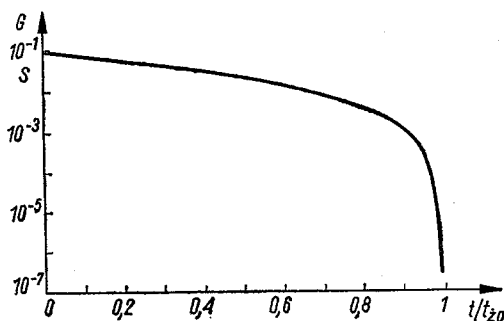
gdzie A_0 — jest współczynnikiem proporcjonalności wynikającym z zależności (19). Obliczenia numeryczne na mikrokomputerze C — 64 wykonano dla mikrostryku aluminiowego, przyjmując za [2] energię aktywacji E procesów chemicznych równą $60,76 \cdot 10^3$ J/mol. Odstęp czasu Δt wybrano równy $0,001 t_{z0}$.

Wyniki obliczeń przedstawiono na rys. 4—7. Zgodnie z oczekiwaniami, w miarę upływu czasu t , obserwuje się coraz większe przyrosty ΔT temperatury mikrostryku, tym większe,

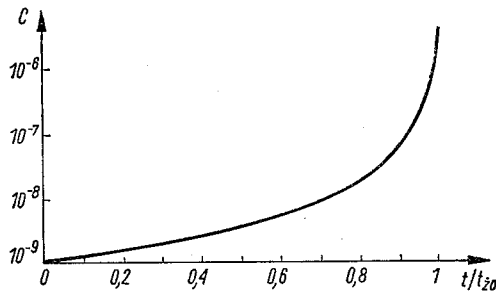
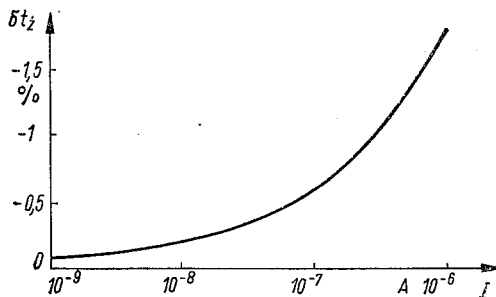


Rys. 4. Zależność $\Delta T = f(t/t_{z0})$

im większa jest wartość natężenia prądu (rys. 4). Wzrostowi temperatury mikrostryku towarzyszy zmniejszanie się jego przewodności G (rys. 5) i wzrost względnej intensywności C szumów typu $1/f$ (rys. 6). Zmiany G i C zachodzą według tej samej krzywej, niezależnie od natężenia I prądu przepływającego przez mikrostryk, przy czym wzrostowi natężenia



Rys. 5. Zależność $G = f(t/t_{z0})$

Rys. 6. Zależność $C = f(t/t_{z0})$ Rys. 7. Zależność $\delta t_z = f(I)$

prądu towarzyszy niewielkie skrócenie czasu życia mikrostryku. Zależność względnej zmiany czasu życia σt_z mikrostryku, zdefiniowanej jako

$$\sigma t_z = \frac{t_z - t_{z0}}{t_{z0}}, \quad (30)$$

gdzie t_{z0} — czas życia mikrostryku przy prądzie o natężeniu równym zeru,

t_z — czas życia mikrostryku przy prądzie o natężeniu I ,

od natężenia prądu przepływającego przez mikrostryk pokazano na rys. 7. Zmianie parametrów elektrycznych mikrostryków będzie oczywiście towarzyszyła zmiana parametrów elektrycznych złącza wielostykowego.

4. PODSUMOWANIE

Z rozważań przeprowadzonych w niniejszej pracy wynika, że

- podstawowe parametry elektryczne modelu złącza wielostykowego zależą od wymiarów mikrostryków tworzących złącze, przy czym najlepszymi właściwościami charakteryzuje się model złożony z mikrostryków o stałej długości,
- aczkolwiek dla rzeczywistych złączy metal-metal należy oczekiwać większych przyrostów temperatury niż to wynika z przeprowadzonych obliczeń, to należy jednak raczej odrzucić hipotezę o możliwości zniszczenia złącza wielostykowego w wyniku przepalania się kolejnych mikrostryków pod wpływem przepływającego prądu, poczynając od mikrostryku o najmniejszym przekroju poprzecznym; za proces degradacji złącza byłyby odpowiedzialne głównie czynniki fizyko-chemiczne,

- natężenie prądu przepływającego przez złącze w niewielkim stopniu wpływa na czas życia mikrostryków a tym samym i na czas życia złącza,
- procesowi degradacji mikrostryku towarzyszy wzrost względnej intensywności szumów typu $1/f$; charakter zmian intensywności szumów typu $1/f$ wywołany procesem degradacji złącza wielostykowego, będzie zależał od wymiarów mikrostryków tworzących to złącze.

BIBLIOGRAFIA

1. E. Fortuna, J. Jabłoński: *Przegląd czynników wpływających na jakość połączeń drutowych w układach scalonych oraz metod testowania tych połączeń*. Wydawnictwo PIE, I.S.P., luty 1976
2. Praca pod red. H. Schneidera: *Nadieżność elektronnych elementów i sistem*. Mir, Moskwa, 1977
3. Praca zbiorowa: *Nadieżność i diagnostyka półprzewodnikowych przyborów*. Sztinica, Kiszyniów, 1984
4. Praca zbiorowa pod kierunkiem S. Stępnia: *Poradnik konstruktora sprzętu elektronicznego*. Warszawa: WKiŁ, 1981
5. L. K. J. Vandamme, R. P. Tijburg: $1/f$ noise measurements for characterizing multispot low ohmic contacts. *Journal of Applied Physics*, t. 47, nr 5, 1976
6. L. H. F. Ortman, L. K. J. Vandamme: *Characterization of Impulse-Fritting Procedures of Contacts nby Measuring $1/f$ Noise*, *Applied Physics*, t. 9, ss. 147—151, 1976
7. A. Kusy: *Struktura, mechanizm przewodnictwa oraz szumy typu $1/f$ rezystywnych warstw grubych*. Wydawnictwo Politechniki Rzeszowskiej, 1979
8. E. P. Honig: *$1/f$ noise of bodies of arbitrary shape and of point contacts*. *Philips. Res. Repts*, 29, 253—260, 1974
9. A. N. Tichonow, A. A. Samarski: *Równania fizyki matematycznej*. Warszawa: PWN, 1963

H. GIERASIMOWICZ

ELECTRICAL PROPERTIES OF MULTICONTACT METAL-METAL JUNCTION MODELS

Summary

The electrical properties of various metal-metal junction models composed of many contacts are compared. The analysis of some factors which may cause the junction degradation enables to conclude that the dominant role is played by physico-chemical factors and that the occurrence of the junction destruction by the current flow is less probable. The junction degradation process is accompanied by changes of the relative $1/f$ noise intensity. This fact may be applied to the elimination of potentially less reliable junctions.

H. GIERASIMOWICZ

PROPRIÉTÉS ÉLECTRIQUES DES MODÈLES Á PLUSIEURS CONTACTS
DE LA JONCTION MÉTAL-MÉTAL

Résumé

Dans cet article on a comparé les propriétés électriques des modèles des jonctions métal-métal composées de plusieurs micro-contacts de taille variable. En examinant les facteurs, qui peuvent causer la dégradation des jonctions, on a constaté que les facteurs physico-chimiques jouent le rôle dominant, par

contre il est peu probable que la jonction soit détériorée par le courant électrique y passant. Le processus de la dégradation est accompagné par la variation de l'intensité relative du bruit $1/f$, ce qui peut être utilisé pour l'élimination des jonctions à fiabilité inférieure.

H. GIERASIMOWICZ

ELEKTRISCHE EIGENSCHAFTEN MEHRPOLIGER MODELLE DER METALL-METALL-VERBINDUNGEN

Zusammenfassung

In dieser Bearbeitung wurden die elektrischen Eigenschaften mehrpoliger Modelle von Metall-Metall-Verbindungen, die aus Mikrokontakten verschiedener Größen bestehen, miteinander verglichen. Mit Hilfe der Analyse von Faktoren, die zur Degradation der Verbindungen führen können, wurde festgestellt, daß dabei die Hauptrolle die physikalisch-chemischen Faktoren spielen. Dagegen erscheint es weniger wahrscheinlich, daß die Zerstörung der Verbindungen als Folge des durchfließenden Stromes auftritt. Während des Degradationsprozesses erfolgt gleichzeitig eine Veränderung der relativen Rauschintensität des Typs $1/f$, die zur Eliminierung der potentiell unstätigen Verbindungen genutzt werden kann.

Х. ГЕРАСИМОВИЧ

ЭЛЕКТРИЧЕСКИЕ СВОЙСТВА МНОГОКОНТАКТНЫХ МОДЕЛЕЙ СОЕДИНЕНИЯ МЕТАЛЛ-МЕТАЛЛ

Резюме

Сравнены электрические свойства многоконтактных моделей соединений металл-металл, которые состоят из микроконтактов различных размеров. Анализируя факторы, которые могут вести к деградации соединения установлено, что главную роль играют физико-химические факторы, зато мало вероятно, чтобы протекающий через соединение ток мог вести к уничтожению контакта. Деградацию соединения сопровождает изменение относительной интенсивности $1/f$ шума, что можно использовать при браковке потенциально ненадежных соединений.

Algorytmy detekcji lokalnych wad obrazów

ZBIGNIEW M. WÓJCIK

Przemysłowy Instytut Elektroniki, Warszawa

Otrzymano 1985.10.02

Autoryzowano do druku 1987.06.05

Artykuł zawiera skondensowany przegląd metod detekcji lokalnych wad obrazów. Metody opisane w tej pracy polegają na badaniu nieprawidłowości obrazu wewnątrz pewnego założonego kolistego podobszaru operatora. Przedstawiono w nim również ulepszoną wersję metody porównywania obrazów ze wzorcem.

1. WPROWADZENIE

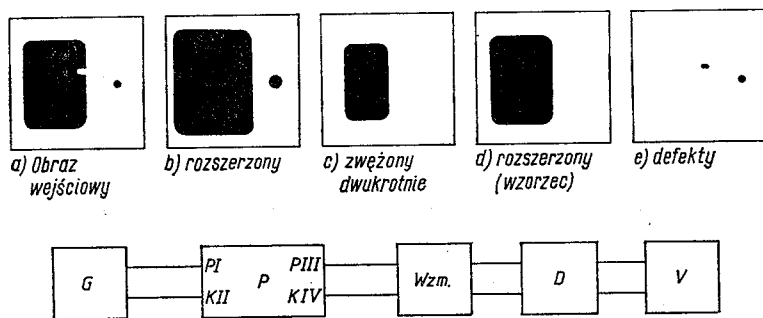
Metody opisane w tej pracy polegają na badaniu nieprawidłowości obrazu wewnątrz pewnego założonego kolistego podobszaru operatora. Operator taki jest przesuwany po całym obrazie, punkt po punkcie, zapewniając w ten sposób przebadanie całej powierzchni. Wykrywane są tu tylko nieprawidłowości lokalne, które mogą być dostrzeżone wewnątrz pola operatora. Nie można więc sprawdzać zachowania wymiarów całych obiektów, znacznie przekraczających rozmiarami pole operatora. Są natomiast łatwo wykrywalne wszystkie drobne defekty, np. plamki i prześwity obiektów oraz postrzępienia brzegów. Przedstawione zostaną następujące metody detekcji wad obrazów:

- a) przez rozszerzanie — zwężanie obiektów obrazu (Eijri, [5]);
- b) przez badanie spójności operatora kolistego z jego centrum (Danielsson, [1]);
- c) przez pomiar liczby obiektów operatora kolistego (jego liczby Eulera), (Kruse, [1]);
- d) przez pomiar liczby obiektów znajdujących się na brzegu operatora (tj. liczby Eulera brzegu operatora), (Kruse, [1]);
- e) technikę wirującego sąsiedztwa.

Natomiast wszystkim przedstawionym metodom wykrywającym lokalne nieprawidłowości obrazu przeciwstawiona zostanie klasyczna metoda porównywania obrazu badanego ze wzorcem na poziomie sygnałów.

2. ROZSZERZANIE — ZWĘŻANIE OBIEKTÓW

W metodzie rozszerzania-zwężania obiektów obrazu (Eijri [5]), wzorzec, z którym porównywany jest na poziomie sygnałów obraz badany, jest otrzymywany z obrazu badanego przez następujące po sobie kolejno operacje rozszerzania — zwężania — zwężania — rozszerzania obiektów. Rys. 1 przedstawia ideę tej metody.

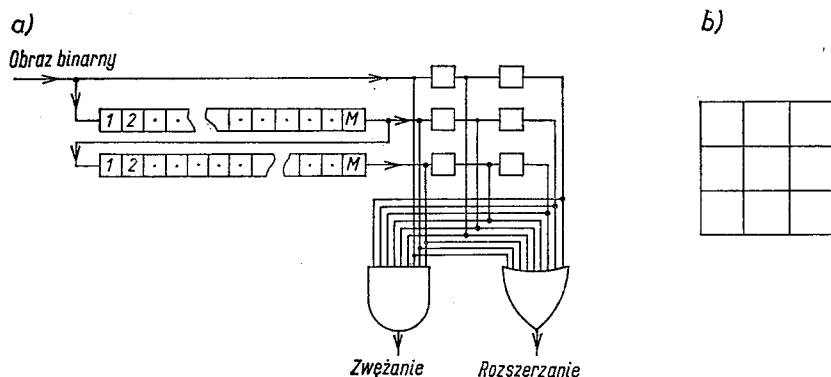


Rys. 1. Kolejne etapy detekcji defektów metodą rozszerzania — zwężania obiektów obrazu

Obiekty obrazu wejściowego zawierającego dwa defekty obrazu: plamkę i prześwit na brzegu obiektu, są rozszerzane we wszystkich kierunkach o wielkość przewyższającą połowę rozmiaru (np. promień) największego spodziewanego defektu. W wyniku tego rozszerzenia znikają prześwity w obiektach (rys. 1b). Rozszerzone obiekty są następnie dwukrotnie zwężane we wszystkich kierunkach o tą samą wielkość. W wyniku tego zwężenia znikają plamy w tle obiektów (rys. 1c). Po rozszerzeniu obiektów o tę samą wielkość otrzymywany jest obraz wzorcowy, pozbawiony defektów mniejszych od założonego rozmiaru (rys. 1d). Po porównaniu obrazu wejściowego (rys. 1a) z tym wzorcem (rys. 1d), np. za pomocą operacji *exclusive-or*, „ $\oplus$ ”, wydzielane są defekty obrazu binarnego (rys. 1e).

Zasadniczym mankamentem tej metody jest traktowanie jako defekty ostrych naroży oraz drobnych postrzępień brzegów obiektów, nie będących rzeczywistymi defektami, spowodowanych ucyfrowianiem obrazu, co nie pozwala na stosowanie jej do wykrywania wad masek lub układów scalonych.

Interesująca jest realizacja elektroniczna tej metody. Rys. 2 przedstawia ideę układu cyfrowego przeprowadzającego operacje rozszerzenia i zwężania obiektów obrazu binarnego. Wejściowy obraz dwuwartościowy (tj. sprowadzalny do obrazu czarno-białego) jest przekazywany na zestaw dwóch rejestrów przesuwających M-bitowych połączonych szeregowo (M jest liczbą komórek w jednej linii obrazu). Wejście tego obrazu oraz wyjście



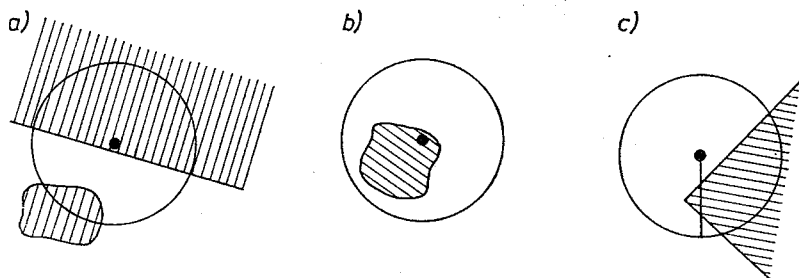
Rys. 2. Wyszczególniony układ elektroniczny zwężania i rozszerzania obiektów obrazu (a) przez mnożenie i sumowanie logiczne wartości cyfrowych komórek operatora dziewięciokomórkowego (b)

każdego z tych dwóch długich rejestrów przekazywane są na zestaw sześciu rejestrów przesuwających 1-bitowych, których wszystkie wyjścia stanowią operator dziewięciokomórkowy (rys. 2b). Operator ten przesuwa się więc po całym obrazie element po elemencie w takt impulsów synchronizujących przetwornika optyczno-elektrycznego (np. kamery TV). Operacja iloczynu logicznego wartości cyfrowych tego operatora daje w wyniku zwęźanie obiektów obrazu o jedną komórkę, natomiast suma logiczna rozszerza obiekty.

3. BADANIE SPÓJNOŚCI OBIEKTÓW OPERATORA I CENTRUM OPERATORA

W metodzie badania spójności obiektów operatora kolistego z jego centrum (Daniels-son [1]) analizuje się obraz za pomocą kolistego operatora. Wykrywane są odległości pomiędzy obiektami lub szerokości obiektów mniejsze od promienia operatora. Defekt obrazu jest wykrywany, gdy centrum operatora nie jest spójne z jakimś obiektem operatora o tej samej wartości cyfrowej do centrum. Na przykład, na rys. 3a wykrywany jest defekt obrazu, ponieważ niewielki obiekt nie jest spójny z centrum operatora.

Mankamentem tej metody jest niewykrywanie drobnych defektów mieszczących się całkowicie w polu operatora (rys. 3b: istnieje defekt mniejszy od promienia operatora,



Rys. 3 Przypadki wykrywania defektów metodą badania spójności obiektów pokrytych operatorem kulistym: a) defekt jest wykrywany, ponieważ jego odległość od dużego obiektu jest mniejsza od promienia operatora; b) defekt mniejszy od operatora (jak również każdy większy) nie jest wykrywany, gdy jego odległość od jakiegoś dużego obiektu jest większa od promienia operatora; c) nieprawidłowe traktowanie ostrego naroża prawidłowego obiektu jako defektu

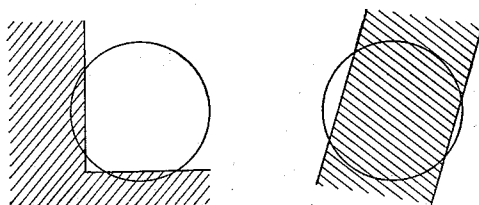
jednakże wszystkie punkty tego obiektu są spójne z centrum). Drugim mankamentem metody jest duża złożoność procesu badania spójności komórek obrazu. Uproszczone sprawdzanie spójności polega na badaniu połączenia centrum z komórkami rozmieszczonymi wzdłuż promienia operatora, jednakże w realizacji takiej jako defekty wykrywane są także ostre naroża prawidłowych obiektów (rys. 3c).

4. ANALIZA LICZBY EULERA W POLU OPERATORA

W metodzie wyznaczania liczby Eulera (tj. liczby obiektów) w polu operatora lub liczby obiektów na obwodzie operatora (Kruse [1]) wykrywane są odległości pomiędzy obiektami lub szerokości obiektów mniejsze od średnicy operatora kolistego. Jeżeli liczba

obiektów (albo liczba fragmentów tła) jest większa od 1, wykrywana jest wada obrazu. W analogiczny sposób, jeżeli liczba spójnych odcinków obwodu operatora przekracza 1, wykrywany jest także defekt obrazu. Na przykład, liczba obiektów w polu operatora lub liczba odcinków obwodu operatora na rys. 3c jest równa 1, defekt obrazu nie jest więc wykrywany. Natomiast na rys. 3a liczba obiektów w polu operatora (lub liczba spójnych odcinków obwodu operatora lub liczba spójnych odcinków obwodu operatora reprezentujących tło) wynosi 2, wykrywany jest więc defekt. Na rys. 4a wykrywana jest zbyt wąska ścieżka, ponieważ liczba fragmentów tła (lub liczba spójnych odcinków obwodu operatora reprezentujących tło) wynosi 2.

Mankamentem metody jest niewykrywanie drobnych defektów obrazu (jak na rys. 3b), mieszczących się całkowicie w polu operatora: liczba Eulera w polu operatora jest równa 1.



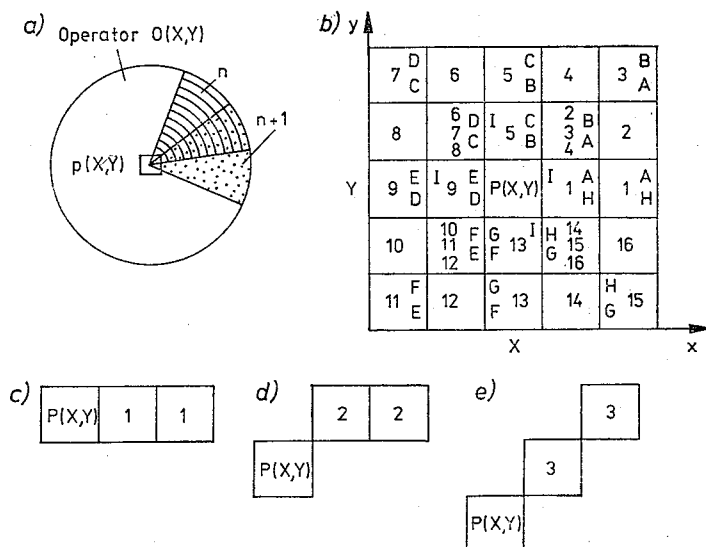
Rys. 4 Wykrycie ścieżki węższej od średnicy operatora (a) przez wyznaczenie liczby Eulera w polu operatora (liczby fragmentów tła); nieprawidłowe traktowanie ostrych naroży obiektów jako defektów (liczba obiektów w polu operatora większa od jedności)

Wykrywane są także jako defekty ostre naroża obiektów (rys. 4b: liczba obiektów w polu operatora wynosi 2). W przypadku płytek drukowanych, kolejnym mankamentem metody jest wykrywanie jako defekty wszystkich ścieżek prawidłowych, które są cieńsze od badanych aktualnie, albo niemożliwość wykrywania przewężeń ścieżek szerokich (przewężeń szerszych od ścieżek najcieńszych) gdy średnica operatora jest mniejsza od szerokości ścieżek najcieńszych. Narzuca to konieczność osobnego śledzenia ścieżek o różnych rozmiarach, co w konsekwencji uniemożliwia stosowanie metody do badań ścieżek układów scalonych lub płytek drukowanych: metoda ta rozwija się w inną metodę. Analogiczną niedogodność ma metoda wykrywania defektów przez badanie spójności obiektów, przykrytych operatorem, względem jego centrum.

5. TECHNIKA WIRUJĄCEGO SĄSIEDZTWA

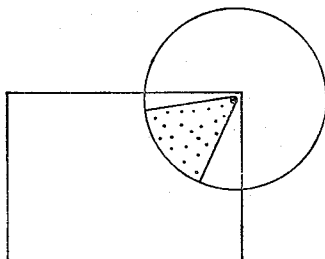
Technika zachowująca brzegi obiektów polega na badaniu jednorodności stopni szarości pewnego sąsiedztwa, wirującego wewnątrz kolistego operatora, względem centrum tego operatora (rys. 5). Jednym z badanych sąsiedztw, oprócz wirującego, jest także sąsiedztwo symetryczne względem centrum operatora i posiadające taką samą powierzchnię jak to wirujące (rys. 5).

Lemat 1. Technika wirującego sąsiedztwa zachowuje brzegi wszystkich tych obiektów, których minimalne rozmiary są większe od rozmiarów wirującego sąsiedztwa.



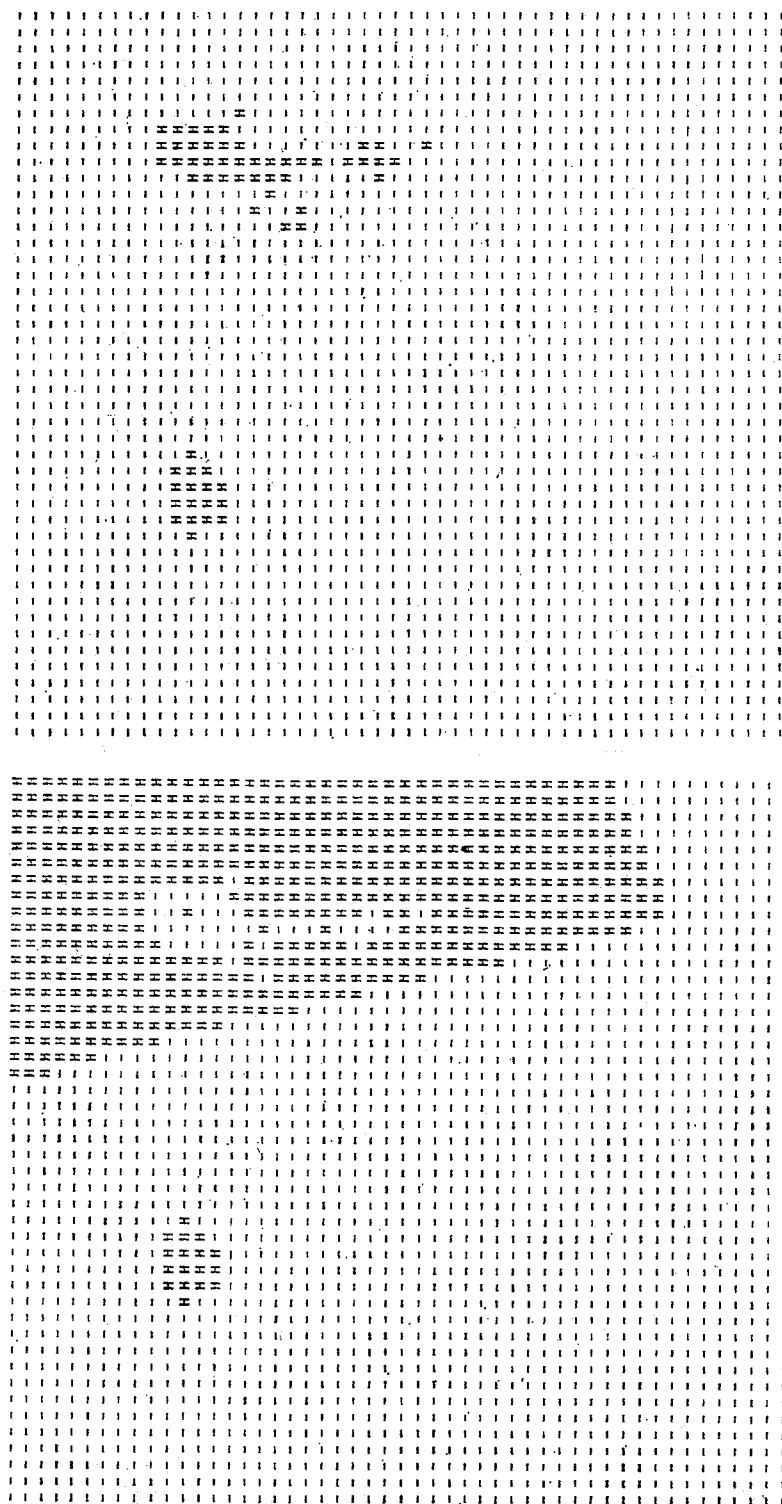
Rys. 5 Dwie pozycje wirującego sąsiedztwa (wycinka koła) (a) wewnątrz operatora kolistego oraz 25-komórkowy operator (b) z 16 pozycjami sąsiedztwa 2-komórkowego zaznaczonymi cyframi od 1 do 16. Duże litery od A do H oznaczają 8 pozycji operatora 4-komórkowego stosowanego do detekcji wad obrazu dwuwartościowego. Typowe formy sąsiedztwa 2-komórkowego są pokazane na rys. (c), (d) i (e)

Dowód. Weźmy pod uwagę najbardziej niekorzystny przypadek, gdy centrum operatora zawierającego wirujące sąsiedztwo znajduje się w rogu obiektu (rys. 6). Centrum operatora należy więc także do obiektu i wymiary wirującego sąsiedztwa są mniejsze od wymiarów obiektu. Istnieje zatem takie położenie wirującego sąsiedztwa, że w przypadku braku zakłóceń i wad obrazu spełniona będzie zasada jednorodności centrum operatora z sąsiedztwem w tym położeniu. Ponieważ pozycje wirującego sąsiedztwa są rozłożone symetrycznie wokół centrum operatora i badany jest cały aktywny obszar obrazu, brzegi obiektów są zachowywane w całym aktywnym obszarze obrazu. Spełniony jest więc podstawowy wa-



Rys. 6 Najbardziej niekorzystny przypadek badania jednorodności centrum operatora, znajdującego się w ostrych narożach regularnego obiektu, z wirującym sąsiedztwem którego rozmiary są mniejsze od najmniejszych rozmiarów obiektu

runek reprezentowania obrazów: wszystkie mniej niekorzystne przypadki od prezentowanego na rys. 6 są zatem uwzględniane na całym obrazie. Technika wirującego sąsiedztwa nie traktuje więc komórek brzegowych obiektów jako nieprawidłowości obrazu, cbdu.



Rys. 7 Ucyfrowiony fragment (a) maski półprzewodnikowej zawierającej przeswit i plamkę oraz wynik automatycznej detekcji tych defektów (b) metodą wirującego sąsiedztwa

Wniosek. Jedynie w przypadku, gdy badana komórka różni się pod względem badanej właściwości (np. stopni szarości) od wirującego sąsiedztwa dla wszystkich jego pozycji, zasada jednorodności nie jest spełniona (wykrywane jest wtedy zakłócenie lub defekt: badana komórka reprezentuje nieprawidłowość obrazu).

Rys. 7a przedstawia fragment maski półprzewodnikowej z dwoma defektami, a rys. 7b pokazuje wynik automatycznej detekcji tych defektów za pomocą wirującego sąsiedztwa [3, 4]. Z porównania tych rysunków wynika, że wirujące sąsiedztwo wykrywa prawidłowo drobne defekty mieszczące się w polu operatora oraz nie traktuje, jako defektów, naroży obiektów oraz drobnych postrzępień brzegów obiektu.

6. PORÓWNYWANIE OBRAZU ZE WZORCEM NA POZIOMIE SYGNAŁÓW

Najbardziej naturalne w sensie elektronicznym jest porównywanie bezpośrednio sygnałów pochodzących z przedmiotu badanego i z jego wzorca. Sygnał elektroniczny reprezentujący obraz badany jest porównywany bezpośrednio za pomocą relacji *exclusive-or* „ $\oplus$ ” z sygnałem reprezentującym wzorec. Obraz badany i wzorec mogą być wprowadzane do komputera za pomocą dwóch oddzielnych przetworników optyczno-elektrycznych, zsynchronizowanych ze sobą. W wyniku porównania otrzymywane są różnice pomiędzy tymi dwoma obrazami, a więc np. defekty i zakłócenia obrazu badanego.

W odróżnieniu od rozpoznawania, po którym następuje porównanie struktury jakościowej (np. grafu odzwierciedlającego uporządkowanie cech obiektu) i ilościowej (miar cech obiektu) z analogicznymi strukturami wzorców, porównanie obrazu ze wzorcem na poziomie sygnałów nie wymaga wyznaczania cech obiektów. Niezbędne są natomiast:

- a) scentrowanie obiektu badanego z obiektem wzorcowym dokładniejsze od rozmiaru jednej komórki;
- b) liniowość i powtarzalność pracy przetwornika optycznoelektrycznego lepsze od rozmiaru pojedynczej komórki.

Oznacza to, że gdy jedna komórka obrazu ucyfrowionego odpowiada jednemu mikrometrowi kwadratowemu obrazu badanego, liniowość i powtarzalność pracy przetwornika optyczno-elektrycznego muszą być lepsze niż $1 \mu\text{m}^2$ po stronie obrazu badanego. Technicznie precyzja taka jest wyjątkowo trudna do uzyskania.

Niespełnienie warunku (a) lub (b) powoduje błędne sygnalizowanie defektu obrazu w wyniku niedoskonałości skanowania. Można temu częściowo zapobiec, stosując np. zwężanie obiektów obrazu wejściowego i rozszerzanie obiektów wzorca w celu wykrywania plam (wad tła) oraz następnie operacji wynikania dla porównania obrazu z wzorcem:

$$\begin{array}{ccc} (\text{obraz wejściowy} \Rightarrow \text{wzorec}) = \text{defekty tła} & & \\ (\text{zwężony}) & (\text{rozszerzony}) & (1) \end{array}$$

Proces porównywania ilustruje tabela 1. W wyniku powyższego porównania nie mogą więc być wykrywane defekty w obszarze obiektów, ponieważ z wartości 0 nic nie można wynioskować.

Tabela 1

Wartość cyfrowa komórki		Wynik
obrazu wejściowego	wzorca	
1	1	brak defektu
1	0	defekt tła
0	0	—
0	0	—

W celu wykrywania prześwitów (defektów w obszarze obiektów) należy zastosować niezależnie od wykonywania operacji (1) rozszerzanie obiektów obrazu wejściowego i zważanie obiektów wzorca, a następnie operację wynikania (tabela 2):

$$\begin{aligned}
 &(\text{wzorec} \Rightarrow \text{obraz wejściowy}) = \text{defekty obiektów} \\
 &(\text{zważony}) \quad (\text{rozszerzony})
 \end{aligned}
 \tag{2}$$

Tabela 2

Wartość cyfrowa komórki		Wynik
wzorca	obrazu badanego	
1	1	brak defektu
1	0	defekt obiektu
0	1	—
0	0	—

Operacje zważania i rozszerzania obiektów uodporniają do pewnego stopnia wyniki porównywania obrazu ze wzorcem na nieliniowość przetwornika optyczno-elektrycznego lub na niescentrowanie obrazów o sumę promieni zważania i rozszerzania, jednocześnie jednak kompensują defekty tła o rozmiarach mniejszych od promienia zważania obrazu wejściowego i defekty obiektu o rozmiarach mniejszych od promienia rozszerzania obrazu badanego. Przeprowadzenie dodatkowej operacji badania niejednorodności brzegów obiektów pozwala jednak na wykrywanie tych defektów.

Za pomocą metody porównywania obrazu ze wzorcem na poziomie sygnału można najłatwiej wykrywać duże, regularne defekty. Metodę tę należy jednak stosować w połączeniu z metodami wykrywania i usuwania lokalnych nieprawidłowości obrazu dla uchronienia się przed możliwościami potraktowania niedoskonałości przetwornika optyczno-elektrycznego jako defektów obrazu.

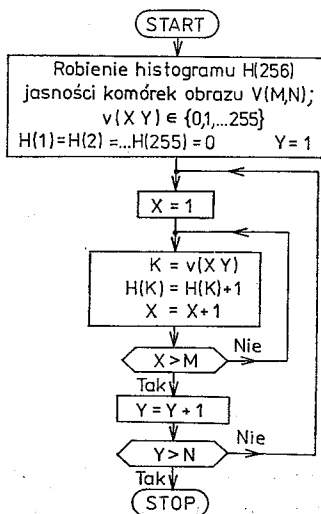
7. ANALIZA OBRAZÓW WIELOWARTOŚCIOWYCH

Przedstawione powyżej metody odnoszą się bezpośrednio do obrazów czarno-białych lub sprowadzalnych do takich obrazów. Również obrazy posiadające wiele poziomów szarości (np. obrazy mikrostruktur półprzewodnikowych) dają się reprezentować w postaci szeregu obrazów dwuwartościowych. Poziom progu kwantowania jest ustalony pomiędzy kolejnymi poziomami jasności (lub szarości) występującymi w badanym obrazie. Możliwe

jest automatyczne ustalanie wielkości progu na podstawie komputerowej analizy histogramu stopni szarości obrazu, tj. wykresu liczebności komórek obrazu dla każdego możliwego stopnia zaczernienia obrazu.

Algorytm wykonywania histogramu, zapisywanego w tablicy H , jest przedstawiony na rys. 8. Każda z możliwych wartości $v(X, Y)$ komórki jest pozycją tablicy H (adresem), której zawartość jest zwiększana o 1 po każdorazowym odczycie $v(X, Y)$.

