TECHNICAL UNIVERSITY – SOFIA
Department Theoretical Electrical Engineering

8TH SUMMER SCHOOL

ADVANCED ASPECTS
OF THEORETICAL ELECTRICAL ENGINEERING

Sozopol'10

Part 1. Plenary Lectures

Edited by Valeri Mladenov

in the framework of

DAYS OF SCIENCE
OF THE TECHNICAL UNIVERSITY OF SOFIA

Sozopol'10, BULGARIA, 19-22.IX.2010

TECHNICAL UNIVERSITY – SOFIA
Department Theoretical Electrical Engineering

8TH SUMMER SCHOOL

ADVANCED ASPECTS
OF THEORETICAL ELECTRICAL ENGINEERING

Sozopol'10

Part 2. Regular papers

Edited by Valeri Mladenov

in the framework of

DAYS OF SCIENCE
OF THE TECHNICAL UNIVERSITY OF SOFIA

Sozopol'10, BULGARIA, 19-22.IX.2010

ORGANIZATION

The Summer School is organized by the Department of Theoretical Electrical Engineering at the Technical University of Sofia in the framework of the “Days of Science of the Technical University of Sofia”, Sozopol, Bulgaria, September 2010

TECHNICAL **U**NIVERSITY OF **S**OFIA, **B**ULGARIA

under the patronage of the INTERNATIONAL SYMPOSIUM ON THEORETICAL ELECTRICAL ENGINEERING (ISTET) and it is a regular ISTET event

INTERNATIONAL **S**YMPOSIUM ON **T**HEORETICAL **E**LECTRICAL ENGINEERING (**ISTET**)

SUPPORT

The main sponsor of the Summer School is:

RESEARCH & **D**EVELOPMENT **S**ECTOR, **T**ECHNICAL **U**NIVERSITY – **S**OFIA

Other sponsors:

IEEE BULGARIA CAS CHAPTER

THE WORLD SCIENTIFIC AND ENGINEERING ACADEMY AND SOCIETY (WSEAS)

TECHNICAL UNIVERSITY OF SOFIA

FACULTY AUTOMATION

DEPARTMENT THEORETICAL ELECTRICAL ENGINEERING

8th SUMMER SCHOOL

ADVANCED ASPECTS OF THEORETICAL ELECTRICAL ENGINEERING – SOZOPOL'10

in the framework of

THE DAYS OF SCIENCE OF THE TECHNICAL UNIVERSITY OF SOFIA, SOZOPOL, BULGARIA, SEPT. 2010

Part I: Plenary Lectures

Edited by Valeri Mladenov

ISSN: 1313-9479

Organizing Committee

Honorary Chairmen:

V. Georgiev – Bulgaria
S. Farchy – Bulgaria
L. Kolev – Bulgaria
S. Papasow – Bulgaria
V. Savov – Bulgaria

Chair:

V. Mladenov, TU-Sofia, Bulgaria

Co Chairs:

S. Guninski, TU-Sofia, Bulgaria
J. Georgiev, TU-Sofia, Bulgaria
I. Uzunov, TU-Sofia, Bulgaria

Members:

K. Brandisky, TU-Sofia, Bulgaria
A. Chervenkov, TU-Sofia, Bulgaria
K. Ivanov, TU-Sofia, Bulgaria
N. Petkova, TU-Sofia, Bulgaria
S. Petrakieva, TU-Sofia, Bulgaria
N. Radev, TU-Sofia, Bulgaria
I. Tabahnev, TU-Sofia, Bulgaria
S. Terzieva, TU-Sofia, Bulgaria
K. Todorova, TU-Sofia, Bulgaria
G. Toshev, TU-Sofia, Bulgaria
S. Vladov, TU-Sofia, Bulgaria
I. Trushev, TU-Sofia, Bulgaria
G. Tsenov, TU-Sofia, Bulgaria
I. Yacheva, TU-Sofia, Bulgaria

International Programme Committee

Honorary Chairmen:

L. Chua, Berkeley-USA
N. Mastorakis, Greece
R. Sikora, Poland

Members:

D. Baldomir, Spain
A. Bossavit, France
A. Brykalski, Poland
H. Butterweck, Netherlands
T. Chady, Poland
K. Demirtchyan, Russia
S. Gratkowski, Poland
K. Hameyer, Germany
W. John, Germany
L. Klinkenbusch, Germany
L. Kolev, Bulgaria
A. Kost, Germany
Z. Leonowicz, Poland
W. Mathis, Germany
V. Mladenov, Bulgaria
M. Ogorzałek, Poland
S. Osowski, Poland
L. Pichon, France
B. Reljin, Serbia
L. Sumichrast, Slovakia
J. Sykulski, UK
M. Tadeusiewicz, Poland
R. Tetzlaff, Germany
H. Toepfer, Germany
H. Uhlmann, Germany
R. Weigel, Germany

PREFACE

These Proceedings are organized in two parts and contain the plenary lectures and the regular papers presented at the 8th Summer School *Sozopol'10,* which took place in Sozopol, Bulgaria, between 19 and 22 Sept. 2010 in the framework of the "Days of the Science of the Technical University of Sofia". The Summer School covers the advanced aspects of Theoretical Electrical Engineering and it is a platform for post-graduate training of Ph.D. students and young scientists. During the Summer School well-known experts presented some advanced aspects of circuits and systems theory, electromagnetic field theory and their applications. Apart from the educational part of the Summer School a presentation of original authors' papers took place.

The main topics of the Summer School *Sozopol'10* include Circuits and Systems Theory, Signal Processing and Identification Aspects, Electromagnetic Fields, Theoretical Concepts, Applications and Innovative Educational Aspects.

The Summer School *Sozopol'10* has been organized by the Department of Theoretical Electrical Engineering of the Technical University of Sofia with the main sponsorship of the Research and Development Sector of the Technical University of Sofia. This has been the eighth edition of the event, after the Summer Schools in 1986, 1988, 2001, 2002, 2005, 2007 and 2009. The Summer School is under the patronage of the International Symposium on Theoretical Electrical Engineering (ISTET) and it is a regular ISTET event.

There were 56 participants from 5 different countries at the Summer School this year. There were 13 plenary lectures and 30 regular papers that are published in these Proceedings. Providing the recent advances in Theoretical Electrical Engineering the Proceedings will be of interest to all researchers, educators and Ph.D. students in the area of Electrical Engineering.

Special thanks are due to the Research and Development Sector, Faculty of Automation and the Section of Social Services of the Technical University of Sofia about the overall support of the event. We also want to thank to the IEEE Bulgaria CAS Chapter and the World Scientific and Engineering Academy and Society (WSEAS), which partially sponsored the event. We hope to meet again in the following edition of the Summer School to continue the good tradition and collaboration in the field of Theoretical Electrical Engineering.

Organizing Committee
Sofia, Oct. 2010

CONTENTS

SYSTEMATIC WRITING OF STATE EQUATIONS FOR CIRCUITS WITH EXCESS ELEMENTS

Paul Dan Cristea, Rodica Tuduce

Biomedical Engineering Center, University "Politehnica" of Bucharest,
Spl.Independentei 313, 060042 Bucharest, Romania,
phone: +40 21 316 9569, e-mail: {pcristea, trodica}@dsp.pub.ro

Abstract: *A circuit has excess elements when it contains capacitor loops – closed contours containing only ideal capacitors and (possibly) ideal independent voltage sources, and/or inductor cut-sets – closed surfaces cutting only ideal inductors and (possibly) ideal independent current sources. For such circuits it is not possible to chose all capacitor voltages and all inductor currents as state variables, as they are not mutually independent. Therefore, sub-sets of essential capacitor voltages and essential inductor currents are selected as state variables. The remaining capacitor voltages and inductor currents are considered dependent variables, as they can be expressed in terms of the essential variables.*

Keywords: State variable equations, Systematic equation formulation, Excess circuit elements, Essential circuit elements

1. INTRODUCTION

The state variables method is a valuable tool for the description and analysis of large classes of systems, in particular electrical circuits [1], [2]. One intrinsic advantage is that the state variables method is a natural time domain method, with direct computer (numeric) implementations for linear, parametric or nonlinear systems.

In the case of linear time invariant systems (LTISs), the closed form solution of the state variable equations naturally provides separate terms for the "free component" – the component determined by the system's initial conditions, and for the "forced component" of the system solution – the one determined by the excitations. For LTISs excited by a set of combined constant (DC) or harmonic (AC) inputs (*e.g.*, for electric circuits, independent voltage and current sources and/or, equivalently, voltages and currents applied at circuit terminals), superposition can be used to combine the state variable approach with specific DC and/or AC methods. One uses the DC and/or AC methods to separately and efficiently compute the forced component of the system solution, while the state variable method is used only to determine the free component.

For non-linear or parametric systems, small signal linear methods can be used around each stationary point of functioning to determine the instantaneous behavior and evolution of the system in a short enough time interval.

The state equations of a LTIS have the standard form:

$$\dot{x}\rceil = \lceil A \rceil x\rceil + \lceil B \rceil u\rceil \tag{1}$$

where $x\rceil \in R^{n\times 1}$ is the (column) state variables vector, $u\rceil \in R^{m\times 1}$ is the (column) excitation vector (sources and inputs), $[A]\in R^{n\times n}$ is the system coefficient matrix and $[B]\in R^{n\times m}$ – the input coefficient matrix[1].

The corresponding solution for LTIS is given by:

$$x(t)\rceil = e^{[A](t-t_{0+})}\, x(t_{0+})\rceil + \int_{t_0}^{t} e^{[A](t-t_{0+})}\, [B]\; u(\tau)\rceil d\tau \qquad (2)$$

where $x(t)\rceil$ is the state variable vector at a moment t, $(t_0 \le t < \infty)$, whereas $x(t_{0+})\rceil$ is its initial value, at the moment

For electric circuits, a natural choice of the state variables is the vector of all capacitor voltages and all inductor currents:

$$x\rceil = \begin{matrix} u_C\rceil \\ i_L\rceil \end{matrix}\Bigg\rceil, \qquad (3)$$

where $u_C\rceil$ is the vector of capacitor voltages and $i_L\rceil$ - the vector of inductor currents. This choice is based on the hypothesis that these variables are mutually independent.

Correspondingly, the state co-variables is the vector which contains the complementary variables:

$$X\rceil = \begin{matrix} I_C\rceil \\ U_L\rceil \end{matrix}\Bigg\rceil, \qquad (4)$$

where $I_C\rceil$ is the vector of capacitor currents and $U_L\rceil$ - the vector of inductor voltages.

The excitations for an electrical circuit are given by the vector:

$$u(t)\rceil = \begin{matrix} e(t)\rceil \\ j(t)\rceil \end{matrix}\Bigg\rceil, \qquad (5)$$

comprising $e(t)\rceil$– the vector of the *emf*-s of all independent voltage sources (equivalently, the possible independent voltages applied to the circuit) and $j(t)\rceil$ – the vector of the currents of the independent current sources (equivalently, the possible independent currents injected into the circuit).

One implicit condition when selecting the state variables $x\rceil$ of a system is to be able to compute all system functions, such as the total energy of the system, in function of $x\rceil$ and system parameters.

The total energy stored in the reactive elements of an electric circuit, *i.e.*, in its capacitors and inductors, can readily be expressed in terms of the state variables:

[1] Notice the notation with double square brackets for square or rectangular matrices, and with a single square bracket placed on the right for column matrices (vectors). This notation has the advantage of suggesting directly the shape of a matrix(square/rectangular or column).

$$w = w_e + w_m = \sum_{k=1}^{n_C} \frac{C_k u_{C_k}}{2} + \sum_{k=1}^{n_L} \frac{L_k i_{L_k}}{2} + \frac{1}{2}\sum_{j=1}^{n_L}\sum_{k=1}^{n_L} L_{jk}\, i_{L_j} i_{L_k} \quad . \tag{6}$$

By defining the matrices of capacitances, inductances and reactive parameters:

$$[C] = \mathrm{diag}\left(C_1, C_2, \cdots, C_{n_C}\right),\ [L] = \begin{bmatrix} L_1 & L_{12} & \cdots & L_{1n_L} \\ L_{21} & L_2 & \cdots & L_{2n_L} \\ \vdots & \vdots & \ddots & \vdots \\ L_{n_L 1} & L_{n_L 2} & \cdots & L_{n_L} \end{bmatrix} \text{ and } [P] = \begin{bmatrix} [C] & \\ & [L] \end{bmatrix}, \tag{7}$$

the equation (5) takes the more compact matrix form:

$$w = w_e + w_m = \frac{1}{2} u_C]^{\mathrm{T}} [C]\, u_C] + \frac{1}{2} i_L]^{\mathrm{T}} [L]\, i_L] = \frac{1}{2} x]^{\mathrm{T}} [P]\, x] \tag{8}$$

In the following, we show how the state variable equations (1), for circuits without or with excess elements, can be established systematically, by reducing the problem to the computing of transfer functions in a resistive multiport [3].

2. STATE VARIABLE EQUATIONS OF CIRCUITS WITHOUT EXCESS ELEMENTS

If the circuit has neither *capacitor loops* (loops consisting only of ideal capacitors and, possibly, of ideal independent voltage source), nor *inductor cut-sets* (closed surfaces cutting only ideal inductors and, possibly, ideal independent current sources), all $x]$ variables defined above are mutually independent from the interconnection point of view and it is possible to chose the state variables and co-variables as shown in equations (3) and (4). In this case, all capacitors and inductors are called *essential circuit elements*, and there are no *excess circuit elements*.

To systematically establish the state variable equations of such a circuit without excess elements, let us consider the circuit symbolically represented as in Fig. 1, in which the capacitors, inductors, ideal voltage sources and ideal current sources are shown connected to the terminals of a resistive passive multiport. By construction, this resistive multiport contains no reactive elements and no independent sources (in general, excitations), but it might contain any type of (resistive) controlled sources (voltage controlled voltage sources, current controlled voltage sources, voltage controlled current sources, and current controlled current sources)

It is a simple resistive circuit problem to compute the state co-variables $I_C]$ and $U_L]$ from the equations of the multiport, in terms of the state variables $u_C]$ and $i_L]$, and of the independent voltage and current sources parameters $e(t)]$ and $j(t)]$, respectively. As the state variables, $u_C]$ and $i_L]$, are independent variables, they can

be considered as the parameters of ideal voltage and ideal current sources connected at multiport terminals, as shown in Fig. 2.

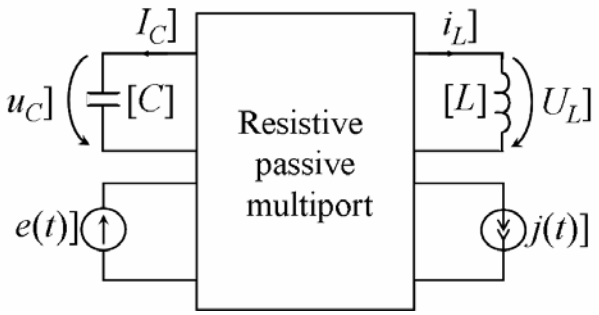


Fig. 1. Electric circuit represented as a resistive passive multiport, possibly containing controlled sources, at which terminals are connected all the capacitors, inductors and independent sources in the circuit.

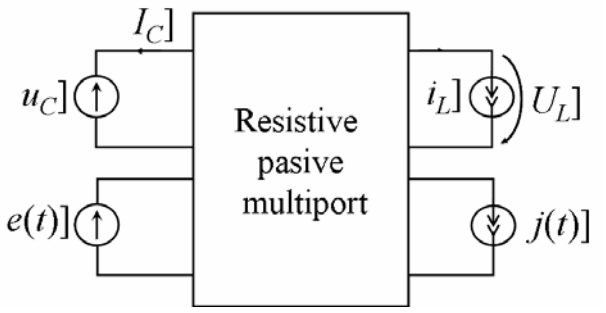


Fig. 2. The circuit in Fig. 1, in which the focus is on computing the currents of all capacitors and the voltages of all inductors, when knowing the excitations at multiport terminals

The state co-variables are computed as the output variables at multiport terminals:

$$\begin{bmatrix} I_C] \\ U_L] \end{bmatrix} = \begin{bmatrix} G_{CC} & B_{CL} \\ A_{LC} & R_{LL} \end{bmatrix} \begin{bmatrix} u_C] \\ i_L] \end{bmatrix} + \begin{bmatrix} G_{Ce} & B_{Cj} \\ A_{Le} & R_{Lj} \end{bmatrix} \begin{bmatrix} e] \\ j] \end{bmatrix}, \tag{9}$$

or, more compactly:

$$X] = [a]\, x] + [b]\, u], \tag{10}$$

where the hybrid transfer function matrices [a] and [b] contain the transfer conductances G, transfer resistances R, voltage gains A and current gains B, explicitly shown in (8).

The capacitors and inductors are described by their constitutive equations:

$$I_C] = \dot{q}_C] = [C]\, \dot{u}_C] \text{ and } U_L] = \dot{\varphi}_L] = [L]\, \dot{i}_L], \tag{11}$$

which allow to express the state co-variables vector $X]$ in function of the time derivative of the state variables vector $\dot{x}]$, and the reverse:

$$X] = [P]\, \dot{x}] \text{ and } \dot{x}] = [P]^{-1} [X]. \tag{12}$$

Combining (12) with (10), one finally gets the state variable equations (1):

$$\dot{x}] = [P]^{-1} [X] = [P]^{-1} ([a]\, x] + [b]\, u]) = [A]\, x] + [B]\, u], \tag{13}$$

in which the system coefficient matrix [A] and the input coefficient matrix [B] are given by:

$$A] = [P]^{-1} [a] \text{ and } B] = [P]^{-1} [b] \tag{14}$$

3. STATE VARIABLE EQUATIONS OF CIRCUITS WITH EXCESS ELEMENTS

Consider now a circuit in which there are *capacitor loops* and/or *inductor cut-sets*. In this case, there is a link between the capacitor voltages, on one hand, and between the inductor currents, on the other, imposed by the circuit topology (from the circuit interconnection point of view). One capacitor voltage can be expressed in terms of the other capacitor voltages for each independent *capacitor loop*. Such a voltage must be eliminated from the vector $u_C]$ of independent capacitor voltages and be introduced in a vector $U_{C'}]$ comprising dependent capacitor voltages. Similarly, one inductor current can be expressed in terms of the other inductor currents for each independent *inductor cut-set.* This current must be eliminated from the vector $i_L]$ of independent inductor currents and be introduced in a vector $I_{L'}]$ comprising the dependent inductor currents.

The state variables vector $x]$ has the same structure as in (3), but comprises only independent capacitor voltages and independent inductor currents, which belong to the *essential circuit elements*. The vectors of dependent variables $U_{C'}]$ and $I_{L'}]$ are parts of the *excess circuit element co-variable vector* $X']$, as shown in (15).

The variables and co-variables of the *essential* and *excess* elements are:

$$x] = \begin{bmatrix} u_C] \\ i_L] \end{bmatrix}, \quad X] = \begin{bmatrix} I_C] \\ U_L] \end{bmatrix}; \quad x'] = \begin{bmatrix} i_{C'}] \\ u_{L'}] \end{bmatrix}, \quad X'] = \begin{bmatrix} U_{C'}] \\ I_{L'}] \end{bmatrix} \tag{15}$$

Consider now the circuit represented in Fig. 3, in which the capacitors, inductors, ideal voltage sources and ideal current sources are shown connected to the terminals of a resistive passive multiport, which can contain any type of controlled sources. In this case, there are two types of capacitors and inductors, represented distinctly. To reduce the problem to solving a resistive multiport, the reactive elements will be replaced by ideal voltage or current sources, giving the same inputs to the resistive multiport, as shown in Fig. 4. The essential capacitors have independent voltages $u_C]$ that can be chosen arbitrarily, so that these elements will be replaced by ideal voltage sources giving the same voltage at the terminals. This is not possible for the excess capacitors, for which the voltages $U_{C'}]$ depend on the voltages $u_C]$. But the currents of these elements have no interconnection restrictions, so that the excess capacitors can be replaced by ideal current sources giving the same currents $i_{C'}]$. Similarly, the essential inductors are replaced with ideal current sources giving $i_L]$, whereas the excess inductors are replaced with ideal voltage sources giving $u_{L'}]$.

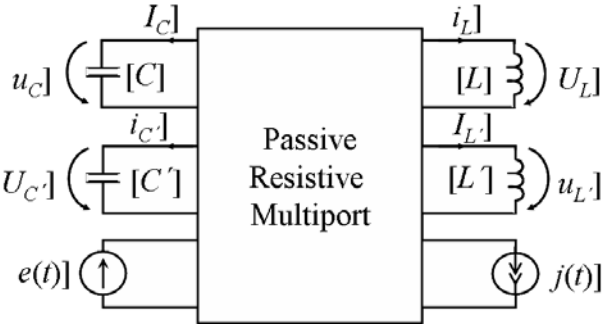


Fig. 3. Electric circuit with excess elements represented as a resistive passive multiport, possibly containing controlled sources, having connected at its terminals all the essential and excess reactive elements (capacitors and inductors), as well as all the independent sources of the circuit.

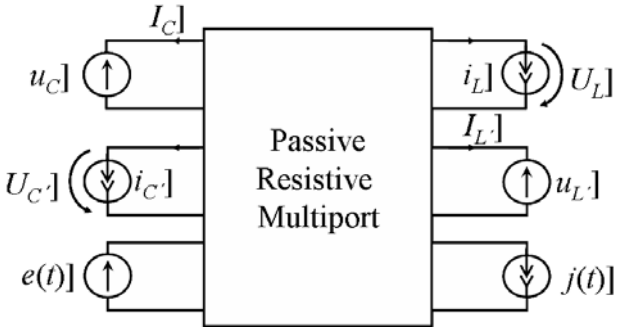


Fig. 4. The electric circuit in Fig. 3, in which the capacitors and inductors have been replaced by ideal sources as explained in the text. The problem has been reduced to the solving of a resistive multiport, possibly containing controlled sources.

Solving the resistive circuit, the dependent variables at the terminals of the resistive multiport result:

$$\begin{bmatrix} I_C \\ U_L \end{bmatrix} = \begin{bmatrix} G_{CC} & B_{CL} \\ A_{LC} & R_{LL} \end{bmatrix} \begin{bmatrix} u_C \\ i_L \end{bmatrix} + \begin{bmatrix} G_{Ce} & B_{Cj} \\ A_{Le} & R_{Lj} \end{bmatrix} \begin{bmatrix} e \\ j \end{bmatrix} + \begin{bmatrix} B_{CC'} & B_{CL'} \\ R_{LC'} & A_{LL'} \end{bmatrix} \begin{bmatrix} i_{C'} \\ u_{L'} \end{bmatrix} \tag{16}$$

$$\begin{bmatrix} U_{C'} \\ I_{L'} \end{bmatrix} = \begin{bmatrix} A_{C'C} & 0 \\ 0 & B_{L'C} \end{bmatrix} \begin{bmatrix} u_C \\ i_L \end{bmatrix} + \begin{bmatrix} A_{C'e} & 0 \\ 0 & B_{L'j} \end{bmatrix} \begin{bmatrix} e \\ j \end{bmatrix} \tag{17}$$

Or, written more compactly:

$$[X] = [a][x] + [b][u] + [c][x'], \tag{18}$$

$$[X'] = [d][x] + [f][u]. \tag{19}$$

The matrices $[d]$ and $[f]$ can have only elements with the values +1, -1 and 0, as they correspond to links between capacitor voltages belonging to the same capacitor loop, or to links between inductor currents belonging to the same inductor cut-set.

By defining the parameter matrices for the essential and the excess elements:

$$[P] = \begin{bmatrix} [C] & \\ & [L] \end{bmatrix} \text{ and } [P'] = \begin{bmatrix} [C'] & \\ & [L'] \end{bmatrix}, \tag{20}$$

the reactive elements' constitutive equations can be written:

$$[X] = [P][\dot{x}] \text{ and } [x'] = [P'][\dot{X}']. \tag{21}$$

Reversing the first equation above and replacing the derivative of (19) in the second, one gets:

$$[\dot{x}] = [P]^{-1}[X] \text{ and } [\dot{X}'] = [d][\dot{x}] + [f][\dot{u}] \tag{22}$$

From the equations (18), (19), (21), and (22) it results successively:

$$\begin{aligned}\dot{x}] &= [P]^{-1} X\ ] \\ &= [P]^{-1} \{[a]\ x] + [b] u] + [c][x']\} \\ &= [P]^{-1} \{[a]\ x] + [b] u] + [c][P'] ([d]\ \dot{x}] + [f]\ \dot{u}])\}\end{aligned} \tag{23}$$

from where the terms containing $\dot{x}]$ can be separated:

$$([P] - [c][P'][d])\ \dot{x}] = ([a]\ x] + [b] u] + [c][P'][f]\ \dot{u}])\ . \tag{24}$$

Defining the *equivalent reactive parameters* matrix:

$$[\Pi] = [P] - [c][P'][d],\ . \tag{25}$$

the state variable equations for the circuit with excess elements is obtained:

$$\dot{x}] = [\Pi]^{-1} ([a]\ x] + [b] u] + [c][P'][f]\ \dot{u}])\ . \tag{26}$$

This equation can be put in the standard form (1), with:

$$[A] = [\Pi]^{-1} [a]\,, \qquad [B] = [\Pi]^{-1} [\,[b] \quad [c][P'][f]\,] \tag{27}$$

and replacing $u(t)]$ with a new vector of excitation:

$$U(t)] = \begin{bmatrix} u(t)] \\ \dot{u}(t)] \end{bmatrix}\ , \tag{28}$$

which comprises both the vector of the sources and its *time derivative*.

4. EXAMPLE

In this section we illustrate the method of systematically establishing the state variable equations of a circuit with excess elements presented in Section 3, by considering the circuit in Fig. 5.

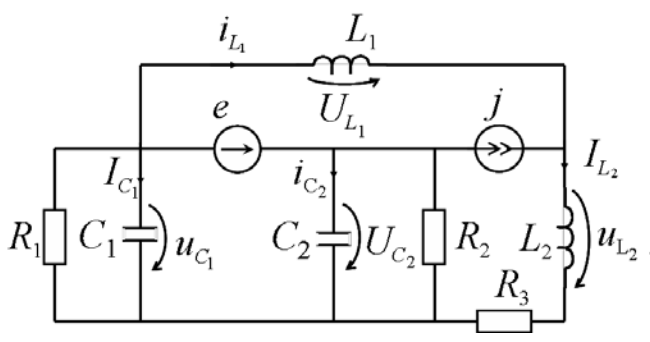


Fig. 5. Example of electric circuit with *essential* (C_1, L_1) and *excess* (C_2, L_2) elements.

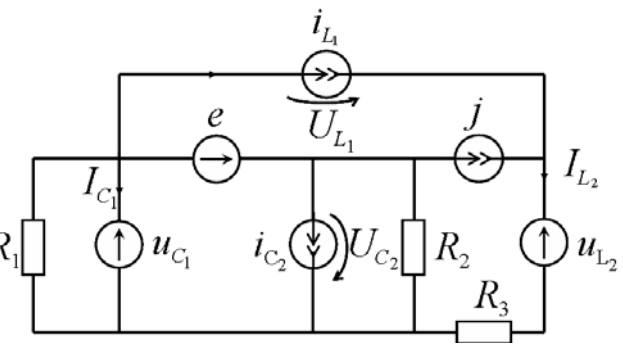


Fig. 6. Resistive multiport of the circuit in Fig. 5, after replacing the capacitors and inductors with ideal independent sources, as explained in the text.

The circuit comprises the capacitor loop ($C_1 - e - C_2$) and the inductor cut-set ($L_1 - L_2 - j$). We select C_1, L_1 as essential elements, and C_2, L_2 as excess elements. Accordingly, in the resistive circuit of Fig.6, the essential capacitor C_1 is replaced an ideal voltage source u_{C_1}, the essential inductor L_1 with an ideal current source i_{L_1}, whereas the excess capacitor C_2 is replaced with an ideal current source i_{C_2}, and the excess inductor L_2 with an ideal voltage source u_{L_2}.

A superposition approach is used to compute successively the transfer functions in equations (16) and (17) for the resistive multiport in Fig.6. One after the other, each source in the circuit is kept active alone, in sequence, while all the other sources are passivated (short circuit for ideal voltage sources, and open circuit for ideal current sources). To passivate an ideal voltage source, its *emf* must be made zero, *i.e.*, the source must be replaced with a short circuit (s.c.). Similarly, an ideal current source is passivated by making zero its current, *i.e.*, by replacing it with an open circuit (o.c.). The method is illustrated for two instances: In Fig. 7 only the source giving the voltage u_{C_1} is maintained active, whereas in Fig. 8 only the source giving the current i_{L_1} is active. In each case, the other sources are passivated. The outputs I_{C_1}, U_{L_1}, U_{C_2} and I_{L_2}, are calculated, which gives the corresponding transfer functions by a simple (symbolic) division.

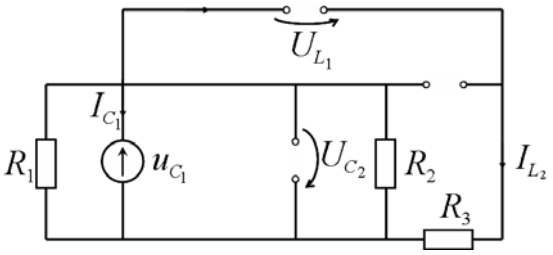


Fig. 7. Example of computing the resistive multiport transfer functions for the source u_{C_1}, the other sources being passivated.

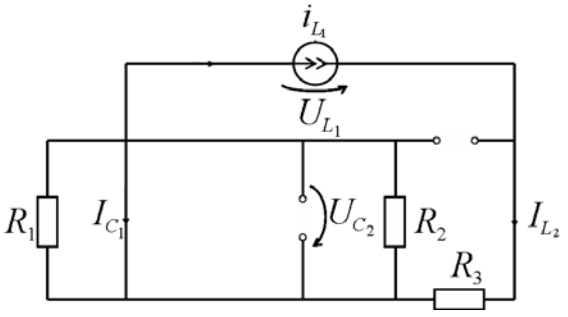


Fig. 8. Example of computing the resistive multiport transfer functions for the source i_{L_1}, the other sources being passivated.

The equations (16) and (17) which give the dependent variables at the terminals of the resistive multiport in Fig. 6 result:

$$\begin{bmatrix} I_{C_1} \\ U_{L1} \end{bmatrix} = \begin{bmatrix} -\dfrac{R_1+R_2}{R_1R_2} & -1 \\ 1 & -R_3 \end{bmatrix} \begin{bmatrix} u_{C_1} \\ i_{L_1} \end{bmatrix} + \begin{bmatrix} -\dfrac{1}{R_2} & -1 \\ 0 & -R_3 \end{bmatrix} \cdot \begin{bmatrix} e \\ j \end{bmatrix} + \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} i_{C_2} \\ u_{L_2} \end{bmatrix} \tag{29}$$

$$\begin{bmatrix} U_{C2} \\ I_{L2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{C_1} \\ i_{L_1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e \\ j \end{bmatrix}. \tag{30}$$

The parameters of the essential and excess circuit elements are, respectively:

$$[P]=\begin{bmatrix} C_1 & 0 \\ 0 & L_1 \end{bmatrix} \text{ and } [P']=\begin{bmatrix} C_2 & 0 \\ 0 & L_2 \end{bmatrix}. \tag{31}$$

The matrix $[\Pi]$ given by equation (25) is a diagonal matrix for this circuit:

$$[\Pi]=[P]-[c][P'][d]=\begin{bmatrix} C_1 & 0 \\ 0 & L_1 \end{bmatrix}-\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} C_2 & 0 \\ 0 & L_2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}=\begin{bmatrix} C_1+C_2 & 0 \\ 0 & L_1+L_2 \end{bmatrix} \tag{32}$$

so that its inverse results simply:

$$[\Pi]^{-1}=\begin{bmatrix} \dfrac{1}{C_1+C_2} & 0 \\ 0 & \dfrac{1}{L_1+L_2} \end{bmatrix} \tag{33}$$

and the matrices in the state variable equation are:

$$[A]=[\Pi]^{-1}[a]=\begin{bmatrix} -\dfrac{R_1+R_2}{R_1R_2(C_1+C_2)} & -\dfrac{1}{C_1+C_2} \\ \dfrac{1}{L_1+L_2} & -\dfrac{R_3}{L_1+L_2} \end{bmatrix},$$

$$[B]=[\Pi]^{-1}[[b] \quad [c][P'][f]]=\begin{bmatrix} -\dfrac{1}{R_2(C_1+C_2)} & -\dfrac{1}{C_1+C_2} & -\dfrac{C_2}{C_1+C_2} & 0 \\ 0 & -\dfrac{R_3}{L_1+L_2} & 0 & -\dfrac{L_2}{L_1+L_2} \end{bmatrix}, \tag{34}$$

$$U(t)]=\begin{bmatrix} e(t) & j(t) & \dfrac{\mathrm{d}e(t)}{\mathrm{d}t} & \dfrac{\mathrm{d}j(t)}{\mathrm{d}t} \end{bmatrix}$$

Notice that the vector of excitations $U(t)]$ contains not only the parameters of the voltage and current sources $(e(t)$ and $j(t))$, but also their time derivatives $(\dfrac{\mathrm{d}e(t)}{\mathrm{d}t}$ and $\dfrac{\mathrm{d}j(t)}{\mathrm{d}t})$. This is a typical property of the circuits with excess elements and can determine discontinuous time variation of (idealized) capacitor voltages or inductor currents. Such variables are continuous in time for circuits without excess elements.

5. CONCLUSIONS

An efficient algorithm to systematically establish the state variable equations for circuits with or without excess elements is given. The resistive part, which can

contain any type of controlled sources, is separated from the capacitor, inductor and source (excitation) components. As shown in the paper, the state variable equations of the circuit are found by computing the transfer functions of the purely resistive multiport.

State variable equations are a powerful tool for studying switching in circuits with inconsistent initial conditions [4], [5], [6], [7].

This lecture is based on a revised form of the paper submitted by the authors at the previous edition of the Sozopol Summer School, Bulgaria, 2009 [8].

REFERENCES

[1] Paul M. DeRusso, Rob J. Roy, Charles M. Close, Alan A. Desrochers, "State Variables for Engineers", 2nd Edition, Wiley Knowledge for Generations, 1998.

[2] M. Preda, P. D. Cristea, F. Spinei *et al*., "Circuit Analysis – State Variables" (in Romanian), University "Politehnica" Publishing House, Bucharest, 1973.

[3] M. Preda, P. D. Cristea et al, "Fundamentals of Electrotechnics" (in Romanian), Didactic and Pedagogic Publishing House, Bucharest, vol. 2, 920 pages, 1980

[4] J. Vlach, J.M. Wojciechowski and A. Opal, "Analysis on nonlinear networks with inconsistent initial conditions", IEEE Transactions on Circuits and Systems, vol. 42, no. 4, April 1995, pp. 195-200.

[5] D. Biolek, "Initial conditions of linear switched networks", ECCTD'95 European Conference on Circuit Theory Design, pp. 115-118, 1995, I.T.Ü.

[6] P. D. Cristea and Rodica Tuduce, "State variable analysis of linear switched networks with excess elements and inconsistent initial conditions", Rev. Roum. Sci. Techn. - Électrotechn. et Énerg., vol. 42, no.2, Bucharest, pp.209-218, 1997.

[7] P. D. Cristea and Rodica Tuduce, "Switching Processes in Idealized Networks with excess Elements and Inconsistent Initial Conditions", 5th Electronic Devices and Systems International Conference, Brno, Czech Republic, June 11-12, 1998, http://www.dsp.pub.ro/articles/switchingproces/brno98.htm

[8] P. D. Cristea and Rodica Tuduce, "Systematic Formulation of State Variable Equations for Circuits with Excess Elements", Advanced Aspects of Theoretical Electrical Engineering, September 20– 23, Sozopol, Bulgaria, pp.13, 2009.

AN APPROACH TO THE STATE-SPACE SYNTHESIS OF HIGH-ORDER LOG-DOMAIN FILTERS

Georgi A. Nenov

Higher School of Transport "T.Kableshkov", Sofia,
1574 Sofia, 158 G.Milev Str., Bulgaria
phone: +35929709232; +359896617007
e-mail: gnenov1@gmail.com

Abstract: *The topic of log-domain filter synthesis is an interesting and promising area of the global network theory and practice. During the last two decades it marks an extensive development and as a result of the research activity a variety of new methods arose. Almost of them are based on the use of bipolar or MOS transistors as translinear elements and lead to current processing structures. A significant part of the synthesis methods represent mainly attempts to simulate state-space equations of low-sensitivity passive filters. In the paper presented here one compares the well known nodal equation with the modified state-space equation of a suitable passive prototype in order to include exponential relationships of the bipolar transistors. In the realized structures one implements some constant coefficients in the equation set by current sources. The experimental results confirm the theory relationships.*

Keywords: Network Synthesis, Translinear Circuits, Log-Domain Filters.

1. INTRODUCTION

Since the beginning of twenty's century the filter theory and practice are the main motive force of the circuit theory evolution. Their path is marked by an introduction of new concepts and new circuit elements arising in consequence of the technology progress and of the obtained research theoretical results. After the intensive investigations in the field of the passive and active RC-filters the tendency to avoid not only inductors but resistors too brought to the theory of the operational transconductance amplifier–capacitor (OTA-C), switched-capacitor (SC), switched-current (SI) and during the last two decades of the translinear (TL) filters. The typical representatives of the TL-filters are log-domain filters. The main advantage of this class of frequency selective devices consists in the possibility to realize them on the base of the well elaborate integral technology without use of resistors and in the considerably reduced power consumption. The last peculiarity is a result of some exponential characteristics of the bipolar or MOS-transistors used in the translinear filter circuits.

The paper presented is an attempt to realize high-order log-domain filters on the base of a suitable description of passive prototypes by a set of their state-space equations. This approach is not new and it was widely used first at all in active RC-filter synthesis. Later this method was transferred at applied to the synthesis of SC-, SI, OTA-C- and finally to translinear filters [1-3, 5-8]. A special interest are the procedures, described in [1, 2, 6, 7], where one transforms the set of the previously derived

space-state equations in such a way that the new set can be realize by using translinear tranconductors [1, 2] or by introducing of the suitable translinear loops [5 - 8]. However the examples given in these papers concern to first- or second-order filters only and it is no enough clear, especially in [6], how the approach proposed can be applied to higher-order log-domain filters. That is way below we use the combination of the methods in [1, 2, 4, 6] on purpose to improve and widen the possibility to design high-order log-domain filters.

2. NONLINEAR ELEMENT

Usually the basic nonlinear characteristic used in the synthesis of log-domain filters is the function between the collector current $i_C(t)$ and the base-emiter voltage $u_{BE}(t)$ of a bipolar (BJ) NPN- or PNP- transistor [3]:

$$i_C(t) = I_s e^{\frac{u_{BE}(t)}{U_T}}, \tag{1}$$

where I_s is the reverse saturation current and U_T is the thermal voltage ($I_s = 10^{-15} \div 10^{-12}$, A; $U_T \sim 26.10^{-3}$, V).

From the expression (1) one obtains

$$\ln i_C(t) = \ln I_s + \frac{u_{BE}(t)}{U_T}. \tag{2}$$

This characteristic allows to find the relationship between the transistor collector currents in a translinear loop according to the static translinear principle [3].

3. STATE-SPACE EQUATIONS

The common form of the state-space equations of a system (particularly of the passive filter prototype) is

$$\left.\begin{aligned} \frac{d\mathbf{x}(t)}{dt} &= \mathbf{A}.\mathbf{x}(t) + \mathbf{B}.\mathbf{f}(t); \\ \mathbf{y}(t) &= \mathbf{C}.\mathbf{x}(t) + \mathbf{D}.\mathbf{f}(t). \end{aligned}\right\}.$$

Here

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) & \dots & x_n(t) \end{bmatrix}_t$$

is the vector of the state variables. In the case of an electrical network it is formed exclusively from the inductor currents and the capacitor voltages,

$$\mathrm{f}(t) = \begin{bmatrix} f_1(t) & \dots & f_m(t) \end{bmatrix}_t,$$

is the vector of the external independent variables (currents of independent current sources and voltages of the independent voltage sources) and

$$\mathrm{y}(t) = \left\lfloor y_1(t) \quad \ldots \quad y_r(t) \right\rfloor_t \tag{6}$$

is the vector of quantities not belonging to the state variables. The elements of the matrices A of size ($n \times n$), B of size ($n \times m$), C of size ($r \times n$) and D of size ($r \times m$) depend on the circuit element parameters.

In order to transform properly further the equation (1) one assumes [1, 2, 6]

$$x_i(t) = I_s.e^{\frac{v_i(t)}{U_T}}; i = 1, 2, ..., n; \;\; f_k(t) = I_s.e^{\frac{w_k(t)}{U_T}}; k = 1, 2, ..., m, \tag{7}$$

where $v_i(t)$ and $w_k(t)$ are circuit node potentials. Then we have

$$\frac{dx_i(t)}{dt} = \frac{I_s}{U_T}.e^{\frac{v_i(t)}{U_T}}.\frac{dv_i(t)}{dt}; \quad . \tag{8}$$

Now the state-space equations (3) can be rewritten as

$$\left.\begin{aligned} &\mathrm{KE}(t).\frac{d}{dt}\mathrm{v}(t) = U_T \mathrm{A.e}_v(t) + U_T.\mathrm{KB.e}_w(t); \\ &\mathrm{y}(t) = I_s.\mathrm{C.e}_v(t) + I_s.\mathrm{D.e}_w(t) \end{aligned}\right\}, \tag{9}$$

for

$$\left.\begin{aligned} &\mathrm{e}_v(t) = \left[e^{\frac{v_1(t)}{U_T}} \quad \ldots \quad e^{\frac{v_n(t)}{U_T}} \right]_t; \;\; \mathrm{e}_w(t) = \left[e^{\frac{w_1(t)}{U_T}} \quad \ldots \quad e^{\frac{w_m(t)}{U_T}} \right] \\ &\mathrm{E}(t) = diag\left\{ e^{\frac{v_1(t)}{U_T}} \quad \ldots \quad e^{\frac{v_n(t)}{U_T}} \right\}; \\ &\mathrm{K} = diag\left\{ C_1 \quad \ldots \quad C_n \right\}; \;\; \mathrm{v}(t) = \left\lfloor v_1(t) \quad \ldots \quad v_n(t) \right\rfloor_t \end{aligned}\right\}, \tag{10}$$

where C_1,…, C_n are capacity values. After multiplying both sides of (10) by the matrix E^{-1} from the first equation of (9) it follows

$$.\mathbf{K}.\frac{d\mathbf{v}(t)}{dt} = U_T.\mathbf{KE}(t)^{-1}.\mathbf{A}.\mathbf{e}_v(t) + .U_T.\mathbf{KE}(t)^{-1}\mathbf{B}.\mathbf{e}_w(t). \tag{11}$$

Then the detailed form of the i^{th} equation in the set (11) is

$$C_i.\frac{dv_i(t)}{dt} = C_i.U_T.\left(\sum_{j=1}^{n} a_{ij}.e^{\frac{v_j(t)-v_i(t)}{U_T}} + \sum_{k=1}^{m} b_{ik}.e^{\frac{w_k(t)-v_i(t)}{U_T}}\right) \tag{12}$$

or

$$C_i.\frac{dv_i(t)}{dt} = \sum_{j=1}^{n} K_{a,ij}.e^{\frac{v_j(t)-v_i(t)}{U_T}} + .\sum_{s=1}^{m} K_{b,ik}.e^{\frac{w_k(t)-v_i(t)}{U_T}}; K_{a,ij} = C_i.U_T.a_{ij}; K_{b,ik} = C_i.U_T.b_{ik}$$

and it corresponds to the global current flowing into the capacitor C_i forced by the potentials $v_i(t)$, $v_j(t)$ and $w_k(t)$.

4. REALIZATION ASPECTS

The practical circuit implementation of the sums in (13) depends mainly on their signs. In [1, 2] it is demonstrated an use of positive and negative transconductors for that purpose, whereas the first-order filter examples given in [6] need only positive coefficients in these sums. The disadvantage of the transconductor subsircuits, given in [1, 2] is that the delivered current $i(t)$ depends on the voltage difference $(v_2(t) - v_1(t))/2U_T$ instead of the difference $(v_2(t) - v_1(t))/U_T$ – this peculiarity limits the application of the mentioned subcircuits together with other similar devices. The transconductor units proposed in [4] however do not possess this drawback. In Fig. 1 such a negative transconductor is shown where

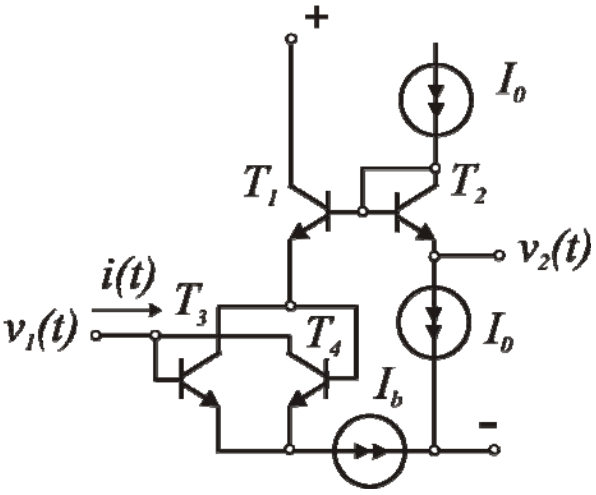


Fig. 1. Negative transconductor [4]

$$i(t) = I_0.e^{\frac{v_2(t)-v_1(t)}{U_T}}. \tag{14}$$

An other circuit solution in [6] uses translinear loops in order to obtain the potential differences in the exponent powers. However it seems the approach proposed in [6] is not directly applicable for the log-domain realization of higher-order filters. In what follows we try to combine the approaches, given in [4, 6] and to adapt them for the higher-order log-domain filter realization. Indeed we can obtain the positive coefficients in the modified state-space equation set by the simpler circuits described in [6] and the negative coefficients by using negative transconductors [4]. As a result an element number reduction in the circuit is possible.

5. EXAMPLE

Let's track out the synthesis of a third-order log-domain Chebyshev low-pass filter without transfer function zero and with following specifications [9]: cut-off frequency f_c= *150 kHz*; bandpass ripple ΔA = 0,177 dB; equal load resistances r = 10 Ω. In the case the passive prototype in Fig. 3 satisfies these conditions for r = 10 Ω; c_1 = 126,2, nF; l_2 = 122,4, μH; c_3 = 126,2, nF. The filter circuit state-space equations are

$$\frac{d}{dt}\begin{bmatrix} u_1(t) \\ i_2(t) \\ u_3(t) \end{bmatrix} = \left\{ \begin{bmatrix} -\frac{1}{rc_1} & -\frac{1}{c_1} & 0 \\ \frac{1}{l_2} & 0 & -\frac{1}{l_2} \\ 0 & \frac{1}{c_3} & -\frac{1}{c_3} \end{bmatrix} . \begin{bmatrix} u_1(t) \\ i_2(t) \\ u_3(t) \end{bmatrix} + \begin{bmatrix} \frac{1}{c_1} \\ 0 \\ 0 \end{bmatrix} . u_{in}(t); u_{out}(t) = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} . \begin{bmatrix} u_1(t) \\ i_2(t) \\ u_3(t) \end{bmatrix} \right\} \quad (15)$$

Then by the substitution

$$\left.\begin{aligned} u_{in}(t) = I_s e^{\frac{w_1(t)}{U_T}}; u_1(t) = I_s e^{\frac{v_1(t)}{U_T}}; i_2(t) = I_s e^{\frac{v_2(t)}{U_T}}; \\ u_3(t) = u_{out}(t) = I_s e^{\frac{v_3(t)}{U_T}} \end{aligned}\right\} \quad (16)$$

and having in mind (13) from (10) and (15) one obtains

$$\left.\begin{aligned} \mathrm{e}_v(t) = \begin{bmatrix} e^{\frac{v_1(t)}{U_T}} & e^{\frac{v_2(t)}{U_T}} & e^{\frac{v_3(t)}{U_T}} \end{bmatrix}_t ; \ \mathrm{E}(t) = diag\left\{ e^{\frac{v_1(t)}{U_T}} \quad e^{\frac{v_2(t)}{U_T}} \quad e^{\frac{v_3(t)}{U_T}} \right\}; \\ \mathrm{e}_w(t) = e^{\frac{w_1(t)}{U_T}} ; \ \mathrm{K} = diag\{C_1 \quad C_2 \quad C_3\}; \ \mathrm{v}(t) = \begin{bmatrix} v_1(t) & v_2(t) & v_3(t) \end{bmatrix}_t \end{aligned}\right\}. \quad (17)$$

Consequently the modified set of state-space equations is

$$\left.\begin{aligned} C_1 . \frac{dv_1(t)}{dt} = K_{a,11} + K_{a,12} . e^{\frac{v_2(t)-v_1(t)}{U_T}} + K_{b,1i} . e^{\frac{w_1(t)-v_1(t)}{U_T}} \\ C_2 \frac{dv_2(t)}{dt} = K_{a,21} . e^{\frac{v_1(t)-v_2(t)}{U_T}} + K_{a,23} . e^{\frac{v_3(t)-v_2(t)}{U_T}} \\ C_3 \frac{dv_3(t)}{dt} = K_{a,32} . e^{\frac{v_2(t)-v_3(t)}{U_T}} + K_{a,33} \end{aligned}\right| \quad (18)$$

for

$$\left.\begin{aligned} K_{a,11} &= -\frac{U_T.C_1}{r.c_1}; K_{a,12} = -\frac{U_T.C_1}{c_1}; K_{b,1i} = \frac{U_T.C_1}{r.c_1}; \\ K_{a,21} &= \frac{U_T.C_2}{l_2}; K_{a,23} = -\frac{U_T.C_2}{l_2}; \\ K_{a,32} &= \frac{U_T.C_3}{c_3}; K_{a,33} = -\frac{U_T.C_3}{r.c_3} \end{aligned}\right\}. \quad (19)$$

According to [6] we express the coefficients $K_{b,1i}$, $K_{a,21}$ and $K_{a,32}$ as follows

$$K_{b,1i} = I_s.e^{\frac{V_{b,1i}}{U_T}}; K_{a,21} = I_s.e^{\frac{V_{a,21}}{U_T}}; K_{a,32} = I_s.e^{\frac{V_{a,32}}{U_T}} \quad (20)$$

and by replacing them in (18) the equation set yields:

$$\left|\begin{aligned} C_1.\frac{dv_1(t)}{dt} &= K_{a,11} + K_{a,12}.e^{\frac{v_2(t)-v_1(t)}{U_T}} + I_s.e^{\frac{V_{b,1i}+w_1(t)-v_1(t)}{U_T}} \\ C_2\frac{dv_2(t)}{dt} &= I_s.e^{\frac{V_{a,21}+v_1(t)-v_2(t)}{U_T}} + K_{a,23}.e^{\frac{v_3(t)-v_2(t)}{U_T}} \\ C_3\frac{dv_3(t)}{dt} &= I_s.e^{\frac{V_{a,32}+v_2(t)-v_3(t)}{U_T}} + K_{a,33} \end{aligned}\right. . \quad (21)$$

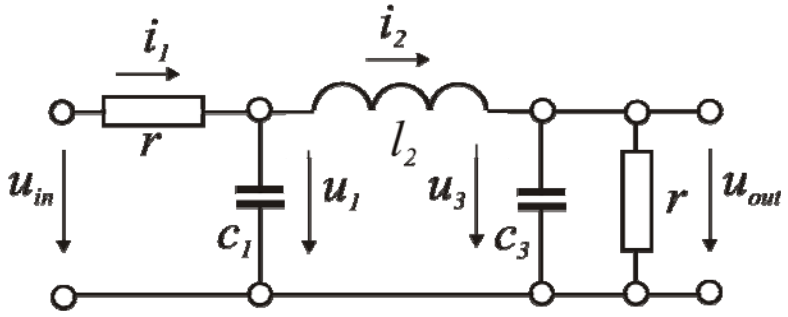


Fig. 2. Prototype third-order low-pas Chebyshev filter

For the calculation of the coefficients (19) we choose $I_s = 10^{-14}$, A; $U_T = 23.10^{-3}$, V; $c_1 = c_2 = c_3 = 100$ nF and then

$$\left.\begin{array}{l} K_{a,11} = -2{,}06.10^{-3}, A; K_{a,12} = -2{,}06.10^{-2}\ , V; K_{b,1i} = 2{,}06.10^{-3}\ , A; \\ K_{a,21} = 2{,}124.10^{-4}, Q/H; K_{a,23} = -2{,}124.10^{-4}, Q/H; \\ K_{a,32} = 2{,}06.10^{-2}, V; K_{a,33} = -2{,}06.10^{-3}\ , A \end{array}\right\}. \qquad (22)$$

As it can be seen from (21) and (22) the coefficients $K_{b,1i}$, $K_{a,21}$ and $K_{a,32}$ have positive values, whereas the coefficient values of the other terms are negative. This means that we are able to apply for the positive coefficients the approach based on a creating of suitable translinear loops. The realization of the negative coefficients $K_{a,12}$ and $K_{a,23}$ will be implemented by using the negative transconductor unit in Fig. 1. The constant coefficients $K_{a,11}$ and $K_{a,\,33}$ will be realized by independent current sources.

The circuit of the synthesized according to (21) log-domain filter is shown in Fig.3. The implementation of the positive coefficients $K_{b,1i}$, $K_{a,21}$ and $K_{a,32}$ follows from the translinear loop equations [3]:

$$\left.\begin{array}{l} u_{BE,1}(t) + u_{BE,2}(t) - v_1(t) = u_{BE,3}(t); \\ v_1(t) + u_{BE,4}(t) - v_2(t) = u_{BE.5}(t); \\ v_2(t) + u_{BE,6}(t) - v_3(t) = u_{BE,7}(t) \end{array}\right\}, \qquad (23)$$

respectively, where $u_{BA,k}(t)$; k = 1, 2, …, 7 are the base-emiter voltages of the corresponding transistors. On the other hand in connection with (21) we have

$$\left.\begin{array}{l} u_{BE,1}(t) = w_1(t); u_{BE,2}(t) = v_{b,1i}(t); u_{BE,3}(t) = U_T.[\ln i_{C,3}(t) - \ln I_s]; \\ u_{BE,4}(t) = v_{a,21}(t); u_{BE,5}(t) = U_T.[\ln i_{C,5}(t) - \ln I_s]; \\ u_{BE,6}(t) = v_{a,32}(t); u_{BE,7}(t) = U_T.[\ln i_{C,7}(t) - \ln I_s] \end{array}\right\}. \qquad (24)$$

The replacing of (24) in (23) yields:

$$\left.\begin{array}{l} i_{C,3}(t) = I_s.e^{\frac{v_{b,1i}(t)+w_1(t)-v_1(t)}{U_T}}; \\ i_{C,5}(t) = I_s.e^{\frac{v_{a,21}(t)+v_1(t)-v_2(t)}{U_T}}; \\ i_{C,7}(t) = I_s.e^{\frac{v_{a,32}(t)+v_2(t)-v_3(t)}{U_T}} \end{array}\right\} \qquad (25)$$

and this result coincides with (21).

In order to simplify the circuit in Fig. 3, it is possible also to unite the parallel connected current sources $K_{a,11}$ and $K_{a,21}$.

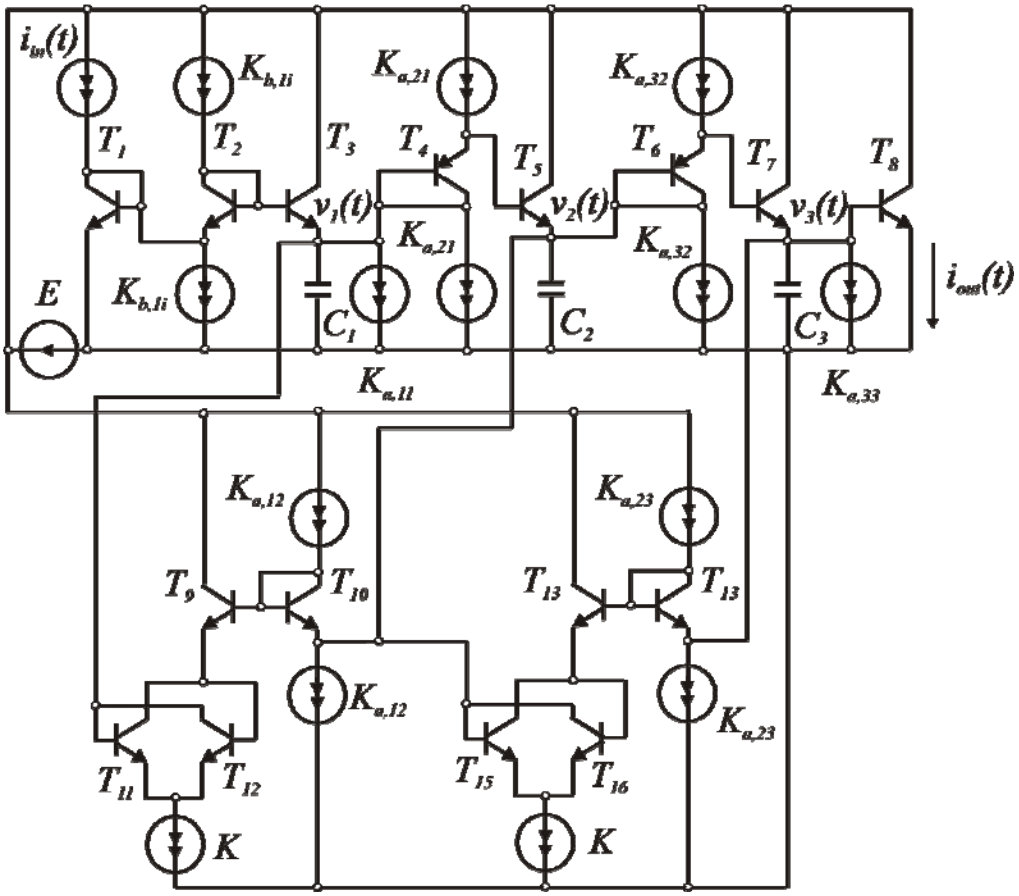


Fig. 3. Synthesized log-domain third-order low-pass Chebyshev filter

The simulation of the synthesized log-domain filter was performed on the base of the program *Multisim 2001* for the power voltage E=3 V and virtual NPN- and PNP-transistors. In order to ensure the positive values of variables a 2 mA offset input current was added. The value of the biasing current sources of negative transconductors is K = 8 mA. Fig. 4a represents the attenuation characteristic of the prototype filter, whereas Fig. 4b and Fig. 4c represent the attenuation and phase characteristics of the simulated filter versus frequency, respectively.

6. CONCLUSIONS

A systematic synthesis procedure of high-order log-domain filters on the base of the modified state-space equations of the passive prototype is proposed. The circuit implementation of the equation coefficients uses an introduction of suitable translinear loops and negative tranconductor units. The theoretical results are applied to the synthesis of a third-order log-domain Chebyshev low-pass filter and the corresponding computer simulation is carried out.

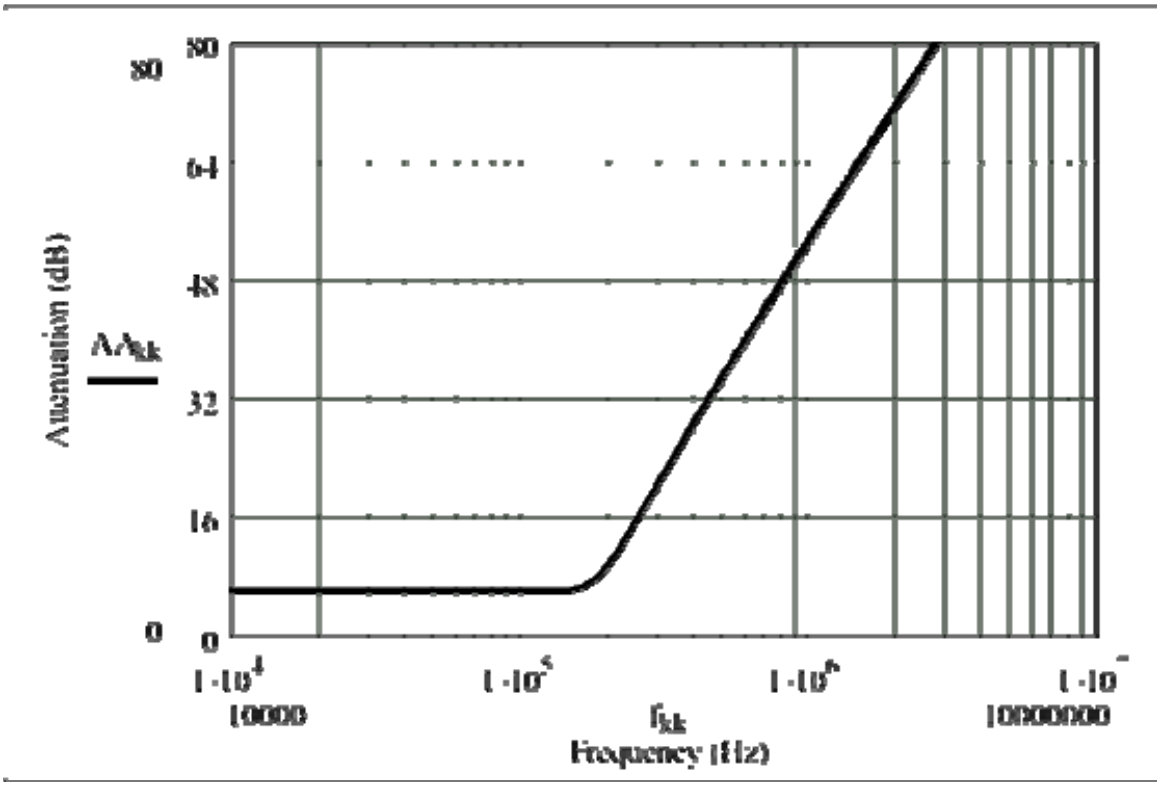


Fig. 4a. Attenuation vs frequency of the prototype filter

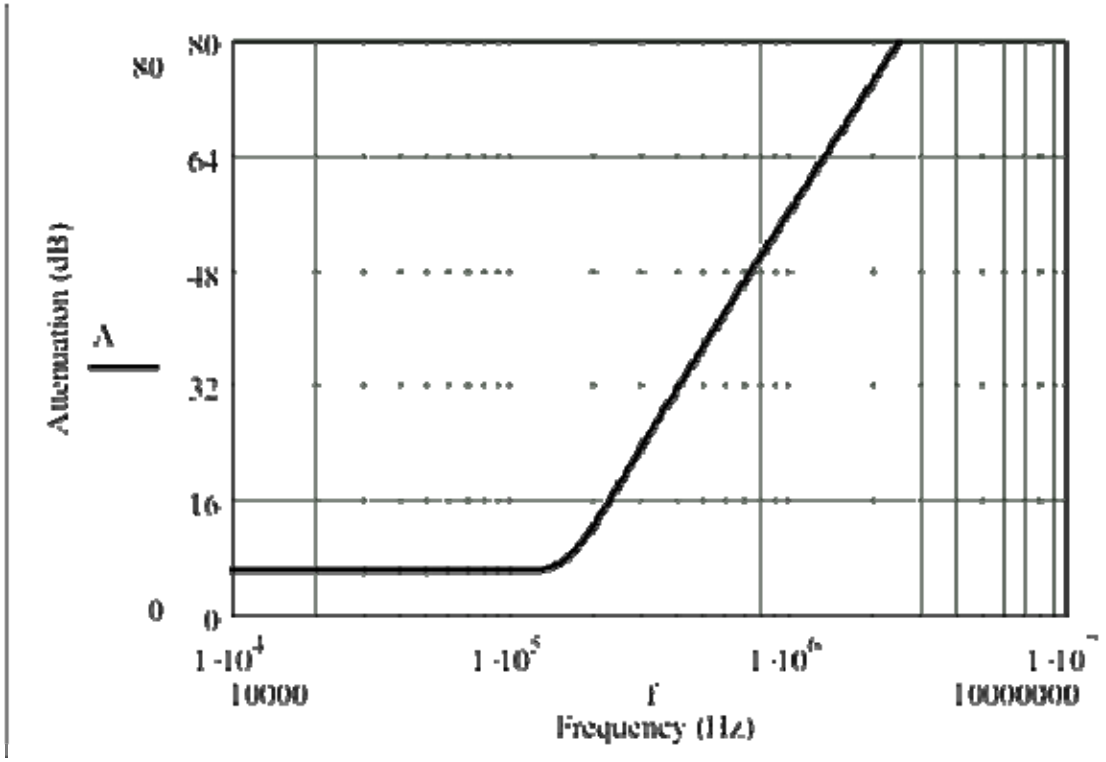


Fig. 4b. Attenuation vs frequency of the simulated filter

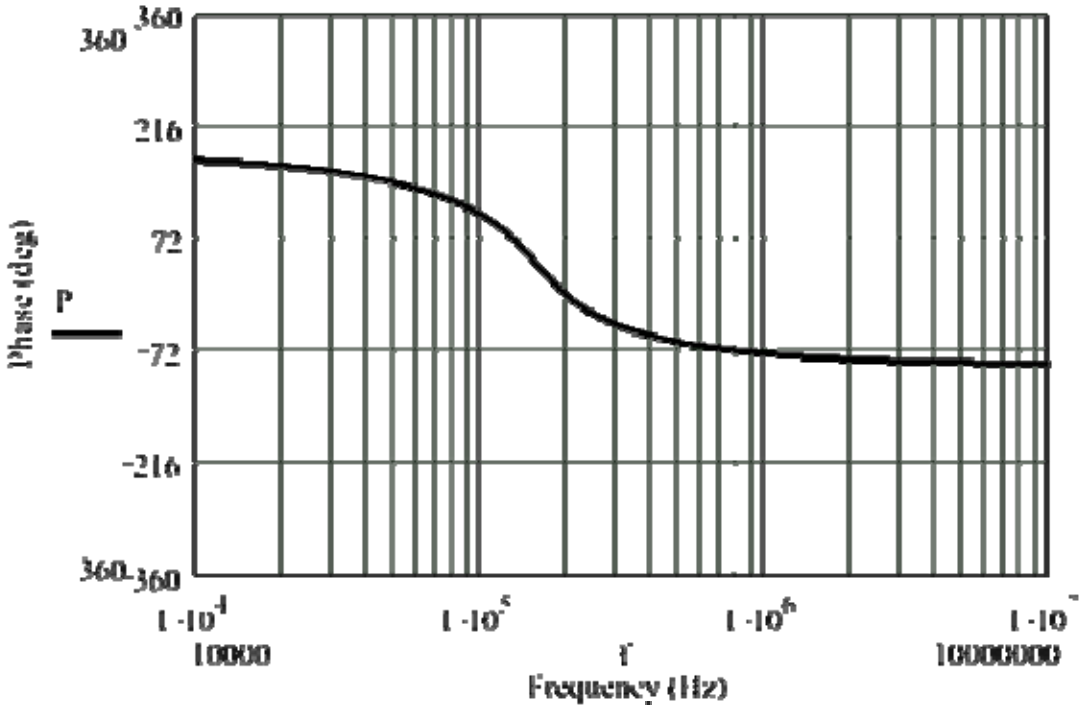


Fig. 4c. Phase vs frequency of the simulated filter

REFERENCES

[1] D. R. Frey, "Exponential State Space Filters: A Generic Current Mode Design Strategy", IEEE Trans. on Circuits and Systems, Part I, vol. 43, 1, 1996, pp.34-42

[2] D. R. Frey, "State-Space Synthesis and Analysis of Log-Domain Filters", IEEE Trans. on Circuits and Systems, Part II, vol. 45, 9, 1998, pp. 1205-1211

[3] C. Toumazou, F. J. Lidgey, D. G. Haigh – Editors, "Analogue IC Design: the Current-Mode Approach", Peter Pelegrinus Ltd, 1998

[4] E. M. Drakakis, A. J. Pajne, C. Toumazou, "Log-Domain Staye-Space": A\Systemati Transistor- Level Approach for Log-Domain Filtering", IEEE Tr. on Circuits and Systems, Part II, vol.46, 3 1999, pp. 290-305

[5] S. Subramanian, D.V.Anderson, P. Hasler, B. A. Minch, "Synthesis of MITE Log-Domain Filters with Unique Operating Points", IEEE Int. Symposium on Circuits and Systems, 2005, Kobe, Japan, pp. 996-999

[6] A. Kircay, U. Cam, "State-Space Synthesis of Current-Mode First-Order Log-Domain Filters", Turk. J. Elec. Engin., vol. 14, 3, 2006, pp. 399-416

[7] A. Kircay, U. Cam, "A Novel Log-Domain First-Order Multifunction Filter, ETRI Journal, vol. 28, 2, 2006, pp. 401-404

[8] R. Arslanalp, A.T.Tola, "State Space Representation for Log-Domain Filtering Synthesis", Indian Journal of Pure & Applied Physics, vol. 47, October, 2009, pp. 745-752

[9] Г. Ханзел, "Справочник по расчету фильтров", Сов. Радио, Москва, 1974.

STATISTICAL APPROACH TO FAULT OBSERVABILITY INVESTIGATION IN ANALOG CIRCUITS

Elissaveta Dimitrova Gadjeva[*], Dimitar Yordanov Shikalanov[**], Anton Georgiev Atanasov[***]

[*]Department of Electronics and Electronic Technologies, Technical University of Sofia, 8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria, e-mail: egadjeva@tu-sofia.bg

[**]Department of Informatics, New Bulgarian University, Montevideo 21 Str., 1635 Sofia, Bulgaria, e-mail: dys@nbu.bg

[***]XPEQT Ltd., 2, Samokovsko Shosse, 1138 Sofia, Bulgaria, e-mail:antonatan@abv.bg

Abstract: *An approach is developed in the present paper to automated fault observability investigation in analog circuits based on statistical analysis. Diagnosis macromodels are developed for n-fold fault generation. The proposed approach allows the fault coverage and fault observability investigation in presence of design and process tolerances. The computer realization of the diagnosis approach is performed using the Cadence PSpice circuit simulator and Cadence Probe graphical analyzer. As a result, the optimal test node set and test quantities, maximizing the fault observability, can be selected. Macrodefinitions in Probe for the functional and I_{dd} testing are described. An example is given to illustrate the possibilities of the proposed approach.*

Keywords: Fault observability, Monte Carlo simulation, Analog circuits, Cadence PSpice simulator

1. INTRODUCTION

The increasing complexity of the analog and mixed-signal VLSI circuits and the limited number of accessible nodes complicate significantly the fault diagnosis. Fast and efficient fault diagnosis approaches are of significant importance to reduce the cost of testing analog circuits. A number of approaches are proposed to increase the fault observability and the fault coverage of the circuit under test (CUT) based on sensitivity analysis [1], statistical simulation [2], oscillation based testing [3,4], etc. In [5] the fault observability investigation is based on the output voltage and supply current monitoring in the presence of catastrophic and parametric faults. The fault coverage using functional and supply current (I_{dd}) testing are compared. The oscillation based testing (OBT) strategy is often used to achieve satisfactory fault coverage. The principle of OBT consists in reconfiguring the CUT in test mode to an oscillator [3,4] (Fig. 1a). The existence of catastrophic and parametric faults in the CUT results in deviation of the oscillation frequency from its normal value or in loss of oscillation. The fault coverage can be increased by monitoring the oscillation frequency, the amplitude of the oscillation voltage signal, as well as the supply current.

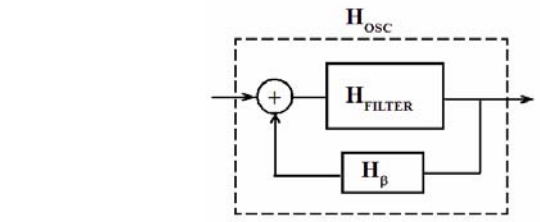
An approach is developed in the present paper to automated fault observability investigation in analog circuits based on statistical analysis. Diagnosis macromodels are developed for *n*-fold fault generation. The fault coverage is obtained using *Monte*

Carlo simulation of the diagnosis model taking into account the design tolerances. The proposed approach allows the fault coverage and fault observability investigation in presence of design and process tolerances. As a result, the optimal test node set and test quantities, maximizing the fault observability, can be selected.

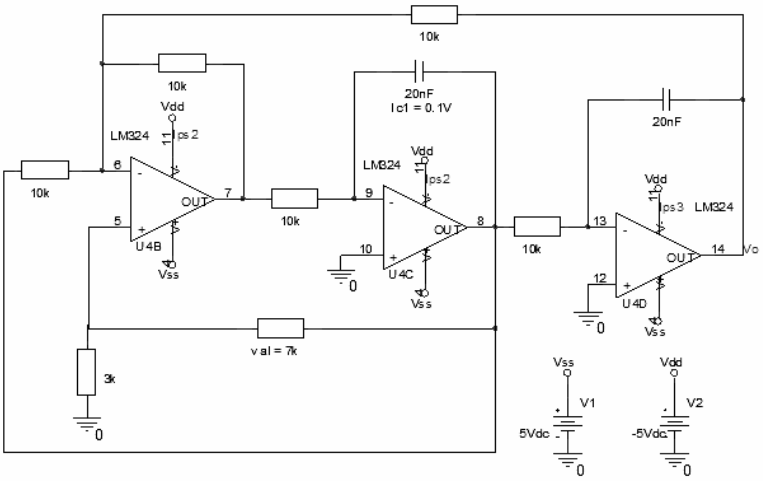
2. STATISTICAL OBT APPROACH TO FAULT COVERAGE INVESTIGATION

Oscillation based testing of the CUT is performed combining functional and I_{dd} testing. The oscillation frequency and the amplitude of the output voltage is monitored for the functional testing. The average and the peak-to-peak value of the current through the voltage supply I_{dd} are monitored for the I_{dd} testing. The modeled parametric faults of -50% and +100% and design tolerances of 5% are defined for the example benchmark circuit shown in Fig. 1b [3].

The signals N_{Faults} and *TEST* are defined in the diagnosis model (Fig. 2 and Fig. 3). The signal $N_{Faults} = 0$ if the circuit is correct (all element values are in the design tolerance limits $\pm\varepsilon\%$). $N_{Faults} \neq 0$ if the circuit is faulty. In this case N_{Faults} is equal to the number of modeled faults for the current run (N_{Faults} = 1 for single fault, N_{Faults} = 2 for double fault, N_{Faults} = *n* for *n*-fold fault). The signal *TEST* is defined in the following way: *TEST* = 0 for the correct circuit and *TEST* = 1 for the faulty circuit. *TEST* is used to evaluate the fault coverage using the selected test quantity.



a)



b)

Fig. 1. The benchmark circuit in test mode

3. COMPUTER REALIZATION OF THE APPROACH

The models of the faulty circuit elements are created. The computer model of the faulty resistor is presented in Fig. 4. R_1 models the resistor element of nominal value {@val}. The tolerance deviation is defined by **Distribution** statement as shown in Fig. 5. It is described in the form:

```
.DISTRIBUTION DIST (-1,0)(-0.5,0)(-0.5,1)(-0.45,1)(-0.45,0)(-0.05,0)
(-0.05,1)(0.05,1)(0.05,0)(0.95,0)(0.95,1)(1,1)
```

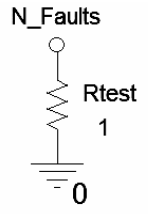


Fig. 2. Creating the signal N_{faults}

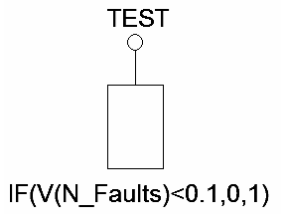


Fig. 3. Creating the signal TEST

$R_{1a} = 1\Omega$ is defined to test if the statistically generated value R_1 is faulty or non-faulty. R_1 and R_{1a} are defined with simultaneous (corellated) variation using the mode **LOT**. If $|R_{1a} - R_1| \le \varepsilon\%$ the resistor is nonfaulty, otherwise the resistor is faulty. The voltage controlled current source (VCCS) G3 of IVALUE type is used to generate the test current I_1 .