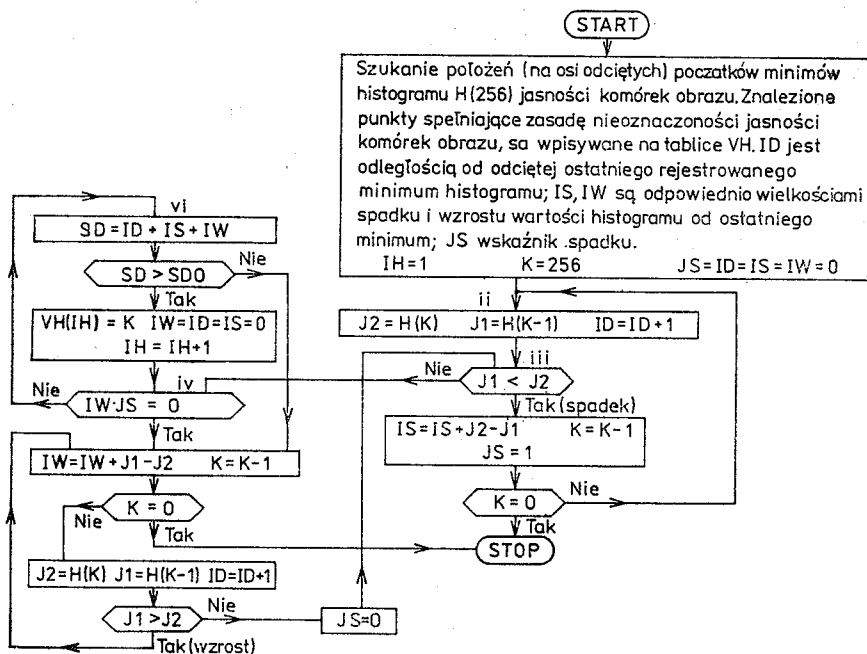
Poziomy kwantowania można ustalać automatycznie na wartościach odciętych początków minimów histogramu jasności komórek obrazu.



Rys. 8 Algorytm robienia histogramu wartości komórek obrazu

Lemat 2. Wydzielanie odciętych początków minimów histogramu jasności komórek (lub ogólnie: kolejnych minimów funkcji) jest wykonywane przez (rys. 9):

- i. Odczyt ostatniej i przedostatniej wartości histogramu.
Ustawienie na zero:
 - a) odległości od odciętej ostatniego rejestrowanego minimum histogramu;
 - b) łącznego spadku i wzrostu rzędnych liczonych od ostatniego zarejestrowanego minimum histogramu;
 - c) wskaźnika spadku rzędnej histogramu;
- ii. Zwiększenie o jeden odległości od odciętej ostatniego zarejestrowanego minimum histogramu;
- iii. a) jeśli bieżąca rzędna histogramu nie jest mniejsza od poprzedniej rzędnej — oznacza to wzrost lub stałą wartość histogramu — następuje skok do (iv);
 b) jeżeli bieżąca rzędna histogramu jest mniejsza od poprzedniej rzędnej — oznacza to spadek histogramu — do dotychczasowego spadku histogramu dodawany jest bieżący spadek, wskaźnik spadku ustawiany jest na jeden i wyznaczany jest indeks następnego punktu histogramu (indeks ten zmniejszany jest o 1);
 c) jeżeli wszystkie punkty histogramu zostały przeanalizowane, koniec analizy histogramu, w przeciwnym przypadku skok do początku (ii);



Rys. 9 Algorytm wyszukiwania początków minimów histogramu

- iv. Jeżeli iloczyn wzrostu histogramu przez wskaźnik spadku jest różny od zera (oznacza to identyfikację początku minimum lokalnego histogramu) — skok do (vi);
- v. a) Do dotychczasowego wzrostu histogramu od ostatniego minimum dodawany jest bieżący wzrost i wyznaczany jest indeks następnego punktu histogramu. Jeżeli cały histogram został przebadany, koniec pracy algorytmu;
- b) zwiększenie o jeden odległości od odciętej ostatniego minimum histogramu;
- c) jeżeli bieżąca rzędna histogramu jest większa od poprzedniej rzędnej, oznacza to wzrost histogramu i następuje skok do początku (v);
- d) wskaźnik spadku histogramu jest gaszony i następuje skok do (iii);
- vi. Jeżeli łączna odległość od ostatniego minimum (będąca sumą odległości od odciętej ostatniego minimum oraz spadku i wzrostu wartości histogramu od ostatniego minimum) jest większa od założonego progu, rejestrowana jest bieżąca odcięta histogramu jako odcięta zidentyfikowanego początku minimum histogramu, następnie zerowane są wszystkie składniki łącznej odległości od ostatniego minimum i następuje skok do (iv); w przeciwnym razie skok do (v).

Dowód. Początek minimum funkcji, w tym przypadku histogramu, jest identyfikowany, jeżeli wartość iloczynu wzrostu histogramu przez wskaźnik spadku histogramu nie jest równy zero (punkt (iv)). Oznacza to bowiem w tej konfiguracji algorytmu, że wskaźnik spadku jest zapalony (tj. równy 1; gdyby wskaźnik spadku był równy zero, rozpatrywany iloczyn byłby także równy zero) oraz że nie ma już spadku, ponieważ w poprzedzającej operacji w (iii. a) stwierdzony został wzrost lub stała wartość histogramu. Niezerowy wzrost histogramu wykazywany przez rozpatrywany iloczyn świadczy o tym, że wzrost

ten musiał mieć miejsce przed zakończonym aktualnie spadkiem. Identyfikacja końca tego spadku jest więc równoznaczna z identyfikacją minimum lokalnego histogramu. Porównywanie odległości tego wyznaczonego minimum od poprzedniego minimum histogramu z zadaną wartością progową umożliwia pomijanie minimów mało odległych od poprzednio zidentyfikowanego i zarejestrowanego minimum lokalnego, cbdu.

WNIOSKI

Za pomocą przedstawionych operatorów analizujących lokalne nieprawidłowości obrazu można wykazywać jedynie drobne defekty lub postrzępienia powstałe w wyniku dużych braków lub plam obrazu. Nieobecność całego podobszaru nie jest wykrywana. Nie wykrywa się też dużej wady o regularnych brzegach. Defekty takie można natomiast wykrywać przez bezpośrednie porównywanie obrazu ze wzorcem, wadą takiego podejścia jest jednak trudność natury technicznej — konieczność niezwykle precyzyjnego scentrowania wzorca z obrazem.

BIBLIOGRAFIA

1. P. Danielsson, B. Kruse: *Distance Checking Algorithms*. Comp. Graph. Image Proc., 1979 vol. 4, s. 339—376
2. Z. M. Wójcik: *Cyfrowa analiza lokalnych nieprawidłowości obrazów masek i płytek drukowanych*. Elekotrnizacja, 1984, z. 12.
3. Z. M. Wójcik: *Automatic Detection of Semiconductor Mask Defects*. Microelectronics and Reliability, Perg. Press, 1976, vol. 15, s. 585—593
4. Z. M. Wójcik: *A System for an Automatic Detection of Defects of Semiconductor Masks and PC Boards*. Electron Technology, Warszawa, 1977, vol. 10, no 3, s. 79—96
5. M. Ejiri, T. Uno, M. Mese, S. Ikeda: *A Process for Defects in Complicated Patterns* Comp. Graph. Image Proc., 1973, vol. 2, s. 326—339

Z. M. WÓJCIK

DETECTION ALGORITHMS OF LOCAL DEFECTS OF IMAGES

S u m m a r y

The paper contains a condensed review of detection methods of local defects of images. The methods described rely on investigation of image irregularities within some circular subarea assumed. An improved version of a method of comparing images with pattern is displayed.

Z. M. WÓJCIK

ALGORITHMES DE DÉTECTION DES DÉFAUTS LOCAUX DES IMAGES

R é s u m é

L'étude contient la description des méthodes de détection des défauts locaux des images. Les méthodes décrites dans cet article consistent en examen des irrégularités des images dans un certain sous-espace circulaire. On a présenté aussi la méthode perfectionnée de comparaison des images avec l'éta lon.

Z. M. WÓJCIK

ALGORITHMUS DER DETEKTION LOKALER BILDMÄNGEL

Zusammenfassung

Der Artikel enthält einen konzentrierten Überblick über die Detektionsmethoden lokaler Bildmängel. Die beschriebene Methode beruht auf der Untersuchung von Bildunregelmäßigkeiten innerhalb eines angenommenen kugelförmigen Unterbereichs des Operators. Es wurden auch verbesserte Varianten der Methode für den Vergleich der Bilder mit dem Modell dargestellt.

З. М. ВУЙЦИК

АЛГОРИТМЫ ДЕТЕКТИРОВАНИЯ ЛОКАЛЬНЫХ ДЕФЕКТОВ ИЗОБРАЖЕНИЙ

Резюме

Приведен конденсированный обзор методов детектирования локальных дефектов изображений. Описанные методы заключаются в исследовании дефекта изображения внутри определенной кругообразной подобласти оператора. Представлена также модифицированная версия метода сравнения изображений с образцом.

Algorytm kodowania i rozpoznawania kształtu obiektu

ZBIGNIEW M. WÓJCIK

Przemysłowy Instytut Elektroniki, Warszawa

Otrzymano 1985.09.20

Autoryzowano do druku 1987.06.10

W pracy przedstawiono pełny algorytm rozpoznawania kształtu obiektu linearnego lub konturu. Jego efektywność potwierdzają reprezentowane wyniki praktyczne. Algorytm bada tylko te obszary obrazu, w których istnieje obiekt, co zaoszczędza czas komputera. W pracy zaproponowano również kodowanie lokalnych cech obiektów przez liczbę Eulera.

1. WPROWADZENIE

Kodowanie kształtów obiektów jest najważniejszym etapem redukcji danych obrazowych. W procesie tym stopnie szarości dowolnego odseparowanego obiektu (np. w procesie segmentacji [2]) są zastępowane przez ogólne nazwy i atrybuty pewnych charakterystycznych fragmentów obiektów. Nazwy te i ich atrybuty są porządkowane w bazie danych poprzez kształty obiektów. Kodowanie jest tym procesem, który umożliwia przekształcanie stopni szarości komórek reprezentujących odseparowane obiekty w pewne uporządkowania nazw i atrybutów odpowiadające występowaniu cech i relacji w badanym obrazie. Kodowanie może być dokonywane poprzez śledzenie obiektów za pomocą kolistego operatora, a uporządkowania nazw relacji i atrybutów reprezentujących obiekty przybiera postać grafów.

Transformacja obrazu w grafy dokonywana jest przez śledzenie brzegu każdego wydzielonego obiektu lub obiektu o kształcie linii za pomocą kolistego operatora [2—7]. Dla każdej pozycji operatora na linii (lub na konturze) generowany jest węzeł grafu, który jest oznaczany nazwą cechy obrazu (tj. jednym ze słów: „*odcinek linii*”, „*zakończenie linii*”, „*rozwidlenie*”, „*skrzyżowanie*”, „*przerwa linii*”). Nazwa ta jest wytwarzana na podstawie liczb Eulera^{*)} obliczonych w polu operatora: liczby fragmentów linii przykrytych

^{*)} Liczba Eulera służy do charakteryzowania powierzchni dwustronnych. Jest to największa liczba linii (krzywych) zamkniętych, które można poprowadzić na badanej powierzchni, nie rozcinając jej na oddzielne części. Pole operatora kolistego jest traktowane tu jako powierzchnia dwustronna, tj. kula: brzeg operatora reprezentuje tu więc niewidoczną część kuli, a obiekt reprezentuje otwór w kuli. Jeżeli na powierzchni operatora nie ma żadnego obiektu, liczba Eulera równa się zero, ponieważ każda linia przeprowadzona przez operator (tj. sieczna) przetnie go na dwie części. Jeżeli na powierzchni operatora znajduje się jeden obiekt, jedna linia poprowadzona przez operator od brzegu operatora do tego obiektu nie podzieli go na oddzielne części: liczba Eulera wynosi zatem jeden. W przypadku dwóch obiektów można

operatorem oraz liczby części, na które jest podzielone tło wewnątrz operatora przez linie lub na podstawie liczby części, na jakie jest podzielony brzeg operatora przez analizowaną linię. Operator jest przesuwany wzdłuż linii automatycznie, a kolejne węzły odpowiadające tej samej linii są łączone łukiem indeksowanym symetryczną relacją dwuargumentową „przylega do”, lub „sąsiaduje z”. Uporządkowanie węzłów i łuków w grafie reprezentującym obraz odpowiada więc w wyniku śledzenia uporządkowaniu cech i relacji na tych cechach, wykrytych w tym obrazie.

Obiekt przekształcony w graf (w wyniku rozpoznania) otrzymuje nazwę wzorca, jeżeli uporządkowanie nazw cech i relacji w grafie tego obiektu jest identyczne jak w grafie wzorcowym, tzn. jeżeli obydwa te grafy posiadają identyczną strukturę jakościową. W pewnych przypadkach wymagana jest także zbliżona (podobna) struktura ilościowa tych dwóch grafów np. zbliżone zmiany nachylenia odcinków linii, jednostajności zmian nachyleń odcinków, długości linii. Grafy wzorcowe mogą być otrzymywane przez rozpoznanie obiektów wzorcowych.

Założenie 1. Proste, elementarne cechy obrazów są wykrywane przez pewne funkcje charakterystyczne tych cech obliczone w polu operatora i przyjmujące założone wartości, jeśli rozkłady zaczerpień obrazu na powierzchni operatora są zgodne z przyjętymi opisami badanych cech oraz inne wartości w przeciwnych przypadkach.

Lemat 1. Wynik rozpoznawania jest zgodny z podstawowym warunkiem reprezentowania scen, jeżeli wartość przyjętej funkcji charakterystycznej w polu operatora jest niezależna względem obrotów obrazu wokół centrum tego operatora oraz jeśli wszystkie komórki obrazu są analizowane za jej pomocą.

Dowód. Niezależność wyniku pewnej przyjętej funkcji charakterystycznej operatora względem obrotów obrazu wokół centrum operatora i analiza za pomocą takiego symetrycznego operatora całego obrazu jest warunkiem koniecznym reprezentowania obrazów [3], ponieważ przemieszczenia operatora są równoważne przemieszczeniom obrazu (pod warunkiem, że rozpatrywany obraz nie wychodzi poza fotoczulą mozaikę systemu w trakcie przemieszczeń).

Wynikiem funkcji charakterystycznej jest konkretna wartość liczbową: należy jeszcze dowiedzieć, że odpowiada ona nazwie poszukiwanej cechy, jeżeli operator natrafia na badaną cechę oraz że wartość funkcji charakterystycznej jest poza założonym przedziałem liczbowym jeżeli operator nie leży na poszukiwanej cesze, co byłoby warunkiem wystarczającym reprezentowania badanej cechy. Rozpatrywany wynik spełnia warunek konieczny reprezentowania obrazów, nie spełnia jednak warunku wystarczającego reprezentowania obrazów: jest zatem tylko zgodny z podstawowym warunkiem reprezentowania scen, c.d.t.

Założenie 2. Zgodność wartości funkcji charakterystycznych z nazwami cech jest stwierdzana poprzez porównywanie otrzymywanych wyników działania tych funkcji na obrazy z założonymi definicjami cech.

Zamiast „funkcja charakterystyczna” stosuje się wymiennie pojęcie „operator”, Słowo „operator” ma więc podwójne znaczenie.

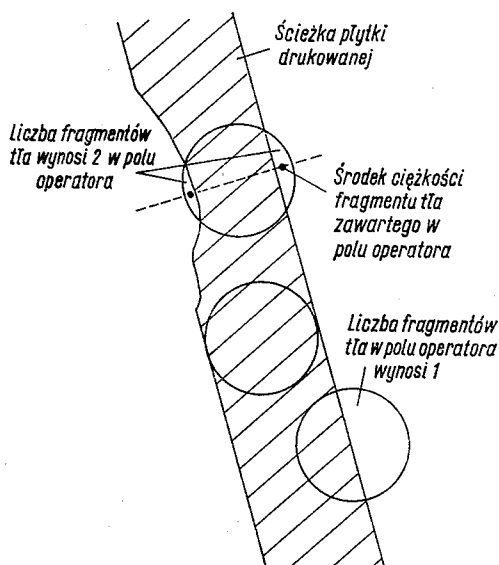
poprowadzić dwie sieczne łączące brzeg operatora z każdym z tych obiektów i nie dzielące operatora na oddzielne części: liczba Eulera równa się więc dwa, itd.

Twierdzenie 1. Wynik rozpoznawania spełnia podstawowy warunek reprezentowania obrazów, jeżeli wynik ten reprezentuje badany obiekt lub zdarzenie i jeżeli wynik ten jest zgodny z podstawowym warunkiem reprezentowania scen.

Dowód. Zgodność wyniku rozpoznawania z podstawowym warunkiem reprezentowania scen oznacza spełnianie warunku koniecznego reprezentowania scen przez ten wynik (lemat 1) dopiero po uzgodnieniu uzyskiwanych wartości założonych funkcji charakterystycznych z nazwami poszukiwanych cech (założenie 2). Otrzymuje się bowiem wtedy konkretne nazwy wydzielanych cech i właściwości obrazu, co automatycznie prowadzi do dowodu warunku wystarczającego reprezentowania scen, cbdu.

Podstawowy warunek reprezentowania obrazów, stosowany na wszystkich etapach przetwarzania i rozpoznawania obrazów nakłada pewne ograniczenia na algorytmy rozpoznawania, zapewnia jednak w zamian poprawność wyników rozpoznawania. Podstawowy warunek reprezentowania obrazów dostarcza także metodologicznej spójności w podejściu do problemów przetwarzania i rozpoznawania obrazów.

Jeden z algorytmów sprawdzania odległości podany przez B. Kruse [1] wykorzystuje ideę wyznaczania liczby Eulera (liczby obiektów) wewnątrz kolistego pola operatora. Algorytm ten jest przeznaczony dla wizualnego badania ścieżek płytek drukowanych. Jeżeli liczba części tła pola operatora jest większa od 1, algorytm wydaje decyzję, że ścieżka jest zbyt wąska (rys. 1). Średnica pola operatora nie może być większa od szerokości spraw-



Rys. 1. Trzy pozycje kolistego operatora o średnicy równej szerokości ścieżki na płycie drukowanej

dzanej ścieżki. Jeśli średnica pola operatorowego jest zwiększona powyżej szerokości linii, staje się możliwe wykrycie odcinka linii w operatorze, jeżeli liczba części tła jest równa 2 wewnątrz operatora. Spostrzeżenie to [4, 5], uzupełnione o przypadki analizy wszystkich typów linii, rozwinięte zostało w uniwersalną metodę kodowania kształtu obiektów [2—6], powiązaną ponadto ściśle z grafową reprezentacją zdań opisujących rozpoznawane kształty.

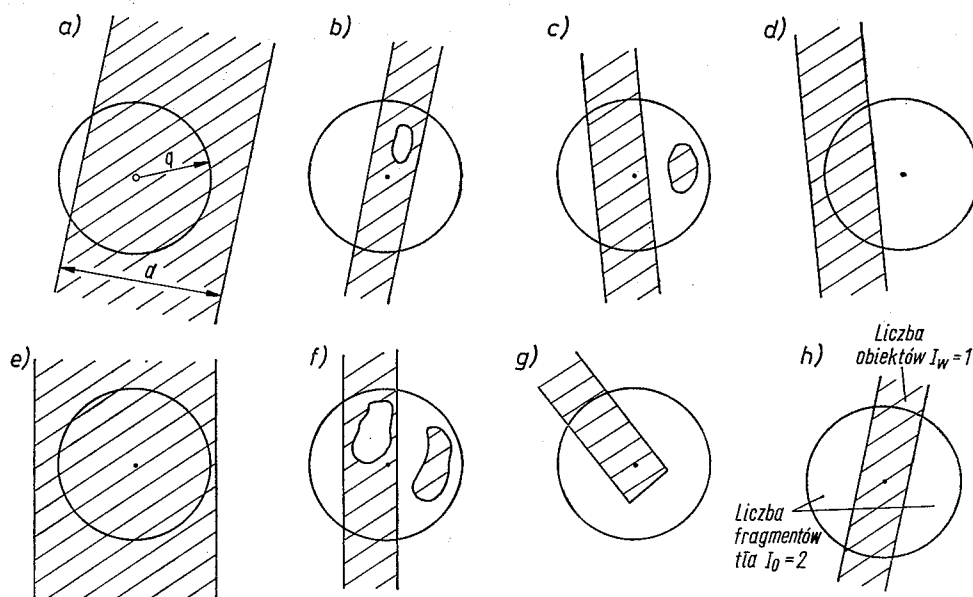
Kierunek odcinka linii może być zmierzony przez obliczenie środków ciężkości tych dwóch części tła oddzielnie i poprzez wyznaczenie linii prostopadłej do linii przechodzącej przez te środki (rys. 1). Linia może być śledzona poprzez przemieszczanie pola operatora wzdłuż wyznaczonego kierunku.

2. ZWIĄZKI POMIĘDZY LICZBAMI EULERA A NAZWAMI CECH

Założenie 3. Każda kodowana linia jest spójna, tj. pole operatora nie zawiera ani podobszaru tła otoczonego polem linii (rys. 2b) ani żadnego regionu o tej samej jasności co linia (rys. 2 c, f).

Wniosek 1. Należy przeprowadzić właściwą wstępną obróbkę dla uniknięcia niespójności w obrazach linii.

Założenie 4. Promień q operatora kolistego jest większy od średnicy d badanej linii (rys. 2a).



Rys. 2. Możliwe przypadki linii wewnątrz kolistego operatora: a) linia, której średnica „ d ” jest większa od promienia q operatora; b) linia niespójna zawierająca dziurę; c) zakłócenie tła w polu operatora; d) linia nie przecinająca centrum operatora; e) operator zawarty całkowicie w grubej linii; f) operator zawierający linię z dziurą i zakłócenie tła; g) zakończenie linii; h) prawidłowy odcinek linii wewnątrz operatora

Założenie 5. Badana linia jest dłuższa od średnicy pola operatora.

Wniosek 2. Promień koła operatora jest dopasowany do średnicy każdej badanej linii.

Założenie 6. Linia przecina centrum operatora.

Założenie 7. Linia nie ma przerwań (rys. 4).

Twierdzenie 2. Pojedyncza linia przechodzi przez pole operatora jeśli liczba I_w obiektów o poziomie szarości odpowiadającym jasności linii wynosi 1, a liczba I_0 obiektów tła wynosi 2 w polu operatora (rys 2h).

Dowód. Gdyby liczba I_w obiektów odpowiadających linii była równa 0, nie byłoby obiektów w polu operatora. Gdyby liczba I_w była równa 2, albo pole operatora zawierałoby dwa obiekty (tj. odcinki linii — patrz rys. 3c oraz 4), albo założenie 3 nie byłoby spełnione (rys. 2c, f). Przejście linii przez pole operatora może być rozpoznane, jeżeli $I_w = 1$ (rys. 2h, e, a, g). Jeśli dla $I_w = 1$ liczba I_0 obiektów tła jest równa 0, otrzymywana jest sprzeczność z założeniem 4, ponieważ operator jest umieszczony wewnątrz pola linii, która jest ponad dwa razy grubsza od promienia pola operatora (rys. 2e). Jeśli dla $I_w = 1$ liczba I_0 obiektów tła jest równa 1, tylko część linii przechodzi przez operator (rys. 2a) lub operator zawiera zakończenie linii (rys. 2g). Jeśli założenia 3—7 są spełnione, pojedyncza linia dzieli tło pola operatora na dwie części (rys. 2h) i dlatego liczba części tła jest równa 2. Jeśli dla $I_w = 1$ liczba $I_0 = 3$, tło pola operatora jest podzielone na trzy części (rys. 3a) a nie na dwie części: sytuacja taka oznacza rozwidlenie linii. Liczba I_w musi być wtedy równa 1 dla pojedynczej linii, a kiedy $I_w = 1$, $I_0 = 2$, odcinek pojedynczej linii jest stwierdzany w polu operatora, cbdu.

Twierdzenie 3. Zakończenie pojedynczej linii jest rozpoznawane w polu operatora jeśli $I_w = 1$ i $I_0 = 1$.

Dowód. Ponieważ zakończenie linii leżące w polu operatora nie może podzielić jego powierzchni, $I_0 = 1$ (rys. 2g). Jeśli $I_0 = 2$, odcinek linii jest rozpoznawany (twierdzenie 1).

Liczby $I_w = 1$ i $I_0 = 1$ nie odróżniają zakończenia linii od przypadku, kiedy pole operatora jest umieszczone na krawędzi linii (porównaj rys. 2g, a, d). Jednakże, przypadki pokazane na rys. 2a, d) są niedopuszczalne: linia z rys. 2a jest grubsza niż promień operatora (sprzeczność z założeniem 4), a linia z rys. 2d nie przecina centrum operatora (sprzeczność z założeniem 6). Aby pojedyncze zakończenie było rozpoznane, I_0 musi być jednocześnie równe 1, cbdu.

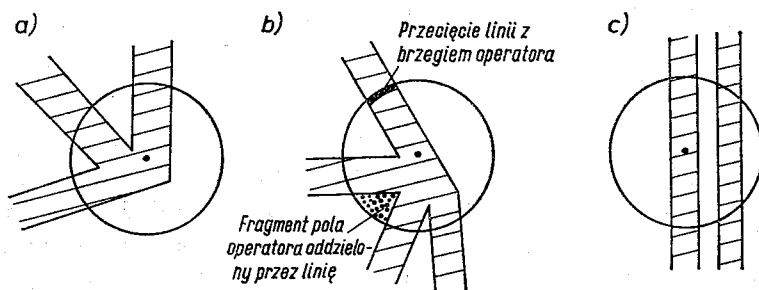
Twierdzenie 4. Pojedyncze rozwidlenie jest rozpoznawane w polu operatora, jeśli $I_w = 1$ i $I_0 = 3$.

Dowód. Jeśli $I_w = 1$ i $I_0 = 2$, odcinek linii jest rozpoznawany (twierdzenie 2). Jeżeli $I_w = 1$ i $I_0 = 4$, stwierdzone jest skrzyżowanie (rys. 3b). Pojedyncze rozwidlenie ($I_w = 1$) dzieli więc pole operatora na trzy części i dlatego $I_0 = 3$ (rys. 3a), cbdu.

Twierdzenie 5. Pojedyncze skrzyżowanie jest rozpoznawane w polu operatora, jeśli $I_w = 1$ i $I_0 = 4$.

Dowód. Jeżeli $I_w = 1$ i $I_0 = 3$, rozwidlenie jest rozpoznawane (twierdzenie 3). Skoro pojedyncze skrzyżowanie ($I_w = 1$) dzieli pole operatora na cztery części (rys. 3b), liczba I_0 wynosi 4, cbdu.

Lemat 2. Jeżeli $I_w = 2$ i $I_0 = 3$ wtedy dwie nieprzecinające się linie przechodzą przez pole operatora (rys. 3c).



Rys. 3. Przypadki linii w polu operatora dla $I_0 > 2$: a) rozwidlenie; b) skrzyżowanie; c) dwie linie nieprzecinające się

Dowód. Pojedyncza linia ($I_w = 1$) dzieli pole operatora na dwie części (Twierdzenie 2), tj. $I_0 = 2$. Druga linia ($I_w = 2$) dzieli jedną z dwóch części tła operatora na dwie części, w wyniku czego $I_0 = 3$, cbdu.

Twierdzenie 6. Jeżeli $I_w = n$ oraz $I_0 = n + 1$, wtedy n nieprzecinających się linii przechodzi przez pole operatora.

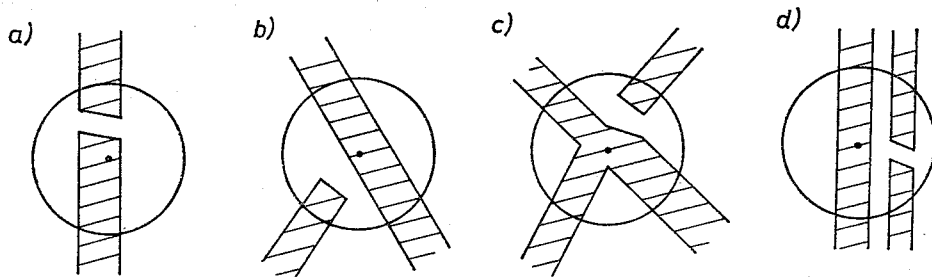
Dowód (przez indukcję). Dla $n = I_w = 1$ oraz dla $I_0 = n + 1 = 1 + 1$, twierdzenie 6 jest równoważne twierdzeniu 2. Dla $n = I_w = 2$ i $I_0 = n + 1 = 2 + 1$, twierdzenie 6 jest równoważne lematowi 2. Będzie udowodnione, że jeśli twierdzenie 6 jest prawdziwe dla n linii, $n \geq 2$, to jest tak samo prawdziwe dla $n + 1$ linii. Zakładamy więc, że n nieprzecinających się linii dzieli tło operatora na $n + 1$ części. Nowa $n + 1$ linia dzieli $n + 1$ część tła operatora na dwie nowe części w wyniku czego $I_0 = n + 2$, cbdu.

Twierdzenie 7 sformułowane poniżej dotyczy przypadków kiedy założenie 7 nie jest spełnione.

Twierdzenie 7. Pojedyncza przerwa linii leżąca w polu operatora zmniejsza liczbę I_0 części tła o 1 oraz zwiększa liczbę I_w obiektów o 1.

Dowód. Wiążąc twierdzenie 7 z twierdzeniami 2—6, otrzymuje się następujące lematy 3—6:

Lemat 3. Jeżeli pole operatora pokrywa pojedynczą przerwę w jednej linii, wtedy $I_w = 2$ i $I_0 = 1$ (rys. 4a).



Rys. 4. Przypadki przerw linii; a) w pojedynczej linii; b) w rozwidleniu; c) w skrzyżowaniu; d) w jednej z dwóch linii nieprzecinających się

Lemat 4. Jeżeli pole operatora jest umieszczone na pojedynczym rozwidleniu z pojedynczą przerwą, wtedy $I_w = 2$ i $I_0 = 2$ (rys. 4b).

Lemat 5. Jeśli pole operatora jest umieszczone na pojedynczym skrzyżowaniu z jedną przerwą, wtedy $I_w = 2$ i $I_0 = 3$ (rys. 4c).

Lemat 6. Jeżeli pole operatora jest umieszczone na dwóch nieprzecinających się liniach z pojedynczą przerwą, wtedy $I_w = 3$ i $I_0 = 2$ (rys. 4).

Dowód Twierdzenia 7 oraz Lematów 3 do 6:

Pojedyncza przerwa jednej linii w polu operatora dzieli linię na dwie części w wyniku czego liczba I_w ulega zwiększeniu o 1. To samo powoduje połączenie dwóch sąsiednich części tła operatora dając w wyniku zmniejszenie liczby I_0 o 1, cbdu.

Twierdzenie 8. Jeżeli $I_w \geq 2$, wtedy przerwy linii powodują niejednoznaczności w interpretacji linii znajdującej się w polu operatora.

Dowód. Następujący lemat 7 zostanie wyróżniony w twierdzeniu 8.

Lemat 7. Zarówno w przypadku nieprzecinających się linii (rys. 3c) i w przypadku pojedynczej przerwy w jednym skrzyżowaniu (rys. 4c) położonych w polu operatora, liczby $I_w = 2$ i $I_0 = 3$. Dowód wynika bezpośrednio z lematów 2 i 5. Lemat 7 dowodzi co najmniej jednej niejednoznaczności w interpretacji linii wewnątrz pola operatora kiedy $I_w = 2$. Dla $I_w = 1$ nie może być żadnej niejednoznaczności ponieważ jeżeli pojawia się przerwanie linii wtedy jednocześnie $I_w \geq 2$ porównaj twierdzenia 2 do 5 z twierdzeniem 7, cbdu.

Aby uniknąć powyższej niejednoznaczności, pole operatora o mniejszej średnicy powinno być użyte tak że powinna się w nim mieścić wyłącznie jedna linia.

Kontury wydzielone jako brzegi regionów o zbliżonym poziomie szarości [2, 6] są wolne od przerw. Brzegi występują jako linie o jednolitej grubości wynoszącej około jednej komórki. W ten sposób niejednoznaczności wyrażone twierdzeniem 8 nie zdarzają się, jeżeli kontury są wydzielane jako brzegi obiektów.

Lemat 8. Liczby I_0 i I_w spełniają podstawowy warunek reprezentowania obrazów, jeśli linia poddana badaniu znajduje się wewnątrz pola operatora.

Dowód. Pole operatora jest symetryczne w odniesieniu do swojego centrum. Dowolne obroty lub przesunięcia badanej linii (włączając skrzyżowania, rozwidlenia, zakończenia, itd.) w polu operatora mogą powodować pojawienie się i znikanie pewnych lokalnych minimów i maksimów, jednakże liczba I_w i liczba I_0 pozostają niezmiennicze [7, 5], cbdu.

Twierdzenie 9. Jeśli $I_w = 1$, wtedy liczba I_0 jest równa liczbie przecięć linii z brzegiem operatora.

Dowód. Każdy odcinek obwodu operatora oddzielony liniami przechodzącymi przez ten obwód odpowiada jednej części pola operatora oddzielonej przez te linie (rys. 3b). W ten sposób liczba odcinków obwodu odpowiada liczbie I_0 . Zatem liczba I_0 równa się liczbie przecięć linii z obwodem operatora, ponieważ obwód operatora tworzy linię zamkniętą, cbdu.

Wniosek 3. Jeśli w polu operatora $I_w = 1$, wtedy pojedyncza linia może być reprezentowana jednoznacznie zarówno liczbą I_0 części pola operatora oddzielonych przez linię oraz przez liczbę I_p przecięć z obwodem operatora ($I_p = I_0$):

e
 LICZBY NI PRZECIEC BRZEGU Z OBIEKTEM
 2 2 1
 POZYCJE PRZECIEC OPERATORA Z OBIEKTEM
 10.00 11.50 11.00
 6.33 11.75 13.44
 NACHYLENIA SEGMENTOW (W RAD)
 1.21 1.43 -1.28
 1 OZNACZA ODCINEK LINII; 2 -ZAKONCZENIE
 1 1 1
 DLUGOSCI ODCINKOW LINII (LICZBY KOMOREK)
 3.38 2.27 1.77

```

-----
- - - - -
- - - - - H H - - -
- - - - - H H H - - -
- - - - - H H H - - -
- - - - - H - - -
- - - - - * * * - - -
- - - - - * - - - * - - -
- - - - - * - - - H - * - - -
- - - - - Q H H H - * - - -
- - - - - H Q H H - * - - -
- - - - - H H H H H Q * - - -
- - - - - H H H H H H - - -
- - - - - H H H H H H - - -
- - - - - H H H H - - -
- - - - -
-----

```

f
 LICZBY NI PRZECIEC BRZEGU Z OBIEKTEM
 2 2 1 1
 POZYCJE PRZECIEC OPERATORA Z OBIEKTEM
 10.00 11.50 11.00 8.78
 6.33 11.75 13.44 15.44
 NACHYLENIA SEGMENTOW (W RAD)
 1.21 1.43 -1.28 -0.73
 1 OZNACZA ODCINEK LINII; 2 -ZAKONCZENIE
 1 1 1 1
 DLUGOSCI ODCINKOW LINII (LICZBY KOMOREK)
 3.38 2.27 1.77 2.99

```

-----
- - - - -
- - - - - H H - - -
- - - - - H H H - - -
- - - - - H H H - - -
- - - - - H - - -
- - - - - * * * - - -
- - - - - * - - - * - - -
- - - - - * - - - H - * - - -
- - - - - H Q H H H - * - - -
- - - - - H H H H H H - - -
- - - - - H H H H H Q * - - -
- - - - - H H H H - - -
- - - - -
-----

```

g
 LICZBY NI PRZECIEC BRZEGU Z OBIEKTEM
 2 2 1 1 1
 POZYCJE PRZECIEC OPERATORA Z OBIEKTEM
 10.00 11.50 11.00 8.78 6.30
 6.33 11.75 13.44 15.44 16.70
 NACHYLENIA SEGMENTOW (W RAD)
 1.21 1.43 -1.28 -0.73 -0.47
 1 OZNACZA ODCINEK LINII; 2 -ZAKONCZENIE
 1 1 1 1 1
 DLUGOSCI ODCINKOW LINII (LICZBY KOMOREK)
 3.38 2.27 1.77 2.99 2.78

```

-----
- - - - -
- - - - - H H - - -
- - - - - H H H - - -
- - - - - H H H - - -
- - - - - H - - -
- - - - - * * * - - -
- - - - - * - - - * - - -
- - - - - * - - - Q H H - - -
- - - - - H Q H H H - * - - -
- - - - - H H Q H - - -
- - - - - * * * - - -
- - - - -
-----

```

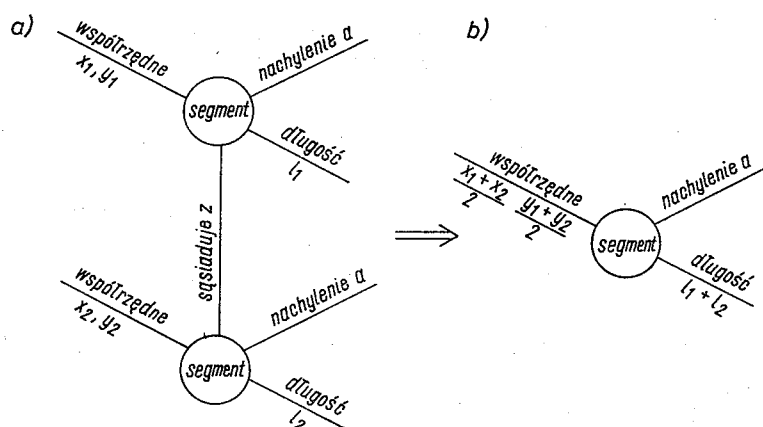
h
 LICZBY NI PRZECIEC BRZEGU Z OBIEKTEM
 2 2 1 1 1 0
 POZYCJE PRZECIEC OPERATORA Z OBIEKTEM
 10.00 11.50 11.00 8.78 6.30 4.33
 6.33 11.75 13.44 15.44 16.70 17.67
 NACHYLENIA SEGMENTOW (W RAD)
 1.21 1.43 -1.28 -0.73 -0.47 0.00
 1 OZNACZA ODCINEK LINII; 2 -ZAKONCZENIE
 1 1 1 1 2
 DLUGOSCI ODCINKOW LINII (LICZBY KOMOREK)
 3.38 2.27 1.77 2.99 2.78 2.19

```

-----
- - - - -
- - - - - H H - - -
- - - - - H H H - - -
- - - - - H H H - - -
- - - - - H - - -
- - - - - * * * - - -
- - - - - * - - - * - - -
- - - - - * - - - * - - -
- - - - - * - - - H - * - - -
- - - - - H Q H H H - * - - -
- - - - - H H H H H - - -
- - - - - H H H H H Q * - - -
- - - - - H H H H - - -
- - - - -
-----

```


wzorcowych. Rozpoznany obiekt (tzn. zakodowany w graf) otrzymuje nazwę tego grafu wzorcowego, który jest jakościowo identyczny z grafem rozpoznanego obiektu, tzn. jeżeli uporządkowanie nazw cech i relacji w grafie otrzymanym przez rozpoznanie obiektu jest identyczne jak w grafie wzorcowym. Warto zauważyć, że dowolny graf wzorcowy można otrzymać przez rozpoznanie obiektu wzorcowego.



Rys. 6. Redukcja do jednego węzła (b) dwóch kolejnych węzłów (a) o zbliżonych do siebie nachyleniach

Graf, przedstawiony na rys. 6b może być reprezentowany w języku LISP za pomocą następującej listy:

(odcinek, (kierunek, α), (długość, d),
(współrzędne, x , y), kontur, prosty, n).

(1)

Na rys. 5 wydrukowano pozycje przecięć brzegu operatora kołowego z kodowanym obiektem, nachylenia i długości segmentów linii zawartych w operatorze w trakcie śledzenia i kodowania. Parametry te zostały obliczone na podstawie środków ciężkości wyznaczonych przecięć. Dla ułatwienia obliczeń cała powierzchnia operatora jest zerowana po wyznaczeniu I_p oraz parametrów. Kilka kroków operatora na początku śledzenia pozwala umieścić operator we właściwej pozycji na kodowanym obiekcie. Ciągła zmiana nachylenia linii w trakcie kodowania w jednym tylko kierunku (patrz dół rys. 5j) i całkowita suma zmian nachylenia są podstawą do zidentyfikowania półprzestrzeni.

Jeśli rozpoznawany obiekt nie posiada kształtu linearnego, można śledzić i kodować grafowo jego kontur. Operator użyty do kodowania grafowego powinien posiadać średnicę większą od 3 komórek, ale mniejszą od dowolnego rozmiaru kodowanego obiektu. Im większa jest średnica użytego operatora, tym proces kodowania grafowego jest szybszy, ale precyzja tego kodowania maleje. W trakcie kodowania grafowego brzegu odseparowanego obiektu liczba I_p jest zawsze mniejsza niż trzy. W przypadku jednak analizy obrazu, będącego sumą topologiczną dwóch obrazów złożonych z wydzielonych brzegów, liczba I_p może być większa od dwóch. W ten sposób można przeprowadzić analizę połączeń konturów obiektów.

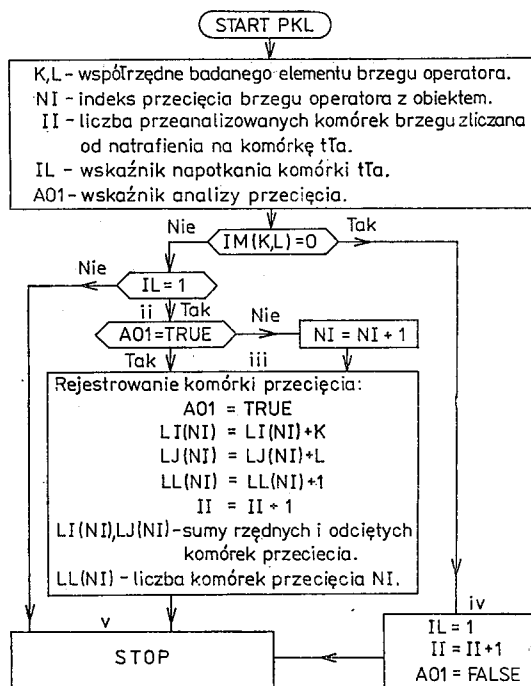
4. SZCZEGÓŁOWE ALGORYTMY KODOWANIA KSZTAŁTU

Założenie 8. Pełna analiza brzegu operatora kodującego obiektu wymaga:

- ustalenia przynależności każdej komórki brzegowej do obiektu albo do jego tła;
- zaindeksowania każdego oddzielnego przecięcia obiektu z brzegiem operatora;
- prawidłowego przypisania każdego elementu brzegowego reprezentującego obiekt do właściwego przecięcia;
- zarejestrowania rzędnej i odciętej każdego takiego przypisanego elementu w pozycji bazy danych odpowiadającej oddzielnemu przecięciu;
- zliczania komórek każdego oddzielnego przecięcia.

Lemat 9. Analiza i rejestrowanie przynależności dowolnego elementu brzegowego operatora do przecięć obiektu z brzegiem operatora jest wykonywana przez (rys. 7):

- odczyt wartości cyfrowej tego elementu. Jeżeli element ten nie reprezentuje przecięcia, skok do (iv), w przeciwnym przypadku badany jest wskaźnik napotkania tła. Jeśli wskaźnik ten nie jest zapalony, trwa poszukiwanie pierwszej komórki tła przez skok do (v);
- jeśli wskaźnik analizy przecięcia obiektu z brzegiem operatora nie jest zapalony, zwiększenie liczby przecięć o 1;
- rejestracja elementu przecięcia;
— zapalenie wskaźnika przecięcia,



Rys. 7. Analiza i rejestrowanie przynależności pojedynczego elementu brzegowego operatora do przecięć brzegu operatora z obiektem (procedura PKL)

- dodanie rzędnej i odciętej elementu do sumy rzędnych i odciętych aktualnego przecięcia,
 - dodanie jedynki do liczby rejestrowanych elementów brzegowych operatora,
 - skok do (v);
- iv. zapalenie wskaźnika napotkania tła i dodanie jedynki do liczby rejestrowanych komórek operatora oraz zgaszenie wskaźnika przecięcia,
- v. koniec analizy i rejestracji bieżącego elementu brzegowego operatora.

Dowód. Rozpatrzmy wszystkie możliwe przypadki elementu brzegowego operatora:

- a) element reprezentuje obiekt ale wskaźnik napotkania tła nie jest zapalony: element reprezentuje przecięcie, ale nie może zostać zarejestrowany jako element przecięcia, ponieważ dla uniknięcia zbędnych komplikacji rejestracja rozpoczyna się od natrafienia na początek przecięcia;
- b) element nie reprezentuje obiektu: wskaźnik napotkania tła jest zapalony i każdy element reprezentujący obiekt będzie potem rejestrowany aż do zakończenia analizy całego brzegu operatora (patrz następny lemat). Wskaźnik przecięcia jest gaszony, aby umożliwić zwiększenie liczby przecięć z chwilą natrafienia na początek przecięcia;
- c) element reprezentuje obiekt i wskaźnik tła jest zapalony: zgaszony wskaźnik analizy przecięcia jest równoznaczny z natrafieniem na początek przecięcia, a więc z koniecznością zwiększenia indeksu (bieżącej liczby) przecięć o 1.

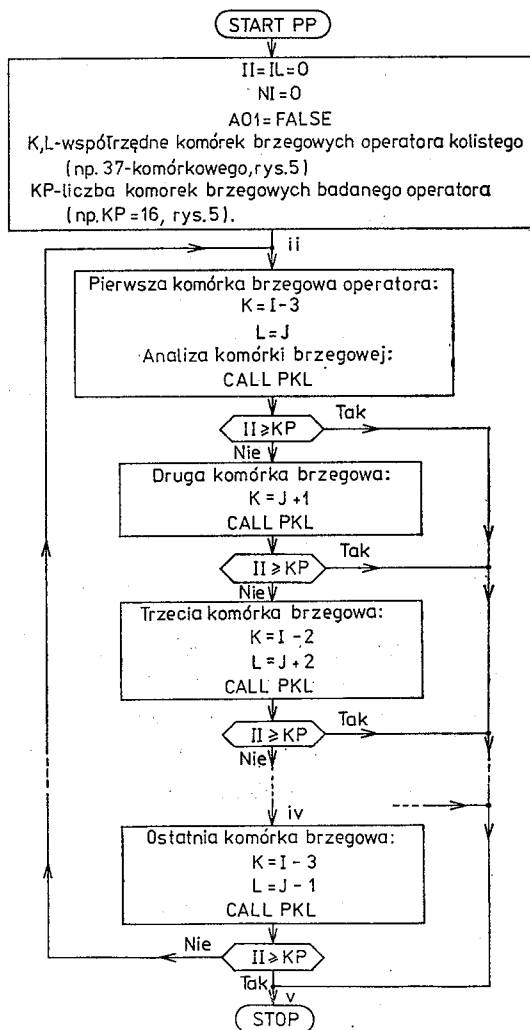
Przy zapalonym wskaźniku analizy przecięcia następuje normalna rejestracja elementu przecięcia pod bieżącym indeksem parametrów tego przecięcia.

Algorytm klasyfikuje więc prawidłowo (zgodnie z założeniem 8) wszystkie możliwe przypadki komórki brzegowej operatora i rejestruje poprawnie elementy reprezentujące przecięcia, cbdu.

Lemat 10. Analiza całego brzegu operatora i rejestracja przecięć kodowanego obiektu z brzegiem operatora jest przeprowadzona poprzez (rys. 8):

- i. zgaszenie wskaźnika napotkania indeksu przecięcia i liczby rejestrowanych komórek brzegowych operatora na zero;
- ii. ustawienie współrzędnych pierwszego elementu brzegu operatora. Analiza i rejestrowanie przynależności tego elementu do przecięcia. Jeśli liczba zarejestrowanych komórek jest co najmniej równa liczbie wszystkich elementów brzegowych operatora, skok do (v);
- iii. ustawienie współrzędnych następnego elementu brzegu operatora. (Analiza kolejnych elementów operatora jak w (ii));
- iv. ustawienie współrzędnych ostatniego elementu brzegu operatora. Analiza i rejestrowanie przynależności tego elementu do przecięcia. Jeśli liczba zarejestrowanych komórek jest mniejsza od liczby wszystkich elementów brzegowych operatora, skok do (ii);
- v. koniec analizy całego brzegu operatora.

Dowód. Z elementów brzegowych utworzona jest w tym algorytmie pierścieniowa baza danych przez porównywanie liczby zarejestrowanych komórek brzegowych z liczbą wszystkich elementów brzegu operatora i przez ewentualny skok po odczycie ostatniego ele-



Rys. 8. Procedura PP analizy i rejestrowania przecięć całego brzegu operatora z kodowanym obiektem

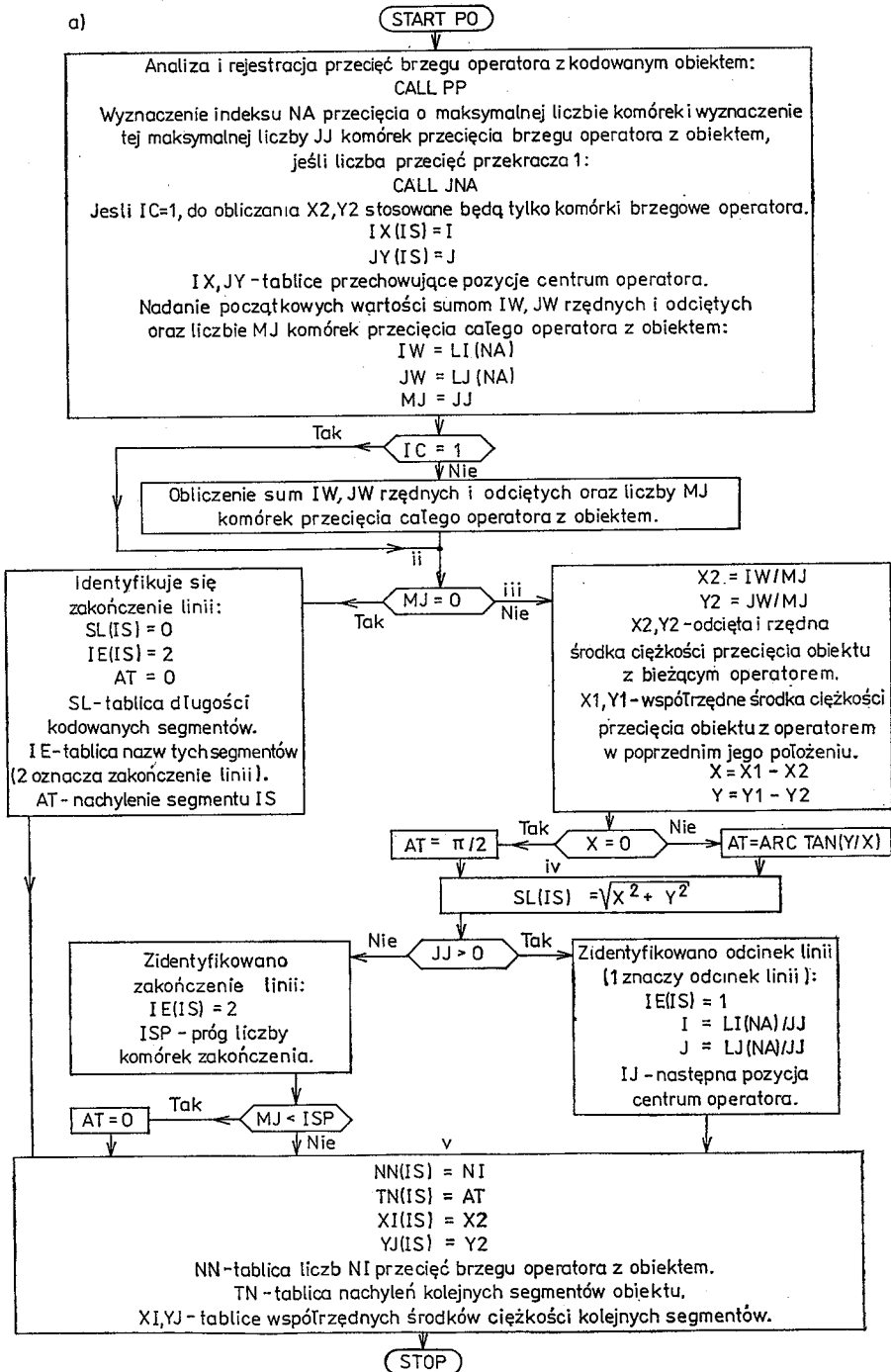
mentu brzegu do analizy i rejestracji pierwszego elementu brzegowego. Zakończenie analizy brzegu operatora odbywa się więc po rejestracji wszystkich komórek brzegowych zliczanych od momentu natrafienia na początek pierwszego przecięcia obiektu z brzegiem operatora (patrz poprzedni lemat), cbdu.

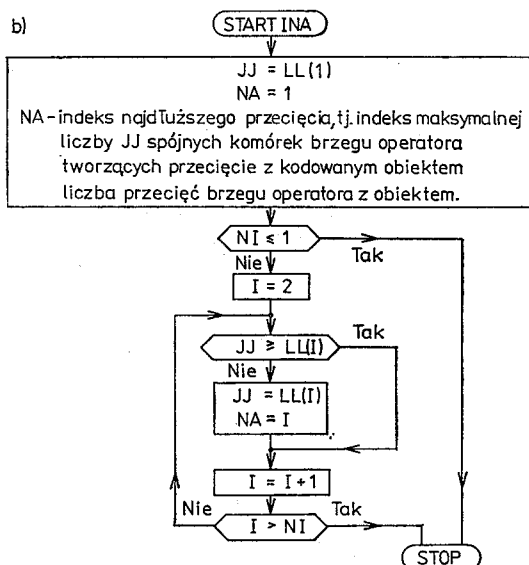
Założenie 9. Pełna analiza całej powierzchni operatora kodującego obiekty przez śledzenie wymaga:

- nadania właściwej nazwy fragmentowi obiektu znajdującemu się w polu tego operatora;
- wyznaczenie parametrów specyfikujących każdy fragment obiektu identyfikowany w polu operatora;
- wyznaczenia następnej pozycji operatora na kodowanej linii.

Lemat 11. Proces nadawania nazwy *odcinek linii* albo *zakończenie linii* oraz wyznaczania parametrów fragmentu obiektu mieszczącego się w polu operatora odbywa się w następujący sposób (rys. 9):

1. — Analiza i rejestracja przecięć brzegu operatora z kodowanym obiektem (Lemat 10);
 - Jeśli obliczona podczas tej analizy liczba przecięć przekracza jeden, wyznacza się indeks przecięcia o maksymalnej liczbie komórek i wyznacza się tę liczbę (rys. 9b);
 - Rejestracja bieżącej pozycji centrum operatora pod indeksem badanego segmentu pokrytego operatorem;
 - Nadanie sumom rzędnych i odciętych komórek przecięcia całego operatora z obiektem początkowych wartości sum rzędnych i odciętych komórek przecięcia brzegu operatora z obiektem (lematy 10 i 9);
 - Nadanie liczbie komórek przecięcia całego operatora z obiektem początkowej wartości równej liczbie komórek przecięcia brzegu operatora z obiektem (lematy 10 i 9);
 - Jeśli odbywa się początkowy proces osadzania operatora na obiekcie (rys. 5b, c), skok do (ii);
 - Obliczenie sum rzędnych i odciętych komórek przecięcia całego operatora z obiektem oraz całkowitej liczby tych komórek;
- ii. — Jeśli liczba komórek przecięcia całego operatora z obiektem nie jest równa zero, skok do (iii);
 - Identyfikuje się zakończenie linii. Pod indeks aktualnego segmentu wpisuje się do bazy danych (rys. 5j);
 - a) nazwę *zakończenie*,
 - b) długość równą zero,
 - c) nachylenie równe zero,
 - d) skok do (v);
- iii. — Obliczenie i rejestracja środka ciężkości bieżącego segmentu linii przykrytego operatora przez podzielenie sum rzędnych i odciętych komórek przecięcia całego operatora z obiektem przez całkowitą liczbę komórek tego przecięcia;
 - Obliczenie różnicy Y pomiędzy rzędną środka ciężkości poprzedniego przecięcia całego operatora z obiektem a rzędną aktualnego przecięcia. Obliczenie takiej samej różnicy X dla odciętych;
 - Jeżeli różnica X pomiędzy odciętymi środkami ciężkości poprzedniego i bieżącego przecięcia całego operatora z obiektem wynosi zero, nachylenie badanego segmentu wynosi $\pi/2$ i następuje skok do (iv);
 - Nachylenie bieżącego segmentu pokrytego przez operator wynosi $\arctan(Y/X)$;
- iv. — Długość bieżącego segmentu równa jest odległości pomiędzy środkami ciężkości poprzedniego i bieżącego przecięcia (tj. $\sqrt{X^2 + Y^2}$).
 - Rejestracja długości tego segmentu;
 - Jeżeli liczba komórek przecięcia brzegu operatora z obiektem wynosi 0, identyfikowane jest zakończenie linii:
 - a) pod indeks aktualnego segmentu wpisuje się nazwę *zakończenie*,
 - b) jeśli liczba komórek przecięcia całego operatora z obiektem jest mniejsza od progowej liczby zakończenia linii, nachylenie wykrytego zakończenia wynosi zero (lub jest równe nachyleniu poprzedniego segmentu linii),





Rys. 9. Procedura PO (a) nadawania nazwy „segment linii” albo „zakończenie linii” oraz wyznaczania parametrów fragmentu obiektu mieszczącego się w polu operatora; b) pomocnicza procedura INA wyznaczania indeksu NA przecięcia posiadającego maksymalną liczbę komórek i wyznaczania tej maksymalnej liczby JJ komórek przecięcia brzegu operatora z obiektem

— c) skok do (v)

— Identyfikowany jest odcinek linii:

a) pod indeks bieżącego segmentu wpisuje się nazwę *odcinek*

b) oblicza się następną pozycję elementu centralnego operatora przez podzielenie sum rzędnych i odciętych przecięcia brzegu operatora z obiektem przez liczbę komórek tego przecięcia;

v. Rejestracja pod indeksem bieżącego segmentu:

a) liczby przecięć brzegu operatora z obiektem,

b) nachylenie zidentyfikowanego segmentu,

c) rzędnej i odciętej środka ciężkości przecięcia całego operatora z obiektem.

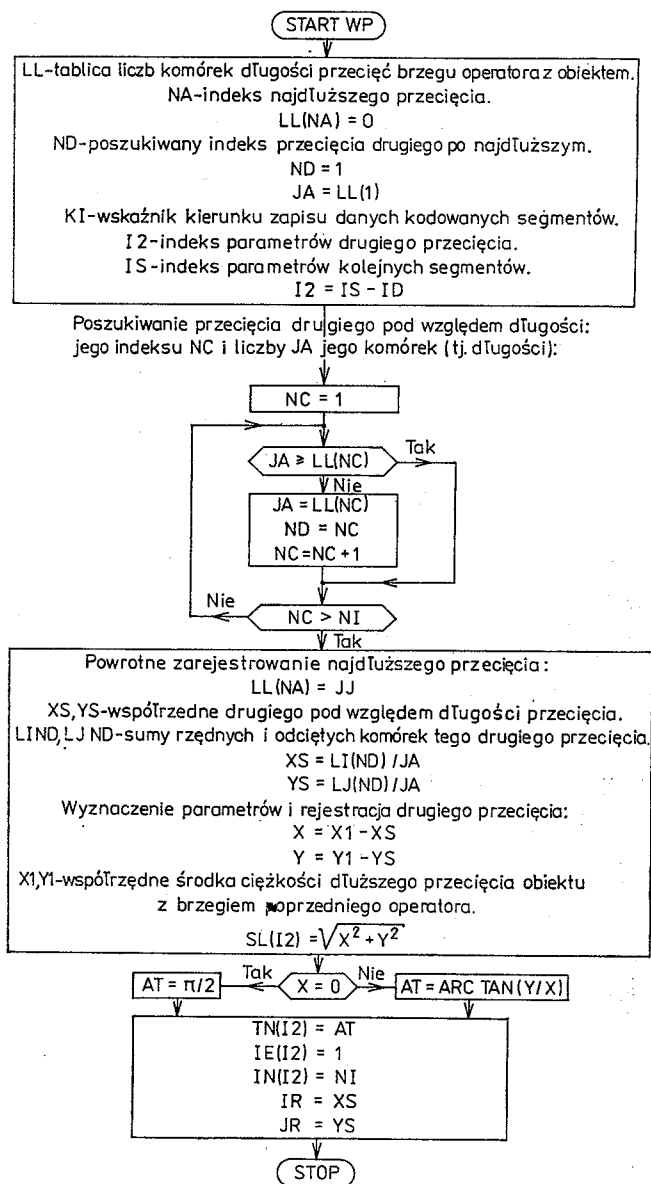
Dowód. Algorytm ten nadaje śledzonemu fragmentowi obiektu nazwę *zakończenie linii* jeżeli:

a) Proces osadzania operatora na obiekcie dobiegł już końca,

b) w trakcie śledzenia stwierdza się brak komórek przecięcia obiektu z całym operatorem (rys. 5j oraz punkt (ii)) lub z brzegiem operatora (rys. 5h oraz punkt (iva)). Liczba NI przecięć brzegu operatora wynosi wtedy zero i jest ona prawidłowa, ponieważ przed każdym kolejnym krokiem śledzenia operator jest zerowany, w wyniku czego NI jest o 1 mniejsze od liczby I_p przecięć obiektów z brzegiem operatora:

$$NI = I_p - 1 \quad (2)$$

Równość $NI = I_p$ zachodzi podczas usadawiania operatora na linii (rys. 5b, c) i w pierwszym kroku kodowania grafowego (rys. 5d), gdy operator nie jest zerowany. Nazwa



Rys. 10. Procedura WP wyznaczania parametrów drugiego pod względem długości przecięcia brzegu operatora z obiektem i współrzędnych IR , JR dla powrotu operatora po prześledzeniu całego pierwszego odcinka obiektu

zakończenie linii jest więc nadawana właściwie. Nazwa *odcinek linii* jest nadawana w pozostałych przypadkach położenia operatora na pojedynczej linii bez rozwidleń i skrzyżowań.

Wszystkie parametry kodowanych segmentów są wyznaczane ze znanych, prostych zależności geometrycznych:

- a) Jeśli operator jest w trakcie kodowania odcinka pojedynczego obiektu nie zawierającego rozwidleń ani skrzyżowań, do obliczania nachylenia odcinka śledzonej linii bierze się pod uwagę wszystkie komórki tej linii zawarte aktualnie w polu operatora (rys. 5a, f...j);
- b) W czasie osadzania operatora na linii (rys. 5b, c) oraz w pierwszym kroku kodowania grafowego (rys. 5d) do obliczania nachylenia odcinka linii bierze się pod uwagę tylko komórki przecięcia linii z brzegiem, cdbu.

Lemat 12. Parametry wejściowe dla kodowania kształtu fragmentu obiektu, pozostałego po zakodowaniu głównego fragmentu linii, są wyznaczane przez następującą procedurę (rys. 10):

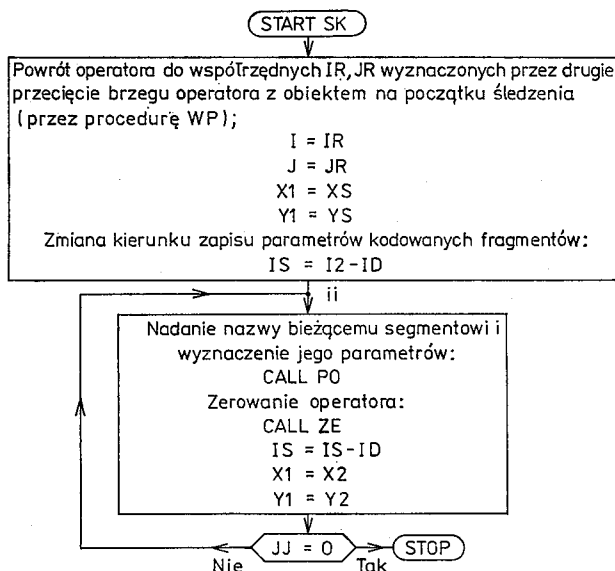
- i. wyznaczenie indeksu zapisu parametrów drugiego przecięcia obiektu z operatorem rozpoczynającym śledzenie;
- ii. wyszukanie drugiego (po najdłuższym) przecięcia brzegu operatora z obiektem przez tymczasowe usunięcie z bazy danych przecięcia najdłuższego spośród wszystkich przecięć obiektu z operatorem rozpoczynającym śledzenie, a następnie przez poszukiwanie przecięcia najdłuższego wśród pozostałych danych;
- iii. wyznaczenie środka ciężkości tego drugiego przecięcia. Jest on jednocześnie pozycją, do której powraca operator po zakończeniu kodowania głównego fragmentu obiektu, aby rozpocząć śledzenie fragmentu pozostałego;
- iv. wyznaczenie i zarejestrowanie odległości pomiędzy środkiem ciężkości tego drugiego przecięcia (rys. 5d) a środkiem ciężkości najdłuższego przecięcia obiektu z brzegiem operatora osadzonego (rys. 5c); obliczenie i zarejestrowanie nachylenia odcinka łączącego te dwa środki ciężkości.

Dowód. Usadowiony operator, znajdujący się w pierwszym kroku kodowania (rys. 5d) rozdziela obiekt na dwie części. Jedno z przecięć brzegu operatora z obiektem jest pierwszym i wyznacza ono początkowy kierunek kodowania linii. Wpis parametrów kodowych odcinków rozpoczyna się w okolicy środka bazy danych (tablic). Drugie przecięcie określa pozycję operatora i kierunek kontynuowania śledzenia obiektu po zakończeniu kodowania linii w kierunku wyznaczonym przez pierwsze przecięcie. Ponieważ obydwa przecięcia leżą na brzegu jednego operatora osadzonego na tej samej linii, parametry obydwu przecięć powinny być zapisywane obok siebie. Dlatego indeks parametrów drugiego przecięcia różni się o jeden od indeksu parametrów pierwszego przecięcia, a zapis parametrów obiektu, kodowanego od tego drugiego przecięcia, odbywa się (rys. 5i) w kierunku przeciwnym do kierunku zapisu pierwszego fragmentu obiektu.

Dla uniknięcia zapętlenia, osadzony operator jest przesuwany do pozycji dłuższego przecięcia obiektu z brzegiem operatora (lema 11 i rys. 5c, d). Dlatego pierwszym przecięciem jest przecięcie dłuższe, a krótsze przecięcie musi być następnie odszukane w bazie danych. Potrzebne parametry tego drugiego przecięcia obliczane są po podstawieniu odszukanych danych wejściowych do prostych związków geometrycznych, cdbu.

Lemat 13. Śledzenie końcowego odcinka linii bez rozwidleń (rys. 5, i, j) jest wykonywane w następujący sposób (rys. 11):

- i. — nadanie elementowi centralnemu operatora współrzędnych środka ciężkości drugiego przecięcia obiektu z brzegiem operatora (np. rozpoczynającego śledzenie, rys. 5d);



Rys. 11. Procedura SK śledzenia i kodowania pozostającego, końcowego odcinka obiektu o kształcie linii nie posiadającego rozwidleń lub skrzyżowań

- ustalenie indeksu początkowego i zmiana kierunku zapisu danych,
- ii. — nadanie nazwy bieżącemu segmentowi przykrytemu operatorem i wyznaczenie parametrów tego segmentu (lemat 11);
- zerowanie fragmentu badanego obrazu zawartego przez operator;
- wyznaczenie indeksu dla wpisu parametrów następnego segmentu do bazy danych;
- zarejestrowanie współrzędnych środka ciężkości przecięcia obiektu z całym operatorem;
- jeśli brzeg operatora zawiera komórki reprezentujące obiekt, skok na początek (ii), w przeciwnym razie koniec śledzenia.

Dowód. Parametry każdego bieżącego segmentu, wyznaczone zgodnie z Lematem 11, zapewniają prawidłowe przesuwanie operatora w położenie środka przecięcia brzegu operatora z obiektem w każdym nowym kroku kodowanie grafowego kształtu, aż stwierdzone zostanie zakończenie linii. Zerowanie pola operatora po każdym kroku wyznaczania parametrów nie dopuszcza do zapętlenia programu i do powtarzania obliczeń na tych samych komórkach reprezentujących obiekt. Parametry kolejnych odcinków linii są wpisywane do bazy danych zgodnie z kierunkiem śledzenia, przeciwnie do kierunku rejestrowania danych głównego fragmentu obiektu i począwszy od indeksu rozpoczęcia zapisu parametrów tego głównego fragmentu linii, cbdu.

Lemat 14. Proces śledzenia i kodowania obiektu w kształcie linii bez rozwidleń i skrzyżowań jest następujący (rys. 12);

- i. ustalenie współrzędnych pierwszego napotkanego elementu obiektu, skąd rozpocznie się osadzanie operatora śledzącego;

- odciętej środka ciężkości poprzedniego przecięcia, parametry kolejnych segmentów pierwszego odcinka obiektu będą zapisywane w kierunku wzrastających adresów bazy danych, a drugiego odcinka obiektu odwrotnie (patrz lemat 13);
- iv. nadanie nazwy bieżącemu segmentowi przykrytemu operatorem i wyznaczenie jego parametrów (lemat 11). Jeśli nie jest to początek kodowania grafowego obiektu, skok do (vi);
 - v. jeśli liczba przecięć obiektu z brzegiem operatora przekracza jeden i operator znajduje się nadal w pierwszym kroku śledzenia, wyznacza się miejsce powrotu operatora po zakończeniu kodowania pierwszego odcinka obiektu w celu kontynuowania śledzenia w drugim kierunku przez obliczenie parametrów drugiego pod względem długości przecięcia obiektu z brzegiem operatora (patrz lemat 12);
 - vi. zerowanie całej powierzchni operatora,
 - jeśli liczba komórek najdłuższego przecięcia obiektu z brzegiem operatora wynosi zero, skok do (vii),
 - odcięta i rzędna środka ciężkości poprzedniego przecięcia obiektu z operatorem przyjmuje obliczone (zgodnie z Lematem 11) współrzędne środka ciężkości bieżącego przecięcia;
 - wyznaczenie indeksu parametrów następnego segmentu końcowego obiektu, zgodnie z kierunkiem rejestrowania ustalonym w (iii);
 - skok do (iv);
 - vii. śledzenie pierwszego fragmentu obiektu zostało zakończone. Śledzenie pozostałego fragmentu obiektu zgodnie z lematem 13.

Dowód. Przedstawiony powyżej proces śledzenia obiektu o kształcie linii obejmuje:

- a) Osadzanie operatora na obiekcie,
- b) Ustalanie kierunku śledzenia i rejestracji danych na początku kodowania grafowego oraz wyznaczanie miejsca powrotu dla operatora, który zakończy śledzenie pierwszego fragmentu obiektu;
- c) Nadawanie nazwy każdemu segmentowi obiektu zawartemu w polu operatora, wyznaczanie parametrów tego segmentu zgodnie z lematem 11;
- d) Zerowanie pola tego operatora po zakończeniu analizy każdego segmentu, dzięki czemu unika się zapętlenia programu lub powtórnego badania tego samego fragmentu obiektu oraz zapewnia się przemieszczanie operatora do właściwej, nowej pozycji (wyznaczanej zgodnie z lematem 11);
- e) Wyznaczanie właściwego indeksu dla wpisu do bazy danych parametrów każdego segmentu obliczonych zgodnie z lematem 11;
- f) Powrót do rozpoczęcia śledzenia pozostającego drugiego odcinka linii po zakończeniu kodowania odcinka pierwszego (patrz lemat 13), cbdu.

Wynik rozpoznawania konturów (tj. graf) spełnia warunek podstawowy reprezentowania obrazów: jeżeli kontur istnieje w obrazie, to jego graf jest otrzymywany niezależnie od jego położenia. Zmieniają się jedynie szczegółowe parametry konturu; jego współrzędne i kierunki poszczególnych odcinków, przy czym względne (wzajemne) kierunki pozostają niezmiennione.

5. IDENTYFIKACJA ROZPOZNANEGO OBIEKTU

Intuicyjnie odróżniamy od siebie pojęcia „rozpoznawanie” i „identyfikacja”, a np. rozpoznajemy gwiazdy na niebie, ponieważ je widzimy jako świecące punkty, ale nie możemy zwykle ich zidentyfikować, tzn. przypisać tym gwiazdom poprawnych nazw, jeżeli nie zajmujemy się astronomią.

Założenie 10. Każdy wzorzec (np. graf) stosowany do identyfikacji rozpoznawanych obiektów jest uzyskiwany przez rozpoznawanie obiektu wzorcowego za pomocą algorytmów, których wyniki spełniają podstawowy warunek reprezentowania obrazów.

Wniosek 9. Zbiór wzorców jest rozstrzygalny semantycznie [2], ponieważ każdy wzorzec spełnia podstawowy warunek reprezentowania scen, tzn. istnieje efektywny sposób rozstrzygnięcia, czy wzorzec ten jest czy nie jest prawdziwy.

Definicja 1. Identyfikacja jest procesem nadawania nieznanemu obiektowi nazwy tego wzorca, z którym jest identyczny jakościowo wynik rozpoznawania tego obiektu.