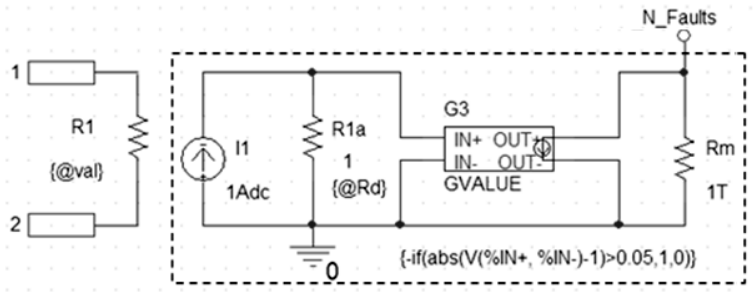


Fig. 4. Computer model of the faulty resistor

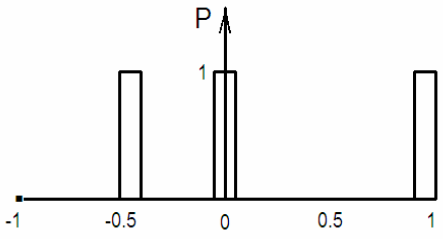


Fig. 5. Statistical distribution for defining the parameter value

The current $I_1 = 0$ for the nonfaulty resistor and $I_1 = 1$A for the faulty resistor. I_1 is defined using the IF-THEN-ELSE statement in the form:

```
{-IF(ABS(V(%IN+,%IN-)-1)>0.05,1,0)}
```

Similarly, the computer model of the faulty capacitor shown in Fig. 6 can be used for diagnosis in the frequency domain.

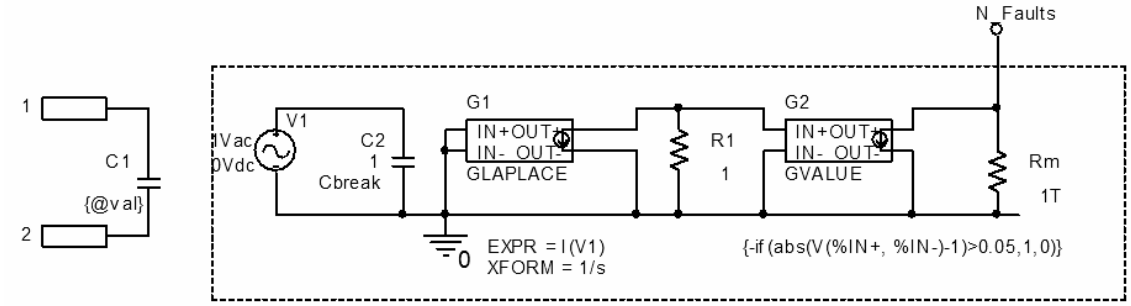


Fig. 6. Computer model of the faulty capacitor in the frequency domain

In the general case, the computer diagnosis model of the faulty capacitor shown in Fig. 7 is constructed. It is obtained from the diagnosis capacitor model proposed in [6] where the block shown in dashed line is added to create the signal N_{faults}.

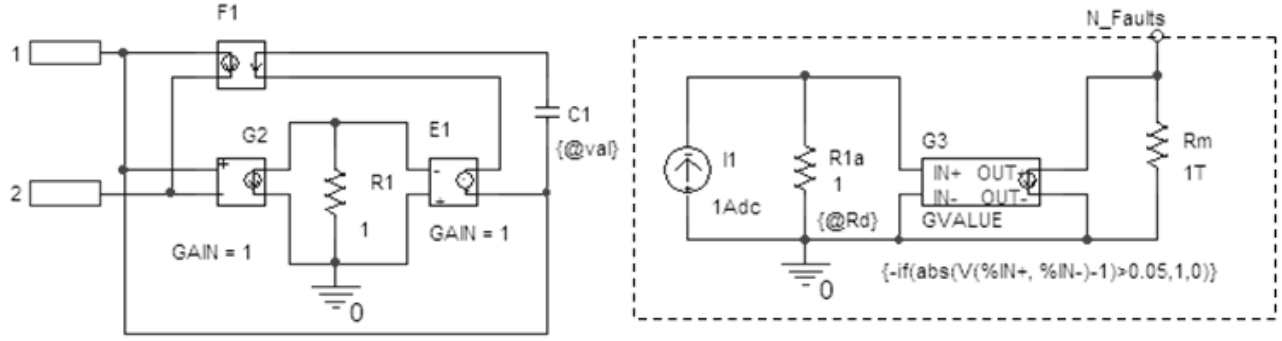


Fig. 7. Generalized computer model of the faulty capacitor

The signal N_{faults} is created in the form of node voltage of the node *N_Faults* defined as a global node using a global port (Fig. 2). In this way all test currents generated for each of the circuit elements, are connected in parallel with the element $R_{test} = 1\Omega$. As a result, the signal *V(N_faults)* is equal to the number of simultaneously generated faults:

$$V_{N_{faults}} = \sum_i R_{test} I_{test,i} \qquad (1)$$

The signal *TEST* is created using the *ABM* element from the library *ABM.lib* (Fig. 3). It is defined using the *IF-THEN-ELSE* statement:

```
IF(V(N_FAULTS)<0.1,0,1)
```

The faulty resistor and capacitor models are realized in *PSpice* using block and using subcircuit definition.

The computer realization of the model of faulty resistor using block **RES_d** is shown in Fig. 4. The block name and the attribute Rd = **@Reference** are transferred to the block in order to define the model name *Rd* for statistical distribution. The *PSpice* template is defined in the form:

```
R^@REFDES %1 %2 ?TOLERANCE|@RD| @VALUE ?TOLERANCE|\N.MODEL @RD RES
R=1 LOT/DIST2=@TOLERANCE%|
```

The model statement

```
.MODEL @RD RES R=1 LOT/DIST2=@TOLERANCE%|
```

is introduced in the *PSpice* template of the resistor elements R_1 and R_{1a}. The element R_{1a} is of **Rbreak** type with the same tolerance model name which coincides with the block name. Its *PspiceTemplate* is in the form:

```
R^@REFDES %1 %2 @MODEL @VALUE
```

The diagnosis models of the faulty elements can be also defined using subcircuit definitions. In this case, the *PSpiceTemplate* for the resistor model is in the form:

```
X^@REFDES %1 %2 @MODEL PARAMS: VAL={@VAL}
```

The *PSpiceTemplate* for the capacitor model is in the form:

```
X^@REFDES %1 %2 @MODEL PARAMS: VAL={@VAL} IC1={@IC1}
```

whete IC1 defines the initial condition.

The subcircuit descriptons for the resistor and capacitor elements are in the form:

```
*RESISTOR MODEL
.SUBCKT RES_DIAG 1 2 OPTIONAL: N_FAULTS=$G_N_FAULTS PARAMS: VAL = 1
R_R1 1 2 RGR1 {VAL}
I_I1 0 5 DC 1ADC AC 1AAC
R_R2 0 N_FAULTS 1T
G_G3 N_FAULTS 0 VALUE {{-IF(ABS(V(5, 0)-1)>0.05,1,0)}}
R_R1A 0 5 RGR1 1
.MODEL RGR1 RES R=1 LOT/DIST 100%
.ENDS
*CAPACITOR MODEL
.SUBCKT CAP_DIAG 1 2 OPTIONAL:N_FAULTS=$G_N_FAULTS PARAMS:VAL=1 IC1=1
G_G2 3 0 1 2 1
E_E1 1 4 0 3 1
C_C1 1 6 {VAL} IC={IC1}
X_F1 6 4 1 2 CAP_D_F1
R_R1 0 3 RGROUP 1
I_I1 0 5 DC 1ADC AC 1AAC
R_R2 0 N_FAULTS 1T
G_G3 N_FAULTS 0 VALUE {{-IF(ABS(V(5, 0)-1)>0.05,1,0)}}
R_R1A 0 5 RGROUP 1
.MODEL RGROUP RES R=1 LOT/DIST=100%
.ENDS
```

The monitored test quantities for the functional test are the oscillation frequency and the amplitude of the output voltage. The test quantities for the I_{dd} test are the average value of I_{dd} and the peak-to-peak value of Idd. They are calculated using the following macrodefinitions in *Probe*:

```
FO = 1/PERIOD(V(VO))
AMPL = (MAX(V(VO))-MIN(V(VO)))/2
IDD_AVG = -MAX(-AVG(I(V2))*(TIME)/TMAX)
IDDPP = MAX(I(V2)*TIME/TMAX)-MIN(I(V2)*TMAX/TIME)
```

The parameters *TEST_Fo*, *Test_ampl*, *TEST_Idd_avg* and *TEST_Iddpp* are obtained in *Probe*. The parameter *TEST_Fo* = 0 if the generated variant is non-faulty, *TEST_Fo* = 1 for the nonrecognized faulty variant and *TEST_Fo* = 2 for the recognized faulty variant. The parameters *Test_ampl*, *TEST_Iddavg* and *TEST_Iddpp* are defined similarly. They are calculated using the following macros in *Probe*:

```
COR_AMPL = 0.5*(SGN(AMPL_MIN-AMPL)+1)+0.5*(SGN(AMPL-AMPL_MAX)+1)
TEST_AMPL = COR_AMPL+MAX(V(TEST))
COR_FO = 0.5*(SGN(FOMIN-FO)+1)+0.5*(SGN(FO-FOMAX)+1)
TEST_FO = COR_FO+MAX(V(TEST))
```

```
COR_V = 0.5*(SGN(COR_FO+COR_AMPL-0.1)+1)
TEST_V = COR_V+MAX(V(TEST))
COR_IDDAVG = 0.5*(SGN(IDD_AVGMIN-IDD_AVG)+1)+0.5*(SGN(IDD_AVG-
IDD_AVGMAX)+1)
TEST_IDDAVG = COR_IDDAVG+MAX(V(TEST))
COR_IDDPP = 0.5*(SGN(IDDPPMIN-IPSPP)+1)+0.5*(SGN(IDDPP-IDDPPMAX)+1)
TEST_IDDPP = COR_IDDPP+MAX(V(TEST))
COR_I = 0.5*(SGN(COR_IDDAVG+COR_IDDPP-0.1)+1)
TEST_I = COR_I+MAX(V(TEST))
COR_IV = 0.5*(SGN(COR_I+COR_V-0.1)+1)
TEST_IV = COR_IV+MAX(V(TEST))
```

The hystograms of *TEST_Fo*, *Test_ampl*, *TEST_Iddavg* and *TEST_Iddpp* are built for the functional and the I_{dd} testing. The fault coverage FC is defined as a ratio of the recognized faults to the total number of modeled faults.

$$\mathrm{FC} = \frac{T_2}{T_1 + T_2} \tag{2}$$

where T_1 is the number of nonrecognized faults and T_2 is the number of recognized faults. The non-faulty limits of the test quantities are obtained using *Monte Carlo* simulation defining uniform distribution for the resistors and capacitors with design tolerances of 5%. The histograms for the test quantities are obtained. The total number of runs is 1000 for the considered example. The histogram of *Test_Iddpp* is presented in Fig. 8. The number of recognized faulty variants is 760, the nonrecognized faulty variants is 210. Hence the fault coverage *FC* is 78.35%. The histogram of *Test_iv* for the combined test is shown in Fig. 9. 170 runs are skipped due to the loss of oscillations. The fault coverage is 98.94%.

The results for the functional test, the I_{dd} test and the combined (functional and I_{dd}) test are presented in Table 1. As a result, the optimal test node set and test quantities, maximizing the fault coverage and the fault observability, can be selected.

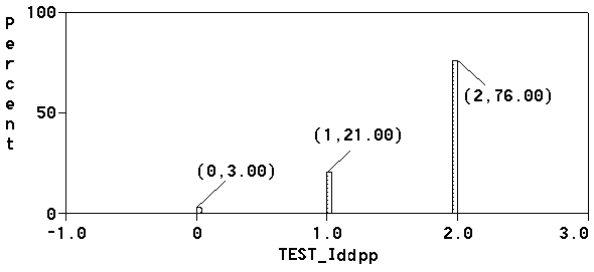


Fig. 8. The histogram of TEST_Iddpp

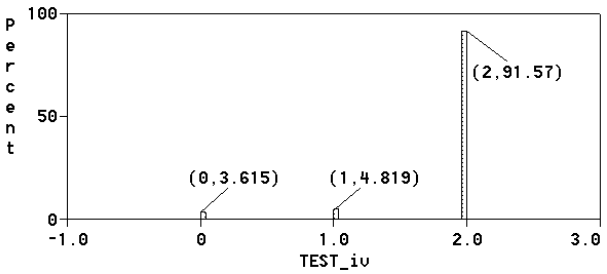


Fig. 9. The histogram of Test_iv

Table 1. Test results

Functional test		Idd test		Combined test	
Test_Fo	91.75%	*Test_Iddavg*	60.8%		
Test_ampl	70.1%	*Test_Iddpp*	78.35%		
Test_V	**93.8%**	*Test_I*	**82.47%**	Test_iv	**98.94%**

4. CONCLUSIONS

A computer-aided approach has been developed to automated fault observability investigation in analog circuits based on statistical analysis. The fault coverage is obtained using *Monte Carlo* simulation of the diagnosis model taking into account the design tolerances. The computer realization of the diagnosis approach is performed using the *Cadence PSpice* circuit simulator and *Cadence Prob*e graphical analyzer. Passive elements from the *BREAKOUT* library with user-defined tolerance deviations are used in order to control the multiplicity of generated faults and to select the correct variants. Based on macrodefinitions and postprocessing in *Cadence Probe*, the percentage of the correct, faulty and recognized faulty variants is automatically obtained for the defined design specifications. In this way, the fault observability can be compared for different test node selections and different design specifications.

REFERENCES

[1] M. Slamani and B. Kaminska, “Fault Observability Analysis of Analog Circuits in Frequency Domain”, IEEE Transactions on Circuits and Systems II. Analog and Digital Signal Processing, Feb. 1996, Vol. 43, No. 2, pp. 134 – 139.

[2] D. Papakostas, V. Kosmidis and A. Hatzopoulos, “Analog Fault Detectability Based on Statistical Circuit Analysis”, Proceedings of the Third IEEE Intern. Conf. on Electronics, Circuits, and Systems ICECS'96, 13-16 Oct. 1996, Rodos, Greece, Vol.2, pp. 1076 – 1079.

[3] A. Raghunathan, J. Chun, J. Abraham and A. Chatterjee, “Quasi-Oscillation Based Test for Improved Prediction of Analog Performance Parameters”, Proceedings of the International Test Conference, ITC 2004, 26-28 Oct. 2004, pp. 252 – 261, 2004.

[4] K. Arabi and B. Kaminska, ”Oscillation-Test Methodology for Low-Cost Testing of Active Analog Filters”, IEEE Transactions on Instrumentation and Measurement, Vol. 48, No. 4, pp. 798 – 806. 1999.

[5] D. Papakostas and A. Hatzopoulos, “Analogue Fault Detectability Comparison Between Power Supply Current and Output Voltage Magnitude and Phase Spectrum Components”, Electronics Letters, Vol. 40, No. 8, pp. 457 – 458, 2004.

[6] E. Dimitrova, E. Gadjeva, A. Van den Bossche and V. Valchev, ”A Model-Based Approach to Automatic Diagnosis Using General Purpose Circuit Simulators”, IEEE International Symposium on Industrial Electronics, ISIE 2006, Montreal, Que., Canada, 9-13 July 2006, pp. 2972 – 2977.

MULTIDISCIPLINARY APPROACH FOR TEACHING COMPUTER-AIDES DESIGN IN COMMUNICATIONS USING MULTIPLE SOFTWARE TOOLS

Galia Marinova

Department Name, Technical University - Sofia, bul. Kliment Ohridski No.8, Sofia 1000, Bulgaria, phone: +3592 965 31 88, e-mail: gim@tu-sofia.bg

***Abstracts**: The paper deals with the application of a multidisciplinary approach for increasing the motivation of students in Telecommunications faculty in Technical University – Sofia by helping them to find out the connections between Computer-aided design (CAD) and the other courses in the Bachelor degree program. Multidisciplinary approach helps students to combine different CAD methods and tools to solve complex tasks in system design. It gives them the abilities to compare different tools and techniques and to determine the best set for solving a concrete practical task. The multidisciplinary approach is achieved by enlarging the existing Knowledge base with new techniques and tools. The diversity of tools permits to cover better the Telecommunications syllabus.*

Keywords: FilterCAD, PAC Designer, MMICAD

1. INTRODUCTION

The Computer-aided design laboratory in the Telecommunications faculty in Technical University – Sofia has long years experience with teaching ORCAD/ PSpice tools for analog, digital and A/D mixed circuit design, device modeling and statistical design. The tool ORCAD/Layout is also traditionally used for PCB design. A Knowledge base for designing different types of circuits is supported to help learning different applications. It contains data for more than 250 circuits and helps the project-based learning approach described in [6]. During the last years a special course for Computer-aided design of digital communication circuits using VHDL is proposed for bachelor degree students and it is accepted with interest and satisfaction by them. This course focuses on the application of tools as ISE Webpack (XILINX), Quartus II (ALTERA) and WARP (CYPRESS). A Learning environment was made available on the web. It helps students in this course applying the model-based design approach described in [1].

The goal is to provide the CAD courses with a larger Knowledge base in the Learning environment which corresponds to the main courses in the Bachelor degree program in Telecommunications. The courses for CAD in communications consider as acquired some basic notions and methods as well as some practical experience with the PSpice tool, taught in the course *"Theoretical Electrical Engineering"*. The application of a multidisciplinary approach in the courses for CAD in communications helps students in solving design problems close to real practical tasks.

The multidisciplinary approach in Computer-aided design in communications was achieved by enlarging the Knowledge base described in [2-5] by several new techniques and software tools.

2. MULTIDISCIPLINARY APPROACH

Table 1 shows the correspondence between the software tools taught in the courses for CAD in communications and the other subjects in the Telecommunications Bachelor degree syllabus [16].

Table 1

Software tool in the Knowledge base	**Application**	**Subject taught in Telecommunications Bachelor degree program**
ORCAD – Capture, PSpice A/D, Layout, Versions: Eval and 10.5	Analog, digital and analog-to-digital mixed simulation	Electrotechnics Analog circuits and systems Power supply design Communication nets (Digital filters design) Pulse and digital circuits and systems
	PCB design	Construction and reliability of communication equipment Computer-aided manufacturing
	Statistical design, Design for manufacturability	Computer-aided manufacturing Construction and reliability of communication equipment
AMPredictor	Electromagnetic compatibility and Signal integrity estimation	Construction and reliability of communication equipment
PAC Designer (LATTICE)	Design of programmable analog circuits	Analog circuits and systems Measurement in communications
FilterCAD	Analog filters synthesis	Communication nets (Analog filter design)
MMICAD	RF and microwave device and circuit design	Radiocommunications Radiocommunications systems Microwave devices and circuits
MICROWIND	Microelectronics layout design and simulation, CMOS technology	Microelectronics
ISE Webpack (XILINX) Quartus II (ALTERA) WARP, PSoC (CYPRESS)	Design of digital programmable circuits – PLD, CPLD, FPGA	Signals and systems Digital signal processing Communication nets (Digital filters design) Pulse and digital circuits and systems Digital communications
MATLAB	Specification definition and verification Model based design	Digital signal processing

3. KNOWLEDGE BASE WITH COMPUTER-AIDED DESIGN TECHNIQUES AND TOOLS

The paper presents the e-content in the Knowledge base developed for a set of software tools used for design of communication circuits:

- ORCAD – Capture, PSpice A/D, Layout, Versions: Eval and 10.5,
- AMPredictor Signal Integrity Analyzer,
- PAC Designer (Lattice) [14],
- FilterCAD (Linear Technology) [9],
- MMICAD [10], MICROWIND [8],
- ISE Webpack (XILINX) [11],
- Quartus II (ALTERA) [12],
- WARP (CYPRESS),
- PSoC (Programmable system-on-a-chip design (CYPRESS) [13],
- MATLAB [15].

The e-content for each software tool is structured as follows:

- Main possibilities,
- Step-by-step guide for the main applications,
- Library with examples for design of communication circuits,
- Methodology and examples for design of communication systems using multiple software tools.

The main characteristics of the new tools in the Knowledge base (PAC Designer, FilterCAD and MMICAD) are discussed in brief.

The software system PAC Designer permits three types of programmability for analog circuits:

- Programmable functions – Amplification, Conversion and Filtering;
- Programmable interconnect – Single/multiple architectures, reconfigurable architectures and 100% routing on-chip;
- Programmable characteristics – Gain, Bandwidth, Offset and Thresholds.

The tool ISP [7] permits in-system-programming of the analog integrated circuits (ICs) from the families ispPAC10, ispPAC20 and ispPAC80. The ICs ispPAC10 and ispPAC20 are composed from Instrument amplifier PACell and Output amplifier PACell forming a Filter/Summation PACblock. The IC ispPAC80 is a Programmable low pass filter. The main applications of the programmable analog circuits designed with PAC Designer are gain/attenuation, sum/difference, filtering/integrating or other linear signal processing. For example they can be used as attenuators in robotics, lossy integrators in imaging systems or biquad filters in industrial control.

Some typical designs with PAC Designer are:

- Biquad filter implementation,

- Precision ladder filter generation,
- Interface for bridge sensor to 12-bit ADC,
- Bridge measurement,
- Law cost temperature measurement,
- Advanced voltage monitoring
- Temperature monitoring.

PAC Designer and ISR tools can be very useful to illustrate some topics from the courses *"Analog circuits and systems"* and *"Measurement in communications"*.

The software system FilterCAD permits the synthesis of Lawpass, Highpass, Bandpass and Notch filters given the Stopband/Passband parameters. The approximations used are Butterwort, Bessel, Elliptic, Chebychev and Min Q Elliptic. The order of the filter can be predefined or calculated for the different approximations. The power supply is specified by the user and a set of ICs from Analog device (for example 1066, 1067, etc.) are proposed for implementation. The schematic of the filter realized is visualized and realistic values of the resistors and their tolerances can be chosen by the user. The circuit realized is simulated in frequency and time domain. The input signal in time domain can be Step, Pulse or Sine.

FilterCAD can be a very useful tool to illustrate analog filter synthesis which is taught in the course *"Communication nets"*.

The software system MMICAD performs linear analysis and circuit design in radiocommunications. The system integrates a set of tools with different possibilities:

- **Coaxial Resonator Design** is a tool for coaxial transmission line element design. Inductors and resonators can be designed.
- **Dielectric Resonator Design** is a tool, which given the resonance frequency of the dielectric resonator to be designed, permits the selection of the resonator geometry (Disc or Cylindrical), the resonator support, the resonator material with a specific dielectric constant and the cavity type (Ideal cavity or Thin substrate).
- **Capacitor Design** is a tool for design of capacitors which permits the user to select the model (Dicap, Border Cap, GapCap and Multilayer), the circuit configuration (Series, Shunt and TLine), the material, the dimensions (Length, Width and Thickness) and the temperature. Six parameters of the capacitor can be simulated in frequency area: resistance R, Reactance X, C Leq, Q Factor, S11 and S12.
- **EEZMatch** is a tool for design of matching networks and similar circuits using the Smith Chart. Multiple topologies are proposed and can be tested. The sensitivities to component values and the frequency effects can be estimated. The netlist output of the matching circuit is generated in MMICAD format.

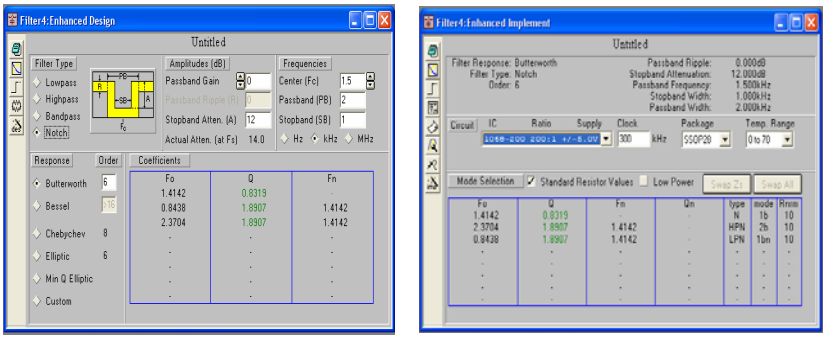
- **MMICAD Layout** is a tool for topology design of microwave and radio-communication circuits.
- **MMICAD Yield Analysis** is a tool for yield prediction, optimization and yield driven design.

MMICAD and its tools is a very suitable CAD system for illustrating topics in the courses *"Radiocommunications", "Radiocommunications systems"* and *"Microwave devices and circuits"*.

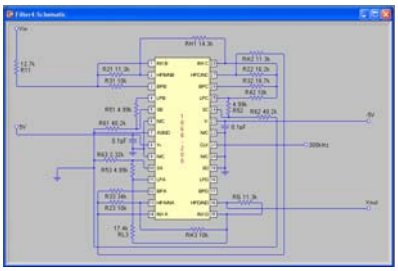
4. EXAMPLES OF THE E-CONTENT IN THE KNOWLEDGE BASE WITH COMPUTER-AIDED DESIGN TECHNIQUES AND TOOLS

Four examples are given to illustrate the application of the tools PAC Designer, FilterCAD and MMICAD for solving tasks coming from different courses in Tele-communications syllabus.

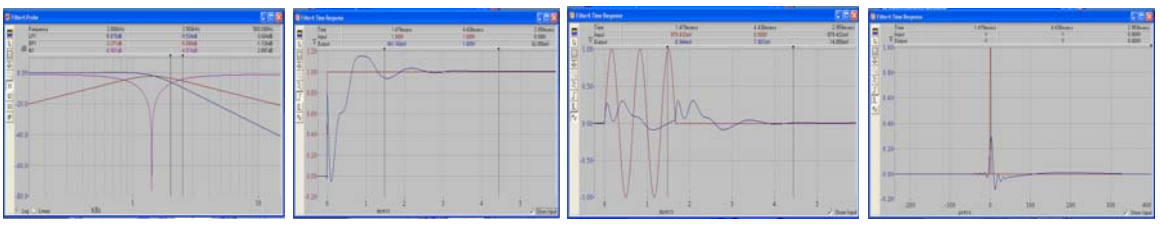
Example 1. Design of a Notch filter in FilterCAD



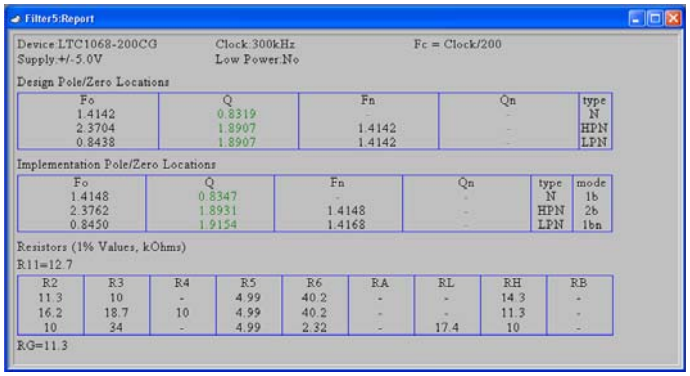
Specification, approximation and synthesis of a notch filter and realization of the Notch filter on IC 1068-4002



Schematic of the notch filter designed



Simulation results in frequency area and in time area – sine, pulse and step response



Report from FilterCAD

Fig. 1. Design of a Notch filter in FilterCAD

Fig.1 illustrates the application of the tool FilterCAD for designing a notch filter with center frequency Fc = 1.5 kHz. The specification, the approximation, the realization, the schematic generation, the parameter values tuning and the simulation results in frequency and time area are shown. The final report from FilterCAD gives the information for the implementation IC LTC1068-200CG, the power supply ±5V, the center frequency to clock frequency ratio Fclock/Fc = 200, the design pole/zero locations and the implementation pole/zero locations, the resistor values and the 1% tolerance chosen for the resistors.

Example 2. Design of Lowpass and Bandpass filters with programmable analog circuits with PAC Designer

Fig. 2 presents the design of Lowpass and Bandpass filters with programmable analog circuits with PAC Designer and the Gain and Phase curves of both filters in frequency domain. The design is based on biquad filter and it's realized on the IC ispPAC10. The parameters programmed are the gain, the amplifier feedback and the interconnections. A multistage architecture is used.

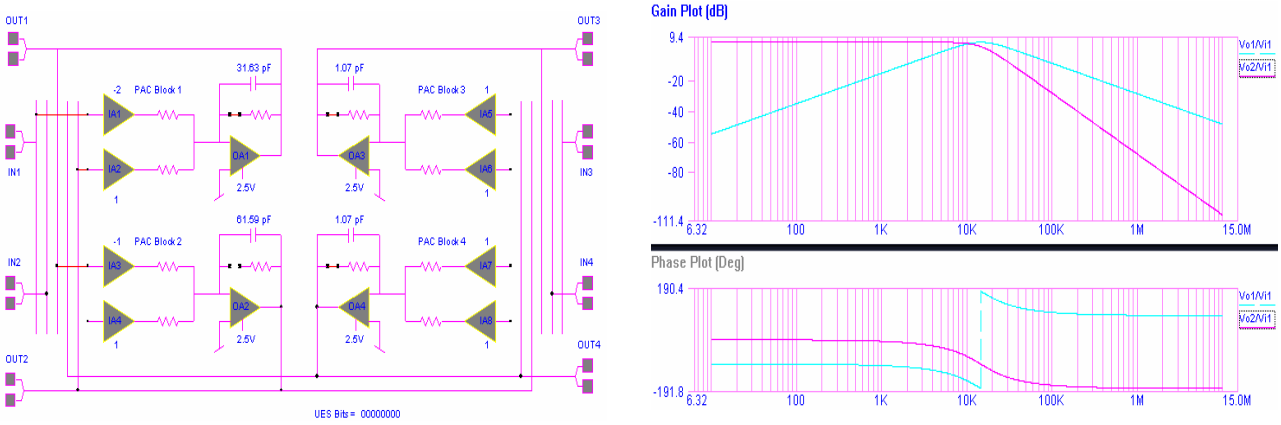


Fig. 2. Design of Lowpass and Bandpass filters with programmable analog circuits with PAC Designer – circuit and simulation results

Example 3. Design of inductor, resonators and capacitor in MMICAD

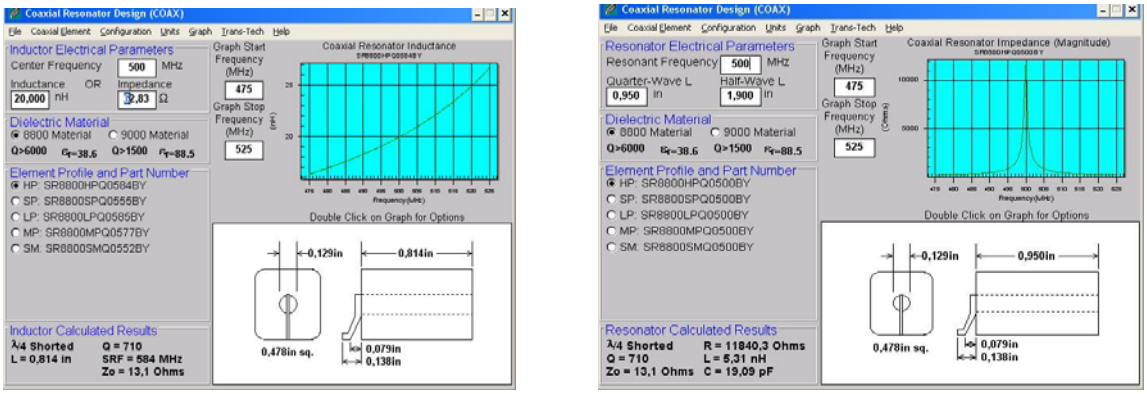


Fig. 3. Coaxial resonator design – inductor and resonator

Fig. 3 shows two applications of coaxial resonator design in MMICAD - inductor and resonator. The inductor designed has an inductance of 20nH and the resonator designed has a resonant frequency of 500MHz. Simulation curves in frequency area for the inductor and the resonator, are also visualized.

Fig. 4 illustrates the Cavity configuration for dielectric resonator design in MMICAD. The calculation of dimensions and the simulation of Resonant frequency over Cavity tuning F(Lr) are visualized.

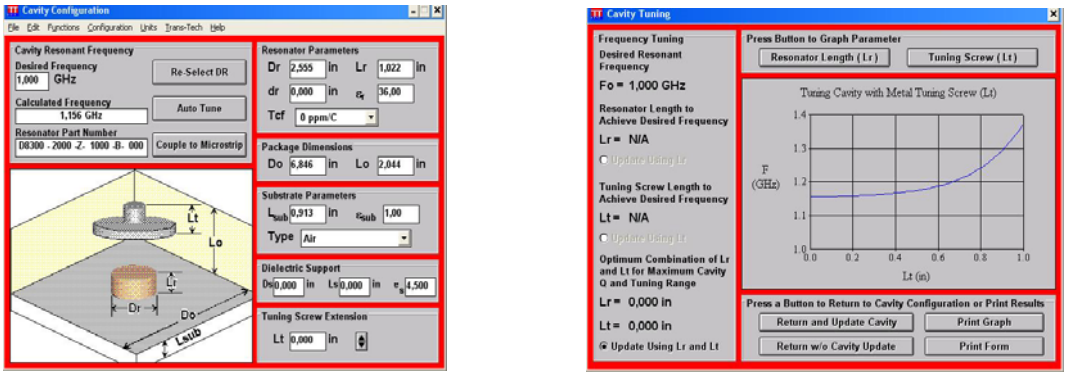


Fig. 4. Dielectric resonator design – calculation of dimensions and simulation of resonant frequency over cavity tuning F(Lr)

Fig. 5 illustrates the specification and the simulation results for Capacitor design in MMICAD. The capacitor is modeled in Shunt Transmission line configuration. For the capacitor designed, six simulation curves in frequency area are visualized (R, X, C Leq, Q Factor, S11, S12).

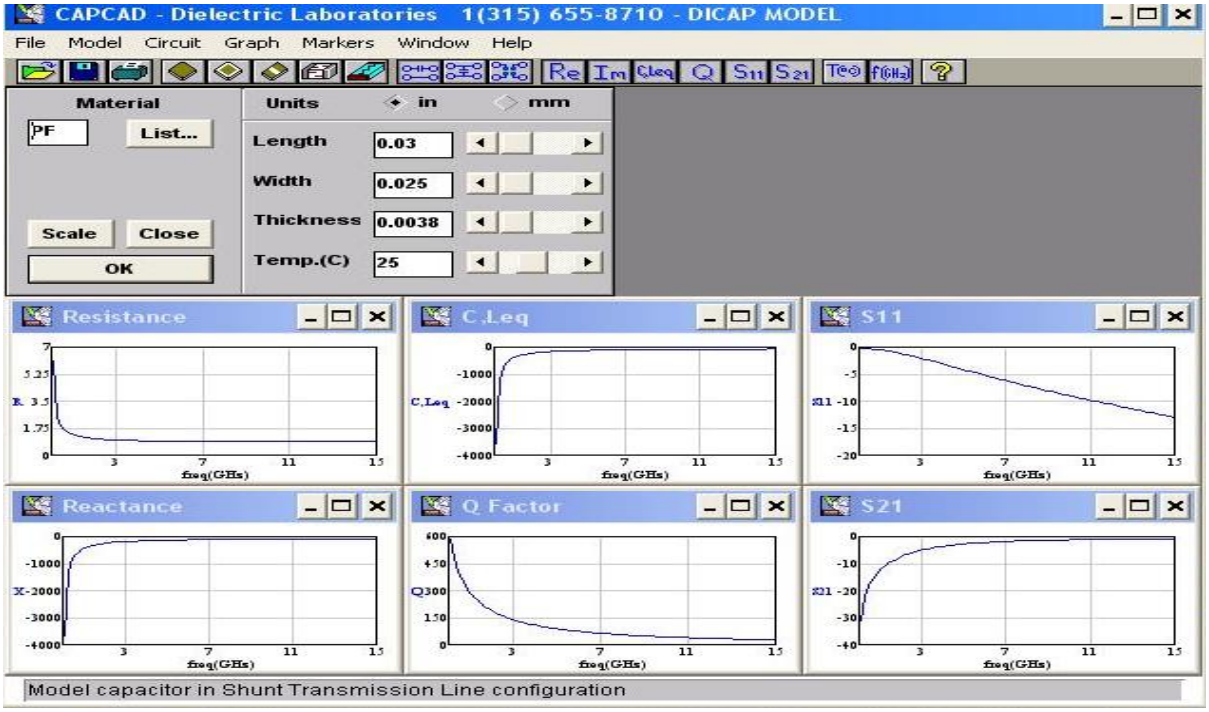


Fig. 5. Capacitor design – geometry definition and simulation results

Example 4. Impedance matching with EEZ Match

The Smith chart diagram of a matching circuit, the netlist and the models extracted from the Smith chart diagram are shown on Fig.6. The netlist can be exported for MMICAD simulation. EEZMatch permits to experience multiple matching circuit solutions and to select the most suitable for the design purposes.

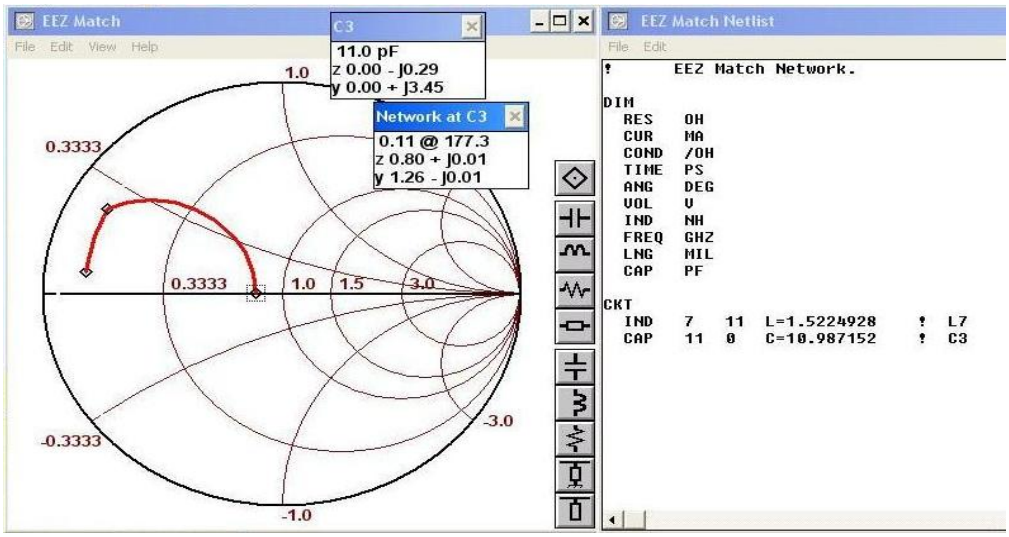


Fig. 6. Impedance matching with EEZ Match

5. CONCLUSION

The application of the e-content developed permits to enlarge the knowledge of students from Telecommunications faculty in Technical University – Sofia with additional software tools and methods, besides the PSpice A/D simulator and the development systems for FPGA and CPLD design as ISE Webpack, Quartus and Warp, which are taught in the courses for computer-aided design in the Telecommunications Bachelor degree syllabus. Versions in Bulgarian, English, German and French languages are foreseen to help teaching in foreign language faculties in Technical University - Sofia.

The influence of multidisciplinary approach on students' motivation will continue to be estimated using different feedbacks as survey, number of students choosing optional courses for CAD in communications, discussions with students and students' forums.

REFERENCES

[1] L. Belanger, "DSP as the key technology driver for embedded systems electronics", CTO, Lyrtech, IEEE, MWS/NEWCAS Circuits and Systems 2007. Montreal, Canada

[2] G. Marinova and D. Dimitrov, "Learning optimal synthesis of voltage regulator circuits through comparative study in PSpice", XV International Symposium on Theoretical Electrical Engineering, 22-24 June 2009, Lubeck, Germany, pp. 211-215

[3] G. Marinova, "Learning environment for design and verification of communication circuits realized on FPGA", XV International Symposium on Theoretical Electrical Engineering, 22-24 June 2009, Lubeck, Germany, pp. 224-226

[4] G**.** Marinova, "E-content in the Learning Environment for Design and Verification of Communication Circuits", TELFOR 2009, Belgrade, pp. 110-114.

[5] G. Marinova, "E-content for computer-aided design of Communication circuits with PSpice simulator", Proc. of ICEST'2010, 23-26 June 2010, Ohrid, Macedonia, pp.943-946

[6] P. Tobin, "PSpice and project-based learning at ordinary degree level", ICEE 2008, Second International Conference on Engineering education, 27-31 July 2008, Pecs-Budapest, Hungary, www.compeng.dit.ie/staff/ptobin/ICEE2008tobin.pdf

[7] New Dimensions in ISP, Programmable analog circuits, Lattice semiconductor, September 1999

[8] http://www.microwind.net/, C-MOS design and simulation tool

[9] http://www.linear.com/designtools/software/filtercad.jsp

[10] http://www.optotek.com/mmicbro7.htm, MMICAD linear simulator

[11] http://www. xilinx.com

[12] http://www.altera.com

[13] http://www.cypress.com

[14] http://www.latticesemi.com/products/designsoftware/pacdesigner/index.cfm,PAC-Designer

[15] http://www.mathworks.com

[16] Bachelor degree program in Telecommunications, Telecommunications faculty, Technical University–Sofia,http://fktt.tu-sofia.bg/new_site/documents/creditBC_BG_5_23012008.pdf

RECURRENT NEURAL NETWORKS FOR LINEAR AND QUADRATIC OPTIMIZATION WITH k-WTA APPLICATION

Qingshan Liu
School of Automation, Southeast University, Nanjing 210096, China
email: qsliu@seu.edu.cn

Jun Wang
Department of Mechanical and Automation Engineering
The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
email: jwang@mae.cuhk.edu.hk

Over the past two decades, recurrent neural networks have been shown to be very powerful computational models for optimization. A number of recurrent neural network models have been designed for solving linear and quadratic programming problems. Based on the Karush-Kuhn-Tucker optimality conditions, recurrent neural networks are designed to obtain the optimal solutions. The present chapter presents several recurrent neural networks with discontinuous activation functions for solving linear and quadratic programming problems. Moreover, as a special application, the k-winners-take-all (k-WTA) operation is investigated and some recurrent neural networks are introduced to perform this problem.

Keywords: Recurrent neural networks, Linear programming, Quadratic programming, k-winners-take-all

1. INTRODUCTION

Optimization problems arise in numerous of science and engineering applications, such as robot control, manufacturing system design, signal and image processing, and pattern recognition [2, 3]. Over the years, a variety of numerical algorithms have been developed for solving linear and quadratic programming problems, such as the simplex methods, active set methods, and interior point methods [2]. As parallel computational models for optimization, recurrent neural networks have received substantial attention over the past two decades (e.g., see [10][16][22][30][42] and references therein). In many applications, real-time optimal solutions are usually imperative, such as on-board signal processing [8] and robot motion planning and control [43].

In the past two decades, recurrent neural networks for optimization and their engineering applications have been widely investigated. Tank and Hopfield [30] proposed a recurrent neural network for solving linear programming problems which motivated the development of neural networks for optimization. In 1988, the dynamical canonical nonlinear programming circuit (NPC) was introduced by Kennedy and Chua [16] for nonlinear programming by utilizing a finite penalty parameter, which can generate the approximate optimal solutions. From then on the research

on NPC has been well developed and plenty of neural network models have been designed for optimization problems (e.g., see [5, 26, 32, 33, 37, 38, 41] and references therein). Among them, Zhang and Constantinides [47] proposed the Lagrangian network based on the Lagrangian method which had a two-layers structure and this neural network was globally convergent to an optimal solution if only the objective function was strictly convex. The deterministic annealing neural network by Wang [32, 33] was developed for linear and convex programming. Based on the primal-dual method, Xia [36, 39, 40] proposed some primal, dual, and primal-dual neural networks for solving linear and quadratic programming problems. Later on, the projection method [12] was introduced to construct recurrent neural networks for linear and nonlinear optimization, such as the models proposed by Xia *et al.* [37], which are globally convergent to exact optimal solutions of convex optimization problems. Recently, Forti *et al.* [10, 11] investigated the generalized NPC (G-NPC) for non-smooth optimization, which can be considered as a natural extension of NPC. The delayed neural networks [6, 19, 24, 44] were developed for solving convex quadratic programming problems. In order to reduce the model complexity, the dual, simplified, and improved dual neural networks were introduced for solving convex programming problems with dual variables only [14, 25]. More recently, some one-layer recurrent neural networks [21, 22] were developed for solving linear and quadratic programming problems. The one-layer recurrent neural networks, which number of neurons is equal to that of decision variables in the programming problems, have more simply architecture complexity than the other neural networks such as Lagrangian network and primal-dual networks.

The k-winners-take-all (k-WTA) operation selects the k largest inputs out of n inputs ($1 \leq k \leq n$), and it is a generalization of winner-take-all (WTA) operation, which is to select the maximum from a collection of input signals. The k-WTA networks have many applications, such as associative memories [13], signal processing [1], cooperative models of binocular stereo [29], Fukushima's neocognitron for feature extraction, etc. [46]. Selecting the k largest elements from a data set of numbers is a key task in decision making, pattern recognition, and competitive learning networks (see [17, 31, 45], for example). The k-WTA networks find many applications in telecommunications, particularly for controlling data packet switches [4]. As the number of inputs increases and/or the selection process should be operated in real time, parallel algorithms and hardware implementation are desirable [16, 30, 32]. For these reasons, there have been many attempts to design very large scale integrated (VLSI) circuits to perform k-WTA operations [15, 27, 35].

Recently, the k-WTA networks based on circuit implementation have been widely investigated (see, e.g., [14, 20, 23, 25, 28, 34]). Generally, the k-WTA problems can

be formulated into linear or quadratic programming problems, then the corresponding neural networks are designed to solve the problems. This chapter will introduce some k-WTA networks constructed from the linear and quadratic programming problems.

2. Related Definitions and Lemmas

Definition 1. *Suppose $\mathcal{D} \subset \mathbb{R}^n$. $F : x \mapsto F(x)$ is called a set-valued map from $\mathcal{D} \hookrightarrow \mathbb{R}^n$, if to each point x of a set $\mathcal{D}$, there corresponds to a nonempty closed set $F(x) \subset \mathbb{R}^n$.*

Definition 2. *A function $\varphi : \mathbb{R}^n \to \mathbb{R}$ is said to be Lipschitz near $x \in \mathbb{R}^n$ if there exist $\varepsilon, \delta > 0$, such that for any $x', x'' \in \mathbb{R}^n$ satisfies $\|x' - x\|_2 < \delta$ and $\|x'' - x\|_2 < \delta$, we have $|\varphi(x') - \varphi(x'')| \leq \varepsilon \|x' - x''\|_2$, where $\|\cdot\|_2$ denotes the l_2-norm in $\mathbb{R}^n$. If φ is Lipschitz near any point $x \in \mathbb{R}^n$, then φ is also said to be locally Lipschitz in $\mathbb{R}^n$.*

Assume that φ is Lipschitz near x. The generalized directional derivative of φ at x in the direction $v \in \mathbb{R}^n$ is given by

$$\varphi^0(x; v) = \limsup_{\substack{y \to x \\ s \to 0^+}} \frac{\varphi(y + sv) - \varphi(y)}{s}.$$

The Clarke's generalized gradient of f is defined as

$$\partial\varphi(x) = \{y \in \mathbb{R}^n : \varphi^0(x; v) \geq y^T v, \forall v \in \mathbb{R}^n\}.$$

When φ is locally Lipschitz in $\mathbb{R}^n$, φ is differentiable for almost all (a.a.) $x \in \mathbb{R}^n$ (in the sense of Lebesgue measure). Then, the Clarke's generalized gradient of φ at $x \in \mathbb{R}^n$ is equivalent to

$$\partial\varphi(x) = K\{\lim_{n \to \infty} \nabla\varphi(x_n) : x_n \to x, x_n \notin \mathcal{N}, x_n \notin \mathcal{D}\},$$

where $K(\cdot)$ denotes the closure of the convex hull of the corresponding set, $\mathcal{N} \subset \mathbb{R}^n$ is an arbitrary set with measure zero, and $\mathcal{D} \subset \mathbb{R}^n$ is the set of points where φ is not differentiable.

Definition 3. *A function $\varphi : \mathbb{R}^n \to \mathbb{R}$, which is locally Lipschitz near $x \in \mathbb{R}^n$, is said to be regular at x if there exists the one-sided directional derivative for any direction $v \in \mathbb{R}^n$ which is given by*

$$\varphi'(x; v) = \lim_{\xi \to 0^+} \frac{\varphi(x + \xi v) - \varphi(x)}{\xi},$$

and we have $\varphi^0(x; v) = \varphi'(x; v)$. The function φ is said to be regular in $\mathbb{R}^n$ if it is regular for any $x \in \mathbb{R}^n$.

Regular function is very important in the Lyapunov approach and non-smooth analysis, which has been studied in the literature (see, e.g. [7, 9] for references). In particular, a non-smooth convex function on $\mathbb{R}^n$ is regular at any $x \in \mathbb{R}^n$. For a finite family of functions $\varphi_i (i = 1, 2, \ldots, n)$, which are regular at x, we have $\partial(\sum_{i=1}^n \varphi_i)(x) = \sum_{i=1}^n \partial\varphi_i(x)$.

To tackle an optimization problem by using a recurrent neural network, the key lie in the construction of its dynamic equations such that the equilibrium points correspond to the desired optimal solutions. In general, a neural network can be described by the following dynamical equation governed by

$$\frac{dx}{dt} = \psi(x), \quad x(t_0) = x_0. \tag{1}$$

If $\psi(x)$ is discontinuous on $\mathbb{R}^n$, it is necessary to explain what is meant by a solution of (1). A possible definition is that of Filippov [9].

Definition 4. *A set-valued map:* $\phi : \mathbb{R}^n \hookrightarrow \mathbb{R}^n$ *is defined as*

$$\phi(x) = \bigcap_{\delta>0} \bigcap_{\mu(\mathcal{N})=0} K[\psi(\mathcal{B}(x, \delta) - \mathcal{N})],$$

where $\mathcal{N}$ *is an arbitrary set with measure zero,* $\mu(\mathcal{N})$ *is the Lebesgue measure of set* $\mathcal{N}$*, and* $\mathcal{B}(x, \delta) = \{y \in \mathbb{R}^n : \|y - x\| \leq \delta\}$*. A Filippov solution of* (1) *is an absolutely continuous function* $x(t)$ *defined on an interval* $[t_0, t_1](t_0 \leq t_1 \leq +\infty)$*, which satisfies* $x(t_0) = x_0$ *and differential inclusion:*

$$\frac{dx}{dt} \in \phi(x), \quad \text{a.a. } t \in [t_0, t_1].$$

Definition 5. $\bar{x}$ *is said to be an equilibrium point of system* (1) *if*

$$0 \in \phi(\bar{x}). \tag{2}$$

Definition 6. *The neural network* (1) *is said to be globally convergent to an equilibrium point if for any trajectory* $x(t)$ *of the neural network with initial point* $x(t_0) \in \mathbb{R}^n$*, there exists an equilibrium point* $\bar{x}$ *such that* $\lim_{t\to+\infty} x(t) = \bar{x}$*.*

3. For Linear Programming

Consider the following linear programming (LP) problem as:

$$\begin{array}{ll} \text{minimize} & c^T x, \\ \text{subject to} & Ax = b, \\ & l \leq x \leq u, \end{array} \tag{3}$$

where $x \in \mathbb{R}^n$ is the vector of decision variables; $A \in \mathbb{R}^{m\times n}$ is a full row-rank matrix (i.e., rank$(A) = m$, $m < n$); $c, l, u \in \mathbb{R}^n$ and $b \in \mathbb{R}^m$. Here, if $l = -\infty$ or $h = \infty$, the inequality constraints are one-side.

3.1 Model Description and Convergence Results

According to the Karush-Kuhn-Tucker (KKT) conditions [2], x^* is an optimal solution of (3), if and only if there exist $y^* \in \mathbb{R}^m$ and $z^* \in \mathbb{R}^n$ such that $(x^*, y^*, z^*)^T$ satisfies the following optimality conditions:

$$c + A^T y + z = 0, \tag{4}$$

$$Ax = b, \tag{5}$$

$$\begin{cases} z_i \geq 0, & \text{if } x_i = u_i, \\ z_i = 0, & \text{if } x_i \in (l_i, u_i), \\ z_i \leq 0, & \text{if } x_i = l_i. \end{cases} \tag{6}$$

From (4), we have

$$x = x - c - A^T y - z. \tag{7}$$

Substituting (7) into (5), we have

$$Ax - Ac - AA^T y - Az = b.$$

Since A is full row-rank, AA^T is invertible. Then

$$y = (AA^T)^{-1}(Ax - Az - Ac - b). \tag{8}$$

Substituting (8) into (4), we have

$$c + A^T(AA^T)^{-1}(Ax - Az - Ac - b) + z = 0. \tag{9}$$

Let $P = A^T(AA^T)^{-1}A$ and $q = -(I - P)c + A^T(AA^T)^{-1}b$. Then (9) can be written as

$$Px + (I - P)z - q = 0, \tag{10}$$

where I is an identity matrix. The matrix P, called the projection matrix, has the following properties by simple calculation.

Lemma 1. (i) *P is symmetric, $P^2 = P$ and $(I - P)^2 = I - P$;*

(ii) *For any $\zeta > -1$, $I + \zeta P$ is invertible and $(I + \zeta P)(I - P) = I - P$.*

Based on above analysis, a recurrent neural network for solving (3) is described as following dynamical system:

$$\epsilon \frac{dx}{dt} = -Px - \sigma(I - P)g(x) + q, \tag{11}$$

where ϵ is a positive scaling constant, σ is a nonnegative gain parameter, and $g(x) = (g_1(x_1), g_2(x_2), \ldots, g_n(x_n))^T$ with $g_i(x_i)$ defined as the following discontinuous ac-

tivation function

$$g_i(x_i) = \begin{cases} 1, & \text{if } x_i > u_i, \\ [0,1], & \text{if } x_i = u_i, \\ 0, & \text{if } x_i \in (l_i, u_i), \\ [-1,0], & \text{if } x_i = l_i, \\ -1, & \text{if } x_i < l_i. \end{cases} \quad (i = 1, 2, \ldots, n) \tag{12}$$

The following lemma gives a necessary condition for the existence of equilibrium point of neural network (11).

Lemma 2. *[21] If there exists an equilibrium point of neural network* (11)*, then*

$$\sigma \geq \begin{cases} 0, & \text{if} \quad (I-P)c = 0; \\ \dfrac{c^T(I-P)c}{\|(I-P)c\|_1}, & \text{if} \quad (I-P)c \neq 0, \end{cases}$$

where $\|\cdot\|_1$ *denotes the* l_1*-norm in* $\mathbb{R}^n$*.*

In general, the equilibrium point of neural network (11) may not always be an optimal solution of problem (3). However, the following lemma reveals the relationship between the optimal solution of problem (3) and the equilibrium point of neural network (11).

Lemma 3. *[21]*

(i) *Any optimal solution of problem* (3)*, denoted as* x^**, is an equilibrium point of neural network* (11) *if* $\sigma \geq 0$ *when* $(I-P)c = 0$*, or one of the following conditions holds when* $(I-P)c \neq 0$*:*

(a) $\sigma \geq \|(I-P)c\|_2 / \min^+_{\xi \in X} \|(I-P)\xi\|_2$*, or*

(b) $\sigma \geq c^T(I-P)c / \min^+_{\xi \in X}\{|c^T(I-P)\xi|\}$,

where $\min^+_{\xi \in X}\{c^T(I-P)\xi\}$ *denotes the minimum positive value of the corresponding function and* $X = \{\xi = (\xi_1, \xi_2, \ldots \xi_n)^T \in \mathbb{R}^n : \xi_i = -1,\ 0$ *or* $1, i = 1, 2, \ldots, n\}$.

(ii) *Any equilibrium point of neural network* (11)*, denoted by* $\bar{x}$*, is an optimal solution of problem* (3)*, if one of the following conditions holds:*

(a) $\bar{x} \in \Omega = \{x \in \mathbb{R}^n : l \leq x \leq u\}$*, or*

(b) $\bar{x}$ *is unique and* σ *satisfies the conditions in* (i).

The global convergence of neural network (11) is guaranteed by the following theorem.

Theorem 1. *[21] If the equilibrium point set of neural network* (11) *is not empty, the neural network is globally convergent to an equilibrium point.*

From Lemma 3 and Theorem 1, the optimal solutions of problem (3) can be guaranteed by the following two corollaries.

Corollary 1. *The neural network* (11) *is globally convergent to an optimal solution of problem* (3)*, if the equilibrium point set of* (11) *is included in the set* Ω *defined in Lemma 3*(ii).

Corollary 2. *The neural network* (11) *is globally convergent to an optimal solution of problem* (3)*, if it has a unique equilibrium point and* $\sigma \geq 0$ *when* $(I-P)c = 0$ *or one of the following conditions holds when* $(I - P)c \neq 0$*:*

(i) $\sigma \geq \|(I - P)c\|_2 / \min^+_{\xi \in X} \|(I - P)\xi\|_2$*, or*

(ii) $\sigma \geq c^T(I - P)c / \min^+_{\xi \in X}\{|c^T(I - P)\xi|\}$,

where X *is defined in Lemma 3*(i).

3.2 Simulation Results

Example 1. Consider the following linear programming problem:

$$\begin{array}{ll} \text{minimize} & -2x_1 + x_2 - 4x_3 - x_4, \\ \text{subject to} & 1.2x_1 + 2x_2 - 2x_3 - 1.5x_4 = -4, \\ & x_1 - x_2 + 2.5x_3 + 2x_4 = 2, \\ & -2 \leq x_1, x_2, x_3, x_4 \leq 2. \end{array}$$

The neural network in (11) is utilized to solve this problem. By Lemma 2, a necessary condition for existence of the equilibrium point of (11) is $\sigma \geq 1.9634$. According to the conditions in Corollary 2, to get the optimal solution, the lower bound of σ is estimated as 7.3. Simulation results are depicted in Fig. 1 with $\epsilon = 10^{-5}$ and $\sigma = 8$. It shows that the state variables of the neural network are globally convergent to the unique optimal solution $x^* = (-0.3125, -1.3125, 2, -2)^T$ from 10 random initial points.

4. For Quadratic Programming

In this section, we are concerned with a quadratic programming (QP) problem as follows:

$$\begin{array}{ll} \text{minimize} & f(x) = \dfrac{1}{2}x^T Q x + c^T x, \\ \text{subject to} & Ax = b, \\ & l \leq x \leq u, \end{array} \tag{13}$$

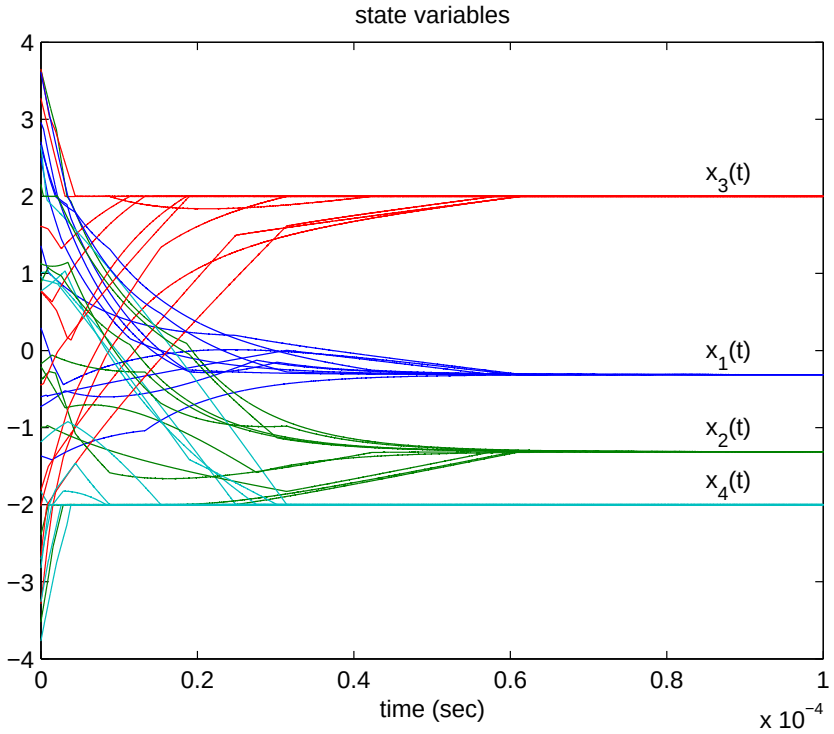


Figure 1: Transient behaviors of the state variables of neural network (11) with $\sigma = 8$ in Example 1.

where $x \in \mathbb{R}^n$ is the vector of decision variables, $Q \in \mathbb{R}^{n\times n}$ is symmetric but not necessarily positive semidefinite, A, b, c, l and u are defined in (3).

4.1 Model Description

According to the Karush-Kuhn-Tucker (KKT) conditions [2], if x^* is a optimal solution of (13), then there exist $y^* \in \mathbb{R}^m$ and $z^* \in \mathbb{R}^n$ such that $(x^*, y^*, z^*)^T$ satisfies the following optimality conditions:

$$Qx + c + A^T y + z = 0, \tag{14}$$

$$Ax = b, \tag{15}$$

$$\begin{cases} x_i = u_i, & \text{if } z_i > 0, \\ l_i \le x_i \le u_i, & \text{if } z_i = 0, \\ x_i = l_i, & \text{if } z_i < 0. \end{cases} \tag{16}$$

If (13) is a convex programming problem, the KKT conditions (14)-(16) are both necessary and sufficient.

From (14), we have

$$x = (I - Q)x - c - A^T y - z. \tag{17}$$

Substituting (17) into (15), we have

$$A(I - Q)x - Ac - AA^T y - Az = b.$$

Because A is full row-rank, AA^T is invertible. Then

$$y = (AA^T)^{-1}(A(I-Q)x - Az - Ac - b). \tag{18}$$

Substituting (18) into (14), we have

$$Qx + c + A^T(AA^T)^{-1}(A(I-Q)x - Az - Ac - b) + z = 0. \tag{19}$$

Let $P = A^T(AA^T)^{-1}A$ and $s = -c + Pc + A^T(AA^T)^{-1}b$, then, (19) can be written as

$$(I-P)z + [(I-P)Q + P]x - s = 0. \tag{20}$$

According to Lemma 1, $I + \zeta P$ is invertible as $\zeta > -1$. By multiplying $I + \zeta P$ in both sides of (20), it follows that

$$(I-P)z + [(I-P)Q + \alpha P]x - q = 0, \tag{21}$$

where $\alpha = 1+\zeta > 0$ is a parameter and $q = (I+\zeta P)s = -c+Pc+\alpha A^T(AA^T)^{-1}b$.

Based on (21), the proposed recurrent neural network model for solving (13) is described as follows:

- State equation

$$\epsilon\frac{dz}{dt} = -(I-P)z - [(I-P)Q + \alpha P]h(z) + q, \tag{22}$$

- Output equation

$$x = ((I-P)Q + \alpha P)^{-1}(-(I-P)z + q), \tag{23}$$

where ϵ is a positive scaling constant and $h(z) = (h_1(z_1), h_2(z_2), \ldots, h_n(z_n))^T$, in which $h_i(z_i)$ is the hard-limiting activation function defined as

$$h_i(z_i) = \begin{cases} u_i, & \text{if } z_i > 0, \\ [l_i, u_i], & \text{if } z_i = 0, \\ l_i, & \text{if } z_i < 0. \end{cases} \tag{24}$$

Here we assume that there exists $\alpha > 0$ such that $(I-P)Q + \alpha P$ is invertible. In fact, if the objective function in (13) is strictly convex on the set defined by the equality constraints, the matrix $(I-P)Q + \alpha P$ is always invertible, which can be guaranteed by the results listed in the next subsection.

From above analysis, for any $z \in \mathbb{R}^n$, the output vector x of the neural network always satisfies the equality constraints. If x^* is an optimal solution of (13), there exists $z^* \in \mathbb{R}^n$ such that z^* is an equilibrium point of system (22). Conversely, if the objective function of (13) is convex on the set $\mathcal{S} = \{x \in \mathbb{R}^n : Ax = b\}$ and z^* is an equilibrium point of system (22), then $x^* = ((I-P)Q + \alpha P)^{-1}(-(I-P)z^* + q)$ is an optimal solution of (13).

In the following, some convergence results of neural network (22) are listed for solving problem (13). The main results have been published in the reference [22].

4.2 Convergence Results

Let $[E]^S$ be the symmetric part of matrix E; i.e., $[E]^S = (E + E^T)/2$. Based on the Lyapunov method, the stability and global convergence of (22) can be obtained.

Theorem 2. *[22] The state vector $z(t)$ of neural network* (22) *is stable in the sense of Lyapunov and globally convergent to an equilibrium point if there exists a constant $s \geq \alpha$ such that $2s[(I - P)Q + \alpha P]^S - Q(I - P)Q - \alpha^2 P$ is positive semidefinite.*

As a special case, if $(I-P)Q+\alpha P$ is positive definite, the conditions in Theorem 2 hold if $s \geq \max\{\alpha, \lambda_{\max}(Q(I - P)Q + \alpha^2 P)/(2\lambda_{\min}([(I - P)Q + \alpha P]^S))\}$, where $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ are respectively the maximum and minimum eigenvalues of corresponding matrixes. Then the following corollary holds.

Corollary 3. *The state vector $z(t)$ of neural network* (22) *is stable in the sense of Lyapunov and globally convergent to an equilibrium point if $(I - P)Q + \alpha P$ is positive definite.*

Remark 1. According to the results in Theorem (2) and Corollary (3), if (13) is a convex quadratic programming problem (i.e., the objective function $f(x)$ is convex on the equality constraints), the output vector $x(t)$ in (23) is globally convergent to an optimal solution if only the conditions in Theorem 2 or Corollary 3 are satisfied by choosing proper α. Moreover, if (13) is a strictly convex quadratic programming problem (i.e., the objective function $f(x)$ is strictly convex on the equality constraints), the output vector $x(t)$ in (23) is always globally convergent to an optimal solution if only α is large enough. We show the details in the rest part of this subsection.

According to the definition of P, the rank of P is m (i.e., rank$(P) = m$). Since $P^2 = P$, there exists an orthogonal matrix Γ such that

$$\Gamma^T P \Gamma = \begin{pmatrix} I & O \\ O & O \end{pmatrix},$$

where O is a zero matrix. Suppose

$$\Gamma^T Q \Gamma = \bar{Q} = \begin{pmatrix} \bar{Q}_1 & \bar{Q}_2 \\ \bar{Q}_2^T & \bar{Q}_3 \end{pmatrix},$$

where $\bar{Q}_1 \in \mathbb{R}^{m\times m}$, $\bar{Q}_2 \in \mathbb{R}^{m\times(n-m)}$ and $\bar{Q}_3 \in \mathbb{R}^{(n-m)\times(n-m)}$.

If the matrix Q in (13) is positive definite, we can choose a sufficiently large value of the parameter α to guarantee the stability of the neural network according to the following theorem.

Theorem 3. *[22] Assume that Q is positive definite. If*

$$\alpha \geq \lambda_{\max}(\bar{Q}_1)/2,$$

then the state vector $z(t)$ of neural network (22) *is globally convergent to an equilibrium point and the output vector $x(t)$ in* (23) *is globally convergent to an optimal solution of problem* (13)*.*

Since $\lambda_{\max}(\bar{Q}_1) \leq \lambda_{\max}(\bar{Q}) = \lambda_{\max}(Q)$, the following corollary is more convenient to estimate the lower bound of parameter α.

Corollary 4. *Assume that Q is positive definite. If*

$$\alpha \geq \lambda_{\max}(Q)/2,$$

then the state vector $z(t)$ of neural network (22) *is globally convergent to an equilibrium point and the output vector $x(t)$ in* (23) *is globally convergent to an optimal solution of problem* (13)*.*

Denote trace(Q) as the trace of Q; i.e., trace$(Q) = \sum_{i=1}^{n} q_{ii}$, where q_{ii} is the ith diagonal element of Q. Then, $\lambda_{\max}(Q) \leq$ trace(Q).

Corollary 5. *Assume that Q is positive definite. If*

$$\alpha \geq \text{trace}(Q)/2,$$

then the state vector $z(t)$ of neural network (22) *is globally convergent to an equilibrium point and the output vector $x(t)$ in* (23) *is globally convergent to an optimal solution of problem* (13)*.*

If the objective function $f(x)$ of problem (13) is not convex everywhere, but strictly convex on the set $\mathcal{S} = \{x \in \mathbb{R}^n : Ax = b\}$, the neural network in (22) and (23) can still solve the problem with proper parameter α based on the following results.

Theorem 4. *[22] Assume that the objective function $f(x)$ of problem* (13) *is strictly convex on the set $\mathcal{S} = \{x \in \mathbb{R}^n : Ax = b\}$. If*

$$\alpha > \lambda_{\max}(\bar{Q}_2\bar{Q}_3^{-1}\bar{Q}_2^T)/4,$$

then the state vector $z(t)$ of neural network (22) *is globally convergent to an equilibrium point and the output vector $x(t)$ in* (23) *is globally convergent to an optimal solution of problem* (13)*.*

Since

$$\begin{aligned}
\lambda_{\max}(\bar{Q}_2\bar{Q}_3^{-1}\bar{Q}_2^T) &\leq \lambda_{\max}(\bar{Q}_2\bar{Q}_2^T)\lambda_{\max}(\bar{Q}_3^{-1}) \\
&\leq \lambda_{\max}(\bar{Q}^2)\lambda_{\max}(\bar{Q}^{-1}) \\
&= \lambda_{\max}(Q^2)\lambda_{\max}(Q^{-1}),
\end{aligned}$$

the following corollary holds.

Corollary 6. *Assume that the objective function* $f(x)$ *of problem* (13) *is strictly convex on the set* $\mathcal{S} = \{x \in \mathbb{R}^n : Ax = b\}$. *If*

$$\alpha > \lambda_{\max}(Q^2)\lambda_{\max}(Q^{-1})/4,$$

then the state vector $z(t)$ *of neural network* (22) *is globally convergent to an equilibrium point and the output vector* $x(t)$ *in* (23) *is globally convergent to an optimal solution of problem* (13).

Since $\lambda_{\max}(Q^2) \leq \text{trace}(Q^2)$, and $\lambda_{\max}(Q^{-1}) = 1/\lambda^+_{\min}(Q)$, where $\lambda^+_{\min}(Q)$ is the minimum positive eigenvalue of Q, we can obtain the following corollary.

Corollary 7. *Assume that the objective function* $f(x)$ *of problem* (13) *is strictly convex on the set* $\mathcal{S} = \{x \in \mathbb{R}^n : Ax = b\}$. *If*

$$\alpha > \frac{\text{trace}(Q^2)}{4\lambda^+_{\min}(Q)},$$

then the state vector $z(t)$ *of neural network* (22) *is globally convergent to an equilibrium point and the output vector* $x(t)$ *in* (23) *is globally convergent to an optimal solution of problem* (13).

As a special case, when $Q = \eta I$, where $\eta > 0$, according to Corollary 6, the output vector $x(t)$ of the neural network is globally convergent to an optimal solution if $\alpha \geq \eta/4$.

4.3 Simulation Results

Example 2. Consider the following quadratic programming problem:

$$\begin{array}{ll} \text{minimize} & f(x) = -x_1^2 + 0.5x_2^2 + x_3^2 + x_4^2 - x_1x_2 + 2x_1x_3 \\ & \qquad +x_1x_4 + 1.5x_2x_4 + 2x_1 - x_2 + x_3 - x_4, \\ \text{subject to} & x_1 + x_2 - x_3 = 2, \\ & x_1 - x_2 - x_4 = -2, \\ & -2 \leq x_1, x_2, x_3, x_4 \leq 2. \end{array} \tag{25}$$

As

$$Q = \begin{pmatrix} -2 & -1 & 2 & 1 \\ -1 & 1 & 0 & 1.5 \\ 2 & 0 & 2 & 0 \\ 1 & 1.5 & 0 & 2 \end{pmatrix}$$

is not positive definite, the objective function is not convex everywhere. However, if we substitute $x_3 = x_1 + x_2 - 2$ and $x_4 = x_1 - x_2 + 2$ into the objective function,

then we have $\tilde{f}(x_1, x_2) = 4x_1^2 + x_2^2 + 1.5x_1x_2 + 2x_1 - 5x_2 + 2$ is strictly convex. Here, $\lambda_{\max}(Q^2) = 11.2315$ and $\lambda_{\max}(Q^{-1}) = 3.0330$. According to Corollary 6, let's choose $\alpha = 9 > \lambda_{\max}(Q^2)\lambda_{\max}(Q^{-1})/4 = 8.5163$ and $\epsilon = 10^{-5}$ in the neural network model. Fig. 2 depicts the transient behaviors of output vector $x(t)$ with 5 random initial points, from which we can see that the output variables are globally convergent to the optimal solution $x^* = (-0.375, 2, -0.375, -0.375)^T$.

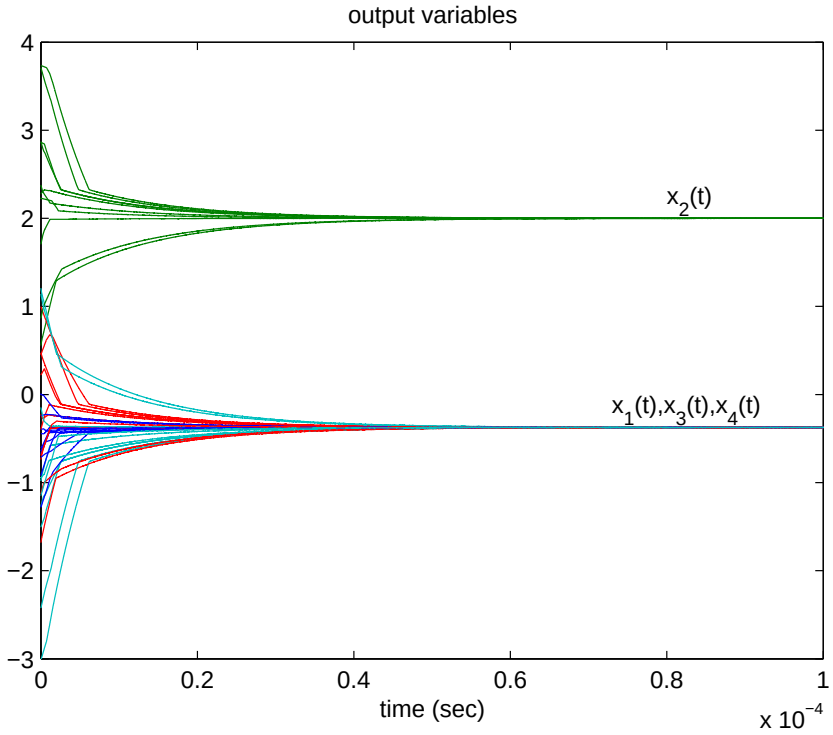


Figure 2: Transient behaviors of the output variables of neural network in (22) and (23) with $\alpha = 9$ in Example 2.

5. APPLICATION TO k-WINNERS-TAKE-ALL

Generally, the k-WTA operation can be defined as the following function

$$x_i = f(v_i) = \begin{cases} 1, & \text{if } v_i \in \{k \text{ largest elements of } v\}, \\ 0, & \text{otherwise}, \end{cases} \tag{26}$$

where $v = (v_1, v_2, \ldots, v_n)^T$ is the input vector and $x = (x_1, x_2, \ldots, x_n)^T$ is the output vector.

The solution to the k-WTA problem can be determined by solving the following

linear integer programming problem:

$$\begin{aligned} \text{minimize} \quad & -\sum_{i=1}^{n} v_i x_i, \\ \text{subject to} \quad & \sum_{i=1}^{n} x_i = k, \\ & x_i \in \{0,1\}, \quad i = 1,2,\ldots,n. \end{aligned} \tag{27}$$

According to [25], if the kth and $(k+1)$th largest elements of v (denoted as $\bar{v}_k$ and $\bar{v}_{k+1}$ respectively) are different, problem (27) is equivalent to the following linear programming problem:

$$\begin{aligned} \text{minimize} \quad & -v^T x, \\ \text{subject to} \quad & \sum_{i=1}^{n} x_i = k, \\ & 0 \le x_i \le 1, \quad i = 1,2,\ldots,n, \end{aligned} \tag{28}$$

or the following quadratic programming problem:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\eta x^T x - v^T x, \\ \text{subject to} \quad & \sum_{i=1}^{n} x_i = k, \\ & 0 \le x_i \le 1, \quad i = 1,2,\ldots,n, \end{aligned} \tag{29}$$

where $\eta \le \bar{v}_k - \bar{v}_{k+1}$ is a positive constant.

In the following, based on the linear and quadratic programming problems, some recurrent neural networks will be introduced for solving the k-WTA problems.

5.1 k-WTA Network (I)

According to the neural network (11) for linear programming in Section 3, the dynamic equation of the LP-based k-WTA network is described as follows [23]:

$$\epsilon \frac{dx}{dt} = -Px - \sigma(I - P)g(x) + q, \tag{30}$$

where $P = ee^T/n$, $q = v - Pv + ke/n$, $e = (1,1,\ldots,1)^T \in \mathbb{R}^n$, ϵ is a positive constant, σ is a nonnegative gain parameter, and $g(x) = (g(x_1), g(x_2), \ldots, g(x_n))^T$ is the activation function defined in (12) with $l_i = 0$ and $u_i = 1$ $(i = 1,2,\ldots,n)$.