Twierdzenie 10. Identyfikacja jest przeprowadzalna przez zbadanie rozstrzygalności semantycznej wyniku rozpoznawania obiektu na zbiorze wzorców.

Dowód. Zarówno wzorce jak i wyniki rozpoznawania są rozstrzygalne w sensie możliwości efektywnego sprawdzenia, czy są one prawdziwe. Z prawdziwości tej wynika, że zarówno każdy obraz wzorcowy jak i obraz badany są reprezentowane przez symbole o skończonych liczbach znaków (np. każdy za pomocą grafu o skończonej liczbie węzłów i łuków), ponieważ symbole te spełniają podstawowy warunek reprezentowania scen. Ponieważ istnieje metoda efektywnego sprawdzania identyczności każdych dwóch symboli o skończonych liczbach znaków, istnieje tym samym metoda rozstrzygnięcia, czy jeden ze wzorców jest identyczny z wynikiem rozpoznawania nieznanego obiektu, tzn. czy jeden ze wzorców jest wynikiem rozpoznawania, cbdu.

6. STEROWANIE PROCESEM KODOWANIA KSZTAŁTU

Jako bardziej złożony przykład rozpatrzmy proces kodowania kształtu struktury grzebieniowej, zawierającej rozwidlenia, jak na rys. 14a.

Lemat 15. W przypadku linii bez pętli, każde rozwidlenie daje jeden nowy odcinek do kodowania.

Dowód. W trakcie śledzenia linii operator jest przemieszczany przez rozwidlenie wzdłuż jednej z linii dochodzącej do rozwidlenia i następnie wzdłuż drugiej odchodzącej od rozwidlenia.

Pozostaje więc jeszcze jedna linia rozwidleniowa do prześledzenia, cbdu.

Twierdzenie 11. W linii bez pętli i bez skrzyżowań liczba G rozwidleń jest o 1 mniejsza od liczby F kodowanych linii:

$$F = G + 1 \quad (3)$$

Dowód (przez indukcję). Dla linii bez rozwidleń liczba linii równa jest jeden. Dla linii z jednym rozwidleniem liczba linii kodowanych oddzielnie wynosi dwa (patrz lemat 15). Wzór (3) jest więc zachowany. Załóżmy teraz, że liczba rozwidleń linii wynosi G i założmy prawdziwość wzoru $F = G + 1$. Na mocy lematu 15 dodanie nowego rozwidlenia ($G + 1$) zwiększa dotychczasową liczbę F kodowanych oddzielnie linii o 1:

$$(F+1) = (G+1)+1$$

cbdu.

Lemat 16. W przypadku linii bez pętli, każde skrzyżowanie dodaje dwa nowe odcinki do kodowania.

Dowód. Operator jest przemieszczany wzdłuż jednej z linii dochodzącej do skrzyżowania, a następnie wzdłuż innej linii odchodzącej od skrzyżowania. W wyniku zerowania (kasowania) zawartości operatora, pozostałe dwie linie odchodzące od skrzyżowania zostają rozłączone, tworząc dwa nowe odcinki do kodowania, cbdu.

Twierdzenie 12. W linii bez pętli liczba F kodowanych oddzielnie odcinków wynosi:

$$F = G + 1 + K \times 2 \quad (4)$$

gdzie K jest liczbą skrzyżowań.

Dowód (przez indukcję). Prawdziwość wzoru (4) dla $K = 0$ (brak skrzyżowań) została udowodniona w *Twierdzeniu 11*. W przypadku jednego skrzyżowania prawdziwość (4) wynika z lematu 16. Zakładając prawdziwość (4) dla K skrzyżowań i dodając jedno nowe skrzyżowanie uzyskujemy natomiast na mocy LEMATU 16 zwiększenie liczby F linii o 2;

$$(F+2) = G+1+(K+1) \times 2$$

cbdu.

Twierdzenie 13. Każda pojedyncza pętla zmniejsza liczbę kodowanych linii o jeden.

Dowód. Linia bez rozwidleń i skrzyżowań nie może zawierać pętli, ponieważ pętla tworzy rozwidlenie lub skrzyżowanie. W przypadku jednego rozwidlenia i jednej pętli ($L = 1$) liczba kodowanych linii wynosi jeden (rys. 13a):

$$F = G + 1 - L \quad (5)$$



Rys. 13. Przykład pętli rozwidleniowej i pętli skrzyżowaniowej

Dla jednej pętli i jednego skrzyżowania liczba kodowanych linii wynosi dwa, ponieważ w wyniku zerowania pola operatora, operator przemieszczając się przez skrzyżowanie oddziela drugi odcinek linii od skrzyżowania:

$$F = G + 1 + K \times 2 - L \quad (6)$$

LICZBY NI PRZECIEC PRZEGU Z ORIEKTEM

7 2
POZYCJE PRZECIEC OPERATORA Z OBIEKTEM

8.00 6.00

0.32 0.32

1 1

3.16 3.16

3.16 3.16



14b

LICZBY NI PRZECIEC BRZEGU Z ORIEKTEM

1 1 1 1 2 2
MASY SEGMENTOW (W LICZBACH KOMOREK):

15 5 4 4 7 2
POZYCJE PRZECIEC OPERATORA Z OBIEKTEM

29.53 27.80 25.00 22.25 21.00 15.00

14.33 12.00 10.50 9.25 8.00 6.00

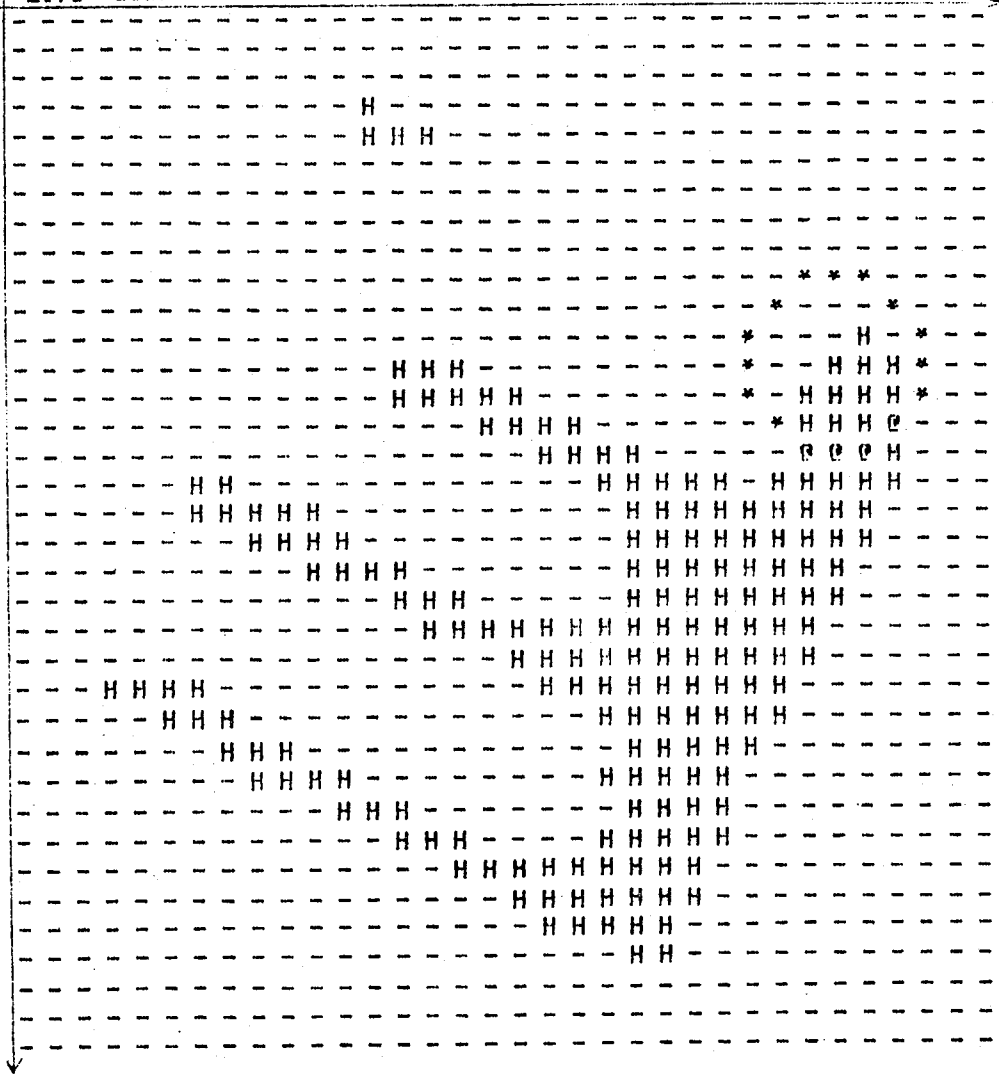
NACHYLENIA SEGMENTOW (W RAD)

0.93 0.49 0.43 0.79 0.32 0.32

1 OZNACZA ODCINEK LINII; 2 -ZAKONCZENIE; 3 -ROZWIDLENIE:

1 1 1 1 1 1
DLUGOSCI ODCINKOW LINII (W LICZBACH KOMOREK):

2.91 3.18 3.02 1.77 3.16 3.16



14c

LICZBY NI PRZECIEC BRZEGU Z ORIEKTEM

1 1 1 1 1 2 2
MASY SEGMENTOW (W LICZBACH KOMOREK):

11 15 5 4 4 7 2

POZYCJE PRZECIEC OPERATORA Z ORIEKTEM

29.27 29.53 27.80 25.00 22.25 21.00 15.00

17.55 14.33 12.00 10.50 9.25 8.00 6.00

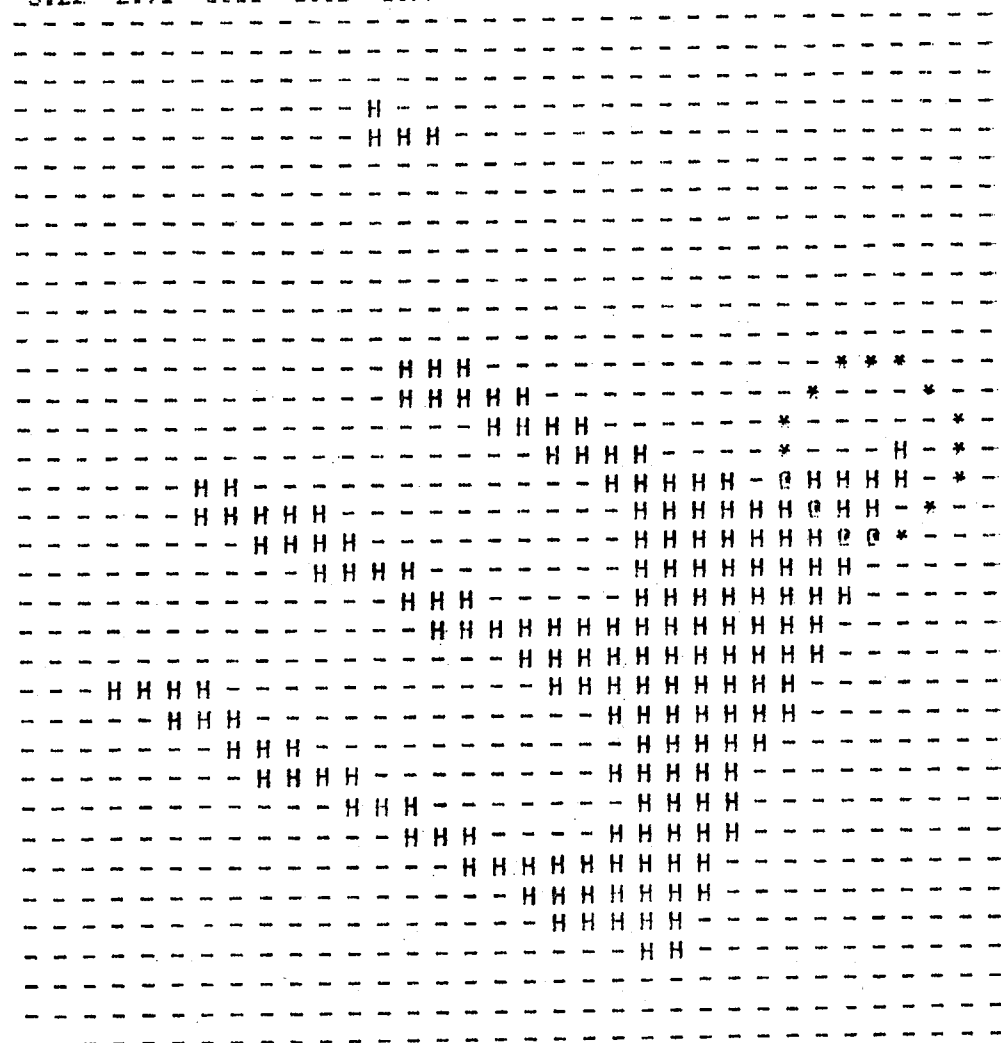
NACHYLENIA SEGMENTOW (W RAD)

-1.49 0.93 0.49 0.43 0.79 0.32 0.32

1 OZNACZA ODCINEK LINII: 2 -ZAKONCZENIE: 3 -ROZWIDLENIE:

1 1 1 1 1 1 1
DLUGOSCI ODCINKOW LINII (W LICZBACH KOMOREK):

3.22 2.91 3.18 3.02 1.77 3.16 3.16



14d

LICZBY NI PRZECIEC BRZEGU Z OBIEKTEM

1 1 1 1 1 1 2 2

MASY SEGMENTOW (W LICZBACH KOMOREK):

10 11 15 5 4 4 7 2

POZYCJE PRZECIEC OPERATORA Z OBIEKTEM

27.50 29.27 29.53 27.80 25.00 22.25 21.00 15.00

19.50 17.55 14.33 12.00 10.50 9.25 8.00 6.00

NACHYLENIA SEGMENTOW (W RAD)

-0.83 -1.49 0.93 0.49 0.43 0.79 0.32 0.32

1 OZNACZA ODCINEK LINII; 2 -ZAKONCZENIE; 3 -ROZWIDLENIE:

1 1 1 1 1 1 1 1

DLUGOSCI ODCINKOW LINII (W LICZBACH KOMOREK):

2.64 3.22 2.91 3.18 3.02 1.77 3.16 3.16

[illegible]

[illegible]

14 f

LINEARNY FRAGMENT 1 WYKŁAD ZREDUKOWANEGO ROZPOZYCHANIA LINII:

MASY ROZPOZYCHANIA CECI LINII: 1 OZNACZA DŁUGIEC LINII
2 OZNACZA ZAKOŃCZENIE: 3 - ROZWIDLENIE: 4 - LUK: 5 - NAROZIE:

	2	1	5	4	3	5	1	2
SREDNIE PUZYCJE X-Y ROZPOZYCHANIA CECI LINII:								
4.50 11.95 19.00 21.83 27.16 26.00 29.77 29.53 23.00 19.40	13.75							
24.00 27.45 31.20 31.42 26.38 20.52 17.55 14.33 17.00 7.90	4.75							
DŁUGOSCI ROZPOZYCHANIA FRAGMENTOW LINII (W LICZBACH KOHOREK):								
2.38 13.94 2.84 5.05 7.00 4.42 3.22 2.91 3.18 12.88	0.00							
SREDNIE NACHYLENIA TYCH FRAGMENTOW (RAD):								
0.00 0.48 0.08 -1.51 -0.49 -0.74 -1.49 0.93 0.49 0.40	0.00							
CAŁKOWITE ZMIANY NACHYLEN TYCH FRAGMENTOW (RAD):								
0.00 -0.01 -0.46 1.56 0.52 0.25 -0.75 -0.72 -0.44 -0.06	0.00							
SREDNIE GRUBOSCI TYCH FRAGMENTOW (W LICZBACH KOHOREK):								
0.84 1.51 1.76 2.38 4.57 13.11 3.41 5.16 1.57 1.63	0.00							
MASY ROZPOZYCHANIA CECI LINII (W LICZBACH KOHOREK):								
2 21 5 12 32 58 11 15 5 21 0								

WSPÓLRZĘDNE ROZWIĄDLENIA 1: $X=26,00$; $Y=20,52$

Rys. 14. Ważniejsze etapy (a — j) kodowania grafowego linii głównej obszaru metalizacji emitera — obiektu zawierającego dwa rozwidlenia oraz ostateczny wynik (k) rozpoznawania zakodowanej linii głównej po redukcji segmentów tworzących odcinki proste i po identyfikacji łuków i naroży

16a

LICZBY NI PRZECIEC BRZEGU Z OBIEKTEM

2 2 1
MASY SEGMENTOW (W LICZBACH KOMOREK):

0 2 4
POZYCJE PRZECIEC OPERATORA Z OBIEKTEM

22.16 18.00 16.00

26.38 22.50 21.75

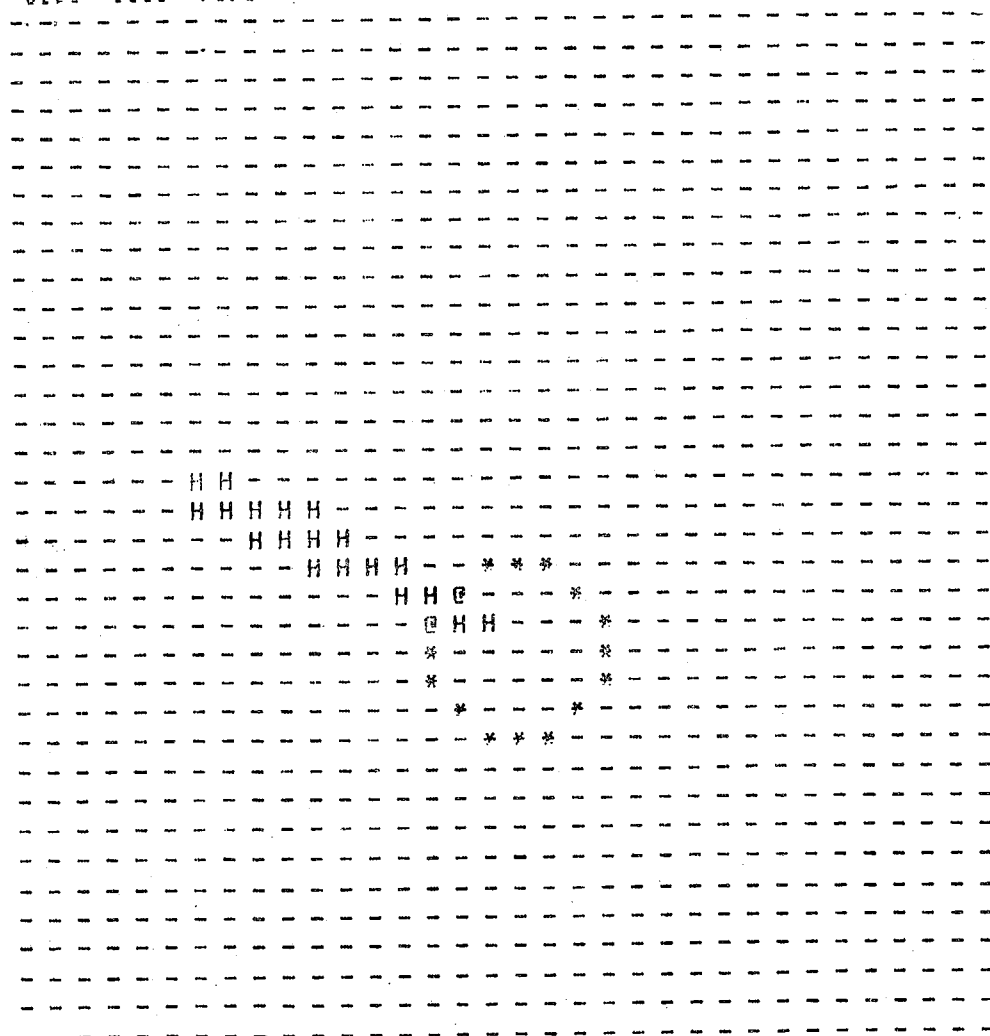
NACHYLENIA SEGMENTOW (W RAD)

0.00 0.00 0.00

1 OZNACZA ODCINEK LINII: 2 -ZAKONCZENIE: 3 -ROZWIDLENIE:

3 1 1
DLUGOSCI ODCINKOW LINII (W LICZBACH KOMOREK):

0.00 0.00 7.70



16b

LICZBY NI PRZECIEC BRZEMU Z ORBIKTEM
 2 2 1 1 1 1 0
 MASY SEGMENTOW (W LICZBACH KOMOREK):
 0 2 4 3 5 5 4
 POZYCJE PRZECIEC OPERATORA Z ORBIKTEM
 22.16 18.00 16.00 14.33 11.00 9.00 7.50
 26.38 22.50 21.75 20.67 19.60 18.40 17.50
 NACHYLENIA SEGMENTOW (W RAD)
 0.00 0.00 0.00 0.58 0.40 0.54 0.00
 1 OZNACZA ODCINEK LINII: 2 -ZAKONCZENIE: 3 -ROZWIDLENIE:
 3 1 1 1 1 1 2
 DLUGOSCI ODCINKOW LINII (W LICZBACH KOMOREK):
 0.00 0.00 7.70 1.99 2.75 2.33 2.47

ZMIANY W NACHYLENIU ODCINKOW LINII
 0.00 0.00 0.00 0.00 -0.18 0.14 0.00

16c

LINEARNY FRAGMENT 3 GRAFU ZPRZEDUKOWANEGO ROZPOZNANEJ LINII:
 NAZWY ROZPOZNANYCH CECH LINII: 1 OZNACZA ODCINEK LINII
 2 OZNACZA ZAKONCZENIE: 3 -ROZWIDLENIE: 4 -LUK: 5 -NAROZE:
 3 1 2
 SREDNIE POZYCJE X:Y ROZPOZNANYCH CECH LINII:
 22.16 17.99 7.50
 26.38 20.58 17.50
 DLUGOSCI ROZPOZNANYCH FRAGMENTOW LINII (W LICZBACH KOMOREK):
 0.00 14.77 2.47
 SREDNIE NACHYLENIA TYCH FRAGMENTOW (RAD):
 0.00 0.51 0.00
 CALKOWITE ZMIANY NACHYLEN TYCH FRAGMENTOW (RAD):
 0.00 -0.04 0.00
 SREDNIE GRUBOSCII TYCH FRAGMENTOW (W LICZBACH KOMOREK):
 0.00 1.29 1.62
 MASY ROZPOZNANYCH CECH LINII (W LICZBACH KOMOREK):
 0 19 4
 PARAMETRY GLOBALNE KAZDEJ Z 3 LINII SKLADOWYCH ORBIKTU:
 LINIA NR.: 1 2 3
 DLUGOSCI LINII: 57.82 13.87 17.24
 LICZBY KOMOREK (MASY): 182.0 19.0 23.0

Rys. 16. Pierwszy (a) i końcowy (b) etap śledzenia drugiej linii rozwidleniowej metalizacji emitera oraz ostateczny wynik (c), rozpoznawania tej linii rozwidleniowej

Dla każdej liczby $G > 1$ lub $K > 1$ wzory (5) i (6) pozostają słuszne na mocy twierdzeń 11 i 12, ponieważ zarówno rozwidlenia jak i skrzyżowania są addytywne, cbdu.

Obiekt grzebieniowy (będący emiterym tranzystora), wydzielony w wyniku segmentacji [2], jest kodowany wzdłuż linii głównej otaczającej dwa środkowe występy. Rysunki 14a—e przedstawiają ważniejsze, kolejne etapy śledzenia tej linii, z pominięciem pozycji operatora na prostych odcinkach. Jeżeli długość przecięcia brzegu operatora z obiektem jest większa lub równa średnicy tego operatora (w liczbach komórek), generowany jest operator o większej średnicy. Długość tego przecięcia jest mierzona w każdym kroku kodowania.

Do parametrów rozwidlenia dołączone są parametry segmentu bezpośrednio poprzedzającego rozwidlenie aby wyrównać długości wszystkich trzech odcinków tworzących rozwidlenie. W momencie napotkania na rozwidlenie ($NI = 2$) zapamiętywane są w oddzielnej tablicy współrzędne krótszego (lub drugiego z kolei) przecięcia brzegu operatora z obiektem oraz pozycja (indeks) parametrów rozwidlenia w bazie danych. Ze współrzędnych tych rozpoczyna się śledzenie linii rozwidleniowych (rys. 15a i 16a).

Pierwszą pozycję parametrów linii rozwidleniowych stanowią parametry rozwidlenia, przepisane z tablic parametrów linii głównej zgodnie z zapamiętaną pozycją danych rozwidlenia. Drugą pozycję danych stanowią parametry krótszego (lub drugiego) przecięcia rozwidlenia z brzegiem operatora (rys. 15a oraz 16a). Dopiero trzecią pozycją danych są parametry linii wyznaczone dla pierwszego położenia operatora na linii. Nachylenie jest wpisywane dopiero od czwartej pozycji danych, ponieważ położenie rozwidlenia zwykle odbiega od kierunku linii rozwidleniowych.

Poza przedstawionymi powyżej zasadami generowania operatora o większej średnicy przetwarzania parametrów rozwidlenia, śledzenie linii głównej (rys. 14a—e) odbywa się zgodnie z lematem 11 (rys. 12), natomiast linie rozwidleniowe (rys. 15a, b, oraz 16a, b) są traktowane tak samo jak końcowe odcinki linii (lemat 13 i rys. 11).

7. ALGORYTM REDUKCJI GRAFU REPREZENTUJĄCEGO ROZPOZNANY OBIEKT

Po zakończeniu kodowania linii głównej (rys. 14e) korygowane są nachylenia segmentów sąsiadujących bezpośrednio z operatorem osadzonym (tj. rozpoczynającym kodowanie). Na początku nachylenia te były wyznaczane w stosunku do przecięcia brzegu operatora osadzonego z obiektem; nachylenia te korygowane są przez odniesienie do przecięcia wnętrza operatora osadzonego z obiektem (porównaj wydruk u dołu rys. 14e). Korygowane są także nachylenia końcowych segmentów linii przez uśrednienie (z odpowiednio przyjętymi wagami) z sąsiednimi segmentami. Ostatnim etapem rozpoznawania jest redukcja segmentów tworzących linie proste i identyfikacja naroży.

Lemat 17. Redukcja odcinków linii prostych i identyfikacja naroży oraz łuków jest przeprowadzana w następujący sposób (rys. 17):

- i. odczyt nazwy pierwszego segmentu linii. Jest to zakończenie albo rozwidlenie. Rejestracja tego zakończenia albo rozwidlenia.
- ii. Zerowanie parametrów cząstkowych linii prostej;

iii. Podstawienie aktualnej nazwy segmentu linii redukowanej pod nazwą segmentu poprzedniego.

Odczyt nazwy następnego segmentu redukowanego.

Jeśli odczytana nazwa jest odcinkiem, skok do (iv).

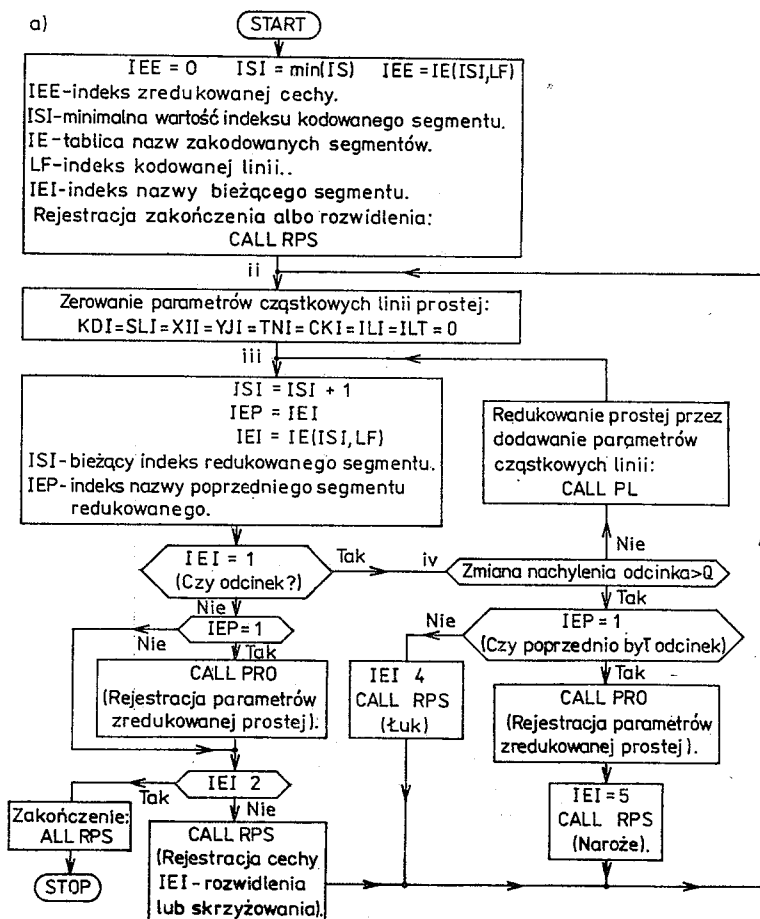
Jeśli poprzednim segmentem był odcinek, redukcja odcinka prostego dobiegła końca i rejestruje się parametry zredukowanej prostej.

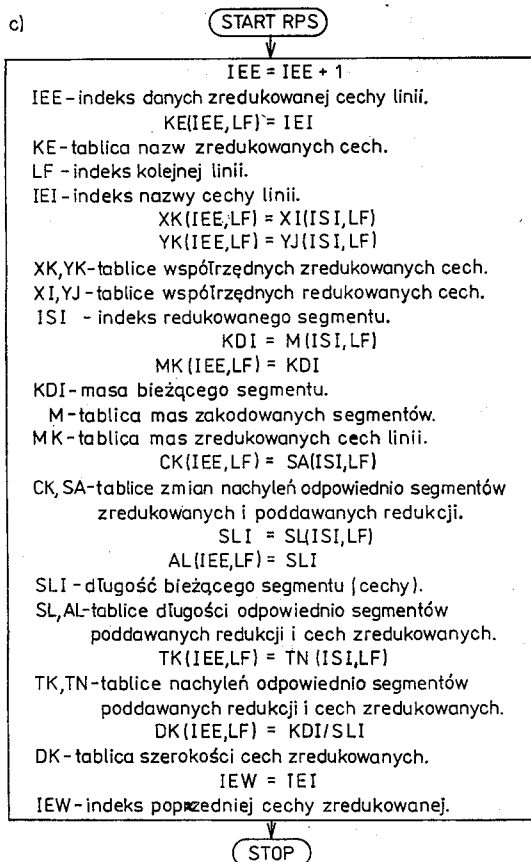
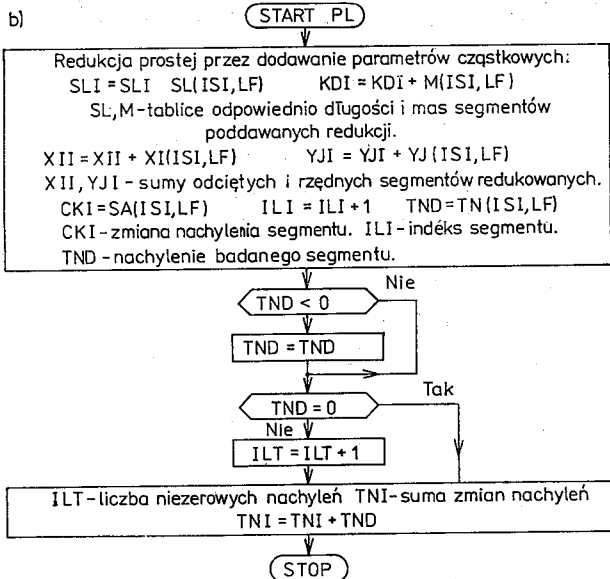
Jeśli bieżącym segmentem jest zakończenie, odbywa się jego rejestracja i algorytm kończy pracę.

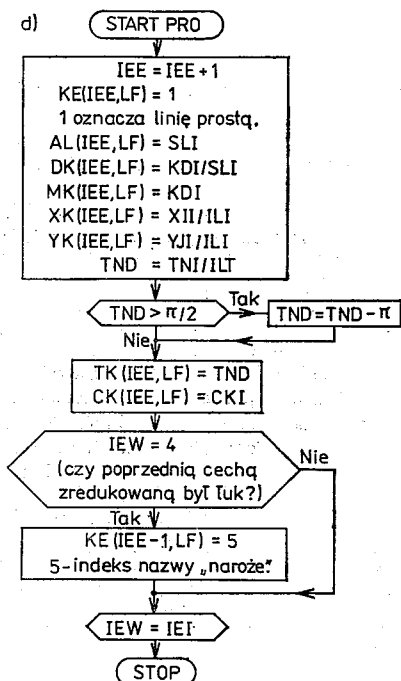
Jeżeli bieżącym segmentem nie jest zakończenie, rejestruje się nazwę tego segmentu (będącego rozwidleniem albo skrzyżowaniem) i następuje skok do (ii);

iv. Jeżeli zmiana nachylenia bieżącego odcinka nie przekracza progowej zmiany nachylenia dla łuku linii, odbywa się redukcja parametrów linii prostej przez dodanie długości, masy, współrzędnych i nachylenia bieżącego odcinka do odpowiedniej, dotychczasowej sumy długości, masy, współrzędnych i nachylenia prostej oraz skok do (iii);

Jeżeli ta zmiana nachylenia przekracza progową zmianę nachylenia dla łuku linii,





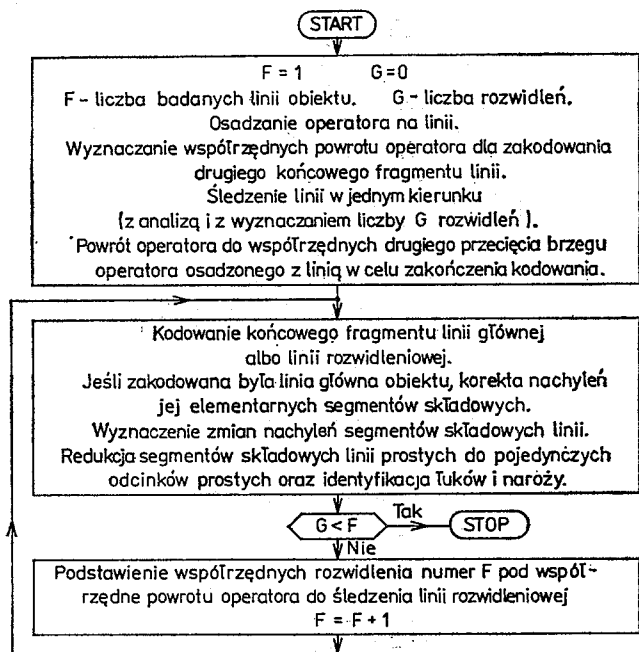


Rys. 17. Algorytm redukcji segmentów tworzących odcinki linii prostych (b) oraz identyfikacji łuków naroży linii oraz redukcji całego grafu obiektu (a) oraz procedura pomocnicza RPS (c) rejestracji parametrów rozwidlenia, zakończenia, łuku, naroża albo skrzyżowania w zredukowanej bazie danych i PRO (d) rejestracji zredukowanego odcinka prostej i identyfikacji naroża występującego przed prostą

identyfikowany jest łuk jeśli poprzednią cechą zredukowaną nie był odcinek, albo kończona jest redukcja prostej przez rejestrację jej parametrów i identyfikowane jest naroże przez wpis indeksu naroża i rejestrację jego parametrów. Po zidentyfikowaniu łuku albo naroża i rejestracji jego parametrów następuje skok do (ii).

Dowód. Algorytm czyta nazwy kolejnych cech zredukowanego obiektu i wpisuje je do bazy danych zredukowanego opisu (reprezentacji) tego obiektu, sprowadzając wszystkie kolejne odcinki tej samej prostej do jednej, właściwej prostej i identyfikując łuk albo naroże linii, jeśli zmiana nachylenia bieżącego odcinka jest właściwa dla łuku albo dla naroża. Naroże jest identyfikowane zamiast łuku, jeżeli zmiana nachylenia właściwa dla łuku następuje po linii prostej (rys. 17a) albo przed linią prostą (rys. 17b), cbdu. Zredukowane reprezentacje grafowe obiektu grzebieniowego (emitera) z rys. 14a są przedstawione na rys. 14f (linii głównej) i na rys. 15c oraz 16c (linii rozwidleniowych).

Zakończenie procesu rozpoznawania obiektu jest wyznaczane na podstawie twierdzeń 11–13. Po przeprowadzeniu redukcji grafu, reprezentującego obiekt o kształcie linii, robiony jest bilans liczby rozpoznanych linii i liczb rozwidleń, skrzyżowań i pętli. Jeżeli liczba linii, tworzących spójny obiekt, przekracza liczbę rozwidleń, podwójną liczbę skrzyżowań minus liczba pętli, następuje koniec rozpoznawania (rys. 18).