The following two corollaries are directly from the results in Section 3.

Corollary 8. *The network* (30) *can perform the k-WTA operation if the equilibrium point set* $\Omega^e \subset \{x \in \mathbb{R}^n : 0 \le x \le 1\}$.

Corollary 9. *The network* (30) *can perform the k-WTA operation if it has a unique equilibrium point and* $\sigma \ge 0$ *when* $(I - ee^T/n)u = 0$ *or one of the following conditions holds when* $(I - ee^T/n)u \ne 0$*:*

(i) $\sigma \geq n\sqrt{\frac{\sum_{i=1}^{n}(v_i-\sum_{j=1}^{n}v_j/n)^2}{n(n-1)}}$, *or*

(ii) $\sigma \geq \frac{\sqrt{\sum_{i=1}^{n}(v_i-\sum_{j=1}^{n}v_j/n)^2}}{\min^{+}_{\xi_i\in\{-1,0,1\}}\left\{|\sum_{i=1}^{n}(v_i-\sum_{j=1}^{n}v_j/n)\xi_i|\right\}}$.

5.2 k-WTA Network (II)

According to the neural network in (22) and (23) for quadratic programming in Section 4, the QP-based k-WTA network is described as follows [23]:

- State equation

$$\epsilon\frac{dz}{dt} = -(I-P)z - [\eta I + (1-\eta)P]h(z) + q, \tag{31}$$

- Output equation

$$x = -\frac{1}{\eta}(I-P)z + \frac{q}{\eta} + \frac{k(\eta-1)}{n\eta}e, \tag{32}$$

where P, q and ϵ are defined in (30), $h(z) = (h(z_1), h(z_2), \ldots, h(z_n))^T$ is the activation function defined in (24) with $l_i = 0$ and $u_i = 1$ $(i = 1, 2, \ldots, n)$.

From the results in Section 4, we get the following two corollaries.

Corollary 10. *The system* (31) *with any* $\eta > 0$ *is stable in the sense of Lyapunov and any trajectory is globally convergent to an equilibrium point.*

Corollary 11. $x^* = -(I-P)z^*/\eta + q/\eta + (\eta-1)ke/(n\eta)$ *is an optimal solution of* k*-WTA problem* (29)*, where* z^* *is an equilibrium point of system* (31).

From the results of Corollaries 10 and 11, the network in (31) and (32) is globally convergent and the output vector is globally convergent to an optimal solution of problem (29). Thus, if the kth and $(k+1)$th largest elements of the input signals $v_i (i = 1, 2, \ldots, n)$ are different, the network in (31) and (32) is capable of guaranteeing the k-WTA operation.

5.3 k-WTA Network (III)

In general, the networks in above two subsections are capable of solving the k-WTA problem if the kth and $(k+1)$th largest elements of v take different values. For this class of k-WTA networks, some other models can be found in [14, 18, 25]. Among them, a much simpler k-WTA network with a single state variable only was developed based on the improved dual neural network [14], and it's discrete-time counterpart was presented in [18]. However, these networks are not capable of handling the ill cases, which have same input values. Here we introduce a k-WTA network with a single state variable only for solving the general problems including the ill cases [20, 34].

According to the Karush-Kuhn-Tucker (KKT) conditions for linear programming [2], x^* is an optimal solution of (28), if and only if there exist $y^* \in \mathbb{R}$ and $z^* \in \mathbb{R}^n$ such that (x^*, y^*, z^*) satisfies the following optimality conditions:

$$-v + ey + z = 0, \tag{33}$$

$$e^T x = k, \tag{34}$$

$$\begin{cases} x_i = 1, & z_i > 0, \\ 0 \le x_i \le 1, & z_i = 0, \\ x_i = 0, & z_i < 0, \end{cases} \tag{35}$$

where $e = (1, 1, \ldots, 1)^T \in \mathbb{R}^n$.

From (35), we define a discontinuous function $x = g(z) = (g(z_1), g(z_2), \ldots, g(z_n))^T$ with its components as follows

$$x_i = g(z_i) = \begin{cases} 1, & z_i > 0, \\ [0, 1], & z_i = 0, \\ 0, & z_i < 0. \end{cases} \tag{36}$$

The function $g(z_i)$ is monotone in $\mathbb{R}$ and discontinuous at $z_i = 0$.

According to (33) and (36), we have

$$x = g(z) = g(v - ey). \tag{37}$$

Substituting (37) into (34), we have

$$e^T g(v - ey) = k. \tag{38}$$

Based on (38), the one-neuron recurrent neural network is described as follows:

$$\epsilon \frac{dy}{dt} = e^T g(v - ey) - k, \tag{39}$$

where ϵ is a positive constant.

The global convergence of neural network (39) can be guaranteed in finite time as the following theorem.

Theorem 5. *[20] For any initial point $y(0) \in \mathbb{R}$, the state variable $y(t)$ of* (39) *is globally convergent to an equilibrium point y^* in finite time.*

Furthermore, the equilibrium point of neural network (39) is determined by the kth and $(k+1)$th largest elements of input v. We state it as the following theorem.

Theorem 6. *[20] Assume that the kth and $(k+1)$th largest elements of input v are $v_{\sigma(k)}$ and $v_{\sigma(k+1)}$ respectively. Then the equilibrium point set of neural network* (39) *is the interval $[v_{\sigma(k+1)}, u_{\sigma(k)}]$ if $v_{\sigma(k)} \neq v_{\sigma(k+1)}$. In particular, if $v_{\sigma(k)} = v_{\sigma(k+1)}$, there is a unique equilibrium point $y^* = v_{\sigma(k)}$ for neural network* (39).

Combining the results in Theorems 5 and 6, the following theorem is easily to be derived.

Theorem 7. *[20] If the initial value satisfies* $y(0) > v_{\sigma(k)}$*, the neural network* (39) *is globally convergent to the* k*th largest element* $v_{\sigma(k)}$ *of input* v *in finite time. If the initial value satisfies* $y(0) < v_{\sigma(k+1)}$*, the neural network* (39) *is globally convergent to the* $(k+1)$*th largest element* $v_{\sigma(k+1)}$ *of input* v *in finite time. If the initial value satisfies* $v_{\sigma(k+1)} \leq y(0) \leq v_{\sigma(k)}$*, the state variable is steady at an equilibrium point. In particular, if* $v_{\sigma(k)} = v_{\sigma(k+1)}$*, the neural network* (39) *is globally convergent to* $v_{\sigma(k)}$ *(or* $v_{\sigma(k+1)}$*) in finite time.*

To design a k-WTA network to get the k-selection of input v, the neural network in (39) plays a key role. The block diagram of the k-WTA network is shown in Fig. 3. It shows that the k-WTA network is composed of two parts. The signal v is the input to the network. The output of Part A is simultaneously the input of Part B. The output x get the k-selection of the input v.

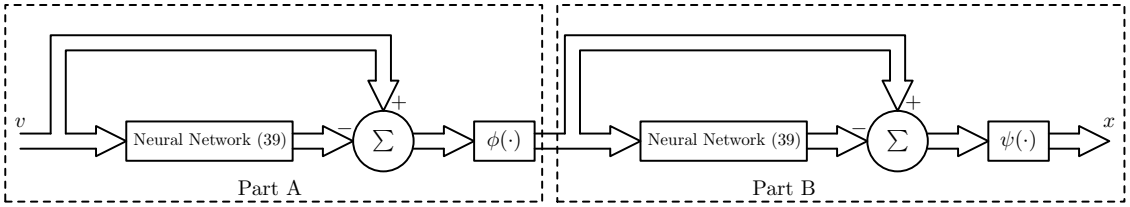


Figure 3: Block diagram of the k-WTA network.

For the k-WTA network, the output function in Parts A is defined as $\phi(z) = (\phi(z_1), \phi(z_2), \ldots, \phi(z_n))^T$, where $\phi(z_i)$ is as follows

$$\phi(z_i) = \begin{cases} \Delta, & z_i > 0, \\ \delta_i, & z_i = 0, \\ 0, & z_i < 0, \end{cases} \tag{40}$$

where Δ is a large positive constant, and δ_i is defined to separate the kth and $(k+1)$th largest elements with $0 < \delta_i < \Delta$. There are many methods to define δ_i. For example, $\delta_i = i\Delta/(n+1)$, $\delta_i = \Delta - i\Delta/(n+1)$, or δ_i takes a random value in $(0, \Delta)$. We note that δ_i takes different valves in $(0, \Delta)$ if $v_{\sigma(k)} = v_{\sigma(k+1)}$.

Combining the neural network (39) and the output function (40), the system equations of Part A in the k-WTA network is described as follows:

• State equation:

$$\epsilon \frac{dy}{dt} = e^T g(v - ey) - k. \tag{41a}$$

• Output equation:

$$w = \phi(v - ey). \tag{41b}$$

For any initial point $y(0) \in \mathbb{R}$, assume that y^* is an equilibrium point of system (41a) and w^* is the output vector corresponding to y^*; i.e., $w^* = \phi(v - ey^*)$. Let $w^*_{\sigma(k)}$ and $w^*_{\sigma(k+1)}$ be the kth and $(k+1)$th largest elements of w^* respectively. In the processing of the k-WTA network in Part A, the output vector w^* gives a k-selection for any input vector v, and the kth and $(k+1)$th largest elements of w^* are different.

After getting a k-selection of input vector v, next in Part B, the k-selection is determined if the output of the k-WTA network takes 1. Otherwise, the output takes 0. For Part B in the k-WTA network, the output function is defined as $\psi(z) = (\psi(z_1), \psi(z_2), \ldots, \psi(z_n))^T$, where $\psi(z_i)$ is as follows

$$\psi(z_i) = \begin{cases} 1, & z_i \geq 0, \\ 0, & z_i < 0. \end{cases} \tag{42}$$

Combining the neural network (39) and the output function (42), the system equations of Part B in the k-WTA network is described as follows:

• State equation:

$$\epsilon \frac{dy}{dt} = e^T g(w - ey) - k. \tag{43a}$$

• Output equation:

$$x = \psi(w - ey). \tag{43b}$$

According to Theorem 7, if $y(0) \geq \Delta$, the neural network in (43a) is globally convergent to the kth largest element of w in finite time. Since the kth and $(k+1)$th largest elements of w are different by the k-selection in Part A, the output vector x determines the k largest elements of w using the output function defined in (42). Furthermore, since the k largest elements of w correspond to a k-selection of input vector v, thus the output vector x determines the k largest elements of v. That is for any initial point $y(0) \geq \Delta$, the output vector x gets the solution of the k-WTA problem.

Similar to above analysis, if the output function in (42) is defined as

$$\psi'(z_i) = \begin{cases} 1, & z_i > 0, \\ 0, & z_i \leq 0, \end{cases} \tag{44}$$

for any initial point $y(0) \leq 0$, the output vector x also determines the k largest elements of input vector v.

5.4 Simulation Results

Example 3. We select 801 points evenly from $[-2, 2]$ and the inputs are generated from a simple function in $\mathbb{R}$ as follows:

$$v = f(x) = 0.5 \times \frac{\sin(20x/\pi)}{20x/\pi} + \rho,$$

where ρ is an additive Gaussian noise with zero mean and standard deviation $\sigma = 0.6$.

Let $\Delta = 1$ and δ_i in (40) takes random values, the simulation results are shown in Fig. 4. The line in Fig. 4 takes the value of the kth largest element of input v. The points above the line are k winners and the others under the line are $(n - k)$ losers.

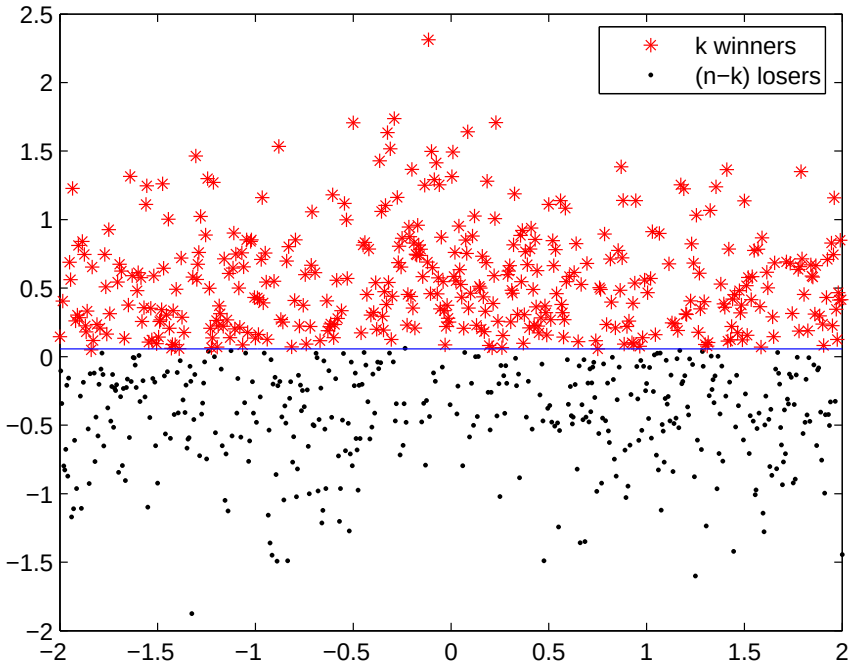


Figure 4: Simulation results of the $k(= 400)$ largest elements of the input signals in Example 3. The points on the line take the value of the kth largest element of input v. The points above the line (including some on the line) are k winners and the others under the line (including some on the line) are $(n - k)$ losers.

Example 4. Let $k = 2$ and consider a set of five input signals with the following instantaneous values $v_i(s) = 10 \sin[2\pi(s + 0.2(i - 1))]$ $(i = 1, 2, 3, 4)$, $v_5(s) = 0$ if $v_1(s) \geq 0$ and $v_5(s) = v_1(s)$ otherwise, where $s \in [0, 1]$. In the five input signals, when s is within approximately $[0.5, 0.55]$, any one of v_1 and v_5 can be considered as the second largest element since they take the same values. We give the simulations of the k-WTA network with $\Delta = 1$ and different form of δ_i in (40).

Fig. 5 depicts the simulation results with $\delta_i = i\Delta/6$ $(i = 1, 2, 3, 4, 5)$. It shows that the k-WTA network selects v_5 as the second largest element when inputs $v_1 = v_5$ as $s \in [0.5, 0.55]$. If δ_i in (40) takes random values, the simulation results are shown in Fig. 6. We can see that the k-WTA network selects randomly v_1 or v_5 as the the second largest element when $s \in [0.5, 0.55]$. The simulation results show that the k-WTA network can generate the two largest elements in real time.

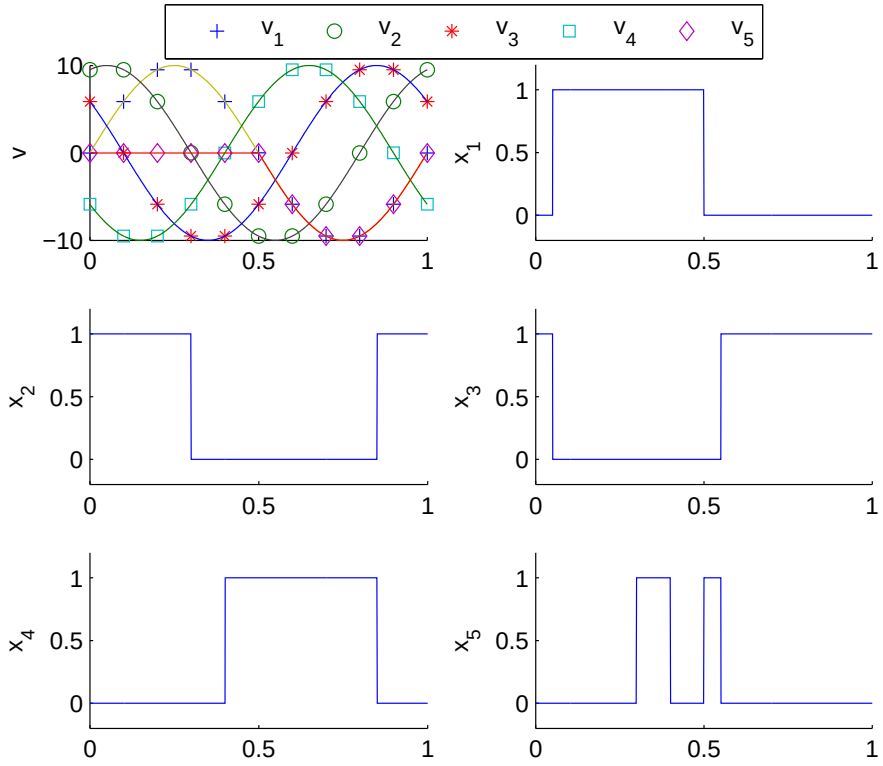


Figure 5: Inputs and outputs of the k-WTA network as $\delta_i = i\Delta/6$ in Example 4. It shows that the k-WTA network selects v_5 as the second largest element when inputs $v_1 = v_5$ as $s \in [0.5, 0.55]$.

6. Concluding Remarks

Recurrent neural networks are computational models for solving optimization problems. Ever since the seminal work of Hopfield and Tank on neurodynamic optimization, numerous recurrent neural networks have been developed for optimization. The recurrent neural networks have been improved with desirable properties such as exact optimality, global convergence, and simple architecture.

This chapter is devoted to recurrent neural network approaches to linear and quadratic programming problems with the k-winners-take-all application. Several recurrent neural networks with discontinuous activation functions are introduced

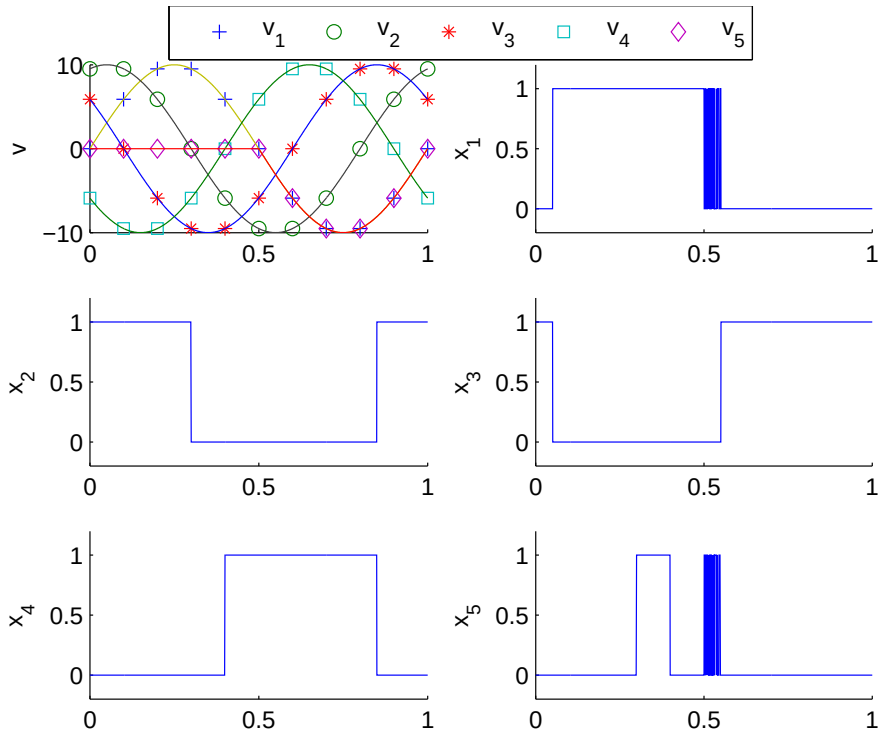


Figure 6: Inputs and outputs of the k-WTA networks as δ_i taking random values in Example 4. It shows that the k-WTA network selects randomly v_1 or v_5 as the the second largest element when inputs $v_1 = v_5$ as $s \in [0.5, 0.55]$.

with detailed model description, design procedures, convergence results and complexity analysis. For linear and convex quadratic programming problems, two one-layer recurrent neural networks have been introduced. Moreover, some k-WTA networks are presented for solving the k-winners-take-all problems.

7. REFERENCES

[1] A. Andreou, K. Boahen, P. Pouliquen, A. Pavasovic, R. Jenkins, and K. Strohbehn, "Current-mode subthreshold mos circuits for analog vlsi neural systems," IEEE Transactions on Neural Networks, vol. 2, no. 2, pp. 205–213, Mar. 1991.

[2] M. Bazaraa, H. Sherali, and C. Shetty, Nonlinear Programming: Theory and Algorithms (3rd Ed.). Hoboken, New Jersey: John Wiley & Sons, 2006.

[3] D. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods. New York: Academic, 1982.

[4] L. Binh and H. Chong, "A neural-network contention controller for packet switching networks," IEEE transactions on neural networks, vol. 6, no. 6, pp. 1402–1410, Nov. 1995.

[5] A. Bouzerdoum and T. Pattison, "Neural network for quadratic optimization with bound constraints," IEEE Transactions on Neural Networks, vol. 4, no. 2, pp. 293–304, Mar. 1993.

[6] Y. Chen and S. Fang, "Neurocomputing with time delay analysis for solving convex quadratic programming problems," IEEE Transactions on Neural Networks, vol. 11, no. 1, pp. 230–240, Jan. 2000.

[7] F. Clarke, Optimization and Nonsmooth Analysis. Wiley, New York, 1983.

[8] A. Cochocki and R. Unbehauen, Neural Networks for Optimization and Signal Processing. John Wiley & Sons, Inc. New York, NY, USA, 1993.

[9] A. Filippov, Differential Equations with Discontinuous Righthand Sides. Mathematics and its applications (Soviet series). Boston: Kluwer Academic Publishers, 1988.

[10] M. Forti, P. Nistri, and M. Quincampoix, "Generalized neural network for nonsmooth nonlinear programming problems," IEEE Transactions on Circuits and Systems-I, vol. 51, no. 9, pp. 1741–1754, Sept. 2004.

[11] ——, "Convergence of neural networks for programming problems via a nonsmooth Łojasiewicz inequality," IEEE Transactions on Neural Networks, vol. 17, no. 6, pp. 1471–1486, Nov. 2006.

[12] T. Friesz, D. Bernstein, N. Mehta, R. Tobin, and S. Ganjalizadeh, "Day-to-day dynamic network disequilibria and idealized traveler information systems," Operations Research, vol. 42, no. 6, pp. 1120–1136, 1994.

[13] J. Hertz, A. Krogh, and R. Palmer, Introduction to the Theory of Neural Computing. Massachusetts: Addison Wesley, 1991.

[14] X. Hu and J. Wang, "An improved dual neural network for solving a class of quadratic programming problems and its k-winners-take-all application," IEEE Transactions on Neural Networks, vol. 19, no. 12, pp. 2022–2031, Dec. 2008.

[15] Jayadeva and S. Rahman, "A neural network with o(n) neurons for ranking n numbers in o(1/n) time," IEEE Transactions on Circuits and Systems-I, vol. 51, no. 10, pp. 2044–2051, Oct. 2004.

[16] M. Kennedy and L. Chua, "Neural networks for nonlinear programming," IEEE Transactions on Circuits and Systems, vol. 35, no. 5, pp. 554–562, May 1988.

[17] M. Lemmon and B. Kumar, "Competitive learning with generalized winner-take-all activation," IEEE Transactions on Neural Networks, vol. 3, no. 2, pp. 167–175, Mar. 1992.

[18] Q. Liu, J. Cao, and J. Liang, "A discrete-time recurrent neural network with one neuron for k-winners-take-all operation," in Proc. 6rd Int. Symposium on Neural Networks (ISNN2009), Wuhan, China, June 2009, pp. 272–278.

[19] Q. Liu, J. Cao, and Y. Xia, "A delayed neural network for solving linear projection equations and its analysis," IEEE Transactions on Neural Networks, vol. 16, no. 4, pp. 834–843, July 2005.

[20] Q. Liu, C. Dang, and J. Cao, "A novel recurrent neural network with one neuron and finite-time convergence for k-winners-take-all operation," IEEE Transactions on Neural Networks, vol. 21, no. 7, pp. 1140–1148, July 2010.

[21] Q. Liu and J. Wang, "A one-layer recurrent neural network with a discontinuous activation function for linear programming," Neural Computation, vol. 20, no. 5, pp. 1366–1383, 2008.

[22] ——, "A one-layer recurrent neural network with a discontinuous hard-limiting activation function for quadratic programming," IEEE Transactions on Neural Networks, vol. 19, no. 4, pp. 558–570, Apr. 2008.

[23] ——, "Two k-winners-take-all networks with discontinuous activation functions," Neural Networks, vol. 21, no. 2-3, pp. 406–413, 2008.

[24] Q. Liu, J. Wang, and J. Cao, "A delayed lagrangian network for solving quadratic programming problems with equality constraints," in Lecture Notes in Computer Science (ISNN2006), vol. 3971. Springer, 2006, pp. 369–378.

[25] S. Liu and J. Wang, "A simplified dual neural network for quadratic programming with its kwta application," IEEE Transactions on Neural Networks, vol. 17, no. 6, pp. 1500–1510, Nov. 2006.

[26] C. Maa and M. Shanblatt, "Linear and quadratic programming neural network analysis," IEEE Transactions on Neural Networks, vol. 3, no. 4, pp. 580–594, July 1992.

[27] W. Maass, "On the computational power of winner-take-all," Neural Computation, vol. 12, no. 11, pp. 2519–2535, 2000.

[28] C. Marinov and J. Hopfield, "Stable computational dynamics for a class of circuits with o(n) interconnections capable of kwta and rank extractions,"

IEEE Transactions on Circuits and Systems-I, vol. 52, no. 5, pp. 949–959, May 2005.

[29] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," Science, vol. 194, no. 4262, pp. 283–287, 1976.

[30] D. Tank and J. Hopfield, "Simple neural optimization networks: An a/d converter, signal decision circuit, and a linear programming circuit," IEEE Transactions on Circuits and Systems, vol. 33, no. 5, pp. 533–541, May 1986.

[31] K. Urahama and T. Nagao, "K-winners-take-all circuit with o(n) complexity," IEEE Transactions on Neural Networks, vol. 6, no. 3, pp. 776–778, May 1995.

[32] J. Wang, "Analysis and design of a recurrent neural network for linear programming," IEEE Transactions on Circuits and Systems-I, vol. 40, no. 9, pp. 613–618, Sept. 1993.

[33] ——, "A deterministic annealing neural network for convex programming," Neural Networks, vol. 7, no. 4, pp. 629–641, 1994.

[34] ——, "Analysis and design of a k-winners-take-all model with a single state variable and the heaviside step activation function," IEEE Transactions on Neural Networks, vol. 21, no. 9, pp. 1496–1506, Sept. 2010.

[35] W. Wolfe, D. Mathis, C. Anderson, J. Rothman, M. Gottler, G. Brady, R. Walker, G. Duane, and G. Alaghband, "K-winner networks," IEEE Transactions on Neural Networks, vol. 2, no. 2, pp. 310–315, Mar. 1991.

[36] Y. Xia, "A new neural network for solving linear and quadratic programming problems," IEEE Transactions on Neural Networks, vol. 7, no. 6, pp. 1544–1548, Nov. 1996.

[37] Y. Xia, H. Leung, and J. Wang, "A projection neural network and its application to constrained optimization problems," IEEE Transactions Circuits and Systems-I, vol. 49, no. 4, pp. 447–458, Apr. 2002.

[38] Y. Xia and J. Wang, "Neural network for solving linear programming problems with bounded variables," IEEE Transactions on Neural Networks, vol. 6, no. 2, pp. 515–519, 1995.

[39] ——, "Primal neural networks for solving convex quadratic programs," in International Joint Conference on Neural Networks (IJCNN 1999), vol. 1, 1999, pp. 582–587.

[40] ——, "A dual neural network for kinematic control of redundant robot manipulators," IEEE Transactions on Systems, Man and Cybernetics-B, vol. 31, no. 1, pp. 147–154, 2001.

[41] ——, "A general projection neural network for solving monotone variational inequalities and related optimization problems," IEEE Transactions on Neural Networks, vol. 15, no. 2, pp. 318–328, Mar. 2004.

[42] ——, "A recurrent neural network for nonlinear convex optimization subject to nonlinear inequality constraints," IEEE Transactions on Circuits and Systems-I, vol. 51, no. 7, pp. 1385–1394, July 2004.

[43] Y. Xia, J. Wang, and L. Fok, "Grasping-force optimization for multifingered robotic hands using a recurrent neural network," IEEE Transactions on Robotics and Automation, vol. 20, no. 3, pp. 549–554, June 2004.

[44] Y. Yang and J. Cao, "Solving quadratic programming problems by delayed projection neural network," IEEE Transactions on Neural Networks, vol. 17, no. 6, pp. 1630–1634, Nov. 2006.

[45] J. Yen, J. Guo, and H. Chen, "A new k-winners-take-all neural network and its array architecture," IEEE Transactions on Neural Networks, vol. 9, no. 5, pp. 901–912, Sept. 1998.

[46] A. Yuille and D. Geiger, "Winner-take-all networks," in The Handbook of Brain Theory and Neural Networks (2nd ed.). MIT Press Cambridge, MA, 2003, pp. 1228–1231.

[47] S. Zhang and A. Constantinides, "Lagrange programming neural networks," IEEE Transactions on Circuits and Systems-II, vol. 39, no. 7, pp. 441–452, July 1992.

APPLICATIONS OF ARTIFICIAL NEURAL NETWORKS IN MOBILE ROBOTS NAVIGATION

V. Tiponut*, I. Bogdanov*, T. Slavici**

*Applied Electronics Department
**Mechanical Machines and Transportations
POLITEHNICA University of Timisoara
Bvd. Vasile Parvan 2, 300223 Timisoara
ROMANIA
virgil.tiponut@etc.upt.ro, ivan.bogdanov@etc.upt.ro, titus.slavici@mec.upt.ro

Abstract: *In the present paper, it will be reported original results concerning the application of Artificial Neural Networks (ANN) and in particular Cellular Neural Networks (CNN) in mobile robot control. In the first part of the paper, a complete solution, using ANN, for obstacles detection and robot navigation in unstructured environment is presented. The rest of the paper is entirely devoted to applications implemented using CNN. Different methods for path finding to the target and trajectory planning, for a single robot and mobile robot collectivities, are analyzed and the obtained results are reported.*

Keywords*:* - artificial neural networks, mobile robot, robot collectivities, path finding, trajectory planning.

1. INTRODUCTION

ANN and CNN are valuable tools, widely used today in autonomous mobile robot navigation. Obstacles detection in unknown working environment, path finding and trajectory planning, mobile robot collectivities cooperation, are the main applications where ANN and CNN have successfully proved their huge potential.

The artificial neural network paradigm is very attractive for applications in mobile robotics because of some favorable characteristics:

- High speed of data processing, due to the ANN's massive parallelism,
- ANN are capable to handle noisy and incomplete data, specific to the sensorial system of mobile robots,
- There is no necessary need for a precise mathematically model of the process, in order to implement an ANN based application and,
- ANN are capable of generalization, a feature specific only to the human and living beings behavior.

In the most cases, the multilayer perceptron – trained by using the well-known error back-propagation algorithm, are preferred; in some cases, a single layer perceptron can solve the problem. Other types of ANN (Kohonen, ART1, ART2, Winner-take-all, Genetic Algorithm, etc.) are, also, taken into account, depending of the nature of the application.

The success of an ANN based application strongly depends on decisions taken in the design phase: type, size and number of layers of the ANN that best fits to that ap-

plication, the preprocessing procedure of the input data and the quality of the learning process. However, in some particular cases, the ANN solutions are the only and, therefore, the best solution.

In the last years, the interest of researchers has focused on CNN, a special type of ANN. Unlike the other type of ANN, CNN have a limited local connectivity, and therefore can be integrated in VLSI technology. These hardware implementations promise the highest processing speed available today.

The basic CNN has a two-dimensional rectangular structure, composed of identical, nonlinear analog circuits (called cells). This arrangement suggests their capability to be used in image and other two-dimensional signals processing. For this purpose, appropriate operators (Templates) have to be associated to the basic CNN structure, according to the operation that should be performed. A complex application may need a chain of processing operations, which are performed serially, assigning appropriate templates to the same basic CNN.

There is a huge library of templates, available to the whole scientific community, whose content is continuously updated. The difficult task in an implementation of a CNN application is to choose the most appropriate series of templates, corresponding to the processing operations to be performed.

The rest of the paper is organized as follows. In section 2, the implementation of a mobile robot entirely controlled by ANN is presented. The next two sections are devoted to applications of CNN in mobile robotics: single robot navigation (section 3) and mobile robot collectivities (section 4). Some conclusions and suggestions for further research are presented in the final section.

2. AUTONOMOUS MOBILE ROBOT CONTROLLED BY ANN

The mobile robot that will be presented here, implemented entirely by ANN, has been developed in our early works [1], [2], [3], [4]. The only mission of the robot is to wonder in the working space, with obstacles avoidance. For this purpose, an ultrasonic sensorial system has been included in the robot's structure. Unlike the commonly used in robotics ultrasonic systems, which are using pulse-echo method for obstacles detection and can include up to 32 sensors placed around the robot, in the proposed system only three transducers are used, bioinspired from the echo locator of the bats.

The ultrasonic transducers are disposed as is shown in Fig. 1, imitating the mouth (transducer placed in the bottom position) and the ears of a bat (the two upper transducers).

The investigated field in front of the subject, corresponding to this arrangement, is depicted in Fig. 2.

The presence of an obstacle can be detected in the area represented in grey, which is the common place of the individual fields investigated by each reception sensor. A single transmitted pulse and the resulting echoes are needed in order to locate all the obstacles in front of the subject. In this way, the presented method for obstacles de-

tection is in the same time simple (only three transducers are necessary for implementation) and fast.

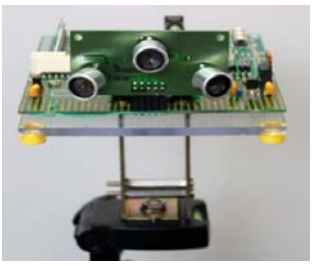


Fig. 1. The placement of the transducers in the bioinspired ultrasonic system

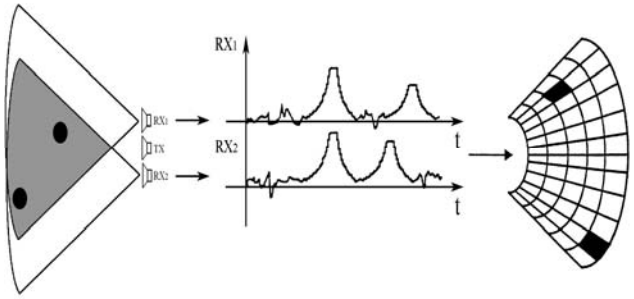


Fig. 2. The principle of obstacles detection using bioinspired ultrasonic system

In Fig. 2, the echoes corresponding to the presence of two obstacles in the supervised area are represented.

The amplitude of these pulses is a measure of the distance from the subject to the obstacles, while the difference in time of the moment of occurrences of these pulses give the azimuth, relative to the median axis of the sensorial system.

The acquired data are processed entirely using ANN, connected as in Fig. 3. ANN1 extract the relevant information for the mobile robot movement: the distance and the azimuth of each obstacle in front of the robot and is capable, in this way, to indicate the presence of obstacles in the investigated area.

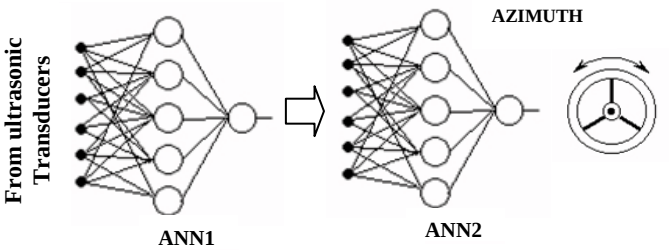


Fig. 3. The ANN used to process sensorial data

Please also notice that ANN1 has a limited resolution, i.e. is capable to indicate the presence or absence of obstacles in a limited number of small discrete areas in front of the robot. These discrete areas result by dividing the supervised area using an appropriate grid, as is illustrated in Fig. 2.

Actually, the output levels of ANN1 network reflects the probability of existing obstacles in each of the discrete area mentioned above. This information has to be processed again, in order to obtain the value of the azimuth of the mobile robot trajectory which avoidances obstacles.

The processing unit for this task includes two neural networks, represented here by ANN2. The first ANN acts as a classification engine, and is capable to recognize different scenarios, with different positions of obstacles relative to the position of mobile robot. Based upon this information, the second ANN issues to its output the signed value of the azimuth, that allow the movement of the mobile robot with obstacles avoidance.

All the above mentioned ANN have been developed using simple perceptron type networks, and can easily software implemented on a microcontroller system.

3. CNN BASED SINGLE ROBOT NAVIGATION

The first our results on application of CNN in mobile robotics have been reported in [5]. Many other papers in this area were published after that, some of them being focused on a single robot navigation using CNN methods [6],[7],[8],[9],[10] and [11].

In this section it will presented, as an example, a CNN algorithm for guidance the movement of a mobile robot in a labyrinth. For this purpose, we will consider a working space having the structure represented in Fig. 4. A video camera, connected to a general purpose computer, acquires periodically the image of the workspace. This information is processed then on the personal computer, using CNN based algorithms and appropriate signals are generated, in order to control the movement of the robot.

Actually, the mission of the mobile robot is to move in a labyrinth, with continuously walls, and to reach a target totally unknown for the robot.

There is a simple solution to this problem: the robot has to move in the labyrinth keeping a small, fixed distance to the one wall of the corridor, until the target will be reached (we will assume in the following that the target is placed beside the wall). Please notice that there are two possible trajectories to the target, as is depicted in Fig. 4: having the wall on the left or on the right in the sense of movement.

Anyway, the wall following is a simple procedure to reach the target. All the robot has to do is to “keep contact” with one wall of the labyrinth. In mobile robotics, this type of activity is called local path planning, and usually it is performed by processing the information acquired by the sensorial system of the robot.

In the present application, local path planning is performed by processing the images of the whole workspace, acquired periodically before to each next step of the robot. These images include the actual position of the robot, the target and the obstacles (the walls).

Each image is converted first in a binary form, having values corresponding to the CNN domain: +1 (black) representing obstacles (the wall) and -1 (white) representing the space free of obstacles. In the same image both, mobile robot and target are represented by only one pixel (the central point of each of them). It should be point out that

all these preprocessing operation are performed using simple CNN algorithms [5].

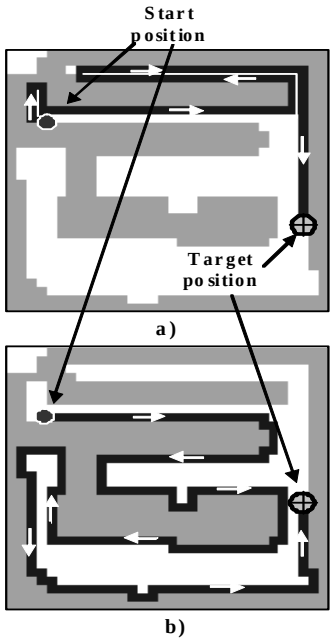


Fig. 4. The structure of the labyrinth and the possible trajectories of the robot's movement: having the wall a)on the left and b) on the right

The most important task in this application is, however, the determination of the direction of the movement before each next step. There are eight possible directions of movement, as is shown in Fig. 5.

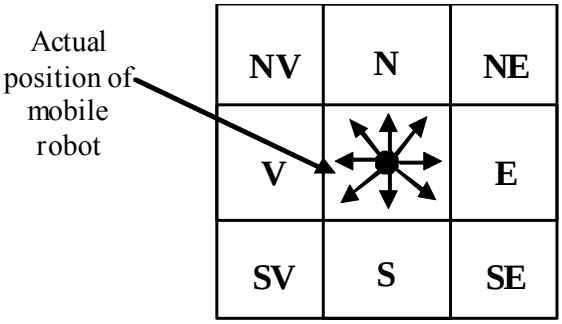


Fig. 5. The possible moving directions of the mobile robot

The direction of movement is chosen in such a way that the next position of the robot has to be free of obstacles and in the wall vicinity (the right or left wall, as is shown in Fig. 4). According to this simple rules, the determination of the direction for the next step take place as is depicted in the flowchart given in Fig. 6.

In the above algorithm we considered the robot is moving having the wall on the right; a similar flowchart can be developed for the movement of robot having the wall on his left [7].

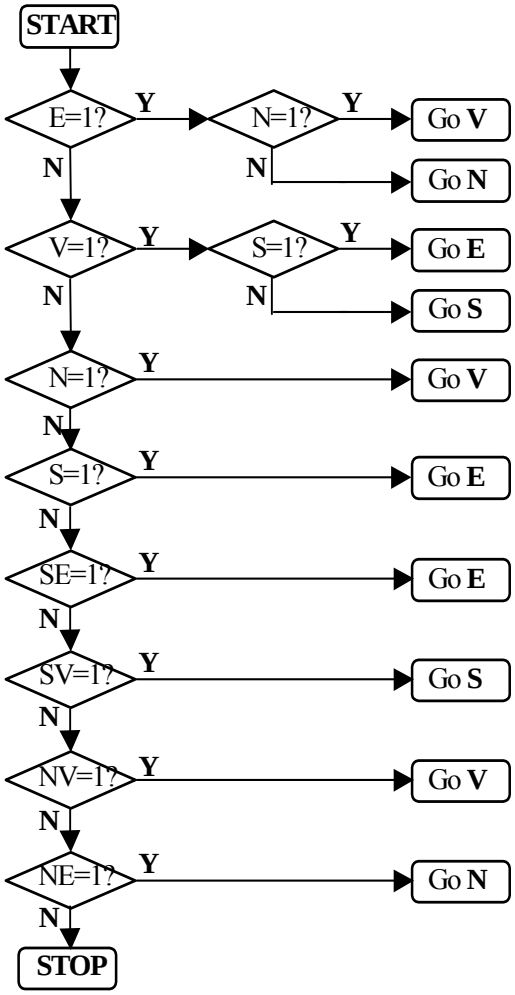


Fig. 6. Algorithm for the determination of the movement direction

The proposed algorithm have been implemented and tested in Extended Analogic Macro Code (AMC) [5] using the template shift.tem, particularized for the four main directions. Fig. 7 shows the movement of the robot along the right wall, according to the flowchart presented in Fig. 6.

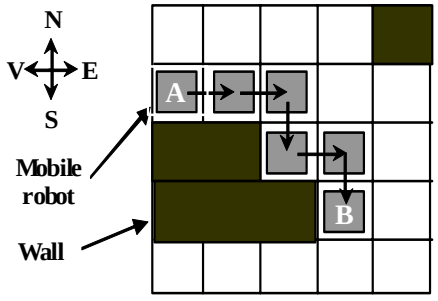


Fig. 7. The robot's movement along the right wall

4. APPLICATIONS OF CNN TO MOBILE ROBOT COLLECTIVITIES

The mobile robot collectivities have been intensively studied during the years, because of their huge potential in different fields of application (industry, services, etc.).

In this section it will be presented an application of CNN in controlling a collectivity that includes two mobile robots [12]. Other results of authors' research, in this field, can be finding in [13] and [14].

In the following, we will assume that the workplace is a common 2-dimensional environment with unknown, stationary, obstacles, as is illustrated in Fig. 8.

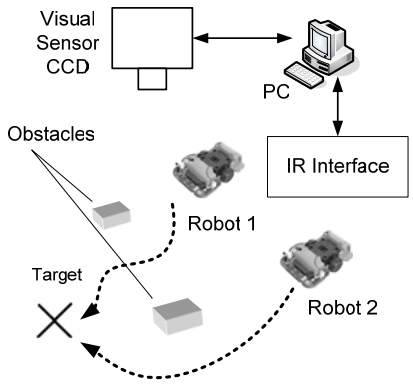


Fig. 8. Arrangement of the mobile robot collectivity

CNN are used here in order to implement the global path planning task, which include the following main activities:

– To find possible paths from the actual positions of robots to the target, free of obstacles, and

– Based upon this information to generate a valid trajectory for each mobile robot, which satisfies some constrains (usually, the shortest way to the target).

For this purpose, the image of the working space is acquired by a CCD camera, and converted into a binary image. In this image, compatible with CNN levels [5], pixels having +1 value (black) represent obstacles, while the pixels with values of -1 (white) are available for robot's movement. The binary image is processed then, using CNN procedures, in order to obtain appropriate signals to control the movement.

In order to avoid collision between robots, a simple principle is used here: during the movement of one robot the other robot is stopped, and becomes a static obstacle. This principle is implemented in the flowchart depicted in Fig. 9. All tasks included here are entirely performed by CNN algorithms, using 3*3 dimension templates [12]. For path planning, a plane wave is generated first in the image plane, with the origin of the source in the target point. In this way, it will be obtained an image with the value of both target and obstacles pixels unchanged, while the values of the rest of pixels are increased proportionally with the distance from target to that pixels (see Fig. 10). This operation can be performed by using template expore.tem [12].

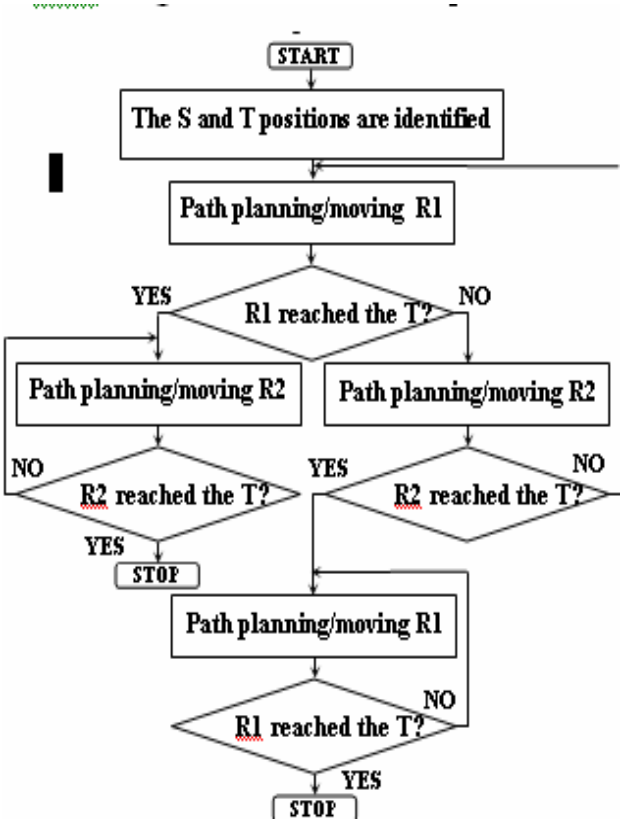


Fig. 9. Motion planning algorithm for two mobile robots (**R1, R2**), when they are moving from start (**S**) to the same target (**T**).

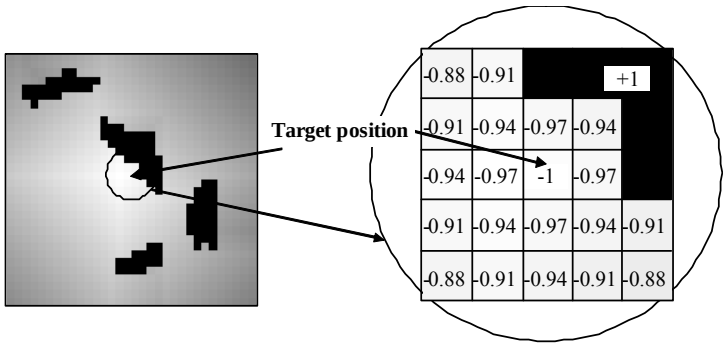


Fig. 10. The determination of the distance between the target and points within the working space

The obtained image allows now to generate a multitude of trajectory to the target, free of obstacles. For this purpose, the robot should move from a cell to another cell, in immediate its vicinity, and having the corresponding value smaller than the value of the current position (i.e. the destination cell is closer to the target). Moreover, the robot may keep the same direction as much as the above mentioned condition is satisfied.

However, if a trajectory with the minimum length is imposed, the above rules have to be changed: the destination cell should have the smallest value within all cells

included in the immediate vicinity of the current cell. For a 3*3 template there are eight possible directions to be investigated (see Fig. 5) and template select.tem can be used for this purpose [12].

The path planning must also include the collision avoidance of a robot with other robots within the same collectivity. For this purpose, it can be defined a neighboring area around each robot, forbidden for other robots. This neighborhood may be obtain by generating a wave around the pixel representing the mobile robot's position, using the template dilatation.tem [12].

The motion planning algorithm, presented above, has been tested by simulation, using "CadetWin" environment. The final results of this simulation are illustrated in Fig. 11.

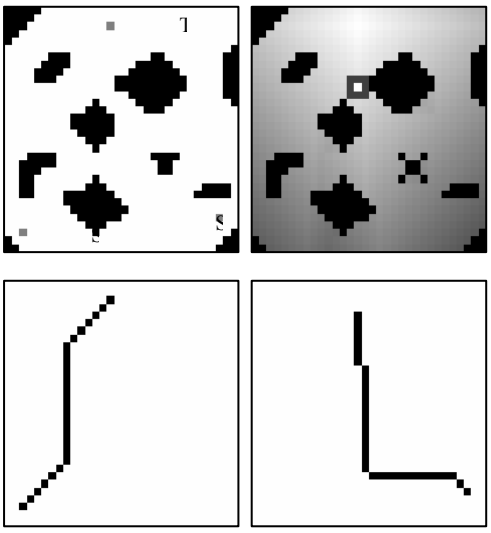


Fig. 11. Simulation results of CNN proposed algorithm: a) start and target positions after the preprocessing step; b) evaluation of distances from the target to points within the workspace; the trajectories of Robot 1c) and Robot 2 d).

5. CONCLUSION

The mobile robot implementations, presented in the previous sections, allow concluding that ANN represents viable solutions for data processing in mobile robotics. Furthermore, some special properties of ANN (the generalization, their capability to learn a process and to handle noisy and incomplete data), specific only to human and living beings, suggests the idea that ANN based implementations are, more or less, "bioinspired".

To day, the solutions developed in mobile robotics and other related fields aim to get closer to the biology, in the form of neuromorphic circuits [16]. Such circuits not only process the information using bioinspired procedures, like ANN, but in the same time their structure try to imitate the structure of the biological model that inspired them.

The transition from simple ANN based implementations to neuromorphic circuits is best illustrated having in mind the ultrasonic sensorial system presented in section

2. In our research [1], for the first time has been presented an ultrasonic sensorial system bioinspired from the bat's echolocator. The same topic was extensively approched later in [15], and finalized then in the form of a neuromorphic circuit [16].

As a final remark, we will observe that in the present paper only original results of author's team have been presented. There is a huge topic in mobile robotics and the obtained results are remarkable. In despite of this diversity, there is only one way for further developments: the bioinspired systems.

REFERENCES

[1] T. Botos, Researches on Adaptive Control of Autonomous Mobile Robots, *PhD Thesis, Politehnica University Timisoara,* Romania, 2003.

[2] T. Botos, V. Tiponut, A Mobile Robot Control Method with Artificial Neural Networks (I), *Transactions on Electronics and Communications,* Politehnica University Timisoara, Romania, Tom 48(62), Vol. 2, 2003, pp. 3-6.

[3] T. Botos, V. Tiponut, A Mobile Robot Control Method with Artificial Neural Networks (II), *Transactions on Electronics and Communications,* Politehnica University of Timisoara, Romania, Tom 48(62), Vol. 2, 2003, pp. 7-10.

[4] T. Botos, V. Tiponut, P.Gurka, Simulation Environment for Mobile Robots Guided by Neural Networks, *Proceedings of the Symposium on Electronics and Telecommunication "Etc.'98",* Timisoara, Sept. 17-18 1998, pp. 219-222.

[5] Gacsadi, Contributions to the Adaptive Control of Robots by Visual Information Processing using Cellular Neural Networks, *PhD Thesis, Politehnica University Timisoara*, Romania, 2001.

[6] Gavrilut, Contributions to the Autonomous Mobile Robots Navigation Using Cellular Neural Networks, *PhD Thesis, Politehnica University Timisoara*, Romania, 2006.

[7] Gavrilut, A. Gacsadi, L. Tepelea, V. Tiponut, Target Search by Mobile Robot in a Labyrinth by Using Cellular Neural Networks, *Proceedings of the International Conference on Engineering of Modern Electric Systems (EMES'2005),* ISSN 1454-9239, Oradea 2005, pp. 54-59.

[8] Gavrilut, A. Gacsadi, C. Grava, V. Tiponut, Vision Based Algorithm for Path Planning of a Mobile Robot by Using Cellular Neural Networks, *Int. Conf. On Automation, Quality &Testing, Robotics, AQTR 2006,* Cluj-Napoca, pp. 306-311.

[9] Gavrilut, V. Tiponut, A. Gacsadi, An Integrated Environment for Mobile Robot Navigation Based on CNN Images Processing, Proceedings of the 11th WSEAS International Conference on SYSTEMS, Agios Nicolaos, Crete, Greece, ISSN: 1790-5117, ISBN: 978-960-8457-90-4, July 2007, pp. 81-86.

[10] Gavrilut, V. Tiponut, A. Gacsadi, Path Planning of Mobile Robots by Using Cellular Neural Networks, *Proceddings of the IEEE International Workshop on Cellular Neural Networks and Their Applications (CNNA 2006),* Istambul, Turkey, 2006, pp. 234-239.

[11] Gacsadi, T. Maghiar, V. Tiponut, Path Planning for a Mobile Robot in an Environment with Obstacles Using Cellular Neural Network, *Intern. Workshop on CNN and their Applications, (CNNA 2002)*, Frankfurt/Main, 2002, pp. 188-194.

[12] Gavrilut, A. Gacsadi, L. Tepelea, V. Tiponut, Motion Planning for Two Mobile Robots in an Environment with Obstacles by Using Cellular Neural Networks, *The 7-th Intern. Symposium on Signals, Circuits and Systems, (ISSCS 2005)*, Iaşi, 2005, pp. 801-804.

[13] V. Tiponut, I. Gavrilut, C. Caleanu, A. Gacsadi, Neural Network Guided Robot Collectivity - An Experimental Setup, *Proceedings of the 7th WSEAS International Conference on Neural Networks*, Cavtat, Croația, 2006, pp. 41-45.

[14] V. Tiponut, I. Gavrilut, C. Caleanu , A. Gacsadi, Development of a Neural Network Guided Mobile Robot Collectivity, *WSEAS Transactions on Circuits and Systems,* ISSN 1109-2734, Issue 6, Vol. 5 June 2006, pp. 805-812.

[15] M. Aytekin, Sound Localization by Echolocating Bats, *PhD Thesis,* 2007. [16] R. Z. Shi, T. K. Horiuchi, A Neuromorphic VLSI Model of Bat Interaural Level Difference Processing for Azimuthal Echolocation, *IEEE Trans. Circuits and Systems,* vol. 54, 2007, pp. 74-88.

FRESNEL ZONE PLATE ANTENNAS

S. Savov[1], L. Kambourov[1], J. Urumov[1], H. Hristov[2]

[1]Technical University of Varna, Varna, Bulgaria
[2]Universidad Tecnica Federico Santa Maria, Valparaiso, Chile
9010 Varna, Bulgaria, phone: +359 52 383 408, e-mail: savovsv@yahoo.com

Abstract: *Microwave/mm-wave Fresnel zone plate (FZP) antennas, conformal to commonly used rotational curvilinear surfaces are studied and contrasted in this paper. Their vector diffraction theory is based on the conical-segment lens profile approximation proposed by the authors. The resulted in far-field equation has been explored for computer analysis of the main FZP antenna characteristics: co- and cross-polarization patterns, directive gain and aperture efficiency. The far-field theory for rotational FZP antenna is then extended for FZP antenna with lens conformal to a general curvilinear surface. FZP antennas, based on compound of different lens surfaces are also examined. These antenna constructions refer to combination of surfaces (plane-curvilinear or curvilinear only) that coincide with the geometry of real objects (plastic radome, edifice wall, roof, plane or satellite body surface, etc.).*

Keywords: antennas, far-field calculation, gain, polarization pattern

1. INTRODUCTION

The Fresnel zone plate (FZP) antennas originated in optics as a plane construction of a focusing lens [1] - [4]. This antenna uses half-open/half-blocked Fresnel zones. The plane FZP lens has advantage to be planar (2-D) structure with a small thickness, lighter and easier to manufacture. However, in some applications the zone plate surface has to be synthesized with a given curvilinear shape, or to be *conformal* to some manmade formation, rigid, collapsible or inflatable, as for example, building cupola, dome construction or space vehicle. In general, the FZP lens can be made on an arbitrary surface. However, more practical are the zone plates with an axially symmetry, like: spherical, parabolic and conical. In the FZP Antennas the increasing of the number of the zones in the plane variant results in strong decreasing the width of these zones. This fact impedes their practical realization. The application of curved surfaces provides possibilities for getting wider zones for same other conditions. Of important practical interest are the axial symmetric *focussing systems* that are smooth second order surfaces. It is assumed that the source and the focus of the Fresnel lenses both lie on the surface's axis. It is also accepted that the antennas could operate like transmitters/receivers, where the receiver is in the far-zone of the transmitter.

Some applications of FZP microwave / millimeter wave lenses and antennas *up to date*:

1. Lens / reflector antennas for reception of broadcast satellite signals.
2. Outdoor radio-relay communications antennas and passive repeaters.

3. Doppler radar antennas.

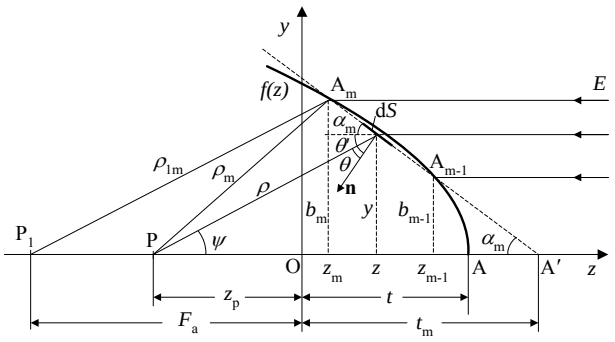
4. Measurement instrumentation: Michelson interferometer, measurement of dielectric materials, etc.

2. CONFORMAL FZP LENS ANTENNA ANALYSIS

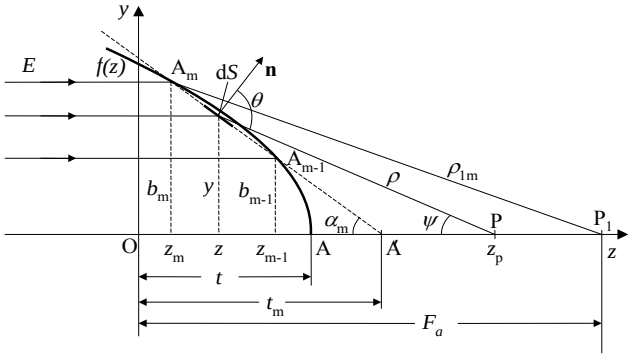
2.1. Geometry of the problem

In Fig. 1 are presented the half-section geometries of two types of curvilinear zone plates in the plane YOZ: *convex-side* illuminated (Fig. 1a) and *concave-side* illuminated (Fig. 1b). In both cases a *plane-wave* traveling in $\mp z$ direction, respectively, is assumed as a source. The wave is diffracted by the open Fresnel zone apertures and partially collimates at the primary focal point P_1.

The half-open Fresnel zone plate comprises circular curvilinear zones alternatively open and blocked. Here is assumed that the *odd zones* are open, while the even zones are blocked. The actual radiation of the curvilinear FZP lens includes *all open zones* on the lens surface. All these zones are projected in the *base aperture* (the plane $z = 0$).



(a)



(b)

Fig. 1. Geometry of curvilinear FZP lens: (a) convex-side illuminated and (b) concave-side illuminated by plane-wave front.

The field diffracted by the open zones is simplified by use of the *conical segment interpolation method* [2], [6] (see Fig. 1). It is also assumed that the supporting lens structure has a negligible effect on the lens electromagnetic behavior. The open curvi-

linear zones after the first one are presented by conical rings approximately. To make the lens shorter the first curvilinear zone can be replaced by its equivalent circular aperture (truncated FZP lens). Similar curve-linearization procedure is also applied to the blocked zones, which are covered by conical metal rings.

The design of the geometrical dimensions of one FZP lens includes determination of the radii b_m and the coordinates z_m, assuming that the following parameters are given:

1) *lens profile* function $f(z)$ – plane, spherical, conical, parabolic, or a surface compound of them;
2) *maximum sizes of the lens* in y – direction ($b_{\max}$) and in z-direction ($z_{\max}$);
3) *focal distance of the lens* F_a – between the focal point P_1 and the aperture plane ($z = 0$);
4) *wavelength* λ.

In the general case of symmetrical curvilinear surface of the lens, the parameters b_m are the radii of the projections of the mth zones on the plane $z = 0$, while z_m are the projections of the mth zones on the z-axis (Fig. 1). In the process of the determination of the radii and coordinates, an infinite thin lens is assumed for simplicity. The boundaries of the mth Fresnel zone are determined by the condition that the phase difference between the electromagnetic wave, arriving in the focal point from the internal circle of the zone and the direct electromagnetic wave (on the z-axis) to be $m\lambda/2$.

According to Fig. 1 for the convex and concave cases of FZP the design equations are respectively

$$\rho_{1m} + (t_m \mp z_m) = F_a + t_m + m\lambda/2, \quad \text{or} \quad \rho_{1m} \mp z_m = F_a + m\lambda/2 \tag{1}$$

where the *upper sign* is related to *convex case* (Fig. 1a), while the *lower sign* is related to the *concave one* (Fig. 1b), t_m is the thickness (height) of the cone, ρ_{1m} is the distance between the focal point and the boundary of the mth zone.

From these formulas for the radius of the mth zone a system of two simultaneous equations has to be solved to find the values of the two unknowns b_m and z_m is obtained

$$\begin{vmatrix} b_m^2 - m\lambda(F_a \pm z_m) - (m\lambda/2)^2 = 0 \\ b_m = f(z_m) \end{vmatrix} \quad (m = 0,\ 1,\ 2, \ldots, M) \tag{2}$$

For the particular case of a conical Fresnel lens, the corresponding function is $f(z) = (t - z)\tan\alpha$, where α is the semi-angle of the conical tip. For a *conical half-wavelength lens* the radii and the positions of the zones are respectively

$$b_m = \sqrt{2a(F_a + t) + a^2 + \left(\frac{a}{\tan\alpha}\right)^2} \pm \frac{a}{\tan\alpha}; \tag{3}$$

$$z_m = t - \frac{b_m}{\tan\alpha} \tag{4}$$

where $a = m\alpha/2$.

Hence, on the aperture of an arbitrary axial symmetric curvilinear surface can be placed *the same number* of M planar Fresnel zones.

As a consequence of this the following *properties* of Fresnel antennas with an axial symmetry can be formulated:

1. The *number of the Fresnel zones M*, that can be placed on an arbitrary smooth symmetrical surface for given wavelength depends on the diameter only and the distance of the last one to the source point *does not depend on the shape of the surface*;
2. The boundaries of the Fresnel zones on the curvilinear surface can be obtained from the cross-section of this surface with the *elliptical trajectories* that describe the boundaries of the Fresnel zones *on a plane*, orthogonal to the focal axis, when the last one moves from the aperture to the apex of the curvilinear surface.

2.2. Vector far-field calculation

The transparent for electromagnetic field Fresnel zones can be approximated for the symmetrical case by simpler truncated conical surfaces with different angles of the tip (Fig. 1). This approximation very simplifies the analysis of the radiation field of the FZP lens antenna.

The configuration of a curvilinear Fresnel antenna with an arbitrary surface is presented in Fig.2. Consider a point source situated in the focal point P_1. The axial-symmetric surface has a focal length $F = F_a + t$ (t is the lens thickness). This surface is illuminated from the the concave side by a *spherical wave* of the point source (here small horn antenna). The Fresnel zones are segments from curvilinear surface. The projections of the boundaries of these segments on the base aperture (the plane XOY) are concentric circles. The base aperture has a diameter $D_a = \overline{A_a A_a'}$. An actual radiating surface is the projections of the open zones on this plane. The points $A_{m,m-1}$ defines the borders of the Fresnel zones on the lens surface in the plane XOZ and have coordinates $x_{m,m-1} = b_{m,m-1}$ and $z_{m,m-1}$. A point from the elementary surface dS is defined with a radius-vector $\boldsymbol{\rho}$ which has a spherical coordinates (ρ, ψ, ς) in a coordinate systems with an origin at the focal point P_1 (Fig. 2). The field distribution of the electromagnetic wave radiated by the source on the aperture plane is described in the same coordinate system. A point P in the far zone is defined by the radius-vector

r and is described in other spherical coordinate system (r,θ,φ) with an origin at the point O.

By application of the *vector Kirchhoff's diffraction theory*, after a conversion from the Cartesian coordinates (x,y,z) to the *spherical coordinates* (r,θ,φ) the following equations for the *far-field* components are obtained [5], [6]:

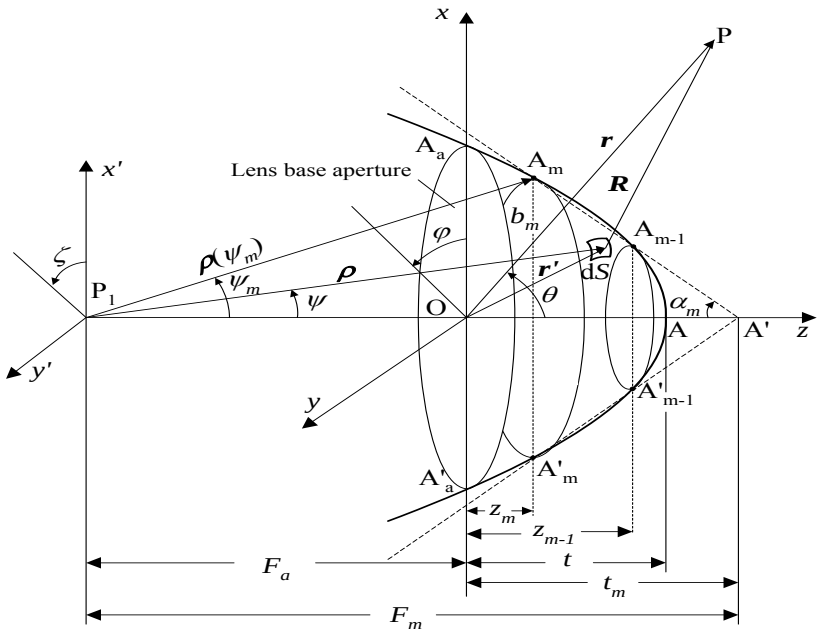


Fig. 2. Geometry of an axial-symmetrical Fresnel antenna

$$E_r(\theta,\varphi) \approx 0 \tag{5}$$

$$E_\theta(\theta,\varphi) = \pi C \cos\varphi \sum_{m=1}^{3,5,\ldots} \int_{\psi_{m-1}}^{\psi_m} Q(\psi) e^{M(\theta,\psi)} \left[-S'(\psi,\alpha_m) J_0\left(N(\theta,\psi)\right) + S''(\psi,\alpha_m) J_2\left(N(\theta,\psi)\right) \right] \mathrm{d}\psi \tag{6}$$

$$E_\varphi(\theta,\varphi) = \pi C \sin\varphi \sum_{m=1}^{3,5,\ldots} \int_{\psi_{m-1}}^{\psi_m} Q(\psi) e^{M(\theta,\psi)} \left\{ \left[S'(\psi,\alpha_m) J_0\left(N(\theta,\psi)\right) + S''(\psi,\alpha_m) J_2\left(N(\theta,\psi)\right) \right] \cos\theta + j2\cos\alpha_m \sin\theta J_1\left(N(\theta,\psi)\right) \right\} d\psi \tag{7}$$

where the constant is $C = \dfrac{j\beta e^{-j\beta r}}{2\pi r}\sqrt{60P_t}$ (here r is the distance to the far zone, and P_t is the transmitted power), the summation is performed on the *odd indices only* (open zones), and the single integrals are taken *numerically*. Here $J_0(.), J_1(.), J_2(.)$ are Bessel functions, first kind, order 0, 1 and 2, and the auxiliary functions under the integrals are:

$$M(\theta,\psi) = -j\beta\left[\frac{F_m \sin\alpha_m}{\sin(\psi+\alpha_m)} - \left(\frac{t_m \sin\alpha_m \cos\psi - F_a \cos\alpha_m \sin\psi}{\sin(\psi+\alpha_m)}\right)\cos\theta\right]; \quad (8)$$

$$N(\theta,\psi) = \frac{\beta F_m \sin\alpha_m \sin\theta \sin\psi}{\sin(\psi+\alpha_m)}; \quad (9)$$

$$S'(\psi,\alpha_m) = \sin(\psi+\alpha_m) + \sin(\alpha_m); \quad (10)$$

$$S''(\psi,\alpha_m) = \sin(\psi+\alpha_m) - \sin(\alpha_m); \quad (11)$$

$$Q(\psi) = \sqrt{G_f(\psi)}\,\frac{F_m \sin\alpha_m \sin\psi}{\sin^2(\psi+\alpha_m)}, \quad (12)$$

where the feed gain is $G_f(\psi)$.

The *directive gain* for the free space of the FZP antenna is given by the following expression [1]

$$\boxed{G(\theta,\varphi) = \frac{r^2}{60P_t}|E(\theta,\varphi)|^2} \quad (13)$$

The *co-polarization* field component $E_{co}(\theta,\varphi)$ and the *cross-polarization* field component $E_{xp}(\theta,\varphi)$ are expressed according to [5] as

$$E_{co}(\theta,\varphi) = \cos\varphi E_\theta(\theta,\varphi) - \sin\varphi E_\varphi(\theta,\varphi) \quad (14)$$

$$E_{xp}(\theta,\varphi) = -\sin\varphi E_\theta(\theta,\varphi) - \cos\varphi E_\varphi(\theta,\varphi) \quad (15)$$

The present theory is based on Kirchhoff's diffraction theory, which may produce inaccurate results for zone apertures that are very narrow in terms of the wavelength, as in the case of very high-numbered zones. For the planar half-open FZP antenna this problem is studied in [5] and [8],[9], where it is concluded that a combination of the Kirchhoff's diffraction theory for the main lobe region with the *Uniform Theory of Diffraction* (UTD) for the side lobes of the radiation pattern leads to much closer theoretical and experimental results. This more complex antenna analysis can be also applied to the curvilinear FZP antenna. Nevertheless, the microwave/millimeter FZP lenses consist in practice of no more than 20-30 zones and the classical Kirchhoff's method may be well suited for the corresponding FZP antennas.

The general far-field equations for the half-open curvilinear FZP lens antennas, given in this paper, are easily specified for the conical and plane lens profiles. In the *conical* FZP lens the opening semi-angle $\alpha_m = \alpha$ is constant, while for the *plane* FZP lens $\alpha = \pi/2$, $t = 0$ and $F_m = F_a = F$. Thus, the general radiation equations for the

curvilinear half-open FZP antennas, presented here, are simply converted to the corresponding equations of the plane half-open FZP antenna derived in [6],[10].

3. NUMERICAL RESULTS FOR CONFORMAL FZP ANTENNAS

3.1. Simple surfaces

Our first numerical study is carried out for a set of one planar and three curvilinear FZP antennas of spherical, parabolic or conical lens shape, designed for the millimeter wavelength $\lambda = 2.143$ mm that corresponds to the frequency of 140 GHz. Each curvilinear lens has *same circular base aperture* in the origin coordinate plane $z = 0$, and an apex A located at $z = t$ (we call these case Set I). The planar FZP lens coincides with the lens base aperture and is centered at the origin point O. All lenses are equal in diameter D_a ($D_a = \overline{A_a A_a'} = 2b_M = 200$ mm). The thickness of the plane FZP lens t_{pl} is assumed negligibly small and its focal length F_{pl} is equal to its diameter, i.e. $F_{pl} = D_a = F_a$. The curvilinear lenses are equal in thickness ($t = t_{sph} = t_{par} = t_{con} = 50$ mm) and focal length ($F = F_a + t = 250$ mm). The spherical, parabolic and conical antenna lenses are characterized by a sphere radius $R = 125$ mm, a shape focal length $F_p = 100$ mm and opening half angle $\alpha = 63.4^0$, respectively.

For the above dimensions and positioning all FZP antenna lenses (curved and plane) have the same number of zones $M = 23$, regardless of their different in shape surface profiles. This is in accordance with the corollary proved recently by the present authors in [10], which states that for a given wavelength *all FZP lenses independent of their shape, comprise the same number of zones* if they share a common transverse aperture, and have the same diameter and aperture distance to the focal point.

The odd zones are open (radiating) and the even zones are closed by thin metal or absorbing conical rings. The last closed Fresnel zone has a radius $b_{M-1} = b_{22} = 100$ mm, i.e. the physical lens aperture diameter is $D_a = 2b_{M-1}$ and the last open zone with a number $M = 23$ is considered as a virtual half-radiation zone. The feed horns of all antennas in Set I have identical axial-direction gain $G_f = 15$ dB, and the same aperture diameters and positioning, which ensure equal outer-edge illumination of their last closed Fresnel zones (-8 dB relative to G_f). For comparison the lens semi-profiles are sketched together in Fig. 3. As seen, the curvilinear FZP lenses of Set I are enclosed in a cylindrical volume 200 mm in diameter, and 50 mm in length. The plane lens occupies a negligible part of this volume, while the conical and the other two curvilinear lenses fill at least half of it.

Fig. 4 shows the half-portions of the 2D co-polarization radiation patterns of the FZP antennas with plane (dash-dotted line), conical (dashed line), spherical (solid line) and parabolic (dotted line) lenses*. As seen in Fig. 4, the *co-polarization radiation*

patterns of all curvilinear FZP lenses *are very similar*. Because the lens profiles of the spherical and parabolic lenses are very close in Fig. 5 shows in a much smaller angular span ($\theta = 0 \div 1$ deg). The planar FZP lens pattern is somewhat *different*, especially in the side-lobe region. Similar observations are valid for the *cross-polarization gain patterns* (Fig. 3). The maximum cross-polarization levels of all FZP lens antennas are more than 30 dB down with respect to the maximum co-polarization maximum gain (39 dB) as it can be seen from Figs. 4 and 5. The best cross-polarization discrimination of all FZP lens antennas is exhibited by the *planar* FZP lens: its maximum cross-polarization level is down by 37 dB relative to the co-polarization maximum gain.

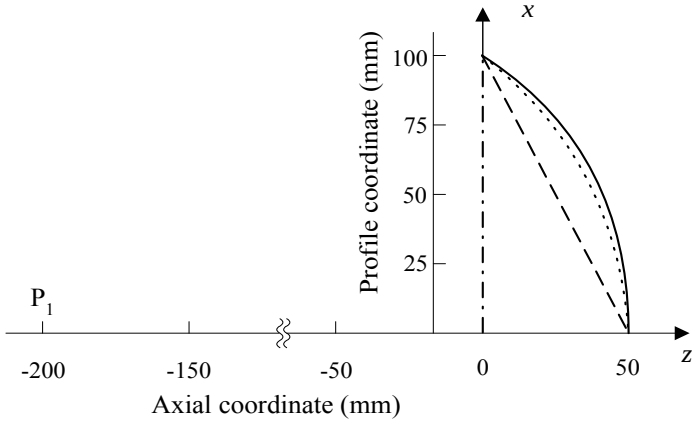


Fig. 3. Lens profile curves of FZP antennas: spherical (solid line), parabolic (dotted line), conical (dashed line) and planar (dash-dotted line).

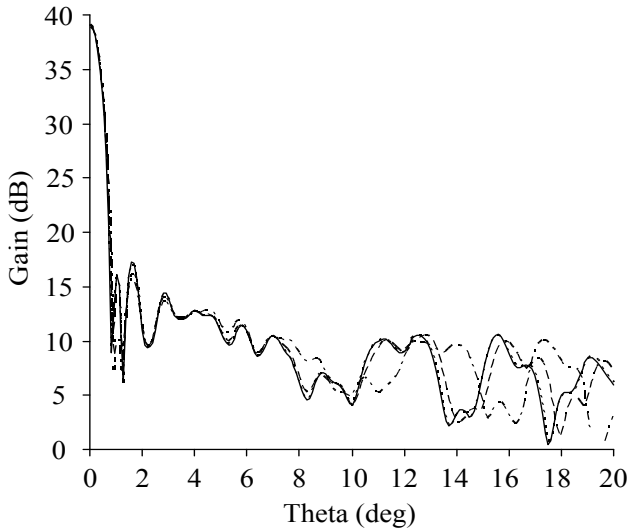


Fig. 4. Co-polarization gain radiation patterns of spherical (solid line), parabolic (doted line), conical (dashed line) and planar (dash-doted line) antennas.

The 3-D graph of Fig. 6 offers a better visualization of the co-polarization pattern shape for the case of the *spherical* FZP lens antenna. We have computed and compared the normalized theoretical gain plots versus the frequency (not shown in the paper) for the studied half-open plane and curvilinear FZP antennas. The gain-

frequency characteristics of all antennas virtually overlap and their -3 dB level *bandwidth* is about 10.6 GHz, which relatively to the design frequency is about 8 %.

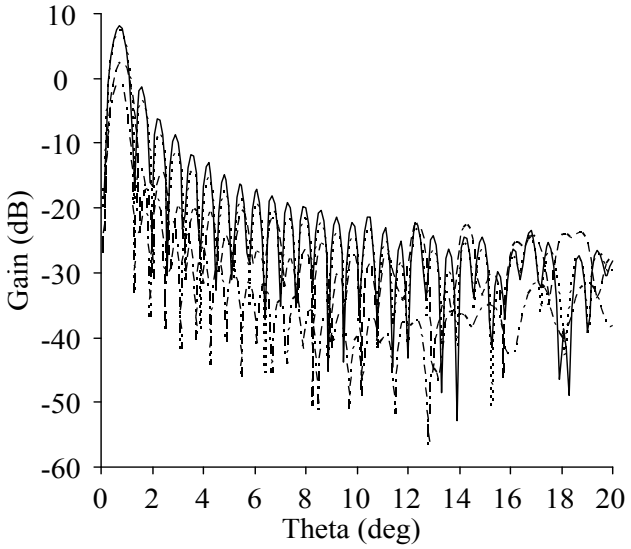


Fig. 5. Cross-polarization gain radiation patterns of spherical (solid line), parabolic(doted line), conical (dashed line) and planar (dash-doted line) antennas.

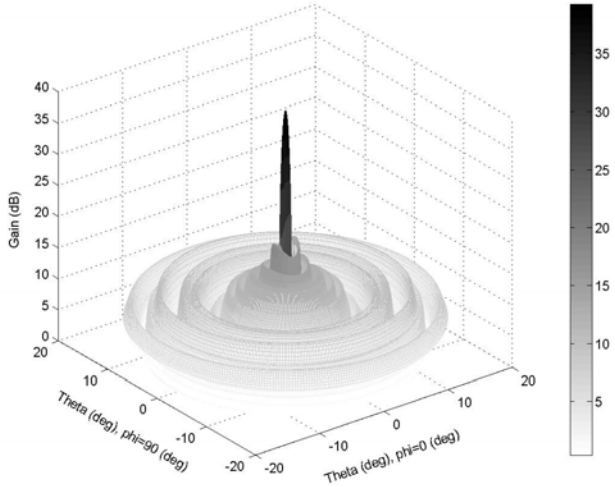


Fig. 6. Co-polarization 3-D radiation pattern of spherical FZP antenna

3.2. Compound surfaces

Now a set of compound antennas consist of Fresnel zones comprised on two different in shapes simple surfaces (one truncated cone and plane) as shown in Fig. 7. They are designed in according to the general equations considered in Section II. Here we are interested in compound antennas with common base aperture again (see the half profiles in Fig. 7). They have the same focal length and diameter and equal number of Fresnel zones (M=75). All the lenses are also illuminated by the same feed – *horn*, with gain $G_f = 11.3dB$ (Table 1).

As expected, in Fig. 8 *very good matching* between the radiation patterns (co- and cross-polar) of the truncated cone (compound) lens and the corresponding simple (not truncated) FZP lens *is observed*. In conclusion, the much *shorter* compound antenna (truncated conical FZP lens) has the same focusing action as the usual, non-truncated conical lens. All radiation patterns in this paper have been calculated for the $\varphi = 45^0$ diagonal plane cut.

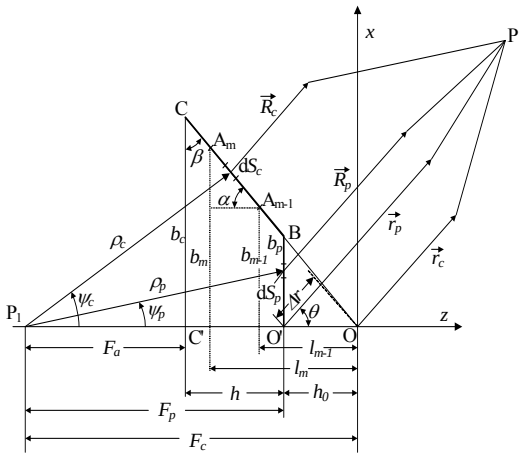


Fig.7. Radiation geometry of FZP lens antenna conformal to truncated conical dome

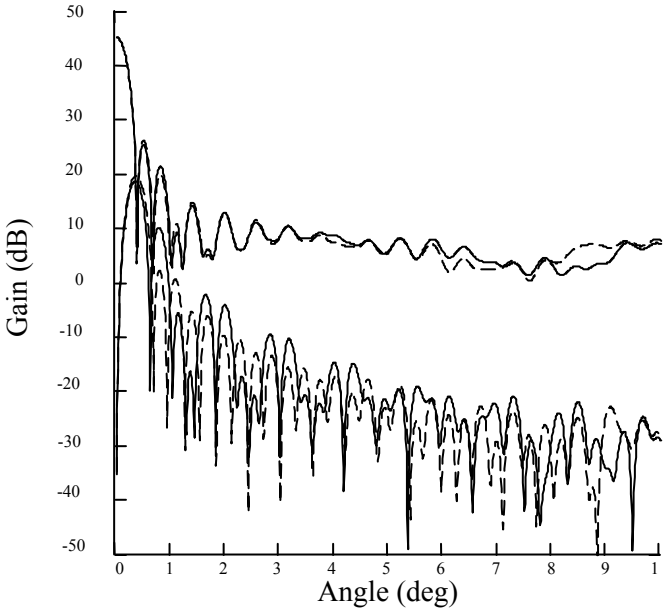


Fig. 8. Gain radiation patterns: *co-polar* (upper graphs) and *cross-polar* (lower graphs), for compound (*truncated conical lens*) (solid line) and for simple (*not truncated conical lens*) (dash line), $\alpha = 75^0$

The basic antenna radiation parameters of this compound FZP antenna are summarized in the Table 1 [13]. Here α is the half-angle of the conical tip, SL is the *side-lobe level*, G_{co} is the co-polarization gain, G_{xp} is the cross-polarization gain; G_f is the feed-horn gain, BW is the *beam-width* of the pattern's main lobe and η is the *aperture efficiency*. Here the frequency f=11.7GHz, the diameter of the aperture D=5m and

the focal distance is $F_a = 2.75\text{m}$. It is obvious that for this set of antennas (with a common base aperture), where the number of the Fresnel zones is constant $M=75$, many of the parameters are similar, for example the half-power beamwidth BW(0.33^0), and the co-polar gain G_{co} (almost constant, around 45 dB). With increasing the conical angle α, however, the side-lobe level SL tends to increase (from -25 dB to -15 dB) and the G_{xp} varies in wide range. The lowest cross-polarization is obtained for the case $\alpha = 75^0$ (see the table). The aperture efficiency η of the FZP antenna slightly decreases with the increasing of the conical angle α (from 9 to 8 %).

Table 1. Basic Antenna Radiation Parameters

α (deg)	G_{co} (dB)	SL (dB)	G_{xp} (dB)	G_f (dB)	BW (deg)	η (%)
45	45.2	-25.6	18.7	11.3	0.33	9.0
60	45.1	-23.0	10.8	11.3	0.33	8.6
75	**44.9**	**-19.8**	**-2.3**	**11.3**	**0.33**	**8.3**
90	44.7	-15.0	13.6	11.3	0.33	7.9

4. CONCLUSIONS

1. A comparison in height and complexity with the classical planar FZP antenna does not favor them; however, the antennas comprising compound FZP lenses can be in many cases the most natural choice because they are 3-D configurations and have more degree of freedom for design and optimization.

2. There is a possibility to increase the low aperture efficiency of this compound antenna, by covering the truncated conical surface with suitable inhomogeneous dielectric (for example, from 9% to 29% approximately). The most popular case is this of a binary dielectric layer that yields also increasing of the directive gain of the antenna in some extent (for example, from 45 dB to 50 dB approximately). However, this dielectric covered truncated cone FZP antenna becomes more difficult for fabrication, especially when the number of the zones tends to increase. This is just one way to extend the degrees of freedom of antenna design and optimization. More information about this particular case of conformal FZP antenna can be found in [14].

3. Compared to the ordinary antennas the conformal FZP configurations have small aperture efficiency but similar radiation resolution. Practically, the efficiency can be increased up to 70-80 % by use of zone multi-phase correction.

4. The FZP antennas are used like small in size microwave/mm-wave antennas or very large (solid or inflatable) dome-like telescopes.

5. The FZP antennas maintain the beauty of urban and rural environment by assembling them to be conformal to some arbitrary in shape, man-made or natural formation.

Potential Future Applications:

1. Antennas for fixed and mobile communication systems (rigid and collapsible).

2. Antennas for ground and space communications, radars or power transporting antennas.

3. Antennas for indoor wave focusing and specific field distribution.

4. Antennas for holography and medical diagnostics.

5. Antennas for plasma focusing and diagnostics.

6. Antennas for security applications: systems for sensing and detection of hidden objects (arms, mines, etc.).

REFERENCES

[1] J. Ojeda-Castañeda and C. Gomes-Reino, (Editors), "*Selected Papers on Zone Plates*", SPIE Milestone Series, 1996.

[2] H. D. Hristov, "*Fresnel Zones in Wireless Links, Zone Plate Lenses and Antennas*", *Artech House*, 2000.

[3] Y. J. Guo and S. K. Burton, "*Fresnel Zone Antennas*", *Kluwer Academic Publishing*, 2002.

[4] Minin and O. Minin, "*Difractional Optics of Millimeter Waves*", CRC Press, IoP, 2004.

[5] L. Leiten and M. H. A. J. Herben, "Vector far-field of the Fresnel-zone plate antenna", *Microwave and Opt. Tech. Letts*, vol. 5, no. 2, pp. 49-56, 1992.

[6] H.D. Hristov, L.P. Kamburov and J.R. Urumov, "Focusing characteristics of curvilinear half-open Fresnel zone plate lenses: plane wave illumination", *IEEE Trans. Antennas & Propagat.*, vol. 53, no. 6, pp. 1912 – 1919, 2005.

[7] A.C. Ludwig, "The definition of cross-polarization", *IEEE Trans. Antennas & Propagat*., vol. 21, no. 1, pp. 116-119, 1973.

[8] L.C.J. Baggen, and M.H.A.J. Herben, "Calculating the radiation pattern of a Fresnel zone plate antenna: A comparison between UTD/GTD and PO, *Electromagnetics*, vol. 15, pp. 321-345, 1995.

[9] J. Sluijter, M.H.A.J. Herben and O.J.G. Vullers, "Experimental validation of PO/UTD applied to Fresnel zone plate antennas", *Microwave Opt. Tech. Letts*, vol. 9, no. 2, pp. 111-113, 1995.

[10] L.P. Kamburov L.P., H.D. Hristov, J.R. Urumov and R. Feick, "Curvilinear Fresnel-Zone Plate Lens Antenna: Vector Radiation Theory", *International Journal of Infrared and Millimeter Waves*, vol. 26, no. 11, pp. 1593 – 1611, 2005.

[11] F. Sobel, F. L. Wentworth, and J.C. Wiltse, "Quasi-optical surface waveguide and other components for the 100-to 300-Gc region", *IRE Trans. Microwave Theory Tech*., vol. 9, pp. 512-518, 1961.

[12] D.N. Black, and J.C. Wiltse, "Millimeter-wave characteristics of phase-correcting Fresnel zone plates", *IEEE Trans. on Microwave Theory and Tech*., vol. 35, pp. 1122-1129, 1987.

[13] H. D. Hristov, L. P. Kamburov, J. R. Urumov, "Microwave antenna with compound Fresnel zone- plane lens conformal to truncated conical dome", *Proc. of 4th European Workshop on Conformal Antennas* (EWCA-05), May 2005, Stockholm, Sweden, pp. 113-116.

[14] H. D. Hristov, L. P. Kamburov, J. R. Urumov, R. Feick, A. I. Atanassov and W. Grote, "Conical double-dielectric Fresnel-zone lens and antenna", *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 5, pp. 325-327, 2007.

ON THE APPLICATION OF THE MODEL OF THE EFFECTIVE BANDWIDTH OF THE AMPLIFIED SPONTANEOUS EMISSION IN THE INVESTIGATION OF EDFA

Ivan M. Uzunov

Department of Applied Physics, Technical University of Sofia
8 Kl. Ohridski Blvd., Sofia 1000, Bulgaria, e-mail: ivan_uzunov@tu-sofia.bg

Abstract: *The purpose of this work is to explore the possibilities of the model of the effective bandwidth of the amplified spontaneous emission in the investigation of the following important issues related to EDFA: a) comparison of characteristics of EDFA with fixed fiber length for a 980-nm and 1480-nm pumps and a 1550-nm signal in the case of codirectional pumping; b) calculation of the dependence of the EDFA gain on fiber length for a 980-nm and 1480-nm pumps and a 1550-nm signal in the case of codirectional pumping, and c) comparison of three different pump configurations for pumping a EDFA with fixed fiber length for a 980-nm and a 1550-nm signal. It is established that the numerical results obtained in this work correlate qualitatively well with the corresponding results obtained by means of complete treatment of amplified spontaneous emission presented in [2].*

Keywords: optical amplification, EDFA, fiber-optic communication systems

1. INTRODUCTION

The most common optical amplifier today is the erbium-doped fiber amplifier (EDFA) operating in the C-band $(1525 \div 1565\,nm)$ due to its high gain $(30 \div 50\,dB)$, wide waveband $(\approx 35\,nm)$, low noise figure $(3 \div 5\,dB)$, and high level of the output signal $(10 - 20\,dBm)$ [1-3].

Three different pump configurations are possible for pumping an EDFA. Usually the pump light is injected from the same direction as the signal flow, which is known as codirectional pumping. It is also possible to inject the pump power in the opposite direction to the signal flow, which is known as counterdirectional pumping. The case of simultaneous codirectional and counterdirectional pumping is known as bidirectional pumping.

Modeling of EDFA includes solution of rate and propagation equations [1-2, 4]. Rate equations describe the effects of absorption, stimulated emission, and spontaneous emission on the population densities N_2 and N_1 of the ground energy level $^4I_{15/2}$ and metastable energy level $^4I_{13/2}$ in a two level approximation. Often population density of metastable energy level N_2 is taken in the steady-state approximation (see below) [1-2, 4]. The propagation equations describe the propagation of the powers of the beams through the fiber. [1-3]. Existing methods for numerical solutions of rate and propagation equations are summarized in [1-2, 4].

The full model of ASE needs to take the spectral shape of ASE, spanning from 1450 nm to 1650 nm, with channels spaced by 1nm [4,2]. By means of this model a complex relationship between amplified signal, pumps, ASE, and population density of metastable energy level can be completely analyzed. Results of such an analysis concerning our three chosen problems are summarized in [2].

A simple way to tackle ASE consists in merely modeling the ASE as a single signal with an effective bandwidth centered at a signal frequency. In this case, the problem is simple to treat numerically as there are only two ASE signals (with equal effective bandwidth) to track, on for the forward direction P_{ASE}^{+}, and one for the backward direction P_{ASE}^{-} [2].

The purpose of this work is to explore the possibilities of the model of the effective bandwidth of the amplified spontaneous emission in the investigation of following practically important issues related to EDFA: a) comparison of characteristics of EDFA with fixed fiber length for a 980-nm and 1480-nm pumps and a 1550-nm signal in the case of codirectional pumping; b) calculation of the dependence of the EDFA gain on fiber length for a 980-nm and 1480-nm pumps and a 1550-nm signal in the case of codirectional pumping, and c) comparison of three different pump configurations for pumping a EDFA with fixed fiber length for a 980-nm and a 1550-nm signal. The three different issues have been chosen due to their practical importance and to the fact that they give good possibility to examine the basic physics of interaction between amplified signal, pumps, ASE and upper state population. So the main question which will be answered in this paper is can the model of the effective bandwidth of the amplified spontaneous emission successfully describe the physics of the process of amplification.

2. BASIC EQUATIONS

In this work I will study amplification of signal with wavelength $\lambda_S = 1550\,nm$, whose initial power will be noted as $P_S(0)$. The initial power of the pump for codirectional pumping will be noted as $P_P^{+}(0)$, and the one for counterdirectional pumping as $P_P^{-}(0)$. The ASE will be described by the model of the effective bandwidth, whose value will be $\Delta\lambda = 10\,nm$ [4].

The power propagation equations for the amplified signal $P_S(z)[W]$, forward $P_P^{+}(z)[W]$ pump, backward $P_P^{-}(z)[W]$ pump, forward ASE component $P_{ASE}^{+}(z)[W]$, and backward ASE component $P_{ASE}^{-}(z)[W]$ are [1-2, 4]:

$$\begin{aligned}
\frac{dP_S}{dz} &= \Gamma_S\left(\sigma_S^E N_2 - \sigma_S^A(N_0 - N_2)\right)P_S \\
\frac{dP_P^{\pm}}{dz} &= \pm\Gamma_P\left(\sigma_P^E N_2 - \sigma_P^A(N_0 - N_2)\right)P_P^{\pm} \\
\frac{dP_{ASE}^{\pm}}{dz} &= \pm(\Gamma_S\sigma_S^E N_2 P_0 + \Gamma_S\left(\sigma_S^E N_2 - \sigma_S^A(N_0 - N_2)\right)P_{ASE}^{\pm})
\end{aligned} \tag{1}$$

where Γ_S are Γ_P are the overlap factors between the light field modes of the signal and pumps and the erbium distribution. $\sigma_S^E, \sigma_P^E \left[m^2\right]$ и $\sigma_S^A, \sigma_P^A \left[m^2\right]$ are the emission and absorption cross sections for the signal and pump wavelengths. $N_0 \left[ions/m^3\right]$ is the Er^{3+} concentration of the fiber core. $N_2 \left[ions/m^3\right]$ is the concentration of active Er^{3+} ions. $P_0 = 2h\nu_S \Delta\nu$ is the noise power in a bandwidth $\Delta\nu$, which corresponds to spontaneous emission, h is the Plank's constant, ν_S and is the frequency of the amplified signal. Following [4], we assume that $\Delta\nu = 1250\,GHz\ (\Delta\lambda = 10nm)$. Note that in Eq.(1) the fiber losses at the signal, pump, and ASE components wavelengths are not taken into account, as they can be neglected for typical amplifier lengths of $10 \pm 20m$.

The rate equation in the steady-state approximation relates the ion population density in the upper level N_2 with the field powers and the total ion density N_0 [1-2]:

$$N_2 = \frac{\frac{\tau\sigma_S^A}{Ah\nu_S}\Gamma_S P_S + \frac{\tau\sigma_S^A}{Ah\nu_S}\Gamma_S\left(P_{ASE}^+ + P_{ASE}^-\right) + \frac{\tau\sigma_P^A}{Ah\nu_P}\Gamma_P\left(P_P^+ + P_P^-\right)}{\frac{\tau\left(\sigma_S^E + \sigma_S^A\right)}{Ah\nu_S}\Gamma_S P_S + \frac{\tau\left(\sigma_S^E + \sigma_S^A\right)}{Ah\nu_S}\Gamma_S\left(P_{ASE}^+ + P_{ASE}^-\right) + \frac{\tau\left(\sigma_P^E + \sigma_P^A\right)}{Ah\nu_P}\Gamma_P\left(P_P^+ + P_P^-\right) + 1} N_0 \quad (2)$$

where $\tau \approx 10\,ms$ is the lifetime of the metastable energy level ${}^4I_{13/2}$ of Er^{3+}, and A is the effective cross- sectional area of the distribution of erbium ions.

Numerical treatment of the system of equations (1-2), requires solving of a dual value problem. The forward ASE component is zero at the beginning of the fiber $\left(P_{ASE}^+(0) = 0\right)$, and the backward ASE component is zero at the end of the fiber $\left(P_{ASE}^-(L) = 0\right)$. Note that L is the length of the fiber. The numerical solution of the coupled system of equations (1-2), is done by iterative procedure that uses Runge-Kutta method on Mathematica. The computation involves solving the equations by propagating the light fields forward and backward along the fiber, using the boundary conditions for the signal P_S, pumps P_P^+ and P_P^-, forward ASE component P_{ASE}^+ and backward ASE component P_{ASE}^-, until a self-consistent solution is found. The presence of both backward ASE component P_{ASE}^-, and backward pump P_P^- requires such an iterative procedure.

In the numerical simulations erbium doped fiber with the following parameters has been used [2]: the total ion density $N_0 = 7\times10^{24}\ ions/m^3$, the effective cross- sectional area of the distribution of erbium ions $A = 7{,}065\times10^{-12}\,m^2$, the overlap factors between the light field modes of the signal and pumps and the erbium distribution: $\Gamma_S = 0{,}40$ (for $1535-1565\ nm$), $\Gamma_P = 0{,}43$ (for $1480\ nm$), $\Gamma_P = 0{,}64$ (for $980\ nm$). For the 980 nm wavelength pump, the absorption and emission cross sections were assumed to be $\sigma_P^A = 2{,}7\times10^{-25}\,m^2$ and $\sigma_P^E = 0$, correspondingly [2]. For the 1480 nm

wavelength pump, the absorption and emission cross sections were determined to be $\sigma_P^A = 1{,}5\times10^{-25}\,m^2$ and $\sigma_P^E = 0{,}5\times10^{-25}\,m^2$, correspondingly [2]. Finally for 1550 nm signal the absorption and emission cross sections were determined to be $\sigma_S^A = 1{,}9\times10^{-25}\,m^2$ and $\sigma_S^E = 2{,}6\times10^{-25}\,m^2$, correspondingly [2].

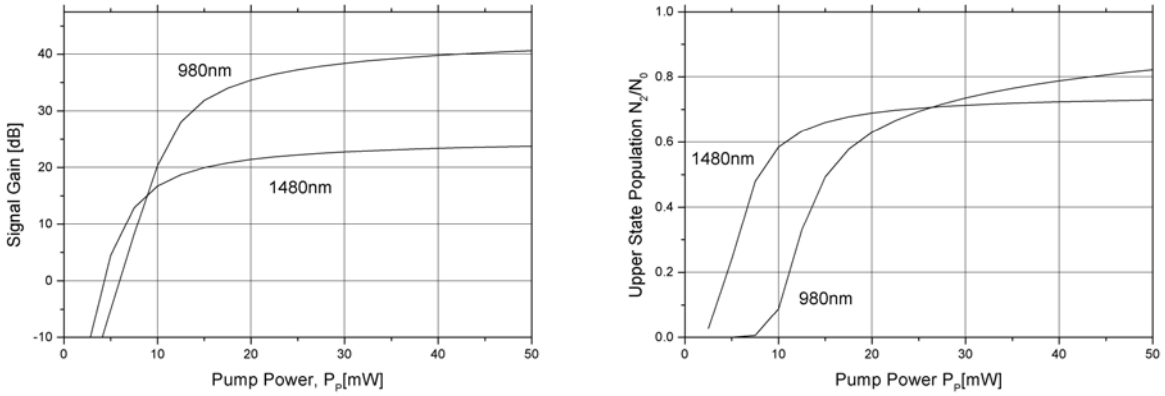
Gain in the case of codirectional pumping for a $L = 14\,m$ length of fiber pumped at $\lambda_P = 980\,nm$ and $\lambda_P = 1480\,nm$

Following [2], I study the gain of $L = 14\,m$ length of fiber, of injected signal with $P_S(0) = -40dBm$ and wavelength $\lambda_S = 1550\,nm$ (small- signal gain). The gain of erbium-doped fiber with length L (in dB) is defined as:

$$G[dB] = 10Log\, P_S(L)/P_S(0),$$

where $P_S(0)$ and $P_S(L)$ are the powers of signal in the beginning and in the end of the fiber, correspondingly.

At Fig.1 (a, b) are presented the dependencies of the gain $G[dB]$ and the population of the upper state N_2/N_0 as a function of pump power P_P^+ for two different pumping wavelengths: $\lambda_P = 980\,nm$ land $\lambda_P = 1480\,nm$.

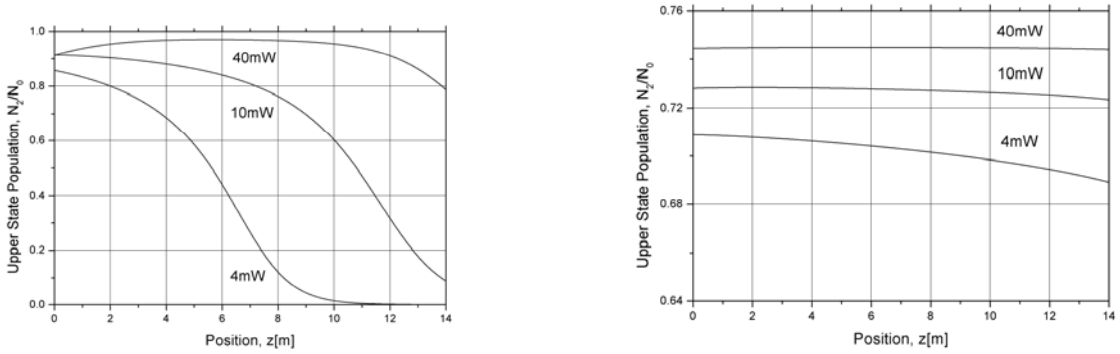


a) b)

Fig. 1(a, b) Gain $G[dB]$ (a) and population in the upper state N_2/N_0 (b) as a function of pump power P_P^+ (mW) for a 14m length of fiber pumped at $\lambda_P = 980\,nm$ and $\lambda_P = 1480\,nm$.

As can be seen from Fig. 1(a), and in agreement with [2], in the case of codirectional pumping, the $\lambda_P = 980\,nm$ pump wavelength yields higher gains than $\lambda_P = 1480\,nm$. As follows from the Fig. 1(b), this is related to the fact that pump with $\lambda_P = 980\,nm$ has achieved a higher inversion than pump with $\lambda_P = 1480\,nm$. The incomplete inversion due to nonzero emission cross section at $\lambda_P = 1480\,nm$, ($\sigma_P^E = 0{,}5\times10^{-25}\,m^2$) should be also mentioned.

At Fig. 2(a, b) is presented the dependence of the population in the upper state N_2/N_0 as a function of position, pumped at $\lambda_P = 980\,nm$ (a) and $\lambda_P = 1480\,nm$ (b), with the initial powers $P_P^+(0) = 4\,mW, 10\,mW, 40mW$. The same initial signal is amplified.

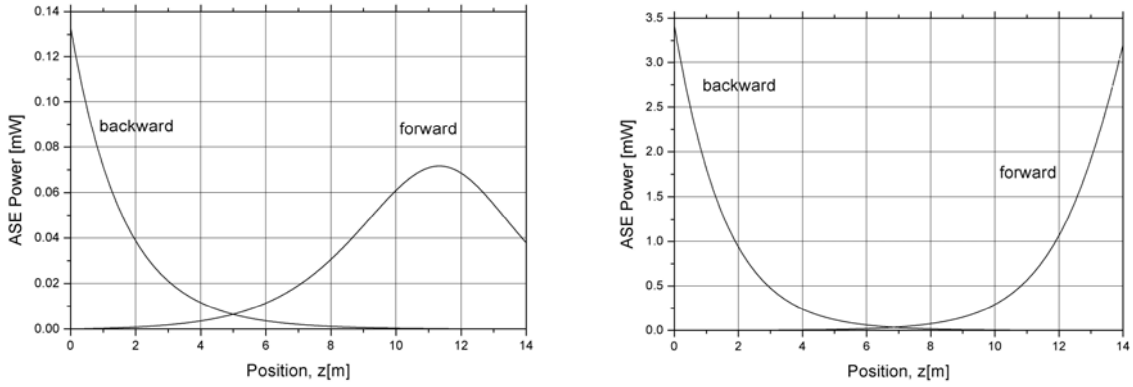


a) b)

Fig. 2(a, b) Population in the upper state N_2/N_0 as a function of position along the fiber pumped at $\lambda_P = 980\,nm$ (a) and $\lambda_P = 1480\,nm$ (b), with powers $P_P^+(0) = 4\,mW, 10\,mW, 40mW$

As can be seen from Fig. 2(a), at low power $P_P^+(0) = 4\,mW$, pump at $\lambda_P = 980\,nm$ is practically absorbed in the first part of the fiber, creating higher population in the upper state N_2/N_0 in comparison with the case of $\lambda_P = 1480\,nm$. Fig. 2(b) shows in the case of pump at $\lambda_P = 1480\,nm$ distribution of population in the upper state N_2/N_0 along the fiber is flatter, but also the population is smaller. This behavior of the population is also related to the nonzero emission cross section at $\lambda_P = 1480\,nm$ ($\sigma_P^E = 0{,}5 \times 10^{-25}\,m^2$).

How the large value of the population N_2/N_0 between 4 and 10m at pump at $\lambda_P = 980\,nm$ with pump power $P_P^+(0) = 40mW$ at Fig. 2(a) can be explained? Following the idea of analysis presented in [2] we have presented at Fig. 3(a, b) the forward ASE contribution P_{ASE}^+ and backward ASE contribution P_{ASE}^- as a function of position in the fiber pump at $\lambda_P = 980\,nm$ and pump powers: a) $P_P^+(0) = 10\,mW$, and b) $P_P^+(0) = 40mW$, respectively.



a) b)

Fig. 3(a, b) Forward ASE contribution P_{ASE}^+ and backward ASE contribution P_{ASE}^- as a function of position in the fiber pumped at $\lambda_P = 980\,nm$ with pump powers: a) $P_P^+(0) = 10\,mW$, and b) $P_P^+(0) = 40mW$, respectively

By comparing Fig. 3(a) and Fig 3(b) it can be clearly seen that, increase in the pump power leads to the strong increase in the forward ASE power P_{ASE}^{+} and backward ASE power P_{ASE}^{-}. Fig. 3(b) shows that, the large value of the population N_2/N_0 between 4 and 10m at pump at $\lambda_P = 980\,nm$ with pump power $P_P^{+}(0) = 40mW$ at Fig. 2(a) is related with the fact that between 4 and 8 m complete ASE power is minimal. Moreover, the lower value of the population N_2/N_0 in the beginning of the fiber for pump power $P_P^{+}(0) = 40mW$ in the same figure is connected with the large value of the backward ASE power P_{ASE}^{-} – see Fig.3 (b). So, smaller value of the population N_2/N_0 for high pump power leads to smaller gain. From Fig. 3(a) and Fig. 3(b) it can also be seen that, backward ASE power P_{ASE}^{-} is larger than the forward ASE power P_{ASE}^{+}. This observation can be explained by the fact that backward ASE power P_{ASE}^{-} travels more over well inverted fiber.

The ASE power plays the role in the formation of distribution of population N_2/N_0, which influences the gain of the signal that can be seen from the space evolution of pump powers along the fiber shown in Fig. 4.

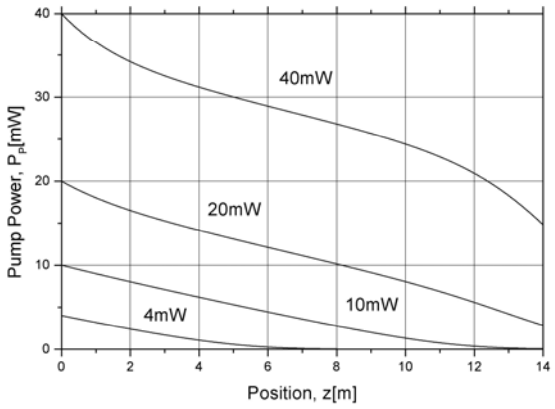


Fig. 4. Pump power as a function of position along the fiber for pump at $\lambda_P = 980\,nm$ and pump powers $P_P^{+}(0) = 4\,mW, 10\,mW, 20\,mW, 40\,mW$ during amplification of the same signal.

As can be seen from Fig. 4, the places where the power pump $P_P^{+}(0) = 40\,mW$ drops rapidly – in the beginning and the end of the fiber are due to the depletion of the pump from the backward ASE P_{ASE}^{-} and forward P_{ASE}^{+}, respectively (see Fig. 3(b)).

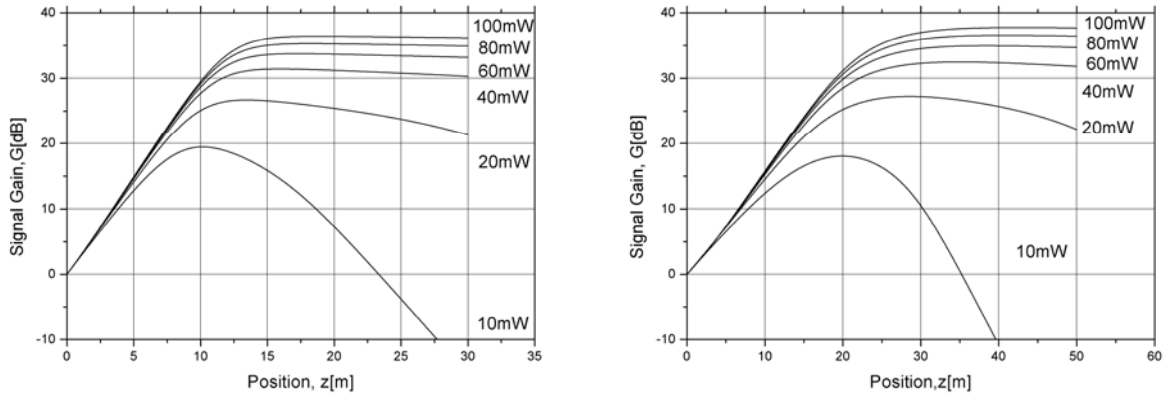
In conclusion, the larger gain in the case of codirectional pumping for a $L = 14\,m$ length of fiber pumped at $\lambda_P = 980\,nm$ at high powers in comparison with $\lambda_P = 1480\,nm$ is due to the larger value of the population in the upper state N_2/N_0. Important role in the formation of the population at high pump powers, starts to play the total ASE power, and in particular the backward ASE component P_{ASE}^{-}.

The results obtained in this paragraph correlate well with those obtained in chapter 6 of [2], obtained by complete description of ASE.

Gain as a function of fiber length in the case of codirectional pumping for a pump at $\lambda_P = 980\,nm$ and $\lambda_P = 1480\,nm$

I study the following characteristics of the amplification of an injected signal with wavelength $\lambda_S = 1550\,nm$: a) gain as a function of fiber length in the case of codirectional pumping for the input power of an injected signal $P_S(0) = -20dBm\ (10^{-5}W)$ and the following values of the initial pump power $P_P^+(0) = 10mW;\ 20mW;40mW;\ 60mW,80mW;100mW$ and b) gain as a function of pump power for the following fixed values of the initial power of the amplified signal $P_S(0) = 10^{-2}mW;10^{-1}mW;\ 1mW;10mW$. Pumping at $\lambda_P = 980\,nm$ and $\lambda_P = 1480\,nm$ is considered.

Fig. 5 shows gain $G[dB]$ as a function of fiber length in the case of codirectional pumping for the input power of an injected signal $P_S(0) = -20dBm\ (10^{-5}W)$ and the following pump powers: $P_P^+(0) = 10mW;\ 20mW;40mW;60mW,80mW;100mW$.

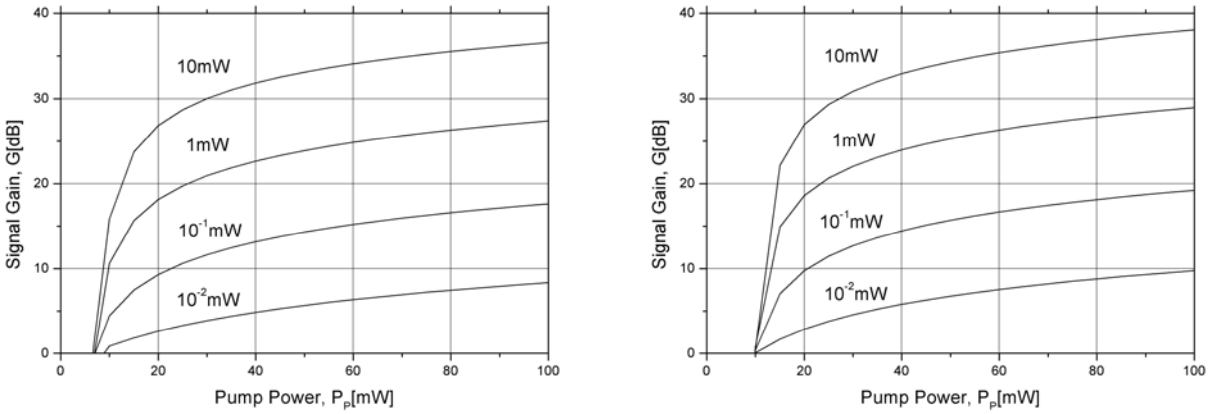


a) b)

Fig. 5. Gain $G[dB]$ as a function of fiber length in the case of codirectional pumping for the input power of signal $P_S(0) = -20dBm\ (10^{-5}W)$ and the pump powers: $P_P^+(0) = 10mW;\ 20mW;40mW;60mW,80mW;100mW$ pumped at: a) $\lambda_P = 980\,nm$; b) $\lambda_P = 1480\,nm$

As it is well known [3] and can be seen from Fig.5 for a given small pump power ($\leq 10mW$), the amplifier gain becomes maximum at an optimum value of position (optimum length of the doped fiber) and drops sharply when the distance exceeds the optimum value. As can be expected from the previous paragraph, till 15 m the gain for the pump at $\lambda_P = 980\,nm$ is larger than this at $\lambda_P = 1480\,nm$. As can be seen from Fig. 5, for the pump at $\lambda_P = 980\,nm$, effective gain (i.e. gain without saturation) appears till around $15m$ from the fiber, while for the pump at $\lambda_P = 1480\,nm$, it is till $30m$. So, the optimal length of the erbium-doped fiber for the pump at $\lambda_P = 980\,nm$ is around $15m$, while the optimal length of the erbium-doped fiber for the pump at $\lambda_P = 1480\,nm$ is around $30m$. Simultaneously, the achieved gain at $\lambda_P = 1480\,nm$ is larger than ones at $\lambda_P = 980\,nm$ [5-6].

Fig. 6 shows gain $G[dB]$ as a function of pump power $P_P^+(0) = 5\,mW \div 100\,mW$ for the given values of the initial power of the amplified signal $P_S(0) = 10^{-2}\,mW$; $10^{-1}\,mW$; $1mW$; $10mW$.



a) b)

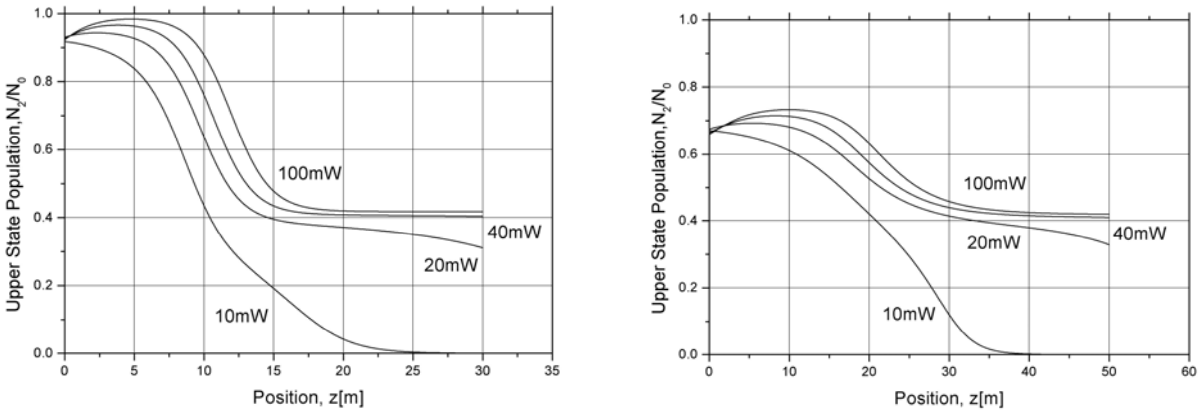
Fig. 6. Gain $G[dB]$ as a function of pump power $P_P^+(0) = 5\,mW \div 100\,mW$ for the following values of the initial power of the amplified signal $P_S(0) = 10^{-2}\,mW; 10^{-1}\,mW; 1mW; 10mW$: a) pump at $\lambda_P = 980\,nm$ and fiber length $L = 15\,m$; б) pump at $\lambda_P = 1480\,nm$ and fiber length $L = 35\,m$

As can be seen from Fig. 6, the amplifier gain initially increases exponentially with the pump power, but the increase becomes much smaller when the pump power exceeds a certain value around $15 - 20\,mW$. As mentioned previously the achieved gain at $\lambda_P = 1480\,nm$ and fiber length $L = 35\,m$ is larger than the ones at $\lambda_P = 980\,nm$ and fiber length $L = 15\,m$. Since the optimum length depends on the pump power it's necessary to choose both properly.

The reason for this is related to the peculiarity of distribution of the population of the upper state N_2/N_0 along the doped fiber. The first circumstance that should be taken into account is the condition for the existing of gain: $\sigma_S^E N_2 - \sigma_S^A N_1 = (\sigma_S^E - \sigma_S^A)N_2 - \sigma_S^A N > 0$, or $\frac{N_2}{N} > \frac{\sigma_S^A}{\sigma_S^E + \sigma_S^A}$. For amplified signal at $\lambda_S = 1550\,nm$, it leads to the condition $N_2/N > 0{,}42$. The second circumstance is the larger value of the absorption section $\sigma_P^A = 2{,}7 \times 10^{-25}\,m^2$ at $\lambda_P = 980\,nm$, in comparison with $\sigma_P^A = 1{,}5 \times 10^{-25}\,m^2$ at $\lambda_P = 1480\,nm$. The larger value of the absorption section leads to the larger population of the upper state (N_2/N_0). The larger value of the population of the upper state suggests faster redistribution of the pump power. The third circumstance is related to the influence of ASE on the population of the upper state N_2/N_0. As mentioned previously in the former paragraph (see Fig. 3) for $L = 14\,m$ doped fiber pumped at $\lambda_P = 980\,nm$, by increasing the pump power the generated total ASE grows. By increasing the length of the doped fiber, the growth of backward ASE

component P_{ASE}^{-} is particularly important. P_{ASE}^{-} depletes the pump power (see Fig.4), reducing the gain of the signal. It turned out [2] that by increasing the length of the doped fiber the generated backward ASE component P_{ASE}^{-}, obtained by pumping at $\lambda_P = 980\,nm$ is larger than the one generated by pumping at $\lambda_P = 1480\,nm$ [2].

The population of the upper state (N_2/N_0) as a function of position along the fiber as a result of the simultaneous influence of all these factors is shown at Fig. 6.



a) b)

Fig. 7. The population of the upper state (N_2/N_0) as a function of position along the fiber (for some of the cases presented at Fig. 5)

As can be seen from Fig. 7(a), the population of the upper state (N_2/N_0), pumped at $\lambda_P = 980\,nm$, after $15m$ in the fiber, is below or very close to the minimal for amplification. At the same time, Fig. 7(b) shows that, the population of the upper state (N_2/N_0), pumped at $\lambda_P = 1480\,nm$, maintains the necessary inversion levels over significantly longer lengths than the $\lambda_P = 980\,nm$ pump. This is the reason for the observed larger gain in the case of $L = 35\,m$ doped fiber pumped at $\lambda_P = 1480\,nm$ in comparison with the gain of $L = 15\,m$ doped fiber pumped at $\lambda_P = 980\,nm$.

The results obtained here are in a good agreement with those presented in [2].

A. Comparison of three different pump configurations for pumping EDFA

I will examine the amplification of initial signal with wavelength $\lambda = 1550\,nm$ and power $P_S(0) = -20dBm$ in EDFA with length $L = 15\,m$, pumped at $\lambda = 980\,nm$ and pump powers: $P_P^{+} = 50mW$ and/ or $P_P^{-} = 50mW$. The following characteristics of pump configurations will be compared: a) gain of signal P_S; b) space evolution of the powers of forward P_{ASE}^{+} and backward P_{ASE}^{-} ASE contributions; and c) population of upper state N_2/N_0 as a function of position in fiber.

The results for the signal power P_S as a function of position along the fiber for the three different pump configurations for pumping EDFA are presented in Fig. 8.

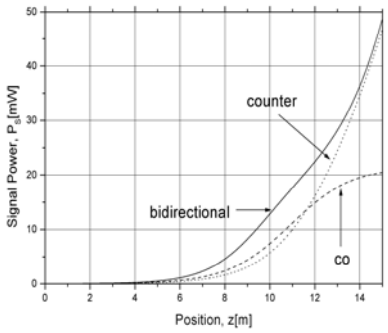


Fig. 8. Signal power P_S as a function of position along the fiber for the three different pump configurations for pumping EDFA

As can be seen from the Fig.8 the gain of the signal in the case of counter and bidirectional pumping is practically the same and larger than the one in the case of codirectional pumping.

The results for the powers of both components of ASE P_{ASE}^{+} and P_{ASE}^{-} as a function of position along the fiber for the three different pump configurations for pumping EDFA are presented in Fig. 9.

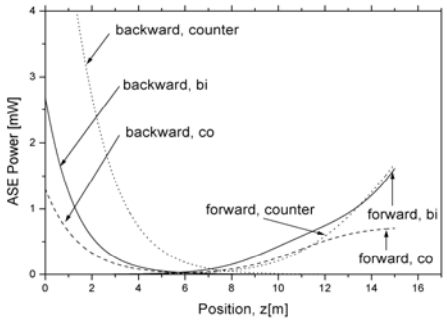


Fig. 9. Powers of forward P_{ASE}^{+} and backward P_{ASE}^{-} ASE components as a function of position along the fiber for the three different pump configurations for pumping EDFA

Regarding the power of the forward P_{ASE}^{+} ASE component, it is the largest in the case of counterdirectional, and the smallest in the case of codirectional pumping. The power of the forward P_{ASE}^{+} ASE component in the case of bidirectional pumping is between the values of codirectional and counterdirectional pumping but closer to the later.

The results for the population of the upper state (N_2/N_0) as a function of position along the fiber for the three different pump configurations for pumping EDFA are presented in Fig. 10.

As can be seen from the Fig.10, the bidirectional pumping provides the flattest spatial distribution of the upper state population. Such an observation correlates well with the observed maximal signal gain in Fig. 8 and the intermediate values of powers of forward P_{ASE}^{+} and backward P_{ASE}^{-} ASE components in this case are obtained in Fig. 9.

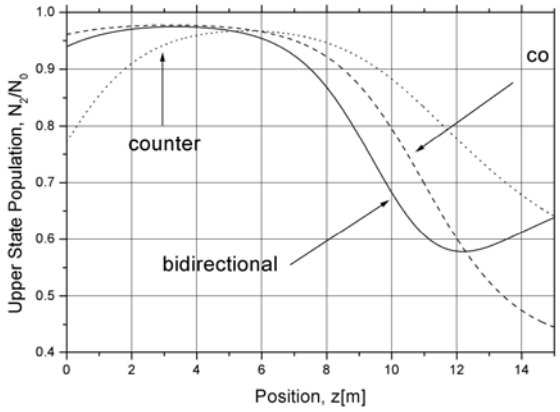


Fig. 10. Population of the upper state (N_2/N_0) as a function of position along the fiber for the three different pump configurations for pumping EDFA

Let me summarize the obtained results in this paragraph. For the fiber under consideration it was shown that the codirectional pumping provides the minimum signal gain, but yields the smallest values of the powers of forward P_{ASE}^{+} and backward P_{ASE}^{-} ASE components. Meanwhile, the application of counterdirectional pumping provides high signal gain, but is also responsible for the largest values of the powers of forward P_{ASE}^{+} and backward P_{ASE}^{-} ASE components. The bidirectional pumping is more efficient than them. It provides the high gain, and yields reduction of total ASE power, in particular substantial reduction of the power of backward P_{ASE}^{-} ASE component.

The presented in this paragraph results are in a good agreement with numerical and experimental results presented in [1, 2].

4. CONCLUSION

In this work the possibilities of the model of the effective bandwidth of the amplified spontaneous emission in the investigation of the following important issues related to EDFA have been explored: a) comparison of characteristics of EDFA with fixed fiber length for a 980-nm and 1480-nm pumps and a 1550-nm signal in the case of codirectional pumping; b) calculation of the dependence of the EDFA gain on fiber length for a 980-nm and 1480-nm pumps and a 1550-nm signal in the case of codirectional pumping, and c) comparison of three different pump configurations for pumping a EDFA with given fiber length for a 980-nm and a 1550-nm signal. Great attention has been paid to each of considered issues to clarify the interaction between signal, pumps, ASE contributions and the population in the upper state.

It is established that the results obtained in this work correlate qualitatively well with the corresponding results obtained by means of complete treatment of amplified spontaneous emission presented in [2].

So we can conclude that independently of its simplicity the model of the effective bandwidth of the amplified spontaneous emission can be useful in the investigation of

practically important issues related to EDFA. Of course, in the case of analysis of WDM systems it should be replaced with the model that includes complete description of ASE.

5. ACKNOWLEDGMENTS

This research was supported by the Project 102 ни 122-20 with the TU-Sofia, Bulgaria.

REFERENCES

[1] E. Desurvire, “Erbium-Doped Fiber Amplifiers – Principal and Applications”, John Wiley & Sons, Inc., 1994.

[2] P. Becker, N. Olsson, and J. Simpson, “Erbium-Doped Fiber Amplifier - Fundamental and Technologies”, Academic Press, 1997.

[3] G. P. Agrawal, “Fiber-optic communication systems”, John Wiley&Sons, Inc., 2002.

[4] C. R. Giles and E. Desurvire, “Modeling erbium-doped fiber amplifiers”, Journal of Lightwave Technology, vol. 9, pp. 271-283, 1991.

[5] J. H. Povsen, A. Bjarklev, O. Lumholt, H. Vendeltorp-Pommer, K. Rotwitt, and T. Rasmussen, “Optimizing gain and noise performance of EDFA’s with insertion of a filter or an isolator”, in Fiber Laser Sources and Amplifiers, M.J. Digonnet, and E. Snitzer, Ed. Proc. SPIE 1581, pp. 107-113, 1991.

[6] L. Jordanova, and V. Topchiev, “Optimizing the Parameters of Amplifiers used in the optical channel for CATV Systems” Telecom 2008, Conference Proceedings, pp. 206-211, Varna, Bulgaria, 2008.

FEMTOSECOND FIBER LASERS: BASIC CHARACTERISTICS AND OPERATION REGIMES

I. G. Koprinkov

Department of Applied Physics, Technical University of Sofia,
8 Kl. Ochridski Blvd, 1000 Sofia, Bulgaria, phone:+02 965 3072, e-mail:igk@tu-sofia.bg

Abstract: *The main advantages of the optical fibers as a propagation and gain medium are summarized. The absorption and the emission characteristics of erbium and ytterbium active media are presented. The pulse propagation along passive and active optical fibers is considered. The most widely applicable mode-locked methods are given. The main operation regimes of the femtosecond fiber lasers: soliton regime, dispersion managed soliton regime, self-similar regime and all normal dispersion regime, are presented. The dispersion profile of the laser cavity and the pulse behavior at the various mode-locking regimes are shown.*

Keywords: optical fiber lasers, mode locking, femtosecond pulses

1. INTRODUCTION

The ultrafast science cannot be considered anymore as an exotic and curious field of research, instead, it now sets the standards in many branches of the modern science and technology. The wide spreading of the ultrafast methods substantially depends on the creation of reliable femtosecond light sources. The femtosecond fiber lasers (FSFL) become the most promising candidate for such a role in an increasing number of applications. It is based on the unique reliability and operation stability accompanied by compactness and relatively low cost of those laser systems. Although the former generation of femtosecond lasers, the titan sapphire lasers, remains the cutting edge technology for the generation of ultrashort high-energy pulses, the FSFL still have great potential for development and continuously shorten the gap with their titan sapphire counterparts. Generation of 30 fs pulses and up to few tens nJ pulse energy directly from the fiber oscillators is now possible. Recently, single-cycle pulses at 1500 nm lasting 4.3 fs only have been reported by external compression of optical supercontinuum obtained by a nonlinear conversion of erbium oscillator-amplifier emission in highly nonlinear fibers [1].

Pulse energy of the FSFL can further be increased using fiber amplifiers. Chirped pulse amplification technique, initially introduced in the bulk lasers, has been successfully applied in the fiber lasers. Using photonic-crystal fibers as a gain medium allows increasing the single mode field diameter up to 50 μm and more, which is nearly a hundred times the mode area that of the standard single mode fibers. This increases the onset level of the nonlinearities by the same factor, so the amplifier emission can reach much higher power/energy. In contrast to bulk solid state laser systems, thermal lensing in the fiber lasers and amplifiers is strongly reduced,

allowing average power scaling to the kW range, keeping in the same time the compact size of the system. Thus, *e.g.*, average power of $835\,W$ in trains of $660\,fs$ pulses [2] and pulse energy of $8.2\,\mu J$ in trains of $375\,fs$ pulses [3] have been achieved. The average power and pulse energy of the fiber lasers and amplifiers has been strongly increased since the introduction of double-clad fibers.

The FSFL exploiting rare-earth-doped fibers allow generation of femtosecond pulses in the spectral range not directly covered by the other femtosecond lasers. The most widely used fiber lasers are based on ytterbium or erbium doped fibers that have broad enough gain bandwidth needed for generation of femtosecond pulses. Ytterbium doped FSFL emitting from 1020 to 1120 nm offer the highest powers. Erbium fiber lasers are less powerful and efficient but operate near 1550 nm that allows using highly reliable telecom components. The above two gain media have favorable material properties, such as wide-bandwidth, low quantum defect and compatibility with telecom grade pump laser diodes. FSFL emitting at longer wavelengths have also been demonstrated using thulium, and holmium doped fibers. A number of mode-locking regimes has been developed in order to generate ultrashort pulses in the different fiber active media: soliton regime [4-6], dispersion managed soliton (DMS) regime (also known as stretched mode regime) [7, 8], self-similar (SS) regime [9, 10], and all-normal dispersion (ANDi) regime [11, 12].

The aim of the present lecture is to make a short introduction in some of the basic properties and operation principles of the FSFL. The advantages of the optical fibers as propagation and gain medium will be presented. The optical pulse propagation along passive and gain fibers will be considered. The main working regimes of the femtosecond fiber oscillators will be given.

2. ADVANTAGE OF THE FIBER LASER TECHNOLOGY

In this section, we will summarize the main advantages of the optical fibers as a gain and propagation medium. The optical fibers in the fiber lasers can be classified as *active* or *doped* fibers and *passive* fibers. The core of the active fibers is doped by ions of the respective active elements (ytterbium, erbium, thulium, etc.) that allow achieving optical gain at suitable pump conditions. The passive fibers serve as a propagation medium for the generated light thus closing the laser cavity. For the case of the FSFL, however, the passive fibers play, in addition, the important role to manage the cavity dispersion and, thus, the mode-locking regime.

The advantages of the fiber lasers, based on the advantages of the optical fibers as gain and propagation medium, can be summarized as:

1) compactness, flexibility and small size.

2) environmentally stable operation – the influence of the external factors is strongly reduces due to the confinement of the light inside the fiber core.

3) excellent and stable transversal light distribution if single-mode fibers are used.

4) excellent mode matching between the pump and the generated emission is automatically provided in the fiber core.

5) long gain medium that ensures efficient operation.
6) reliable and long lived telecom components can be used.
7) excellent thermal properties and simplified cooling.
8) alignment and maintenance free operation.
9) relatively low cost.

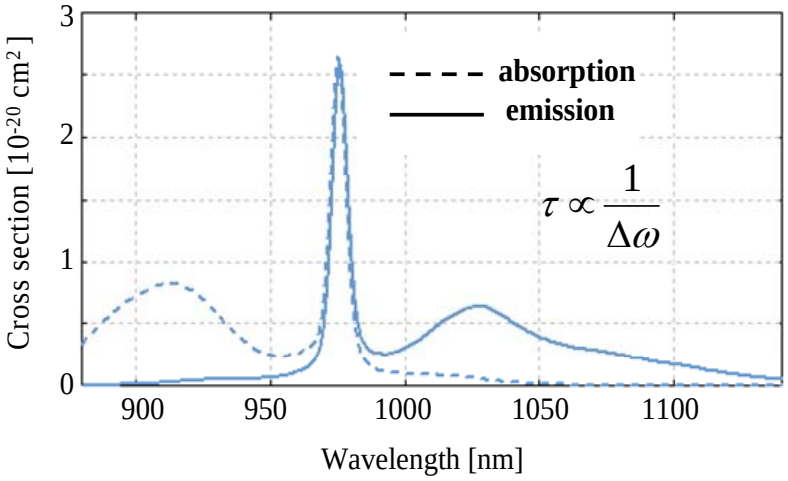
The above properties cannot be usually met in the other femtosecond laser systems.

3. RADIATION CHARACTERISTICS OF YTTERBIUM AND ERBIUM DOPED FIBERS

The energy structure and radiation characteristics of ytterbium and erbium doped fibers will be considered here as they are most widely used media in the fiber lasers.

3.1. Ytterbium doped fibers

Ytterbium ions have two strong absorption bands in the near-infrared regions that can be used for optical pumping, a broad absorption band around $910\,nm$ and a narrow absorption band around $975\,nm$, Fig.1. The emission also consists of two intense bands, a narrow one centered around $975\,nm$ and a broad one covered roughly 1000–1120 nm spectral range. In the broader emission band, the emission cross-section strongly dominates the absorption cross-section. The relevant energy level structure includes the ground-state $^2F_{7/2}$ and the excited state $^2F_{5/2}$ manifolds, Fig.2.



Fig, 1. Absorption and emission cross sections of Yb^{3+} ion in the near infrared region.

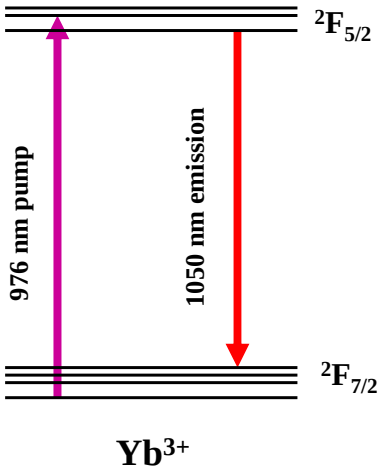


Fig. 2. Energy level diagram of Yb^{3+} ion.

Pumping and emission involve transitions between different but closely spaced sublevels of the ground-state and excited-state manifolds, $^2F_{5/2} \rightarrow {}^2F_{7/2}$ transitions. It

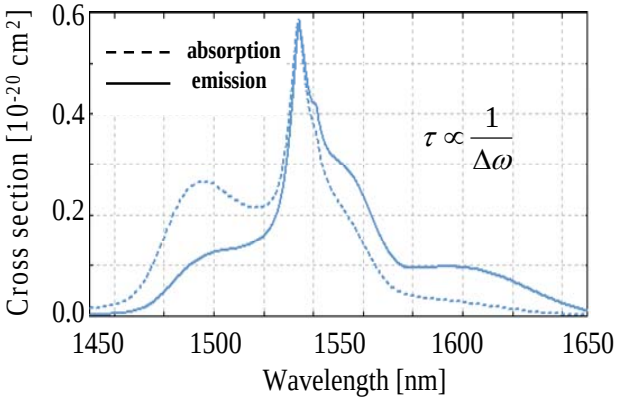
results in a small quantum defect between the pump and emission photons and becomes important for the high efficiencies of Yb fiber lasers as well as for the reduced thermal effects. On the other hand, the small quantum defect leads to some unwanted consequence, *i.e.*, strong quasi-three-level behavior for lasing around 1030 nm. Nearly four-level behavior and very little reabsorption is observed beyond 1080 nm. It requires an increased pump intensity to reach the laser threshold, especially in the shorter wavelength region. Another difficulty arises from the cavity design: a cavity mirror (for end-pumped fiber lasers) must have a high-reflectivity at the laser wavelength and a high-transmission at the slightly shorter pump wavelength. Similar problems arise for the wavelength division multiplexer (WDM) used in the ring fiber lasers. Such components ate difficult to make.

The gain bandwidth of the laser transitions is typically very large, up to about one hundred nanometers, which allows for generating femtosecond pulses in mode-locking regime. The lifetime of the upper-state $^2F_{5/2}$ of the ytterbium ion is long, about $(1-2)\,ms$, which is beneficial for Q switching. The wavelength range around $1\,\mu m$, emitted by Yb-doped fiber lasers, falls in the normal dispersion regime of the ordinary fibers. This imposes specific difficulties in the generation of femtosecond pulses and new regimes of mode-locking have been developed for Yb-doped fiber lasers [9-12].

3.2. Erbium doped fibers

Erbium ions have two strong absorption bands in the near-infrared regions centered around $975\,nm$ and $1450\,nm$. They can be used for optical pumping of Er-doped fiber lasers. The emission and absorption spectrum of the Erbium ions around $1550\,nm$ is shown in Fig. 3. The emission of Er doped fiber lasers is centered around $1550\,nm$ and involves $^4I_{13/2} \rightarrow {}^4I_{15/2}$ transition. The pump transition at $975\,nm$ terminates on $^4I_{11/2}$ level that, after a non-radiation transition, populates the upper laser level $^4I_{13/2}$, Fig. 4. The pump transition at $1450\,nm$ populates directly sublevels of the upper laser level $^4I_{13/2}$, which results in a very low quantum defect and a high pumping efficiency. It leads to the standard three-level problems already mentioned for the Yb fiber lasers.

The gain bandwidth of the laser transitions is large, about few tens nanometers, which allows for generating femtosecond pulses in mode-locked regime. The lifetime of the upper-level $^4I_{13/2}$ of the erbium ion is very long, about $(8-10)\,ms$, which is beneficial for Q switching. The wavelength range around $1550\,nm$, emitted by Er-doped fiber lasers, falls in the range of anomalous dispersion of the ordinary fibers. It allows a soliton operation regime of the FSFL.



Fig, 3. Absorption and emission cross sections of Er^{3+} ion in the near infrared region.

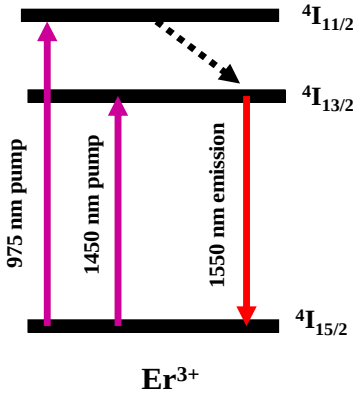


Fig. 4. A simplified energy level diagram of Er^{3+} ion.

4. MODE-LOCKING REGIME OF THE FIBER LASERS

All working regimes known from bulk lasers can be realized in the fiber lasers. Among these, the mode-locking regime is one that leads to the generation of femto-second pulses. The mode-locking can be realized in continuous wave (CW) or Q-switching modes of operation of the laser. In the CW mode-locking, the laser generates a uniform sequence of identical pulses (pulse train). In the Q-switching mode-locking, the output of the laser consists of relatively short train of pulses of variable duration having bell shape envelope of their amplitudes. The shortest pulses are generated in the CW mode-locking that is why it is the preferable regime of operation of the laser oscillators.

The pulse duration τ is limited from below from the bandwidth $\Delta\nu$ of the generated laser spectrum by means of a time-bandwidth product relation, $\tau = C/\Delta\nu$, where the coefficient C is in the order of unity and depends on the shape of the pulse.

For Gaussian pulse, $C = 0.441$, for sech-pulse, $C = 0.315$, etc. Thus, the generation of short pulses requires large gain bandwidth $\Delta\nu$ of the active medium, *i.e.*, a large number of excited laser cavity modes, Fig. 5. This condition is well satisfied for the fiber lasers, Fig. 1, 3. The broad spectrum, however, is not sufficient to generate short pulses. If the excited cavity modes appear with random phases, the emission will have noise-like character. To obtain short pulses, a definite phase relationship of the cavity modes must be ensured. This is achieved in the mode-locking regime.

Two general types of mode locking exist, passive and active ones. Each of these can be realized by a number of ways listed below.

Passive mode-locking

- Saturable absorber mode-locking
- Kerr lens mode-locking
- Colliding pulse mode-locking
- Nonlinear polarization rotation mode-locking
- Additive pulse mode-locking

Active mode-locking

- Amplitude modulation mode-locking
- Frequency modulation mode-locking
- Synchronous pumping mode-locking

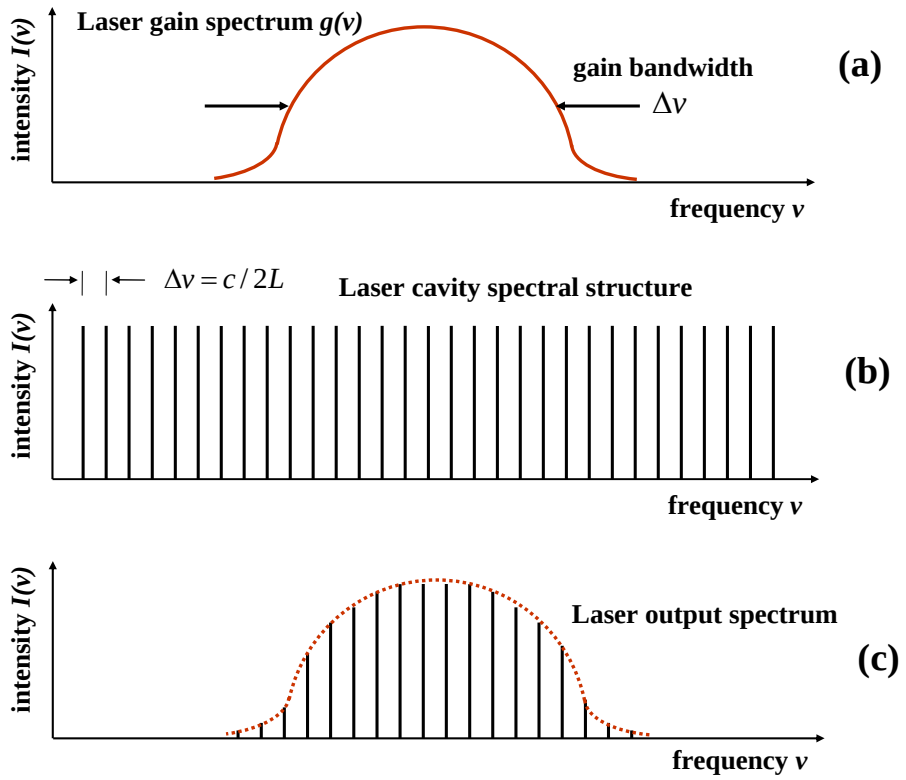


Fig. 5. Laser gain spectrum (a), laser cavity spectral structure (b) and the generated spectrum (c)

The passive mode-locking is much more frequently used, especially in the fiber lasers because of its simplicity and better achieved results. That is why, passive mode-locking only will be considered here. The most widely used passive mode-locking methods in the fiber lasers are based on the semiconductor saturable absorber mirror (SESAM) and the nonlinear polarization rotation (NPR). Both methods represent particular implementations of the saturable absorber (SA) passive mode locking method: a real absorber - for the case of SESAM, and a virtual absorber – for the case of NPR. The basic idea of the saturable absorber mode-locking is to create high losses for the low-intensity radiation and low losses for the high-intensity radiation. In practice, the low-intensity radiation appears in a form of CW radiation or long pulses, whereas the short pulses usually have high-intensity. The high-intensity pulses, more particularly, the leading part of the pulse may easily saturate the absorber, which results in low losses for the subsequent part of the pulse. This is not the case for long or CW light of low-intensity. Due to the short relaxation time of the absorber material, multiple absorption cycles may take place for the low-intensity light that results in high losses. All these cause automatic selection of the short pulses and suppress the generation of long pulses or CW radiation.

4.1. Sesam

The structure of SESAM is shown in Fig. 6. It consists of two main components: a saturable absorber and a total reflectance mirror (Bragg mirror) deposited on a suitable substrate. The saturable absorber consists of a semiconductor material with a band gap slightly lower than the photon energy of the laser emission. It can be fabricated in a form of single or multiple thin, *e.g.*, quantum wells, or thicker absorption layers. Low temperature GaAs or InGaAs are typically used as a SA material. The Bragg mirror consists of AlAs and GaAs quarter-wave layers. The basic parameters of the SESAM as a mode-locking device are: recovery time, modulation depth, bandwidth, saturation intensity and non-saturable losses.

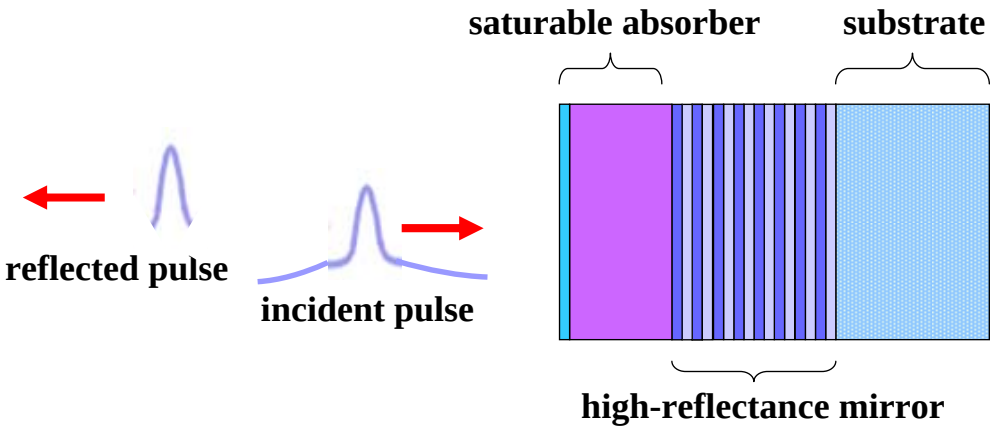


Fig. 6. Structure and operation principle of SESAM

The SESAM operates in the following way. The SA of SESAM absorbs photons from the laser emission, which results in an excitation of electrons from the valence band to the conductive band and generation of holes in the valence band. If the photon flux inside the cavity has low-intensity, the SESAM is in a linear absorption regime, which leads to strong losses. As a result, the high-reflectance mirror becomes “hidden” behind the SA and the laser cavity is closed. Thus, laser generation cannot develop. On the other hand, high-intensity light inside the cavity may strongly excite the SA and to reduce the number of the electrons in the valence band. It reduces the absorption of the SESAM and, at given level of excitation, the absorption saturates and the absorber is bleached. At this stage, the pulse “sees” the high-reflectance mirror and the cavity becomes open. Now, laser generation can develop inside the cavity. The electron-hole pairs recombine after a mean relaxation time (recovery time) due to different effects, *e. g.*, phonon or photon emission, and the SESAM restores again its ability to absorb emission. It again closes the cavity. Hence, SESAM tolerates the pulsed operation and suppresses the CW operation. Stable mode-locking is achieved at shorter relaxation time. On the other hand, extremely short relaxation time may cause excess losses of the pulse. In addition, it may also strongly affect the pulse dynamics (if the pulse is chirped) at the impact with the SESAM. Fast non-radiative

relaxation of the electron-hole carriers can be fabricated introducing lattice defects in the absorber layer of SESAM.

The mode-locked lasers with saturable absorbers may tend to Q-switch. It may become stronger expressed if the modulation depth of absorber is high. Q-switched mode-locking can be suppressed by decreasing the laser spot size on the SESAM, or increasing the average power inside the cavity (by decreasing the output coupling or increasing the pump power). The peak intensity of the pulse is limited by the damage threshold of the SESAM, which is about $300\,MW/cm^2$.

4.2. Nonlinear polarization rotation

Another widely used mode-locking method is based on the NPR effect in conjunction with a polarizer. The NPR works in a similar way as the SESAM, although no real absorption element exists in this case. The NPR relies on the following operation principle, Fig. 7. The optical fibers have some natural birefringence that originates from the fabrication process. The ordinary optical fiber can be considered as an effective Kerr medium whose local birefringence change along the fiber length. That is why, the birefringence tends to change the state of the polarization of the light propagating along the fiber. The mode-locking scheme of such laser includes the elements shown in Fig. 7. Let the light passes through a polarizer P_1, which results in a linear polarization of the transmitted light. By a sequence of three wave-plates, $\lambda/4-\lambda/2-\lambda/4$, that together form the so-called polarization controller PC, any state of polarization of the light can be created from the initial one. The PC can be done either by using discrete elements or entirely from optical fiber. The latter case is in accordance to the all-fiber concept of the mode-locked lasers. Due to the linear law of rotation of the polarization of the light in the PC, the low-intensity and the high-intensity light have same state of polarization after the PC. This, however, is not the case along the optical fiber. The Kerr effect leads to a nonlinear change of the polarization state inside the fiber. It results in different rotation of the polarization ellipse for the low-intensity and the high-intensity light after the fiber as a result of the intensity dependent nonlinear phase acquired by the different polarization components, Fig. 7. In this way, the low-intensity and the high-intensity light will experience different losses while passing through a second polarizer P_2. Both polarizers are usually one and the same polarizer that the light hits twice while making round trips inside the cavity. By control of the polarization of the light by the PC, we may tune the polarization state so as favor either the CW or the mode-locked operation of the laser.

Independently on the type of the mode-locking, the laser emits a series of regularly spaced short (picosecond or femtosecond) pulses, a pulse train, similar to that shown in Fig. 8. The time separation of the pulses is equal to the cavity round trip time and depends on the length of the cavity.

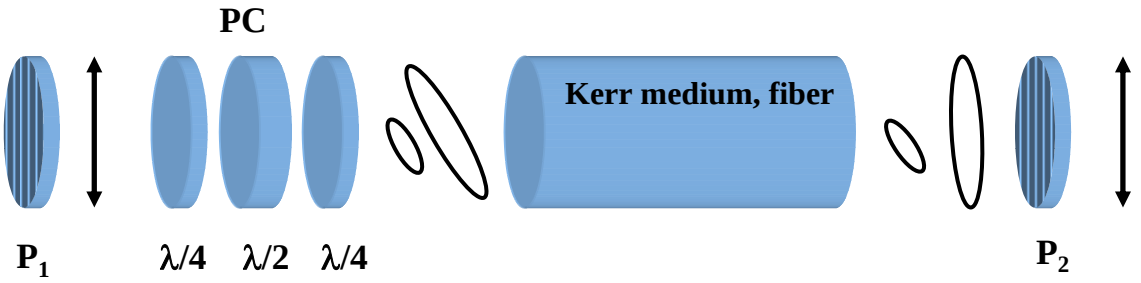


Fig. 7. Principle of operation of the nonlinear polarization rotation, NPR

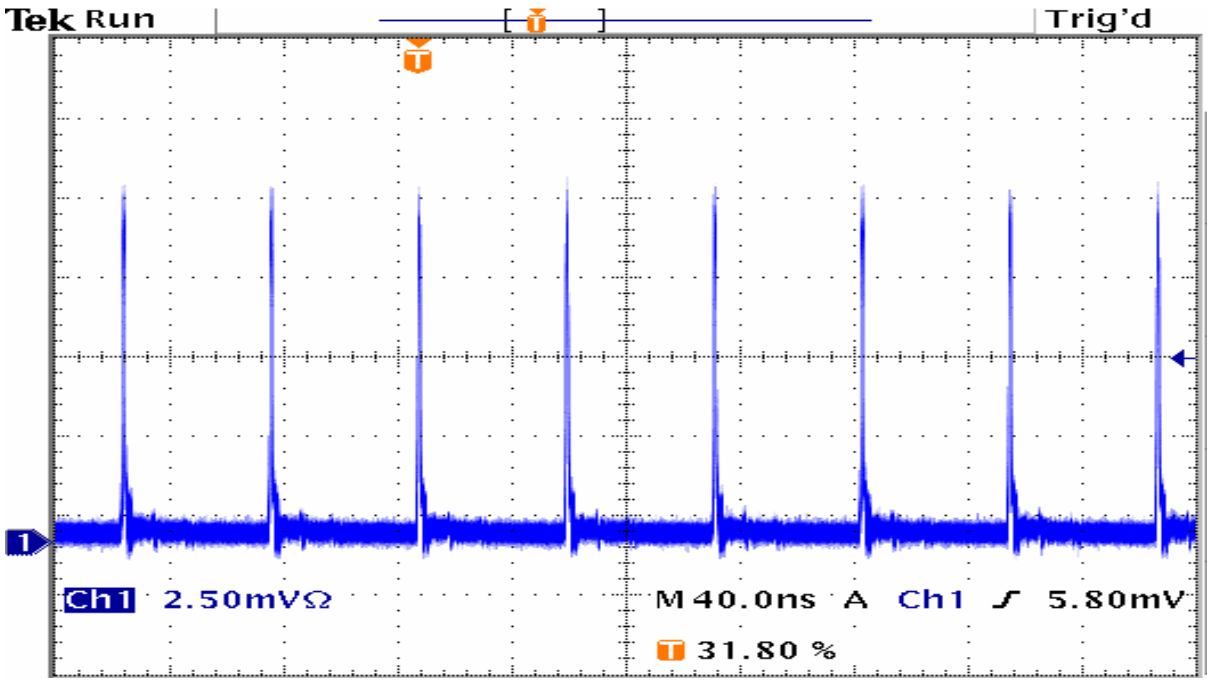


Fig. 8. A CW mode-locking pulse train

5. PULSE PROPAGATION IN THE FIBER LASER

The mode-locking mechanism generates short pulses at the mode-locker. During the propagation along the fiber, however, the optical pulse is a subject to a number of dispersive or nonlinear processes that may lead to substantial distortion, broadening and splitting of the initial pulse thus limiting the capabilities of the mode-locking operation of the fiber laser. That is why, the final generation of ultrashort pulses requires a stable or, at least, a well controllable way of propagation of the pulse inside the fiber laser cavity.