Rys. 18. Algorytm rozpoznawania obiektu zawierającego linie proste, rozwidlenia, zakończenia, łuki i naroża, w którym zastosowano twierdzenie łączące liczbę rozwidleń z liczbą kodowanych linii dla wyznaczania końca rozpoznawania

WNIOSKI

Przedstawiony w pracy pełny algorytm rozpoznawania kształtu obiektu linearnego lub konturu jest efektywny, co potwierdzają reprezentowane wyniki praktyczne. Algorytm ten bada tylko te obszary obrazu, w których istnieje obiekt, co zaoszczędza czasu komputera.

Istnieje tu pewna analogia sposobu pracy algorytmu z właściwością oka człowieka, który ogląda obiekt przesuując wzrok wzdłuż istotnych fragmentów obserwowanego obiektu.

Dużą zaletą prezentowanej metody jest niewrażliwość na zmiany grubości analizowanych linii oraz rozpoznawanie przerw w linii (mogących powstać wskutek zakłóceń). Stosowana dotychczas metoda kodowania łańcuchowego (ang. *chaine-coding*) [8] jest bardzo wrażliwa na przerwy, a linie o grubości większej od jednego elementu wymagają wcześniejszego pocienienia do grubości jednej komórki.

BIBLIOGRAFIA

1. P. E. Danielsson, B. Kruse: *Distance Checking Algorithms*. Comp. Graph. Image Proc, 1979, vol. 4, s. 349—376
2. Z. M. Wójcik: *Segmentacja obrazu i rozpoznanie głowicy silnika*. Gliwice, Zeszyty Naukowe Politechniki Śląskiej, 1984, z. 74

3. Z. M. Wójcik: *Conclusiveness of Natural Languages and Recognition of Images*. Hemisphere Publ. Corp., Washington, Cybernetics and Systems, 1983, vol. 14, s. 1—31
4. Z. M. Wójcik: *Algorytmy rozpoznawania linii i znaków alfanumerycznych*. Rozprawy Elektrotechniczne. 1981, z. 1, s. 241—277
5. Z. M. Wójcik: *An Approach to Recognition of Contours and Line-shaped Objects*. Comp. Vision, Graphics and Image Proc. 1984, vol. 25, s. 184—204
6. Z. M. Wójcik: *Metoda detekcji i rozpoznawania konturów obrazu*. Rozprawy Elektrotechniczne, 1983, z. 3, s. 959—983
7. Z. M. Wójcik: *Metoda rozpoznawania liczby obiektów i automatycznej detekcji obrazów wielowartościowych*. Rozprawy Elektrotechniczne, 1981, z. 4, s. 1197—1214
8. H. Freeman: *Analysis of Line Drawings, Digital Image Processing and Analysis*. Nordhoff, Crooningen, 1977

Z. M. WÓJCIK

OBJECT SHAPE CODING AND RECOGNITION ALGORITHMS

Summary

A full algorithm of shape recognition of a linear object or contour is presented. Its efficiency is confirmed by practical results quoted here. This algorithm scans only those areas of the image in which an object exists saving computer time. A coding method of local features of objects by means of Euler numbers is proposed.

Z. M. WÓJCIK

ALGORITHME DE CODAGE ET DE RECONNAISSANCE DE LA FORME DE L'OBJET

Résumé

Dans cet article on a présenté le plein algorithme de reconnaissance de la forme de l'objet ou de son contour. Son effectivité est confirmée par les résultats pratiques. L'algorithme n'examine que ces régions de l'image, dans lesquelles se trouve l'objet, ce qui contribue à épargner le temps de l'ordinateur. Dans cet article on a proposé aussi le codage des signes locaux des objets par les nombres d'Euler.

Z. M. WÓJCIK

ALGORITHMUS FÜR DIE KODIERUNG UND ERKENNUNG DER OBJEKTFORM

Zusammenfassung

Im Aufsatz wurde der komplette Algorithmus für die Erkennung der linearen Objektform oder des Umrisses geschildert. Seine Effektivität bezeugen die dargestellten praktischen Ergebnisse. Der Algorithmus untersucht nur diejenigen Bildbereiche, worin sich das Objekt befindet, wodurch die Zeit des Digitalrechners gespart wird. Es wurde auch die Kodierung der Lokaleigenschaften von Objekten mittels der Euler-Zahl vorgeschlagen.

З. М. ВУЙЦИК

АЛГОРИТМ КОДИРОВАНИЯ И РАСПОЗНАВАНИЯ ФОРМЫ ОБЪЕКТА

Резюме

Представлен полный алгоритм распознавания формы линейного объекта или очертания. Его эффективность подтверждается представленными практическими результатами. Этот алгоритм проверяет только те области образа, в которых существует объект, что экономит время компьютера. Предложено также кодирование локальных черт объектов числами Эйлера.

A quasilogarithmic data compression method for an x-ray spectrometer

MAREK HŁOND

Centrum Badań Kosmicznych (Warszawa)

Received 1987.03.24

Authorized 1987.09.10

A quasilogarithmic method of compression of digital data resulting from photon counters is presented. Such a method will be adapted for an X-ray spectrometer to be installed on board a solar satellite.

1. INTRODUCTION

Instruments placed aboard satellites, such as the X-ray solar photometers and spectrometers, usually measure the intensity of the radiation flux with photon counters. Their signals recorded in digital form by the pulse counters are accumulated and periodically transmitted to Earth by telemetry. In the case of such physical events as solar flares, the measured flux intensity can change during a short time by several orders of magnitude. Satellite telemetric systems usually have a narrow digital transmission channel (most often 8-bit) and a small data transmission rate, caused by a great number of the scientific instruments (telemetric channels) used. This restricts strongly the dynamical range of the measurements. To overcome these limitations in satellite's scientific instruments the compression of transmitted data is commonly used.

2. METHODS OF DATA COMPRESSION

The data from measurements based on the counting of random events, such as from a photon counter, have a statistical character. The standard deviation of such data equals the square root of these counts. Therefore if data (obtained in the accumulation interval Δt) are transmitted from the instrument to Earth in their full form of a n -bit counting number N this information is too precise as compared with the applicable statistical error ($\sqrt{N}$) affecting this number. The latter conclusion that the statistically insignificant bits of the data resulting from the photon counters can be neglected during telemetric transmission, is the basis of application of the compression process. In instruments used for space research three data compression methods are applied: square root, logarithmic and flo-

ating-point compressions. The advantages of these methods have been pointed out by several authors [1, 2, 3]. In general, these compression processes can be represented as a certain kind of transformation (F) of the count number N to the number $F(N)$ [2].

$$N \xrightarrow{F} F(N) \pm \delta_F \quad (1)$$

where δ_F is an extra compression error.

2.1. THE SQUARE ROOT METHOD

This method is based on the conversion of the counting number N to number $\lfloor \sqrt{N} \rfloor$ (the nearest integer to the square root of N). The extra compression error caused by this compression algorithm has a value ± 0.5 and equals the standard deviation δ_N of the source data. This means that while for the transmission of a maximum number N the n -bit counter is necessary a $n/2$ -bit counter would suffice for the transmission of the compressed number $F(N)$. If an extra compression error is too large to accept this method can be modified by increasing the bit size of the counter of compressed data by l bits. This reduces the compression error by the ratio $1/2^l$:

$$N \xrightarrow{F} 2^l \lfloor \sqrt{N} \rfloor \pm \delta_N / 2^l \quad (2)$$

where most often $l = 1$.

In practically available instruments this compression method can be used in the real time during the inflow of information, or indirectly by a transformation process after the count N has been completed [1, 2]. This method is useful for large expected values of N (high counting rates).

2.2. THE LOGARITHMIC METHOD

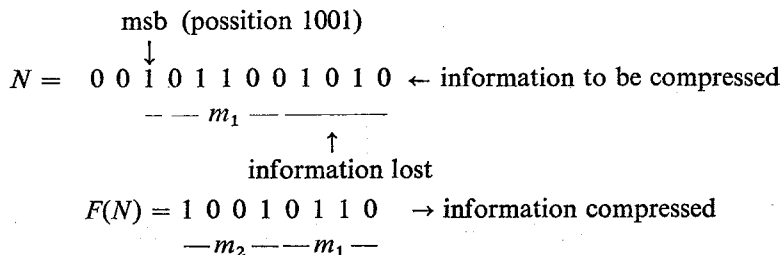
In the logarithmic method each count number N received from the instrument during the time Δt is transformed into the integer nearest to the value $\lfloor B \log_2 N \rfloor$ with an extra error of ± 0.5 . B is a scale coefficient considering the highest possible value of the compressed data.

For instance, if the count number N is to be transmitted from a 12-bit counter by a 8-bit telemetric channel the equation (1) takes the form:

$$N \xrightarrow{F} \lfloor 21.33 \log_2 N \rfloor \pm 0.5 \quad (3)$$

2.3. THE FLOATING-POINT METHOD

In the floating-point method the transmitted count number N from the n -bit counter is presented by the m -bit number $F(N)$ composed of a m_1 -bit mantissa and a m_2 -bit exponent [4, 5]. The exponent indicates the position of the most significant bit (MSB) in the source data while the mantissa contains bits following MSB. For example, the count number N from a 12-bit counter compressed to a 8-bit number $F(N)$, has the form:



3. THE QUASILOGARITHMIC COMPRESSION METHOD USED IN AN X-RAY SPECTROMETER

The method presented below allows for the transformation of data requiring a 12-bit word length to a 8-bit word for the telemetry channel. The block diagram of this device is presented in Figure 1.

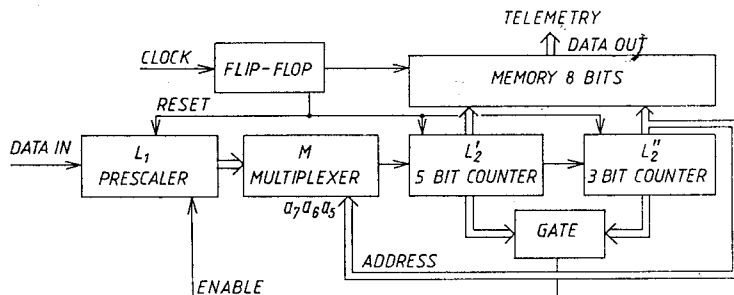


Figure 1. Block diagram of quasilogarithmic data compression device

The system contains two 8-bit counters (scalers). The first one (L_1) is used as divider prescaler and the second (L_2) as the proper counter, cooperating with the memory and the telemetry. The highest bits of L_2 (a_7, a_6, a_5) are simultaneously used for a multiplexer M as an address activating successively its inputs. These inputs are connected to the outputs of the prescaler L_1 acting as an adjustable input divider. The principle of operation of the presented system is to count of the „pack” of pulses. The number of pulses in the pack increases in geometrical manner according to the successive activation of multiplexer inputs by addresses $a_7a_6a_5$ from the L_2 prescaler. Thanks to this course of action, the information about the number of pulses N is subject to compression in the approximately logarithmic scale.

We can determine a count number N recorded by the device from the contents of scaler L_2 , according to equations:

$$N = \begin{cases} \sum_{n=0}^{n=4} 2^n a_n & (\text{for } a_7 a_6 a_5 = 000) \\ 32\Delta k + \sum_{n=0}^{n=4} 2^n a_n \Delta k & (\text{for remaining addresses}) \end{cases} \quad (4)$$

where values of Δk are taken from the Table.

Values of parameter Δk

address ($a_7a_6a_5$)	000	001	010	011	100	101	110	111
Δk	1	2	2	4	8	16	32	64

The described data compression device allows for recording the highest intensity of 40.960 pulse/sec (for 10 Hz frequency of the clock pulses) with an additional compression error displayed in Figure 2. In consequence the extra error related to this method becomes

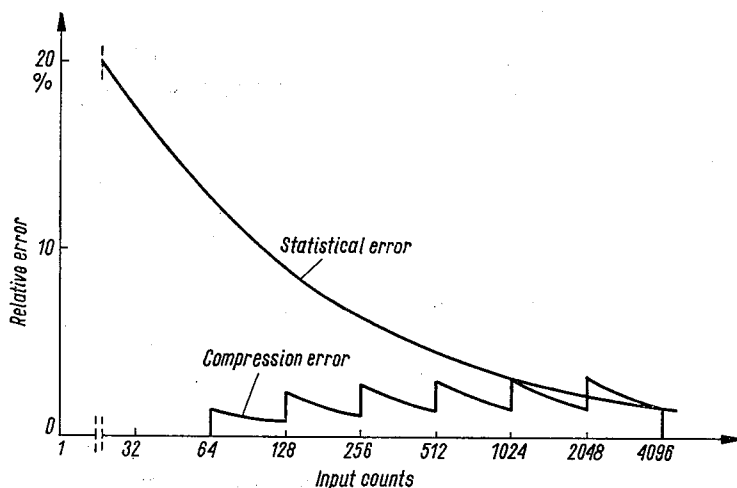


Figure 2. Maximum compression error versus input counts compared to Poissonian error

essential (comparable with the statistical error) for counts of the order 1×10^3 pulses. On the contrary to the square root compression method, the method presented here is more advantageous for lower counts where accuracy is determined by the statistical error δ_N .

The described compression method will be applied in the X-ray spectrometer being developed by our Laboratory for a solar satellite.

REFERENCES

1. G. Green, H.G. Hasler, H. Kunow: *An adaptive data compression method for a cosmic ray experiment on board a space probe*, Nucl. Instr. Metho., 1970, No. 86, p. 213
2. J.M. Beckers: *The compression of data resulting from photon counters*. Space Sci. Instr., 1975, No. 1, p. 153
3. D.W. Jurhens: *Sigma data compression*. IEE Trans. Nucl. Sci., 1975, Vol. NS-22, p. 603.
4. I.I. Van Rooijen, G.I. Van Dijen, H.T. Lafleur, P. Lawes: *A low energy photon spectrometer for directional distribution measurements in space*. Space Sci. Instr. 1979, No. 4, p. 373.
5. A. Balogh, T. Iversen: *Digital on board processor for flux measurements*. Space Sci. Instr. 1977, No. 3, p. 187

M. HŁOND

METODA QUASILOGARYTMICZNEJ KOMPRESJI DANYCH
W SPEKTROMETRZE RENTGENOWSKIM

Streszczenie

Przedstawiono metodę quasilogarytmicznej kompresji danych pomiarowych uzyskiwanych z liczników fotonów. Metoda ta będzie wykorzystana w spektrometrze rentgenowskim konstruowanym dla satelity słonecznego.

M. HŁOND

MÉTHODE DE COMPRESSION QUASI-LOGARITHMIQUE DES DONNÉES
POUR LE SPECTROMÈTRE AUX RAYONS X

Résumé

Dans cet article on a présenté la méthode de compression quasi-logarithmique des données obtenues à partir des compteurs des photons. Cette méthode sera adaptée pour le spectromètre aux rayons X, qui sera installé sur un satellite solaire.

M. HŁOND

QUASILOGARITHMISCHE METHODE FÜR DIE DATENKOMPRESSION
IM RÖNTGENSPEKTROMETER

Zusammenfassung

In dem Artikel wurde eine quasilogarithmische Methode für die Kompression von Meßangaben dargestellt, die von Photonenzählern gewonnen werden. Diese Methode wird bei einem für einen Sonnensatelliten konstruierten Röntgenspektrometer ihre Anwendung finden.

М. ХЛОНД

МЕТОД КВАЗИЛОГАРИФМИЧЕСКОГО СЖАТИЯ ИНФОРМАЦИИ ДЛЯ
РЕНТГЕНОВСКОГО СПЕКТРОМЕТРА

Резюме

Представлен метод квазилогарифмического сжатия измерительных данных получаемых со счетчиков фотонов. Указанный метод будет применен в рентгеновском спектрометре предназначенном для солнечного спутника.

Modelowanie zjawisk fizycznych w laserach złączowych

WŁODZIMIERZ NAKWASKI

Instytut Fizyki, Politechnika Łódzka

Otrzymano 1987.01.27

Autoryzowano do druku 1987.11.27

CZĘŚĆ I — MODELOWANIE ZJAWISK ELEKTRYCZNYCH

W artykule przedstawiono kompletny model zjawisk elektrycznych w biheterozłączowym laserze paskowym. W modelu wyznaczono rozkłady potencjału w obszarze lasera oraz rozkłady koncentracji elektronów i dziur w obszarze czynnym.

1. WPROWADZENIE

Ogromnemu postępowi, zachodzącemu w ostatnich latach w technologii wytwarzania laserów złączowych, cały czas towarzyszył analogiczny postęp w pracach teoretycznych zajmujących się modelowaniem zjawisk fizycznych w laserze. Prace te miały za zadanie lepsze poznanie działania przyrządu oraz umożliwienie jego optymalizacji dla szczególnych zastosowań.

W niniejszej pracy zostaną przedstawione zasady właściwego modelowania zjawisk fizycznych w nowoczesnych laserach złączowych. Pojemność i szybkość pracy obecnych komputerów czynią niemożliwym rozwiązanie pełnego układu podanych równań. Należy więc korzystać z odpowiednich metod przybliżonych.

Realizacje powyższych zasad modelowania dla konkretnych konstrukcji laserów nazywać będziemy modelami tych laserów. Mogą się one wzajemnie znacznie różnić, gdyż różne konstrukcje laserów umożliwiają stosowanie odmiennych przybliżeń.

Niniejsza praca jest w znacznej mierze oparta na dwóch wybitnych publikacjach: pracy Wilta i Yariva [1] — głównie w I części o zasadach modelowania zjawisk elektrycznych, oraz pracy Buusa [2] — głównie w II części [3] o zasadach modelowania zjawisk optycznych. Trzecia część [4] pracy stanowi porównanie znanych do tej pory modeli laserów złączowych.

2. ROZŁAD POTENCJAŁU W WARSTWACH OGRANICZAJĄCYCH

W większości modeli laserów (np. [5—9]), zjawiska elektryczne modelowano w prosty sposób poprzez przyjęcie zwykłej diodowej charakterystyki prądowo-napięciowej (także dla pracy powyżej progu akcji laserowej):

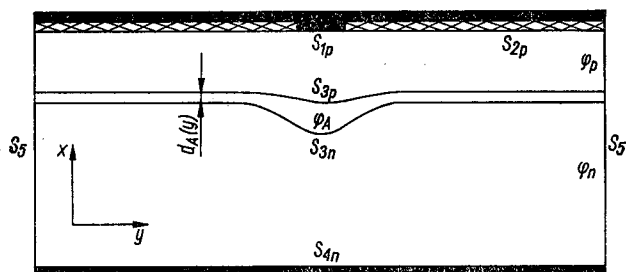
$$j = j_s \exp\left(\frac{eU}{n_s k_B T}\right) \quad (1)$$

gdzie j jest gęstością prądu, e — ładunkiem elementarnym, U — spadkiem napięcia na złączu p - n , k_B — stałą Boltzmann, T — temperaturą w skali bezwzględnej, a j_s i n_s są parametrami danej konstrukcji (n_s może jeszcze zależeć od wartości prądu zasilania).

Zastosowanie związku (1) do modelowania zjawisk elektrycznych w nowoczesnej (a więc paskowej) strukturze lasera złączowego powoduje zaniedbanie dwóch istotnych zjawisk. Pierwszym z nich jest rozptył prądu w warstwie między paskowym kontaktem i obszarem czynnym oraz dyfuzja nośników w płaszczyźnie równoległej do złącza p - n w tym obszarze, drugim — nasycenie koncentracji nośników (a więc i spadku napięcia na złączu) związanej z osiągnięciem przez nią wartości progowej akcji laserowej.

Znacznie bardziej poprawnym założeniem jest przyjęcie dla lasera biheterozłączowego ciągłości quasi-poziomów Fermiego na obu heterozłączach [1]. Założenie to w sposób naturalny uwzględnia nasycenie spadku napięcia na złączu dla progu akcji laserowej, wymaga jednakże znacznie dokładniejszej analizy zjawisk elektrycznych w laserze.

Strukturę typowego biheterozłączowego lasera paskowego pokazano schematycznie na rys. 1.



Rys. 1. Struktura typowego lasera biheterozłączowego z kontaktem paskowym

Modelowanie zjawisk fizycznych w takiej strukturze można podzielić na cztery powiązane wzajemnie zagadnienia: dwa problemy rozptyłu prądu w warstwach ograniczających oraz dwa zagadnienia ciągłości w obszarze czynnym.

Rozkłady potencjałów φ_p i φ_n w warstwach ograniczających typu p i n można znaleźć za pomocą równań Laplace'a [1]:

$$\nabla \cdot (\sigma_p \nabla \varphi_p) = 0 \quad (2)$$

$$\nabla \cdot (\sigma_n \nabla \varphi_n) = 0, \quad (3)$$

gdzie σ_p i σ_n są przewodnościami elektrycznymi powyższych obszarów. Równania te należy uzupełnić o odpowiednie warunki brzegowe. Przyjmujemy więc, że kontakty są omowe i ekwipotencjalne:

$$\varphi_p|_{S_{1p}} = \varphi_{p0} \quad (4)$$

$$\varphi_n|_{S_{4n}} = \varphi_{n0} \quad (5)$$

pomijamy przepływ prądu przez powierzchnię górną poza kontaktem paskowym:

$$\left. \frac{\partial \varphi_p}{\partial x} \right|_{S_{2p}} = 0 \quad (6)$$

i przez boczne ścianki przyrządu:

$$\left. \frac{\partial \varphi_n}{\partial y} \right|_{S_5} = 0 \quad (7)$$

oraz zakładamy ciągłość potencjału na krawędziach obszaru czynnego:

$$\varphi_p|_{S_{3p}} = \Phi_p(y) \quad (8)$$

$$\varphi_n|_{S_{3n}} = \Phi_n(y). \quad (9)$$

Rozwiązując układ równań (2—9), otrzymujemy rozkłady potencjału elektrostatycznego $\varphi_p(x, y)$ i $\varphi_n(x, y)$ wewnątrz odpowiednich warstw ograniczających, będące funkcjami brzegowych rozkładów $\Phi_p(y)$ i $\Phi_n(y)$, oraz gęstości prądów dziurowego i elektronowego wstrzykiwanych do obszaru czynnego:

$$j_{h, \text{inj}}(y) = -\sigma_p \left. \frac{\partial \varphi_p}{\partial x} \right|_{S_{3p}} \quad (10)$$

$$j_{e, \text{inj}}(y) = -\sigma_n \left. \frac{\partial \varphi_n}{\partial x} \right|_{S_{3n}}. \quad (11)$$

3. ROZKŁADY KONCENTRACJI NOŚNIKÓW W OBSZARZE CZYNNYM

Model pasmowy lasera biheterozłączonego pokazano na rys. 2. Quasi-poziomy Fermiego przedstawiono na nim w postaci linii prostych w wyniku przyjęcia założenia, że obszar czynny jest cienki w porównaniu z długością drogi dyfuzji:

$$d_A \ll L_D \quad (12)$$

Wówczas równanie Poissona dla potencjału φ_A wewnątrz obszaru czynnego można zapisać następująco [1]:

$$\nabla^2 \varphi_A(y) = \frac{\varrho_A(y)}{\varepsilon \varepsilon_0}, \quad (13)$$

gdzie ϱ_A jest gęstością ładunku w obszarze czynnym, a ε_0 i ε są przenikalnościami dielektrycznymi odpowiednio próżni i względna.

Typowa długość drogi Debye'a jest dla laserów złączowych równa około 100 Å; z warunku quasineutralności otrzymujemy więc:

$$p(y) + N_D^+(y) = n(y) + N_A^-(y), \quad (14)$$

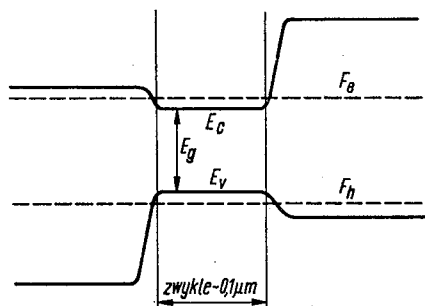
gdzie N_D^+ i N_A^- są koncentracjami zjonizowanych donorów i akceptorów, a koncentracje dziur i elektronów w obszarze czynnym, tj. p i n , można z wiązać z quasi-poziomami Fermiego F_h i F_e następująco:

$$p(y) = \frac{2}{\sqrt{\pi}} N_v F_{1/2} \left(\frac{E_v - F_h(y)}{k_B T} \right) \quad (15)$$

$$n(y) = \frac{2}{\sqrt{\pi}} N_c F_{1/2} \left(\frac{F_e(y) - E_c}{k_B T} \right) \quad (16)$$

$$E_c - E_v = E_g \quad (17)$$

$$F_e(y) - F_h(y) = e[\Phi_p(y) - \Phi_n(y)] \quad (18)$$



Rys. 2. Model pasmowy lasera biheterozłączonego spolaryzowanego wysokim napięciem w kierunku przewodzenia, E_c i E_v — krawędzie odpowiednio pasma przewodnictwa i pasma walencyjnego, E_g — szerokość przerwy energetycznej, wszystkie trzy wielkości dotyczą materiału obszaru czynnego, F_e i F_h — quasi-poziomy Fermiego odpowiednio elektronów w pasmie przewodnictwa i dziur w pasmie walencyjnym

W powyższych związkach E_c i E_v są krawędziami odpowiednio pasma przewodnictwa i pasma walencyjnego, E_g jest szerokością przerwy energetycznej w obszarze czynnym, F_e i F_h są quasi-poziomymi Fermiego odpowiednio elektronów w pasmie przewodnictwa i dziur w pasmie walencyjnym, $F_{1/2}$ jest funkcją Fermiego:

$$F_{1/2}(\eta_f) = \int_0^{\infty} \frac{\sqrt{\eta} d\eta}{1 + \exp(\eta - \eta_f)} \quad (19)$$

N_c i N_v są efektywnymi gęstościami stanów w pasmie przewodnictwa i pasmie walencyjnym:

$$N_c = \frac{2}{h^3} (2\pi m_e k_B T)^{3/2} \quad (20)$$

$$N_v = \frac{2}{h^3} (2\pi m_h k_B T)^{3/2} \quad (21)$$

a m_e i m_h są masami efektywnymi elektronu w pasmie przewodnictwa i dziury w pasmie walencyjnym. Powyższe równania pozwalają wyznaczyć rozkłady koncentracji elektronów i dziur w obszarze czynnym w funkcji skokowej zmiany potencjału na jego krawędziach, tj. $\Phi_p(y) - \Phi_n(y)$.

Rozkłady gęstości prądów wstrzykniętych do obszaru czynnego można związać z rozkładem potencjału za pomocą równań ciągłości:

$$\frac{dp(y)}{dt} = G_h(y) - R_h(y) - \frac{1}{e} \nabla \cdot \vec{J}_h(y) \quad (22)$$

$$\frac{dn(y)}{dt} = G_e(y) - R_e(y) + \frac{1}{e} \nabla \cdot \vec{J}_e(y) \quad (23)$$

gdzie G_h i G_e są szybkościami generacji dziur i elektronów, R_h i R_e są odpowiednimi szybkościami rekombinacji, a J_h i J_e są gęstościami odpowiednio prądu dziurowego i prądu elektronowego (unoszenie i dyfuzja), płynących w obszarze czynnym w kierunku równoległym do płaszczyzny złącza.

4. GENERACJA I REKOMBINACJA NOŚNIKÓW W OBSZARZE CZYNNYM

Pomijając znikomą generację cieplną nośników, człony generacyjne występujące w równaniach (22) i (23) można łączyć wprost z gęstościami prądów dziur i elektronów wstrzykniętych do obszaru czynnego:

$$G_h(y) = \frac{1}{ed_A(y)} j_{h, \text{inj}}(y) \quad (24)$$

$$G_e(y) = \frac{1}{ed_A(y)} j_{e, \text{inj}}(y) \quad (25)$$

gdzie d_A jest grubością warstwy czynnej.

W członach rekombinacyjnych należy uwzględnić zarówno rekombinację niepromienistą — tj. rekombinację związaną z obecnością pułapek (indeks: trap) i rekombinację powierzchniową (surface), jak i rekombinację promienistą, w tym spontaniczną (spont) i stymulowaną (stim):

$$R_{h, \text{trap}} = R_{e, \text{trap}} = \frac{pn}{p\tau_e + n\tau_h} \quad (26)$$

$$R_{h, \text{surface}} = R_{e, \text{surface}} = \frac{S' \delta(y - y_s) pn}{p + n} \quad (27)$$

$$R_{h, \text{spont}} = R_{e, \text{spont}} = B_r pn = \frac{n}{\tau_r} \quad (28)$$

$$R_{h, \text{stim}} = R_{e, \text{stim}} = \sum_i \frac{P_i g_i}{\hbar \omega_i} \quad (29)$$

W powyższych zależnościach τ_e i τ_h są efektywnymi czasami życia nośników mniejszościowych (odpowiednio elektronów i dziur) związanymi z rekombinacją niepromienistą (można przy ich wyznaczaniu uwzględnić ewentualnie ponadbarierową ucieczkę nośników z obszaru czynnego), S' jest szybkością rekombinacji powierzchniowej, y_s jest współrzędną powierzchni, B_r jest stałą rekombinacji promienistej, τ_r jest czasem życia nośników związanym z rekombinacją promienistą, P_i jest gęstością mocy optycznej wewnątrz obszaru czynnego, g_i jest lokalnym wzmocnieniem promieniowania, a $\hbar \omega_i$ jest energią fotonów, z tym, że ostatnie trzy wielkości odnoszą się do podłużnego modu promieniowania rzędu i .

Wygodnym przybliżeniem, zgodnym z wynikami teoretycznymi [10], jest przyjęcie założenia, że wzmacnienie lokalne dla danej długości fali promieniowania jest liniową funkcją koncentracji nośników [11, 12]:

$$g_i(n) \equiv g(n, \lambda_i) = a_i n(y) - b_i, \quad (30)$$

gdzie λ_i jest długością fali rozpatrywanego modu promieniowania, a a_i i b_i są stałymi.

5. ROZPŁYW PRĄDU

Przyjmując, że grubość obszaru czynnego $d_A(y)$ zmienia się stosunkowo wolno, można z zasady zachowania ładunku otrzymać jawną postać ostatnich wyrazów w równaniach (22) i (23) [1]:

$$-\frac{1}{e} \nabla \cdot \vec{J}_h = -\frac{1}{e} \left(\frac{1}{d_A(y)} \frac{d[d_A(y)]}{dy} + \frac{d}{dy} \right) \left(p^{\mu_h} \frac{dF_h}{dy} \right) \quad (31)$$

$$-\frac{1}{e} \nabla \cdot \vec{J}_e = \frac{1}{e} \left(\frac{1}{d_A(y)} \frac{d[d_A(y)]}{dy} + \frac{d}{dy} \right) \left(n^{\mu_e} \frac{dF_e}{dy} \right), \quad (32)$$

gdzie μ_h i μ_e są ruchliwościami odpowiednio dziur i elektronów. Korzystając ze związków oczywistych dla półprzewodnika niezdegenerowanego:

$$\frac{dp(y)}{dy} = \frac{p(y)}{k_B T} \frac{dF_h(y)}{dy} \quad (33)$$

$$\frac{dn(y)}{dy} = \frac{n(y)}{k_B T} \frac{dF_e(y)}{dy} \quad (34)$$

oraz wzorów Einsteina:

$$k_B T = \frac{eD_e}{\mu_e} = \frac{eD_h}{\mu_h} \quad (35)$$

można zależności (31) i (32) przekształcić do postaci [2]:

$$-\frac{1}{e} \nabla \cdot \vec{J}_h = -D_h \left(\frac{\partial^2 p}{\partial y^2} + \frac{1}{d_A(y)} \frac{d[d_A(y)]}{dy} \frac{dp}{dy} \right), \quad (36)$$

$$\frac{1}{e} \nabla \cdot \vec{J}_e = D_e \left(\frac{\partial^2 n}{\partial y^2} + \frac{1}{d_A(y)} \frac{d[d_A(y)]}{dy} \frac{dn}{dy} \right), \quad (37)$$

gdzie D_e i D_h są stałymi dyfuzji odpowiednio elektronów i dziur.

Powyższą analizę uzupełnia wzór na całkowity prąd I_T płynący przez diodę. Można go obliczyć całkując gęstości prądów wstrzykiwanych do obszaru czynnego oraz (ewentualnie) gęstości prądów wynikających z ponadbarierowej ucieczki elektronów i dziur:

$$I_T = 2L \int_0^{w/2} [j_{h, \text{inj}}(y) + j_{e, \text{inj}}(y) + j_N(y) + j_P(y)] dy, \quad (38)$$

gdzie L i W są długością i szerokością kryształu diody, a j_N i j_P — gęstościami prądów ucieczki ponad odpowiednio barierą po stronie N i barierą po stronie P .

BIBLIOGRAFIA

1. D. P. Wilt, A. Yariv: *A self-consistent static model of the double-heterostructure laser*. IEEE J. Quantum Electron., QE-17 (1981), pp. 1941—1949
2. J. Buus: *Principles of semiconductor laser modelling*. IEE Proc., (1985) 132, Pt. J, pp. 42—51
3. W. Nakwaski: *Modelowanie zjawisk fizycznych w laserach złączowych, II. Modelowanie zjawisk optycznych*. Rozprawy Elektrotechniczne, 1988, t. 34, z. 3 ss. 987-992.
4. W. Nakwaski: *Modelowanie zjawisk fizycznych w laserach złączowych, III. Porównanie modeli*. Rozprawy Elektrotechniczne, 1988, t. 34, z. 3 ss. 992-1002.
5. J. Buus: *A model for the static properties of DH lasers*. IEEE J. Quantum Electron., (1979) QE-15 pp. 734—749
6. K. A. Shore, T. E. Rozzi, G. H. in't Veld: *Semiconductor laser analysis: General method for characterising devices of various cross sectional geometries*. IEE Proc., (1980) 127, Pt. I, 221—229
7. N. Chinone: *Nonlinearity in power-output-current characteristics of stripe-geometry injection lasers*. J. Appl. Phys., (1977) 48 3237—3243
8. R. Lang: *Lateral transverse mode instability and its stabilization in stripe-geometry injection lasers*. IEEE J. Quantum Electron., (1979) QE-15 pp. 718—726
9. W. Streifer, D. R. Scifres, R. D. Burnham: *Above threshold analysis of double heterostructure lasers with laterally tapered active regions*. Appl. Phys. Lett., (1980) 37, pp. 877—879
10. F. Stern: *Calculated spectral dependence of gain in excited GaAs*, J. Appl. Phys., (1976) 47 5382—5386
11. W. Streifer, D. R. Scifres, R. D. Burnham: *Diode laser threshold current density and lasing wavelength as function of active region thickness*. Appl. Phys. Lett., (1983), 42 pp. 401—403
12. J. Machač: *Static model of DH laser*. IEE Proc., (1983) 130, Pt. I, pp. 61—66

W. NAKWASKI

MODELLING OF PHYSICAL PHENOMENA IN INJECTION LASERS
Part I. ELECTRICAL PHENOMENA MODELLING

Summary

A complete model of electrical phenomena in a double-heterostructure stripe laser is presented. The potential distribution inside the laser area as well as the distribution of the electron and hole concentrations in the active layer are determined.

W. NAKWASKI

MODÈLAGE DES PHÉNOMÈNES PHYSIQUES DANS LES LASERS CONJUGES
Partie I. MODÈLAGE DES PHÉNOMÈNES ÉLECTRIQUES

Résumé

Dans l'étude on a présenté le modèle complet des phénomènes électriques dans un laser longitudinal bihétéroconjugué. Dans le modèle on a déterminé les répartitions du potentiel dans la zone du laser et les répartitions de la concentration des électrons et des trous dans la zone active.

W. NAKWASKI

MODELLIEREN DER PHYSIKALISCHEN ERSCHEINUNGEN IN DIODENLASERN
Teil I. MODELLIEREN DER ELEKTRISCHEN ERSCHEINUNGEN

Zusammenfassung

In dem Aufsatz wird ein vollständiges Modell der elektrischen Erscheinungen in doppel-heterostrukturellen Streifendiodelnlasern geschildert. Bei dem Modell werden Potentialverteilungen im Laserraum und Konzentrationsverteilungen von Elektronen und Löchern im aktiven Raum bestimmt.