The simplest case of stable pulse propagation along the optical fiber occurs if a balance between the group velocity dispersion (GVD) and the self-phase modulation (SPM), being the leading orders of dispersion and nonlinearity, is ensured. The GVD and the SPM act in opposite direction in a medium of negative GVD and may exactly cancel each other at certain conditions. As a result, the temporal and spectral shape of the pulse remains constant during the propagation thus leading to soliton formation.

In a medium of positive GVD, both processes act in the same direction, which leads continuous broadening of the pulse. The pulse propagation in that simplest case is ruled by the Nonlinear Schrödinger equation (NLSE) for slowly-varying electric field amplitude $A(z,t)$ (in a frame moving with the group velocity of the pulse) [13].

$$\frac{\partial A}{\partial z} + i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} - i\gamma\left|A\right|^2 A = 0, \tag{1}$$

where β_2 is the GVD, $\gamma = kn_2 / A_{eff}$ is the nonlinear parameter, k is the wave number, n_2 is the nonlinear refractive index, A_{eff} is the effective cross section of the fundamental fiber mode, z is the propagation distance along the fiber, and t is the local time of the pulse. The first order optical soliton of the NLSE is given by the famous expression $A(z,t) = A_0 \sec h(t/\tau)\exp(iz/2)$.

The dispersion and the nonlinearity can be characterized in terms of dispersion $L_D = \tau^2 \Big/ \left|\beta_2\right|$ and nonlinear $L_N = 1/\gamma P$ lengths, where τ is the soliton width (τ is related to the full width at half maximum duration τ_{FWHM} of the soliton by $\tau = \tau_{FWHM}/1.76$), and P is the peak power of the soliton. In terms of the characteristic lengths, the first order *soliton threshold* can be expressed as $L_D = L_N$. The latter condition reveals the existence of a strong relationship between the soliton parameters and the parameters of the propagation medium/fiber

$$P\tau^2 = \frac{\beta_2 A_{eff}}{kn_2}, \tag{2}$$

or, equivalently

$$W\tau \approx \frac{\beta_2 A_{eff}}{kn_2}, \tag{2'}$$

taking into account that peak power-time duration product equals the energy W of the pulse, $W \approx P\tau$.

The propagation dynamics of ultrashort pulses around and below 100 *fs*, that are generated directly from some of the fiber oscillators, is substantially more complex than that one described by the simplest form of the NLSE, Eq. (1). The basic optical phenomena accompanying the propagation of such pulses along the optical fibers can be described by the modified (perturbed) NLSE [13].

$$\frac{\partial A}{\partial z} + i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} - \frac{\beta_3}{6}\frac{\partial^3 A}{\partial t^3} - i\gamma\left|A\right|^2 A - \frac{i}{\omega_0}\frac{\partial}{\partial t}\left(\left|A\right|^2 A\right) - T_R A\frac{\partial\left|A\right|^2}{\partial t} = 0, \tag{3}$$

where the third order dispersion, self-steepening and intra-pulse Raman scattering (third, fifth and sixth terms in (3)) have to be taken into account.

The pulse propagation in real fiber mode-locked lasers becomes even more complex due the existence of loss and gain of definite bandwidth and a SA. They do not cause the phase modulation but the amplitude modulation in the time and the spectral domain. Such physics can be described within the nonlinear Ginzburg-Landau equation (NLGLE), which is a generalization of the NLSE when switching from the conservative (Hamiltonian) systems to the dissipative systems. A rather good form of the NLGLE to model the real mode-locked lasers is [14].

$$\frac{\partial A}{\partial z} + i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} - i\gamma\left|A\right|^2 A - \frac{1}{\Omega}\frac{\partial^2 A}{\partial t^2} - gA - \alpha\left|A\right|^2 A - \delta\left|A\right|^4 A = 0, \qquad (4)$$

where g is net gain and loss, Ω is related to the filter bandwidth, α and δ are cubic and quintic SA terms. Stable localized light structures due to flow of energy, gain and loss, known as dissipative solitons, exist for the NLGLE [15].

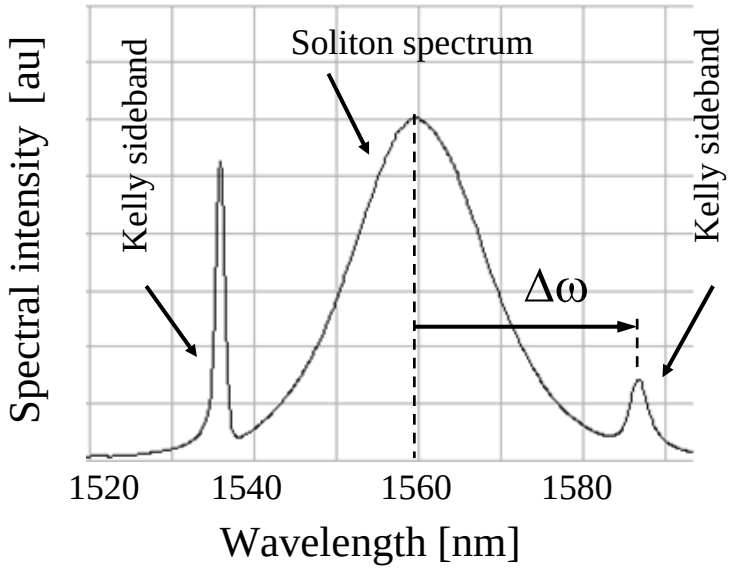


Fig. 9. The soliton spectrum and the Kelly sidebands

The main advantage of the soliton is in its remarkable stability. If the parameters of the initial pulse do not exactly match those of the soliton, the pulse automatically evolves toward the soliton shape while shedding the excess energy into a dispersive wave [16-18]. The intensity of the dispersive wave increases if the shape of the initial pulse departures stronger from the exact soliton shape. A dispersive wave can also be formed when the soliton is perturbed due to a rapid gain or a sudden loss (*e.g.*, at the laser output, causing a deviation from the soliton condition by suddenly reducing the pulse energy), or by a transition into a fiber with modified parameters. Similar effects occur for quasi-soliton pulses circulating in a cavity of different localized positions of

the dispersion and the nonlinearity. The periodic perturbations couple the soliton to the copropagating dispersive wave. The soliton is also stable against the change of the medium parameters if the changes occur over distances that are long with respect to the soliton period, $L_0 = L_D\pi/2$. Thus, the soliton can adiabatically adapt its profile to some slowly varying parameters of the medium. The soliton instabilities arise if the length of the laser cavity L_C exceed 8 times the soliton period, $L_C \geq 8L_0 = 4\pi\tau^2/\beta_2$. The soliton can also accommodate some amount of higher-order dispersion while keeping its stability.

In the time domain, the dispersive wave represents a temporally spreading background due to the GVD. Such spreading is not compensated by the nonlinearity due to the low intensity of the dispersive wave. In the spectral domain, the dispersive wave appears as the so-called Kelly sidebands, Fig. 9. The Kelly sidebands become better pronounced as the pulse duration is decreased. The frequency offset $\Delta\omega$ of the Kelly sideband with respect to the peak of the soliton spectrum is given by, Fig. 8,

$$\Delta\omega = \pm\frac{1}{\tau}\sqrt{m\frac{8L_0}{L_C} - 1}, \tag{5}$$

where m is an integer.

If the relation between the pulse energy and the fundamental soliton energy is the square of an integer number, the so-called higher-order soliton occurs. Such pulses do not have a constant shape, but periodically change their shape with a period equal to the soliton period L_0. The higher-order solitons are much less stable than the fundamental solitons and can split into fundamental solitons under the influence of higher-order dispersion and other disturbing effects.

6. MAIN REGIMES OF OPERATION OF THE FEMTOSECOND FIBER LASERS

Within the mode-locking operation, different regimes of pulse formation in the fiber lasers exist depending on the parameters of the laser cavity. The variety of the regimes of the mode-locked fiber lasers is enforced by the gain efficiency of the active fibers and the low losses of the system as a whole, which makes it possible to operate with fairly low pump powers as well as to use intracavity elements with losses that usually cannot be tolerated in the bulk lasers. One may distinguish four main regimes of the FSFL: soliton regime, dispersion managed soliton (DMS) regime, self-similar (SS) regime, and all normal dispersion (ANDi) regime. In the following, the cavity characteristics and the general behavior of the pulse in the above regimes will be shortly outlined.

6.1. Soliton regime

The soliton regime requires negative GVD, $\beta_2 < 0$, in any space point of the cavity (the ideal case) or, at least, net negative total cavity dispersion, $D = \sum_i \beta_{2i} l_i < 0$, where β_{2i} is the GVD of each cavity element and l_i is the length of that element. The above requirement determines the spectral range and the type of the fiber laser, in which the soliton regime can be realized. Based on standard silica fibers, the soliton regime can be realized at $\lambda \geq 1.3\,\mu m$, *i.e.*, above the zero-dispersion point of the fibers. It also means that the erbium fiber lasers, generating around $\lambda \approx 1.55\,\mu m$, may operate in soliton regime, but not the ytterbium fiber lasers, that operate around $\lambda \approx 1\,\mu m$. In contrast to the standard fibers, the photonic crystal fibers may provide negative GVD at a shorter wavelength thus approaching the range of operation of the ytterbium fiber lasers.

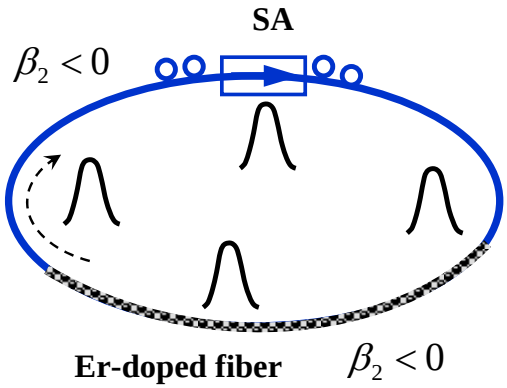


Fig. 10. Schematic representation of a soliton ring fiber laser cavity

A soliton fiber laser is shown schematically in Fig. 10. It consists of a standard single mode fiber (SMF) of negative GVD at the erbium wavelength, and an erbium doped (single mode) fiber. For simplicity, the pump input and the laser output of this and the subsequent types of lasers are not shown in the respective figures. Depending on the fiber parameters, the doped fiber may have negative GVD (the preferable case) or positive GVD. In any case, however, large negative total dispersion of the cavity must be provided in order to have a soliton regime. Although the value of the cavity dispersion is not exactly specified, the following range of cavity dispersion is usually accepted in the soliton regime: $D \approx -(0.05 \div 0.1)\,ps^2$.

The main advantage of the lasers working in soliton regime [4-6] is the unique stability of the generated pulses. For the ideal case of constant negative GVD inside the cavity and small loss/gain perturbations within a cavity roundtrip, the soliton fiber laser generates transform limited (chirp-free) pulses that keep a constant shape and time duration in each cavity point, Fig. 11. The main disadvantage of the soliton lasers is that they typically generate relatively long pulses of relatively low energy. In addition, the spectrum has parasitic Kelly sidebands. The best reported results with Er

fiber soliton lasers are several-hundred femtoseconds for the pulse duration and tens of picojoules for the pulse energy. The shortest pulses obtained directly from a soliton oscillator have time duration of 167 *fs* [19]. Recently, 48 *fs* pulses by compression through amplification of the emission of a soliton oscillator have been measured [20].

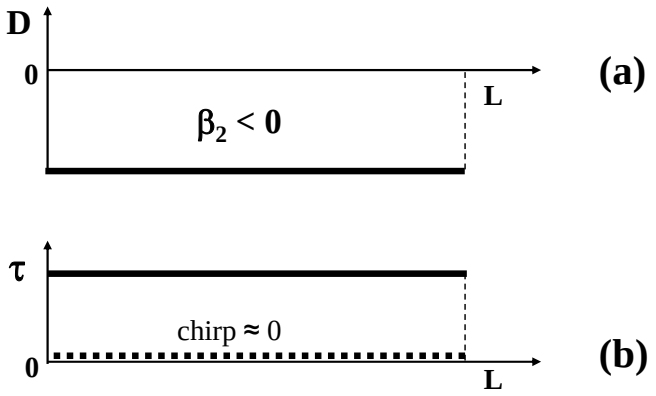


Fig. 11. Dispersion map (a) , and pulse duration (full line) and chirp (broken line) (b) for a soliton fiber laser.

The main limitation for the pulse shortening and increasing of the pulse energy (that together increase the intensity I of the pulse) arises from nonlinear phase shift $\Phi_{NL} = \omega c^{-1} \int_0^L n_2 \, I \, dz$. If the nonlinear phase variation across the pulse is too large, it cannot be controlled by the dispersion. The pulse cannot reproduce itself within a cavity roundtrip and the laser output will fluctuate. The excess nonlinear phase may lead to strong distortion and irreversible breaking of the pulse. The soliton pulse can tolerate only small amount of nonlinear phase shift, $\Phi_{NL} << \pi$. Such disadvantages of the soliton fiber lasers can be strongly reduced applying a dispersion control on the pulse propagation, as in the dispersion managed soliton lasers.

6.2. Dispersion managed soliton regime

The dispersion managed soliton (DMS) fiber laser is shown schematically in Fig. 12. It consists of sections of positive and negative dispersion. The section of positive dispersion includes a doped gain fiber, while that one of negative dispersion includes a standard passive SMF. The dispersion map, evolution of the pulse duration and chirp along the cavity are shown in Fig. 13. The total cavity dispersion of the DMS fiber lasers may vary from small negative to small positive value. The following dispersion range is typical for the DMS regime, $D = (-0{,}01 \div +0{,}01)\ ps^2$. The DMS is periodically stretched and compressed during the propagation along the cavity that is also accompanied by inversion of the chirp, Fig.12, 13. The average pulse duration over a cavity round trip substantially increases in that propagation manner, which results in reduced nonlinear effects. Thus, the DMS fiber laser can tolerate much larger

nonlinear phase shifts than the soliton laser, *e.g.*, $\Phi_{NL} \approx \pi$. It allows decreasing of the pulse duration in parallel with the increasing of the pulse energy. As a result, $sub-100\ fs$ pulses in the nJ -energy range can be generated in this operation regime [7, 8].

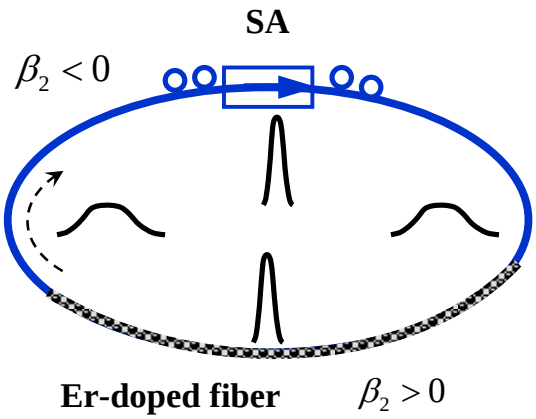


Fig. 12. Schematic presentation of a dispersion managed soliton ring fiber laser cavity

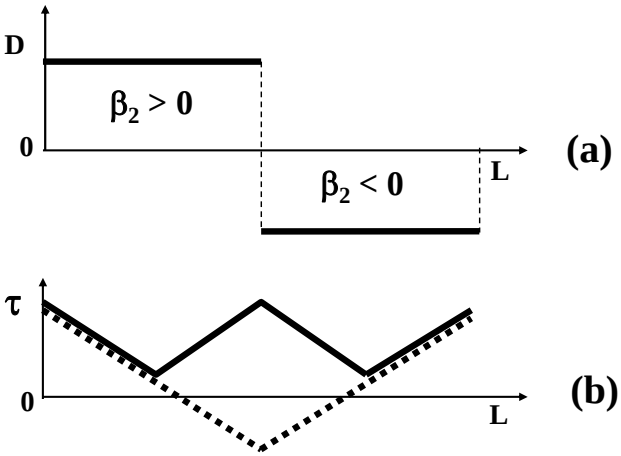


Fig. 13. Dispersion map (a) , and pulse duration (full line) and chirp (broken line) (b) for a dispersion managed soliton fiber laser

The next two regimes are referred to the spectral range and fiber lasers that operate in the positive group velocity dispersion domain of the cavity fibers.

6.3. Self-similar (similariton) regime

The generation of ultrashort laser pulses in positive GVD medium, *i.e.*, for the case of ytterbium doped fiber lasers, is a much more challenging problem. The self-similar (SS) regime [9, 10] seems to be one of the promising solutions for wave-breaking free operation of such fiber lasers and amplifiers. They can generate chirped energetic pulses that can be dechirped down to the $100\,fs$ range. The SS regime is based on the prediction that wave-breaking in positive GVD medium can be avoided if the pulse acquires a monotonic chirp during the propagation [21]. It takes place if the pulse has parabolic intensity and phase profiles

$$A(t,z) = A_0\sqrt{1-(t/\tau(z))^2}\,\exp(ib(z)t^2). \quad (6)$$

Such a pulse preserves its intensity and phase profile during the propagation, although its amplitude and time duration may change. Pulses having such properties are called self-similar pulses, or similaritons. The evolution of a similariton along a passive and a gain fiber is shown in Fig. 14 (a) and (b), respectively. A particular realization of a SS fiber oscillator is shown schematically in Fig. 15. The laser cavity consists of a long SMF of positive GVD, a gain fiber of positive GVD, a SA (real or virtual), and a dispersive delay line (DDL) of negative GVD. The total cavity dispersion $D = \sum_i \beta_{2i} l_i$ of the SS laser is near zero or slightly positive. The pulse acquires large positive linear chirp and strongly extends in time during the propagation along the passive SMF. It reduces the peak intensity and prevents the pulse from breaking. As a result, the SS fiber lasers can tolerate much stronger nonlinear phase, $\Phi_{NL} >> \pi$. The acquired chirp is removed in the DDL, which compresses the pulse down to its initial time duration, Fig. 16, although the pulse is far from transform-limited. The DDL can be formed by diffraction gratings (G_1-G_2-G_3-G_4, as in Fig. 15), chirped fiber Bragg gratings, etc.

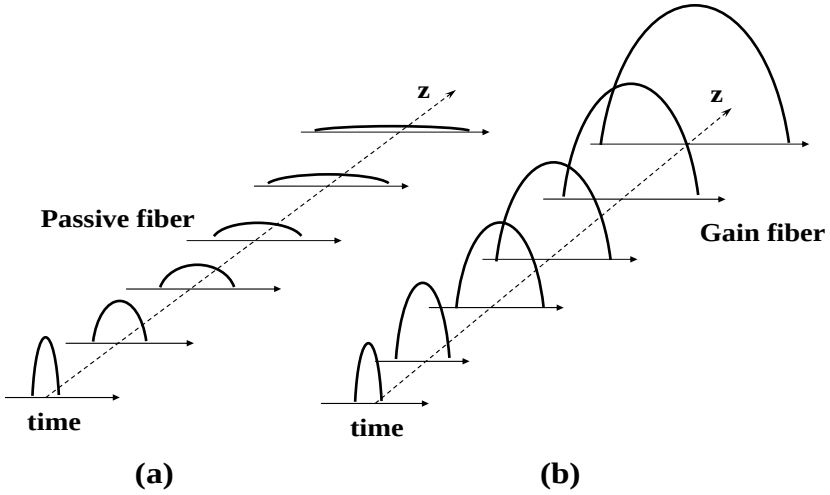


Fig. 14. Similariton propagation along a passive (a) and a gain fiber (b)

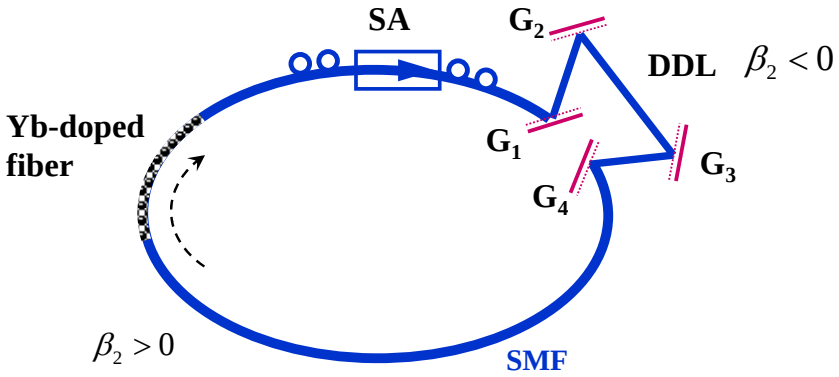


Fig. 15. Schematic presentation of a self-similar ring fiber laser

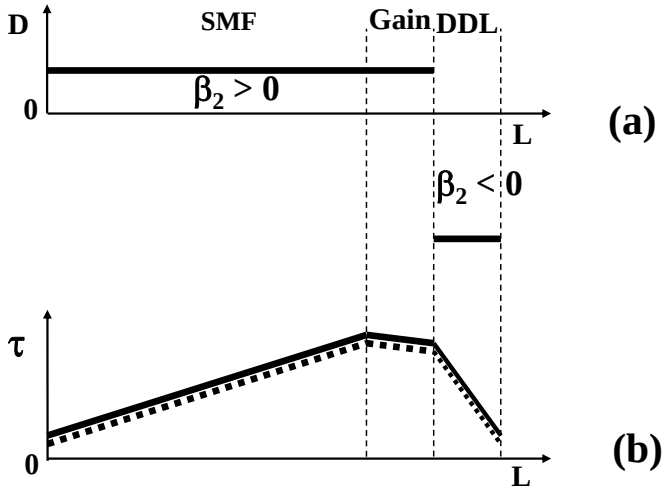


Fig. 16. Dispersion map (a) , and pulse duration (full line) and chirp (broken line) (b) for a self-similar fiber laser

6.4. All normal dispersion (ANDi) regime

Generation of ultrashort laser pulses in positive GVD medium can be obtained even without any negative dispersion inside the cavity [11, 12]. Such lasers are referred to as all-normal-dispersion (ANDi) fiber lasers. The ANDi lasers exploit bandpass spectral filtering to produce self-amplitude modulation in order to stabilize the pulse. In the ANDi lasers, dissipative processes as gain and loss, SA, and, especially spectral filtering play crucial role in the pulse formation rather than the interplay between the dispersion and the nonlinearity.

The ANDi laser cavity is shown schematically in Fig. 17. It consists of a long section of passive SMF, followed by a segment of ytterbium doped gain fiber, a short section of passive SMF, a spectral bandpass filter (BPF) and a SA. The dispersion map of the ANDi laser and the evolution of the pulse duration and the frequency chirp are shown in Fig. 18 (a) and (b), respectively. The operation of the BPF in the pulse formation mechanism can be explained in the following way. The ANDi lasers usually have large positive cavity dispersion, from about $+0.1\,ps^2$ up to about $+0.5\,ps^2$. It results in strongly-chirped pulses. The far frequency components are situated in the far temporal wings of the leading and the trailing edge of the pulse. The spectral filter cults off the frequency components outside the transmittance band of the BPF, which, in fact, cuts off the temporal wings of the pulse, Fig. 19. In this way, the BPF acts as an effective SA and causes an additional amplitude modulation. The pulse formation mechanism in the ANDi fiber lasers allows tolerating strong nonlinear phase, $\Phi_{NL} >> \pi$. The generated pulses from the ANDi fiber oscillators are highly-chirped and usually reach few ps in time duration. They, however, can be dechirped in an external pulse compressor down to $100\,fs$ range. The pulse energy of the ANDi oscillators may reach $10\,nJ$ level.

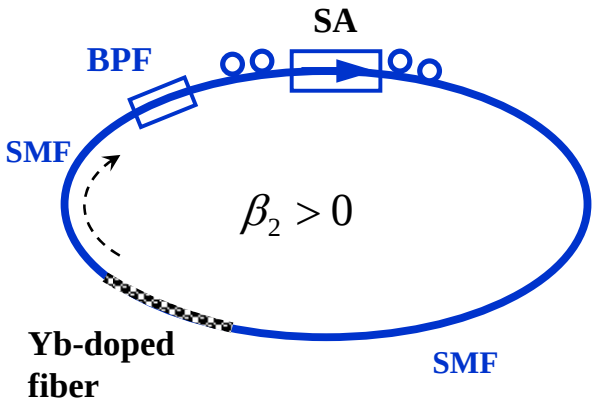


Fig. 17. Schematic presentation of an ANDi ring fiber laser

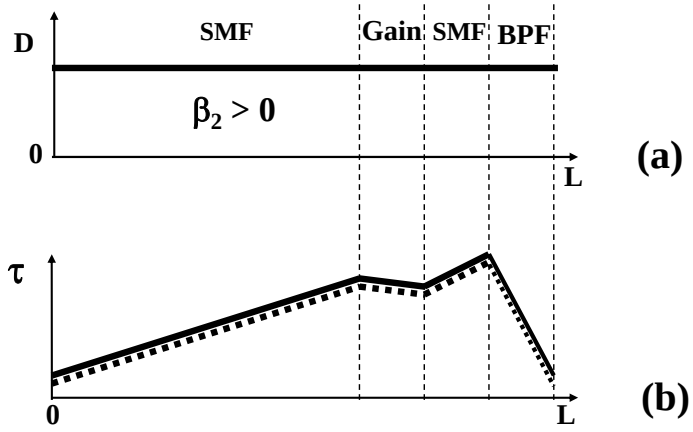


Fig. 18. Dispersion map (a) , and pulse duration (full line) and chirp (broken line) (b) for an ANDi fiber laser

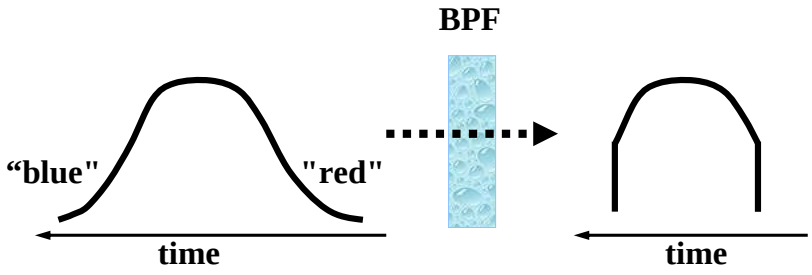


Fig. 19. Pulse formation mechanism of the bandpass filter

7. CONCLUSIONS

The main advantages of the optical fibers as a propagation and gain medium are summarized. The absorption and the emission characteristics of erbium and ytterbium active media are presented. The description of the pulse propagation along passive

and active optical fibers is given. The most widely applicable mode-locked methods are considered. The main operation regimes of the femtosecond fiber lasers: soliton regime, dispersion managed soliton regime, self-similar regime and all normal dispersion regime, are presented. The dispersion profile of the laser cavity and the pulse behavior at the various mode-locking regimes are shown.

REFERENCES

[1] G. Krauss, S. Lohss, T. Hanke, A. Sell, S. Eggert, R. Huber, A. Leitenstorfer, Nature Photonics **4**, 33 (2010)

[2] T. Eidam et al., Advanced Solid State Photonics conference, Feb. 3, 2010, paper AWA2

[3] T. Eidam et al., IEEE J. Sel. Topics Quantum Electronics, **15**, 187 (2009)

[4] J. D. Kafka, T. Baer, Opt. Lett. **14**, 1269 (1989)

[5] I. N. Duling III, Opt. Lett. **16**, 539 (1991)

[6] D. J. Richardson, R. I. Laming, D. N. Payne, M. W. Phillips, V. J. Matsas, Electron. Lett. **27**, 730 (1991).

[7] K. Tamura, E. P. Ippen, H. A. Haus, L. E. Nelson, Opt. Lett. **18**, 1080 (1993)

[8] K. Tamura, L. E. Nelson, H. A. Haus, E. P. Ippen, Appl. Phys. Lett., **64**, 149 (1994).

[9] M. E. Fermann, V. I. Kruglov, B. C. Thomsen, J. M. Dudley, J. D. Harvey, Phys. Rev. Lett., **84**, 6010 (2000)

[10] F. Ö. Ilday, J. R. Buckley, W.G. Clark, F.W.Wise, Phys. Rev. Lett. 94, 213902 (2004)

[11] A. Chong, J. Buckley, W. Renninger, F. Wise, Opt. Express **14**, 10095 (2006)

[12] A. Chong, W. H. Renninger, F. W. Wise, Opt. Lett. **32**, 2408 (2007)

[13] G. P. Agrawal, Nonlinear Fiber Optics, Academic Press (2001)

[14] F. W. Wise, A. Chong, W. H. Renninger, Laser&Photon. Rev., **2**, 58 (2008)

[15] W. H. Renninger, A. Chong, F. W. Wise, Phys. Rev. A 77, 023814 (2008)]

[16] S. M. Kelly, Electron. Lett. **28**, 806 (1992)

[17] N. J. Smith, K. J. Blow, I. Andonovic, IEEE J. Lightwave Technol. **10**, 1329 (1992)

[18] J. P. Gordon, J. Opt. Soc. Am. **B 9**, 91 (1992)

[19] J. Chen, J. W. Sickler, E. P. Ippen, F. X. Kärtner, Opt. Lett., 32, 1566 (2007).

[20] I. G. Koprinkov, K. A. Stankov, Advanced Aspects of Theoretical Electrical Engineering Sozopol 2010, (to be published).

[21] D. Anderson, M. Desaix, M. Karlsson, M. Lisak, M. L. Quiroga-Teixeiro, J. Opt. Soc. Am. **B 10**, 1185 (1993).

A GENETIC ALGORITHM FOR INVERSE EIT PROBLEMS

Vassil Georgiev Guliashki

Decision Support Systems, Institute of Information and Communication Technologies - BAS,
"Acad. G. Bonchev" str. Bl. 2, 1113 Sofia, Bulgaria, phone: (+359 2) 8700 391,
e-mail: vggul@yahoo.com

Abstract. *A hybrid sequential genetic algorithm, named EITIPGA, is proposed. It belongs to the interior path algorithms and is designed to solve the inverse nonlinear problem of Electrical Impedance Tomography (EIT) in 2D domain. The algorithm is hybridized including two different direct search procedures with the aim to overcome the slowly convergence of the interior path methods. The performance of EITIPGA is stable using data up to 3% noise level. Finite Element Method is used to solve the forward EIT problem regarding the nodal scalar potentials and current density values. The inverse problem is solved by means of the variational approach.*

Keywords: Electrical Impedance Tomography, Genetic Algorithms

1. INTRODUCTION

The Electrical Impedance Tomography (EIT) is a Technique reconstructing the electrical field image inside a studied object, based on the conductivity distribution in it. Small currents are injected in an inhomogeneous object and low-frequency voltages are measured by means of electrodes on its boundary. Then the interior electrical conductivity of the object is calculated. At the end EIT gets the electric field image of the object. In this way EIT makes non-destructive testing of materials, geophysical explorations, as well as biomedical diagnoses like diagnosis for breast cancer, investigation of chest organs and cerebral hemorrhaging (brain stroke).

Usually, the reconstruction of an EIT image consists of two parts: 1) Solving a *forward problem* – where the scalar potentials (voltages), as well as the current density values inside the object are calculated, given an approximate conductivity distribution, boundary voltages and currents for a known geometry of the studied volume; 2) Solving an *inverse problem* – where an adequate estimation of the conductivity distribution, based on the calculated (known) scalar potentials and current density values is received. The second is extremely sensitive when boundary values are not fully known, measurements are noisy, calculations are subjected to round-off errors, etc. In addition the intrinsic ill-posed nature of the problem is also associated with irregularities on the optimization hyper-surfaces, the so-called pathology of the optimization problem, which can mislead or restraint the convergence of the numerical procedure. Among such pathological irregularities, which generally surround the global minimum, are multiple local minima.

The inverse problem is a nonlinear ill-posed problem (see [21, 23]), since it involves the inversion of one nonlinear and compact operator, mapping the conductiv-

ity σ of the studied object into the Dirichlet-to-Neumann map. It is known, that the inversion of a compact operator is an ill-posed problem, i.e. it either does not admit an unique solution or its solution does not depend continuously upon the data. Even an operator is invertible, the inverse of a compact operator defined onto an infinite-dimensional space is not continuous. *One way* to overcome this critical feature is to use a regularization method (see [6, 7]). The regularization scheme uses a family of continuous operators approximating the inverse of the operator, which is to be inverted. *Another way* to overcome the lack of continuous dependence of the solution upon the data is to use the quasi-solution, obtained by minimization of an error functional $\boldsymbol{F}(\sigma)$, depending on the measured data. The introduction of the error functional $\boldsymbol{F}(\sigma)$ arises naturally, because usually there is available a contamination of the data by noise. The problem of computing the quasi-solution is well posed under suitable conditions, but the minimization of $\boldsymbol{F}(\sigma)$ is difficult, because there are local minima of $\boldsymbol{F}(\sigma)$ as a result of the nonlinearity of the compact operator and because of the above mentioned valleys or plateaus (regions, where the error functional is almost flat), resulting from the ill-posedness of the original problem. Here is chosen the second way and the inverse EIT problem is converted to an optimization problem.

We use the Finite Element Method (FEM) (see [19, 22]) to solve the forward problem. To solve the inverse problem by means of an exact method is not realistic taking into account the ill-posedness of the problem and the fact that the complexity of the exact optimization methods for such problems grows exponentially with the number of the unknown parameters. On the one hand the calculation of first and second order derivatives of the objective function is a slow procedure for such large size problems. On the other hand the noise in the data and the availability of plateaus of objective function values leads to very slow convergence to the optimal solution. To overcome these shortcomings many hybrid algorithms have been proposed (see for example [1, 2, 8, 12, 13, 14, 17, 18, 21]). Usually they apply a technique, which does not require local gradient information (like genetic algorithms) and hybridize it by means of a direct search or gradient-type method.

For optimization problems with smooth unimodal objective function there are well developed exact optimization methods, which have good convergence speed, and which use information about the objective function derivatives (first order, or first and second order) to calculate the search direction at the current iteration. For problems with complex, non smooth and multimodal objective function, where the information of objective function derivatives is not accessible in an analytic form and/or the gradient evaluation by means of finite differences is impossible due to the large size of the problem have been developed global heuristic algorithms, as well as evolutionary (population based) algorithms, such like Genetic algorithms (see [10]), Scatter search, Ant systems and Particle swarm optimization algorithms. Their main common features are as follows:

- They memorize solutions (or characteristics of solutions) in a population of individuals. Each individual is associated with a feasible solution of the problem at hand.

- They include a generating solutions search procedure, which uses the information (implicitly) stored in the population.

In this paper is proposed a genetic algorithm. To overcome the problem with the plateaus, two different direct search procedures are included in the algorithm scheme.

2. EXPERIMENTAL SETUP

We consider 2D-objects. The data measurement is performed by means of 16 electrodes located on the boundary of the investigated object.

The illustration of the experimental setup for the EIT problem is shown in Fig. 1.

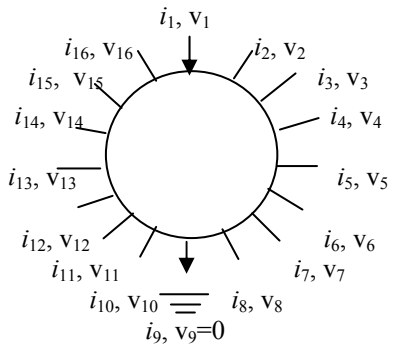


Fig. 1. Experimental setup for the EIT problem

The studied object is considered as an inhomogeneous, conducting body having a known overall shape Ω. For simplicity here is chosen the domain Ω to be a circle. It is divided by an uniform triangularization into 64 triangles. This mesh is assumed to be fine enough, so that the FEM numerical calculations are relatively accurate. The application of mesh with 256 or 576 triangles is also possible. Direct currents i1 (input current) and i9 (output current) are applied to the object. The injected current between these two electrodes has value about 20 mA. The potentials (voltages) are measured between pairs of electrodes, where one of the electrodes in each pair is the grounded electrode. Usually the voltage at the injection electrode cannot be measured reliably and for this reason it is not included in the data set. The measured voltages have values about 1 V. Each electrode can be held to be equi-potential and the contact impedance is neglected.

For the conductivity problem there are two distinct variational principles (see for example [11, 15, 16]): the Dirichlet's principle:

Equation (1) $$\int_{\Omega} \sigma(x) \,|\, \nabla\Phi(x) \,|^2 \; dx \geq \mathrm{P},$$

where P is the power dissipated into heat (the measured power) in the true conductivity medium Ω, and its dual – the Thompson's variational principle, which takes the form:

Equation (2) $$\int_{\Omega} \sigma^{-1}(x) \mid J(x) \mid^2 \, dx \geq \mathrm{P}.$$

These two constraints allow us to obtain upper and lower bounds on the feasible domain of the space that contains the solution to the inverse problem (for details see [3, 4]).

3. DIRECT PROBLEM FORMULATION

The direct EIT problem can be decomposed as two quadratic optimization problems: The first one has the form:

Equation (3) $$\text{Min} \int_{\Omega} \sigma(x) \mid \nabla \Phi(x) \mid^2 \, dx,$$

subject to:

Equation (4) $$\Phi(x) = V(x), \text{ for } x \in \partial\Omega,$$

where V(x), $x \in \partial\Omega$, are the measured potentials on the boundary of the specimen, the voltages V(x), $x \in \Omega$ are unknown, and the conductivities σ(x) are known positive real numbers, assuming that the studied object is locally isotropic.

The second optimization problem has the form:

Equation (5) $$\text{Min} \int_{\Omega} \frac{1}{\sigma(x)} \mid J(x) \mid^2 \, dx,$$

subject to:

Equation (6) $$-J(x).n(x) = I(x), \text{ for } x \in \partial\Omega,$$

Equation (7) $$\int_{\partial\Omega} I(x)dx = 0,$$

Equation (8) $$\nabla \cdot J(x) = 0, \text{ for } x \in \Omega,$$

where I(x) are the currents on the boundary $\partial\Omega$ and n(x) is the unit outward normal to the boundary $\partial\Omega$.

Assuming that there are N current injection experiments, the power dissipated into heat by Ω is:

Equation (9) $$\mathrm{P} = \sum_{i=1}^{N} \int_{\partial\Omega} \Phi_i(x) J_i(x) n(x) dx.$$

The current density J(x) can be expressed by means of the electric vector potential T(x) (see [23]):

Equation (10) $$J(x) = \nabla \times T(x).$$

Hence the second optimization problem can be written in the form:

Equation (11) $$\text{Min} \int_{\Omega} \frac{1}{\sigma(x)} |\nabla \times T(x)|^2 \, dx,$$

subject to:

Equation (12) $$-(\nabla \times T(x)).n(x) = I(\text{x}), \text{ for } x \in \partial\Omega,$$

Equation (13) $$\int_{\partial\Omega} I(x)dx = 0,$$

Beginning with initial approximate values for $\sigma(x)$, $x \in \Omega$, we solve the optimization problems (3)-(4) and (10)-(12) by means of the Finite Element Method and calculate $\Phi(x)$, $T(x)$ and $J(x)$, $x \in \Omega$. In case $\sigma(x)$ is known, $\Phi(x)$ and $J(x)$ are completely determined correspondingly by the boundary voltage $\Phi|_{\partial\Omega}$ and by the boundary current flux $J.n|_{\partial\Omega}$.

4. INVERSE PROBLEM FORMULATION

We apply the variational approach described in [15, 16] in order to overcome the noise influence during the measurement. The measurement procedure is repeated N times (2N is the number of electrodes) clockwise, injecting current between N distinct pairs of diametraly opposed electrodes (the remaining N pairs are symmetric to the first N pairs). Using data from N different measurements each time with different current injection pair of electrodes we solve N times the quadratic optimization problems (3)-(4) and (10)-(12). Using the calculated $\nabla\Phi(\text{x})$ and J(x) we solve the inverse EIT problem with unknowns $\sigma(\text{x})$, $\text{x} \in \Omega$, minimizing the error functional G, similar to that one proposed in [16]:

Equation (14) $$\text{Min } G = \sum_{i=1}^{N} \left[\int_{\Omega} \sigma(x) \, | \, \nabla\Phi_i(x) \, |^2 \, dx + \int_{\Omega} \frac{1}{\sigma(x)} \, | \, J_i(x) \, |^2 \, dx \right]$$

subject to:

Equation (15) $$\Phi_i(x) = V_i(\text{x}), \; -J_i(x).n(x) = I_i(\text{x}), \; i = 1, \ldots, N;$$

$$\nabla \cdot J_j(x) = 0, \;\; j = 1, \ldots, m; \;\; \sigma_j(x) \geq 0.$$

where m is the number of finite elements in the used mesh.

Minimization of the first term in (14) corresponds to the Dirichlet's variational principle and the minimization of the second term corresponds to the Thompson's variational principle.

5. THE PROPOSED HYBRID GENETIC ALGORITHM EITIPGA

There are known two types of optimization methods: *exterior path* and *interior path* methods. The EIT inverse problem could be solved by means of an exterior path method or by an interior path method. During the search of optimal solution both types of methods attempt to converge to a solution on the boundary of the feasible domain, but the exterior path methods converge from outside the feasible domain, while the interior path methods converge from inside the feasible domain. As pointed in [3], the exterior path methods can achieve convergence quickly for data without errors. The interior path methods have the advantage to be insensitive to data errors and perform stable, but they are often slowly converging.

We propose a hybrid genetic algorithm, named EITIPGA, which belongs to the interior path algorithms. It is designed to solve the inverse EIT problem (14)-(15) overcoming the slowly converging of the interior path methods and the instability of the exterior path methods when the signal/noise ratio is greater than 1%. The algorithm EITIPGA combines the traditional genetic search scheme by two different direct search procedures. To reconstruct the electrical field image it performs iteratively the following procedure:

1) Using the last computed $\sigma(x)$ and the measured voltages, minimize (3) and (10) over $\Phi_i(x)$ and $T_i(x)$ for $i = 1,\ldots,m$.

2) Using the calculated $\Phi_i(x)$ and $J_i(x)$ minimize $\boldsymbol{G}$ from (13) over $\sigma(x)$, and update $\sigma(x)$.

It could be expected, that after a given number of k genetic generations the algorithm will reach a plateau of solutions with equal $\boldsymbol{G}$-values. This may lead to very slow convergence of genetic search procedure. The main idea of our algorithm is the integration of two direct search procedures in it, which are performed consecutively after the performed k generations (or when a plateau of solutions is reached) in order to enhance the search process and to make the algorithm more efficacy by the first one. Then, after an approximation of the global optimal solution is localized, the second direct search procedure serves to find out precisely the exact optimal solution.

The algorithm EITIPGA includes two selection strategies (Roulette wheel and Elite selection) and two kinds of crossover (2-point and blend crossover), giving in this way possibility for applying different strategies. The "Elite selection" strategy uses E% of the best individuals (solutions) among the P individuals of the current population (we use E=20%). These individuals mate each with all others. This selection is not useful, when plateau(s) of individuals with equal $\boldsymbol{G}$-values are available, because the crossover is unable to generate individuals outside the plateau. For this reason it could be used at the initial stage of the search process.

The blend crossover (BLX) operator (see [9]) produces one offspring by generating, for each variable x_j a random value uniformly distributed in the range $[l_j - \delta, u_j + \delta\ ]$ with $\delta = \alpha(u_j - l_j)$, where l_j and u_j are the smallest and the largest values respectively for x_j in the parents individuals, and α is set to 0.3. After the crossover the obtained better offspring individuals (having lower fitness value) replace the worst indi-

viduals (with higher fitness value) in the old population. The line variant of the blend crossover produces an offspring, which is a linear combination of the parents, but with a single random number used for all components.

During the first direct search procedure a search step is calculated like in the Nelder and Mead simplex method (see [20]). The individuals (solutions) in the current population are ordered in increasing order of their ***G***-values. The weight center σ_c of the first five individuals is calculated. Let the last individuals with the worst ***G***-values be denoted by σ_{wi} , i=1,2,…t; Each σ_{wi} is reflected towards σ_c making a step $y = \sigma_c - \sigma_{wi}$ to create a new solution. The constraints $\sigma(x) \geq 0$, (1) and (2) are used to reduce the length of this step if it is necessary. The procedure starts by the last individual in the ordered set and continues until there is achieved an improvement of the error functional value. In case someone of the so generated solutions is better than one of the current population, the better solution replaces the worse. Each successful step leading to better solution is repeated once, taking into account (1), (2) and $\sigma(x) \geq 0$. In case the step y is not successful it is reduced in half. Ending the list of individuals to be reflected, the procedure is repeated once again. We chose this kind of acceleration, because: (i) the Nelder and Mead simplex method is widely known as one of the most successful methods for optimization of non smooth functions, (ii) it does not require derivatives and (iii) it is easy to be implemented.

During the second direct search procedure calculation of search steps is organized like in Powell's method. Good results have previously been reported for this method (see for example [1, 14]. In this procedure each iteration completes after a sequence of m line searches is performed which take as the starting point the minimum achieved in the previous step. These directions coincide with the coordinate axes of the optimization problem. After each iteration the initial and the final points define a composite search direction. The procedure finishes if there is not achieved improvement of the error functional ***G***. The search process of Powell's method for an optimization problem having two variables is illustrated on Fig. 2.

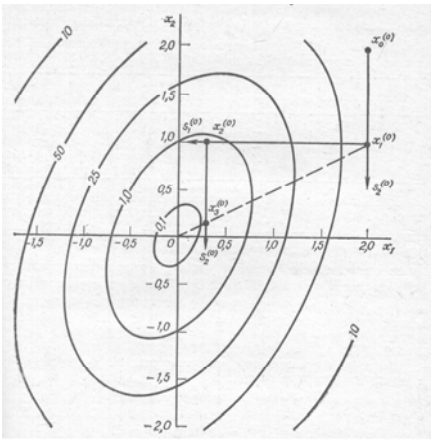


Fig. 2. The Powell's search method

Both direct search procedures are performed only once.

Two termination criteria of genetic procedure are used: a) *limit of generations*; and b) *limit of consecutive generations without essential improving the best-obtained solution*;

The pseudo-code of algorithm EITIPGA proposed here is presented as follows:

Begin
Initialize the population P
Generation
Evaluate the individuals in P
Sort P according to the fitness value
Until stopping criteria have been met,
Repeat
Choose a selection operator.
Apply the chosen selection operator.
Choose a crossover operator
Apply the crossover operator
Select individual x_j to survive
If x_j is better than one individual in P
Replace the old individual in P by x_j.
The check is performed starting by the worst individual.
Endif
End of Generation
Perform Direct search procedure 1.
Perform Direct search procedure 2.
End

To perform test experiments a MATLAB programming code is created and simulated input data are used. Initial conductivity is given for each mesh element. Known data on conductivity of tissues (see [11]) are used to create inhomogenities in the object: 0.006 S/m for bone, 0.06 S/m for muscle (transverse), 0,8 S/m for muscle (longitudinal), 0.1 for lung (expiration), 0.28 S/m for liver, 0.67 S/m for blood and 1.54 S/m for cerebral spinal fluid. Initially areas with different conductivity value are determined. Then on the base of known values for the input and output current the forward problem is solved to obtain the potentials along the boundary. In order to get more realistic results the influence of data noise is considered. Experiments are performed using data contaminated with noise level of 1%, as well as of 3%.

6. CONCLUSIONS AND DIRECTIONS FOR FURTHER RESEARCH

The test results show that the proposed algorithm EITIPGA is better than that one in [12]. Problems with up to 3% noise level in the data and 3 or 4 areas with different conductivity can be solved successfully. The algorithm is relatively insensitive to such levels of noise. The computational efforts to obtain a good reconstructed field image are considerable. In case the performance of the algorithm is slow, especially when finer mesh is used, the limit number k of genetic generations could be reduced and the second direct search procedure could be skipped.

Better efficiency could be achieved by means of a parallel genetic algorithm, like that one proposed in [5].

Another direction for future research is the use of preliminary information about the shape forms of regions in the experimental object having different conductivities. Using as variables shape parameters of studied domain, the number of variables in the problem could be drastically reduced (see [1, 2, 5]).

ACKNOWLEDGEMENT

This study is supported by the Bulgarian National Science Fund, Grant No DTK02/71, "Web-Based Interactive System, Supporting the Building Models and Solving Optimization and Decision Making Problems" and by the contract of the Institute of Information Technologies – BAS, 2010-2012, Code: 010095, "Research, development and application of efficient methods for hard-solvable single- and multi-objective problems".

REFERENCES

[1] Barra, L.P.S., Peters, F. C., Martins, C. P., Barbosa, H. J. C.; „Computational Experiments in Electrical Impedance Tomography"; XXVII Iberian Latin American Congress on Computational Methods in Engineering; Belém/PA; Brazil, September 03-06, 2006; Proceedings of the XXVII Iberian Latin American Congress on Computational Methods in Engineering. http://www.numec.ufjf.br/luisp/docs/papers/06_cilamce_tomo.pdf

[2] Barra, L.P.S., Santos, R.W., Peters, F.C., Santos, E.P., Barbosa, H.J.C., „Parallel Computational Experiments in Electrical Impedance Tomography", HPC LIFE, Workshop on High Performance Computing in the Life Sciences, Ouro Preto, October 17, 2006; http://www.fisiocomp.ufjf.br/hpclife/papers/paper2.pdf

[3] Berryman, J. G., „Some Nonlinear Reconstruction Algorithms for Electrical Impedance Tomography", LLNL, UCRL-JC-143031, *Geometric methods in inverse problems and PDE control*, Minneapolis, Minnesota, July 16-27, 2001, *Conference Proceedings*, http://www.osti.gov/bridge/product.biblio.jsp?osti_id=15013262

[4] Berryman, J. G. and R. V. Kohn, „Variational Constraints for Electrical Impedance Tomography", *Physical Review Letters* 65, 1990, pp. 325-328.

[5] Carosio G. L.C., Rolnik V., Seleghim P. Jr., „Improving efficiency In electrical impedance tomography problem by hybrid parallel genetic algorithm and a priori information", http://www.sbmac.org.br/eventos/cnmac/xxx_cnmac/PDF/539.pdf

[6] Clay M. T. and Ferree T. C., "Weighted Regularization in Electrical Impedance Tomography with Applications to Acute Cerebral Stroke", IEEE Transactions on Medical Imaging, vol. 21, No. 6, June, 2002, pp. 629-637.

[7] Cohen-Bacrie C., Goussard Y. and Guardo R., "Regularized Reconstruction in Electrical Impedance Tomography Using a Variance Uniformization Constraint", IEEE Transactions on Medical Imaging, vol. 16, No. 5, October 1997, pp. 562-571.

[8] Eckel H., „Numerical Study of an Evolutionary Algorithm for Electrical Impedance Tomography", Ph.D. thesis, Göttingen, Germany, 2007. http://webdoc.sub.gwdg.de/diss/2008/eckel/eckel.pdf

[9] Eshelman, L. & Schaffer, J. (1993) "Real coded genetic algorithms and interval schemata". In Whitley, D., ed., Foundations of Genetic Algorithms 2, Morgan Kaufmann, San Mateo, CA, pp. 187-202.

[10] Goldberg D. E. Genetic Algorithms in Search, Optimization and Machine Learning, Addison Wesley, Reading, Mass, 1989.

[11] Gray, G. A., (2002) "A variational study of the Electrical Impedance Tomography problem", Ph.D. Thesis, Rice University, Houston, Texas.

[12] Guliashki V., (2008) „A Quickened Genetic Nonlinear Reconstruction Algorithm for EIT", Proceedings of Papers of XLIII International Scientific Conference on Information, Communication and Energy Systems and Technologies ICEST2008, ISBN: 978-86-85195-60-0, Nis, Serbia, 25-27 June, 2008, Vol. 2, pp. 433-436.

[13] Guliashki V., H. Linkov, (2009) "An accelerated genetic reconstruction algorithm for inverse EIT problems", XV International Symposium on Theoretical Electrical Engineering ISTET 2009, 22-24 June 2009, Lübeck, Germany, ISSN: 0932-6022, pp. 379-382.

[14] Hsiao C., Chahine G. and Gumerov N., „Application of a Hybrid Genetic / Powell Algorithm and a Boundary Element Method to Electrical Impedance Tomography", *Journal of Computational Physics* 173, 2001, pp. 433-454.

[15] Kohn, R. V. and McKenney A., „Numerical Implementation of a Variational Method for Electrical Impedance Tomography", *Inverse Problems*, vol. 6, 1990, pp. 389-414.

[16] Kohn, R. V. and Vogelius M., „Relaxation of a Variational Method for Impedance Computed Tomography", *Communications on Pure and Applied Mathematics* 40, 1987, pp. 745-777.

[17] Marinova G. and V. Gouljashki, "2D Eddy Current Imaging for Steam Generator Tubing Based on Spline Interpolation Mapping and Improved Statistical Filtering", *Proceedings of the 1st Japanese-Bulgarian-Macedonian Joint Seminar on Applied Electromagnetics*, Sofia, BULGARIA, September, 14.-15., 1998, JBMSAEM'98, pp. 145-150.

[18] Olmi, R., Bini, M. and Priori, S., "A genetic algorithm approach to image reconstruction in electrical impedance tomography", IEEE Transactions on Evolutionary Computation, Volume 4, Issue 1, April 2000, pp.:83 – 88.

[19] Papasov, S. P. and Farhy S. L., *Theoretical electro technique. Part* 3, 1988, (Sofia, Technika)

[20] Reklaitis G. V., Ravindran A., Ragsdell K. M., *Engineering Optimization. Methods and Applications*, John Wiley and Sons, 1983.

[21] Rolnik V.P, Seleghim P.Jr., „A Specialized Genetic Algorithm for Electrical Impedance Tomography of Two-Phase Flows", *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. XVIII, № 4, October – Dezember 2006, pp. 378-389.

[22] Strang G. and Fix G., *An Analysis of the Finite Element Method*, 1973 (Englewood Cliffs, NJ: Prentice-Hall)

[23] Tamburino A., S. Ventre and G. Rubinacci, "Electrical Resistance Tomography: Complementarity and Quadratic Models", *Inverse Problems*, vol. 16, 2000, pp. 1585-1618.

ГРАНИЧНИ ЦИКЛИ И ТЕОРИЯ НА МЕЛНИКОВ

Живко Димитров Георгиев

Катедра „Теоретична електротехника”, Технически университет - София,
София 1000, България, E-mail: zhdgeorg@tu-sofia.bg

Резюме: *Понятието граничен цикъл е основно понятие в теорията на нелинейните динамични системи. Различни автоколебателни процеси наблюдавни в технически, физически, химически, биологически и икономически ситеми представляват в действителност гранични цикли.*

В настоящата статия се излагат някои основни положения свързани с анализа на граничните цикли. Дадени са определение и класификация на граничните цикли. Изложени са основните принципи от теорията на Мелников, която се използува за анализ на гранични цикли.

Ключови думи: гранични цикли, функция на Мелников

1. ВЪВЕДЕНИЕ

Периодичните процеси в нелинейните динамични системи намират широки инженерни приложения. Голяма част от тези периодични процеси представляват автоколебателни процеси. Такива автоколебателни процеси възникват при определени условия в различни технически, физически, химически, биологически и икономоически системи. От гледна точка на качествената теория на динамичните системи и диференциалните уравнения автоколебателните процеси представляват гранични цикли във фазовото пространство [10].

Понятието граничен цикъл е въведено от Поанкаре [3], [4], и е едно от основните понятия в теорията на нелинейните динамични системи. Граничният цикъл е математически феномен, който се наблюдава само при нелинейните диференциални уравнения. От друга страна, както беше подчертано по-горе, граничните цикли (или съответствуващите им автоколебателни процеси) представляват и реално съществуващи природни феномени, които се наблюдават много често в различни области. Всичко това, както от теоретична така и от инженерно-приложна гледна точка, налага познаването, изучаването и анализирането на граничните цикли. При това положение изследването на автоколебателните процеси, които възникват в различните по природа системи се свежда до изследване на съответстващите им гранични цикли.

В настоящата статия се излагат някои основни положения и понятия свързани с теорията на граничните цикли. Дава се дефиниция на понятието граничен цикъл и се прави класификация на граничните цикли. Изложени са основните принципи от теорията на Мелников, която се използува за анализ на гранични цикли. Съгласно теорията на Мелников анализът на граничните цикли се свежда до намирането на нулите на функцията на Мелников.

2. ГРАНИЧНИ ЦИКЛИ

2.1. Общо определение на гранични цикли

В настоящата статия ще разглеждаме гранични цикли в планарни (равнинни) системи диференциални уравнения. Нека е дадена системата

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \ , \tag{1}$$

където $\mathbf{x} \in \mathbf{R}^2$, $E \subset \mathbf{R}^2$, E - отворено множество, а $\mathbf{f} \in C^1(E)$. Да означим с $\mathbf{x} = \gamma(t) = \phi_t(\mathbf{x}_0)$ решението на планарната система (1) с начална точка $\mathbf{x}_0$ и нека Γ е съответстващата на това решение фазова траектория.

Нека освен това Γ е периодична фазова траектория на уравнението (1) в $\mathbf{R}^2$. Да означим с $Ext(\Gamma)$ множеството от точки, които принадлежат на неограничента компонента на $\mathbf{R}^2 \setminus \Gamma$ (външността на Γ), а с $Int(\Gamma)$ множеството от точки, които принадлежат на ограничената компонента на $\mathbf{R}^2 \setminus \Gamma$ (вътрешността на Γ).

В литературата съществуват различни определения на понятието граничен цикъл. За нашите цели е удобно следното определение.

Определение: Периодичната фазова траектория Γ се нарича ***граничен цикъл***, ако съществува околност U на Γ, такава че Γ е единствената периодична траектория в околността U. Може да се каже още, че ***граничният цикъл Γ е изолирана периодична фазова траектория.***

Ще отбележим, че не всяка затворена фазова траектория е граничен цикъл. Така например, която и да е затворена фазова траектория от семейство затворени траектории около особена точка тип „център”, не е граничен цикъл, тъй-като такава траектория не е изолирана (в каквато и да е околност на такава траектория съществуват и други затворени траектории). Условието затворената траектория да бъде изолирана е съществено условие за да бъде тя граничен цикъл.

2.2. Изображение на Поанкаре и функция на изместването за планарни системи диференциални уравнения

В този раздел ще дадем някои понятия, които ще използуваме при класификацията и анализа на граничните цикли в планарни системи диференциални уравнения. Нека $\mathbf{x} = \gamma(t) = \phi_t(\mathbf{x}_0)$ е периодично решение на планарната система (1) с период T_0 и начална точка $\mathbf{x}_0$, а Γ_0 е затворената фазова траектория, съответствуваща на това периодично решение.

Определение: Кривата Σ се нарича *трансверзална крива*, или *сечение на Поанкаре* (*Poancare section*) за траекторията Γ_0 (или за траекториите на системата (1), или само за системата (1)), ако са изпълнени следните условия:

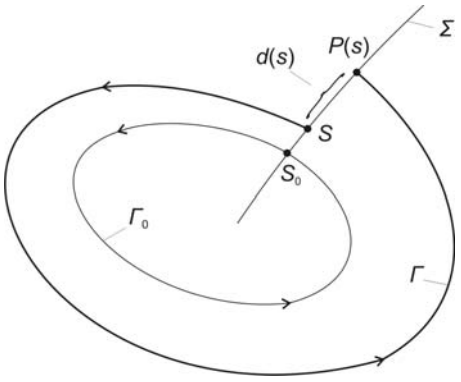
(*i*) Σ е гладка, непресичаща се крива, чийто параметрични уравнения са строго монотонни функции;

(*ii*) Нито една точка от Σ не е особена точка за системата (1);

(*iii*) В която и да е точка от Σ фазовите траектории на системат (1) и кривата Σ не се допират.

Нека Σ е трансверзална крива за (1), минаваща през точката $\mathbf{x}_0$ и нека кривата Σ е параметризирана чрез нарастващия параметър s, като нейното параметрично уравнение е $\mathbf{x}=\mathbf{u}(s)$, $\mathbf{u}=(u_1,u_2)$. В много случаи параметърът s е дължината на дъгата от Σ, но това не е задължително.

Нека на точката $\mathbf{x}_0$, а следователно и на затворената фазова траектория Γ_0, съответствува параметър s_0. Точката s_0 разделя Σ на два сегмента – $\Sigma^+(s_0)$ и $\Sigma^-(s_0)$. За точките от $\Sigma^+(s_0)$ е изпълнено $s \geq s_0$, а за точките от $\Sigma^-(s_0)$ – $s \leq s_0$, т.е. имаме $\Sigma^+(s_0)=\{s \mid s \geq s_0\}$ и $\Sigma^-(s_0)=\{s \mid s \leq s_0\}$. Единият от сегментите $\Sigma^+(s_0)$ и $\Sigma^-(s_0)$ принадлежи на $Ext(\Gamma_0)$, а другият на $Int(\Gamma_0)$.



Фиг. 1. Изображение на Поанкаре и функция на изместването
за планарни системи диференциални уравнения

Да разгледаме решение на системата (1), което има начална точка върху Σ с параметър s и нека Γ е съответната на това решение фазова траектория. Ако точката s е достатъчно близо до s_0, то траекторията Γ на системата (1) с начална точка s ще пресече отново Σ в точка $P(s)$. Изображението

$$s \to P(s) \tag{2}$$

се нарича *изображение*, или *функция на Поанкаре* (*Poincare map*, *Poinacare return map*). За по-кратко функцията $P(s)$ се нарича *функция на Поанкаре*.

Функцията $P(s)$ е скаларна функция на параметъра s и е непрекъсната в някаква околност на s_0. При тази постановка функцията

$$d(s)=P(s)-s \tag{3}$$

се нарича *функция на изместването* (*displacement function*) (фиг. 1). Очевидно, траекторията Γ е затворена, а съответното решение на системата е периодично, само когато е изпълнено условието

$$d(s) = P(s) - s = 0. \tag{4}$$

Функцията на изместването $d(s)$ дава възможност да се изследва характера на траекториите в околност на затворена траектория. Освен това, чрез нея се прави по-фина класификация на граничните цикли в случая, когато системата е аналитична.

Изложението по-нататък се основава на теорията дадена в [1] и [2]. Нека Γ_0 е затворена траектория, която пресича Σ в точка с параметър s_0. Тогава е изпълнено равенството

$$d(s_0) = P(s_0) - s_0 = 0 \ . \tag{5}$$

По отношение на траекториите около Γ_0 а оттук и по отношение на функцията $d(s)$, най-общо могат да съществуват следните два случая:

а) Съществува $\delta > 0$, такова че $d(s) \neq 0$ за всички s, за които е изпълнено $0 < |s - s_0| < \delta$;

б) Каквото и да е $\delta > 0$, съществува s, такова че $0 < |s - s_0| < \delta$ и за което е изпълнено $d(s) = 0$.

Ще разгледаме подробна двата случая.

Случай а): В този случай всички траектории минаващи през точки от кривата Σ, съответстващи на параметъра s, $0 < |s - s_0| < \delta$, са незатворени. Затворената траектория Γ_0 е изолирана, т.е. Γ_0 е *граничен цикъл*. Като се има предвид непрекъснатостта на функцията $d(s)$ се вижда, че са възможни следните четири подслучая:

а1) $d(s) < 0$ при $s \in \Sigma^+(s_0)$ и $d(s) > 0$ при $s \in \Sigma^-(s_0)$;

а2) $d(s) > 0$ при $s \in \Sigma^+(s_0)$ и $d(s) < 0$ при $s \in \Sigma^-(s_0)$;

а3) $d(s) < 0$ при $s \in \Sigma^+(s_0)$ и $d(s) < 0$ при $s \in \Sigma^-(s_0)$;

а4) $d(s) > 0$ при $s \in \Sigma^+(s_0)$ и $d(s) > 0$ при $s \in \Sigma^-(s_0)$;

Подслучай а1): В този случай всички траектории от достатъчно малка околност N на кривата Γ_0 се приближават към Γ_0 при $t \to \infty$ и се отдалечаат от Γ_0 при $t \to -\infty$. Кривата Γ_0 се нарича *устойчив граничен цикъл*. Близките до Γ_0 траектории се навиват спиралообразно към Γ_0.

Подслучай а2): В този случай имаме обратното – всички достатъчно близки до Γ_0 траектории се отдалечават от Γ_0 при $t \to \infty$ и се приближават до Γ_0 при $t \to -\infty$. Кривата Γ_0 се нарича *неустойчив граничен цикъл*.

Подслучай а3): В този случай всички траектории минаващи през точки s, $s \in N \cap \Sigma^+(s_0)$ приближават Γ_0 при $t \to \infty$ и всички траектории минаващи през точки s, $s \in N \cap \Sigma^-(s_0)$ се отдалечават от Γ_0 при $t \to \infty$ (N е достатъчно малка околност на кривата Γ_0).

Подслучай а4): В този случай всички траектории минаващи през точки s, $s \in N \cap \Sigma^{+}(s_0)$ се отдалечават от Γ_0 при $t \to \infty$ и всички траектории минаващи през точки s, $s \in N \cap \Sigma^{-}(s_0)$ приближават Γ_0 при $t \to \infty$.

В случаите **а3)** и **а4)** кривата Γ_0 се нарича *полуустойчив граничен цикъл*.

Случай б): В този случай във всяка околност N на Γ_0 са разположени безкрайно много затворени траектории. Това е случая, когато имаме особена точка тип „център" обградена от непрекъсната фамилия затворени траектории.

Направената дотук класификация на граничните цикли е най-обща. В случая, когато системата (1) е аналитична може да се направи по-фина класификация на граничните цикли. За този именно случай в следващите раздели ще дадем по-подробни и специфични резултати свързани с анализа на гранични цикли.

2.3. Гранични цикли за аналитични планарни системи диференциални уравнения

По-нататък ще предполагаме, че планарната система (1) е аналитична, т.е. функцията $\mathbf{f}(\mathbf{x})$ е аналитична. Нека отново Γ_0 е затворена фазова траектория, а Σ–трансверзална крива за траекторията Γ_0 (или за системата (1)), параметризирана с нарастващия параметър s. Нека траекторията Γ_0 пресича Σ в точка с параметър s_0. Приемаме, че Σ се задава също с аналитични уравнения. При това положение изображението на Поанкаре и функцията на изместването са също аналитични функции. Чрез функцията на изместването ще изследваме траекториите на системата (1) в околността на траекторията Γ_0, а оттук ще направим заключение и за вида на самата траектория Γ_0.

Да развием в ред на Тейлор около точката s_0 функцията на изместването $d(s)$. Като отчетем, че s_0 е параметър от трансверзалната крива Σ, който съответства на затворената траектория Γ_0, получаваме

$$d(s) = \alpha_1 (s - s_0) + \alpha_2 (s - s_0)^2 + \cdots , \tag{6}$$

където

$$\alpha_1 = d'(s_0) = P'(s_0) - 1 , \tag{7а}$$

$$\alpha_n = \frac{1}{n!} d^{(n)}(s_0) = \frac{1}{n!} P^{(n)}(s_0) \ , \ n > 1 . \tag{7б}$$

При анализа на функцията $d(s)$ са възможни следните случаи:

Случай 1: Изпълнени са условията

$$d(s_0) = 0 \ , \tag{8а}$$

$$d'(s_0) = P'(s_0) - 1 \neq 0 \ , \text{ или } P'(s_0) \neq 1 \ . \tag{8б}$$

Тогава очевидно коренът s_0 на функцията $d(s)$ е изолиран, а оттук и затворената траектория Γ_0 е изолирана. В този случай затворената траектория Γ_0 се нарича *прост граничен цикъл*, или *хиперболичен граничен цикъл* (*simple limit cycle*, *hyperbolic limit cycle*). Като се вземе предвид, че в достатъчно малка околност на s_0 знакът на $d(s)$ съвпада със знака на $\alpha_1(s-s_0)=d'(s_0)(s-s_0)$, можем да направим извода, че граничният цикъл Γ_0 е *устойчив*, когато е изпълнено условието

$$d'(s_0)<0 \text{ , или } P'(s_0)<1 \text{ ,} \tag{9а}$$

от което условие следва

$$d(s)<0 \text{ при } s>s_0 \text{ и } d(s)>0 \text{ при } s<s_0 \text{ .} \tag{9б}$$

Неравенствата в (9б) са същите както в случая **а1)** разглеан в раздел 2.2.

Аналогично граничният цикъл Γ_0 е *неустойчив*, когато е изпълнено условието

$$d'(s_0)>0 \text{ , или } P'(s_0)>1 \text{ ,} \tag{10а}$$

от което условие следва

$$d(s)>0 \text{ при } s>s_0 \text{ и } d(s)<0 \text{ при } s<s_0 \text{ .} \tag{10б}$$

Неравенствата в (10б) са същите както в случая **а2)** разглеан в раздел 2.2.

Като резюме можем да кажем, че условията дадени в равенства (8) са условия за възникване на прост граничен цикъл. Нещо повече, условията дадени в равенства (8) се използуват и като дефиниционни условия за прост граничен цикъл, поради което е в сила следното определение.

Определение: Затворената траектория Γ_0 е *прост граничен цикъл*, или *хиперболичен граничен цикъл*, ако удовлетворява условията дадени в равенства (8).

Определение: Простият граничен цикъл е *устойчив*, когато са изпълнени неравенствата (9а) и *неустойчив*, когато са изпълнени неравенствата (10а). Неравенствата дадени в (9а) и (10а) изразяват критериите за устойчивост, или неустойчивост на простия граничен цикъл.

<u>Случай 2</u>: Изпълнени са условията

$$d(s_0)=0 \text{ ,} \tag{11а}$$

$$d'(s_0)=0 \text{ , или } P'(s_0)=1 \text{ ,} \tag{11б}$$

и освен това всички производни на $d(s)$ до някакъв ред са равни на нула при $s=s_0$. Следователно, можем да кажем, че в този случай съществува такова число m, $m>1$, за което е изпълнено

$$d(s_0)=d'(s_0)=\cdots=d^{(m-1)}(s_0)=0 \text{ ,} \tag{12а}$$

$$d^{(m)}(s_0) = P^{(m)}(s_0) = m!\alpha_m \neq 0. \qquad (12б)$$

Тогава за $d(s)$ е в сила разложението

$$d(s) = \alpha_m (s - s_0)^m + \alpha_{m+1}(s - s_0)^{m+1} + \cdots , \qquad (13а)$$

$$\alpha_k = \frac{1}{k!} d^{(k)}(s_0) \ , \ k = m, m+1, \cdots . \qquad (13б)$$

В този случай коренът s_0 на функцията $d(s)$ е също изолиран, а затворената траектория Γ_0 е граничен цикъл, който се нарича *сложен m-кратен граничен цикъл* (*multiple limit cycle of multiplicity m*).

Условията дадени в равенства (12) са условия за възникване на сложен (кратен) граничен цикъл. Нещо повече, условията дадени в равенства (12) се използуват и като дефиниционни условия за сложен (кратен) граничен цикъл, поради което е в сила следното определение.

Определение: Затворената траектория Γ_0 е *сложен m-кратен граничен цикъл*, ако удовлетворява условията дадени в равенства (12).

От равенства (13) следва, че в достатъчно малка околност на точката s_0 знакът на функцията $d(s)$ съвпада със знака на величината $\alpha_m (s - s_0)^m = (1/m!) d^{(m)}(s_0)(s - s_0)^m$. Следователно знакът на $d(s)$, а оттук и характерът на траекториите около Γ_0, зависи съществено от това дали m е нечетно, или четно число. Когато m е нечетно число, то Γ_0 се нарича *нечетно-кратен граничен цикъл* (*multiple limit cycle of odd multiplicity*), а когато m е четно число Γ_0 се нарича *четно-кратен граничен цикъл* (*multiple limit cycle of even multiplicity*). Ще разгледаме последователно двата случая.

<u>Случай 2а</u>: m е нечетно число.

Нека предположим, че е изпълнено условието

$$d^{(m)}(s_0) = P^{(m)}(s_0) = m!\alpha_m < 0. \qquad (14)$$

Тогава са в сила неравенствата

$$d(s) < 0 \ , \text{ или } P(s) < s \ , \text{ при } s > s_0 \ , \qquad (15а)$$

$$d(s) > 0 \ , \text{ или } P(s) > s \ , \text{ при } s < s_0 . \qquad (15б)$$

Следователно последователните пресечни точки на фазовата траектория с трансверзалната крива Σ се приближават към пресечната точка на Σ и траекторията Γ_0. В този случай граничният цикъл Γ_0 се нарича *устойчив нечетно-кратен граничен цикъл* (*stable multiple limit cycle of odd multiplicity*). Лесно се вижда, че неравенствата в (15) са същите както в случая **а1)** разгледан в раздел 2.2.

Аналогично се вижда, че когато е изпълнено условието

$$d^{(m)}(s_0) = P^{(m)}(s_0) = m!\alpha_m > 0 \ , \tag{16}$$

са в сила неравенствата

$$d(s) > 0 \ , \text{ или } P(s) > s \ , \text{ при } s > s_0 \ , \tag{17а}$$

$$d(s) < 0 \ , \text{ или } P(s) < s \ , \text{ при } s < s_0 \ , \tag{17б}$$

и последователните пресечни точки на фазовата траектория и кривата Σ се отдалечават от пресечната точка на Σ и Γ_0. В този случай граничният цикъл Γ_0 се нарича *неустойчив нечетно-кратен граничен цикъл* (*unstable multiple limit cycle of odd multiplicity*). Лесно се вижда, че неравенствата в (17) са същите както в случая **а2)** разгледан в раздел 2.2.

Ще отбележим, че неравенствата дадени в (14) и (16) са критерии за устойчивост и неустойчивост на нечетно-кратен граничен цикъл.

Случай 2б: m е четно число.

В този случай $(s - s_0)^m$ винаги е положителна величина и знакът на $d(s)$ се определя от (или съвпада със) знака на $d^{(m)}(s_0)$.

Когато е изпълнено

$$d^{(m)}(s_0) = P^{(m)}(s_0) = m!\alpha_m < 0 \tag{18а}$$

са в сила неравенствата

$$d(s) < 0 \ , \text{ или } P(s) < s \ , \text{ за всяко } s \ . \tag{18б}$$

Когато е изпълнено

$$d^{(m)}(s_0) = P^{(m)}(s_0) = m!\alpha_m > 0 \tag{19а}$$

са в сила неравенствата

$$d(s) > 0 \ , \text{ или } P(s) > s \ , \text{ за всяко } s \ . \tag{19б}$$

Лесно се вижда, че когато са изпълнени условията (18), траекториите минаващи през точки от сегмента $\Sigma^+(s_0)$ (съответстващи на значения на s, за които $s > s_0$) клонят към Γ_0, а траекториите минаващи през точки от сегмента $\Sigma^-(s_0)$ (съответстващи на значения на s, за които $s < s_0$) се отдалечават от Γ_0. Когато са изпълнени условията (19) имаме обратния случай. И при двата разгледани случаи Γ_0 се нарича *полуустойчив четно-кратен граничен цикъл* (*semi-stable limit cycle of even multiplicity*).

Неравенствата (18б) и (19б) са същите както неравенствата в случаите **а3)** и **а4)** разгледани в раздел 2.2.

Ще отбележим накрая, че четно-кратният граничен цикъл е винаги полуустойчив. Неравенствата (18а) и (19а) са критерии за устойчивост и неустойчивост на полуустойчивия граничен цикъл и дават информация в коя област тра-

екториите се приближават, или отдалечават от граничния цикъл, т.е. в коя област граничният цикъл е устойчив и в коя – неустойчив.

Случай 3: Изпълнено е условието

$$d^{(m)}(s_0) = 0 \text{ за всички } m. \tag{20}$$

В този случай имаме $d(s) \equiv 0$ и всички траектории минаващи достатъчно близо до Γ_0 са затворени, т.е. имаме семейство затворени траектории, но нямаме граничен цикъл.

3. ТЕОРИЯ НА МЕЛНИКОВ

3.1. Функция на Мелников в общ случай

В този раздел ще дефинираме функцията на Мелников за смутени планарни (равнинни) системи диференциални уравнения, зададени в най-общ вид. Да разгледаме равнинната система диференциални уравнения

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + \varepsilon \mathbf{g}(\mathbf{x}, \varepsilon) \ , \tag{21}$$

където $\mathbf{x} \in \mathbf{R}^2$, $\mathbf{f}$ е аналитична функция в E, $E \subset \mathbf{R}^2$, $\mathbf{g}$ е аналитична функция в $G \times I$, $G \subset \mathbf{R}^2$, $I \subset \mathbf{R}$, а ε е малък параметър, $|\varepsilon| << 1$. Функцията $\mathbf{g}(\mathbf{x}, \varepsilon)$ се нарича функция на смущението, а самата система (21) – *смутена система*. Несмутената система се получава при $\varepsilon = 0$ и тя има вида

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}). \tag{22}$$

Казва се още, че смутената система (21) е близка до несмутената система, тъй-като при малко ε и смущението $\varepsilon \mathbf{g}(\mathbf{x}, \varepsilon)$ е малко.

По-нататък ще предполагаме, че несмутената система (22) притежава еднопараметрично семейство затворени фазови траектории, като на всяка затворена траектория Γ_0 съответства определена стойност на параметъра h. При това положение ще означаваме затворената фазова траектория с $\Gamma_0(h)$, а семейството затворени траектории с $\{\Gamma_0(h)\}_{h \in \Omega} = \{\Gamma_0(h) \mid h \in \Omega \subset \mathbf{R}\}$, където Ω е отворено множество, за което траекториите не се израждат в точка, или сепаратрисна крива, $\Omega \subset \mathbf{R}$.

Нека Σ е трансверзална крива за системата (22), която пресича семейството затворени траектории. Ще смятаме, че Σ е параметризирана със същия параметър h, както и семейството затворени траектории.

Нека h_0 е някаква стойност на параметъра h от Ω, която съответствува на дадена затворена траектория $\Gamma_0(h_0)$. Тогава може да се докаже, че ако h е достатъчно близо до h_0 и ε е достатъчно малко, то траекторията $\Gamma_\varepsilon(h)$ на смутената система (21) с начална точка върху Σ с параметър h, ще пресече отново

Σ в точка с параметър $P(h,\varepsilon)$. С други думи, може да се каже, че в достатъчно малка околност на h_0 и при достатъчно малко ε, е дефинирана функцията на Поанкаре $P(h,\varepsilon)$, която вече зависи и от параметъра ε. Тъй-като h_0 беше произволна стойност от Ω, то функцията $P(h,\varepsilon)$ е дефинирана за всяко $h \in \Omega$.

За функцията на изместването получаваме

$$d(h,\varepsilon) = P(h,\varepsilon) - h \ . \tag{23}$$

Ще отбележим, че когато **f** и **g** са аналитични функции, то функциите $d(h,\varepsilon)$ и $P(h,\varepsilon)$ са също аналитични функции. Изпълнението на равенството

$$d(h,\varepsilon) = P(h,\varepsilon) - h = 0 \ , \tag{24}$$

означава, че траекторията $\Gamma_\varepsilon(h)$ на системата (21) минаваща през точка от Σ с параметър h (или с начална точка върху Σ с параметър h) е затворена траектория. По-надолу ще дадем отговор на въпросите при какви условия уравнението (24) има реален корен, изолиран ли е този корен за да имаме граничен цикъл и какъв е типът на граничния цикъл.

Затворените траектории на несмутената система $\{\Gamma_0(h)\}_{h\in\Omega}$ се получават при $\varepsilon = 0$, поради което имаме

$$d(h,0) = P(h,0) - h \equiv 0 \ . \tag{25}$$

Функцията $d(h,\varepsilon)$ е аналитична и може да бъде развита в ред по степените на ε, т.е.

$$d(h,\varepsilon) = M_1(h)\varepsilon + M_2(h)\varepsilon^2 + \cdots \ . \tag{26}$$

Функцията

$$M_k(h) = \frac{1}{k!} d_\varepsilon^{(k)}(h,0) \tag{27}$$

се нарича *функция на Мелников от ред* k.

За аналитични системи функциите $M_k(h)$, $k = 1,2,\cdots$ са аналитични функции. *Функцията на Мелников от първи ред* обикновено се означава само с $M(h)$, т.е.

$$M(h) = M_1(h) = d'_\varepsilon(h,0) . \tag{28}$$

При това положение можем да запишем

$$d(h,\varepsilon) = M(h)\varepsilon + M_2(h)\varepsilon^2 + \cdots \ , \tag{29а}$$

или

$$d(h,\varepsilon) = M(h)\varepsilon + O(\varepsilon^2) . \tag{29б}$$

Анализът на граничните цикли в системата (21) е свързан с намиране на реалните нули на функцията $d(h,\varepsilon)$, които съответстват на затворени траектории

в системата и изследване на функцията $d(h,\varepsilon)$ около тези затворени траектории. В много случаи анализът на функцията $d(h,\varepsilon)$ може да се сведе до анализа на функцията $M(h)$.

Ще отбележим, че функцията на Мелников от втори ред $M_2(h)$ играе роля само когато $M(h) \equiv 0$. В настоящата статия ще разглеждаме само случаи, когато $M(h)$ не е тъждествено равна на нула, т.е. ще анализираме само бифуркации от първи ред.

3.2. Функция на Мелников за смутени хамилтонови системи

По-нататък ще разглеждаме смутени хамилтонови системи (или близки до хамилтонови системи). При такива именно системи условията за съществуване на гранични цикли и функцията на Мелников, която играе основна роля при анализа на граничните цикли, придобиват по-прост и удобен за използуване вид. От друга страна много от уравненията, които описват процесите в автоколебателните системи, могат да се представят като смутени хамилтонови системи. Основната задача, която ще решаваме в този случай ще бъде да намерим условията, при които уравнение (24) има изолирано решение, което означава, че системата има затворена траектория, която е граничен цикъл

Нека е дадена смутената хамилтонова система

$$\left| \begin{array}{l} \dot{x} = \partial H/\partial y + \varepsilon f(x,y,\varepsilon) \\ \dot{y} = -\partial H/\partial x + \varepsilon g(x,y,\varepsilon) \end{array} \right. , \tag{30}$$

където хамилтонианът $H = H(x,y)$ е аналитична функция в $\mathbf{R}^2$, функциите на смущението $f(x,y,\varepsilon)$ и $g(x,y,\varepsilon)$ са аналитични функции в $\mathbf{R}^2 \times \mathbf{R}$, а ε е малък параметър, $|\varepsilon| << 1$. Да означим решението на системата (30) чрез следните аналитични функции (зависещи също аналитично и от ε)

$$\left| \begin{array}{l} x = \varphi(t,\varepsilon) = \varphi_\varepsilon(t) \\ y = \psi(t,\varepsilon) = \psi_\varepsilon(t) \end{array} \right. . \tag{31}$$

При $\varepsilon = 0$ получаваме несмутената система

$$\left| \begin{array}{l} \dot{x} = \partial H/\partial y \\ \dot{y} = -\partial H/\partial x \end{array} \right. , \tag{32}$$

която е хамилтонова. Решението на (32) е

$$\left| \begin{array}{l} x = \varphi_0(t) \\ y = \psi_0(t) \end{array} \right. . \tag{33}$$

Фазовите траектории на системата (32) се дават с уравнението $H(x,y) = const$.

По-нататък ще смятаме, че е в сила следното основно предположение.

Основно предположение: Предполагаме, че хамилтоновата система (32) притежава еднопараметрична фамилия затворени траектории. Всяка затворена траектория съответствува на константна стойност h на хамилтоновата функция. Тогава уравнението на дадена затворена траектория е

$$\Gamma_0(h)\colon H(x,y)=h\ ,\ h\in\Omega\ , \tag{34}$$

където Ω е отворено множество от стойности на параметъра h, за които траекториите не се израждат в точка, или сепаратрисна крива, $\Omega\subset\mathbf{R}$.

Ще въведем някои означения, които ще използуваме нататък. Семейството затворени фазови траектории ще означаваме по следния начин

$$\{\Gamma_0(h)\}_{h\in\Omega}=\{\Gamma_0(h)\mid h\in\Omega\subset\mathbf{R}\}\ . \tag{35}$$

Областта от фазовото пространство, в която са разположени затворените фазови траектории, ще означаваме с $\mathsf{A}\ (h)$ (в англоезичната литература тази област се нарича *period annulus*). При това положение можем да запишем

$$\mathsf{A}\ (h)=\underset{h\in\Omega}{\cup}\Gamma_0(h)=\{(x,y)\mid (x,y)\in\Gamma_0(h),\ h\in\Omega\}\ . \tag{36}$$

Затворената траектория $\Gamma_0(h_0)$ разделя областта $\mathsf{A}\ (h)$ на две подобласти - $\mathsf{A}^{+}(h_0)$ и $\mathsf{A}^{-}(h_0)$ (едната от които е външна, а другата вътрешна по отношение на $\Gamma_0(h_0)$), за които имаме

$$\mathsf{A}^{+}(h_0)=\{(x,y)\mid (x,y)\in\Gamma_0(h),\ h\in\Omega,\ h>h_0\}\ , \tag{37а}$$

$$\mathsf{A}^{-}(h_0)=\{(x,y)\mid (x,y)\in\Gamma_0(h),\ h\in\Omega,\ h<h_0\}\ . \tag{37б}$$

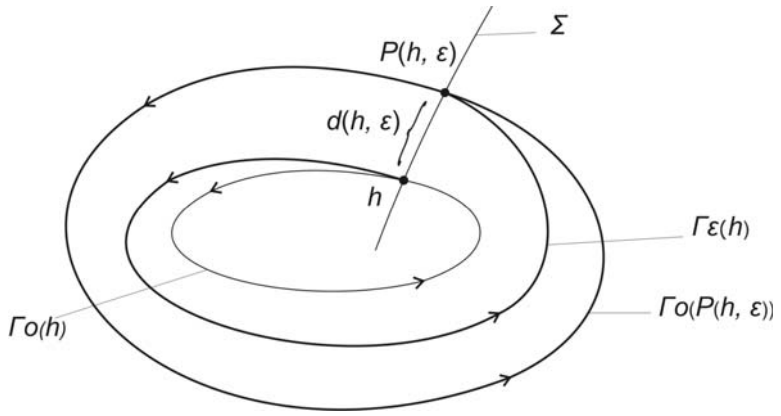
Решенията на хамилтоновата система (32) във времевата област съответствуващи на затворените фазови траектории са периодични функции, в общия случай с различни периоди. Да означим периода на решението във времевата област, съответствуващо на затворената фазова траектория $\Gamma_0(h)$, с $T_0(h)$.

Нека Σ е сечение на Поанкаре, нормално към фамилията затворени траектории и да приемем, че Σ е параметризирано също чрез константните стойности на хамилтоновата функция, т.е. чрез стойностите $h\in\Omega$.

Да означим с $\Gamma_\varepsilon(h)$ фазовата траектория на смутената система (30) с начална точка върху Σ с параметър h. Посоката на изменение на кривата (или посоката на движение по кривата) $\Gamma_\varepsilon(h)$ е по нарастващите стойности на времето. Нека след време $T_\varepsilon(h)=T(h,\varepsilon)$ траекторията $\Gamma_\varepsilon(h)$ да пресече отново (за първи път) сечението Σ в точка, която да има параметър $P(h,\varepsilon)$ (фиг. 2). Съгласно теоремата за съществуване на функцията на Поанкаре [5], точката $P(h,\varepsilon)$ съществува при малки ε, тъй-като траекторията $\Gamma_\varepsilon(h)$ е близка до $\Gamma_0(h)$. При това положение изображението $h\to P(h,\varepsilon)$ дефинира функцията на Поанкаре

$P(h,\varepsilon)$ построена върху сечението Σ. Функцията на изместването е

$$d(h,\varepsilon) = P(h,\varepsilon) - h \ . \tag{38}$$



Фиг. 2. Фазови траектории на хамилтоновата система и смутената хамилтонова система

При $\varepsilon = 0$ функцията $d(h,0)$ съответства на затворената фазова траектория $\Gamma_0(h)$ на несмутената система (32) и следователно имаме

$$d(h,0) = P(h,0) - h = 0 \ . \tag{39}$$

По-нататък ще получим точна формула за функцията $d(h,\varepsilon)$, след което ще намерим първия член в развитието на $d(h,\varepsilon)$ в ред по степените на ε, който именно член дава функцията на Мелников. Ще отбележим, че началната и крайната точка от траекторията $\Gamma_\varepsilon(h)$ принадлежащи на Σ имат съответно параметри h и $P(h,\varepsilon)$ и освен това лежат върху затворените траектории на хамилтоновата система $\Gamma_0(h)$ и $\Gamma_0(P(h,\varepsilon))$ (фиг. 2). При това положение получаваме

$$d(h,\varepsilon) = P(h,\varepsilon) - h = H\big(\varphi_\varepsilon(T_\varepsilon),\psi_\varepsilon(T_\varepsilon)\big) - H\big(\varphi_\varepsilon(0),\psi_\varepsilon(0)\big) =$$

$$= \int_h^{P(h,\varepsilon)} \mathrm{d}\,H = \int_0^{T_\varepsilon(h)} \frac{\mathrm{d}\,H}{\mathrm{d}\,t}\,\mathrm{d}\,t = \int_0^{T_\varepsilon(h)} (H_x\dot{x} + H_y\dot{y})\,\mathrm{d}\,t =$$

$$= \int_0^{T_\varepsilon(h)} \big[(-\dot{y} + \varepsilon g)\dot{x} + (\dot{x} - \varepsilon f)\dot{y}\big]\mathrm{d}\,t = \int_0^{T_\varepsilon(h)} (\varepsilon g\dot{x} - \varepsilon f\dot{y})\,\mathrm{d}\,t \ .$$

Полученият резултат може да се запише по следните еквивалентни начини

$$d(h,\varepsilon) = \varepsilon \int_0^{T_\varepsilon(h)} \big[g(\varphi_\varepsilon,\psi_\varepsilon,\varepsilon)\dot{\varphi}_\varepsilon - f(\varphi_\varepsilon,\psi_\varepsilon,\varepsilon)\dot{\psi}_\varepsilon\big]\mathrm{d}\,t \ , \tag{40а}$$

$$d(h,\varepsilon) = \varepsilon \int_{\Gamma_\varepsilon(h)} \big[g(x,y,\varepsilon)\,\mathrm{d}\,x - f(x,y,\varepsilon)\,\mathrm{d}\,y\big] \ . \tag{40б}$$

Формулите в равенства (40) са търсените точни формули за функцията на изместването.

Функцията на Мелников $M(h)$ се дефинира като първия член в развитието на $d(h,\varepsilon)$ в ред по степените на ε, т.е. имаме

$$M(h) = d'_{\varepsilon}(h,0)\ , \tag{41а}$$

$$d(h,\varepsilon) = \varepsilon M(h) + O(\varepsilon^2)\,. \tag{41б}$$

Величината $\varepsilon M(h)$ представлява приближение на $d(h,\varepsilon)$. Основното изследване на граничните цикли и различните класификации на граничните цикли, се правят с помощта на функцията $d(h,\varepsilon)$. В много случаи изследането на функцията $d(h,\varepsilon)$ може да се сведе до изследване на функцията на Мелников $M(h)$, поради което намирането на конкретен израз за функцията на Мелников е от първостепенно значение.

Функциите φ_ε, ψ_ε, T_ε, f и g са аналитични и могат да се представят чрез редове по степените на ε. Така например

$$T_\varepsilon(h) = T(h,\varepsilon) = T(h,0) + T'_\varepsilon(h,0)\varepsilon + \cdots = T_0(h) + O(\varepsilon)\ ,$$

където $T_0(h) = T(h,0)$. Аналогични редове са в сила и за останалите функции. След внасяне на тези редове в равенство (40а) и след известни преобразувания, се получава първия член от развитието на $d(h,\varepsilon)$ в ред по степените на ε, което представлява функцията на Мелников. Окончателно ***функцията на Мелников*** (или ***интеграл на Мелников***) може да се запише по два еквивалентни начина – чрез интеграл, в който интегрирането е по времето и чрез интеграл, в който се интегрира по фазова траектория, т.е.

$$M(h) = \int_0^{T_0(h)} \left[g(\varphi_0,\psi_0,0)\dot{\varphi}_0 - f(\varphi_0,\psi_0,0)\dot{\psi}_0 \right] \mathrm{d}\,t\ , \tag{42а}$$

$$M(h) = \oint_{\Gamma_0(h)} \big(g(x,y,0)\,\mathrm{d}\,x - f(x,y,0)\,\mathrm{d}\,y \big). \tag{42б}$$

Получените изрази за функцията на Мелников са най-често използвани при анализа на гранични цикли. Ще отбележим, че това са най-общи изрази. Пресмятането на функцията на Мелников и намирането на нейните нули за конкретните случаи представлява отделен математически проблем, който в много случаи е свързан с големи трудности.

3.3. Анализ на гранични цикли и функция на Мелников

Като използваме въведените дотук понятия и означения, можем да формулираме основните теореми, които играят ключова роля при изследването на граничните цикли в автоколебателните системи.

Теорема 1 (локална бифуркация на прост граничен цикъл): Нека смутената хамилтонова система (30) е аналитична. Допускаме, че хамилтоновата система (32) удовлетворява основното предположение, т.е. притежава фамилия затворени траектории $\{\Gamma_0(h)\}_{h\in\Omega}$, параметризирана с параметър h съответстващ на константните стойности на хамилтоновата функция. Нека Σ е сечение на Поанкаре, нормално към семейството затворени траектории $\{\Gamma_0(h)\}_{h\in\Omega}$ и нека Σ е параметризирано със същия параметър h, а $M(h)$ е функцията на Мелников построена върху Σ.

Нека освен това h_0 е проста нула за функцията на Мелников, т.е. изпълнени са условията

$$M(h_0)=0\ ,\ M'_h(h_0)\neq 0\,. \tag{43}$$

Тогава са в сила твърденията:

а) За достатъчно малко $\varepsilon\neq 0$, в $O(\varepsilon)$-околност на h_0 съществува стойност h_ε, такава че смутената система (30) има единствен прост (хиперболичен) граничен цикъл $\Gamma_\varepsilon(h_\varepsilon)$, чиято начална точка върху Σ има параметър h_ε. Граничният цикъл $\Gamma_\varepsilon(h_\varepsilon)$ е разположен в $O(\varepsilon)$-околност на затворената траектория $\Gamma_0(h_0)$ и клони към $\Gamma_0(h_0)$ при $\varepsilon\to 0$;

б) Граничният цикъл $\Gamma_\varepsilon(h_\varepsilon)$ е устойчив, когато е изпълнено условието $\varepsilon M'_h(h_0)<0$ и неустойчив – когато $\varepsilon M'_h(h_0)>0$;

в) Ако е изпълнено условието $M(h_0)\neq 0$, то за достатъчно малко $\varepsilon\neq 0$ системата (30) няма граничен цикъл в $O(\varepsilon)$-околност на затворената траектория $\Gamma_0(h_0)$.■

Теорема 2 (локална бифуркация на сложен граничен цикъл): Нека смутената хамилтонова система (30) е аналитична. Допускаме, че хамилтоновата система (32) удовлетворява основното предположение, т.е. притежава фамилия затворени траектории $\{\Gamma_0(h)\}_{h\in\Omega}$, параметризирана с параметър h съответстващ на константна стойност на хамилтоновата функция. Нека Σ е сечение на Поанкаре, нормално към семейството затворени траектории $\{\Gamma_0(h)\}_{h\in\Omega}$ и нека Σ е параметризирано със същия параметър h, а $M(h)$ е функцията на Мелников, построена върху Σ.

Нека освен това h_0 е нула от кратност m, $m>1$, за функцията на Мелников, т.е. изпълнени са условията

$$M(h_0)=M'_h(h_0)=\cdots=M_h^{(m-1)}(h_0)=0\ , \tag{44а}$$

$$M_h^{(m)}(h_0)\neq 0\,. \tag{44б}$$

Тогава са в сила твърденията:

а) За достатъчно малко $\varepsilon\neq 0$, в $O(\varepsilon)$-околност на h_0 съществува стойност h_ε, такава че смутената система (30) има сложен (кратен) граничен цикъл

$\Gamma_\varepsilon(h_\varepsilon)$ от кратност m, чиято начална точка върху Σ има параметър h_ε. Граничният цикъл $\Gamma_\varepsilon(h_\varepsilon)$ е разположен в $O(\varepsilon)$-околност на затворената траектория $\Gamma_0(h_0)$ и клони към $\Gamma_0(h_0)$ при $\varepsilon \to 0$;

б) Ако m е нечетно число, то $\Gamma_\varepsilon(h_\varepsilon)$ е нечетно-кратен граничен цикъл, който е устойчив, когато е изпълнено условието $\varepsilon M_h^{(m)}(h_0) < 0$ и неустойчив – когато $\varepsilon M_h^{(m)}(h_0) > 0$;

в) Ако m е четно число, то $\Gamma_\varepsilon(h_\varepsilon)$ е четно-кратен граничен цикъл, който е полуустойчив. В този случай съществува околност N на $\Gamma_0(h_0)$, такава че ако е изпълнено неравенството $\varepsilon M_h^{(m)}(h_0) < 0$, то $\Gamma_\varepsilon(h_\varepsilon)$ е устойчив в областта $N \cap \mathsf{A}^{+}(h_0)$ и неустойчив в областта $N \cap \mathsf{A}^{-}(h_0)$. Ако е изпълнено неравенството $\varepsilon M_h^{(m)}(h_0) > 0$, то обратните твърдения са в сила. ■

Като резюме можем да кажем, че условията дадени в равенства (44) са условия за възникване на сложен (кратен) граничен цикъл $\Gamma_\varepsilon(h_\varepsilon)$, който е разположен в достатъчно малка околност на затворената траектория $\Gamma_0(h_0)$.

Формулираните по-горе теорема 1 и теорема 2 дават връзка между възникващите в смутената хамилтонова система (30) гранични цикли (със съответните кратности) и нулите (със съответните кратности) на функцията на Мелников. Ще отбележим още, че теорема 1 може да се разглежда като частен случай на теорема 2 при $m = 1$.

Ще дадем още една важна теорема свързана с кратните гранични цикли.

Теорема 3 (Perko [5], стр. 371, теорема 2): Нека Γ е m-кратен граничен цикъл за дадена планарна аналитична система. Тогава са в сила следните твърдения:

а) Съществува $\delta > 0$ и $\varepsilon > 0$, такива че всяка система ε-близка по отношение на C^m-норма до дадената аналитична система има най-много m гранични цикъла в δ-околност $N_\delta(\Gamma)$ на Γ;

б) За всяко $\delta > 0$ и $\varepsilon > 0$ съществува аналитична система, която е ε-близка по отношение на C^m-норма до дадената и която има точно m прости гранични цикъла в околността $N_\delta(\Gamma)$.■

Дефиниция на понятието близки системи по отношение на зададена норма, което се използва в тази теорема е дадено в [5].

Теорема 1 и теорема 2 дават условията за възникване на гранични цикли в системата (30) около някаква затворена траектория, т.е. описват локалното поведение на системата. Следващата теорема описва глобалното поведение на системата, т.е. поведението на системата в някаква област.

Теорема 4 (глобална бифуркационна теорема): Нека са изпълнени условията дадени в теореми 1 и 2. Тогава, ако функцията на Мелников $M(h)$ има k нули $h_{0,1}$, $h_{0,2}$,...., $h_{0,k}$ (със съответните кратности), т.е. $M(h_{0,i}) = 0$, $i = 1, 2, \cdots, k$,

то при достатъчно малко $\varepsilon \neq 0$ системата (30) има в областта $\mathsf{A}\ (h)$ k гранични цикъла $\Gamma_{\varepsilon}(h_{\varepsilon,i})$, $i = 1,2,\cdots,k$ (със съответните кратности), които са разположени съответно в $O(\varepsilon)$-околности на затворените траектории $\Gamma_0(h_{0,i})$, $i = 1,2,\cdots,k$. Ако уравнението $M(h) = 0$ няма решение при $h \in \Omega$, то системата (30) няма гранични цикли в областта $\mathsf{A}\ (h)$. ■

Подробни, прецизни и различни доказателства на всичките теореми изложени дотук могат да се намерят в [5], [6], [7], [8], [9].

4. ЗАКЛЮЧЕНИЕ

В статията са дадени дефиниция и класификация на граничните цикли. Класификацията е направена с помощта на функцията на изместването. След това са изложени основните принципи от теорията на Мелников, която се използва за анализ на гранични цикли. Съгласно теорията на Мелников анализът на граничните цикли се свежда до анализа на функцията на Мелников. Подходът за изследване на граничните цикли чрез изследване функцията на Мелников има големи предимства в сравнение с други методи. Това е така, защото много от дефинициите, класификациите и свойствата на граничните цикли са правят чрез функцията на изместването, а в много случаи изследването на функцията на изместването може да се сведе до изследване на нейното първо приближение– функцията на Мелников.

ЛИТЕРАТУРА

[1] А. А. Андронов, Е. А. Леонтович, И. И. Гордон, А. Г. Майер, „Теория бифуркаций динамических систем на плоскости”, Наука, Москва, 1967.

[2] Н. Н. Баутин, Е. А. Леонтович, „Методы и приемы качественного исследования динамических систем на плоскости”, Наука, Москва, 1976.

[3] А. Пуанкаре, „О кривых, определяемых дифференциальными уравнениями”, ОГИЗ, Москва, Ленинград, 1947.

[4] Г. Дюлак, „О предельных циклах”, Наука, Москва, 1980.

[5] L. M. Perko, “Differential Equations and Dynamical Systems”, Springer Verlag, New York, Third Edition, 2001.

[6] F. Dumortier, J. Llibre, J. C. Artes, “Qualitative Theory of Planar Differential Systems”, Springer Verlag, Berlin, 2006.

[7] C. Chicone, “Ordinary Differential Equations with Applications”, Springer Verlag, New York, 1999.

[8] T. R. Blow, L. M. Perko, “Bifurcation of Limit Cycles from Center and Separatrix Cycles of Planar Analytic Systems”, “SIAM Review”, vol. **36**, № 3, pp 341-376, 1994.

[9] M. A. F. Sanjuan, “Lienard Systems, Limit Cycles, and Melnikov Theory”, “Physical Review E”, vol. **E57**, № 1, pp. 340-344, 1998.

[10] A. Andronow, “Poincare limit cycles and the theory of self-sustaining oscillations”, “Comptes Rendus de l’Academie des Sciences”, vol. **189**, pp. 559-561, 1929.

ВЛИЯНИЕ, НЕОБХОДИМОСТ И ПОЛЕЗНИ РЕЗУЛТАТИ ОТ ПРОВЕЖДАНЕТО НА ОЛИМПИАДИ ПО ТЕОРЕТИЧНА ЕЛЕКТРОТЕХНИКА СЪС СТУДЕНТИТЕ ОТ ЕЛЕКТРОТЕХНИЧЕСКИТЕ СПЕЦИАЛНОСТИ НА ТЕХНИЧЕСКИТЕ ВУЗ

Таня Методиева Стоянова[1], Георги Рашков Георгиев[2], Валери Марков Младенов[3]

[1,2]Катедра "Теоретична и измервателна електротехника",
РУ "Ангел Кънчев", ул. „Студентска" №8,
[3]Катедра "Теоретична електротехника", ТУ – София,
бул. "Климент Охридски" №8
тел.: [1](00359 82) 888 502, [2](00359 82) 888 412, [3](00359 2) 965 23 86
e-mail: [1]tstoyanova@uni-ruse.bg, [2]grashkov@uni-ruse.bg, [3]valerim@tu-sofia.bg

Резюме: *В публикацията се разглеждат влиянието, необходимостта и полезните резултати от провеждането на Олимпиади по Теоретична електротехника със студентите от електротехническите специалности на техническите ВУЗ(вътрешни и републикански кръгове). Провеждането на такива олимпиади допринася за допълнителна подготовка на студентите и по-голям интерес към дисциплината Теоретична електротехника, която е базова за студентите от електротехническите специалности, при тяхното обучението в Образователно-квалификационната степен "Бакалавър". Изучава се в първи и втори курс. Олимпиадите създават състезателен дух у студентите в съответния ВУЗ и между отделните ВУЗ-ове. Събуждат интереса на състезателите към по-нататъшно активно участие, както в няколко поредни олимпиади по ТЕ, така и към студентски научни сесии, към международните проекти. Дават още и допълнителна възможност за получаване на информация за нивото на обучение на студентите в отделните ВУЗ и стремежът на всеки от тях да го подобрява във всяко отношение.*

Ключови думи: Олимпиада по Теоретична електротехника, Вътрешни кръгове, Републикански кръг

1. ВЪВЕДЕНИЕ

Републиканският кръг на Студентската олимпиада по дисциплината Теоретична електротехника (ТЕ) се провежда през месец май. За първи път това е направено през 1981г. във ВМЕИ, гр. Габрово. Основоположници са тогавашният най-голям технически ВУЗ – ВМЕИ „В. Ил. Ленин" – гр. София и неговите преподаватели по ТЕ: проф. д.т.н. С. Фархи, проф. д.т.н. С. Папазов и други техни колеги. В това състезание участват и ВМЕИ - гр. Варна, ВТУ „А. Кънчев" – гр. Русе, Минногеоложкия институт –гр. София, Военните училища в гр. Велико Търново, гр. Шумен, гр. Долна Митрополия, Военноморското училище от

гр. Варна, ПЖИ „Т. Каблешков” – гр. София. Така е всяка година до 1990г. Републикански кръг има и през 1992г., след което следва дълго прекъсване.

През 2003 г., по инициатива на Русенски университет „А. Кънчев”, традицията на Републиканската студентска олимпиада (РСО) по ТЕ се възобновява. Тогава участват само РУ „А. Кънчев”, ТУ – гр. Варна и ВТУ „Т. Каблешков” – гр. София. С едно прекъсване през 2005г., продължава до днес. Участието на Университети не е толкова масово като през първия период, но е важно, че традицията се запазва и РСО по ТЕ се провеждат. ТУ – София, РУ „А. Кънчев” – гр. Русе, ТУ – гр. Варна участват в РСО по ТЕ с по два отбора, тъй като тези ВУЗ имат над 4000 студенти. В РСО по ТЕ се включиха ТУ – София, филиал Пловдив, ХТМУ – гр. София, ВИ по ХВП – гр. Пловдив. Еднократно участваха Военно въздушен университет – гр. Долна Митрополия, Свободен университет – гр. Бургас, Колеж по ТКП – гр. София.

Като домакини РСО по ТЕ се редуват ВУЗ-те участници (с изключение на военните училища). Състезателите в олимпиадите са студенти от електротехническите специалности на тези ВУЗ. Най- често домакини на РСО по ТЕ и преди, и сега са били: РУ „Ангел Кънчев” – Русе, ТУ – София, ТУ – Варна.

2. СЪСТОЯНИЕ НА ПРОБЛЕМА – ВЪТРЕШНИ КРЪГОВЕ, РЕПУБЛИКАНСКИ КРЪГОВЕ – ОРГАНИЗАЦИЯ, ТЕМАТИКА, ПОДГОТОВКА , ПОДБОР НА УЧАСТНИЦИТЕ

Организатори на Републиканския кръг на Олимпиадата по ТЕ в съответната учебна година са преподавателите от катедрата по ТЕ на ВУЗ-а домакин, определен на ротационен принцип предходната година от участниците в нея. Подготовката започва с изпращане на официални покани от ВУЗ-а домакин, до катедрите по ТЕ на всички технически ВУЗ-ове в България. В тези покани се публикуват вече утвърдени от всички досегашни участници:

– правила за провеждане на Републиканското състезание;

– тематика на РСО по ТЕ;

– правила за проверка и оценяване на състезателите, както и за тяхното класиране;

– график за даване на предложения на съответните ВУЗ-ове за техни членове на проверяващото жури;

– правила за подготовка на вариант за задачи за РСО по ТЕ от всеки ВУЗ-участник;

ВУЗ-ът домакин предлага датата за провеждане на РСО по ТЕ и посочва дата, до която останалите ВУЗ трябва да дадат своето съгласие. Когато всички тези въпроси се изяснят, домакинът съставя програмата за провеждане на състезанието и я изпраща на останалите ВУЗ-ове.

Тематиката на Олимпиадата е по избрани раздели от материала, изучаван по дисциплината Теоретична електротехника от студентите от електротехнически-

те специалности на техническите ВУЗ-ове. Те са посочени в поканата, изпратена от ВУЗ-а домакин. На РСО по ТЕ студентите решават три задачи от тези раздели. Условията на задачите са предложени от ВУЗ-овете, участници в РСО. Сутринта, преди започването на Републиканския кръг, журито ги групира във варианти по подходящ начин, така че да се избегне всякаква субективност. Случайно подбран състезател изтегля един от тези варианти, в присъствието на останалите участници в състезанието, което е анонимно като конкурсен изпит. След приключване на състезанието журито, което се състои от три комисии (за всяка от задачите има отделна комисия), включващи преподаватели от различни ВУЗ-ове, проверява работите на състезателите и класира участниците индивидуално и по отбори. Във всеки отбор участват пет студенти, но в неговия състезателен бал се включват точките само на първите четирима състезатели, с максимален брой точки. Победителите до трето място в двете класации получават материални награди и грамоти, осигурени от ВУЗ-а домакин.

След награждаването на победителите, се решава кой ще е домакин на РСО по ТЕ през следващата учебна година.

Подготовката на своите състезатели всеки ВУЗ прави сам, като това се осъществява от преподаватели по ТЕ от едноименните катедри чрез провеждане на допълнителни занятия. Подборът на участниците и селекцията им в отборите става след провеждане на вътрешни кръгове, които са обикновено два. Те предхождат Републиканския кръг.

3. РЕЗУЛТАТИ

Провеждането на РСО по ТЕ довежда до множество полезни резултати, основните от които са:

– Организира допълнителни занятия по ТЕ със студенти, подготвяни за участие в Олимпиадата по ТЕ. Повишава тяхната подготовка по дисциплината Теоретична електротехника и засилва нтересът им към нея, особено като се има предвид, че по учебен план часовете по ТЕ са твърде малко [1, 2, 3]. Студентите, които участват в допълнителната подготовка, са по-голям брой от състезателите в отборите и са едно сериозно ядро в студентската общност;

– Създава състезателен дух у участниците в Олимпиадите по ТЕ, както по време на вътрешните кръгове, така и на Републиканския кръг;

– Увеличаването на знанията на студентите по ТЕ е предпоставка и за разширяването на базата на техните знания, необходима както за следващите електротехнически предмети (като Електрически измервания, Полупро-водникови елементи и т.н.), така и за формирането им като специалисти.

– Стимулира вече участвали студенти в една Олимпиада, да участват и в следващи такива.

– Предизвиква поява на интерес у студентите към изследователската работа и следващо тяхно участие в Научни конференции и други форуми, а и е съвсем

нормално за тях, след като завършат магистърската си степен на обучение, да се насочат и към участие в конкурси за докторанти и асистенти, както по ТЕ, така и по други изучавани вече от тях дисциплини по време на тяхното обучение;

– Събужда се интереса на тези студенти и към участие в международни проекти;

– При Републиканския кръг се осъществяват контакти както между студентите от различните технически ВУЗ-ове, така и между техните преподаватели по ТЕ. Така става възможно да бъде получена допълнителна информация за обучението както по ТЕ, така и по други изучавани дисциплини, за обучението и базата на ВУЗ-а домакин. Дава се възможност да се „сверяват часовниците" по много страни на обучението в техническите ВУЗ-ове в България. Получава се един много полезен обмен на опит.

– В разговорите и дискусиите по време на самите Олимпиади, които не се ограничават само в рамките на няколко състезателни часа, се получава много полезно взаимно опознаване, както и уеднаквяване на позициите по важни проблеми, едно сплотяване на колегията. То е много полезно при възникване на дебати в образователната система и защитата на общите позиции.

4. ИЗВОДИ

1. Безспорно се налага изводът, че Републикански и вътрешни кръгове на Олимпиади по Теоретична електротехника трябва да се продължи, защото води до получаването на редица положителни резултати в образователно, методично, научно и социално направление.

2. Естествено и нормално е с трупането на опит и традиции постоянно да бъдат подобрявани организацията и технологията на провеждане на Олимпиадите по ТЕ.

3. Важно е, създадената вече доста дълголетна традиция, да бъде съхранена, като един здрав корен на обучението по Теоретична електротехника.

ЛИТЕРАТУРА

[1] Стоянова, Т. М., Г. Р. Георгиев, Влияние на перманентното намаление на седмичните хорариуми по основополагащите дисциплини и преместването им в по-предходни семестри върху обучението на бъдещите инженери, SUMMER SCHOOL, Advanced aspects of theoretical electrical engineering, Sozopol 2005, 30 Sept. 2005 - 03 Oct. 2005, Sozopol, Bulgaria, PROCEEDINGS, str. 195…200, 2005.

[2] http://www.uni-ruse.bg/ECTS_package/New/Erasmus_ECTS_InfoPack_FEEA_Bg.pdf

[3] http://fa.tu-sofia.bg

ИЗПОЛЗВАНЕ НА АНАЛИТИЧНИ МЕТОДИ ЗА АНАЛИЗ НА ПЛОСКИ ЕЛЕКТРОМАГНИТНИ ПРЕОБРАЗУВАТЕЛИ, РАЗПОЛОЖЕНИ НАД ПРОВОДЯЩИ ОБЕКТИ

Стефчо Георгиев Гунински

Технически университет – София, кат. „Обща електротехника”,
бул. „Кл. Охридски” 8, София 1000, тел. 965 2318, e-mail: sgg@tu-sofiq.bg

Резюме: *Показани са предимствата на метода с интегралните трансформации, който комбиниран с по-широкото използване на обобщените функции е едно удобно и бързо средство за решаване на електродинамични задачи при взаимодействието на токови контури с проводящи обекти с по-проста форма. По-специално са показани удобствата на метода при определяне на изходния сигнал на електромагнитен преобразувател (ЕМП), който може да се разглежда като система от токови контури, състоящи се от успоредни и перпендикулярни отсечки, а контролираният обект се представя като области, ограничени от успоредни равнини.*

Keywords: магнитен вектор-потенциал, интегрални преобразования, обобщени функции

1. ВЪВЕДЕНИЕ

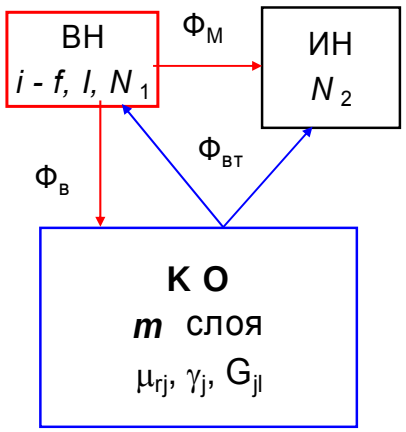
Има много методи за решаване на електродинамичните задачи [1]. Една основна класификация е раделянето им на аналитични и числени. Целта на лекцията е да се покаже, че и в момента аналитичните методи са подходящи за решаването на определени класове задачи. Основно предимства на предлагания метод са простотата, лекотата с която се достига до крайното решение и необходимостта от елементарен, лесно достъпен софтуер. Математическите процедури (решаване на частни диференциални уравнения, използване на обобщени функции и пр.) по своята същност са сложни, но при този метод прилагането им е максимално формализирано и улеснено. Става възможно лесно и бързо да се определя магнитния вектор-потенциал, а чрез него характеристиките на електромагнитното поле (***B, H, E***), напреженията и импедансите на контури (намотки), вихровите токове в проводящи среди др.

Основно се разглежда приложението на метода за целите на електромагнитния контрол и измерване [2], но успешно може да използва и в редица други случаи, когато намотките може да се представят чрез отсечки, а проводящите среди са плоски. В [3] е използвана модификация на метода за цилидрични тела и „криволинейни” правоъгълни намотки.

Основите на метода са положени преди повече от век. За развитието му определен принос имат преподаватели от катедра „Обща електротехника” на Техническия университет – София и особено покойният доц. Димчо Цветков, който е автор на идеята за разширяване на използването на обобщените функции [4] и на естественото продължение на метода – идеята за „конструиране” на решението [5].

2. СЪЩНОСТ НА ЕЛЕКТРОМАГНИТНИТЕ МЕТОДИ ЗА ИЗМЕРВАНЕ И КОНТРОЛ

Електромагнитните методи отдавна успешно се използват за целите на безразрушителен контрол [2] и измерването на някои електрически, магнитни и неелектрически величини. Идеята на метода се вижда от фиг. 1. Във възбудителната намотка (ВН) с N_1 навивки протича променлив ток i , най-често синусоидален с ефективна стойност I и честота f . Създаденият от него магнитен поток Φ_B прониква в проводящия m – слоен контролиран обект (КО). В зависимост от електрофизичните му параметри μ_{rj} и γ_j и формата и геометричните му размери G_{jl} в него протичат вихрови токове, чийто магнитен поток Φ_{BT} индуктира във възбудителната и в измервателната (ИН) намотки електродвижещи напрежения. Ако са налице и двете намотки, електромагнитният пробразувател е трансформаторен и изходен сигнал е е.д.н., индуктирано в ИН. Ако няма измервателна намотка, преобразувателят е параметричен и изходен сигнал е промяната на комплексния импеданс на ВН.



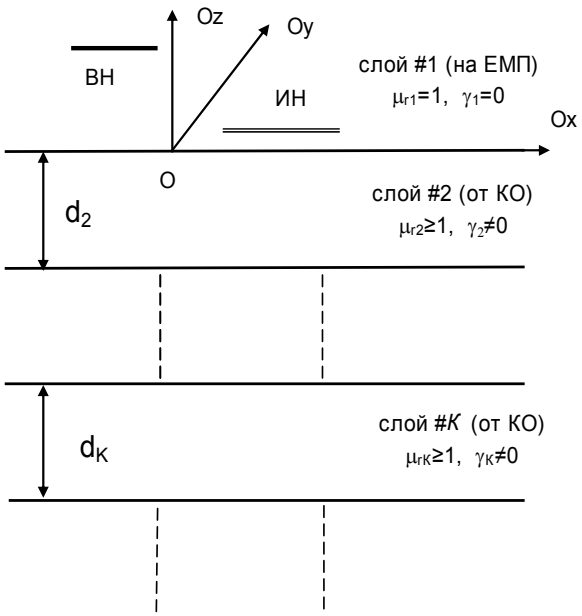
Фиг. 1. Блокова схема на електромагнитния контрол/измерване

3. ЕЛЕКТРОДИНАМИЧНА ЗАДАЧА ПРИ ЕЛЕКТРОМАГНИТНИТЕ МЕТОДИ ЗА ИЗМЕРВАНЕ И КОНТРОЛ

3.1 Модел на системата ЕМП – КО

Моделът на системата ЕМП – КО е показан на фиг. 2. Въведена е декартова координатна система *Oxyz*. Равнини, успоредни на *Oxy*, разделят пространството на *(m+1)* области, номерирани отгоре надолу. Слой #1 ($z > 0$) е областта на ЕМП, състоящ се в общия случай от ВН и ИН и която обикновено е заета от въздух. И двете намотки се разглеждат като линейни контури. Във ВН протича синусоидален ток. Контролираният обект е представен от останалите m слоя (с номера от 2 до *(m+1)*). Те се считат линейни, хомогенни и изотропни със зададени константни специфична електрическа проводимост и магнитна проницае-

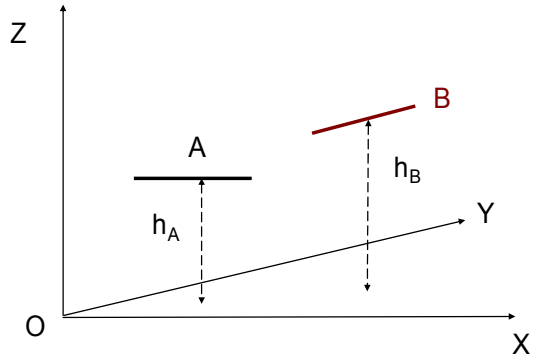
мост. Този модел не води до големи грешки, ако размерите на ЕМП са много по-малки от размерите на КО (главно тези по осите *Ox* и *Oy*) и напречните размери на намотките на ЕМП са много по-малки от габаритите им.



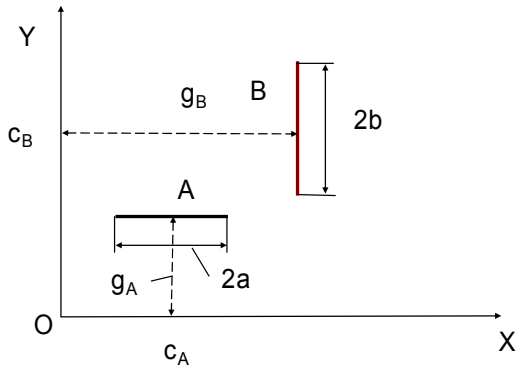
Фиг. 2. Представяне на системата ЕМП - КО

3.2. Вид на възбудителната намотка

Възбудителната намотка е съставена от комбинация от отсечки, които са успоредни на оста *Ox* (тип ***A***) или на *Oy* (тип ***B***) – фиг. 3. Техните проекции в равнината *Oxy* са дадена на фиг. 4.



Фиг. 3. Двата типа съставни отсечки на ВН



Фиг. 4. Проекции в равнината *Oxy* на двата типа съставни отсечки на ВН

3.3 Изходни уравнения

Изхожда се от уравненията на Максуел в комплексен вид [1]. Известни са предимствата от въвеждането на комплексния магнитен вектор-потенциал [1]:

(1) $$\vec{B} = rot\,\vec{A}$$

След известните преобразувания за $i^{\text{тата}}$ област се получава уравнението на Хелмхолц [2]:

(2) $$\nabla^2 \vec{A}_i - k_i^2 \vec{A}_i = -\mu_i \vec{J}_{fi} \quad , \quad i = 1,2...(m+1) \ ,$$

където $k_i^2 = j\omega\mu_i\gamma_i$.

Векторът на токовата плътност външния възбудителен ток има проекции само в първата област (i = 1) и те са от вида:

Фрагмент А:

(3) $$J_{fx} = I\,\delta\,(z\text{-}h_A)\,\delta\,(y\text{-}c_A)[\eta(x\text{-}c_A + a\,) - \eta(x - c_A - a\,)]$$
$$J_{fy} = J_{fz} = 0 \text{ (индексът “1” е изпуснат).}$$

Фрагмент В:

(4) $$J_{fy} = I\,\delta\,(z\text{-}h_B)\,\delta\,(y\text{-}c_B)[\eta(x\text{-}c_B + b\,) - \eta(x - c_B - b\,)]$$
$$J_{fx} = J_{fz} = 0 \text{ (индексът “1” е изпуснат) ,}$$

където

$\delta\,(v - n)$ е функцията на Дирак ,
$\eta(v - n)$ е функцията на Хевисайд [4].

3.4. Решение на векторните частни диференциални уравнения

Във формула (2) всъщност се крие система от 2(m + 1) скаларни частни диференциални уравнения, защото за всяка от (m + 1)$^{\text{те}}$ области има по *едно векторно* уравнение, от което произлизат по *две скаларни* уравнения за проекциите A_x и A_y на комплексния вектор-потенциал. Той няма проекция по оста Oz , защото от (3) и (4) се вижда, че и токовата плътност на възбудителния ток няма такава проекция. Всички проекции обаче са функция и на трите координати x, y и z. Въпросните уравнения са от вида:

(5) $$\nabla^2 \dot{A}_{il} - k_i^2 \dot{A}_{il} = -\mu_i \dot{J}_{fli} \quad , \quad i = 1,2...(m+1) \ , \ l = x, y \ .$$

Много голямо улеснение е, че всяко скаларно уравнение съдържа само една неизвестна проекция, т.е. може да се решава самостоятелно.

Ако се приложи двойна безкрайна Фурие интегрална трансформация с параметър λ по x и параметър υ по y [6]:

(6) $$A^*(\lambda, \nu, z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A(x, y, z)\, e^{-j(\lambda x + \nu y)}\, dx\, dy, \quad A(x, y, z) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} A^*(\lambda, \nu, z)\, e^{j(\lambda x + \nu y)}\, d\lambda\, d\nu$$

вторите частни производни по x и y в хамилтоновия оператор ще се заменят с умножение съответно по λ^2 и υ^2 , т.е. ще остане само втората производна по z и по отношение на образите $A^*(z,\lambda,\upsilon)$ диференциалните уравнения стават обикновени. Уравненията са линейни и за $i \neq 1$ са и хомогенни, т.е. решението им е

елементарно. Единствено в първата област уравненията са нехомогенни. Нехомогенността се дължи на образа на плътността на възбудителния ток във ВН. От (3) и (4) се вижда,че оригиналите на проекциите на $\boldsymbol{J}_{\mathbf{f}}$ са от вида:

$$J_{\mathrm{fl}} = K\,\delta\,(y\text{-}c)[\eta(x - a) - \eta(x - b)]\ ,\ l = x\ ,y \tag{7}$$

или същото, но с разменени места на x и y. От (6) за образите се получават изрази от вида:

$$\begin{aligned} J^{*}_{fl} &= K.\int_{-\infty}^{+\infty}\delta(y-c)\exp(-jy\upsilon)dy\int_{a}^{b}\exp(-jx\lambda)dx = \\ &= j\frac{K}{\lambda}\left[e^{-j(\lambda b+\upsilon c)} - e^{-j(\lambda a+\upsilon c)}\right] \end{aligned} \tag{8}$$

При по-сложна форма на контура на ВН се пести време, ако се намери образа по (8) само на проекцията с по-прост вид, а образа на другата проекция да се определи от израза за дивергенцията на $\boldsymbol{J}_{\mathbf{f}}$**,** който за образите има вида:

$$\mathrm{j}\lambda\, J^{*}_{fx} + \mathrm{j}\upsilon\, J^{*}_{\mathrm{fy}} = 0\ . \tag{9}$$

Методът показва голяма ефективност при многонавивкови ВН. При тях токовата плътност на възбудителния ток е сума от плътностите на отделните навивки. От (8) се вижда, че те от своя страна при образите това са алгебрични суми от експоненти с имагинерен степенен показател (виж средните скоби в (8)), т.е. получават се няколко суми от вида:

$$S = K\sum_{l=1}^{N_1} e^{jal} \tag{10}$$

Но събираемите в дясната страна образуват геометрична прогресия с първи член и частно exp(ja) и сумирането е елементарно:

$$S_1 = K\,e^{ja(N_1+1)/2}\,\frac{\sin aN_1/2}{\sin a/2}\ . \tag{11}$$

Понякога в (10) се сумира по l от 0 до ($N_1 - 1$). Тогава сумата е:

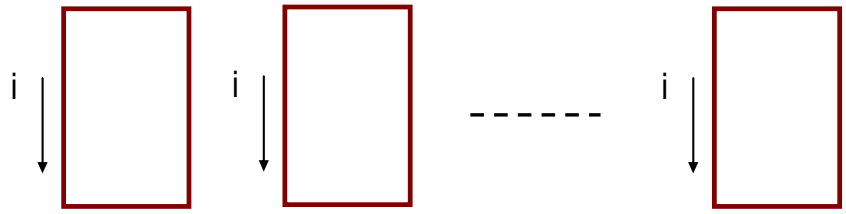
$$S_2 = K\,e^{ja(N_1-1)/2}\,\frac{\sin aN_1/2}{\sin a/2}\ . \tag{12}$$

Очевидно е опростяването на изчисленията, тъй като вместо сумата от (10) ще се изчисляват много по-компактния израз от (11) или (12).

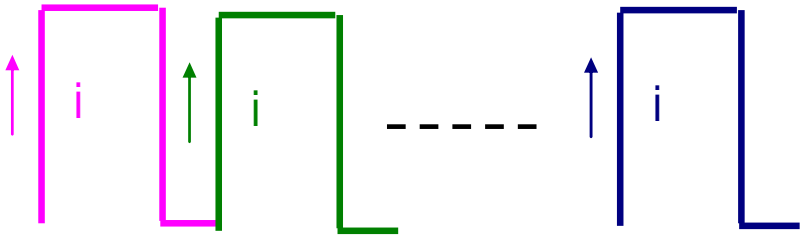
За решаване нехомогенните уравнения е удобно да се приложи метода на Лагранж с вариране на константите. При него намирането на частния интеграл е свързано с решаването на два интеграла, но от (3) и (4) се вижда, че те съдържат z само в делта-функции, а наличието й в подинтегрална функция прави интегрирането лесно.

Посоченото предимство се проявява най-силно при многонавивкови ВН при които:

– във формата има цикличност (фиг. 5. и фиг. 6)

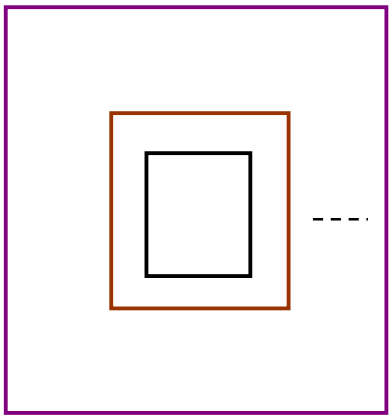


Фиг. 5. Многонавивкова ВН тип „успоредни правоъгълници”

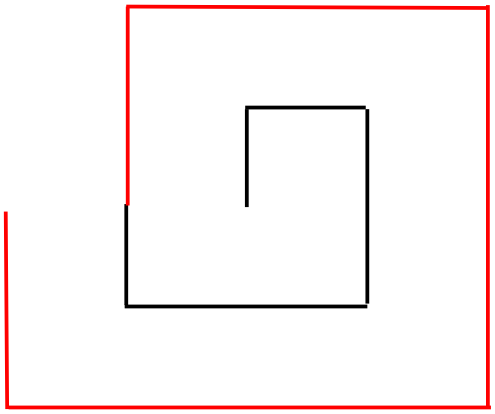


Фиг. 6. Многонавивкова ВН тип „серпантина” [8]

– във формата има “прогресивна” цикличност (фиг. 7. и фиг. 8)



Фиг. 7. Многонавивкова ВН тип „концентрични правоъгълници”



Фиг. 8. Многонавивкова ВН тип „правоъгълна спирала” [7,9]

Решението за всяка от проекциите на образите на магнитния вектор-потенциал съдържа по две комплексни интеграционни константи. Общият им брой за всички области е $2(m + 1)$ и те се определят от обичайните гранични условия за нормалните и тангенциалните компоненти на $\boldsymbol{B}$ и $\boldsymbol{H}$ на границите на слоевете при $z = 0, -d_2 , -d_3 , \ldots , -d_m$.

4. ИЗХОДЕН СИГНАЛ ЕМП

Изходният сигнал на ЕМП се определя чрез магнитния вектор-потенциал [2] в зависимост от вида преобразувателя:

– параметричен (само ВН)

(13) $$Z_{BH} = -\dot{E}_{1BH} / \dot{I} = j\omega N_1 \dot{\Phi}_{BH} / \dot{I} = \frac{j\omega N_1}{\dot{I}} \oint_{\Gamma} \vec{A}_{1BH} d\vec{l}$$

– трансформаторен (ВН и ИН)

(14) $$\dot{U}_{И} = -\dot{E}_{2BH} = j\omega N_2 \dot{\Phi}_{BH} = j\omega N_2 \oint_{\Gamma} \vec{A}_{1BH} d\vec{l}$$

4. ЗАКЛЮЧЕНИЕ

Разгледаният метод съвсем не е универсален. Например при електромагнитнитния безразрушителен контрол той е много удобен и лесно приложим при повечето задачи от дебелометрията и структороскопията, но е неподходящ за повечето задачи при дефектоскопията.

При анализа на електромагнитни полета при взаимодействието на некръгли („ъгловати”) намотки с проводящи (в т. ч. феромагнитни) плоски тела с относително проста форма той има определени предимства – бързина, формализиране на операциите (на практика липсва същинско интегриране), необходимост от конвенционален, лесно достъпен софтуер.

Методът ефикасно може да се използва при решаването на по-прости електродинамични задачи като ефикасна алтернатива на мощните, универсални, но и по-сложни и по-бавни и най-важно изискващи скъп софтуер числени методи като например използващите метода на крайните елементи.

ЛИТЕРАТУРА

[1] Папазов С., С. Фархи, Теоретична електротехника, ч. 3, С. Техника, 1988.
[2] Електромагнитна дефектометрия. Под ред. на Д. Цветков, Техника, София, 1981.
[3] Гунински С., Тримерни задачи на електромагнитния контрол при наличие на движение. Автореф. на докт. дисертация, София, 1991.
[4] Гельфанд И., Г. Шилов, Обобщенные функции и действиями над ними. М., ГИФМЛ, 1976.
[5] Цветков Д., Към въпроса за непосредствено получаване на решенията при вихровотоковата дефектоскопия на плоски изделия, Известия на ВМЕИ„Ленин”, т.XXXIV, кн.2, 1975.
[6] Винер Н., Р. Пэли, Преобразование Фурие в комплексной области, М., Наука, 1962
[7] Гунински С., И. Петров, Изходен сигнал на електромагнитен преобразувател със спирална форма, Научни известия на НТС по Машиностроене, год. XII, бр. 1 (79), юни 2005 год., Дефектоскопия‘2005. XX Национална конференция по безразрушителен контрол с международно участие, Созопол, 13-15 Юни 2005 г., с. 243-247.
[8] S. Mukhopadhyay, Novel high performance planar electromagnetic sensors, NDT.net, Aug 2005, Vol. 10 No 8.
[9] W. Hurley, M. Duffy, S. O’Reilly and S. O’Mathuna. Impedance Formulas for Planar Magnetic Structures with Spiral Windings, IEEE Tranc. Of Industrial Electronics, vol 46, NO 2, April 1999, pp. 271-278.

TECHNICAL UNIVERSITY OF SOFIA

FACULTY AUTOMATION

DEPARTMENT THEORETICAL ELECTRICAL ENGINEERING

8th SUMMER SCHOOL

ADVANCED ASPECTS OF THEORETICAL ELECTRICAL ENGINEERING - SOZOPOL'10

Sozopol'10

in the framework of

THE DAYS OF SCIENCE OF THE TECHNICAL UNIVERSITY OF SOFIA, SOZOPOL, BULGARIA, SEPT. 2010

Part II: Regular Papers

Edited by Valeri Mladenov

ISSN: 1313-9487

Organizing Committee

Honorary Chairmen:

V. Georgiev – Bulgaria
S. Farchy – Bulgaria
L. Kolev – Bulgaria
S. Papasow – Bulgaria
V. Savov – Bulgaria

Chair:

V. Mladenov, TU-Sofia, Bulgaria

Co Chairs:

S. Guninski, TU-Sofia, Bulgaria
J. Georgiev, TU-Sofia, Bulgaria
I. Uzunov, TU-Sofia, Bulgaria

Members:

K. Brandisky, TU-Sofia, Bulgaria
A. Chervenkov, TU-Sofia, Bulgaria
K. Ivanov, TU-Sofia, Bulgaria
N. Petkova, TU-Sofia, Bulgaria
S. Petrakieva, TU-Sofia, Bulgaria
N. Radev, TU-Sofia, Bulgaria
I. Tabahnev, TU-Sofia, Bulgaria
S. Terzieva, TU-Sofia, Bulgaria
K. Todorova, TU-Sofia, Bulgaria
G. Toshev, TU-Sofia, Bulgaria
S. Vladov, TU-Sofia, Bulgaria
I. Trushev, TU-Sofia, Bulgaria
G. Tsenov, TU-Sofia, Bulgaria
I. Yacheva, TU-Sofia, Bulgaria

International Programme Committee

Honorary Chairmen:

L. Chua, Berkeley-USA
N. Mastorakis, Greece
R. Sikora, Poland

Members:

D. Baldomir, Spain
A. Bossavit, France
A. Brykalski, Poland
H. Butterweck, Netherlands
T. Chady, Poland
K. Demirtchyan, Russia
S. Gratkowski, Poland
K. Hameyer, Germany
W. John, Germany
L. Klinkenbusch, Germany
L. Kolev, Bulgaria
A. Kost, Germany
Z. Leonowicz, Poland
W. Mathis, Germany
V. Mladenov, Bulgaria
M. Ogorzałek, Poland
S. Osowski, Poland
L. Pichon, France
B. Reljin, Serbia
L. Sumichrast, Slovakia
J. Sykulski, UK
M. Tadeusiewicz, Poland
R. Tetzlaff, Germany
H. Toepfer, Germany
H. Uhlmann, Germany
R. Weigel, Germany

PREFACE

These Proceedings are organized in two parts and contain the plenary lectures and the regular papers presented at the 8th Summer School *Sozopol'10,* which took place in Sozopol, Bulgaria, between 19 and 22 Sept. 2010 in the framework of the "Days of the Science of the Technical University of Sofia". The Summer School covers the advanced aspects of Theoretical Electrical Engineering and it is a platform for post-graduate training of Ph.D. students and young scientists. During the Summer School well-known experts presented some advanced aspects of circuits and systems theory, electromagnetic field theory and their applications. Apart from the educational part of the Summer School a presentation of original authors' papers took place.

The main topics of the Summer School *Sozopol'10* include Circuits and Systems Theory, Signal Processing and Identification Aspects, Electromagnetic Fields, Theoretical Concepts, Applications and Innovative Educational Aspects.

The Summer School *Sozopol'10* has been organized by the Department of Theoretical Electrical Engineering of the Technical University of Sofia with the main sponsorship of the Research and Development Sector of the Technical University of Sofia. This has been the eighth edition of the event, after the Summer Schools in 1986, 1988, 2001, 2002, 2005, 2007 and 2009. The Summer School is under the patronage of the International Symposium on Theoretical Electrical Engineering (ISTET) and it is a regular ISTET event.

There were 56 participants from 5 different countries at the Summer School this year. There were 13 plenary lectures and 30 regular papers that are published in these Proceedings. Providing the recent advances in Theoretical Electrical Engineering the Proceedings will be of interest to all researchers, educators and Ph.D. students in the area of Electrical Engineering.

Special thanks are due to the Research and Development Sector, Faculty of Automation and the Section of Social Services of the Technical University of Sofia about the overall support of the event. We also want to thank to the IEEE Bulgaria CAS Chapter and the World Scientific and Engineering Academy and Society (WSEAS), which partially sponsored the event. We hope to meet again in the following edition of the Summer School to continue the good tradition and collaboration in the field of Theoretical Electrical Engineering.

Organizing Committee
Sofia, Oct. 2010

CONTENTS

SECTION 1 – CIRCUITS AND SYSTEMS THEORY, APPLICATIONS

SECTION 2 – THEORETICAL CONCEPTS, SIGNAL PROCESSING AND IDENTIFICATION ASPECTS

SECTION 3 – ELECTROMAGNETIC FIELDS

SIGNAL COMPETITION APPROACH FOR SYNTHESIS OF ASYNCHRONOUS HIGH-SPEED DIGITAL CIRCUITS

Simona Filipova*, Valeri Mladenov**

* Department of Theoretical Electrical Engineering, Technical University of Sofia, 1156 Sofia, Bulgaria, E-mail: petrakievas-te@tu-sofia.bg
** Department of Theoretical Electrical Engineering, Technical University of Sofia, 1156 Sofia, Bulgaria, E-mail: valerim@tu-sofia.bg

Abstract: *In this paper, the problem of signal competition at the synthesis of asynchronous high-speed digital circuits is investigated. The synthesis problem consists of determining the delay of all asynchronous gates of the circuit, such that to avoid the signal competition. The proposed approach is based on determining all possible pairs of paths between every pair of gates in the circuit. The conditions for eliminating the signal competitions are transformed into a linear programming problem. The desired delays of the asynchronous gates are obtained as a solution of this problem.*

Keywords: Synthesis of the asynchronous digital circuits, RSFQ circuits, Signal competition, Simplex method for Linear programming

1. INTRODUCTION

The clock frequencies leads to arising the more and more significant problem between the clock distribution and the desire of minimum power consumption within the synchronous digital circuits [1]. This has led to an increasing importance of the asynchronous circuits, whose synthesis has been an attractive area of research since the mid of the previous century.

One of the most promising techniques for realization of such circuits is the Rapid Single Flux Quantum (RSFQ) electronics [2] which use Low-Temperature Superconductors. It relies on quantum effects in superconductors with respect to switching the signals instead of transistors. However, it is very different from the traditional CMOS electronics technology. First, it is based on superconductors, so a cryogenic environment is required. Second, the digital signals are represented by picosecond-duration pulses that travel down superconducting microstrip transmission lines instead of voltage levels in the conventional semiconductor electronics. Third, the pulses are single quanta at the lowest energy level allowed by quantum mechanics in the system, and hence cannot change significantly in transit. They do not lose energy, spread out, or interfere. The quantum pulses are switched by Josephson junctions instead of transistors. The signals in RSFQ circuits cannot be split into multiple outputs without active circuit elements. Both the extremely high operation speed [3] and ultra low-power consumption [4] of these digital circuits make them a promising tool for some practical asynchronous circuits' applications [5].

One of the most attractive areas for application of the ultra high-speed digital asynchronous circuits is the Internet data processing. Due to the lack of global clock-

ing, these circuits must satisfy much harder timing conditions than the synchronous ones. The reasons for them are the signal competitions, which occur if there are different paths, through which a given signal can reach one and the same gate of the circuit. Such signal competitions cause hazards, which in turn lead to wrong circuit operation. In order to avoid the signal competition, the synthesis of the asynchronous digital circuits consists of proper choice of the delays of the asynchronous gates.

In this paper we propose a method for such a synthesis. The method consists of two stages. After the first stage all pairs of competitive signal paths are obtained. The properly values of the delays of the gates included in these paths adjusts in the second stage. The problem of obtaining all pairs of competitive signal paths is solved in [6] and [7]. Based on this, the second problem (for proper adjustments of the delays of the asynchronous gates) is transformed into a linear programming problem, which solves with simplex method. The final solution gives a desired set of delays corresponding to the circuit gates. If there is more than one solution, the optimal set of delays chooses applying an additional criterion. Such a criterion, should take into account some extra requirements for the asynchronous circuit considered.

The paper is organized as follows. In the next section is defined the synthesis problem. Then, in section 3, this problem is transformed into a linear programming problem, related to the desired set of the delays of the asynchronous gates. A numerical example of the digital circuit with 10 gates is given in section 4. In the last section of the paper are made the concluding remarks.

2. PROBLEM STATEMENT

The asynchronous digital circuits transmit and process data pass through high-speed channels. They model by a directed graph, where the graph vertexes present the gates of the circuits studied and the arcs correspond to the signal flow paths in it. The algorithms proposed in [6] and [7] allow obtaining all different path pairs between every pair of vertices. These signal paths do not contain joint intermediate vertexes different from the start and the end vertex of the paths.

The conditions about signal competition in the asynchronous circuit considered are related to the time delays associated with every different path pairs q^1 and q^2 between every pair of vertices i and j of the directed graph model:

$$\left| \sum_{l \in q^1} \tau^l - \sum_{l \in q^2} \tau^l \right| \leq \Delta T, \tag{1}$$

where the sums in (1) are with respect to the gates included in the paths q^1 and q^2 between the same pair of vertexes i and j, τ^l is the delay associated with gate l and ΔT is the period of the input pulses, entered to the analyzed circuit [6], [7].

Then the synthesis problem reduces to obtaining the delays τ^s, $s = 1,2,...,n$ of the asynchronous gates, where n is the total number of the gates in the circuit studied.

3. PROBLEM SOLUTION

The synthesis problem solution gives the values of each gate delay in the asynchronous circuit considered. These delays satisfy the signal competition conditions (1) for every different path pairs q^1 and q^2 between arbitrary pair vertices *i* and *j* of the corresponding directed graph model.

Furthermore, in practise there are some technological conditions concerning to the upper and the lower bounds of the delays

$$\underline{\tau}^i \le \tau^i \le \overline{\tau^i},\ i = 1,...,n. \tag{2}$$

Conditions (1) and (2) are set of inequalities in the *n* dimensional space of the unknown delays. If a region that is a common solution of these inequalities exists, every point from this region corresponds to a set of delays that satisfy both the signal competition conditions (1) and the technological constraints (2). This gives a chance to choose the optimal, in desired sense point, from this region.

For guarantee the high-speed signal processing, the most appropriate decision is to choose the delays with minimum values. Hence, a possible criterion for optimal choice of the delays could be the total sum of the gate delays of the circuit studied to get a minimum value

$$F(\tau) = \sum_{i=1}^{n} \tau^i \to \min. \tag{3}$$

Thus, the basic synthesis problem transforms into a linear programming model with objective function (3) and subject to constrains (1) and (2). The solution of this problem gives the values of delays that satisfy both the signal competition conditions (1) and the technological constraints (2).\

In many practical applications, more appropriate criterion could be the weighted sum of the gate delays to get a minimum value. In this case one can state some preferences between the delays by appropriate choice of the weights.

The matrix of coefficients in constraints (1) and (2) has the following features:

1. Its elements are only 0, 1 and -1.
2. The elements of the columns, associated to the input and the output gates of the asynchronous digital circuit considered, are equal to 0.
3. The total number of 1's and (-1)'s in the rows of the constraints matrix, associated to (1), is:

$$\max\left(\tau^1,\ \tau^2\right) - 2, \tag{4}$$

where

$$\tau^1 = \sum_{l \in q^1} \tau^l,\ \ \tau^2 = \sum_{l \in q^2} \tau^l.$$

The solution of the arisen linear programming problem is obtained by simplex method (see Fig. 1).

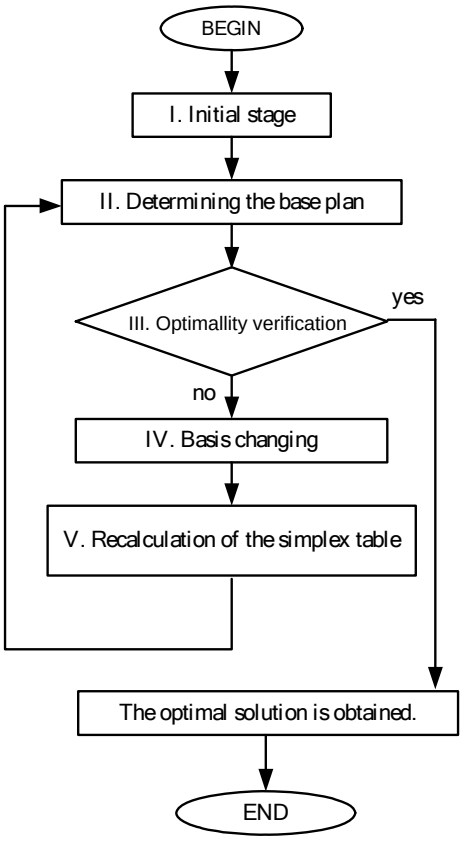


Fig. 1. Simplex method algorithm solving the linear programming problem

The algorithm of the well known simplex method is following.

Step I: Initial stage

I.1. Filling the simplex table.

I.2. Determining the base variables (according to the identity matrix in the system of constraints) and put them into the simplex table.

Step II. Determining the base plan

It sets the value 0 to the free variables (the first $(2n+m)$ variables) and the others (basis variables) take the respective values of the coefficients on the right side of the constraints' system. Then the objective calculates with these values of the variables and it notes in the simplex table.

Step III. Optimality verification of the base plan

- "yes" – the solution of the linear programming problem is found;
- "no" – go to Step IV.

Step IV: Basis changing

Changing the variables included in the based plan with respect to satisfying the optimality criterion defined.

Step V: Recalculation of the simplex table

In this step the simplex table recalculates and it goes to step 2.

4. NUMERICAL EXAMPLE

Let consider an asynchronous digital circuit, presented with the corresponding directed graph shown on Fig. 2 [7]. It consists of 10 gates and 6 multiplexed in time channels, designed with *a*, *b*, *c*, *d*, *e* and *f*, respectively. Two complexes signals $a.\overline{c} + b.d.f + \overline{f}$ and $a.\overline{c} + b.d.f + \overline{e} + \overline{f}$ get as an output of the circuit studied. The input signals *a*, *b*, *c*, *d*, *e* and *f* arrive on the two circuit's inputs according to Fig. 2. Then they demultiplex and proceed through the high-speed channels and finally the mixing signals $a.\overline{c} + b.d.f + \overline{f}$ and $a.\overline{c} + b.d.f + \overline{e} + \overline{f}$ depart the circuit through the two output channels.

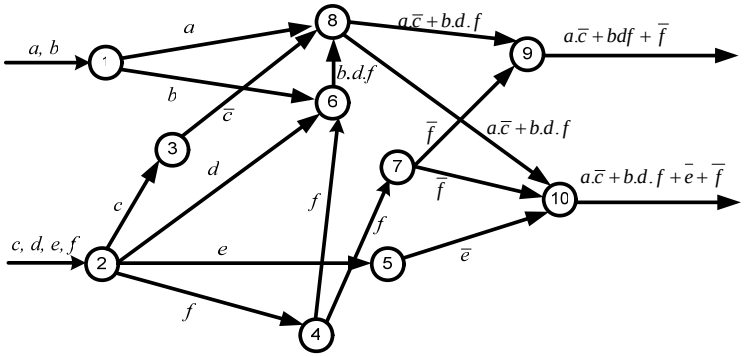


Fig. 2. The directed graph of the asynchronous digital feedforward electric circuit

The delays, associated to the vertexes in the graph model considered, are τ_s, $s = 1,2,...,10$, respectively. The corresponding objective function is:

$$F(\tau) = \sum_{i=1}^{10} \tau^i \rightarrow \min . \tag{5}$$

All possible different pairs of paths between each two arbitrary vertexes in the graph model are obtained in [7]:

$$\begin{array}{lll}
P_{26}^1 = \{V_2, V_6\} & \text{and} & P_{26}^2 = \{V_2, V_4, V_6\} \\
P_{18}^1 = \{V_1, V_8\} & \text{and} & P_{18}^2 = \{V_1, V_6, V_8\} \\
P_{28}^1 = \{V_2, V_3, V_8\} & \text{and} & P_{28}^2 = \{V_2, V_6, V_8\} \\
P_{28}^1 = \{V_2, V_3, V_8\} & \text{and} & P_{28}^2 = \{V_2, V_4, V_6, V_8\} \\
P_{2,10}^1 = \{V_2, V_5, V_{10}\} & \text{and} & P_{2,10}^2 = \{V_2, V_4, V_7, V_{10}\} \\
P_{2,10}^1 = \{V_2, V_5, V_{10}\} & \text{and} & P_{2,10}^2 = \{V_2, V_3, V_8, V_{10}\} \\
P_{2,10}^1 = \{V_2, V_5, V_{10}\} & \text{and} & P_{2,10}^2 = \{V_2, V_6, V_8, V_{10}\} \\
P_{2,10}^1 = \{V_2, V_5, V_{10}\} & \text{and} & P_{2,10}^2 = \{V_2, V_4, V_6, V_8, V_{10}\} \\
P_{49}^1 = \{V_4, V_7, V_9\} & \text{and} & P_{49}^2 = \{V_4, V_6, V_8, V_9\} \\
P_{4,10}^1 = \{V_4, V_7, V_{10}\} & \text{and} & P_{4,10}^2 = \{V_4, V_6, V_8, V_{10}\} \\
P_{29}^1 = \{V_2, V_4, V_7, V_9\} & \text{and} & P_{29}^2 = \{V_2, V_3, V_8, V_9\} \\
P_{29}^1 = \{V_2, V_4, V_7, V_9\} & \text{and} & P_{29}^2 = \{V_2, V_6, V_8, V_9\} \\
P_{2,10}^1 = \{V_2, V_4, V_7, V_{10}\} & \text{and} & P_{2,10}^2 = \{V_2, V_3, V_8, V_{10}\} \\
P_{2,10}^1 = \{V_2, V_4, V_7, V_{10}\} & \text{and} & P_{2,10}^2 = \{V_2, V_6, V_8, V_{10}\}
\end{array} \tag{6}$$

The conditions (1) about the signal competition, related to the time delays associated with the different path pairs q^1 and q^2, between every pair of vertices i and j of the respective directed graph, are:

$$
\begin{aligned}
&\left|\tau_{26}^{1}-\tau_{26}^{2}\right| \le \Delta T\\
&\left|\tau_{18}^{1}-\tau_{18}^{2}\right| \le \Delta T\\
&\left|\tau_{28}^{1}-\tau_{28}^{2}\right| \le \Delta T\\
&\left|\tau_{28}^{1}-\tau_{28}^{2}\right| \le \Delta T\\
&\left|\tau_{2,10}^{1}-\tau_{2,10}^{2}\right| \le \Delta T\\
&\left|\tau_{2,10}^{1}-\tau_{2,10}^{2}\right| \le \Delta T\\
&\left|\tau_{2,10}^{1}-\tau_{2,10}^{2}\right| \le \Delta T\\
&\left|\tau_{2,10}^{1}-\tau_{2,10}^{2}\right| \le \Delta T\\
&\left|\tau_{49}^{1}-\tau_{49}^{2}\right| \le \Delta T\\
&\left|\tau_{4,10}^{1}-\tau_{4,10}^{2}\right| \le \Delta T\\
&\left|\tau_{29}^{1}-\tau_{29}^{2}\right| \le \Delta T\\
&\left|\tau_{29}^{1}-\tau_{29}^{2}\right| \le \Delta T\\
&\left|\tau_{2,10}^{1}-\tau_{2,10}^{2}\right| \le \Delta T\\
&\left|\tau_{2,10}^{1}-\tau_{2,10}^{2}\right| \le \Delta T
\end{aligned}
\qquad (7a)
$$

where $\tau_{i,j}^{k}$ is the k-th path between vertexes i and j and ΔT is chosen 48 ps.