В. НАКВАСКИ

МОДЕЛИРОВАНИЕ ФИЗИЧЕСКИХ ЯВЛЕНИЙ В ЛАЗЕРНЫХ ДИОДАХ
Часть 1. МОДЕЛИРОВАНИЕ ЭЛЕКТРИЧЕСКИХ ЯВЛЕНИЙ**Резюме**

Представлена полная модель электрических явлений в полосковом лазерном диоде с двойной гетероструктурой. В модели определено распределение потенциала в лазерной области а также распределения концентрации электронов и дырок в активной области.

Modelowanie zjawisk fizycznych w laserach złączowych

WŁODZIMIERZ NAKWASKI

Instytut Fizyki, Politechnika Łódzka

Otrzymano 1987.01.27

Autoryzowano do druku 1987.11.27

CZĘŚĆ II — MODELOWANIE ZJAWISK OPTYCZNYCH

Równania opisujące rozkład pola promieniowania i rozkłady nośników w obszarze lasera złączowego są wzajemnie sprzężone, co uniemożliwia ich separację oraz rozwiązanie analityczne bez dodatkowych założeń upraszczających, wpływających jednakże na poprawność wyników. W I części niniejszej pracy przedstawiono zasady modelowania zjawisk elektrycznych biheterozłączowym laserze paskowym (rys. 1). Przedmiotem części II będzie analiza zjawisk optycznych wewnątrz obszaru czynnego i sprzężenie ich ze zjawiskami elektrycznymi.

1. RÓWNANIE KINETYCZNE DLA GĘSTOŚCI FOTONÓW PROMIENIOWANIA

Dla pola elektromagnetycznego, analogicznym do równań (22) i (23) w pierwszej części, jest równanie kinetyczne dla gęstości fotonów S_{ij} skupionych w odpowiednich modach promieniowania S_{ij} w postaci [3]:

$$\frac{dS_{ij}}{dt} = \frac{c}{\bar{n}_R} (G_{ij} - \alpha_s - \alpha_{\text{end}}) S_{ij} + \beta_i R_j \quad (1)$$

gdzie i oraz j są numerami odpowiednio podłużnych i bocznych modów promieniowania, c jest prędkością światła w próżni, $\bar{n}_R$ jest grupowym współczynnikiem załamania [4]:

$$\bar{n}_R = n_{R,A} - \lambda \frac{dn_{R,A}}{d\lambda} \quad (2)$$

$n_{R,A}$ jest współczynnikiem załamania materiału obszaru czynnego, λ jest długością fali promieniowania, G_{ij} jest wzmocnieniem modowym (patrz równanie (13)), α_s reprezentuje straty wewnętrzne spowodowane przez rozpraszanie promieniowania laserowego na nie-

doskonałościach materiału warstwy czynnej oraz zaburzeniach jej grubości, a α_{end} — straty użyteczne związane z emisją części promieniowania przez zwierciadła rezonatora:

$$\alpha_{\text{end}} = \frac{1}{L} \ln \frac{1}{\sqrt{R_1 R_2}}. \quad (3)$$

Przez R_1 i R_2 oznaczono współczynniki odbicia od zwierciadeł rezonatora, które w ogólnym przypadku mogą być różne dla różnych modów bocznych, β_i jest współczynnikiem emisji spontanicznej, przy czym zakłada się, że zależy on jedynie od numeru modu podłużnego i , a R_j jest średnią szybkością emisji spontanicznej (porównaj równanie (10)), zależną z kolei od rozkładu natężenia promieniowania w modzie bocznym, czyli będącą funkcją jego numeru j .

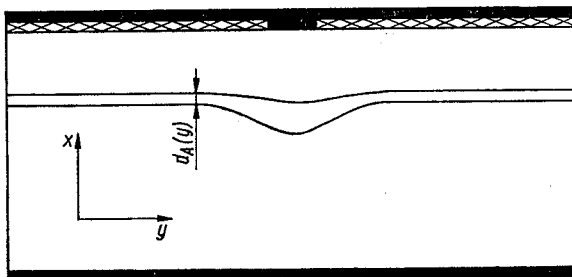
Równanie (1) sformułowano przy założeniu, że zmiany natężenia promieniowania wzdłuż podłużnej osi lasera nie wpływają na jego działanie. Jednakże w przypadku wysokich mocy promieniowania lub niskich współczynników odbicia od zwierciadeł należy dodatkowo uwzględnić efekt nasycenia wzmocnienia [5—7].

2. RÓWNANIE POLA ELEKTROMAGNETYCZNEGO

Zakładając, że rozkład pola promieniowania wzdłuż osi y jest zależny jedynie od rzędu modu bocznego, a nie zależy od rzędu modu podłużnego, można zagadnienie własne dla pola sformułować w postaci następującego równania skalarnego [3]:

$$\frac{\partial^2 E_j(y, t)}{\partial y^2} + [N_{R, \text{eff}}^2(y, t) k^2 - \gamma_j^2(t)] E_j(y, t) = 0, \quad (4)$$

gdzie γ_j jest zespoloną stałą propagacji i gdzie skorzystano z metody efektywnego współczynnika załamania opisanej w pracy [8].



Struktura typowego biheterozłączonego lasera z kontaktem paskowym

Efektywny zespolony współczynnik załamania można dla struktury pokazanej na Rys. 1 przedstawić następująco [3]:

$$N_{R,\text{eff}}^2(y, t) \equiv \varepsilon_{\text{eff}}(y, t) = n_{R,\text{eff}}^2(y) + 2\Gamma(y)n_{R,A}n(y, t) \frac{\partial n_{R,A}}{\partial n} + \\ + 2n_{R,\text{eff}}(y)\Delta n_R(y, t) + \frac{i}{k} \{ \Gamma(y)n_{R,A}[g(n) - \alpha_A(n)] - [1 - \Gamma(y)]n_{R,P}\alpha_P \} \quad (5)$$

gdzie i jest jednostką urojoną, ε_{eff} — zespoloną, efektywną stałą dielektryczną, k — liczbą falową, $k = 2\pi/\lambda$, Γ jest współczynnikiem wypełnienia, $n_{R,A}$ i $n_{R,P}$ są współczynnikami załamania odpowiednio materiału obszaru czynnego i materiału warstw ograniczających, a α_A i α_P są współczynnikami strat w powyższych obszarach.

Efektywny współczynnik załamania można wyznaczyć następująco [8]:

$$n_{R,\text{eff}}^R = bn_{R,A}^2 + (1-b)n_{R,P}^2, \quad (6)$$

gdzie znormalizowana stała propagacji b jest rozwiązaniem równania [9]:

$$v\sqrt{1-b} = \arctan \sqrt{\frac{b}{1-b}} \quad (7)$$

ze znormalizowaną częstotliwością v daną równaniem [8]:

$$v = \frac{1}{2} kd_A(y) \sqrt{n_{R,A}^2 - n_{R,P}^2}. \quad (8)$$

Dla struktur o szybko zmieniającej się grubości obszaru czynnego powyższe postępowanie przestaje być poprawne. Wówczas należy stosować zależności wyprowadzone specjalnie dla tego przypadku w pracy [10].

W równaniu (5) pierwszy wyraz po prawej stronie stanowi niezaburzoną, efektywną stałą dielektryczną, drugi wyraz ilustruje wpływ obecności nośników swobodnych o koncentracji n [11–18], trzeci z kolei opisuje wpływ pozostałych czynników (temperatura [19–22], efekt fotoelastyczny [23] itd.), a część urojona obrazuje wpływ wzmocnienia i strat promieniowania.

Wprowadźmy znormalizowane natężenie promieniowania $I_j(y, t)$:

$$I_j(y, t) = \frac{|E_j(y, t)|^2}{\frac{1}{W} \int_{-\infty}^{\infty} |E_j(y, t)|^2 dy} \quad (9)$$

a stąd szybkość emisji spontanicznej $R_j(t)$:

$$R_j(t) = \frac{1}{W} \int_{-\infty}^{\infty} \frac{n(y, t)}{\tau_r(n)} I_j(y, t) dy, \quad (10)$$

gdzie przez W oznaczono szerokość krysztalu diody, a przez τ_r czas życia nośników mniejszościowych ze względu na rekombinację spontaniczną.

Zależność (9) umożliwia przedstawienie szybkości rekombinacji stymulowanej R_{stim} w następującej postaci [3] por. wzór (29) w [1]:

$$R_{stim} = \frac{c}{n_R} \Gamma(y) \frac{A}{W d_A(y)} \sum_{i,j} g(n, \lambda_i) I_j(y, t) S_{ij}(t), \quad (11)$$

gdzie A jest powierzchnią przekroju poprzecznego obszaru czynnego, d_A jest jego grubością, g jest wzmocnieniem lokalnym, a λ_i jest długością fali promieniowania skupionego w podłużnym modzie i . W powyższym wzorze uwzględniono zależność lokalnego wzmocnienia od numeru modu podłużnego, pominięto natomiast spektralny rozkład koncentracji nośników, czyli efekt częstotliwościowego wypalania dziury w rozkładzie wzmocnienia. Przybliżenie to jest prawdopodobnie uzasadnione, choć jego prawdziwość jest wciąż przedmiotem rozważań [24–29]. Zależność (11) sformułowano jedynie dla modów TE promieniowania. Uwzględnienie modów TM stworzyłoby konieczność dodania podobnego drugiego wyrazu, z tym, że wystąpiłaby w nim inna wartość współczynnika wypełnienia.

3. WZMOCNIENIE MODOWE I MOC EMITOWANEGO PROMIENIOWANIA

Wzmocnienie modowe G_j modu bocznego rzędu j można wyznaczyć z urojonej części stałej propagacji:

$$G_j(t) = 2Im[\gamma_j(t)] \quad (12)$$

Dokładniejsza analiza własności falowodowych wymaga jednakże określenia wzmocnienia modowego dla każdego modu podłużnego (indeks i). Odpowiednia zależność jest wówczas następującej postaci [3]:

$$G_{ij}(t) = Re \left\{ \frac{k}{\gamma_j(t)} \right\} \frac{1}{W} \int_{-\infty}^{\infty} \{ \Gamma(y) n_{R,A} [g(n, \lambda_i) - \alpha_A(n)] - [1 - \Gamma(y)] n_{R,P} \alpha_P \} I_j(y, t) dy, \quad (13)$$

w której uwzględniono zależność lokalnego wzmocnienia od energii fotonu. Dla struktury planarnej zależność (13) redukuje się do prostszego wzoru przybliżonego:

$$G_{ij}(t) \approx \frac{\Gamma}{W} \int_{-\infty}^{\infty} [g(n, \lambda_i) - \alpha_A(n)] I_j(y, t) dy - (1 - \Gamma) \alpha_P. \quad (14)$$

Całkowitą moc promieniowania skupionego w i -tym modzie podłużnym i j -tym modzie bocznym można wyznaczyć następująco:

$$P_{ij}(t) = LAS_{ij}(t) \hbar \omega_i \frac{c}{n_R} \alpha_{end}. \quad (15)$$

Moc promieniowania emitowanego przez poszczególne zwierciadła rezonatora jest wówczas równa [30]:

$$P_{ij,1}(t) = P_{ij}(t)(1 + \xi)^{-1} \quad (16)$$

$$P_{ij,2}(t) = P_{ij}(t)(1 + \xi^{-1})^{-1}, \quad (17)$$

BIBLIOGRAFIA

1. W. Nakwaski, *Modelowanie zjawisk fizycznych w laserach złączowych. I. Modelowanie zjawisk elektrycznych*. Rozprawy Elektrotechniczne, 1988, t. 34 z. 3 981—988
2. W. Nakwaski, *Modelowanie zjawisk fizycznych w laserach złączowych. III. Porównanie modeli*. Rozprawy Elektrotechniczne, 1988, t. 34, z. 3
3. J. Buus: *Principles of semiconductor laser modelling*. IEE Proc., Pt. J 1985 132. 42—51
4. J. Buus, M. J. Adams: *Phase and group indices for double heterostructure klasers*. IEE J., Solid-State and Electron. Devices, 1979, 3 pp. 189—195
5. S. Hasuo, T. Ohmi: *Spatial distribution of the light intensity in the injection lasers*. Jpn. J. Appl. Phys., 1974 13 1429—1434
6. J. E. A. Whiteaway, G. H. B. Thompson: *Optimisation of power efficiency of (GaAl)As injection lasers operating at high power levels*. IEE J., Solid-State and Electron Devices, 1977 1 pp. 81—88
7. R. Baets, J.-P. Van de Capelle, P. E. Lagasse: *Longitudinal abalysis of semiconductor laser of with low reflectivity facets*. IEEE J., Quantum Electron., 1985 QE-21, pp. 693—699
8. J. Buus: *The effective index method and its application to semiconducto lasers*. IEEE J. Quantum Electron., QE-18 (1982) 1083—1089
9. W. O. Schlosser: *Gain induced modes in planar structures*. Bell Syst. Tech. J., 1973 52, pp. 887—908
10. J. Buus: *Application of the effective index method to nonplanar structures*. IEEE J. Quantum Electron., 1984, QE-20 pp. 1106—1109
11. M. R. Matthews, R. B. Dyott, W. P. Carling: *Filaments as optical waveguides in gallium — arsenide lasers*. Electron. Lett., 1972, 8 pp. 570—572
12. G. H. B. Thompson: *A theory for filamentation in semiconductor lasers including the dependence of dielectric constant on injected carrier density*. Opto-electronics, 1972 4 pp. 257—310
13. P. R. Selway, G. H. B. Thompson, G. D. Henshall, J. E. A. Whiteway: *Measurement of the effect of injected carriers on the p-n refractive-index step in single heterostructure diode lasers*. Electron. Lett., 1974 10 pp. 453—454
14. M. Ito, T. Kimura: *Carrier density dependence of refractive index in AlGaAs semiconductor lasers*. IEEE J. Quantum Electron., 1980 QE-16, pp. 910—911
15. J. G. Mendoza-Alvarez, F. D. Nunes, N. B. Patel: *Refractive index dependence on free carriers for GaAs*. J. Appl. Phys. 1980 51 pp. 4365—4367
16. J. S. Manning, R. Olshansky: *Carrier-induced index change in AlGaAs double-heterostructure lasers*. Electron. Lett., 1981 17, pp. 506—507
17. K. Kobayashi, H. Nishimoto, R. Lang: *Experimental observation of asymmteric detuning characteristics in semiconductor laser injection locking*. Electron. Lett., 1982 18, pp. 54—56
18. J. Manning, R. Olshansky, C. B. Su: *The carrier-induced index change in AlGaAs and 1.3 μ m InGaAsP diode lasers*. IEEE J. Quantum Electron., 1983 QE-19, pp. 1525—1530
19. M. Cardona: *Temperature dependence of the refractive index and the polarizability of free carriers in some III—V semiconductors*. Proc. Intern. Conf. Semicond. Physics, Prague 1960, 388—394
20. D. T. F. Marple: *Refractive index of GaAs*. J. Appl. Phys., 1964 35 pp. 1241—1242
21. I. Strzalkowski, S. Joshi, C. R. Crowell: *Dielectric constant and its temperature dependence for GaAs, CdTe, and ZnSe*. Appl. Phys. Lett., 1976 28, pp. 350—352
22. G. A. Samara: *Temperature and pressure dependences of the dielectric constants od semiconductors*. Phys. Rev. B, 27 1983 3494—3505
23. P. A. Kirkby, P. R. Selway, L. D. Westbrook: *Photoelastic waveguides and their effect on stripe geometry GaAa/Ga₁ Al As lasers*. J. Appl. Phys., 1979 50 pp. 4567—4579
24. M. Nakamura, K. Aiki, N. Chinone, R. Ito, J. Umeda: *Longitudinal-mode behaviours of mode-stabilised Al Ga_{1-x} As injection lasers*. J. Appl. Phys., 1978 49, pp. 4644—4648
25. B. Zee: *Broadening mechaning mechanism in semiconductor (GaAs) lasers: limitation to single mode power*. IEEE J. Quantum Electron., 1978, QE-14, 727—736

26. M. W. Flemming, A. Mooradian: *Spectral characterisation of external-cavity controlled semiconductor lasers*. IEEE J. Quantum Electron., 1981, QE-17, pp. 44—59
27. M. Yamada, Y. Suematsu: *Analysis of gain suppression in undoped injection lasers*. J. Appl. Phys., 1981, 52 pp. 2653—2664
28. P. Brosseau, W. W. Rühle, N. B. Patel, J. E. Ripper: *Optical coupling of two injection lasers: a new experimental approach to study the gain broadening mechanism*. IEEE J. Quantum Electron., 1981 QE-17, pp. 714—717
29. R. F. Kazarinov, C. H. Henry, R. A. Logan: *Longitudinal mode self-stabilisation in semiconductor lasers*. J. Appl. Phys., 1982 53, pp. 4631—4644
30. M. Ettenberg, H. S. Sommers, Jr., H. Kressel, H. F. Lockwood: *Control of facet damage in GaAs laser diodes*. Appl. Phys. Lett., 1971 18 pp. 571—573

Part II. OPTICAL PHENOMENA MODELLING

A complete model of optical phenomena in a double-heterostructure stripe laser is presented.

Partie II. MODÈLAGE DES PHÉNOMÈNES OPTIQUES

Dans l'étude on a présenté le modèle complet des phénomènes optiques dans le laser longitudinal bihétéroconjugué.

Teil II. MODELLIEREN DER OPTISCHEN ERSCHEINUNGEN

In dem Aufsatz wird das vollständige Modell der optischen Erscheinungen in doppel-heterostrutturellen Streifendiodenlasern geschildert.

Часть 2. МОДЕЛИРОВАНИЕ ОПТИЧЕСКИХ ЯВЛЕНИЙ

Представлена полная модель оптических явлений в полосковом лазерном диоде с двойной гетероструктурой.

Modelowanie zjawisk fizycznych w laserach złączowych

WŁODZIMIERZ NAKWASKI

Instytut Fizyki, Politechnika Łódzka

Otrzymano 1987.01.27

Autoryzowano do druku 1987.11.27

CZĘŚĆ III — PORÓWNANIE MODELI

Przedstawiono różne procedury obliczeń stosowanych przy rozwiązywaniu równań modelujących zjawiska fizyczne zachodzące w biheterozłączowym laserze paskowym oraz dokonano porównania znanych dotychczas modeli laserów złączowych różnych konstrukcji.

1. WPROWADZENIE

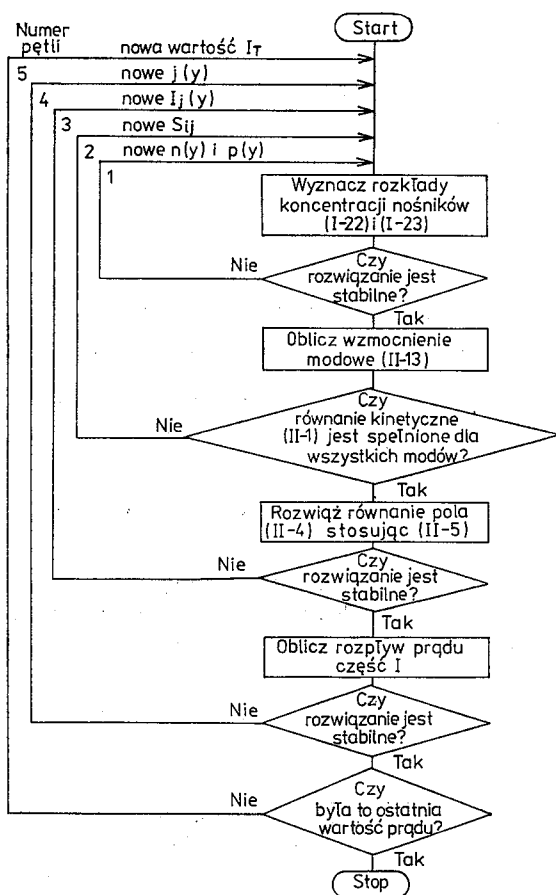
Sformułowanie w części I [1] i II [2] niniejszej pracy kompletnego układu równań opisujących wzajemne oddziaływanie oba pola promieniowania w rezonatorze paskowego lasera heterozłączowego oraz pola nośników w obszarze czynnym tego lasera stanowi dopiero połowę pracy. Układ ten jest zbyt skomplikowany, aby spodziewać się otrzymania rozwiązania analitycznego bez wprowadzania dodatkowych założeń upraszczających, ograniczających w rezultacie dokładność otrzymanych wyników, a nawet wręcz ich poprawność. Pozostają więc metody numeryczne.

W niniejszej pracy przedstawiono różne procedury obliczeń numerycznych stosowanych przy rozwiązywaniu równań modelujących zjawiska fizyczne zachodzące w biheterozłączowym laserze paskowym oraz porównano znane do tej pory modele laserów złączowych, a więc realizację zasad modelowania dla konkretnych konstrukcji laserów.

2. PROCEDURA OBLICZEŃ

2.1. MODELE STATYCZNE

Numeryczne rozwiązanie kompletnego układu równań jest obecnie praktycznie niemożliwe. Nawet ograniczając się do modelu statycznego, w którym pochodne po lewej stronie równań kinetycznych (22) i (23) w [1] oraz (1) w [2] są równo zero, wzajemne powiązanie poszczególnych równań powoduje konieczność stosowania metod iteracyjnych wymagających zwykle stosunkowo długiego czasu obliczeń.



Rys. 1 Schemat blokowy obliczeń w modelu paskowego lasera biheterozłączonego [3]. W punkcie startu zakładamy wstępne rozkłady koncentracji nośników i natężenia promieniowania. I lub II przed numerem wzoru oznacza, w której części pracy znajduje się dana zależność. Przez rozwiązanie stabilne rozumiemy rozwiązanie, które z założoną tolerancją nie zmienia się przez dwie kolejne iteracje.

Jedną z możliwych struktur obliczeń, zawierającą kilka pętli obliczeń iteracyjnych, przedstawiono na rys. 1. W strukturze tej obliczenia wymagające najwięcej czasu są realizowane w pętlach zewnętrznych, aby zminimalizować sumaryczny czas obliczeń. Tym niemniej należy pamiętać, że wielokrotne iteracje mogą być źródłem numerycznych niestabilności. Poza tym ta struktura nie uwzględnia zmian temperatury w obszarze czynnym, które mogą mieć znaczny wpływ na przebieg zjawisk elektrycznych i optycznych, szczególnie w przypadku zasilania lasera stosunkowo wysokim prądem.

2.2. MODELOWANIE ROZPŁYWU PRĄDU

Dokładne rozwiązanie problemu rozptywu prądu między paskowym kontaktem a obszarem czynnym jest zagadnieniem bardzo skomplikowanym nawet bez uwzględniania rekombinacji stymulowanej, czyli poniżej progu akcji laserowej. Zarówno dokładne roz-

wiązania numeryczne [4, 71], jak i wypracowane rozwiązanie analityczne [5], zwykle w postaci szeregu, cechują się bardzo wolną zbieżnością.

Przyjmując, że gęstość wstrzykiwanego prądu jest stała, tj. niezależna od położenia, w części obszaru czynnego bezpośrednio pod paskowym kontaktem, natomiast w pozostałej jego części gęstość prądu jest opisana zwykłą diodową charakterystyką prądowo-napięciową

$$j = j_s \exp\left(\frac{eU}{n_s k_B T}\right), \quad (1)$$

można otrzymać przybliżone rozwiązanie analityczne [6, 7] zagadnienia rozptyłu prądu. W zależności (1) j jest gęstością prądu, e — ładunkiem elementarnym, U — spadkiem napięcia na złączu p - n , k_B — stałą Boltzmanna, T — temperaturą w skali bezwzględnej, a j_s i n_s są parametrami. Zwrotny wpływ rozkładu koncentracji nośników w obszarze czynnym na rozptyw prądu oraz ambipolarną dyfuzję nośników w tym obszarze uwzględniono w pracy [8].

Dokładniejszy model, w którym rozpatrzono nie tylko prąd dyfuzyjny w obszarze czynnym, lecz również prąd przewodzenia płynący w warstwie ograniczającej w kierunku równoległym do płaszczyzny złącza p - n , przedstawiono w pracy [9]. W modelu tym koncentracja nośników w obszarze czynnym jest opisana pojedynczym równaniem nieliniowym. Wówczas dodatkowe założenie, że promienisty czas życia i stałe dyfuzji elektronów i dziur nie zależą od ich koncentracji, oraz pominięcie rekombinacji niepromienistej w obszarze czynnym redukuje równania dyfuzji (patrz równania (22) i (23) w [1]) do prostej postaci liniowej.

Powyżej progu akcji laserowej należy w modelu lasera złączonego dodatkowo uwzględnić silny wpływ rekombinacji stymulowanej na rozkład nośników w obszarze czynnym, tj. zjawisko przestrzennego wypalania dziury w rozkładzie wzmocnienia lokalnego. Zagadnienie to było rozpatrywane w pracach [10—12].

2.3. MODELE DYNAMICZNE

W modelu dynamicznym, opisującym czasowe ewolucje rozkładów gęstości fotonów i koncentracji nośników, nowe wartości gęstości strumienia fotonów i koncentracji nośników w obszarze czynnym otrzymuje się zwykle wprost z równań kinetycznych (patrz równania (22) i (23) w [1] oraz (1) w [2]), podstawiając po ich prawych stronach wartości obliczone dla poprzedniego punktu czasowego. W ten sposób unika się wielokrotnych iteracji, a wystarczającą dokładność obliczeń uzyskuje się już dla kroku czasowego rzędu 0,5—1,0 ps. Ze względu na stosunkowo wolne zmiany rozkładu pola promieniowania można stosować te same wartości gęstości strumienia fotonów dla kilku kolejnych kroków czasowych. W przypadku laserów z samoistnym efektem falowodowym [13—15] należy być ostrożnym w ewentualnym zakładaniu symetrycznej postaci rozwiązania, skoro nawet nieznaczna asymetria struktury może mieć ogromny wpływ na wyniki obliczeń.

Kluczowym punktem analizy zjawisk optycznych w laserze jest rozwiązywanie równania pola elektromagnetycznego (patrz równanie (4) w [2]). Stosowano w tym celu różne metody obliczeń, których krytyczną analizę przedstawiono w tabeli 1. Trudno w niej nie

Tabela 1.

Metody stosowane do rozwiązania skalarnego równania falowego pola elektromagnetycznego [3]

Metoda	Stosowana w *	Szybkość	Dokładność	Złożoność	Transformacja do strefy dalekiej
Dopasowanie rozkładu Gaussa	B J	duża	zależna od struktury	bardzo mała	trywialna
Rachunek zaburzeń	A	średnia	niewielka w przypadku silnego wypalania dziury	średnia	łatwa
Wariacyjna	D	średnia	metoda niepoprawna dla falowodów zespolonych	średnia	trywialna
Różnic skończonych	E I	średnia	dobra	średnia	całkowanie numeryczne
Poprzecznego rezonansu	G	średnia	dobra	średnia	całkowanie numeryczne
Schodkowa	L	średnia	dobra	średnia	Względnie prosta
Bezpośredniego całkowania	H	mała	bardzo dobra	duża	całkowanie numeryczne
Galerkina	C F K	mała	dobra	bardzo duża	łatwa
Promieniowej propagacji (BPM)	M	mała	bardzo dobra	bardzo duża	szybka transformacja Fouriera

* Poszczególne litery stanowią oznaczenia modeli opisanych w Tabeli 2.

zauważyć silnej korelacji pomiędzy czasem obliczeń, ich dokładnością i złożonością modelu.

Wzajemne powiązanie równań opisujących własności elektryczne i optyczne lasera, a będące wynikiem wzajemnego oddziaływania pola nośników w obszarze czynnym i pola promieniowania w rezonatorze, prowadzi niekiedy do wzbudzenia tzw. samopodtrzymujących się pulsacji (*self-sustained pulsations*) [16–21]. Wówczas rozpatrywany układ równań nie ma rozwiązania statycznego.

3. PORÓWNANIE MODELI

3.1. MODELE STATYCZNE

Do tej pory opublikowano już wiele mniej lub bardziej kompletnych modeli zjawisk fizycznych w biheterozłączowych laserach paskowych. Najciekawsze z nich przedstawiono w tabeli 2.

Samouzgodnione modele lasera złączowego [3]

Autor	Równanie dyfuzji	Efekt falowodowy	Metoda rozwiązania równania falowego	Wyniki
(A) Chinone [23]	rekombinacja bimolekularna, prosty rozptyw prądu	tylko wzmocnieniowy, możliwa asymetria	rachunek zaburzeń	nieliniowa charakterystyka emisyjna,
(B) Thompson [24]	stały czas życia, uproszczony rozkład natężenia	obie części, rzeczywista i urojona, proporcjonalne do koncentracji nośników, uwzględnienie falowodu wbudowanego	dopasowanie do przechyłonego i przesuniętego rozkładu Gaussa	nieliniowa charakterystyka emisyjna, przesunięcie pola w strefie bliskiej
(C) Seki i in. [25]	stały czas życia, prosty rozptyw prądu	tylko symetryczny wzmocnienieowy	Galerkina	malenia sprawności kwantowej dla dużej mocy promieniowania
(D) Lang [26, 27]	tak jak C	tak jak B	wariancyjna	tak jak B
(E) Asbeck i in. [28]	stały czas życia, rozważany wpływ złącza na rozptyw prądu	obie części, rzeczywista i urojona, proporcjonalne do koncentracji nośników, tylko przypadek symetryczny	różnic skończonych	nieliniowa charakterystyka emisyjna, zmiany rozkładu pola w strefie dalekiej
(F) Buus [29]	tak jak C	tak jak B + uwzględnienie rozkładu temperatury, tylko przypadek symetryczny	Galerkina	nieliniowa charakterystyka emisyjna, mod boczny pierwszego rzędu
(G) Shore i in. [30–32]	tak jak A, w 31, 32 rozważany wpływ złącza na rozptyw prądu	tak jak E, ale bez wymagania symetryczności	poprzedniego rezonansu	rozwiązania asymetryczne dla struktur symetrycznych
(H) Streifer i in. [33, 34]	tak jak C + nieplanarna warstwa czynna	tak jak E + wbudowany falowód	bezpośredniego całkowania	dokładne porównanie z eksperymentami
(I) Wilt i in. [4]	dokładna analiza rozptywu prądu, nieplanarna warstwa czynna	tak jak H	różnic skończonych	charakterystyka prądowo-napięciowa, mody boczne wyższych rzędów
(J) Su [35]	stały czas życia, stały prąd	tak jak B	podobna do B	tak jak B
(K) Machač [36]	tak jak C	tak jak C	Galerkina	widmo spektralne w funkcji prądu
(L) Buus [37]	tak jak C	tak jak B	schockowa	tak jak F + przesunięcie rozkładu pola w strefie bliskiej
(M) Baets i in. [38]	tak jak C	tak jak B	promieniowej propagacji	uwzględniono zmiany rozkładów koncentracji nośników i natężenia promieniowania wzdłuż rezonatora,

Modele te zawierają różne przybliżenia i wykorzystują różne metody obliczeniowe. Ich poprawność zależy często od specyfiki rozważanej struktury lasera, stąd trudno kategorycznie twierdzić, który z modeli jest lepszy. Ocenę ich jakości najlepiej przeprowadzać poprzez sprawdzenie, czy przewidują one eksperymentalnie obserwowane zjawiska, np. pojawianie się „schodków” (*kinks*) na charakterystykach emisyjnych (moc promieniowania w funkcji prądu) [22], samoistny efekt falowodowy w laserach paskowych bez wbudowanego falowodu [14, 15], niestabilność rozkładu pola promieniowania w strefie bliskiej tych laserów oraz jego znaczne przesunięcie w przypadku istnienia nawet niewielkiej asymetrii [22], samopodtrzymujące się pulsacje itd.

W dwóch z omawianych modeli, a mianowicie w modelach B i J, wyniki podano częściowo w postaci analitycznej, co uzyskano przez wprowadzenie dodatkowych założeń upraszczających. Niektóre z tych wyników są jakościowo zgodne z wynikami bardziej dokładnych modeli numerycznych.

W modelach D i H uwzględniono mod boczny pierwszego rzędu, a w modelach F, I i L — także mody boczne wyższych rzędów. Różne mody podłużne rozpatrzono jedynie w modelu K.

Najbardziej kompletnym opisem zjawisk elektrycznych cechuje się model I. W modelu E i w zmodyfikowanej wersji modelu G, uwzględniono zagadnienie przepływu prądu przez złącze *p-n*, podczas gdy w pozostałych modelach zastosowano jedynie uproszczone zależności opisujące rozptyw prądu między paskowym kontaktem, a obszarem czynnym.

Uśredniania gęstości strumienia fotonów i koncentracji nośników po długości rezonatora uniknięto jedynie w modelu M (porównaj również artykuł [39] w przypadku pracy poniżej progu akcji laserowej). Porównanie modeli L i M przeprowadzone w pracy [40] potwierdziło ich zgodność dla typowej struktury lasera. Jak więc się okazało, powyższe uśrednianie wpływa na pogorszenie dokładności modelu jedynie w przypadku niskich współczynników odbicia od zwierciadeł rezonatora i wysokich mocy emitowanego promieniowania.

Pozostałe modele, nie przedstawione w Tabeli 2, są przeważnie modelami uproszczonymi. W niektórych z nich [41–45] zakłada się np. ustalony rozkład pola promieniowania, co umożliwia separację równań opisujących ten rozkład i równań określających rozkład koncentracji nośników. W innych rozpatruje się jedynie konstrukcje specjalne, np. praca [46] zajmuje się własnościami laserów złączowych z krótkimi rezonatorami. Możliwa obecność nasycających się procesów absorpcyjnych oraz podłużny efekt wypalania dziury w rozkładzie wzmocnienia są analizowane w pracy [47]. Rozkład temperatury w obszarze czynnym został z kolei uwzględniony w modelu [48], a wpływ asymetrycznego rozkładu wzmocnienia był badany w modelu [49].

3.2. MODELE DYNAMICZNE

W większości modeli dynamicznych czasowe zmiany efektu falowodowego były zaniebawane ze względu na wielokrotne wydłużenie czasu niezbędnego dla tych obliczeń. W niektórych modelach [50, 51] uwzględniających ten efekt uwzględniono również wpływ dyfuzji nośników w płaszczyźnie złącza, a więc wpływ zmian rozkładu koncentracji noś-

ników w obszarze czynnym, natomiast w modelach bardziej uproszczonych [52, 53] — pomijano to zjawisko.

Uwzględnienie czasowych zmian rozkładu natężenia promieniowania powoduje znaczne wydłużenie czasu obliczeń w wyniku wielokrotnego rozwiązywania równania pola elektromagnetycznego. Należy wówczas dążyć do rozsądnego kompromisu pomiędzy czasem obliczeń, a ich dokładnością. Pierwszy taki model przedstawiono już ponad dziesięć lat temu [54]. Mimo jego prostoty (uwzględnił jedynie trójstrefowy rozkład pola), otrzymane wyniki były jakościowo zgodne z eksperymentem [55].

Wpływ koncentracji nośników na efekt falowodowy był dokładnie badany w pracy [56] dla wyjaśnienia różnic w zachowaniu się różnych struktur laserowych. Małosygnałową analizę efektu falowodowego zależnego od czasu przeprowadzono w pracach [57, 58]. W pracy [20], w której dla przybliżonego opisu tego efektu w laserze użyto prostego modelu płyty zespolonej, zademonstrowano pojawianie się niestabilności pola promieniowania. Podobny wynik otrzymano również w modelu [59], gdzie równanie pola elektromagnetycznego rozwiązywano metodą wariacyjną z zespoloną funkcją Gaussa jako funkcją próbną. Model ten jednakże nie zawsze daje poprawne wyniki, ponieważ metoda wariacyjna nie jest ściśle poprawna dla falowodów zespolonych [60]. W pracach [61—63] przedstawiono z kolei zastępczy model elektryczny lasera złączowego, w którym zastosowano programy znane z analizy obwodów elektrycznych.