After recalculation of the relations (7a) in term of the gate delays, the following constraints are obtained:

$$
\begin{aligned}
&\tau^4 \le \Delta T\\
&\tau^6 \le \Delta T\\
&\left|\tau^3-\tau^6\right| \le \Delta T\\
&\left|\tau^3-\tau^4-\tau^6\right| \le \Delta T\\
&\left|\tau^5-\tau^4-\tau^6\right| \le \Delta T\\
&\left|\tau^5-\tau^3-\tau^8\right| \le \Delta T\\
&\left|\tau^5-\tau^6-\tau^8\right| \le \Delta T\\
&\left|\tau^7-\tau^6-\tau^8\right| \le \Delta T\\
&\left|\tau^4+\tau^7-\tau^3-\tau^8\right| \le \Delta T\\
&\left|\tau^4+\tau^7-\tau^6-\tau^8\right| \le \Delta T\\
&\left|\tau^5-\tau^4-\tau^6-\tau^8\right| \le \Delta T
\end{aligned}
\qquad (7b)
$$

The lower and upper bounds $\underline{\tau}^s, \overline{\tau^s}, \; s=1,2,...,10$ of the gate delays $\tau_s, \; s=1,2,...,10$ are given in Table 1.

Gate	Logic element	Down bound	Upper bound
1	1x2 demultiplexor	62 ps	82 ps
2	1x2 demultiplexor	62 ps	82 ps
3	1x2 demultiplexor	62 ps	82 ps
4	1x2 demultiplexor	62 ps	82 ps
5	1x2 demultiplexor	62 ps	82 ps
6	AND	24 ps	38 ps
7	1x2 demultiplexor	62 ps	82 ps
8	XOR	25 ps	43 ps
9	AND	24 ps	38 ps
10	XOR	25 ps	43 ps

Table 1. The lower and upper bounds of the gates' delays

Thus, the restrictions resulting from the requirements about the bounds of the gate delays are:

$$
\begin{aligned}
&60 \le \tau^1 \le 82\\
&60 \le \tau^2 \le 82\\
&60 \le \tau^3 \le 82\\
&55 \le \tau^4 \le 82\\
&62 \le \tau^5 \le 100\\
&24 \le \tau^6 \le 38\\
&60 \le \tau^7 \le 82\\
&25 \le \tau^8 \le 45\\
&20 \le \tau^9 \le 38\\
&20 \le \tau^{10} \le 43
\end{aligned}
\qquad (8)
$$

As a result the final linear programming problem consists of the objective function (5) and two types of restrictions: (7b) and (8), which take in mind the corresponding delays in the competitive path pairs of the asynchronous circuit and the technological conditions about the upper and lower bounds of the delays.

This problem is solved with the internal linear programming optimization procedure in Excel, based on the simplex method. The final solution is following:

$$\tau_{opt} = \left[62 \;\; 66.0541 \;\; 62 \;\; 55 \;\; 95 \;\; 31.515 \;\; 63 \;\; 38.936 \;\; 24 \;\; 25\right]^T \; ps \qquad (9)$$

and the corresponding minimum of the objective function is

$$F_{opt}(\tau_{opt}) = 522.506 \; ps\,.$$

5. CONCLUSIONS

The main problem arisen in the asynchronous high-speed digital circuits design is the competitions of the signals, which proceed and transmit through the digital circuit channels.

In this paper the method for determining the delays of the gates, eliminating the signal competitions, is suggested. Firstly, it finds all pairs of competitive signal paths between every pair of gates in the asynchronous circuit considered. Then, the conditions for eliminating the signal competitions are transformed into a linear programming problem, which solution gives the desired delays of the gates. The main advantage of the new method is the fact that it uses the standard simplex method for solving the linear programming problem formulated. The respective procedure of this method is built in well known electronic table Excel and the solution process does not take significant time and calculation resources.

REFERENCES

[1] International Technology Roadmap for Semiconductors, http://public.itrs.net

[2] K. K. Likharev, V.K. Semenov, "RSFQ Logic/Memory Family: A New Josephson-Junction Technology for Sub-Terahertz-Clock-Frequency Digital Systems", IEEE Trans. Appl. Supercond., vol. 1, pp. 3-28, 1991.

[3] W.Chen, A. V. Rylyakov, V. Patel, J. E. Lukens, K. K. Likharev, "Rapid Single Flux Quantum T-Flip Flop Operating up to 770 GHz", IEEE Trans. Appl. Supercond., vol. 9, pp. 3212-3215, 1999.

[4] P. Bunyk, M. Leung, J. Spargo, M. Dorojevets, "FLUX-1 RSFQ Microprocessor: Physical Design and Test Results", IEEE Trans. Appl. Supercond., vol. 13, pp. 433-436, 2003.

[5] B. Dimov, "General Restrictions and Their Possible Solutions for the Development of Ultra High-Speed Integrated RSFQ Digital Circuits", Wissenschaftsverlag Ilmenau, Germany, 2005.

[6] V. Mladenov, K.Filipova, S. Petrakieva , B. Dimov, F.H. Uhlmann, "Analysis of Signal Competition in Asynchronous Ultra High-Speed Digital Circuits", Przegląd Elektrotechniczny, vol. 83, nr 11, pp. 197-200, 2007.

[7] V. Mladenov, S. Filipova-Petrakieva, K. Filipova, "On Signal Competition in Feedforward Asynchronous Ultra High-Speed Digital Electric Circuits", Elektrotechnica&Elektronica, vol. 5, ISSN 0861-4717, pp. 55-62, 2008.

VHDL SIMULATIONS OF RSFQ 8-BIT SHIFT REGISTER INCLUDING INFLUENCE OF JITTER

Tanya Stoyadinova[1], Thomas Ortlepp[2]

[1]Department of Theoretical Electrical Engineering, Faculty of Automatics,
Technical University of Sofia, 8 Kliment Ohridski St, Sofia-1000 BULGARIA,
phone: +35929653319, e-mail: tstoiadinova@abv.bg
[2]Department of Advanced Electromagnetics, Ilmenau University of Technology,
D-98684 Ilmenau, Germany, phone: +493677691185, e-mail: thomas.ortlepp@tu-ilmenau.de

Abstract: *We propose cell-based design methodology using hardware description language to describe RSFQ circuits. Data pulse driven logic is specific characteristic of RSFQ which emphasizes the capabilities of VHDL to describe such circuits. We modeled a couple of logic gates with their logical behavior and existing timing constraints. In order to have an adequate model for high speed operation we include the influence of timing jitter which becomes the main issue in high frequency. All the information that we use in our VHDL model of the cells was extracted from analog simulations. We analyzed 8-bit shift register in terms of influence of jitter to determine the correct shifting between data and clock pulses.*

Keywords: behavior description, shift register, timing jitter, VHDL

1. INTRODUCTION

Rapid Single Flux Quantum (RSFQ) technology is based on superconducting materials using Josephson junctions (JJ) as the basic switching elements. RSFQ circuits can operate at frequencies of hundred of gigahertz while at the same time have low power consumption. In RSFQ circuits data are represented by short voltage pulses instead of voltage level, which is used in semiconductor technology. Classical synchronous RSFQ circuits use a clock signal to translate a pulse on a data line in a "clock window" as a logic "1" and no pulse as a logic "0".

We use Very high speed integrated Hardware Description Language (VHDL) to do a transition from analog simulation to high-level behavior description of the individual basic RSFQ cells. We include the specific features of RSFQ electronics and provide the basis for analyzing more complex circuits. Cell-based technique is used to build shift register which is very important cash memory in RSFQ logic. We analyzed this circuit with the advantage of faster simulations and design. The obtained simulation results are useful for studying the possible influence of timing jitter of the performance of the circuit.

2. VHDL DESCRIPTION OF BASIC RSFQ CELLS

The basic RSFQ cells that take a part of circuit of shift register are Josephson transmission line (JTL), Splitter and D Flip Flop (DFF). A JTL is the simplest RSFQ

gate. It allows an input SFQ pulse to pass through the output. The time from arriving of an input pulse to time of producing an output pulse is the main characteristic of the JTL [1].

The Splitter doubles the SFQ pulses. When there is an input data pulse it produces two output pulses at two different outputs. The Splitter is non-storing cell like JTL. It is a data driven single channel cell. The output delay of the Splitter is about 20 % higher than this of the JTL [2].

The DFF is a bistable clocked basic cell with the ability of data storing. When data pulse appears it results in a stored flux quantum inside the DFF until clock pulse comes and releases it [3]. The timing conditions in DFF are related to the two independent signals – clock and data. There are four minimum times that have to be observing – minimum time between two data signals, minimum time between two clock signals, minimum time between clock before data and minimum time between clock signal after data signal.

The extraction of timing parameters of these cells is provided by analog simulations including thermal noise. For an advanced VHDL model the reflection of supply voltage on readout delay is also considered. For calculation the readout delay a polynomial function provided by analog simulation is used. The influence of supply voltage is participated in the polynomial as a variable. The meaningful values of the supply voltage range from 1.7 mV to 3.3 mV for most of the RSFQ gates. The influence of jitter is implemented in the VHDL models through the digital values of Gaussian distribution of the noise with standard variance σ as a measure for the jitter. We assume that all possible value for the realization of the jitter fall within the interval -6σ to +6σ [4]. The graphic of normal distribution is shown in Fig. 1.

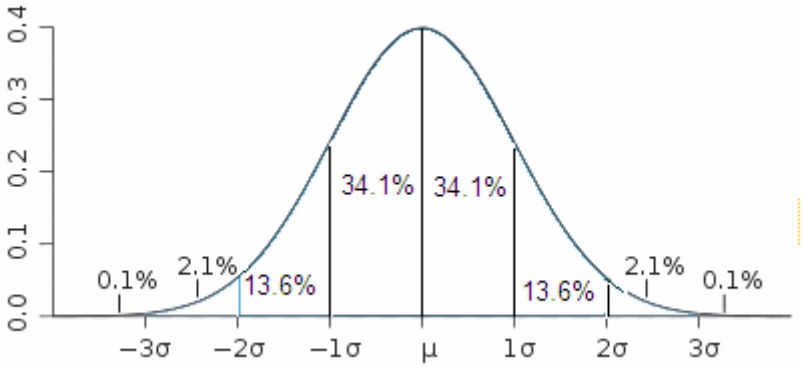


Fig. 1. Graphic of normal distribution

VHDL is one of the programming languages used to model a digital system by dataflow, behavioral and structural style of modeling [5]. It provides a good base for the transition of the logical and timing requirements of the basic RSFQ gates into HDL's model [6]. A design is described in VHDL using the concept of a design entity. A design entity is split into two parts, each of which is called a design unit. The entity declaration represents the external interface to the design entity. The archi-

tecture body represents the internal behavior description of the design entity. These parts of the code are shown in Fig. 2 for the JTL cell.

```
library IEEE;
use IEEE.STD_LOGIC_1164.ALL;
use IEEE.STD_LOGIC_ARITH.ALL;
use IEEE.STD_LOGIC_UNSIGNED.ALL;
entity JTL is
Generic (voltage : real := 2.50;
          jitter : string := "middle delay");
Port ( signai_in : in STD_LOGIC;
          signal_out : out STD_LOGIC );
end JTL;

architecture Behavioral of JTL is
.....
begin
process
begin
variable t_out : time := 0 ps;
signal_out <= transport signal_in after t_out;
end process;
end Behavioral;
```

Fig. 2. Entity section and part of the architecture body for JTL

The specific readout delay for every cell for every value of the supply voltage can be represented by its average, minimum and maximum value. They are restricted to the three possible values because of the way of implementation timing jitter in the VHDL model. The discrete values of the solutions of the polynomial equations are included in the model. A corresponding part of the code showing this transformation is given in Fig. 3.

```
begin
process
variable t_out : time := 0 ps;
begin
wait on signal_in;
 if (choice = "the best delay") then
          if (voltage = 2.05) then
                    t_out := 24.03 ps;
          elsif (voltage = 2.10) then
                    t_out := 22.78 ps;
          elsif (voltage = 2.15) then
                    t_out := 22.01 ps;
          else
                    t_out := 40.0 ps;
          report "ERROR: There is not selected valid value of the bias supply
          voltage in SPLITTER_module" severity FAILURE;
                    end if;
          else
          report "ERROR: There is not selected choice in SPLITTER_module"
          severity FAILURE;
 end if;
```

Fig. 3. Part of the architecture body for JTL including the influence of the supply voltage and timing jitter

3. ARCHITECTURE AND SIMULATION RESULTS OF 8-BIT SHIFT REGISTER

We designed the architecture of shift register hierarchically using the developed VHDL models of the elementary cells (Fig. 4). The cells storing and carrying the data pulses are DFFs. Splitters and JTL are used to build a clock distribution network. The first seven clock pulses load the data into the DFF cells. The eight clock pulse produces the first data output pulse. The issue of this architecture is necessity to prevent timing violations between data and clock in the design process.

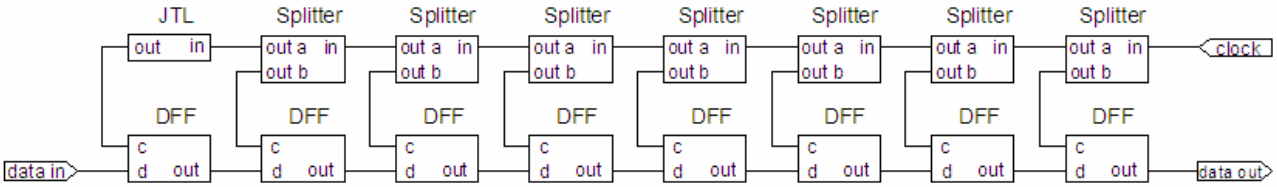


Fig. 4. Block scheme of an 8-bit shift register

This circuit has been analyzed in terms of timing jitter. All the simulations were done with bias supply voltage of 2.50mV at very high frequency of 10GHz. We split the simulations in the three groups. All the cells participate with their minimum delay in the first group, with their mean delay in the second group and with their maximum delay in the third group. The simulation results were achieved by shifting the input data with consideration of the clock period in different steps. The results are shown in Fig. 5, Fig. 6 and Fig. 7.

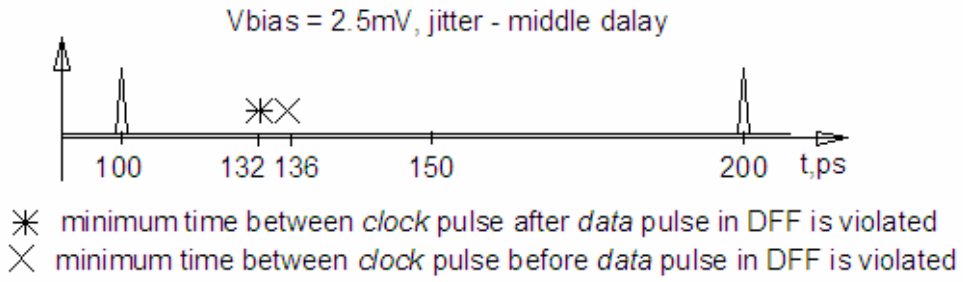


Fig. 5. Simulation result for 8-bit shift register showing the times of violation with middle delay of influence of jitter

The asterisk represents the violated timing constraint – minimum time between clock pulse after data pulse, the cross represents violated minimum time between clock pulse before data pulse in the DFF. It can be seen that the forbidden area is shifted left with minimum cell's delay and right with maximum cell's delay. If the input data pulses come at the forbidden area an error message is reported. The determination of these forbidden areas is important for practical realization of the circuit. The obtained information from the simulations is valuable for adjustment the shifting between data and clock signal in the circuit to get high performance.

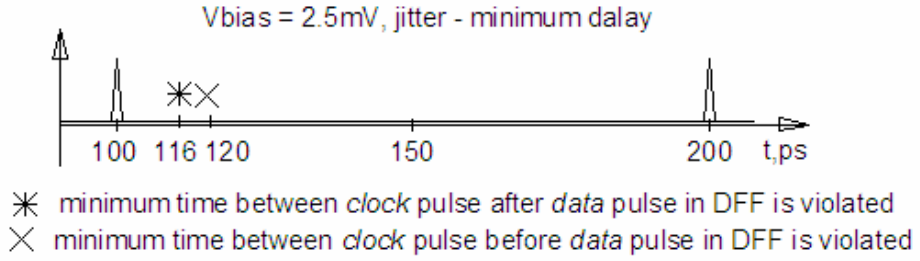


Fig. 6. Simulation result for 8-bit shift register showing the times of violation with minimum delay of influence of jitter

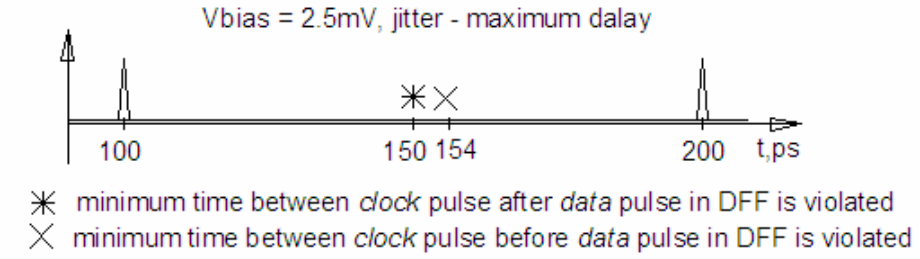


Fig. 7. Simulation result for 8-bit shift register showing the times of violation with maximum delay of influence of jitter

4. CONCLUSION

This article describes RSFQ shift register built by connection of VHDL modeled elementary cells. The capabilities of included specific features of RSFQ logic such as timing constraints and jitter make the developed VHDL models very useful for analyzing more complex circuit. This spreads the application of the implemented models in the different areas of the investigation the circuits. It allows getting useful information doing digital simulations in advance. The advantage of this work remains faster digital simulations.

5. ACKNOWLEDGEMENT

This work was supported by the German-Bulgarian Erasmus program.

REFERENCES

[1] P. Bunyk, K. Likharev and D. Zinoview, "RSFQ technology: physics and devices," in Int. Journal on High Speed Electronics and Systems, vol. 11, No. 1, pp. 257-306, 2001.
[2] J. Weber, Th. Ortlepp, "Untersuchung der Schalt- und Verzögerungszeiten einer supraleitenden Hochgeschwindigkeits elektronik im Pikosekunden-Bereich," master thesis, TU Ilmenau, 2003.
[3] K. K. Likharev and V. K. Semenov, “RSFQ logic/memory family “A new Josephson digital technology for sub-terahertz-clock-frequency digital systems”, in IEEE Trans. On Appl Superconductivity vol. 1, no. 1, pp. 3-28, 1991.
[4] Th. Ortlepp and F. Hermann Uhlmann, “Technology related timing jitter in superconducting electronic” IEEE Trans. on Applied Superconductivity, vol. 17, №. 2, June 2007.
[5] D. Perry, VHDL, 3rd Edition, McGraw-Hill 1998.
[6] M. Sato, M. Tanaka, K. Takagi and N. Takagi, “A verification method of pipelining processing behavior of single-flux-quantum circuits by equivalence checking of timed logic formula, Proc. 3ed S Superconducting SFQ VLSI Workshop (SSV 2010), Yokohama, Japan, pp. 81-85.

A MATHCAD BASED EXAMINATION OF BIFURCATIONS IN THIRD ORDER NONLINEAR CIRCUIT

Galina Petkova Cherneva

Department of Electrical Engineering, Higher School of Transport "Todor Kableshkov"
Geo Milev Str.158, 1574 Sofia, Bulgaria, phone: +359 888 498 478,
e-mail: cherneva@vtu.bg

Abstract: *The paper presents the transition from the limit cycle to the spiral and double-scroll attractor in a third order electrical circuit realization of the Chua Circuit family. A Mathcad program is used to show chaotic behavior. By switching the value of one the components of the circuit, we can see that the circuit's response starts by being periodic, but then it suffers several duplications of period and finally it becomes chaotic behavior.*

Keywords: chaos, nonlinear circuit, attractor, Chua's circuit, bifurcations

1. INTRODUCTION

Chaos is a robust phenomenon which appears in many aspects of modern science. Nonlinear electrical circuit is one of the fields where the chaos was found earliest. Electric circuits with a nonlinear resistor, which are characterized by a piecewise-linear V-I characteristic [1,2,3,5], have emerged as a simple yet powerful experimental and analytical tool in studying chaotic behavior in nonlinear dynamics. Chua's circuit [4,5] is a simple piecewise-linear third order electrical circuit.

Each member of the Chua Circuit family consists of four linear elements: one linear resistor, three linear dynamic elements (capacitors and/or inductors) and one nonlinear resistor characterized by a piecewise-linear V-I characteristic, with at least one segment having a negative slope. Spiral, double scroll, torus, and other attractors among other dynamics phenomena have been observed at different members of this family.

2. STATEMENT OF THE PROBLEM

The equations in dimensionless form governing Chua's circuit (Fig.1) are:

$$\left|\begin{array}{l} \dfrac{dx}{d\tau} = \alpha\left[y - x - f\left(x\right)\right] \\ \dfrac{dy}{d\tau} = x - y + z \\ \dfrac{dz}{d\tau} = -\beta y \end{array}\right. , \qquad (1)$$

where

$$x=\frac{u_1}{E},\quad y=\frac{u_2}{E},\quad z=\frac{i_L R}{E},\quad \tau=\frac{t}{RC_2},\quad \alpha=\frac{C_2}{C_1},\quad \beta=\frac{R^2C_2}{L}, \tag{2}$$

$$f(x)=\begin{cases} a_2x+a_1-a_2, & x>1 \\ a_1x, & -1\le x\le 1 \\ a_2x-a_1+a_2, & x<-1 \end{cases}, \tag{3}$$

$$a_1=RG_b,\quad a_2=RG_a. \tag{4}$$

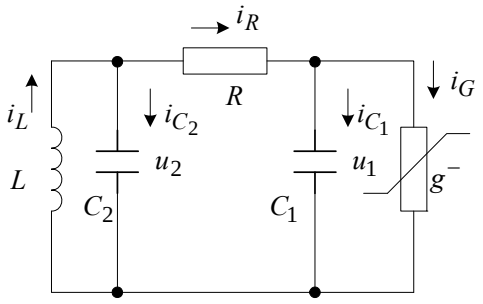


Fig. 1. Chua's circuit

Here G_a and G_b are the slopes of the segments and E denotes the breakpoint as shown in Fig. 2 and $g(u_1)$ is a nonlinear function defined by:

$$g(u_1)=G_bu_1+\frac{1}{2}(G_a-G_b)\left[|u_1+E|-|u_1-E|\right]. \tag{5}$$

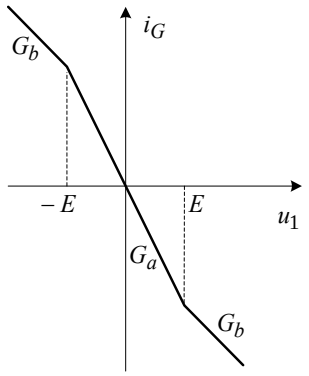


Fig. 2. V-I characteristic of Chua's Diode

The equilibrium points of the system (3), are [6]:

$$P_1(0,\ 0,\ 0) \text{ for } |u_1|\le E;$$

$$P_2\left(\frac{R(a_1-a_2)E}{Ra_1+1},\ 0,\ \frac{(a_2-a_1)E}{Ra_1+1}\right) \text{ for } u_1 < -E;$$

$$P_3\left(\frac{R(a_2-a_1)E}{Ra_1+1},\ 0,\ \frac{(a_1-a_2)E}{Ra_1+1}\right) \text{ for } u_1 > E. \quad (6)$$

The nature of each of the equilibrium points is determined by the three characteristic values of the Jacobian matrix of partial derivatives evaluated at the fixed point in question [6].

There are three cases to consider [6]:

- the three characteristic values are real and unequal;
- the three characteristic values are real and at least two are equal;
- the characteristic values are one real and one complex conjugate pair.

The values used for the simulation were the following [3]:

L = 7,14 mH, C_1 = 5 nF, C_2 = 50 nF, a_1 = 0,241 mS, a_2 = 0,807 mS, E = 1 V. (7)

In that case

$$\alpha = 10. \quad (8)$$

The parameter β, used as the control parameter, kept in the range of 9 to 15.

The used method includes finding a limit cycle (periodic trajectory) and chaotic behavior for a given set of parameters α, β, a_1, a_2.

3. EXPERIMENTAL RESULTS

The system (1) can be solved numerically in Mathcad using integration methods such as Runge-Kutta [7]. In order to visualize the chaotic behavior of the circuit, the program was set to run a total of three simulations, each of them with a different value of the parameter β (resistor R).

For the first simulation was used a value of $\beta = 10$.

Fig. 3 displays the experimental phase portraits of u_1 vs u_2 of the circuit. In this simulation chaos isn't present yet. What we can see here is a periodic response.

In that case two of the three characteristic values of the Jacobian matrix has magnitude unity (two characteristic values are zero and the third one is negative) and this is due to the fact that the trajectory is periodic.

For the second simulation the value of parameter β was set to $\beta = 12$. Fig. 4 displays the voltage in the two capacitors. The chaotic behavior of the circuit can be seen in the form of a spiral attractor.

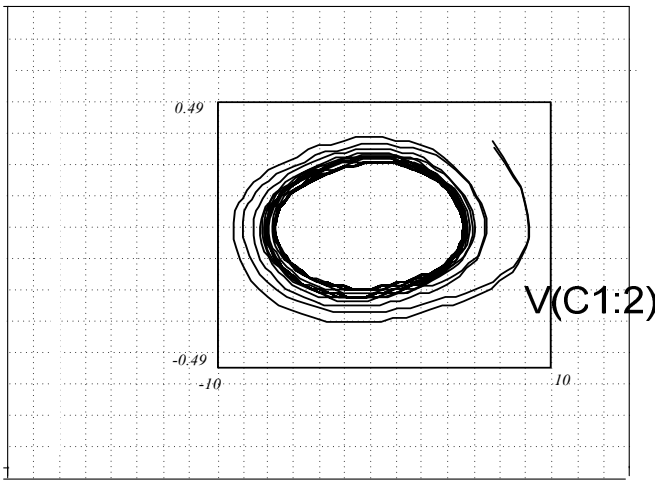


Fig. 3. u_1 vs u_2 with $\beta = 10$

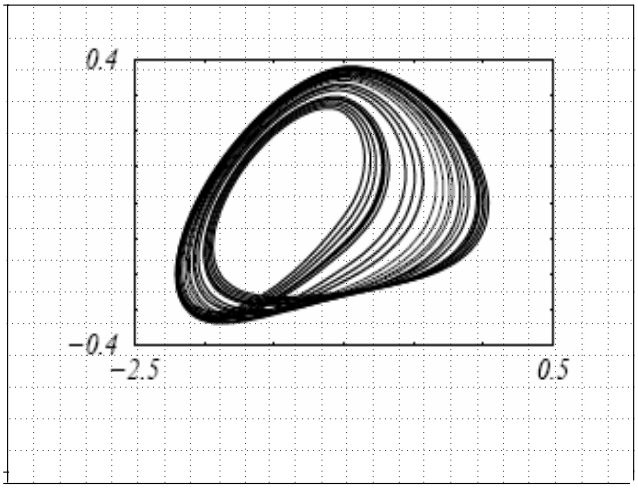


Fig. 4. u_1 vs u_2 with $\beta = 12$

For the third simulation was parameter β set to $\beta = 15$. The chaotic behavior of the circuit (Fig. 5) can be seen in the form of a double scroll.

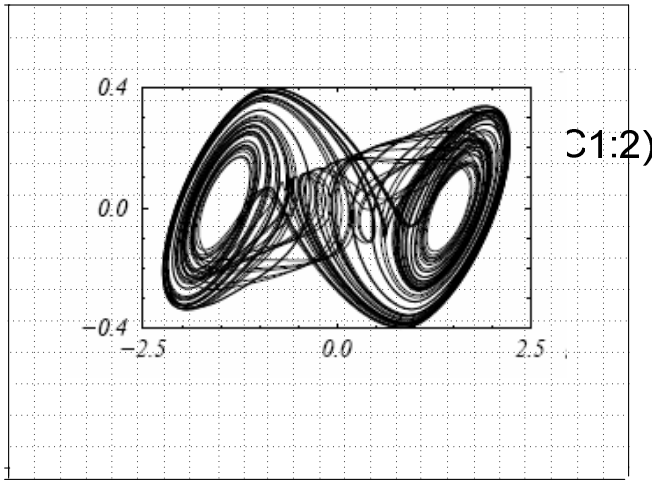


Fig. 5. u_1 vs u_2 with $\beta = 15$

4. CONCLUSIONS

This paper discusses the bifurcations and different chaotic attractors. performed in third order nonlinear circuit. For instance, varying the parameter β (the value of linear resistor R) the circuit goes from limit cycle to a spiral attractor passing through a sequence of period-doubling bifurcations. The results reported in this paper seem relevant in two aspects. First, the Chua's circuit exhibits a wealth of dynamical regimes. We have registered series from its dynamical variables with the parameter β as the bifurcation parameter. Second, the chaotic behavior of the circuit transition from the spiral attractor to the double scroll attractor as a bifurcation parameter is varied.

REFERENCES

[1] Torres, L., L. Aguirre: Inductorless Chua's circuit. Electronic letters, Vol. 36, No 23, 2000.

[2] Khalil H.K.: Nonlinear Systems. Prentice Hall . 2001

[3] L. O. Chua and G. N. Lin: Canonical realization of Chua's circuit family. IEEE Trans. Circuits Syst. I, vol. 37, 1990.

[4] M. P. Kennedy: Robust op amp realization of Chua's circuit. FREQUENZ, vol. 46, 1992.

[5] L. O. Chua and G. N. Lin: Intermittency in a piecewise-linear circuit. IEEE Trans. Circuits Syst. I. vol. 38, 1991.

[6] Cherneva G.,A.Andonov, P. Dimkin. Stability Analysis of States of Third Order Nonlinear Circuit. XXXIII IC-SPETO,Poland.26-29.05.2010.p.89-91

[7] Pritchard Ph. Mathcad: A Tool for Engineers and Scientists. N.J. 2007.

COMPUTER MODELING AND INVESTIGATION OF MEMS ACCELEROMETERS USING SPICE

Peter Rumenov Kanchev[*], Dobromir Georgiev Gaydazhiev[*], Elissaveta Dimitrova Gadjeva[**]

[*]Smartcom, 7th km, Tzarigradsko Chausee Blvd, 1784 Sofia, Bulgaria, phone: +359 29650650,
e-mail: petar_kanchev@smartcom.bg, dobromir_gaydajiev@smartcom.bg
[**]Department of Electronics and Electronic Technologies, Technical University of Sofia,
8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria, e-mail: egadjeva@tu-sofia.bg

Abstract: *A computer model of MicroElectroMechanical Systems (MEMS) is developed in the present paper, based on the equations which describe the behavior of the accelerometer as a function of the physical design variables. A parameterized PSpice model of the accelerometer system is developed. The design variables and the design equations, the geometrical and functional constraints are introduced in the model using the PARAMETER statement. Using computer simulation in DC, frequency- and time domain, the basic characteristics of the accelerometer are investigated. The sensitivity, the minimal detectable acceleration (noise), the maximum detectable acceleration (detecting range), as well the bandwidth, are calculated. Computer electrostatic accelerometer models are constructed and accelerometer performance is evaluated for the open- and closed loop accelerometer operation. The simulated results are compared with the measured results for the MEMS accelerometer.*

Keywords: MEMS Accelerometers, Parameterized *PSpice* model,
Open- and closed loop accelerometer system

1. INTRODUCTION

With the rapid development of MicroElectroMechanical Systems (MEMS) technology, there is a demand of computer-aided tools for design, modeling and simulation of MEMS devices [1-5]. MEMS based accelerometers are widely used in automotive, robotics and other industries. Recently, synthesis tools are developed considering the mechanical structure of the accelerometer, design procedure at the system level and the force feedback control which is used to stabilize the system and increase the maximum detectable acceleration. The development of computer model of a closed loop MEMS accelerometer is a very important consideration in the design of such devices [1].

Computer models are developed in the present paper, based on the equations which describe the behavior of the accelerometer as a function of the physical design variables. They are implemented in the graphical editor *Cadence Capture*. A parameterized *PSpice* model of the accelerometer system is developed. The design variables and the design equations, the geometrical and functional constraints are introduced in the model using the *PARAMETER* statement. The investigations of the accelerometer system are realized using the *Cadence PSpice* circuit simulator. Using computer simulation in DC, frequency- and time domain, the basic characteristics of

the accelerometer are investigated. The sensitivity, defined as a ratio of the output voltage over the input acceleration, is obtained. The minimal detectable acceleration (noise), the maximum detectable acceleration (detecting range), as well as the band-width, are calculated.

Computer electrostatic accelerometer models are constructed and accelerometer performance is evaluated for the open- and closed loop accelerometer operation. The equations describing the blocks of the systems, are defined. The displacement, the time delay and the setting time are analyzed and compared for the open- and closed loop accelerometer systems. Based on DC sweep, the sensitivity and the range of the accelerometer system for open- and closed loop operation is obtained. Using AC sweep, the frequency range is analyzed and the stability is investigated using the Ny-quist criterion. The simulated results are compared with the measured results for the MEMS accelerometer.

2. COMPUTER MODEL REALIZED IN THE ORCAD CAPTURE ENVIRONMENT

The used accelerometer is a prototype designed in PolyMUMPs technology with parameters shown in Table 1. Parameterized models are developed in the OrCAD Capture environment for the open- and closed loop systems. The block scheme in Fig. 1 is represented in OrCAD Capture as shown in Fig. 2.

Table 1. Parameters of the used sensor for acceleration (technology PolyMUMPs)

Parameter	Description	Value
m	proof mass	3.573×10^{-10} kg
B	damping coefficient	1.3312×10^{-7}
k	spring constant	0.576
g0	gap between fingers	3μm
sensitivity		169.5×10^{-18} F/g
n	number of fingers	40
aMAX	range	150 g
Wr	resonant frequency	6.4 kHz
C0	capacitance in steady state	45.784×10^{-15}

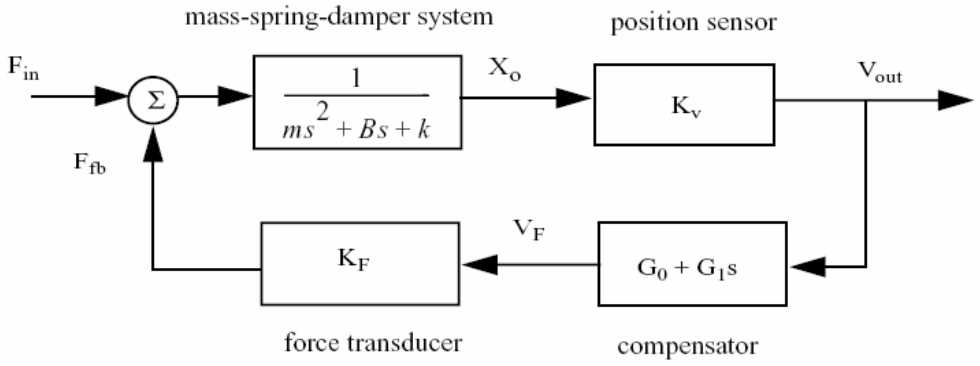


Fig. 1. Analog feedback loop

The closed loop gain is in the form:

$$H(s) = \frac{V_{out}(s)}{F_{in}(s)} = \frac{K_V}{ms^2 + (B + G_1 K_F K_V)s + (K_F K_V G_0 + k)} \tag{1}$$

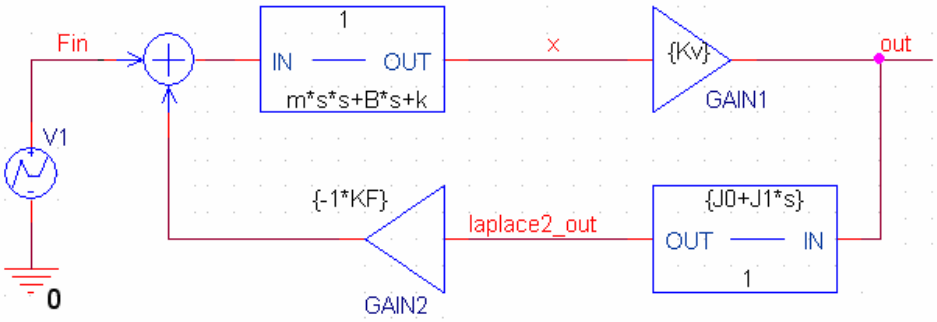


Fig. 2. Representation of the closed loop system in OrCAD Capture

The expressions for G_0, G_1, K_v and K_F are in the form:

$$G_0 = \frac{ma_{in} - kx_{0,steady}}{K_F K_V x_{0,steady}} \; ; \; K_F = \frac{2V_{c,dr}C_F}{g_f}$$

$$K_V = \frac{V_m}{g_0}\left(\frac{2C_0}{2C_0 + C_{para}}\right) \; ; \; G_1 = \frac{-B + 2m\sqrt{\dfrac{a_{in}}{x_{0,steady}}}}{K_F K_V} \tag{2}$$

The statement PARAMETERS describes all parameters and system block descriptions (Fig. 4).

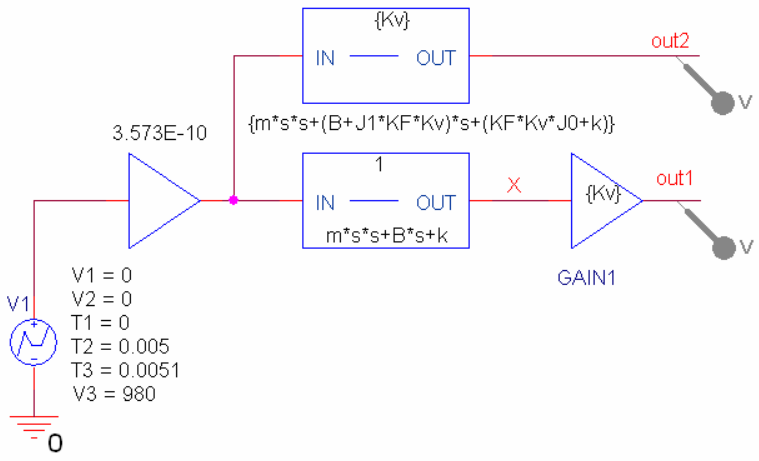


Fig. 3. PSpice model of the open- and closed loop systems

PARAMETERS:
m = 3.573E-10
B = 1.3312E-7
k = 0.576
Kv = {(Vm/g0)*2*C0/(2*C0+Cpara)}
Vm = 1V
g0 = 3um
C0 = 45.784E-15
Cpara = 2pF
J0 = {(m*ain-k*x0_steady)/(KF*Kv*x0_steady)}
ain = {300*G}
G = 9.812865328
KF = {2*Vc_dr*CF/gf}
Vc_dr = 1V
CF = 45.784E-15
gf = 3um
J1 = {(-1*B+2*m*sqrt(ain/x0_steady))/(KF*Kv)}
x0_steady = 1u

Fig. 4. The parameters and expressions describing the system

The equations (2) are introduced using *PARAMETERS* statement in *Cadence Capture* in the form:

J0 = {(m*ain-k*x0_steady)/(KF*Kv*x0_steady)}
Kv={(Vm/g0)*2*C0/(2*C0+Cpara)}
KF={2*Vc_dr*CF/gf}
J1={(-1*B+2*m*sqrt(ain/x0_steady))/(KF*Kv)}

3. TIME DOMAIN SIMULATION

As a result of the time domain simulation, the delay and the setting time of the open- and closed loop systems are compared. The input data is the acceleration in the defined range for both systems (100g). Time Domain (Transient) simulation is defined where T_{max}=30ms and minimal step size is 100ns. Depending on the accelerometer application, the percentage deviation from the steady-state, can vary. For the considered example the decrease is 0.5dB (about 5%). It is seen from the simulation results that the steady-state for the open loop system is reached by output voltage 8.84mV. The steady-state for the closed loop system is reached by output voltage 4.86mV as shown in Fig. 6. The delay of the input signal is 5ms. The obtained delay for the open loop system is 26.8μs and for the closed loop system is 27.4μs (Fig. 7).

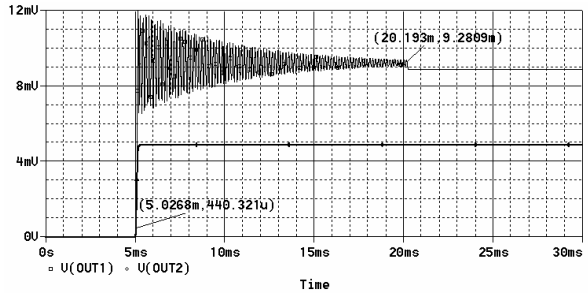


Fig. 5. Setting time for the open loop system (OUT1)

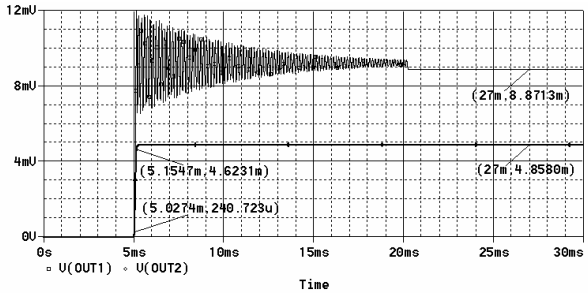


Fig. 6. Setting time for the closed loop system (OUT2)

The obtained result for the setting time is 15.166ms for the open loop system (Fig. 5) and 127.268μs for the closed loop system (Fig. 6). In practice, the applied feedback does not influence on the system delay. The advantage is in the setting time. The peaks in the output voltage (proportional to the mass displacement) are omitted and the closed loop system is settled about 120 times faster.

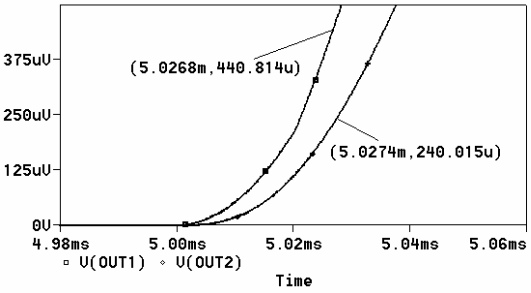


Fig. 7.Time delay determination for the open- (OUT1) and closed loop system (OUT2)

4. SENSITIVITY INVESTIGATION

The sensitivity is defined as variation of the output voltage depending of the applied input acceleration. The results are shown in Fig. 8. The slope defines the sensitivity of the systems. The higher slope defines a higher sensitivity. A DC simulation is performed in the range (-1500V, +1500V). The limits are selected to cover the range of the open loop system 150G (~1500m/s^2). It is seen from Fig. 8 that the dependence is linear, corresponding to $9\mu V/ms^2 = 90\mu V/G$ for the open loop system and $5\mu V/ms^2 = 50\mu V/G$ for the closed loop system. It is concluded from the DC analysis that the application of the feedback in the accelerometer decreases the sensitivity in the same rate as the increasing of the range.

5. ACCELEROMETER RANGE

The accelerometer range is defined as the maximal allowed value for the applied acceleration before proof mass to reach the limiting structures or before to be damaged. The design rules in PolyMUMPs technology allow a distance between two structures of minimum $2\mu m$.The proof mass displacement is obtained at the node *x* in the accelerometer closed loop system (Fig. 2). A DC analysis is used with the values for the acceleration in the range (-3500 m/s^2, 3500m/s^2). The feedback is designed in a way to extends the accelerometer range from 150G to 300G.

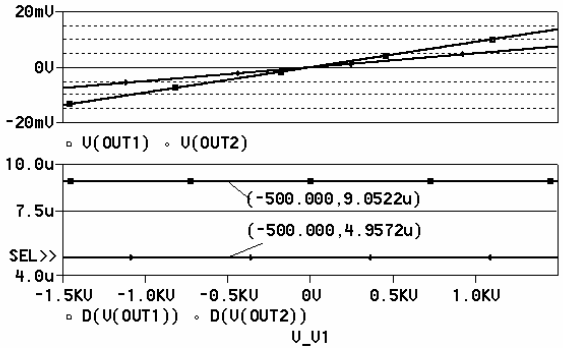


Fig. 8. Sensitivity of the open- (OUT1) and closed loop system (OUT2)

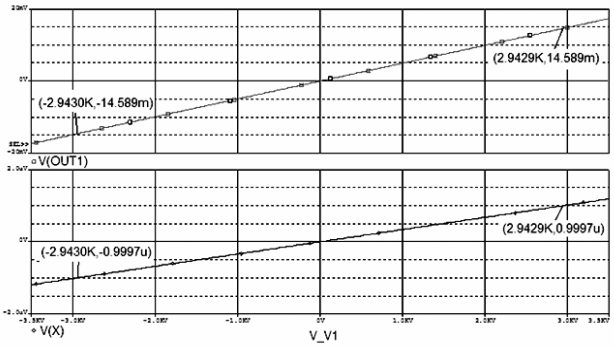


Fig.9. The accelerometer range for the closed loop system

The acceleration corresponding to proof mass displacement is 2943m/s^2. It is seen from the results in Fig. 9 that this acceleration corresponds to a displacement of $0.9997\mu m$. The DC analysis demonstrates that the application of the feedback in the accelerometer leads to increasing the maximum detectable acceleration.

6. FREQUENCY RANGE

The frequency range of the accelerometer is defined as a range in which its characteristic remains nearly constant according to the defined criterion. Different criteria

are used depending on the sensor application as well as on the required accuracy. For the considered case the requirement for the maximal deviation of 5% is defined.

AC analysis is defined in the range (1Hz ÷ 100kHz) by logarithmic scale and 500 Pts/decade. The output voltages are given for the open- and closed loop systems and the frequency ranges are obtained: 1.4kHz for the open loop system and 2kHz for the closed loop system (Fig. 10). The application of the feedback increases the frequency range of the system by approximately 600Hz.

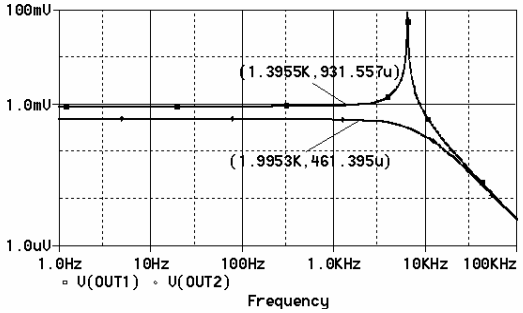


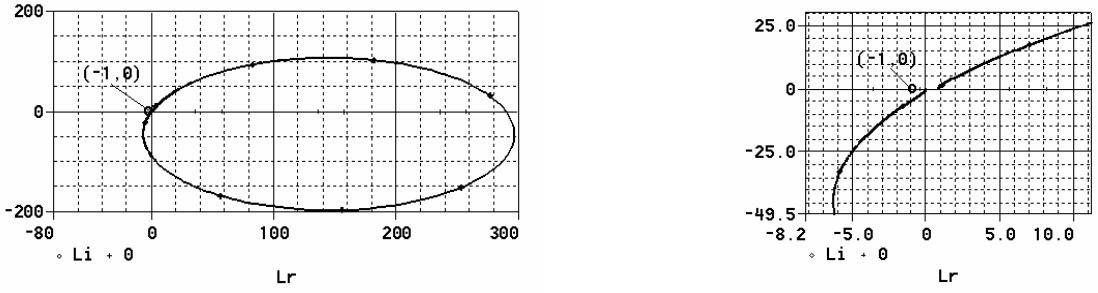
Fig. 10. Frequency range of the open- and closed loop systems

7. STABILITY INVESTIGATION

According to the Nyquist stability criterion, the dependence Im(L) = f (Re(L)) is created, where L is the closed loop gain of the system. The Nyquist plot is created in the graphical analyzer *Probe* using the following macros:

```
L = -V(GAIN2_OUT)/V(LAPLACE_1_IN)
Lr = R(L)
Li=Img(L)
```

The result is shown in Fig. 11a. A zoomed plot around the point (-1,0) is presented in Fig. 11b which confirms the stability of the system.



a) b)

Fig. 11. Stability investigation using Nyquist plot

The influence of the applied feedback on the system can be investigated from the *PSpice* simulations and from the obtained data for the designed accelerometer prototype. The feedback extends the working range of the sensor and decreases significantly the setting time of the system. The peaks of the proof mass displacement are deleted when input acceleration is applied. The application of the feedback increases the maximum detectable acceleration and decreases its sensitivity in the same rate. The feedback extends the frequency range of the system.

8. CONCLUSIONS

Analytical description of the accelerometer behavior is derived and its basic parameters are defined. A synthesis of the feedback of the accelerometer in PolyMUMPs technology is performed in order to increase its working range. A parameterized *PSpice* model of the accelerometer system is developed. Based on computer simulation, the open- and closed loop systems are compared according to the following criteria: setting time of the system, sensor sensitivity, working range of the sensor and the frequency range of the sensor.

9. ACKNOWLEDGEMENT

The investigations are in the framework of DDOKF02/1/13.01.2010 project.

REFERENCES

[1] C. Lewis and M. Kraft, “Simulation of a Micromachined Digital Accelerometer in SIMULINK and PSPICE”, Proceedings of UKACC International Conference on Control, Exeter, UK, Sept. 1996, vol. 1, pp. 205-209

[2] Y. Zhou, “Layout Synthesis of Accelerometers , Department of Electrical and Computer Engineering Carnegie Mellon University, August 1998, http://www.ece.cmu.edu/~mems /pubs/pdfs/ece/ms_thesis/0051_zhou-1998.pdf

[3] G. Zhang, Design and Simulation of A CMOS-MEMS Accelerometer, Tsinghua University, 1994 http://notes.ump.edu.my/fkee/e-Books/MEMS/Design%20And%20Simulation%20Of%20A%20Cmos-Mems%20Accelerometer.pdf.

[4] About MEMS and Nanotechnology , http://www.mems-exchange.org

[5] M. Andrejasic, “MEMS Accelerometers”, University of Ljubljana, Faculty of Mathematics and Physics, Department of Physics, Seminar 2008, http://mafija.fmf.uni-lj.si/seminar/files/2007_2008/MEMS_accelerometers-koncna.pdf

PARALLEL FORM SIGMA-DELTA MODULATOR HIGH ORDER LOOPFILTER REPRESENTATION IN Z AND S DOMAIN

Georgi Tsenov, Valeri Mladenov

Department of Theoretical Electrical Engineering, Faculty of Automatics
Technical University of Sofia 8, Kliment Ohridski St, Sofia-1000
BULGARIA, phone: 9652394, e-mail: gogotzenov@tu-sofia.bg

Abstract: *In this paper is presented the conversion of an arbitrary high order sigma-delta modulator Z-domain loopfilter transfer function into parallel form low order subsections in both S and Z domains. The benefit of this type of filter form representation in comparison with the other widespread types is the existence of theory, which can validate sigma-delta modulator's stability. The S domain representation is suitable for the element value calculation stage that is done according to the utilized transfer function when direct analogue parallel loopfilter implementation is performed.*

Keywords: Sigma-delta modulators; Stability; Analog-to-digital conversion

1. INTRODUCTION

Sigma-Delta modulators are the standard for analog to digital conversion nowadays (ADC). When using high oversampling ratios sigma-delta modulators (SDM) can achieve very high signal to noise ratio (SNR). They shape the noise and push it to frequencies higher than the operational band of interest. Thanks to its simplicity, single bit code shaping SDM are of greatest interest, because they're performance is influenced only by the loopfilter's transfer function and the modulator's oversampling ratio (OSR) [1],[2],[3],[4],[5].

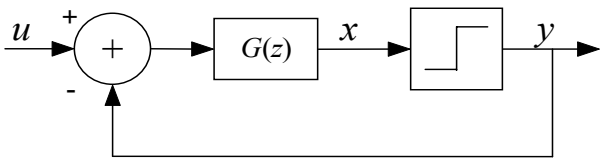


Fig. 1. Basic sigma delta modulator structure

Despite this in practice the modulator's maximal stable DC input signal range is determined mostly by simulations, which also leave a zone of uncertainty. Furthermore a lot of engineers experiment with the loopfilter coefficients in order to achieve more SNR, while keeping the modulator in reasonable stability ranges. All of the realistic loopfilter transfer functions have the poles grouped into a complex conjugate pairs and one real pole when having odd modulator order. There is a theory [6] that allows us to have an approximate estimate on the modulator stability ranges based on loopfilter parallel decomposition, which is very useful for the SDM design.

The paper is organized as follows. In the next section we present the parallel decomposition technique of arbitrary sigma-delta modulator loopfilter in Z-domain. In section three we present the derivation of S-domain parallel filter form representation. Finally, the conclusion remarks are given in the last section.

2. PARALLEL DECOMPOSITION TECHNIQUE IN Z-DOMAIN

Authors in [6] consider Nth order modulator with a loop filter transfer function of the form

$$G(z) = \frac{a_1 z^{-1} + ... + a_N z^{-N}}{1 + d_1 z^{-1} + d_2 z^{-2} + ... + d_N z^{-N}} \tag{1}$$

In the general case the loop filter transfer function have complex conjugated roots. Without loss of generality we will consider only one pair of complex conjugated roots. In this case G(z) becomes

$$G(z) = \frac{b_1 z^{-1}}{1 - \lambda_1 z^{-1}} + ... + G_2(z) = \frac{b_1 z^{-1}}{1 - \lambda_1 z^{-1}} + ... \frac{B_{N-1} z^{-1} + B_N z^{-2}}{1 - d_1 z^{-1} - d_2 z^{-2}} \tag{2}$$

where the coefficients *bi*, $i = 1,2,\ldots, N$ of the fractional components can be found easily using the well known formula

$$b_i = \frac{(1 - \lambda_i z^{-1})}{z^{-1}} G(z)\Big|_{z=\lambda_i} \tag{3}$$

The denominator of the last part of (2) has a complex conjugated pair of roots and therefore (2) becomes

$$G(z) = \frac{b_1 z^{-1}}{1 - \lambda_1 z^{-1}} + ... + \frac{b_{N-1} z^{-1}}{1 - \lambda_{N-1} z^{-1}} + \frac{b_N z^{-1}}{1 - \lambda_N z^{-1}} \tag{4}$$

where

$$\lambda_{N-1} = \alpha + j\beta,\ \lambda_N = \alpha - j\beta$$

$$b_{N-1} = \delta - j\gamma,\ b_N = \delta + j\gamma$$

i.e. λ_{N-1}, λ_N and b_{N-1}, b_N are complex conjugated numbers.

Because of this the parallel presentation given in Fig. 2 of third order modulator is used. The values of the last two blocks are complex, but the output signal of these two blocks is real. They correspond to a second order SDM with complex conjugated poles of the loop filter transfer function G(z). Both signals x_2 and x_3 are complex conjugated (and x2 is real), namely

$$x_2(k+1) = m(k+1) + jn(k+1)$$
$$x_3(k+1) = m(k+1) - jn(k+1)$$

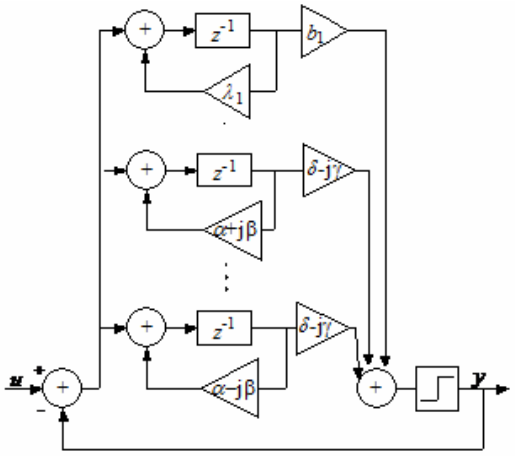


Fig. 2. Block diagram of third order SDM with parallel loopfilter form

Because of this the input of the quantizer is real i.e.

$$(\delta - j\gamma)x_2(k) + (\delta + j\gamma)x_3(k) = 2\delta m(k) + 2\gamma n(k)$$

The modulator could be considered as three first order modulators interacting only through the quantizer function. The connected signals with two modulators are complex, but the input and output signals (u and y) are the "true" signals of the modulator. Both modulators work cooperative, because their signals are conjugated. These modulators do not exist in the real SDM and they are introduced to help the analysis of the behavior of the whole system.

The benefit of this modulator representation is because we can determine whenever the modulator is stable or not by this criterion

$$\frac{(2-\lambda_1)}{\lambda_1}\frac{b_1}{(\lambda_1-1)} > -\sum_{i=2}^{N-2}\frac{|b_i|}{\lambda_i - 1} + \frac{2\,|\,\delta(1-\alpha)+\gamma\beta\,|}{(1-\alpha)^2+\beta^2} \tag{5}$$

and additionally we can also determine the maximal range of input signal ensuring the stability expressed by Δu (DeltaU)

$$\Delta u < \frac{\sum_{i=2}^{N-2}\frac{|b_i|}{\lambda_i - 1} - \frac{2|\delta(1-\alpha)+\gamma\beta|}{(1-\alpha)^2+\beta^2} + \frac{b_1(2-\lambda_1)}{\lambda_1(\lambda_1-1)}}{\frac{b_1}{\lambda_1-1} - \sum_{i=2}^{N-2}\frac{|b_i|}{\lambda_i-1} + \frac{2|\delta(1-\alpha)+\gamma\beta|}{(1-\alpha)^2+\beta^2}} \tag{6}$$

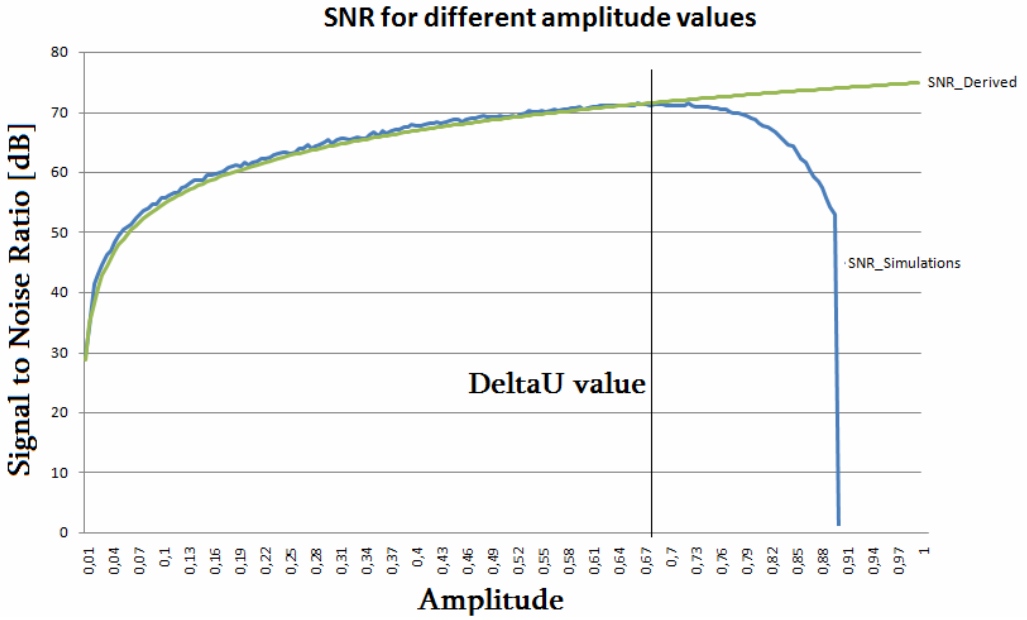


Fig. 3. Relation between DeltaU and the value of the input signal

If we want to make a discrete electronic implementation and have for example loopfilter transfer function:

$$\mathrm{H(z)}=\frac{0.7998z^{-1}-1.3414z^{-2}+0.5522z^{-3}}{1-2.999z^{-1}+2.999z^{-2}-z^{-3}} \tag{7}$$

Can be converted with parallel decomposition into:

$$H(z)=\frac{10.547367000026572z^{-1}}{1-1.0000000000000054z^{-1}}+\frac{(-4.873783500013301-3.9929128994828591i)z^{-1}}{1-(0.999499999999975+0.031618823507504i)z^{-1}}+\frac{(-4.873783500013301+3.9929128994828591i)z^{-1}}{1-(0.999499999999975-0.031618823507504i)z^{-1}}$$

In this case the conjugate pole pairs form one second order section, so for third order modulator we have one first and one second order section.

3. PARALLEL DECOMPOSITION TECHNIQUE IN S-DOMAIN

If we want to make a continuous time electronic implementation we must use S domain transfer functions and build the filter blocks with Active filters. In some countries it is common the S operator to be typed with P. We will use this notation.

If we use the bilinear transform:

$$z=\frac{1+(T/2)p}{1-(T/2)p}\longleftrightarrow p=\frac{2}{T}.\frac{z-1}{z+1} \tag{8}$$

We get for the first order section:

$$H(z)=\frac{b_1z^{-1}}{1-\lambda_1z^{-1}}=\frac{b}{z-\lambda}$$

$$H(p)=\frac{b}{\frac{1+(T/2)p}{1-(T/2)p}-\lambda}=\frac{b-(T/2)bp}{1+(T/2)p-\lambda\left[1-(T/2)p\right]}=$$

$$=\frac{b-(T/2)bp}{p\left((T/2)+\lambda(T/2)\right)+(1-\lambda)}=\frac{\frac{b-(T/2)bp}{1-\lambda}}{\frac{(T/2)(1+\lambda)p}{1-\lambda}+1}= \tag{9}$$

$$=\frac{\left(\frac{b}{1-\lambda}\right)-\left(\frac{Tb}{2(1-\lambda)}\right)p}{\left(\frac{T(1+\lambda)}{2(1-\lambda)}\right)p+1}==\frac{k_1-k_2p}{k_3p+1}$$

$$k_1=\frac{b}{1-\lambda}\qquad k_2=\frac{Tb}{2(1-\lambda)}\qquad k_3=\frac{T(1+\lambda)}{2(1-\lambda)}$$

which is implementable with a first order allpass filter .

Another alterneate mathetical derivation leads to:

$$H(z)=\frac{\left(\frac{2b}{2(1-\lambda)}\right)-\left(\frac{Tb}{2(1-\lambda)}\right)p}{\left(\frac{T(1+\lambda)}{2(1-\lambda)}\right)p+\frac{2(1-\lambda)}{2(1-\lambda)}}=\frac{2b-Tbp}{T(1+\lambda)p+2(1-\lambda)}\frac{1/2b}{1/2b}=\frac{1-\frac{T}{2}p}{\frac{T}{2b}(1+\lambda)p+\frac{(1-\lambda)}{b}} \tag{10}$$

For the complex conjugate pair we have:

$H(z)=\frac{b_{N-1}z^{-1}}{1-\lambda_{N-1}z^{-1}}+\frac{b_Nz^{-1}}{1-\lambda_Nz^{-1}}$, and we can consolidate them into one 2nd order section:

$$H(z)=\frac{(\gamma+\Delta j)z^{-1}}{1-(\alpha-\beta j)z^{-1}}+\frac{(\gamma-\Delta j)z^{-1}}{1-(\alpha+\beta j)z^{-1}}=$$

$$=\frac{(\gamma+\Delta j)(1-\alpha z^{-1}-\beta jz^{-1})z^{-1}+(\gamma-\Delta j)(1-\alpha z^{-1}+\beta jz^{-1})z^{-1}}{(1-\alpha z^{-1}+\beta jz^{-1})(1-\alpha z^{-1}-\beta jz^{-1})}=\frac{2\gamma z^{-1}+2(\beta\Delta-\gamma\alpha)z^{-2}}{1-2\alpha z^{-1}+(\alpha^2+\beta^2)z^{-2}} \tag{11}$$

Here we have the consolidated Z domain loopfilter form into:

$$H(z)=\frac{2\gamma z-2(\gamma\alpha-\beta\Delta)}{z^2-2\alpha z+(\alpha^2+\beta^2)}$$

Then again using bilinear transform and substitution Z with $z=\dfrac{1+(T/2)p}{1-(T/2)p}$ we obtain that: $H(z)=\dfrac{2\gamma z-2\gamma\alpha+2\beta\Delta}{z^2-2\alpha z+(\alpha^2+\beta^2)}\rightarrow$

$$H(p)=\frac{2\gamma\dfrac{1+(T/2)p}{1-(T/2)p}-2\gamma\alpha+2\beta\Delta}{\left(\dfrac{1+(T/2)p}{1-(T/2)p}\right)^2-2\alpha\dfrac{1+(T/2)p}{1-(T/2)p}+(\alpha^2+\beta^2)}=$$

$$=\frac{2\gamma\dfrac{1+(T/2)p}{1-(T/2)p}+(2\beta\Delta-2\gamma\alpha)\dfrac{1-(T/2)p}{1-(T/2)p}}{\left(\dfrac{1+(T/2)p}{1-(T/2)p}\right)^2-2\alpha\dfrac{\left(1+(T/2)p\right)\left(1-(T/2)p\right)}{\left(1-(T/2)p\right)^2}+\dfrac{(\alpha^2+\beta^2)\left(1-(T/2)p\right)^2}{\left(1-(T/2)p\right)^2}}=$$

$$=\frac{\dfrac{2\gamma+2\gamma(T/2)p-2\gamma\alpha+2\gamma\alpha(T/2)p+2\beta\Delta-2\beta\Delta(T/2)p}{1-(T/2)p}}{\dfrac{1+Tp+(T^2/4)p^2-2\alpha+2\alpha(T^2/4)p^2+(\alpha^2+\beta^2)(1-Tp+(T^2/4)p^2)}{\left(1-(T/2)p\right)^2}}= \tag{12}$$

$$=\frac{\left(2\gamma+2\gamma(T/2)p-2\gamma\alpha+2\gamma\alpha(T/2)p+2\beta\Delta-2\beta\Delta(T/2)p\right)\left(1-(T/2)p\right)}{1+Tp+(T^2/4)p^2-2\alpha+2\alpha(T^2/4)p^2+\alpha^2-\alpha^2Tp+\alpha^2(T^2/4)p^2+\beta^2-\beta^2Tp+\beta^2(T^2/4)p^2}=$$

$$=\frac{2\gamma+2\gamma(T/2)p-2\gamma\alpha+2\gamma\alpha(T/2)p+2\beta\Delta-2\beta\Delta(T/2)p-2\gamma(T/2)p-2\gamma(T^2/4)p^2+2\gamma\alpha(T/2)p-2\gamma\alpha(T^2/4)p^2-2\beta\Delta(T/2)p+2\beta\Delta(T^2/4)p^2}{1+Tp+(T^2/4)p^2-2\alpha+2\alpha(T^2/4)p^2+\alpha^2-\alpha^2Tp+\alpha^2(T^2/4)p^2+\beta^2-\beta^2Tp+\beta^2(T^2/4)p^2}=$$

$$H(p)=\frac{(-\gamma\alpha-\gamma+\beta\Delta)(T^2/2)p^2+(\gamma\alpha-\beta\Delta)2Tp+(2\gamma-2\gamma\alpha+2\beta\Delta)}{(2\alpha+1+\alpha^2+\beta^2)(T^2/4)p^2+(1-\alpha^2-\beta^2)Tp+(1-2\alpha+\alpha^2+\beta^2)}$$

This transfer function is implementable with utilization of Kervin-Huelsman-Newcomb, Tow-Thomson, Fleischer-Tow or Fleischer-Laker biquad building blocks. Here for the biquad building blocks we have as shown on Fig. 4:

$$H(s)=\frac{a_2s^2+a_1s+a_0}{s^2+b_1s+b_0}=\frac{a_2s^2+a_1s+a_0}{s^2+\dfrac{\omega_0}{Q}s+\omega_0^2} \tag{13}$$

where $\omega_0^2=\sigma_p^2+\omega_p^2$ $\dfrac{\omega_0}{Q}=-2\sigma_p$ and $H(s)=\dfrac{a_2s^2+a_1s+a_0}{(s-s_p)(s-s_p^*)}$.

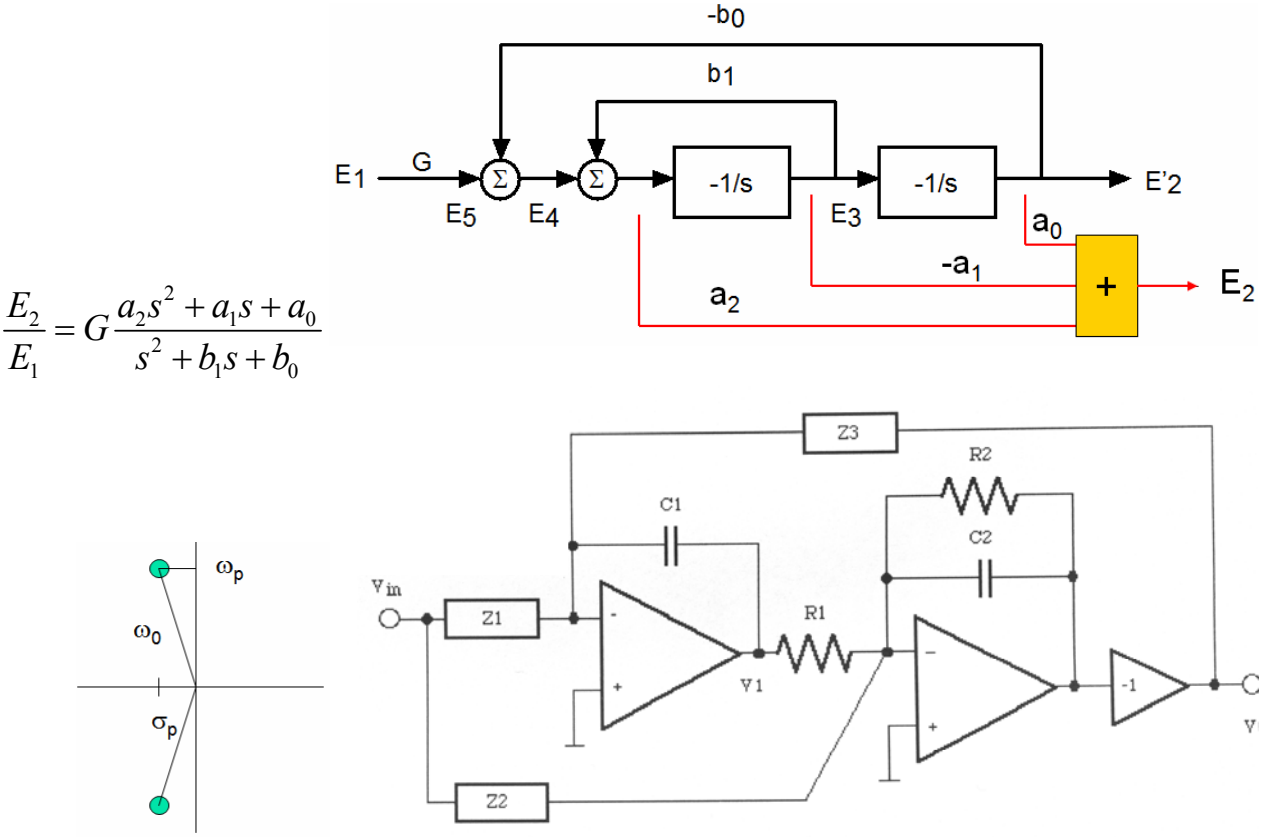


Fig. 4. Example of possible continuous time implementation on the biquad section

4. CONCLUSION

The benefits of Implementing SDM with parallel loopfilter form allows the creation of experimental implementations PCBs with adjustable coefficients, that can be used to design stable high performance modulators. The derivations in this paper are suitable for both discrete and continuous time implementations of parallel loopfilter form SDMs.

REFERENCES

[1] R. Schreier and G. C. Temes, Understanding Delta-Sigma Data Converters. New Jersey: John Wilet & Sons, 2005.
[2] S. R. Norsworthy, R. Schreier, and G. C. Temes, “Delta-Sigma Data Converters : Theory, Design, and Simulation”. Wiley-IEEE Press, 1996.
[3] Arthur H. M. Van Roermund (Editor),Herman Casier (Editor), Michiel Steyaert (Editor),Analog Circuit Design: High-Speed A-D Converters, Automotive Electronics And Ultra-Low Power Wireless, Springer ,2006
[4] Yves Geerts,Michiel Steyaert, and Willy M. C. Sansen, Design of multi-bit delta-sigma A/D converters, Kluwer Academic Publishers, 2002
[5] Aziz, P.M., Sorensen, H.V., Vander Spiegel J. “An overview of sigma-delta converters” Signal Processing Magazine, IEEE Volume: 13, Issue: 1, pp. 61 – 84. Jan. 1996
[6] V. Mladenov, H. Hegt, and A. v. Roermund, "On the Stability Analysis of Sigma-Delta Modulators," 16th European Conference on Circuit Theory and Design ECCTD 2003, Cracow, Poland, pp. I-97-I-100, 2003.

GENERAL SLIDING MODE CONTROL FOR SWITCHED DC/DC BUCK CONVERTER

Ivan Mitkov Trushev

Department of Theoretical Electrical Engineering, Faculty of Automatics,
Technical University of Sofia, 8 Kliment Ohridski St, Sofia-1000
BULGARIA, phone: +35929653319, e-mail: ivant@tu-sofia.bg

Abstract: *Switch-mode power supplies represent a particular class of variable structure systems. Thus, they can take advantage of non-linear control techniques developed for this class of systems. In this paper so called general sliding mode control is reviewed and its application to switched mode power DC/DC buck converter is discussed. It is shown that sliding mode approach can give good results in terms of robustness toward large supply and load variations, while maintaining good dynamic response. Control refinements, like current limitation, constant switched frequency and output voltage steady-state error cancellation are also discussed.*

Keywords: switched DC/DC buck converter, sliding mode control (SMC), variable structure systems, sliding surface (line)

1. INTRODUCTION

Battery operated portable systems such as notebooks, mobile phones etc. require signals processing functions at minimum power consumption with a strong variation in speed requirements. In digital signal processing a clear trade-off can be made between the speed and supply voltage, and supply voltages as low as 0.5V can be expected. From other side the supply voltage requirements for analog circuitry are often dictated by external signal sources and loads. The difference at the requiring levels of the supply voltage for the digital and analog circuitry imposes utilization of switched DC/DC converters. They are essential for efficient conversion of the battery voltage to various supply voltages, needed to perform every function with minimum power drain. An important goal for the converters is achieving of high efficiency.

Switched DC/DC converters are time-varying systems because of their specific switching action [1]. They are nonlinear systems with unstable parameters and inevitable and significant perturbations during operation. Therefore, to stabilize their parameters and to reduce the negative effects of switching the chosen method of control is essential.

The control in high performance DC/DC converters requires not only to ensure system stability but also to achieve a rapid response to sudden changes of the load, to achieve good regulation. There are basically two ways to control the switched DC/DC converters - linear and nonlinear control. Nonlinear control often produces better results but sometimes increases the complexity of the practical implementation of the scheme.

Switch-mode power supplies represent a particular class of variable structure systems. Thus, they can take advantage of non-linear control techniques developed for this class of systems. In this paper so called general sliding mode control is reviewed and its application to switched mode power DC/DC buck converter is discussed. It is shown that sliding mode approach can give good results in terms of robustness toward large supply and load variations, while maintaining good dynamic response. Control refinements, like current limitation, constant switched frequency and output voltage steady-state error cancellation are also discussed.