Modele dynamiczne stosuje się również do badania częstotliwościowej modulacji laserów. Zmiany częstotliwości promieniowania emitowanego przez laser dokonuje się wówczas poprzez odpowiednie zmiany prądu zasilania, np. nakładając na prąd stały prąd sinusoidalnie zmienny. Wówczas rozstęp quasipoziomów Fermiego odtwarza zmiany prądu zasilania, a stąd zmienia się odpowiednio energia (czyli częstotliwość) emitowanego promieniowania. Dewiacja częstotliwości Δf może być w tym przypadku wyznaczona bezpośrednio ze stałej propagacji [3]:

$$\Delta f = -f \frac{\Delta \lambda}{\lambda} = -\frac{f}{\bar{n}_R} \Delta \left[\operatorname{Re} \left(\frac{\gamma}{k} \right) \right] \quad (2)$$

gdzie f — częstotliwość modulacji, λ — długość fali promieniowania, $\bar{n}_R$ — grupowy współczynnik załamania, γ — stała propagacji, a k — liczba falowa. Można tego dokonać nawet dla ustalonego rozkładu natężenia pola stosując parametr R

$$\Delta \left[\operatorname{Re} \left(\frac{\gamma}{k} \right) \right] = \frac{\Delta G}{2k} R, \quad (3)$$

gdzie G jest wzmocnieniem modowym, a R można wyrazić następująco:

$$R \equiv 2k \frac{\partial n_R / \partial n}{\partial g / \partial n} = \frac{2k}{a} \frac{\partial n_R}{\partial n}. \quad (4)$$

W powyższej zależności przez n_R oznaczono współczynnik załamania materiału obszaru czynnego dla długości fali emitowanego promieniowania, n jest koncentracją nośników w obszarze czynnym, g — lokalnym wzmocnieniem, a — współczynnikiem, który wraz z innym współczynnikiem b wiąże rozkład wzmocnienia lokalnego z rozkładem koncentracji nośników:

$$g(y) = an(y) - b \quad (5)$$

Metodę tę zastosowano w pracy [64] dla wyjaśnienia własności spektralnych laserów wykazujących samopodtrzymujące się pulsacje oraz przeskoków częstotliwości (*frequency chirp*) i dynamicznego poszerzenia linii spektralnej. Ostatnio opublikowano też modele teoretyczne laserów o sprzężonych optycznie rezonatorach (*coupled cavity lasers*) [65...68] jak również laserów półprzewodnikowych stosowanych w charakterze optycznych wzmacniaczy [69, 70].

PODSUMOWANIE

Wyboru odpowiedniego modelu teoretycznego należy dokonywać drogą kompromisu między sprawnością, dokładnością i złożonością obliczeń. Należy więc wpięrow ostrożnie rozważyć, który aspekt działania lasera zamierzamy badać, a stąd, jakie przybliżenia można zastosować bez obawy o poprawność otrzymanych wyników. Na wybór modelu wpływa więc zarówno wymagana dokładność modelowania analizowanego problemu, jak i typ komputera, jakim dysponujemy.

BIBLIOGRAFIA

1. W. Nakwaski: *Modelowanie zjawisk fizycznych w laserach złączowych. I. Modelowanie zjawisk elektrycznych*. Rozprawy Elektrotechniczne, 1988, t. 34, z. 3 ss. 981-987
2. W. Nakwaski: *Modelowanie zjawisk fizycznych w laserach złączowych. II. Modelowanie zjawisk optycznych*. Rozprawy Elektrotechniczne, 1988 t. 34 z. 3 ss. 989-994
3. J. Buus: *Principles of semiconductor laser modelling*. IEE Proc., 1985 132. Pt. J 42—51
4. D.P. Wilt, A. Yariv: *A self-consistent static model of the double-heterostructure laser*. IEEE J. Quantum Electron., 1981 QE-17, pp. 1941—1949
5. G. Lengyel, P. Meissner, E. Patzak, K.H. Zschauer: *An analitical solution of the lateral current spreading and diffusion problem in narrow oxide stripe (GaAl)As/GaAs DH lasers*. IEEE J. Quantum Electron., 1982 QE-18, pp. 618—625
6. W.B. Joyce, S.H. Wemple: *Steady-state junction-current distributions in thin resistive films on semiconductor junctions solution to $\nabla^2 v = \pm e^v$* . J. Appl. Phys., 1970 41, pp. 3818—3830
7. H. Yonezu, I. Sakuma, K. Kobayashi, T. Kamejima, M. Ueno, Y. Nannichi: *A GaAs-Alx Ga_{1-x}As double heterostructure planar stripe laser*, Jpn. J. Appl. Phys., 1973 12., 1973, 12, 1585—1592
8. A.S. Logginov, V.E. Solov'ev: *Rasczet raspredelenija koncentracji nositelej toka v inżekcionnyh poloskovykh lazerach*. Vestnik Moskovskogo Universiteta, Serija 3, Fizika, Astronomija, 1977 18 pp. 76—78
9. W.B. Joyce: *Current-crowded carrier confinement in double-heterostructure lasers*. J. Appl. Phys., 1980 51, pp. 2394—2401
10. B.W. Hakki: *GaAs double heterostructure lasing behaviour along the junction plane*. J. Appl. Phys., 1975 46, pp. 292—302
11. Gu Shi-jie, Hou Chong-ru: *Profiles of junction current and carrier-concentration in GaAs injection lasers with filament*. Acta Physica Sinica, 1979, 28, pp. 21—32
12. M. Yano, H. Ishikawa, M. Takusagawa: *Self-consistent solution for spatial hole burning in stripe geometry InGaAsP/InP double heterojunction lasers*. J. Appl. Phys., 1982 53, 3433—3437
13. W. Nakwaski: *Lasery paskowe z samoistnym efektem falowodowym: lasery z kontaktem paskowym i lasery z wysokorezystywnymi warstwami bocznymi*. Rozprawy Elektrotechniczne, 1983 t. 29, ss 611—637

14. W. Nakwaski: *Samoistny efekt falowodowy w laserach paskowych bez wbudowanego falowodu wewnętrznego*. Rozprawy Elektrotechniczne, 1984 30, ss. 245—260
15. W. Nakwaski: *Efekt falowodowy w laserach paskowych*. Rozprawy Elektrotechniczne 1989, t. 35, z. 3
16. B. W. Hakki: *Instabilities in output of injection lasers*. J. Appl. Phys., 1979 50, 5630—5637
17. T. L. Paoli: *Near-threshold behavior of the intrinsic resonant frequency in a semiconductor laser*. IEEE J. Quantum Electron., 1979 QE-15, pp. 807—812
18. J. P. Van der Ziel: *Self-focusing effects in pulsating AlxGa_{1-x} As double-heterostructure lasers*. IEEE J. Quantum Electron., QE-17 1981, pp. 60—68
19. R. P. Brouwer, C. H. F. Velzel, B. S. Yeh: *Lateral modes and self-oscillations in narrow stripe double-heterostructure GaAl-As injection lasers*. IEEE J. Quantum Electron., 1981 QE-17, pp. 694—700
20. B. S. Poh, T. E. Rozzi: *Intrinsic instabilities in narrow stripe geometry lasers caused by lateral current spreading*. IEEE J. Quantum Electron., 1981 QE-17, pp. 723—731
21. J. Buus: *Theory for self-sustained pulsations in semiconductor lasers*. Electron. Lett., 18 1982 718—719
22. W. Nakwaski: *Charakterystyki emisyjne laserów paskowych z samoistnym efektem falowodowym*. Rozprawy Elektrotechniczne, 1984, 30 ss. 261—285
23. N. Chinone: *Non-linearity in power-output-current characteristics of stripe geometry injection lasers*. J. Appl. Phys., 1977 48, pp. 3237—3243
24. G. H. B. Thompson, D. F. Lovelace, S. E. H. Turley: *Kinks in the light/current characteristics and near-field shift in (GaAl)As — heterostructure stripe lasers and their explanation by the effect of self-focusing on a built-in optical waveguide*. IEE J., Solid-State and Electron. Devices, 1978 2, pp. 12—30
25. K. Seki, T. Kamiya, H. Yanai: *Analysis of gain-guiding DH lasers — effects of transverse carrier diffusion*. Trans. Inst. Electron. Commun. Eng. Japan, 1979 E62 pp. 73—80
26. R. Lang: *Lateral transverse mode instability and its stabilization in stripe geometry injection lasers*. IEEE J. Quantum Electron. 1979, QE-15, pp. 718—726
27. M. Ueno, R. Lang, S. Matsumoto, H. Kawano, T. Furuse, I. Sakuma: *Optimum designs for InGaAsP/InP ($\lambda = 1.3 \mu\text{m}$) planoconvex waveguide lasers under lasing conditions*. IEE Proc., 1982, 129, Pt. I pp. 218—228
28. P. M. Asbeck, D. A. Cammack, J. J. Daniele, V. Klebanov: *Lateral mode behaviour in narrow stripe lasers*. IEEE J. Quantum Electron., 1979, QE-15, pp. 727—733
29. J. Buus: *A model for the static properties of DH lasers*. IEEE J. Quantum Electron. 1979 QE-15, pp. 734—739
30. K. A. Shore, T. E. Rozzi, G. H. Int Veld: *Semiconductor laser analysis: general method for characterising devices of various cross-sectional geometries*. IEE Proc., 127 Pt. I, (1980), 221—229
31. K. A. Shore: *Above-threshold current leakage effects in stripe-geometry injection lasers*. Optical and Quantum Electronics, 1983, 15, pp. 371—379
32. K. A. Shore: *Carrier diffusion and recombination influencing gain and current profiles in planar injection lasers*. J. Appl. Phys., 1984, 56, pp. 1293—1297
33. W. Streifer, R. D. Burnham, D. R. Scifres: *Chanelled substrate non planar laser analysis. Part I: Formulation and the plano convex waveguide laser*. IEEE J. Quantum Electron., 1981, QE-17, pp. 736—744
34. W. Streifer, D. R. Scifres, R. D. Burnham: *Channeled substrate non planar laser analysis. Part II: Lasers with tapered active regions*. IEEE J. Quantum Electron. 1981, QE-17, pp. 1521—1530
35. C. B. Su: *An analytical solution of kinks and non-linearities driven by near-field displacement instabilities in stripe geometry diode lasers*. J. Appl. Phys., 52, 1981, 2665—2673
36. J. Machač: *Static model of DH laser*. IEE Proc., 1983, 130, Pt. I pp. 61—66
37. J. Buus: *Models of the static and dynamic behaviour of stripe geometry lasers*. IEEE J. Quantum Electron., QE-19, 1983, 953—960
38. R. Baets, P. E. Lagasse: *Longitudinal static-field model for DH lasers*. Electron. Lett., 1984, 20, pp. 41—42

39. G.P. Agrawal, W.B. Joyce, R.W. Dixon, M. Lax: *Beam-propagation analysis of stripe geometry semiconductor lasers: threshold behaviour*. Appl. Phys. Lett., 1983, 43, pp. 11—13
40. R. Baets, J. Buus: *Comparison of a beam propagation model for laser diodes with other modelling techniques*. IEE Proc., 132, Pt. J, 1985, 205—206
41. J. Pietzsch, T. Kamiya: *Axial mode spectra of planar stripe geometry laser diodes*. Jpn. J. Appl. Phys., 21, 1982, L633—L635
42. J. Pietzsch, T. Kamiya: *Loss and TM mode gain suppression in planar stripe diode lasers with smooth lateral profile*. Jpn. J. Appl. Phys., 22, 1983, pp. 101—108
43. W. Streifer, D.R. Scifres, R.D. Burnham: *Analysis of diode laser properties*. IEEE J. Quantum Electron., 1982, QE-18, pp. 1918—1929
44. W. Streifer, A. Hardy, D.R. Scifres, R.D. Burnham: *Analysis of diode laser properties — Part II*, IEEE J. Quantum Electron., 1983, QE-19, pp. 991—1002
45. G. Coppa, P. Di Vita, M. Potenza: *A simplified model for the analysis of semiconductor lasers*. CSELT Rapporti tecnici, 1983 11, pp. 29—33
46. T.-P. Lee, C.A. Burrus, J.A. Copeland, A.G. Dentai, D. Marcuse: *Short-cavity InGaAsP injection lasers: Dependence of mode spectra and single-longitudinal-mode power on cavity length*. IEEE J. Quantum Electron. 1982, QE-18, 1101—1113
47. J.A. Copeland: *Single-mode stabilization by traps in semiconductor lasers*. IEEE J. Quantum Electron., 1980 QE-16, pp. 721—727
48. T.J.S. Mattos, N.B. Patel, F.D. Nunes: *Calculation of the threshold current of stripe-geometry double-heterostructure GaAs-Ga_{1-x}Al_xAs lasers, including a self-consistent treatment of the current-temperature dependence*. J. Appl. Phys., 1982, 53, pp. 149—155
49. D. Marcuse, F.R. Nash: *Computer model of an injection laser with asymmetrical gain distribution*. IEEE J. Quantum Electron. QE-18, 1982, pp. 30—43
50. J. Buus, M. Danielsen: *Carrier diffusion and higher order transversal modes in spectral dynamics of the semiconductor laser*. IEEE J. Quantum Electron., 1977, QE-13, pp. 669—746
51. M. Ito, T. Ito, T. Kimura: *Dynamic properties of semiconductor lasers*. J. Appl. Phys., 1979 50, pp. 6168—6174
52. K. Petermann: *Theoretical analysis of spectral modulation behaviour of semiconductor injection lasers*. Optical and Quantum Electronics, 1978 10, pp. 233—242
53. M.J. Adams, M. Osinski: *Longitudinal mode competition in semiconductor lasers: rate equations revisited*. IEE Proc., 129, Pt. I 1982, pp. 271—274
54. M. Cross: *A theory of the transient evolution and self-focusing behaviour of lasing filaments in injection lasers*. Phys. Status Solidi (a), 1973, 16 pp. 167—179
55. P.R. Selway, P.A. Kirkby, A.R. Goodwin, G.H.B. Thompson: *Dynamics of self focusing in stripe-geometry semiconductor lasers*. IEE J., Solid-State and Electron. Devices, 1978, 2 pp. 38—40
56. N. Chinone, K. Aiki, M. Nakamura, R. Ito: *Effects of lateral mode and carrier density profile on dynamic behaviour of semiconductor lasers*. IEEE J. Quantum Electron. 1978, QE-14, pp. 625—631
57. M. Yano, K. Seki, T. Kamiya, H. Yanai: *Small signal modulation characteristics of gain-guiding double heterostructure lasers*. Trans. Ins. Electron. Commun. Eng. Jpn., 1978 J61-C, pp. 690—697
58. R. Lang: *Intensity pulsation enhancement by self-focusing in semiconductor lasers*. Jpn. J. Appl. Phys., 1980, 19, pp. L39—L96
59. C.-Z. Guo, K.-G. Wang: *Intrinsic pulsation in stripe-geometry DH semiconductor lasers*. IEEE J. Quantum Electron., 1982, QE-18, pp. 1728—1737
60. J. Buus: *Comparison of two recent semiconductor laser models*. IEEE J. Quantum Electron., 1983, QE-19, pp. 1356—1358
61. R.S. Tucker, D.J. Pope: *Microwave circuit models of semiconductor injection lasers*. IEEE Trans. Microwave Theory and Techn., MTT-31, 1983 pp. 289—294
62. R.S. Tucker, D.J. Pope: *Circuit modelling of the effect of diffusion on damping in a narrow-stripe semiconductor laser*. IEEE J. Quantum Electron., 1983, QE-19 pp. 1179—1183

63. R. S. Tucker, I. P. Kaminow: *High-frequency characteristics of directly modulated InGaAsP ridge waveguide and buried heterostructure lasers*. IEEE J. Lightwave Technol., 1984 LT-2, pp. 385—393
64. M. Osiński, M. J. Adams: *Transient time averaged spectra of rapidly-modulated semiconductor lasers*. IEE Proc., 1985 132, Pt. J, pp. 34—37
65. D. Marcuse, T.-P. Lee: *Rate equation model of a coupled cavity laser*. IEEE J. Quantum Electron., 1983. QE-20 pp. 166—176
66. L. A. Coldren, T. L. Koch: *Analysis and design of couple cavity laser — Part I: Threshold gain analysis and design guidelines*, IEEE J. Quantum Electron., 1984, QE-20 pp. 659—670
67. L. A. Coldren, T. L. Koch: *Analysis and design of coupled cavity lasers — Part II: Transient analysis*. IEEE J. Quantum Electron., 1984. QE-20 pp. 671—682
68. W. Streifer, D. Yevick, T. L. Paoli, R. D. Burnham: *An analysis of cleaved coupled-cavity lasers*. IEEE J. Quantum Electron. QE-20 1984, pp. 754—764
69. D. Marcuse: *Computer model od an injection laser amplifier*. IEEE J. Quantum Electron., 1983 QE-19 pp. 63—73
70. M. J. Adams, J. V. Collins, I. D. Henning: *Analysis of semiconductor laser optical amplifiers*. IEE Proc., 1985 132, Pt. J, pp. 58—63
71. J. E. A. Whiteaway: *Theoretical analysis of current spreading in stripe-geometry injection lasers*. IEE Proc., 1982 129, Pt. I pp. 89—95.

W. NAKWASKI

Part III. COMPARISON OF MODELS

Summary

The calculation procedure of a complete model of Physical phenomena in a diuble-heterostructure stripe laser is presented and a comparison of hitherto existing models of a modern laser dode is accomplished.

W. NAKWASKI

Partie III. COMPARAISON DES MODÈLES

Résumé

Dans l'étude on a présenté les différentes méthodes de calculs utilisées pour la résolution des équations de simulation des phénomènes physiques se produisant dans un laser longitudinal bihétérocon jugué et on a comparé les modèles jusqu'à présent connus des lasers conjugués à constructions divetses.

W. NAKWASKI

Teil III. VERGLEICHUNG VON MODELLEN

Zusammenfassung

In dem Aufsatz werden Berechnungsverfahren geschildert, die physikalischen Erscheinungen der doppel-heterostrutturellen Streifendiodenlaser modullieren, Auch werden die bisher bekannten Lasermodelle von Verbindungslasern verschiedener Konstruktionen mit den modernen Diodenlasern verglichen.

В. НАКВАСКИ

Часть 3. СОПОСТАВЛЕНИЕ МОДЕЛЕЙ

Резюме

Представлена вычисленная процедура полной модели физических явлений в полосковом лазерном диоде с двойной гетероструктурой которая сопоставлена с существовавшими до сих пор моделями современного лаверного диода.

Pomiary pól elektromagnetycznych antenami kierunkowymi

HUBERT TRZASKA

Instytut Telekomunikacji i Akustyki, Politechnika Wrocławska

Otrzymano 1985.03.28

Autoryzowano do druku 1987.09.02

W pracy przeanalizowano wykorzystanie metody wzorcowej anteny odbiorczej do wzorcowania mierników natężenia pola elektromagnetycznego wyposażonych w anteny kierunkowe. Wykazano, że metoda ta może być stosowana wprost, bądź też jej dokładność może być podniesiona przy zastosowaniu stosunkowo prostych środków zaradczych. Krótko przedyskutowano zakres pełnej przydatności mierników z antenami kierunkowymi do pomiarów pola elektromagnetycznego oraz zaprezentowano koncepcję układu pozwalającego na pomiary pól dowolnie złożonych, np. pól o polaryzacji elipsoidalnej.

1. WSTĘP

Współczesna cywilizacja charakteryzuje się olbrzymim rozwojem zastosowań energii wielkiej częstotliwości. Każde z tych zastosowań wiąże się z zamierzonym lub niezamierzonym wypromieniowaniem tej energii. Miarą ilości wypromieniowywanej energii może być degradacja naturalnego środowiska elektromagnetycznego, która w odniesieniu do innych czynników jest relatywnie największa i występuje w skali globalnej.

Rozwojowi zastosowań energii w.cz. towarzyszył rozwój narzędzi do pomiarów pól elektromagnetycznych. Pomiary te są przeprowadzane zwykle w trzech kategoriach zainteresowań:

1. pomiar swobodnie rozchodzących się fal elektromagnetycznych w takich zastosowaniach jak, np.: telekomunikacja bezprzewodowa, radiolokacja, geofizyka. Celem tych pomiarów są badania propagacyjne, badania urządzeń promieniujących, własności ośrodka i in.,
2. pomiar zakłóceń radioelektrycznych z punktu widzenia kompatybilności elektromagnetycznej oraz nieradiotechnicznych źródeł zakłóceń, jak np.: silniki spalinowe i elektryczne, linie energetyczne, wyładowania atmosferyczne, urządzenia domowego użytku,
3. pomiary susceptybilności elektromagnetycznej organizmów, urządzeń i materii, badanie środowiska elektromagnetycznego w jakim one pracują oraz badania skuteczności stosowanych środków ochronnych i to zarówno gdy pole jest wytwarzane celowo (telekomunikacja), jak i wtedy gdy jest ono produktem niepożądanym (zastosowania przemysłowe).

Prace związane z kategorią trzecią a wykonywane w znacznej większości przypadków w polu bliskim, wykazały konieczność stosowania w metrologii pól elektromagnetycznych w polu bliskim anten o sferycznej charakterystyce promieniowania; anten reagujących na sygnały o dowolnej polaryzacji. Konieczność ta wynika z możliwości występowania w polu bliskim pól o polaryzacji sferoidalnej, przy czym zwykle brak jakiejkolwiek informacji a priori dotyczącej przestrzennego położenia wektora natężenia pola elektrycznego lub/i magnetycznego oraz zmian tego położenia w funkcji czasu. Niezależnie od napisanego spotyka się publikacje dotyczące tej dziedziny, a prezentujące wyniki pomiarów w polu bliskim z zastosowaniem anten tubowych, logarytmicznych i innych o kierunkowej charakterystyce promieniowania i określonych własnościach polaryzacyjnych.

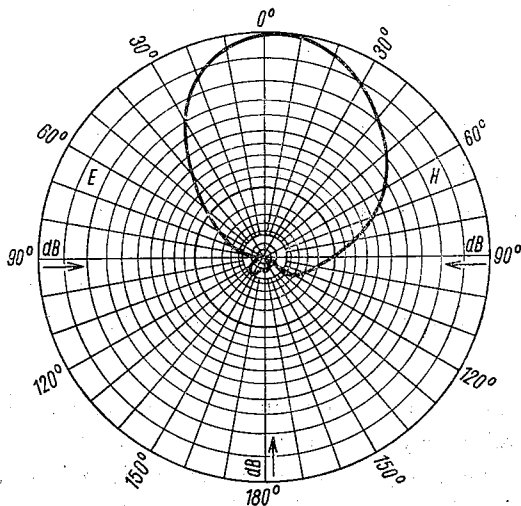
Druga grupa zastosowań dotyczy zarówno pola bliskiego jak i dalekiego. Pewną kontrowersję budzą tu metody pomiarów zakłóceń wytwarzanych, np: przez pojazdy. Pomiarów te są wykonywane, w polu bliskim, antenami rezonansowymi. Jednak pomiary te są traktowane zwykle nie jako pomiary natężenia pola elektromagnetycznego lecz jako pomiary napięcia na wejściu określonej anteny przy ustalonej geometrii układu.

Grupa pierwsza dotyczy w większości pomiarów wykonywanych w polu dalekim. Ze względu na własności tego pola uważano, że nie ma podstaw do formułowania ograniczeń w stosowaniu anten pomiarowych. Za szczególnie korzystne uważano stosowanie anten o kierunkowej charakterystyce promieniowania. Problemowi wykorzystania anten kierunkowych w tej głównie grupie zastosowań poświęcona jest niniejsza praca.

2. MODEL ANTENY

Określenie „*antena kierunkowa*” wymaga pewnego wyjaśnienia gdyż sugeruje ono istnienie anten bezkierunkowych. Można pewnej grupie anten przypisywać własności bezkierunkowe, jedna z takich propozycji zostanie przedstawiona w dalszej części pracy. Ciągłe jednak obowiązuje pogląd, że nie jest możliwym zbudowanie anteny o sferycznej charakterystyce promieniowania a spełniającej zasadę wzajemności. W metrologii pól elektromagnetycznych wykorzystuje się anteny ramowe lub ferrytowe a także monopolowe w zakresie częstotliwości poniżej 30 MHz. Powyżej 30 MHz stosuje się zwykle anteny dipolowe. Wszystkie wymienione anteny są sensu stricto antenami kierunkowymi. Jednak ich własności kierunkowe są słabe, w pewnych warunkach mogą być traktowane jako dookólne a ich własności jako anten o dookólnej charakterystyce promieniowania mogą być poprawiane przez stosowanie kombinacji kilku anten pojedynczych [1]. Ze względu na rozmiary geometryczne anteny o kierunkowej charakterystyce promieniowania są stosowane w metrologii w zakresie częstotliwości powyżej około 200—300 MHz. Kolejno, ze wzrostem częstotliwości, stosuje się anteny Yagi-Uda, logarytmiczno-periodyczne, systemy składające się z kilku anten Yagi, anteny spiralne i tubowe. Własności kierunkowe anten zależą głównie od ich rozmiarów elektrycznych, natomiast własności szerokopasmowe od wyboru typu anteny i jego realizacji. Ze względu na połączenie wystarczających własności kierunkowych z możliwością pracy w szerokim paśmie częstotliwości w zakresie częstotliwości 0,3—1 GHz najchętniej stosuje się w metrologii anteny logarytmiczno-periodyczne. Taką też antenę wybrano jako przedmiot przykładowych rozważań.

Antena typu DAMZ-4/50 jest anteną logarytmicznie-periodyczną przeznaczoną do pracy w zakresie częstotliwości 300—1000 MHz, jej zysk względem dipola półfalowego wynosi około 7 dB a tłumienie promieniowania wstecznego ponad 20 dB. Charakterystykę promieniowania anteny w płaszczyznach E i H pokazano na Rys. 1 [2]. Do celów rachun-



Rys. 1. Charakterystyka promieniowania anteny logarytmiczno-periodycznej w płaszczyznach E i H

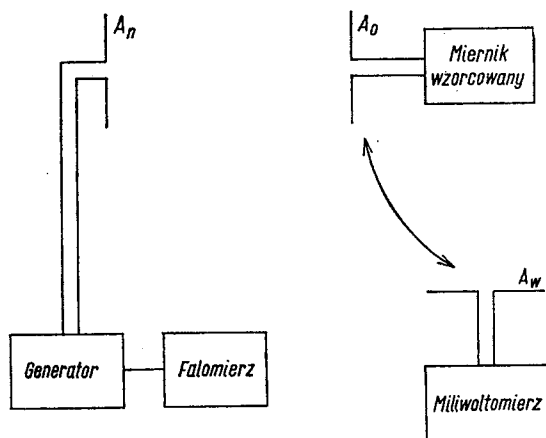
kowych charakterystykę promieniowania w płaszczyźnie H aproksymowano z dokładnością ± 1 dB dla $\varphi < 70^\circ$ funkcją $f(\Theta, \varphi)$ w postaci:

$$f(\Theta, \varphi) = 1 + \cos^2 \varphi. \quad (1)$$

Przedstawiona na Rys. 1 charakterystyka promieniowania jest charakterystyką uśrednioną. Charakterystyka promieniowania anteny logarytmiczno-periodycznej jest zależna od kąta rozwarcia anteny i zmienia się w pewnych granicach w funkcji częstotliwości. Jednak ze względu na decydującą rolę prądów płynących w trzech sąsiednich dipolach charakterystyka ta jest zbliżona do charakterystyki [3] promieniowania trój-elementowej anteny Yagi. W świetle napisanego można przyjąć, że przybliżenie wyrażone wzorem (1) jest wystarczająco dokładne.

3. WZORCOWANIE MIERNIKA NATĘŻENIA POŁA

Wzorcowanie mierników natężenia pola elektromagnetycznego wyposażonych w anteny dipolowe lub kierunkowe prowadzi się metodą wzorcowej anteny nadawczej lub metodą wzorcowej anteny odbiorczej [4]. Zarówno ze względu na dokładność jak i na mniejszą złożoność bardziej rozpowszechniona jest metoda wzorcowej anteny odbiorczej. W metodzie tej (Rys. 2) pole elektromagnetyczne jest wytwarzane przez dowolne źródło. W wybranym punkcie przestrzeni umieszcza się wzorcową antenę odbiorczą — najczęściej w postaci dipola półfalowego i na podstawie pomiaru siły elektromotorycznej indukowanej



Rys. 2. Wzorcowanie metodą podstawienia

w antenie oraz znajomości parametrów anteny wyznacza się natężenie pola w miejscu umieszczenia wzorcowej anteny odbiorczej. Następnie w miejsce anteny wzorcowej podstawia się antenę wzorcowaną przyjmując, że natężenie pola elektromagnetycznego nie ulega przy tym zmianie a jeżeli występują zmiany rozkładu pola w otoczeniu anteny badanej, to będą one podobne w trakcie pomiaru jak i wzorcowania. Metoda taka, przy zachowaniu określonych warunków ograniczających istotność poszczególnych czynników wpływających, nie budzi zastrzeżeń przy wzorcowaniu anten dipolowych o wymiarach zbliżonych do wymiarów anten wzorcowych. Przy wzorcowaniu tą metodą anten o kierunkowej charakterystyce promieniowania, anten dwu- a czasem i trój-wymiarowych, koniecznym staje się przeanalizowanie błędów wzorcowania wynikających z przestrzennej konstrukcji tych anten jak i ich własności kierunkowych.

Podstawowym warunkiem, który musi być spełniony w układach do wzorcowania, jest warunek minimalnej odległości pomiędzy źródłem pola a anteną wzorcową lub wzorcowaną:

$$d > \frac{2D^2}{\lambda} \quad (2)$$

gdzie: d — odległość pomiędzy antenami,

D — maksymalny wymiar liniowy największej z anten.

Spełnienie warunku (2) nie zapewnia jednorodności pola w obszarze zajmowanym przez anteny odbiorcze, rezultatem wyznaczania natężenia pola anteną wzorcową jest podanie wartości uśrednionej w obszarze zajmowanym przez tą antenę, jednak pozwala na stosowanie przybliżeń upraszczających rozważania.

Składową poprzeczną pola elektrycznego wytwarzanego przez dipol półfalowy określa zależność:

$$E = -j30I_m \left(\frac{e^{-jkR_1}}{R_1} + \frac{e^{-jkR_2}}{R_2} \right) \quad (3)$$

gdzie: I_m — amplituda prądu dopływającego do anteny,

$$R_1 = \sqrt{(z - \lambda/4)^2 + d^2}$$

$$R_2 = \sqrt{(z + \lambda/4)^2 + d^2}$$

z — bieżąca długość dipola $z \leq \lambda/4$.

Przyjmując w skrajnym przypadku $z = 0$ oraz dla $d > 2\lambda$, co dla rozważanej anteny logarytmiczno-periodycznej spełnia warunek (2), można przyjąć: $R_1 = R_2 = d$. Ze względu na małą szerokopasmowość oraz brak, korzystnych tu, własności kierunkowych, stosowanie anten dipolowych jako źródeł pola jest ograniczone do częstotliwości poniżej 300 MHz. Powyżej 300 MHz korzystnym będzie zastosowanie jako anteny nadawczej także anteny szerokopasmowej i o pewnych własnościach kierunkowych, np. też anteny logarytmiczno-periodycznej. Na podstawie przeprowadzonego wyżej rozumowania natężenie składowej poprzecznej pola elektrycznego wytwarzanego przez tą antenę określimy w postaci:

$$E = E_0 \frac{e^{-jkd}}{d} F(\Theta, \varphi), \quad (4)$$

gdzie: E_0 — stała obrazująca amplitudę pola, nieistotna z punktu widzenia prowadzonych rozważań,

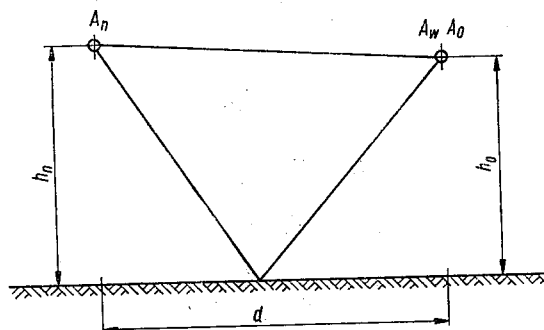
$F(\Theta, \varphi)$ — charakterystyka promieniowania anteny nadawczej (unormowana).

Do dalszych rozważań przyjmujemy, że charakterystyki promieniowania anteny nadawczej i anteny wzorcowej są jednakowe i dane krzywymi pokazanymi na Rys. 1 oraz wzorem (1).

3.1 WZORCOWANIE W WARUNKACH SWOBODNEJ PRZESTRZENI

Geometrię propagacji pokazano na Rys. 3. Przy wzorcowaniu w warunkach swobodnej przestrzeni istnieje jedynie promień bezpośredni. Jeżeli przyjmujemy, że największym wymiarem liniowym w układzie jest długość wzorcowanej anteny logarytmiczno-periodycznej D , oraz obliczymy zmianę modułu natężenia pola wzdłuż anteny otrzymamy:

$$\left| \frac{\Delta E}{E} \right| = \frac{D}{d} \quad (5)$$



Rys. 3. Geometria układu wzorcowania

Zależność (5) wyraża zmianę modułu natężenia pola w stosunku do jednego z końców anteny. Długość anteny DAMZ-4/50 wynosi 70 cm, zatem uzyskanie równomierności natężenia pola wzdłuż anteny nie przekraczającej 1% wymaga rozsunięcia anteny nadawczej i odbiorczej na odległość 70 m. Czy jest to konieczne?

Jak wyżej wspomniano rolę decydującą o pracy anteny logarytmiczno-periodycznej spełniają trzy sąsiadujące dipole. Zatem, z punktu widzenia poprawności wzorcowania, nie interesuje nas jednorodność pola wzdłuż całej anteny a jedynie jego niejednorodność wzdłuż czynnej długości anteny (D'), która dla omawianej anteny nie przekracza $0,25 \lambda$. Przyjmując, dla najmniej korzystnych warunków, $d = 2 \text{ m}$ i $f = 300 \text{ MHz}$ oraz odnosząc wartości natężenia pola do środka jej długości czynnej otrzymamy $|\Delta E|/|E| < \pm 6\%$.

Jednak, jak wynika z wzoru (4), pole w obszarze anteny zmienia się liniowo. Zatem błąd wzorcowania odnoszony do środka czynnej długości anteny powinien się uśredniać i być pomijalnie małym, pomijalnym z punktu widzenia pomiarów współczynnika antenowego. Celem zweryfikowania tej hipotezy pomierzono rozkład natężenia pola w funkcji odległości anteną logarytmicznie-periodyczną i anteną dipolową; przy tym położenia anteny dipolowej odpowiadały położeniu elementu rezonansowego anteny logarytmiczno-periodycznej dla częstotliwości pomiarowej. Różnica napięć zmierzonych na obu antenach, z dokładnością wykonanych pomiarów, dla $d > 1 \text{ m}$ jest wielkością stałą. Uzasadnia to słuszność przeprowadzonego rozumowania i możliwość wykonywania wzorcowań nawet przy $d < 2 \text{ m}$, co bywa stosowane.

3.2 WZORCOWANIE W WARUNKACH POLA POMIAROWEGO

Przy wzorcowaniu miernika pola elektromagnetycznego w warunkach pola pomiarowego sumaryczne natężenie pola elektromagnetycznego w miejscu obserwacji (E_s) jest superpozycją natężeń pola promienia bezpośredniego (E_d) i promienia odbitego od powierzchni ziemi (E_r):

$$|E_s| = \sqrt{|E_r|^2 + |E_d|^2 + 2|E_r||E_d|\cos(\Phi_r - \Phi_d)}, \quad (6)$$

przy czym Φ_r i Φ_d są fazami obu promieni.

E_d można wyznaczyć wprost z zależności (4) podstawiając, dla kierunku maksymalnego promieniowania, $F(\theta, \varphi) = 1$. Natężenie promienia odbitego jest zależne od drogi którą on pokonuje (r), charakterystyk kierunkowych anten nadawczej i odbiorczej oraz współczynnika odbicia na granicy ośrodków (R).

Rozpatrzmy trzy przypadki:

a. antena nadawcza i odbiorcza w postaci dipola

$$E_r = E_0 \frac{e^{-jkr}}{r} |R| \quad (7)$$

b. jedna z anten dipolowa, jedna kierunkowa

$$E_r = E_0 \frac{e^{-jkr}}{r} F(\varphi) R \quad (8)$$

dla polaryzacji poziomej pominięto zależność F od θ

c. obie anteny kierunkowe o jednakowych charakterystykach promieniowania

$$E_r = E_0 \frac{e^{-jkr}}{r} F^2(\varphi) R. \quad (9)$$

Przyjmijmy jednakową wysokość zawieszenia obu anten i policzmy przykładowo dla przypadku b. stosunek $|E_s|/|E_d|$:

$$\frac{|E_s|}{|E_d|} = \sqrt{1 + \left(\frac{d}{r}\right)^2 \left(\frac{1 + \cos^2 \varphi}{2}\right)^2 |R|^2 + 2 \frac{d}{r} \frac{1 + \cos^2 \varphi}{2} |R| \cos \Delta \Phi} \quad (10)$$

gdzie: $\Delta \Phi$ — różnica faz promienia bezpośredniego i odbitego.

Zależność (10) jest analogiczna do stosowanej przez Crawforda [6]. Zależność ta pozwala na wyznaczenie rozkładu natężenia pola w obszarze zajmowanym przez antenę odbiorczą. Ze względu na dostępność maszyn cyfrowych oraz graficzną efektywność takich krzywych są one chętnie prezentowane w licznych publikacjach dotyczących zwłaszcza pola bliskiego. Zastanówmy się nad ich sensem.

Różnica faz $\Delta \Phi$ zależy od łatwych do wyznaczenia czynników określających elektryczną długość obu promieni, tj. kr i kd , oraz od trudniejszych do wyznaczenia faz charakterystyki promieniowania (co zostało bez wyjaśnienia pominięte we wzorze (1)) i współczynnika odbicia. Oczywiście można je wyznaczyć z określoną dokładnością i uwzględnić w obliczeniach, jednak ich zmienność w czasie, przestrzeni i z częstotliwością ogranicza w sposób zasadniczy wiarygodność, a zatem i przydatność, rezultatów takich obliczeń. Zwykle przyjmuje się $R = -1$, jest to jednak w pełni słuszne jedynie przy pomiarach nad ośrodkiem doskonale przewodzącym, co bywa niekiedy stosowane. Jednak, nie tylko z rozpatrywanego punktu widzenia, byłoby korzystnym takie wybranie terenu pomiarowego, aby R był możliwie mały — wtedy faza współczynnika odbicia może być dowolna. Dlatego też umieszcza się układy opisywanego typu nad ośrodkiem o kiepskiej przewodności, a były także koncepcje umieszczania wzorców w naturalnych lub sztucznych jaskiniach, w ośrodku o słabym zawadnieniu. Ponadto, co wymaga przypomnienia, wielkość określona wzorem (10) jest zależna od częstotliwości, co wymaga prowadzenia obliczeń dla wszystkich potrzebnych częstotliwości oddzielnie. Szczegółowa analiza tak zebranych wyników obliczeń jest zatem obciążona błędem stosowanych tu przybliżeń z jednej strony, z drugiej zaś bardzo złożona ze względu na obszerność danych. Dla jej uproszczenia przyjmijmy, że $\Delta \Phi$ może przyjmować wartości dowolne. Pozwala to na znalezienie obwiedni zmian natężenia pola w funkcji odległości, częstotliwości, dla różnych anten i współczynników odbicia. Obwiednię 0 wyznaczmy dzieląc $|E_s|$ przez $|E_d|$ i otrzymamy dla:

a. obu anten dipolowych 0_{dd}

$$O_{dd} = \frac{|E_s|}{|E_d|} = 1 \pm \frac{d}{r} |R| \quad (11)$$

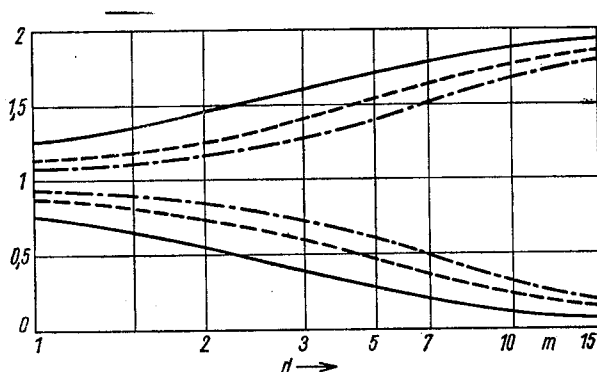
b. anteny dipolowej i logarytmiczno-periodycznej 0_{dl}

$$O_{dl} = 1 \pm \frac{d}{r} \frac{1 + \cos^2 \varphi}{2} |R| \quad (12)$$

c. obu anten logarytmiczno-periodycznych O_{ll}

$$O_{ll} = 1 \pm \frac{d}{r} \left(\frac{1 + \cos^2 \varphi}{2} \right)^2 |R| \quad (13)$$

Przebiegi obwiedni w funkcji odległości d pomiędzy anteną nadawczą i odbiorczą, dla wysokości zawieszenia obu anten $h = 2$ m pokazano na Rys. 4.



Rys. 4. Przebiegi obwiedni: O_{aa} — linia ciągła, O_{al} — linia przerywana, O_{ll} — linia kropkowana

Przedmiotem rozważań w tej części pracy jest dokładność wzorcowania miernika natężenia pola elektromagnetycznego wyposażonego w antenę kierunkową przy umieszczeniu anteny nad powierzchnią odbijającą. A dokładniej wpływ promienia odbitego na dokładność wzorcowania. Wpływ ten nie jest rezultatem istnienia promienia odbitego w ogóle, lecz kierunkowości anteny wzorcowanej i nadawczej. Wzorcową anteną odbiorczą jest zawsze antena dipolowa. Zatem przedstawione na Rys. 4 krzywe pozwalają na przeanalizowanie wpływu promienia odbitego w dwu wariantach: przy wykorzystaniu anteny dipolowej także jako anteny nadawczej (wariant a. + b.) i przy antenie nadawczej logarytmiczno-periodycznej (wariant b. + c.).

Dla zminimalizowania błędu wzorcowania przyjmijmy, że wzorcowanie jest wykonywane w otoczeniu brzośca fali stojącej, t.j. dla $O_{ii} > 1$. Dla dipolowej anteny nadawczej dodatkowy błąd wzorcowania η_d określimy w postaci:

$$\eta_d = \frac{O_{aa} - O_{al}}{O_{aa}} = 1/2 \cdot \sin^2 \varphi \cdot \left(1 + \frac{r}{d} \right)^{-1} \quad (14)$$

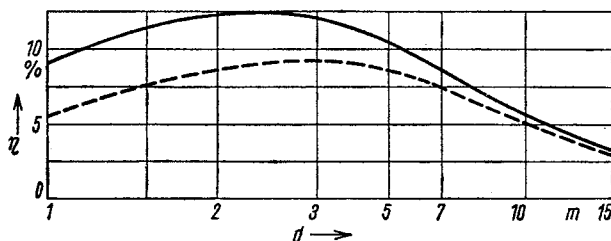
Analogicznie dla logarytmiczno-periodycznej anteny nadawczej dodatkowy błąd wzorcowania η_l wyniesie:

$$\eta_l = \frac{O_{al} - O_{ll}}{O_{al}} \quad (15)$$

Przebiegi η_d i η_l w funkcji odległości pokazano na Rys. 5.

Przy wyprowadzaniu powyższych zależności zostały popełnione trzy nieformalności, są to:

1. Przyjęta jedna wartość E_0 dla dipolowej i kierunkowej anteny nadawczej,
2. prowadzone obliczenia bez uwzględnienia długości anteny logarytmicznej,
3. operowanie wielkością E_s w miejsce SEM indukowanej w antenie odbiorczej.



Rys. 5. Przebiegi η_d — linia ciągła i η_t — linia przerywana

Łatwo zauważyć, że te nieformalności w niczym nie ograniczają wniosków wynikających z analizy zależności (14) i (15), zwłaszcza przy spełnieniu warunku (2).

Można mieć także pewne zastrzeżenia odnośnie stosowania w przedstawionych rozważaniach metody „radiofalowej”. Dokładną analizę własności pola pomiarowego metodą indukowanych SEM przedstawili Kawana i Miyajima [7, 8]. Metodę tą można zastosować także i tu, jednak prowadzi ona do większej złożoności rachunkowej. Zbieżność wyników otrzymanych metodą indukowanych SEM oraz metodą radiofalową, zwłaszcza dla $f > 300$ MHz, wykazali Borsero i Nano [9].

Jak widać z Rys. 5 dodatkowe błędy wzorcowania w wielu zastosowaniach i przy dokładnościach niektórych mierników natężenia pola mogą być pominięte. Pamiętając, że przedstawione krzywe pokazują wartości graniczne, przedstawmy możliwości obniżenia rozważanego błędu. Oto one:

1. Stosowanie kierunkowej anteny nadawczej o możliwie dużym zysku kierunkowym,
2. Dobór odległości pomiędzy antenami,
3. Zastosowanie polaryzacji pionowej celem zmniejszenia R ,
4. Obniżenie R przez:
 - umieszczenie wzorca na terenie o małym R ,
 - pokrycie obszaru odbijającego dielektrykiem pochłaniającym energię w.cz. [10], łatwo zauważyć, że zmniejszenie R do 10% lub poniżej tej wartości rozwiązuje problem radykalnie,
5. Obliczenie i uwzględnianie odpowiednich poprawek; rozważany błąd jest błędem deterministycznym, co umożliwia przeprowadzenie szczegółowych obliczeń dla danej anteny, częstotliwości i geometrii układu oraz odpowiednie skorygowanie danych eksperymentalnych.

Natomiast przy pomiarach bardziej subtelnych, gdzie stosowane środki zaradcze nie dadzą zadowalających wyników, tam gdzie nie są wskazane wzajemne oddziaływania z zewnętrznym środowiskiem elektromagnetycznym, a także gdy istnieje potrzeba prowadzenia badań niezależnie od zewnętrznych warunków klimatycznych, jest koniecznym stosowanie komór bezechoowych.

4. POMIAR NATĘŻENIA POLA

Wiele służb i instytucji zajmuje się pomiarami pola elektromagnetycznego oraz wytwarzaniem pól elektromagnetycznych o znanym natężeniu. Zaprezentowały one wątpliwości dotyczące przydatności dotychczas stosowanych wzorców natężenia pola elektro-

magnetycznego do wzorcowania mierników wyposażonych w anteny kierunkowe. To jest geneza niniejszej pracy. Na podstawie wyników rozumowania przedstawionego w poprzednim rozdziale można stwierdzić, że sytuacja nie wygląda źle, a stosowanie stosunkowo prostych środków zaradczych może ją znacznie poprawić. Stanowi to pełną odpowiedź na istniejące wątpliwości. Można na tym pracę zakończyć. Byłoby to jednak, znane z innych dziedzin życia, odesłanie problemu do innego adresata, bowiem z twierdzenia: „miernik został poprawnie wywzorcowany” nie wynika wniosek: „miernik jest przydatny do wykonywania pomiarów”. Brak tu określenia warunków, w których ten miernik zapewni wymaganą dokładność pomiaru. Dla pełnego przedstawienia zagadnienia przeanalizujemy krótko pomiary pola elektromagnetycznego z wykorzystaniem anten kierunkowych. Rozważania zostaną przeprowadzone dla dwu skrajnych przypadków z odrzuceniem trywialnego, to jest pomiaru w swobodnej przestrzeni.

4.1 POMIAR W WARUNKACH PROPAGACJI MAŁO ZAKŁÓCONEJ

W najogólniejszym przypadku sygnał ze źródła promieniowania dociera do punktu obserwacji dowolną ilością dróg (Rys. 6), przy czym każdy z promieni może ulegać dowolnej ilości odbić. Każdy z tych promieni wytwarza w miejscu obserwacji składową „widzianą” przez antenę odbiorczą E_i , którą opisuje zależność:

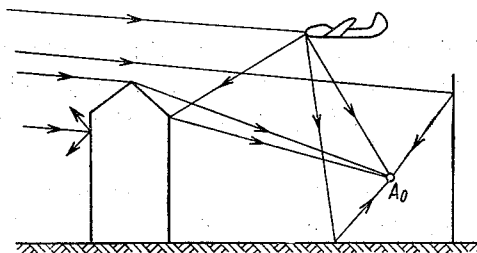
$$E_i = \mathbf{1}_A \cdot \mathbf{1}_i \cdot E_0 \cdot F(\Theta, \varphi) \cdot \frac{e^{-jkr_i}}{r_i} \prod_{j=0}^n R_{ij} \quad (16)$$

gdzie: $\mathbf{1}_A$ i $\mathbf{1}_i$ — wersory płaszczyzny polaryzacji anteny i promienia E_i odpowiednio,

r_i — długość drogi i -tego promienia,

R_{ij} — j -ty współczynnik odbicia i -tego promienia,

$F(\Theta, \varphi)$ — unormowana charakterystyka promieniowania anteny odbiorczej.



Rys. 6. Propagacja wielotorowa

Tymczasem składowa pola elektromagnetycznego w miejscu obserwacji skorelowana z i -tym promieniem, E'_i wynosi:

$$E'_i = E_0 \cdot \frac{e^{-jkr_i}}{r_i} \prod_{j=0}^n R_{ij} \quad (17)$$

Sumując wielkości określone wzorem (17) otrzymamy wypadkowy wektor natężenia pola E' natomiast sumując (po i) wielkości określone wzorem (16) otrzymamy wypadkowe pole E widziane przez antenę. Stosunek E/E' może być miarą błędu pomiaru natężenia pola przy wykorzystaniu anteny o określonej charakterystyce kierunkowej i określonych własnościach polaryzacyjnych.

W najprostszym, mającym praktyczne znaczenie, przypadku do punktu obserwacji docierają dwa promienie: bezpośredni i odbity od powierzchni ziemi. Jeżeli przyjmiemy, że polaryzacja obu promieni jest zgodna, to natężenie pola w punkcie obserwacji można wyznaczyć na podstawie zależności (6). W dużej odległości od źródła promieniowania różnica faz obu promieni zmienia się powoli, a błędy pomiarowe wynikające z kierunkowych własności anteny pomiarowej, dla małych kątów pomiędzy obydwojma promieniami są pomijalnie małe (wynika to także z krzywych pokazanych na Rys. 5). Zatem można stwierdzić, że przy pomiarach natężenia pola w opisanych warunkach wykorzystanie do pomiarów anten kierunkowych nie wprowadza znaczących błędów pomiarowych.

Stwierdzenie powyższe jest bardzo cenne z punktu widzenia zastosowań. Należy jednak podchodzić do niego z dużą ostrożnością, bowiem dotyczy warunków wyidealizowanych tak dotyczących pokrycia i urozmaicenia trasy propagacji jak i polaryzacyjnej zgodności obu promieni — znanym i wykorzystywanym jest zjawisko zmiany polaryzacji sygnału (zwłaszcza przy polaryzacji pionowej) przy propagacji nad ośrodkiem stratnym. Pewnym kryterium poprawności pomiaru może tu być pokrywanie się kierunku do źródła z maksimum promieniowania anteny odbiorczej i maksimum wskazań miernika natężenia pola oraz maksimum wskazań odpowiadających polaryzacji anteny odbiorczej zgodnej z polaryzacją źródła pola.

4.2 POMIAR PÓL SPOLARYZOWANYCH ELIPSOIDALNIE

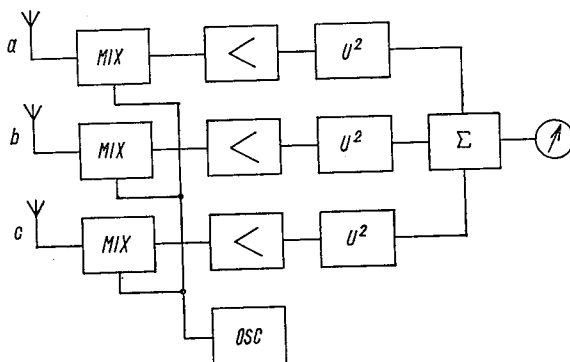
W wielu przypadkach pomiary są prowadzone w terenie urozmaiconym o dużym pokryciu. Sygnał dociera do punktu obserwacji w sposób trudny do ustalenia, niekiedy promienia bezpośredniego w ogóle nie ma. Wskutek odbić od obiektów ruchomych, lub przy wielokrotnych odbiciach sygnałów modulowanych częstotliwościowo może powstać wirowanie płaszczyzny polaryzacji i powstanie sygnału o polaryzacji elipsoidalnej. Istnieje zatem konieczność wykorzystania metod pomiaru pola właściwych dla pola bliskiego [11].

W polu spolaryzowanym eliptycznie lub elipsoidalnie stosunek E/E' może przybierać wartości dowolnie małe w skutek określonych własności kierunkowych i polaryzacyjnych anteny odbiorczej. Należy zatem tak zmodyfikować antenę odbiorczą, aby mogła ona jednakowo reagować na sygnały o dowolnej polaryzacji przestrzennej i dowolnym kierunku wirowania wektora natężenia pola elektrycznego. Koncepcję układu spełniającego powyższe kryterium pokazano na Rys. 7. Układ składa się z trzech synchronicznych torów odbiorczych, z których każdy współpracuje z jedną z trzech wzajemnie prostopadłych anten. Na wyjściu układu znajduje się sumator dołączony do urządzenia wskaźnikowego.

Niech będą dane składowe wektora natężenia pola elektrycznego E_x , E_y , E_z w kartezjańskim układzie współrzędnych:

$$\begin{aligned}
 E_x &= A \cdot \sin \Omega t \\
 E_y &= B \cdot \cos \Omega t \cdot \sin \omega t \\
 E_z &= C \cdot \cos \Omega t \cdot \cos \omega t
 \end{aligned}
 \quad (18)$$

gdzie: Ω — pulsacja przebiegu nośnego,
 ω — pulsacja przebiegu modulującego, niech: $\Omega \gg \omega$.



Rys. 7. Schemat blokowy miernika pola elektromagnetycznego

W środku układu współrzędnych x, y, z znajduje się środek innego kartezjańskiego układu współrzędnych a, b, c utworzonego przez trzy elektrycznie krótkie anteny dipolowe. Napięcia indukowane przez pole w poszczególnych ramionach U_a, U_b, U_c określają zależności:

$$\begin{aligned}
 U_a &= a[A \sin \Omega t \cdot \cos(a, x) + B \cos \Omega t \cdot \sin \omega t \cdot \cos(b, x) + C \cos \Omega t \cdot \cos \omega t \cdot \cos(c, x)] \\
 U_b &= b[A \sin \Omega t \cdot \cos(a, y) + B \cos \Omega t \cdot \sin \omega t \cdot \cos(b, y) + C \cos \Omega t \cdot \cos \omega t \cdot \cos(c, y)] \\
 U_c &= c[A \sin \Omega t \cdot \cos(a, z) + B \cos \Omega t \cdot \sin \omega t \cdot \cos(b, z) + C \cos \Omega t \cdot \cos \omega t \cdot \cos(c, z)]
 \end{aligned}
 \quad (19)$$

gdzie: A, B, C — amplitudy,
 a, b, c — stałe.

Przyjmijmy, że: $A = B = C$, to znaczy mamy szczególny przypadek polaryzacji elipsoidalnej — polaryzację sferyczną. Jeżeli wszystkie tory urządzenia odbiorczego są identyczne to także i $a = b = c$ a wzmocnienie (transmitancja) każdego z torów wynosi K . Przyjmując powyższe założenia i obliczając dla układu przedstawionego na Rys. 7 napięcie wyjściowe sumatora zgodnie z zależnościami (19) otrzymamy:

$$U = K \cdot U_a^2 + K \cdot U_b^2 + K \cdot U_c^2 = a^2 \cdot A^2 \cdot K. \quad (20)$$

Otrzymany wynik stanowi potwierdzenie przydatności zaprezentowanej koncepcji przy pomiarach pól elektromagnetycznych o nieznanymi własnościami polaryzacyjnych.

ZAKOŃCZENIE

Punktem wyjścia niniejszej pracy były wątpliwości związane z przydatnością istniejących wzorców natężenia pola elektromagnetycznego do wzorcowania mierników pola wyposażonych w anteny kierunkowe. W pracy wykazano, że wzorce te mogą być stoso-

wane wprost, bądź też po prostej adaptacji także do wzorcowania mierników z antenami kierunkowymi.

Podstawowym problemem związanym z pomiarami pola miernikami z antenami kierunkowymi jest wielotorowość propagacji sygnału pomiędzy źródłem a punktem obserwacji. W prostych przypadkach propagacji dwu torowej (choć nie zawsze) mogą być stosowane mierniki z dowolnymi antenami. W polach o złożonej konfiguracji czasowo-przestrzennej koniecznym staje się stosowanie anten o sferycznej charakterystyce promieniowania. Koncepcję takiego układu antenowego i elektronicznego zaprezentowano powyżej. Można tu sformułować zastrzeżenia odnośnie przyjętego założenia o elektrycznej długości poszczególnych dipoli, pominięciu wpływu układu zasilającego anteny itp. Oba te czynniki spowodują pewną deformację charakterystyki promieniowania anteny, jednak stosunkowo niewielką i możliwą do skorygowania w stopniu zadowalającym. Podstawową wadą zaprezentowanej koncepcji jest złożoność będąca ceną za jej niewątpliwe zalety i to zarówno w odniesieniu do zastosowań metrologicznych jak i radiokomunikacyjnych.

Jak z powyższego widać istnieją pewne granice w których mogą być wykorzystywane do celów pomiarowych także anteny kierunkowe, niekiedy stosowanie nawet anten dipolowych (czasami traktowanych jako anteny dookólne) będzie ograniczone i koniecznym stanie się stosowanie sposobów i urządzeń bardziej złożonych, jak na przykład zaprezentowana koncepcja. Czy wyczerpuje to wszystkie potrzeby pomiarowe?

Aby odpowiedzieć na to pytanie należałoby sumowanie wielkości określonych wzorami (16) i (17) rozciągnąć także na dziedzinę częstotliwości. Realizacja takiej możliwości byłaby ze wszech miar interesująca a szczególnie istotna dla pomiarów związanych z ochroną środowiska elektromagnetycznego. Istnieją w tej mierze pewne półśrodki, nie są one jednak w pełni zadowalające.

BIBLIOGRAFIA

1. Bem D.J., Teisseyre O.: *Okrętowe urządzenia antenowe* Wyd. Morskie, Gdańsk 1976
2. Zestaw antenowy typu DAMZ-4/50 Instrukcja techniczna, ZZG INCO
3. Bem D.J.: *Anteny i rozchodzenie się fal radiowych* WNT, Warszawa 1973
4. Trzaska H.: *Wzorcowanie mierników natężenia pola elektromagnetycznego*. Prace Inst. Łączn. 1974, vol. 4, s. 3—23
5. Góral A., Grudziński E., Rozwalka K., Trzaska H.: *Wzorzec pola elektromagnetycznego na zakres 300—1000 MHz*. Raport Polit. Wrocław. Nr I-28/R/052/84
6. Crawford M.L.: *Options to open-field and shielded enclosure electromagnetic compatibility measurements*. 1981 Intl. EMC Symp. Proc. s. 383—388
7. Kawana T., Miyajima S.: *Theoretical investigations on site attenuation, propagation characteristics inside the measuring site for the radio interference*. J. of the Radio Res. Labs. vol. 25/1978, s. 105—115
8. Kawana T., Miyajima S.: *Theoretical investigations of site attenuation by means of mutual impedance between antennas, in the case of 3 meters distance between antennas*. J. of the Radio Res. Labs. vol. 26/1979 s. 135—143
9. Borsero M., Nano E.: *Comparison between calculated and measured attenuation of the site recommended by IEC for radiation measurements*. 1981 Intl. EMC Symp. Proc. s. 443—448

10. Micmacher M., Torgowanow W.: *Bezechowyje kamery SWCz Radio i Swiaz'*, Moskwa 1982
11. Trzaska H.: *Pomiary pól elektromagnetycznych w polu bliskim*. Studia i Mat. Monograficzne IMP w Łodzi vol. 4/8/81, s. 168—265

H. TRZASKA

ELECTROMAGNETIC FIELD MEASUREMENTS WITH DIRECTIONAL ANTENNAS

Summary

The applicability of the standard receiving antenna method to calibration of the electromagnetic field strength meters equipped with directional antennas, is considered. It is proven that the method can be applied directly or its accuracy could be improved by the use of relative simple means. The range of full applicability of meters with directional antennas for electromagnetic field measurements is briefly discussed and a concept of a set for measuring arbitrarily complicated fields, eg. for ellipsoidally polarized fields is suggested.

H. TRZASKA

MESURES DES CHAMPS ÉLECTROMAGNÉTIQUES
AVEC LES ANTENNES DIRECTIONNELLES

Résumé

Dans l'article on analyse l'usage de l'antenne réceptrice d'étalonnage pour calibrer les appareils de mesure du champ électromagnétique armés en antennes directionnelles. On a prouvé qu'on peut utiliser cette méthode directement ou avec les moyens préventifs relativement simples. On a discuté brièvement l'étendue de la pleine application aux mesures du champ électromagnétique et on a présenté une conception du circuit de mesure des champs, composés d'une manière quelconque, par ex. les champs à polarisation ellipsoïdale.

H. TRZASKA

ELEKTROMAGNETISCHE FELDSTÄRKEMESSUNGEN MITTELS RICHTUNGSANTENNEN

Zusammenfassung

In dem Beitrag wird die Analyse der Anwendungsmethode für die Standardempfangsantenne zur Kalibrierung der Feldstärkemeßgeräte mittels Richtungsantennen beschrieben. Der Autor hat nachgewiesen, daß es möglich ist, diese Methode direkt oder unter Anwendung relativ einfacher Maßnahmen zu benutzen. Es wurde auch kurz der Bereich der vollen Anwendbarkeit solcher Feldstärkemeßgeräte mit Richtungsantennen diskutiert. Weiterhin präsentiert der Autor eine Netzkonzeption für die Messung beliebig komplizierter Felder, z.B. von Feldern mit einer ellipsoidalen Polarisation.

X. ТРЗАСКА

ИЗМЕРЕНИЯ ЭЛЕКТРОМАГНИТНЫХ ПОЛЕЙ НАПРАВЛЕННЫМИ АНТЕННАМИ

Резюме

Анализируется применение метода образцовой приёмной антенны для калибровки измерителей напряженности электромагнитного поля снабженных направленными антеннами. Доказано, что метод может применяться непосредственно, а его точность можно повышать применяя относительно простые средства. Кратко рассмотрен диапазон полной применимости измерителей снабженных направленными антеннами для измерения электромагнитного поля и представлена идея устройства для измерений довольно сложных полей, н-пр. полей с эллипсоидальной поляризацией.

Zastosowanie kryterium kosztów do oceny pewnego systemu przesyłania informacji z decyzyjnym sprzężeniem zwrotnym

WIESŁAW LUDWIN

Katedra Telekomunikacji, Akademia Górniczo-Hutnicza, Kraków

Otrzymano 1987.07.08

Autoryzowano do druku 1987.11.17

Rozważono system przesyłania informacji cyfrowych z decyzyjnym sprzężeniem zwrotnym. Przebadano wg kryterium ekonomicznego koszty transmisji informacji w tym systemie. W oparciu o zaproponowany model matematyczny systemu i kryterium kosztów, skonstruowano zadanie poszukiwania minimum funkcji kosztów transmisji z ograniczeniami. Przebadano różne warianty rozwiązania tak postawionego zadania. Mając na uwadze ekonomiczne aspekty budowy i eksploatacji systemów przesyłania informacji cyfrowych pokazano, że możliwe jest znalezienie optymalnych systemów transmisji informacji w sensie proponowanego kryterium ekonomicznego.

1. WSTĘP

Tematem przedstawionej pracy jest próba oceny ekonomicznej pewnego systemu przesyłania informacji cyfrowych. Problemy optymalizacji i oceny porównawczej systemów przesyłania informacji oparte o ujednolicone kryteria efektywności i kryterium ekonomiczne rozważał w swojej pracy [3] E. G. Serwinski. Przeprowadził on szczegółową ocenę systemów przesyłania informacji cyfrowych kodowanych kodem beznadmiarowym. Natomiast w prezentowanej pracy rozważono prosty system z decyzyjnym sprzężeniem zwrotnym i kodem nadmiarowym oraz przebadano wg kryterium zaczerpniętego z pracy [3] koszty transmisji informacji w tym systemie.

2. MODEL SYSTEMU

System przesyłania informacji cyfrowych stanowiący przedmiot rozważań niniejszej pracy jest prostym systemem z decyzyjnym sprzężeniem zwrotnym, z oczekiwaniem na potwierdzenie pojedynczej wiadomości. Wiadomości kodowane są kodem nadmiarowym o nieparzystej odległości minimalnej, pozwalającym wykrywać, ewentualnie korygować

skończoną liczbę błędów transmisji. Kanał ziarnisty jest bezpamięciowym, symetrycznym kanałem binarnym o stałych parametrach. Na kanał ten oddziałuje zakłócenie addytywne w postaci szumu białego. Ciąg kodowy moduluje częstotliwość sinusoidalnego drgania nośnego i jest to modulacja z aktywną przerwą. Sygnały na wejściu kanału podkładowego mają równe energie i są ortogonalne w znaczeniu wzmocnionym. Sposób odbioru jest optymalną niekoherentną detekcją. Zapewniona jest synchronizacja rozeznania i dekodowania. Kanał sprzężenia zwrotnego jest kanałem idealnym.

Model matematyczny wyżej opisanego systemu przesyłania informacji cyfrowych powinien uwzględniać podstawowe parametry systemu, a to:

R — prędkość transmisji informacji,

P — prawdopodobieństwo błędnego odbioru wiadomości,

x — prawdopodobieństwo błędu elementarnego w kanale transmisyjnym,

B — pasmo częstotliwości zajmowane przez sygnały w systemie, dalej nazywane pasmem,

C — stosunek mocy średniej sygnału użytecznego do widmowej gęstości mocy szumu na wejściu odbiornika, dalej nazywany stosunkiem mocy sygnał/szum.

Prędkość transmisji informacji możemy wyrazić za pomocą zależności:

$$R = \frac{k}{\bar{T}}, \quad (2.1)$$

gdzie k jest liczbą pozycji informacyjnych użytego kodu nadmiarowego, a $\bar{T}$ jest średnim czasem (wyrażonym w odstępach jednostkowych modulacji) trwania jednej transmisji oraz l retransmisji, po których nastąpi bezbłędne odebranie ciągu kodowego o długości $n > k$. Przez transmisję należy rozumieć przesłanie w kanale podkładowym jednej wiadomości wraz z jej potwierdzeniem odbioru w kanale sprzężenia zwrotnego. Jeżeli założyć, że czas trwania elementarnego symbolu kodowego wynosi T , a czas upływający od nadania ciągu kodowego do odebrania sygnału potwierdzenia przez stację nadawczą wynosi T_0 , to:

$$\bar{T} = (nT + T_0) \sum_{l=1}^{\infty} l \cdot y^{l-1} (1-y), \quad (2.2)$$

gdzie prawdopodobieństwo błędnej transmisji równe jest:

$$y = 1 - (1-x)^n. \quad (2.3)$$

W wyniku przekształcania wzorów (2.2) i (2.3) [4] otrzymamy:

$$\bar{T} = \frac{nT + T_0}{(1-x)^n}. \quad (2.4)$$

Po podstawieniu (2.4) do (2.1), prędkość transmisji informacji w rozważanym przez nas systemie wyniesie:

$$R = \frac{k(1-x)^n}{nT + T_0}. \quad (2.5)$$

Prawdopodobieństwo błędnego odbioru wiadomości oszacujemy za pomocą następującego wzoru [3]:

$$P \leq \sum_{j=2s+1}^n \frac{C_n^{j-s}}{C_j^s} \left(\frac{x}{1-x} \right)^j, \quad (2.6)$$

w którym wielkość P może stanowić parametr porównawczy dla systemów stosujących różne długości ciągów kodowych w przypadku, gdy spełniona jest nierówność $n \cdot x \ll 1$. Wielkość s występująca we wzorze (2.6) jest maksymalną gwarantowaną krotnością błędów korygowanych przez użyty kod nadmiarowy. Prawdopodobieństwo x wystąpienia błędu elementarnego w kanale podkładowym rozważanego przez nas systemu wynosi [3]:

$$x = \frac{1}{2} e^{-\frac{1}{2} C \cdot T}. \quad (2.7)$$

Stosunek mocy sygnał/szum oznaczony przez C może przyjmować wartości z przedziału:

$$0 \leq C \leq C_{max}. \quad (2.8)$$

Pasmo częstotliwości zajmowane przez sygnał niosące informacje i sygnał potwierdzenia oszacujemy wzorem [2]:

$$B = \frac{4}{T} + B_0, \quad (2.9)$$

gdzie B_0 jest pasmem częstotliwości zajmowanym przez sygnał potwierdzenia odbioru przesłanej wiadomości. Pierwszy ze składników wzoru (2.9) na pasmo B , reprezentujący pasmo częstotliwości zajmowane przez sygnały informacyjne w systemie, jest co prawda daleko idącym przybliżeniem, jednakże dość powszechnie stosowanym przy rozpatrywaniu teoretycznych modeli systemów przesyłania informacji [2]. Jako pasmo sygnałów informacyjnych przyjmuje się pasmo zawarte między dwiema częstotliwościami zerowania się widma sygnału zmodulowanego, najbliższymi częstotliwości środkowej. Zakładając, że pasmo częstotliwości B zajmowane przez sygnały w systemie jest ograniczone, możemy zapisać następujące nierówności:

$$B_{min} \leq B \leq B_{max}. \quad (2.10)$$

Wykorzystując równania (2.5) do (2.10) możemy przedstawić model matematyczny rozważanego przez nas systemu za pomocą następującego układu równań i nierówności:

$$\left. \begin{aligned} R &= \frac{k(1-x)^n}{nT + T_0} \\ x &= \frac{1}{2} e^{-\frac{1}{2} C \cdot T} \\ B &= \frac{4}{T} + B_0 \\ P &= \sum_{j=2s+1}^n \frac{C_n^{j-s}}{C_j^s} \left(\frac{x}{1-x} \right)^j \\ B_{min} &\leq B \leq B_{max} \\ 0 &\leq C \leq C_{max}. \end{aligned} \right\} \quad (2.11)$$

Model ten pozwoli nam przeprowadzić próbę oceny kosztów wybranego systemu przesyłania informacji cyfrowych z decyzyjnym sprzężeniem zwrotnym.

3. KRYTERIUM OCENY SYSTEMU

System przesyłania informacji cyfrowych, którego model matematyczny zapisano równaniami (2.11) poddany zostanie ocenie w oparciu o kryterium kosztów postaci [3]:

$$K = K_E + K_B + K_C, \quad (3.1)$$

gdzie: $K_E = f(n, k, s)$ — koszty kodowania i dekodowania,

$K_B = w_B \cdot B$ — koszty pasma częstotliwości,

$K_C = w_C \cdot C$ — koszty stosunku mocy sygnał/szum.

Koszty pasma K_B i koszty mocy K_C są liniowymi funkcjami pasma częstotliwości B i stosunku mocy sygnał/szum C , przy czym współczynniki w_B , w_C mają interpretację kosztów jednostkowych odpowiednio pasma i stosunku mocy sygnał/szum. Koszty kodowania i dekodowania K_E są tylko i wyłącznie funkcją parametrów wybranego kodu nadmiarowego i reprezentują one koszt części cyfrowych rozważanego systemu. Natomiast łączne koszty pasma i mocy stanowią tę część kosztów całkowitych, którą nazywać będziemy kosztami części analogowych systemu. Na podstawie obserwacji współczesnych tendencji dotyczących wzajemnego kształtowania się kosztów części cyfrowych i kosztów części analogowych systemów transmisji informacji [1], można przyjąć, że $K_E \ll (K_B + K_C)$. Tym samym udział kosztów kodowania i dekodowania wiadomości w całkowitych kosztach systemu przesyłania informacji jest prawie zawsze do pominięcia.

W interesującym nas przypadku funkcję kosztów (3.1) zapiszemy, wykorzystując równania (2.5), (2.7) i (2.9), jako funkcję wielu zmiennych postaci:

$$\begin{aligned} K(n, k, s, B_0, T_0, R, x, w_B, w_C) = \\ = f(n, k, s) + w_B B_0 + \frac{2nR}{k(1-x)^n - T_0 R} (2 \cdot w_B - w_C \ln 2x). \end{aligned} \quad (3.2)$$

Poszukiwanie najtańszego systemu w ramach wybranej klasy systemów z decyzyjnym sprzężeniem zwrotnym można sprowadzić do zadania poszukiwania minimum funkcji wielu zmiennych (3.2) przy ograniczeniach (2.11). My jednakże uprościmy to zadanie do problemu poszukiwania minimum funkcji jednej zmiennej z ograniczeniami, przyjmując jednocześnie pozostałe zmienne jako parametry rozważanego systemu. Niech prędkość transmisji informacji $R = R_c$, gdzie R_c jest z góry ustaloną wartością tej prędkości. Niech prawdopodobieństwo błędnego odbioru P nie będzie gorsze niż pewna żądana wartość tego prawdopodobieństwa P_c , tzn. $P \leq P_c$. Załóżmy dalej, że czas odbioru T_0 , a tym samym i pasmo częstotliwości B_0 sygnału potwierdzenia są także ustalone. W tej sytuacji interesującą wydaje się być odpowiedź na pytanie, jak przy przyjętych ustaleniach oraz dla wybranego kodu nadmiarowego o parametrach n , k , s zachowuje się ze względu na argument x funkcja kosztów (3.2) przy różnych wzajemnych relacjach kosztów jednostkowych pasma w_B i stosunku mocy sygnał/szum w_C . Ciekawą rzeczą będzie też sprawdzenie wpływu wybranego kodu nadmiarowego na całkowite koszty systemu transmisji informacji.