2. GENERAL SLIDING MODE CONTROL (GSMC) – THEORETICAL ASPECTS

In sliding mode control the controller [2], [3], [4], employs a sliding surface or line to decide its control input states *u*, which corresponds the turning on and off the power converter's switch, to the system:

$$\sigma = \alpha x_1 + x_2 \tag{1}$$

where α is a positive quantity in some literature called a convergence factor and is taken to be

$$\alpha = \frac{1}{R_L C}$$

Graphically the sliding line (Fig. 1) is a straight line on the state plane with gradient α that determines the dynamic response of the system in sliding mode with a first order time constant $\tau = 1/\alpha$.

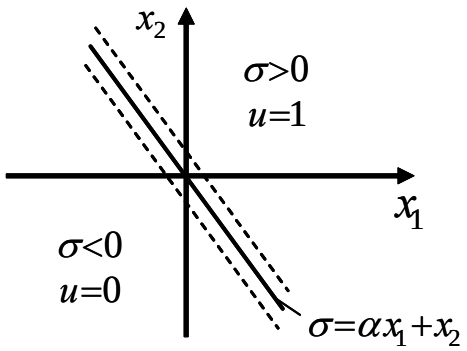


Fig. 1. A diagram of sliding mode control

To ensure that a system follows its sliding surface, a control law must be imposed:

$$u = \begin{cases} 1 = 'ON' & when\ \sigma > 0 \\ 0 = 'OFF' & when\ \sigma < 0 \end{cases}$$

The existing condition for sliding mode is:

$$\dot{\sigma} = \begin{cases} \dot{\sigma} < 0 & for\ \sigma > 0 \\ \dot{\sigma} > 0 & for\ \sigma < 0 \end{cases}$$

Based on this we get the following expression for $\dot{\sigma}$:

$$\dot{\sigma} = \alpha \dot{x}_1 + \dot{x}_2 = \alpha x_2 + \dot{x}_2 = \alpha x_2 - \frac{\beta}{LC}(uV_i) - \frac{1}{LC} x_1 - \frac{1}{R_L C} x_2 + \frac{V_{ref}}{LC}$$

Depending on σ and u the state space is divided into two regions:
region 1: $\sigma > 0$ and $u = 1$

$$\dot{\sigma}_1 = \left(\alpha - \frac{1}{R_L C} \right) x_2 - \frac{\beta}{LC}(uV_i) - \frac{1}{LC} x_1 + \frac{V_{ref}}{LC} < 0$$

region 2: $\sigma < 0$ and $u = 0$

$$\dot{\sigma}_2 = \left(\alpha - \frac{1}{R_L C} \right) x_2 - \frac{1}{LC} x_1 + \frac{V_{ref}}{LC} > 0$$

Sliding mode will only exist on the portion of the sliding line that covers both of the region 1 ($\dot{\sigma}_1 < 0$) and region 2 ($\dot{\sigma}_2 > 0$) [5].

From one side the speed of the system increases with increasing of α (sliding line become steeper), but from other side the existing region of the sliding mode decreases that can cause an overshoot in the voltage response ($\alpha >> 1/R_L C$) [5].

3. SLIDING MODE CONTROL OF DC/DC BUCK CONVERTER – BASIC SCHEME

A basic sliding mode control scheme of DC/DC buck converter is shown in Fig. 2, where V_0 is the regulated output voltage. The comparator switches the input to the buck converter based on the polarity of the compensator output. Unlike in PWM regulators, the switching frequency of the buck converter with sliding control is not fixed by an external source and is a function of a reference voltage. The feedback is highly nonlinear due to the comparator.

However, this kind of system can be intuitively understood by its phase portrait, as shown in Fig. 3.

The buck converter contains two poles, so the feedback loop is a second-order system. The phase portrait (Fig. 3) describes the transient operation of the circuit by the time trajectories of the state variable, (V, dV/dt), with the time variable being implicit.

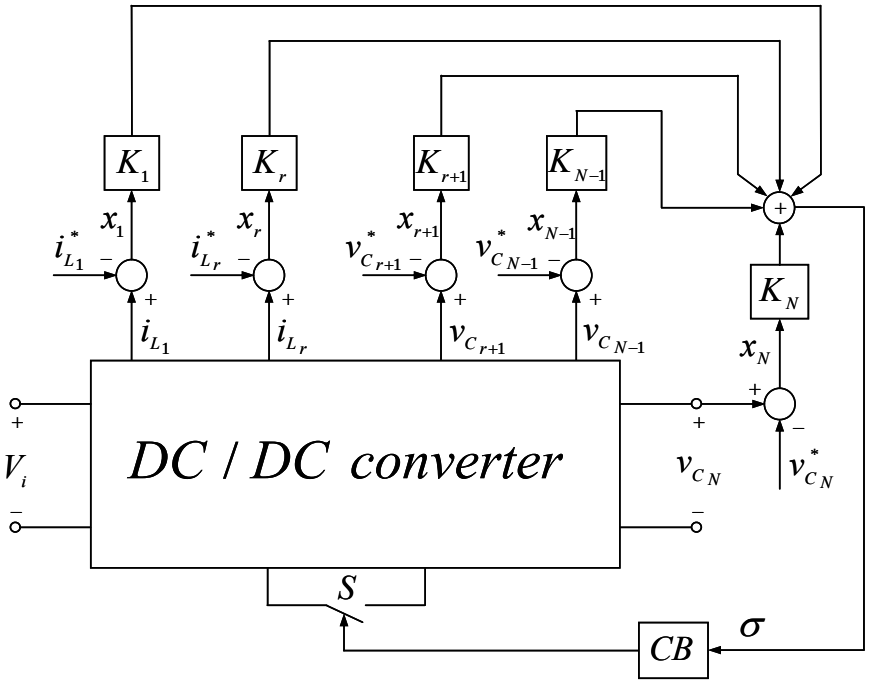


Fig. 2. Sliding mode control for DC/DC buck converter – basic scheme

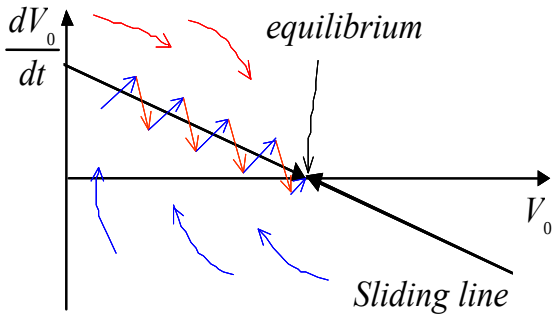


Fig. 3. Sliding mode control – phase portrait

Let V_i and v_{CN} be input and output voltages, respectively, i_{Li} and v_{Cj} ($i = 1 – r, j = r + 1–N–1$) the internal state variables of the converter (inductor currents and capacitor voltages), and N the system order (Fig.2). According to the theory, all state variables are sensed, and the corresponding errors x_i (defined by difference to the steady-state values) are multiplied by proper gains K_i and added together to form the sliding function σ. Then, control block *CB* control the switch so as to maintain function σ near to zero, thus we can write

$$\sigma = \sum_{1}^{N} K_i x_i = K^T x = 0 \qquad (2)$$

where $\mathbf{K}=[K_1, K_2,.., K_N]^T$ is the vector of sliding coefficients (^{T} means transposition).

This equation represents a hyper plane in the phase space passing through the origin (note that (1) is a particular case of (2) for N=2).

This general scheme, although interesting in theory, is not practical. In fact, a control approach representing an effective alternative to standard control techniques must

have the following characteristics - simple design and implementation even for high-order converter topologies, constant switching frequency, no steady-state errors, possibility to implement a current limitation, which are not satisfied in the scheme of Fig. 2.

Moreover, the hysteretic operation causes a variable switching frequency, while steady-state errors can arise from the fact that, in a practical implementation, all state variable errors, except for the output voltage, are computed using high-pass filters, thus they have zero average value. If sliding function σ has non-zero average value, a steady-state output voltage error necessarily appears. Lastly, switch current limitation is not implemented in the scheme shown.

As far as control complexity is concerned, it has been proved that excellent performances can be obtained even with reduced-order controllers, i.e. by sensing only one inductor current (the only inductor current in 2nd-order schemes like buck converter) and the output voltage. This corresponds to choose zero value for some coefficients K_i.

4. PRACTICAL SLIDING MODE CONTROLLER OF DC/DC BUCK CONVERTER

The general-purpose sliding mode controller scheme is shown in Fig. 4. As mentioned above only one inductor current and the output voltage are sensed.

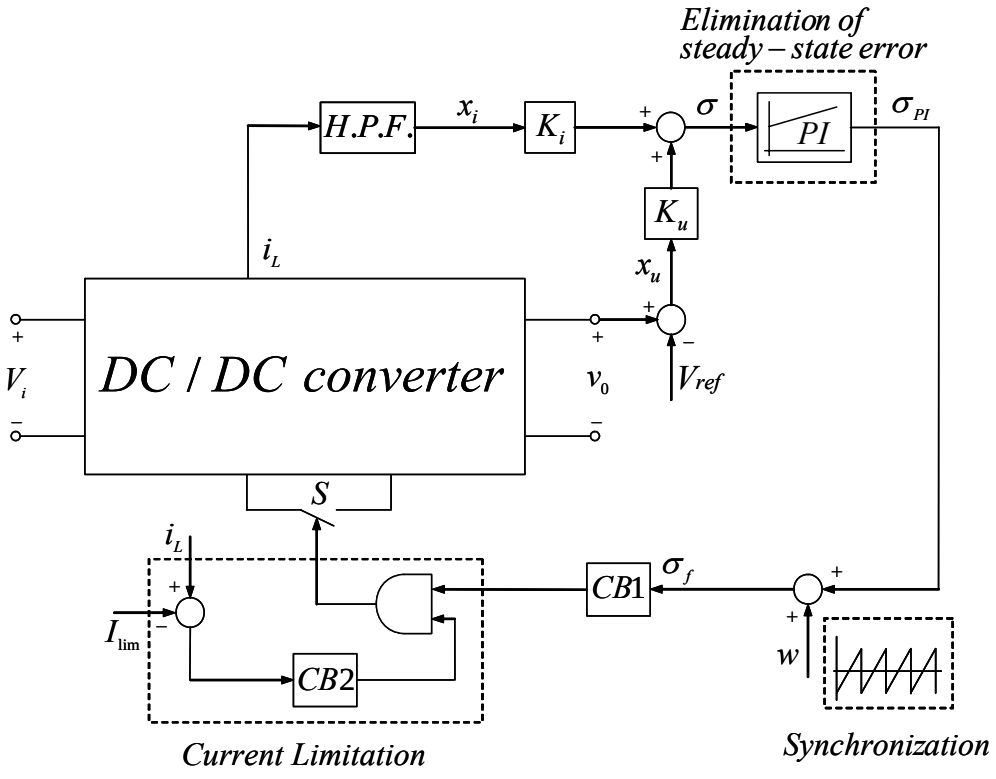


Fig. 4. General purpose sliding mode controller scheme

4.1. Elimination of Steady-State Error

As already mentioned, current error x_i is computed by means of a high-pass filter, while output voltage error x_u is obtained by comparison with reference signal V_{ref}. A PI action is introduced on sliding function σ in order to eliminate its dc value, thus reducing the dc value of the output voltage error to zero. In practice, the integral action of this regulator is enabled only when the system is on the sliding surface; in this way, the system behavior during large transients, when σ can have values far from zero, is not affected, thus maintaining the fast response of sliding mode control.

4.2. Switching Frequency Stabilization

In order to provide stabilization of the switching frequency, a proper ramp signal w at the desired frequency f_w is added to function σ. If in the steady state, the amplitude of w is predominant in σ_f, a commutation occurs at any cycle of w, thus making the switching frequency equal to f_w. This also allows converter synchronization to an external trigger. Instead, under dynamic conditions, error terms x_i and x_u increase, w is overridden, and the system retains the excellent dynamic response of the sliding mode.

4.3. Current Limitation

The current limiter shown in Fig. 4 overrides sliding mode control when the inductor current exceeds I_{lim}. If this happens, the hysteretic control ($CB1$) maintains this current around the value I_{lim}. This corresponds to perform a sliding mode on the inductor current only.

It is important to note that the controller design requires selection of two parameters, the time constant of the high-pass filter used to compute the inductor current error, and the ratio K_i/K_u of the sliding coefficients. However, the general theory does not take into account the additional state variable of the filter. In order to overcome this problem, it is possible to derive a small-signal model of the sliding mode control, which takes into account also the filter time constant, so that the design parameters can be chosen from small-signal analysis as in current-mode control.

5. CONCLUSIONS

Control techniques of variable structure systems find a natural application to switch-mode DC/DC converters. In particular, the sliding mode control represents a powerful tool to enhance performances of power converters.

Sliding mode control is able to ensure system stability even for large supply and load variations, good dynamic response and simple implementation for power DC/DC converters. These features make this control technique a valid alternative to standard control approaches like current-mode control.

Application to DC/DC buck converter is analyzed and provisions to overcome the inherent drawbacks of sliding mode control, i.e. variable switching frequency and possible steady-state errors, are described.

REFERENCES

[1] R. W. Erickson, *Fundamentals of Power Electronics*, 1st ed. NewYork: Chapman and Hall, 1997.

[2] E. Fossas and D. Biel, *A sliding mode approach to robust generation on DC-to-DC nonlinear converters*, in Proceedings,IEEE International Workshop on Variable Structure Systems, pp. 67–71, Dec 1996.

[3] C. Morel, J.-C Guignard, and M. Guillet, *Sliding mode control of DC-to-DC power converters*, in Proceedings, 9th International Conference on Electronics, Circuits and Systems, vol. 3, pp. 971–974, Sep. 2002.

[4] G. Spiazzi and P. Mattavelli, *Sliding-mode control of switched mode power supplies*, Ch. 8, The Power Electronics Handbook, Boca Raton FL: CRC Press LLC, 2002.

[5] V.I. Utkin, *Sliding Modes in Control and Optimization*, Springer-Verlag, 1992

COMPUTER-AIDED EDUCATIONUSING MATLAB IN THE THEORETICAL ELECTRICAL ENGINEERING AND NEURAL NETWORKS COURSE PROJECTS AND LABS AT THE TECHNICAL UNIVERSITY OF SOFIA

Veselin Mladenov, Georgi Komsalov, Georgi Tsenov

Department of Theoretical Electrical Engineering, Faculty of Automatics
Technical University of Sofia 8, Kliment Ohridski St, Sofia-1000

Abstract*. This paper shows the main areas of application of computer-aided education with usage of MATLAB on Theoretical Electrical Engineering course project in the Technical University of Sofia, Bulgaria. The paper also presents the interactive educational system "Virtuallab", an interactive educational windows based system integrated within the MATLAB programming environment, that is used in the lab exercises education process of Neural Networks course in the Technical University of Sofia. The benefits and drawbacks of using computers in course projects and laboratory exercises are outlined. Several examples of course project assignments are given.*

Keywords: Circuit Theory, Electromagnetics, Computer-Aided Education, Neural networks.

1. INTRODUCTION

Theoretical Electrical Engineering (TEE) in Technical University of Sofia is focused on two major areas: electromagnetic field theory and circuit theory. The basic principles, laws and methods of the electrical circuit theory and electromagnetic field theory are fundamental and of great importance for all electrical engineers. Good understanding and deep fundamental knowledge in electromagnetic phenomena are absolutely necessary for the further specialization of all students as it is a common need for the future engineers who will specialize in electronics, communications and computers, as well as for those who will deal with electrical machines and apparatus.

The mathematical basis of the electromagnetic problems is difficult and very often obtaining analytical solution is a long process, requiring great amount of efforts. This makes the course in Theoretical Electrical Engineering not very attractive for the students though as aforementioned it is of fundamental importance. Using computers, tedious mathematical work can be substantially reduced, thus leaving more time for explanation of physical nature of the electromagnetic phenomena and their applications.

Another reason to use extensively personal computers in all aspects of the teaching of Theoretical Electrical Engineering is the contemporary electrical engineering job market that requires engineers with high-level computer literacy.

The basis of this computer application is the department's computer laboratory, equipped with 16 PCs and a laser printer, connected with wireless LAN having Internet access. Portable computers and video projectors are used in the lecture rooms.

The computerized presentation of the lectures is based on the new textbooks. Other, 8 less powerful computers are available in the measuring laboratories to be used as modeling and computational tool in the laboratory exercises, together with a new laboratory manual.

With respect to the high prices of the scientific software and the department's limited budget, the software used in teaching is mostly freeware – like the Student versions of MATLAB package. On the other side this is good for the students as they have the possibility to install the software on their own computers and to use it at home and for other subjects, too.

The major benefits of computer-based education on the TEE are:

- reducing the students' work at lecture time;
- better understanding of the taught material;
- improving student assessment and speeding-up the examination process using computer quizzes;
- continuous, non-interruptible learning process using computers;
- TEE courses become more attractive for the students.

Another contribution to the computerization of the education at TEE department is the availability of Internet-based courses on Circuit Theory and Electromagnetic that run on the department's server. Those courses can be accessed 24 hours – 7 days a week by the students, thus completing the existing textbooks and manuals on the subject. Using those courses the students can improve and check their knowledge and have the possibility to learn the subject in a modern way, reading texts, browsing analytically and numerically solved problems, solving unsolved problems, using interactive Java applets and answering quizzes. This way is especially suited for part-time students and distant learning.

2. COMPUTER APPLICATIONS IN TEE COURSE PROJECTS

The simulation of the processes by using MATLAB package contributes to their easier understanding and to their quantitative and qualitative assessment especially in the reduced time for the topic considered. It is more useful to analyze elementary circuits because it is easy to predict the development of the processes in such circuits in the frequency domain as well as in the time domain. The aim is to demonstrate the application of circuit analysis methods and to compare results of analytical calculation and numerical simulations using MATLAB.

The students are provided with prewritten examples and functions that show how the examples can be solved numerically with algorithm which utilizes Node voltage method. With usage of algorithm problem solving the students are acquiring some basic MATLAB programming skills in the process.

With the examples on electromagnetic field theory students learn how to do electromagnetic field calculations, to calculate electromagnetic parameters – fluxes, energies, inductances, capacitances, etc.

Example on some Student projects on circuit theory

For the circuit shown in Fig. 1 with parameters $Z_1 = 4 - j4\,\Omega$, $Z_2 = 10\,\Omega$, $Z_3 = j10\,\Omega$, $Z_5 = j5\,\Omega$, $Z_6 = -j5\,\Omega$, $\dot{E}_1 = 10 - j10\text{V}$, $\dot{E}_2 = j5\text{V}$, $\dot{E}_4 = 20\text{V}$, $\omega = 100\text{rad/s}$.

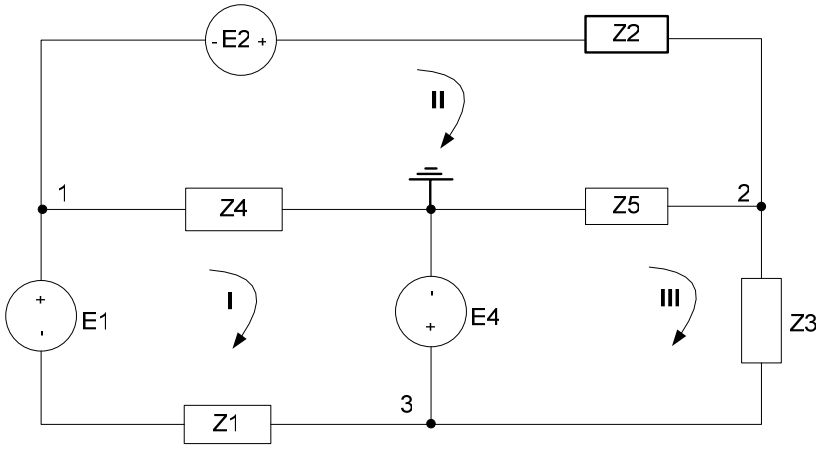


Fig. 1

Obtain:

a) analytically the branch currents by the mesh current method and the node voltage method, and the total power;

b) numerically the branch currents and the total power, using MATLAB. Compare the results.

Node voltage method: Consider again the circuit from Fig. 1. There are 3 nodes available (the choice of the datum node is made). The circuit contains a branch with a single independent voltage source and hence two KCL-equations can be written.

$$\begin{vmatrix} \left(\frac{1}{Z_1}+\frac{1}{Z_2}+\frac{1}{Z_5}\right)\dot{V}_1 - \frac{1}{Z_2}\dot{V}_2 - \frac{1}{Z_1}\dot{V}_3 = \frac{\dot{E}_1}{Z_1} - \frac{\dot{E}_2}{Z_2} \\ -\frac{1}{Z_2}\dot{V}_1 + \left(\frac{1}{Z_2}+\frac{1}{Z_3}+\frac{1}{Z_6}\right)\dot{V}_2 - \frac{1}{Z_3}\dot{V}_3 = \frac{\dot{E}_2}{Z_2} \end{vmatrix} . \qquad (1)$$

The phasor of the node voltage of node 3 is $\dot{V}_3 = \dot{E}_4 = 20\text{V}$.

Once the node voltages are obtained, the branch currents can be obtained from OGL-equations. The results are the same as those obtained by the mesh current method.

In this part, the students practice again KCL and OGL. They can make comparison between two methods.

In the circuit, both the complex generated power and the complex consumed power are:

$$\dot{S}_{src} = \dot{S}_{cons} = 161.95 + j43.65VA. \qquad (2)$$

An m-file is created forming and solving the system of equations for the branch currents and computing the complex power according (3).

The students make a computer program and compare the results obtained in several ways using predefined m-file functions.

Example 2. For circuit with parameters $R_1 = 10\,\Omega$, $L_1 = 4\,\text{mH}$, $L_2 = 3\,\text{mH}$, $L_1 = 4\,\text{mH}$, $C_3 = 200\,\mu\text{F}$, $e(t) = 100\sqrt{2}\sin\omega t$, V, obtain:

a) the corresponding resonant frequencies when the series and the parallel resonance arise – analytically and by MATLAB and compare the results;

b) the frequency response – using MATLAB.

3. COMPUTER APPLICATIONS IN NM DISCIPLINE

In Technical University of Sofia the MATLAB package and its Neural Networks toolbox is also used for the Neural Networks Discipline.

For the ease of the students an window based educational system, which runs within the MATLAB has been created. The system called Virtuallab allows pictures and text to be presented interactively, allowing students on demand to see theory slides and run exemplary code and project variants with just some clicks on the mouse. This system also allows on-line education to be adopted. The Virtuallab system is coupled with student book manual, which includes more step by step simple neural network examples on perceptrons, Hopfield networks, feedforward backpropagation networks, networks based on competition and self organizing maps.

After the completion of the examples included in the Virtuallab module students have a final task to perform a classification on some prerecorded data with usage of the MATLAB Neural network toolbox functions. In the given exercises the students learn and see how they can use the software with arbitrary input data, how the number of neurons in the hidden layers is having impact on the performance and etc.

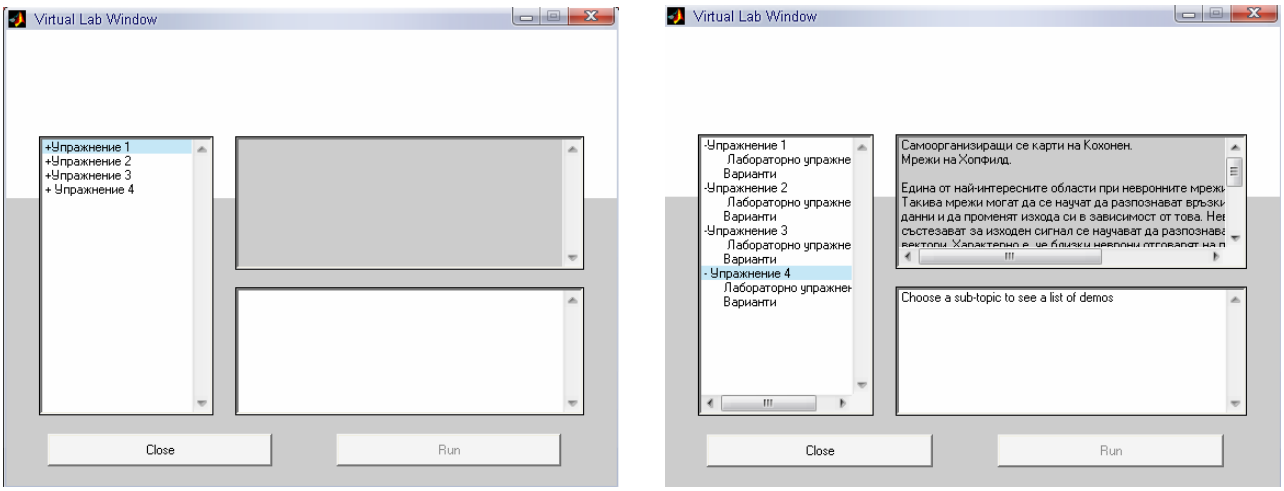
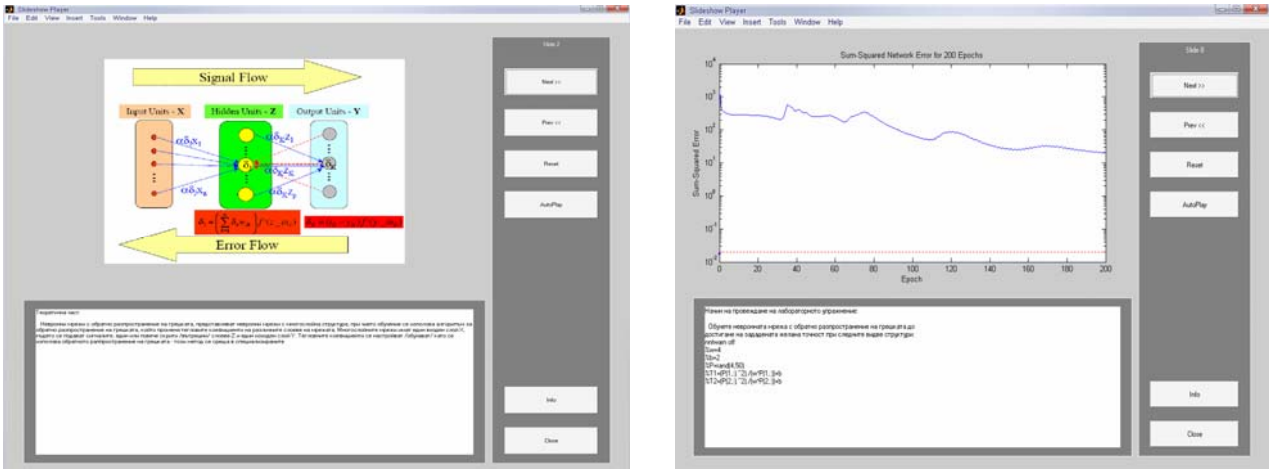


Fig. 2. Students can choose the excersise number, and then open the examples combined with theory or go directly to the assignmens



One example with theory and exaplary code

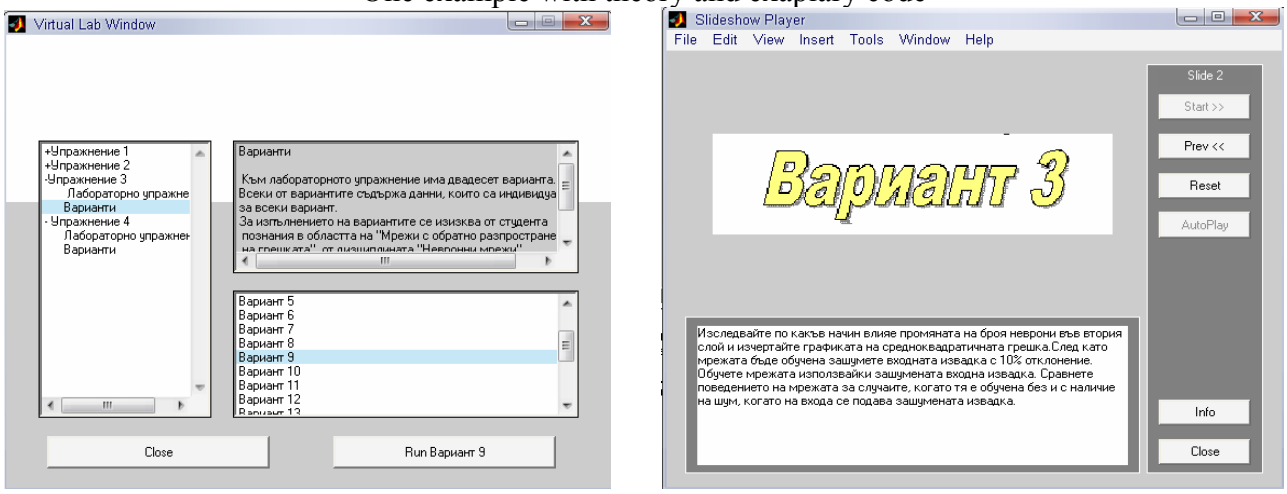


Fig. 3. One example on how students can choose their assignment variant with number specified by the teacher

4. CONCLUSION

The experience with computer-aided education on Theoretical Electrical Engineering and Neural Networks at the Technical University of Sofia shows that the teaching and learning processes have become more effective after their introduction. Students understand better the teaching material for reduced teaching time. The computer-aided course is more attractive for the students and their knowledge of the subject is extended.

REFERENCES

[1] К. Брандиски, В. Младенов, Д. Вълчев, Решаване на задачи по електротехника с MATLAB, София, 2000.
[2] К. Брандиски, Ж. Георгиев, В. Младенов, Р. Станчева, Учебник по Теоретична Електротехника, ч. I, София, 2004.
[3] К. Брандиски, Ж. Георгиев, В. Младенов, Р. Станчева, Учебник по Теоретична Електротехника, ч. II, София, 2004.
[4] С. Йорданова, В. Младенов, Размито управление и невронни мрежи, София,2006 г.
[5] С. Йорданова, В. Младенов, Р. Петрова, Г. Ценов, Ръководство по „Размито управление и невронни мрежи”, София, 2007 г.

СИНТЕЗ И АНАЛИЗ НА ЗАГРАЖДАЩИ АКТИВНИ БИКВАДРАТНИ ФИЛТРОВИ ЗВЕНА С ИЗПОЛЗВАНЕТО НА MATLAB И MICROCAP

Таня Методиева Стоянова[1], Адриана Найденова Бороджиева[2]

[1]Катедра „Теоретична и измервателна електротехника”,
[2]Катедра „Комуникационна техника и технологии”,
Русенски университет „Ангел Кънчев”, България, 7017 Русе, ул. „Студентска” № 8,
тел.: [1](00359 82) 888 502, [2](00359 82) 888 734,
e-mail: [1]tstoyanova@uni-ruse.bg, [2]aborodjieva@ecs.uni-ruse.bg

Резюме: *В тази публикация се синтезират и анализират заграждащи активни биквадратни филтрови звена. Това се извършва по зададена нормализирана предавателна функция по напрежение и при известни нормализиращо съпротивление и средна честота. Синтезът се реализира чрез програмата MATLAB, а анализът на проектираните филтри след избора на стандартни стойности на съпротивленията на резисторите и на капацитетите на кондензаторите в синтезираните вериги се извършва чрез програмата MicroCAP, използвана за симулация на аналогови и цифрови вериги. Резултатите ще бъдат използвани в процеса на обучение по дисциплината „Комуникационни вериги”, изучавана от студентите от специалност „Комуникационна техника и технологии” от образователно-квалификационната степен „Бакалавър”.*

Ключови думи: Анализ, синтез, активни биквадратни филтрови звена.

1. ВЪВЕДЕНИЕ

Разглежданите активни филтрови биквадратни звена [3, 4] са изградени от един операционен усилвател, резистори и кондензатори.

В таблица 1 са дадени нормализираната предавателна функция по напрежение на активни заграждащи филтри (ЗФ) от втори ред, както и тяхната предавателна функция, записана чрез коефициента на усилване в лентата на пропускане k_0, кръговата средна честота ω_0 и полюсния качествен фактор Q на тези филтри [2, 3, 4].

Таблица 1. Предавателни функции по напрежение на ЗФ от втори ред

Активни филтри от втори ред		
Заграждащ филтър (ЗФ)	$T(p)=\dfrac{H\left(p^2+a\right)}{p^2+b_1p+b_0}$	$T(p)=\dfrac{k_0\left(p^2+{\omega_0}^2\right)}{p^2+\dfrac{\omega_0}{Q}p+{\omega_0}^2}$

2. СЪСТОЯНИЕ НА ПРОБЛЕМА – АНАЛИЗ И СИНТЕЗ НА ЗАГРАЖДАЩИ АКТИВНИ ФИЛТРОВИ БИКВАДРАТНИ ЗВЕНА

На фиг. 1 е дадена схемата на заграждащо активно биквадратно филтрово звено. В публикацията се предлагат две методики за синтеза на такива филтри, като са посочени и условията за проектирането им.

В [5] чрез теорията на сигналните графи [1] е изведен аналитичният израз за предавателната функция на проектирания заграждащ филтър:

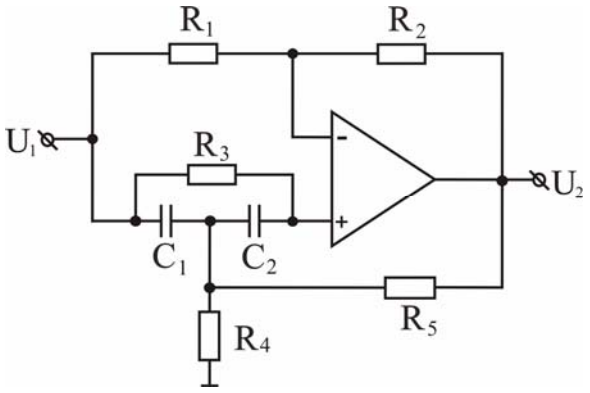
$$T(p)=\frac{p^2C_1C_2.G_2+p.\left[G_2.G_3.(C_1+C_2)-G_1.C_2.(G_4+G_5)\right]+G_2.G_3.(G_4+G_5)}{p^2C_1C_2.G_2+p.\left[G_2.G_3.(C_1+C_2)+C_2.G_2.G_4-C_2.G_1.G_5\right]+G_2.G_3.(G_4+G_5)}. \quad (1)$$

Условието за реализиране на предавателна функция на заграждащ филтър е коефициентът пред p^1 в числителя да е равен на нула:

$$G_2.G_3.(C_1+C_2)=G_1.C_2.(G_4+G_5), \quad (2)$$

при което за предавателната функция на филтъра окончателно се записва:

$$T(p)=\frac{p^2C_1C_2.G_2+G_2.G_3.(G_4+G_5)}{p^2C_1C_2.G_2+p.\left[G_2.G_3.(C_1+C_2)+C_2.G_2.G_4-C_2.G_1.G_5\right]+G_2.G_3.(G_4+G_5)}. \quad (3)$$



Фиг. 1. Заграждащо активно биквадратно филтрово звено

Таблица 2. Количествени съотношения между коефициентите a и b_1

a	b_1	<	2.SQRT(2a)
1	b_1	<	2,8284
2	b_1	<	4,0000
3	b_1	<	4,8990
4	b_1	<	5,6569
5	b_1	<	6,3246

След разделяне на числителя и знаменателя на предавателната функция с $C_1C_2.G_2$ се записва:

$$T(p)=\frac{p^2+\dfrac{G_3.(G_4+G_5)}{C_1C_2}}{p^2+p.\left[\dfrac{G_3.(C_1+C_2)}{C_1C_2}+\dfrac{G_4}{C_1}-\dfrac{G_1.G_5}{C_1.G_2}\right]+\dfrac{G_3.(G_4+G_5)}{C_1C_2}}. \quad (4)$$

Като се отчете общият вид на предавателната функция на ЗФ от Таблица 1 (колона 3) при $k_0 = 1$, може да се състави следната система от три уравнения със седем неизвестни за неговото проектиране:

$$\left|\begin{array}{l} G_2.G_3.(C_1+C_2) = G_1.C_2.(G_4+G_5) \\ \dfrac{G_3.(G_4+G_5)}{C_1C_2} = \omega_0^2 \\ \dfrac{G_3.(C_1+C_2)}{C_1C_2} + \dfrac{G_4}{C_1} - \dfrac{G_1.G_5}{C_1.G_2} = \dfrac{\omega_0}{Q} \end{array}\right. . \qquad (5)$$

Ако предавателната функция на ЗФ е зададена в нормиран вид (Таблица 1, колона 2), то тогава система (5) придобива вида:

$$\left|\begin{array}{l} G_2.G_3.(C_1+C_2) = G_1.C_2.(G_4+G_5) \\ \dfrac{G_3.(G_4+G_5)}{C_1C_2} = a \\ \dfrac{G_3.(C_1+C_2)}{C_1C_2} + \dfrac{G_4}{C_1} - \dfrac{G_1.G_5}{C_1.G_2} = b_1 \end{array}\right. . \qquad (6)$$

В публикацията се предлагат два метода за синтез на заграждащи филтри.

Метод 1:

Приема се: $C_1 = C_2 = C = 1$ и $G_1 = G_2 = G = 1$ (като нормирани стойности). Тогава система (6) се преобразува във вида:

$$\left|\begin{array}{l} 2.C.G.G_3 = G.C.(G_4+G_5) \\ \dfrac{G_3.(G_4+G_5)}{C^2} = a \\ \dfrac{G_3.2C}{C^2} + \dfrac{G_4}{C} - \dfrac{G.G_5}{C.G} = b_1 \end{array}\right. \Rightarrow \left|\begin{array}{l} G_3 = \sqrt{\dfrac{a}{2}} = \dfrac{\sqrt{2.a}}{2} > 0 \\ G_4 = \dfrac{b_1}{2} > 0 \\ G_5 = 2.G_3 - G_4 = \sqrt{2.a} - \dfrac{b_1}{2} > 0 \end{array}\right. . \qquad (7)$$

Условието за съществуване на решение за проектирания заграждащ филтър е: $2.\sqrt{2.a} > b_1 > 0$.

В таблица 2 са показани какви съотношения трябва да изпълнява коефициентът b_1 за различни стойности на коефициента a, за да може да се проектира заграждащият филтър.

Метод 2:

Приема се: $C_1 = C_2 = C$ и $G_4 + G_5 = G$, произволно се избират нормирани стойности за C, G_1 и G_3. Система (6) се трансформира във вида:

$$\left|\begin{array}{l} 2.C.G_2.G_3 = C.G_1.G \\ \dfrac{G_3.G}{C^2} = a \\ \dfrac{2.G_3}{C} + \dfrac{G_4}{C} - \dfrac{G_1.G_5}{C.G_2} = b_1 \end{array}\right| \Rightarrow \left|\begin{array}{l} 2.G_2.G_3 = G_1.G \\ G = \dfrac{a.C^2}{G_3} \\ 2.G_3 + G_4 - \dfrac{G_1}{G_2}.(G - G_4) = b_1.C \end{array}\right| \tag{8}$$

Третото уравнение на система (8) може да се запише във вида:

$$2.G_3 - \frac{G_1}{G_2}.G + G_4.\left(1 + \frac{G_1}{G_2}\right) = b_1.C \Rightarrow G_4 = \frac{b_1.C}{1 + \dfrac{G_1}{G_2}}. \tag{9}$$

Отчитайки първите две уравнения на система (8), уравнение (9) се преобразува до:

$$G_4 = \frac{b_1.C}{1 + \dfrac{G_1}{G_2}} = \frac{b_1.C}{1 + \dfrac{G_1.2.G_3^2}{a.C^2.G_1}} = \frac{a.b_1.C^3}{a.C^2 + 2.G_3^2}, \tag{10}$$

а за G_5 се получава:

$$G_5 = G - G_4 = \frac{a.C^2}{G_3} - \frac{a.b_1.C^3}{a.C^2 + 2.G_3^2} = \frac{a.C^2\left(a.C^2 + 2.G_3^2 - b_1.C.G_3\right)}{G_3.\left(a.C^2 + 2.G_3^2\right)}. \tag{11}$$

Метод 2 съдържа стъпките: произволно се избират нормирани стойности за C, G_1 и G_3, а стойностите на G_2, G_4 и G_5 се изчисляват по формулите:

$$\left|\begin{array}{l} G_2 = \dfrac{a.C^2.G_1}{2.G_3^2} \\ G_4 = \dfrac{a.b_1.C^3}{a.C^2 + 2.G_3^2} \\ G_5 = \dfrac{a.C^2\left(a.C^2 + 2.G_3^2 - b_1.C.G_3\right)}{G_3.\left(a.C^2 + 2.G_3^2\right)} \end{array}\right| . \tag{12}$$

Условието за съществуване на решение на система (12) е:

$$a.C^2 + 2.G_3^2 - b_1.C.G_3 > 0. \tag{13}$$

Израз (13) представлява квадратно неравенство по отношение на G_3. При неговото решаване се разглеждат следните три случая:

Случай 1: $b_1^2 - 8.a > 0$.

Квадратното уравнение $2.G_3^2 - b_1.C.G_3 + a.C^2 = 0$ ще има два реални корена:

$$G_3^{\mathrm{I,II}} = \frac{b_1 C \pm \sqrt{b_1^2 C^2 - 8aC^2}}{2.2} = \frac{C}{4}\left(b_1 \pm \sqrt{b_1^2 - 8a}\right), \tag{14}$$

а неравенство (13) ще има следното решение:

$$G_3 \in \left(0; \frac{C}{4}\left(b_1 - \sqrt{b_1^2 - 8a}\right)\right) \cup \left(\frac{C}{4}\left(b_1 + \sqrt{b_1^2 - 8a}\right); +\infty\right). \tag{15}$$

Случай 2: $b_1^2 - 8.a = 0$.

Квадратното уравнение $2.G_3^2 - b_1.C.G_3 + a.C^2 = 0$ има един двукратен корен:

$$G_3^{\mathrm{I,II}} = \frac{b_1 C}{4}, \tag{16}$$

а решението на неравенство (13) ще е:

$$G_3 \in \left(0; \frac{b_1 C}{4}\right) \cup \left(\frac{b_1 C}{4}; +\infty\right), \text{ т.е. } G_3 \neq \frac{b_1 C}{4}. \tag{17}$$

Случай 3: $b_1^2 - 8.a < 0$.

Квадратното уравнение $2.G_3^2 - b_1.C.G_3 + a.C^2 = 0$ няма реални корени, а неравенство (13) ще е удовлетворено за всяка стойност на G_3.

2. РЕЗУЛТАТИ

Алгоритъмът за синтез и анализ на заграждащи активни филтрови звена съдържа следните стъпки:

1. Въвеждане на коефициентите a, $b_0 = a$ и b_1 в предавателната функция по напрежение $T(p)$ на синтезирания филтър.

2. Изчисляване на нормираните стойности на елементите на синтезирания заграждащ филтър, чиято схема е показана на фиг. 1.

- за ЗФ – метод 1: условие за прилагане на методиката – $2.\sqrt{2.a} > b_1 > 0$ и условие за проектиране – $C_1 = C_2 = C = 1$ и $G_1 = G_2 = G = 1$ (като нормирани стойности).
- за ЗФ – метод 2: условие за проектиране – $C_1 = C_2 = C$ и $G_4 + G_5 = G$; произволно се избират нормирани стойности за C, G_1 и G_3, като при избора на нормирана стойност за G_3 трябва да се спазва определено ограничение в зависимост от коефициентите в числителя и знаменателя на предавателната функция на проектирания филтър: a и b_1.

❖ $b_1^2 - 8.a > 0 \Rightarrow G_3 \in \left(0; \frac{C}{4}\left(b_1 - \sqrt{b_1^2 - 8a}\right)\right) \cup \left(\frac{C}{4}\left(b_1 + \sqrt{b_1^2 - 8a}\right); +\infty\right)$;

❖ $b_1^2 - 8.a = 0 \Rightarrow G_3 \in \left(0; \frac{b_1 C}{4}\right) \cup \left(\frac{b_1 C}{4}; +\infty\right)$, т.е. $G_3 \neq \frac{b_1 C}{4}$;

❖ $b_1^2 - 8.a < 0 \Rightarrow G_3 \in (0; +\infty)$.

Параметрите на останалите елементи се получават от решението на следните системи уравнения:

за ЗФ – метод 1

$$\begin{cases} G_3 = \sqrt{\frac{a}{2}} = \frac{\sqrt{2.a}}{2} > 0 \\ G_4 = \frac{b_1}{2} > 0 \\ G_5 = 2.G_3 - G_4 = \sqrt{2.a} - \frac{b_1}{2} > 0 \end{cases}$$

за ЗФ – метод 2

$$\begin{cases} G_2 = \frac{a.C^2.G_1}{2.G_3^2} \\ G_4 = \frac{b_1.C}{1 + \frac{G_1}{G_2}} = \frac{a.b_1.C^3}{a.C^2 + 2.G_3^2} \\ G_5 = G - G_4 = \frac{a.C^2\left(a.C^2 + 2.G_3^2 - b_1.C.G_3\right)}{G_3.\left(a.C^2 + 2.G_3^2\right)} \end{cases}$$

3. Въвеждане на стойността на нормиращото съпротивление R_N.

4. Въвеждане на средната честота f_0 за заграждащия филтър и изчисляване на нормиращата кръгова честота $\omega_N = 2\,\pi\, f_0$.

5. Изчисляване на денормираните стойности на елементите на двуполюсниците:

- за резисторите – получените стойности за съпротивления $R_k = 1/G_k$ за $k = 1...5$ се умножават с нормиращото съпротивление R_N;

– за кондензаторите – получените стойности за капацитети се разделят на произведението $\omega_N.R_N$ при $a=1$. При $a\neq 1$, получените стойности за капацитети се умножават с $\sqrt{a}/(\omega_N.R_N)$.

6. Избор на стандартни стойности на елементите на филтъра.

7. Извеждане на предавателната функция по напрежение $T(p)$ на заграждащия филтър от фиг. 1 [5]:

$$T(p)=\frac{p^2+\frac{G_3.(G_4+G_5)}{C_1C_2}}{p^2+p.\left[\frac{G_3.(C_1+C_2)}{C_1C_2}+\frac{G_4}{C_1}-\frac{G_1.G_5}{C_1.G_2}\right]+\frac{G_3.(G_4+G_5)}{C_1C_2}}.$$

8. Изчисляване на коефициента на усилване k_0 в лентата на пропускане, на полюсния качествен фактор Q и на средната честота f_0 за заграждащия филтър от фиг. 1, след избора на стандартни стойности на елементите:

$$k_0=1,\ \omega_0=\sqrt{\frac{R_4+R_5}{R_3R_4R_5C_1C_2}},rad/s;\qquad f_0=\frac{1}{2\pi}\sqrt{\frac{R_4+R_5}{R_3R_4R_5C_1C_2}},Hz,$$

$$Q=\frac{\sqrt{\frac{R_4+R_5}{R_3R_4R_5C_1C_2}}}{\frac{C_1+C_2}{R_3C_1C_2}+\frac{1}{R_4C_1}-\frac{R_2}{R_1R_5C_1}}.$$

9. Симулационно изследване на синтезирания филтър с използване на програмния продукт MicroCAP – изчертаване на амплитудно-честотната характеристика, определяне на параметрите k_0, Q и f_0 от снетата амплитудно-честотна характеристика.

Разработени са скриптове на MATLAB [6] за изчисляване на нормализираните и денормализираните стойности на компонентите за заграждащия филтър при зададена нормирана предавателна функция по напрежение.

Пример: Проектиране на заграждащи биквадратни филтри със средна честота f_0 и нормирана предавателна функция по напрежение $T(p)$. Денормализирането по честота и по съпротивление се извършва с нормиращо съпротивление R_N. Резултатите са показани в таблица 3. След това е извършен избор на стандартни стойности по скалата E-24, които се използват при симулацията с MicroCAP [7] за изчертаване на амплитудно-честотните характеристики (в dB) на проектираните филтри. В последната колона са посочени стойностите на коефициента на усилване k_0 в лентата на пропускане, на полюсния качествен

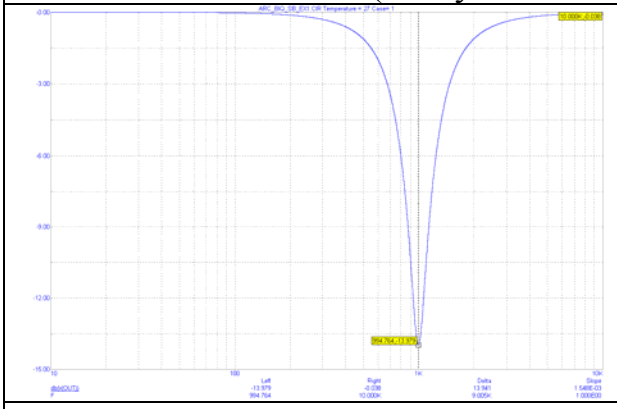
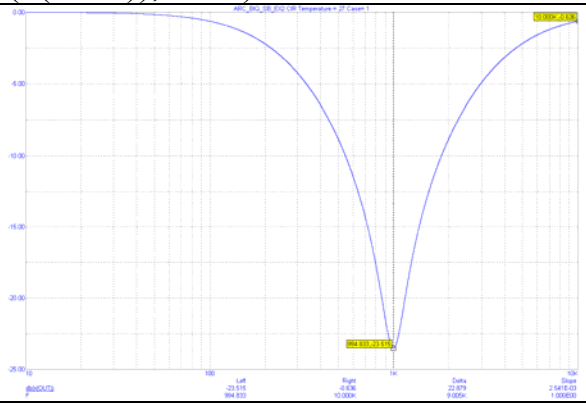
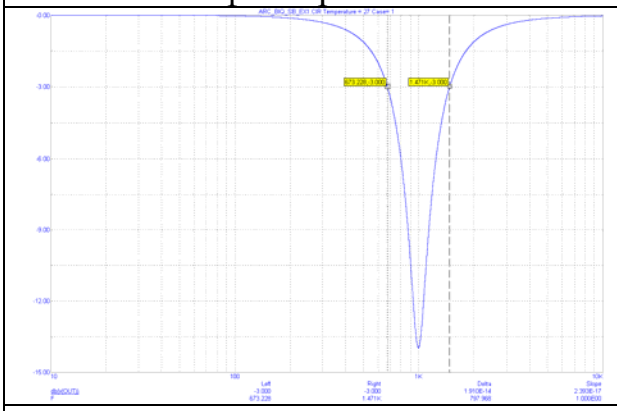
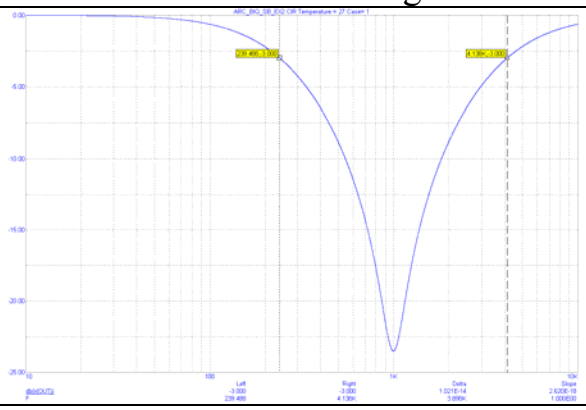
фактор Q и на средната честота f_0 за заграждащия филтър, след избора на стандартни стойности.

Таблица 3. Резултати от проектирането на ЗФ с MATLAB

Тип на филтъра	Нормали-зирани стойности	Денормализирани стойности	Стандартни стойности (Е-24)	Забележка
ЗФ 1	$R_1 = 1$ $R_2 = 1$ $C_1 = 1$ $C_2 = 1$ $R_3 = 1{,}4142$ $R_4 = 2{,}1053$ $R_5 = 1{,}0647$	$R_1 = 100\ k\Omega$ $R_2 = 100\ k\Omega$ $C_1 = 1{,}5915\ nF$ $C_2 = 1{,}5915\ nF$ $R_3 = 141{,}4214\ k\Omega$ $R_4 = 210{,}5263\ k\Omega$ $R_5 = 106{,}4721\ k\Omega$	$R_1 = 100\ k\Omega$ $R_2 = 100\ k\Omega$ $C_1 = 1{,}6\ nF$ $C_2 = 1{,}6\ nF$ $R_3 = 150\ k\Omega$ $R_4 = 200\ k\Omega$ $R_5 = 100\ k\Omega$	$f_0 = 1\ kHz$, $T(p) = \frac{p^2+1}{p^2+0{,}95p+1}$, $R_N = 100\ k\Omega$. --- $f_0 = 994{,}7184 Hz$, $Q = 1{,}2$, $k_0 = 1$.
ЗФ 2	$R_1 = 1$ $R_2 = 8$ $C_1 = 1$ $C_2 = 1$ $R_3 = 0{,}5$ $R_4 = 2{,}25$ $R_5 = 18$	$R_1 = 100\ k\Omega$ $R_2 = 800\ k\Omega$ $C_1 = 1{,}5915\ nF$ $C_2 = 1{,}5915\ nF$ $R_3 = 50\ k\Omega$ $R_4 = 225\ k\Omega$ $R_5 = 1800\ k\Omega$	$R_1 = 100\ k\Omega$ $R_2 = 820\ k\Omega$ $C_1 = 1{,}6\ nF$ $C_2 = 1{,}6\ nF$ $R_3 = 51\ k\Omega$ $R_4 = 220\ k\Omega$ $R_5 = 1800\ k\Omega$	$f_0 = 1\ kHz$, $T(p) = \frac{p^2+1}{p^2+4p+1}$, $R_N = 100\ k\Omega$. --- $f_0 = 994{,}8169 Hz$, $Q = 0{,}2551$, $k_0 = 1$.

В таблица 4 са показани резултатите от симулационното изследване с Micro-CAP, след като е извършен избор на стандартни стойности на елементите на схемата по скалата Е-24.

Таблица 4. Резултати от изследването на заграждащия филтър с MicroCAP

ЗФ 1	ЗФ 2
Определяне на средната честота на заграждащия филтър (Analysis→AC→db(v(OUT)), Peak)	
	
$f_0 = 994{,}764\ Hz \quad T_0 = -13{,}979\ dB$	$f_0 = 994{,}833\ Hz \quad T_0 = -23{,}515\ dB$
Определяне на широчината на лентата на пропускане и на качествения фактор на филтъра Go to Y → −3 → натиска се 2 x Left и 2 x Right	
	
Координати на левия маркер (Left): $f_{01} = 673{,}228\ Hz \quad T_{01} = -3{,}000\ dB$ Координати на десния маркер (Right): $f_{02} = 1{,}471\ kHz \quad T_{02} = -3{,}000\ dB$ Широчина на лентата на пропускане на филтъра: $B = Delta = f_{02} - f_{01} = 797{,}968\ Hz$ Качествен фактор на филтъра: $Q = \frac{f_0}{B} = \frac{994{,}764}{797{,}968} = 1{,}247$	Координати на левия маркер (Left): $f_{01} = 239{,}488\ Hz \quad T_{01} = -3{,}000\ dB$ Координати на десния маркер (Right): $f_{02} = 4{,}138\ kHz \quad T_{02} = -3{,}000\ dB$ Широчина на лентата на пропускане на филтъра: $B = Delta = f_{02} - f_{01} = 3{,}898\ kHz$ Качествен фактор на филтъра: $Q = \frac{f_0}{B} = \frac{994{,}833}{3898} = 0{,}255$
Определяне на коефициента на усилване в лентата на пропускане на филтъра	
$k_0 = 10^0 = 1$	$k_0 = 10^0 = 1$

4. ИЗВОДИ

1. В публикацията е описан алгоритъм, заложен в програмен модул, с използване на MATLAB, създаден за синтез и анализ на заграждащи активни бикводратни филтрови звена по схемата от фиг. 1 с един операционен усилвател.

2. Въз основа на израза за предавателната функция по напрежение на синтезирания филтър са изведени изрази за средната честота, за коефициента на усилване в лентата на пропускане и за полюсния качествен фактор на филтъра.

3. Разработеният програмен модул ще послужи и за автоматизиране на процеса на генериране на варианти на задания за курсови задачи по дисциплината „3110 Комуникационни вериги", включена като задължителна в новия учебен план на специалността „Комуникационна техника и технологии" за образователно-квалификационна степен „Бакалавър".

ЛИТЕРАТУРА

[1] Боянов, Й., Е. Шойкова. Теория на електронните схеми. София, Техника, 1995.

[2] Манукова, А., А. Бороджиева. Комуникационни вериги – ръководство за упражнения. 104 стр., Русе, 2002.

[3] Стоянов, Г. Теоретични основи на съобщителната техника. София, Техника, 1993.

[4] Huelsman, L.P., P.E. Allen. "Introduction to the Theory and Design of Active Filters", Moscow, 1984 (in Russian).

[5] Stoyanova, T., A. Borodzhieva. Synthesis and Analysis of Band-Stop Active Biquads Using MATLAB and MicroCAP. Summer School, Advanced Aspects of Theoretical Electrical Engineering, Sozopol '2010, 19.09. – 22.09.2010, Sozopol, Bulgaria (in Bulgarian, in press).

[6] http://www.mathworks.com.

[7] http://www.spectrum-soft.com/index.shtm.

ПРОГРАМНА СИСТЕМА ЗА ГЕНЕРИРАНЕ НА ЗАДАНИЯ ЗА КУРСОВИ ЗАДАЧИ ПО ДИСЦИПЛИНАТА „КОМУНИКАЦИОННИ ВЕРИГИ”

Таня Методиева Стоянова[1], Адриана Найденова Бороджиева[2]

[1]Катедра „Теоретична и измервателна електротехника”,
[2]Катедра „Комуникационна техника и технологии”, Русенски университет „Ангел Кънчев”,
България, 7017 Русе, ул. „Студентска” № 8,: тел. [1](00359 82) 888 502, [2](00359 82) 888 734,
e-mail: [1]tstoyanova@uni-ruse.bg, [2]aborodjieva@ecs.uni-ruse.bg

Резюме: *В тази публикация е описана програмна система за автоматизирано генериране на задания за курсови задачи по дисциплината „Комуникационни вериги”, изучавана от студентите от специалност „Комуникационна техника и технологии” от образователно-квалификационната степен (ОКС) „Бакалавър”. Системата обхваща следните раздели от материала: трептящи кръгове, електрически филтри и модулации.*

Ключови думи: Комуникационни вериги, трептящи кръгове, електрически филтри, модулации.

1. ВЪВЕДЕНИЕ

Съгласно учебната програма по дисциплината „Комуникационни вериги”, включена като задължителна в действащия учебен план за специалност „Комуникационна техника и технологии”, ОКС „Бакалавър”, в РУ „Ангел Кънчев”, на всеки студент се задава индивидуална курсова задача. Тя съдържа две части: първата обхваща три задачи от разделите, изучавани по дисциплината, а втората част съдържа цялостно проектиране на активен филтър, включително избор на стандартни стойности на резисторите и кондензаторите и симулиране на синтезираната схема чрез програмния продукт MicroCAP.

В публикацията е описана разработената програмна система за генериране на задания за курсови задачи.

2. СЪСТОЯНИЕ НА ПРОБЛЕМА – СЪДЪРЖАНИЕ НА МАТЕРИАЛА ПО ДИСЦИПЛИНАТА „КОМУНИКАЦИОННИ ВЕРИГИ”

Съгласно учебната програма по дисциплината „Комуникационни вериги” се изучават следните раздели:

1) Трептящи кръгове – последователен трептящ кръг; прост паралелен трептящ кръг; паралелни трептящи кръгове с последователно включване на товарното съпротивление на единия от реактивните елементи ($RC \| L$ и $RL \| C$); секционирани трептящи кръгове с индуктивен или с капацитивен делител; трансформатори с един или с два настроени трептящи кръга. В момента заданията за

курсовите задачи съдържат оразмеряването на секционирани трептящи кръгове с индуктивен или с капацитивен делител и трансформатор с един настроен трептящ кръг.

2) Електрически филтри – обща теория на филтрите и честотни преобразувания; реактивни (LC) филтри тип „К” и тип „m” (последователно-производни и паралелно-производни); пасивни и активни RC филтри. Разглеждат се още: кварцови и керамични филтри; амплитудни и фазови коректори. Активните RC филтри са обект на изследване на втората част на курсовата задача. Като възможни схемни варианти са филтри с един операционен усилвател по метода с едноконтурна и многоконтурна отрицателна обратна връзка (нискочестотни, високочестотни и лентови); активни филтри с един източник на напрежение, управляван с напрежение (нискочестотни, високочестотни и лентови); бикваратни активни филтрови звена (нискочестотни, високочестотни и заграждащи). В момента заданията за курсови задачи съдържат: синтез и анализ на лентови и заграждащи LC филтри тип „К”; последователно-производни и паралелно-производни нискочестотни и високочестотни LC филтри тип „m”; пасивни RC филтри – нискочестотни, високочестотни и лентови по схемата на Робинзон-Вин; лентови активни RC филтри с един операционен усилвател по метода с многоконтурната отрицателна обратна връзка.

3) Модулации – амплитудна модулация; ъглова модулация с двете си разновидности – честотна и фазова; въздействие на амплитудно-модулирани сигнали върху трептящи кръгове; импулсни модулации, включително импулсно-кодова модулация.

ПРИМЕРНО ЗАДАНИЕ ЗА КУРСОВА ЗАДАЧА ПО „КОМУНИКАЦИОННИ ВЕРИГИ”

Част I:

1. Да се проектира секциониран трептящ кръг с капацитивен делител така, че да трансформира товарното съпротивление $R_2 = 100\Omega$ в $R_t = 10k\Omega$ при честота $f_0 = 2MHz$, като пропусканата честотна лента на кръга е $B = 20kHz$. Да се реши задачата с използване на: 1) точните формули; 2) приблизителните формули.

2. Да се изчислят елементите на лентово полузвено тип „К” с гранични честоти $f_{c1} = 10kHz$ и $f_{c2} = 15kHz$ и номинално характеристично съпротивление $R = 600\Omega$. Да се определи затихването (в непери и в децибели) на филтъра за честоти $f_1 = 5kHz$, $f_2 = 13kHz$ и $f_3 = 20kHz$. Да се определят стойностите на характеристичните импеданси $Z_{\text{Т}}$ и $Z_{\text{П}}$, както и характерът им за трите посочени честоти. Да се начертаят Г-, Т- и П-схемите на лентовия филтър със съответните стойности на бобините и кондензаторите. Какво товарно съпротивление трябва да се включи откъм Т-страна на филтъра? А откъм П-страна?

3. Антенният ток на предавател за амплитудно-модулирани сигнали, модулиран с дълбочина 40% чрез звуков синусоидален сигнал, е $5{,}2A$. Той нараства

на $6A$ като резултат от модулация и с друга синусоидална вълна. Определете какъв е индекса на модулация, дължащ се на втората вълна и начертайте в подходящ мащаб спектралните диаграми за двата случая, като посочите амплитудите и честотите на съставките. Приемете, че първата модулираща вълна е с по-малка честота от втората.

Част II:

1. Да се синтезира лентов активен филтър с един операционен усилвател по модела с многоконтурна отрицателна обратна връзка със средна честота $f_0 = 750Hz$, реализиращ предавателната функция $T(p) = \dfrac{-2p}{p^2 + 0{,}25p + 1}$, като се използва петелементна конфигурация с C вход (C_1, C_2, R_3, R_4, C_5). Денормирането по съпротивление да се извърши с нормиращо съпротивление $R_N = 100k\Omega$. При синтеза да се обоснове лентовият характер на филтъра при посочената конфигурация за синтезираната схема. Да се извърши избор на стандартни стойности на резисторите и кондензаторите в проектирания филтър.

2. Да се изведе израз за предавателната функция на лентов активен филтър с един операционен усилвател по модела с многоконтурна отрицателна обратна връзка при зададената конфигурация (от точка 1), като се използва теорията на сигналните графи (по избор).

3. Да се определят полюсният качествен фактор на схемата, средната честота и коефициентът на усилване в лентата на пропускане за проектирания филтър след избора на стандартни стойности на резисторите и кондензаторите в схемата.

4. Да се изследва схемата на проектирания филтър след избора на стандартни стойности на резисторите и кондензаторите с използване на програмния продукт MicroCAP, като се снемат характеристиките (АЧХ, ФЧХ и затихване) на филтъра и по снетите характеристики се определят средната честота, коефициентът на усилване в лентата на пропускане и полюсният качествен фактор на схемата, като за целта предварително се измери широчината на честотната лента на филтъра.

3. РЕЗУЛТАТИ

Разработена е програмна система, базирана на MS EXCEL, за генериране на задания за курсови задачи по дисциплината „Комуникационни вериги”. В публикацията са приложени таблици, съдържащи получените резултати.

Програмната система обхваща: **1)** оразмеряването на секционирани трептящи кръгове с капацитивен (таблица 1) или с индуктивен делител (таблица 2), трансформатор с един настроен трептящ кръг (таблица 3); **2)** синтез и анализ на лентови (таблица 4) и заграждащи (таблица 5) LC филтри тип „К”, последователно-производни и паралелно-производни нискочестотни и високочестотни

LC филтри тип „m” (таблица 6 ÷ таблица 9), пасивни RC филтри – нискочестотни (таблица 10), високочестотни (таблица 11) и лентови по схемата на Робинзон-Вин (таблица 12); **3)** два типа задачи за амплитудна модулация (таблица 13 и таблица 14), ъглова модулация с двете си разновидности – честотна (таблица 15) и фазова (таблица 16), въздействие на амплитудно-модулирани сигнали върху трептящи кръгове при настроен (таблица 17) и при разстроен трептящ кръг (таблица 18) и импулсни модулации – амплитудно-импулсна модулация (таблица 19) и импулсно-кодова модулация (таблица 20). От активните RC филтри, обект на изследване на втората част на курсовата задача, са обхванати шест схемни варианта на лентови филтри с един операционен усилвател по метода с многоконтурна отрицателна обратна връзка (таблица 21 ÷ таблица 26). Алгоритмите за синтез и анализ на разглежданите комуникационни вериги са поместени в литературата [1, 2, 3, 4, 5]. Маркираните в таблиците клетки са първоначално зададените стойности.

Таблица 1. Трептящи кръгове с капацитивен делител

										точни				приблизителни	
B, kHz	f_0,MHz	R_t,kΩ	R_2,Ω	Q	ρ,Ω	L,μH	C,pF	N	Q_p	Q_p	C_2,pF	C_{se},pF	C_1,pF	C_2,pF	C_1,pF
20	2	10	100	100	100	7,96	795,77	10	10	9,95	7918,26	7998,23	883,70	7957,75	884,19

Таблица 2. Трептящи кръгове с индуктивен делител

										точни				приблизителни	
B, kHz	f_0,MHz	R_t,kΩ	R_2,Ω	Q	ρ,Ω	L,μH	C,pF	N	Q_p	Q_p	L_2,μH	L_{se},μH	L_1,μH	L_2,μH	L_1,μH
50	5	5	50	100	50	1,59	636,62	10	10	9,95	0,16	0,16	1,43	0,16	1,43

Таблица 3. Трансформатор с настроен трептящ кръг

B, kHz	f_0,MHz	R_t,kΩ	R_2,Ω	Q_p	Q	ρ,Ω	L_0,μH	C,pF	N	$Q_{p\,min}$	k_{min}	k^2	k	L_2,μH	L_1,μH	M,μH
150	15	5	50	10	100	50	0,53	212,21	10	0,99	0,14	0,10	0,32	0,05	0,53	0,05

Таблица 4. LC филтър, тип „K”, лентов – част 1

R,Ω	f_{c1},kHz	f_{c2},kHz	f_1,kHz	f_2,kHz	f_3,kHz	L_1,mH	L_2,mH	C_1,nF	C_2,nF	Ω_1	Ω_2	Ω_3	α_1,Np	α_1,dB	α_2,Np	α_2,dB	α_3,Np	α_3,dB
600	10	15	5	13	20	38,20	1,592	4,42	106,10	5,0000	0,2923	2,5000	2,2924	19,9121	0	0	1,5668	13,6092

Таблица 4. LC филтър, тип „K”, лентов – част 2

										Т	П
Z_{T1},Ω	Z_{T2},Ω	Z_{T3},Ω	$Z_{П1}$,Ω	$Z_{П2}$,Ω	$Z_{П3}$,Ω	0,5.L_1,mH	2.C_1,nF	2.L_2,mH	0,5.C_2,nF	R_L,Ω	R_L,Ω
j 2939,39	573,79	j 1374,77	–j 122,47	627,40	–j 261,86	19,10	8,84	3,18	53,05	480	750

Таблица 5. LC филтър, тип „K”, режекторен – част 1

R,Ω	f_{c1},kHz	f_{c2},kHz	f_1,kHz	f_2,kHz	f_3,kHz	L_1,mH	L_2,mH	C_1,nF	C_2,nF	Ω_1	Ω_2	Ω_3	α_1,Np	α_1,dB	α_2,Np	α_2,dB	α_3,Np	α_3,dB
600	10	20	5	13	15	9,55	4,77	13,26	26,53	0,2857	4,1935	6,0000	0	0	2,1122	18,3463	2,4779	21,5229

Таблица 5. LC филтър, тип „K”, режекторен – част 2

										Т	П
Z_{T1},Ω	Z_{T2},Ω	Z_{T3},Ω	$Z_{П1}$,Ω	$Z_{П2}$,Ω	$Z_{П3}$,Ω	0,5.L_1,mH	2.C_1,nF	2.L_2,mH	0,5.C_2,nF	R_L,Ω	R_L,Ω
574,99	j 2443,54	j 3549,65	626,10	–j 147,33	–j 101,42	4,77	26,53	9,55	13,26	480	750

Таблица 6. LC филтър, тип „m”, последователно-производен, НЧФ – част 1

				К		К, Г-схема		m			m, Г-схема		
R,Ω	f_c,kHz	m	f_∞,Hz	L_1,mH	C_2,nF	0,5.L_1,mH	0,5.C_2,nF	L_{1m},mH	C_{2m},nF	L_{2m},mH	0,5.L_{1m},mH	0,5.C_{2m},nF	2.L_{2m},mH
600	3	0,53	3537,75	63,66	176,84	31,83	88,42	33,74	93,72	21,59	16,87	46,86	43,19

Таблица 6. LC филтър, тип „m”, последователно-производен, НЧФ – част 2

m, П-схема				
0,5.C_{2m},nF	2.L_{2m},mH	1.L_{1m},mH	0,5.C_{2m},nF	2.L_{2m},mH
46,86	43,19	33,74	46,86	43,19

Таблица 7. LC филтър, тип „m”, последователно-производен, ВЧФ – част 1

				К		К, Г-схема		m			m, Г-схема		
R,Ω	f_c,kHz	m	f_∞,Hz	C_1,nF	L_2,mH	2.C_1,nF	2.L_2,mH	C_{1m},nF	L_{2m},mH	C_{2m},nF	2.C_{1m},nF	2.L_{2m},mH	0,5.C_{2m},nF
600	5	0,51	4300,87	26,53	9,55	53,05	19,10	52,01	18,72	73,14	104,02	37,45	36,57

Таблица 7. LC филтър, тип „m”, последователно-производен, ВЧФ – част 2

m, П-схема				
2.L_{2m},mH	0,5.C_{2m},nF	1.C_{1m},nF	2.L_{2m},mH	0,5.C_{2m},nF
37,45	36,57	52,01	37,45	36,57

Таблица 8. LC филтър, тип „m”, паралелно-производен, НЧФ – част 1

				К		К, Г-схема		m			m, Г-схема		
R,Ω	f_c,kHz	m	f_∞,Hz	L_1,mH	C_2,nF	0,5.L_1,mH	0,5.C_2,nF	L_{1m},mH	C_{1m},nF	C_{2m},nF	0,5.L_{1m},mH	2.C_{1m},nF	0,5.C_{2m},nF
600	2	0,55	2394,74	95,49	265,26	47,75	132,63	52,52	84,10	145,89	26,26	168,20	72,95

Таблица 8. LC филтър, тип „m”, паралелно-производен, НЧФ – част 2

m, Т-схема				
0,5.L_{1m},mH	2.C_{1m},nF	1.C_{2m},nF	0,5.L_{1m},mH	2.C_{1m},nF
26,26	168,20	145,89	26,26	168,20

Таблица 9. LC филтър, тип „m”, паралелно-производен, ВЧФ – част 1

				К		К, Г-схема		m			m, Г-схема		
R,Ω	f_c,kHz	m	f_∞,Hz	C_1,nF	L_2,mH	2.C_1,nF	2.L_2,mH	C_{1m},nF	L_{1m},mH	L_{2m},mH	2.C_{1m},nF	0,5.L_{1m},mH	2.L_{2m},mH
600	2,5	0,57	2054,11	53,05	19,10	106,10	38,20	93,07	64,50	33,51	186,15	32,25	67,01

Таблица 9. LC филтър, тип „m”, паралелно-производен, ВЧФ – част 2

m, Т-схема				
2.C_{1m},nF	0,5.L_{1m},mH	1.L_{2m},mH	2.C_{1m},nF	0,5.L_{1m},mH
186,15	32,25	33,51	186,15	32,25

Таблица 10. Пасивен RC филтър, НЧФ – част 1

R,kΩ	f_c,kHz	f_1,kHz	f_2,kHz	f_3,kHz	C,nF	x_1	x_2	x_3	T_1	T_2	T_3	α_1,dB	α_1,Np	α_2,dB	α_2,Np	α_3,dB	α_3,Np
1	2	1	2	3	79,58	0,5	1	1,5	0,8944	0,7071	0,5547	0,9691	0,1114	3,0103	0,3462	5,1188	0,5887

Таблица 10. Пасивен RC филтър, НЧФ – част 2

φ_1,°	φ_2,°	φ_3,°
-26,5651	-45,0000	-56,3099

Таблица 11. Пасивен RC филтър, ВЧФ – част 1

R,kΩ	f_c,kHz	f_1,kHz	f_2,kHz	f_3,kHz	C,nF	x_1	x_2	x_3	T_1	T_2	T_3	α_1,dB	α_1,Np	α_2,dB	α_2,Np	α_3,dB	α_3,Np
100	1	0,5	1	2	1,59	0,5	1	2,0	0,4472	0,7071	0,8944	6,9897	0,8038	3,0103	0,3462	0,9691	0,1114

Таблица 11. Пасивен RC филтър, ВЧФ – част 2

φ_1,°	φ_2,°	φ_3,°
63,4349	45,0000	26,5651

Таблица 12. Пасивен RC филтър, ЛФ по схемата на Робинзон-Вин

		m = 2												
R,kΩ	f_0,kHz	f_1,kHz	f_2,kHz	C,nF	x_1	x_2	T_1	T_2	α_1,dB	α_1,Np	α_2,dB	α_2,Np	φ_1,°	φ_2,°
10	0,05	1	0,1	318,31	20,0	2,0	0,0983	0,4682	20,1494	2,3172	6,5920	0,7581	-78,6625	-20,5560

Таблица 13. Амплитудна модулация – тип I

m_{a1}	m_{a2}	m_{a3}	f_m,Hz	R,Ω	A_c,V	m_a^2	m_a	m_a,%	P_c,W	P_t,W	B_{AM},Hz
0,5	0,7	0,3	300	5	10	0,83	0,911	91,1	10	14,15	1800

Таблица 14. Амплитудна модулация – тип II

m_{a1},%	I_{t1},A	I_t,A	m_{a1}	I_c,A	m_a^2	m_a	m_a,%	m_{a2}^2	m_{a2}	m_{a2},%	$0,5.m_{a1}.I_c$	$0,5.m_{a2}.I_c$
40	5,2	6	0,4	5	0,8757	0,936	93,6	0,7157	0,846	84,6	1,0	2,1

Таблица 15. Ъглова модулация (честотна модулация) – част I

				Беселови функции						Съставки, V					
U_c,V	m_f	f_m,kHz	R,Ω	J_0	J_1	J_2	J_3	J_4	J_5	ω_c	$\omega_c \pm \omega_m$	$\omega_c \pm 2\omega_m$	$\omega_c \pm 3\omega_m$	$\omega_c \pm 4\omega_m$	$\omega_c \pm 5\omega_m$
10	2,5	10	5	-0,05	0,50	0,45	0,22	0,07	0,02	-0,50	± 5,00	4,50	± 2,20	0,70	± 0,20

Таблица 15. Ъглова модулация (честотна модулация) – част II

B_{FM},kHz	Δf_{max},kHz	P_c,W	P_t,W
100	25,0	10,00	5,07

Таблица 16. Ъглова модулация (фазова модулация) – част I

					Беселови функции							Съставки, V				
U_c,V	ω_c,rad/s	ω_m,rad/s	m_p,rad	R,Ω	J_0	J_1	J_2	J_3	J_4	f_c,kHz	f_m,kHz	ω_c	$\omega_c \pm \omega_m$	$\omega_c \pm 2\omega_m$	$\omega_c \pm 3\omega_m$	$\omega_c \pm 4\omega_m$
15	$50 .10^3.\pi$	$0,8 .10^3.\pi$	2	10	0,22	0,58	0,35	0,13	0,03	25	0,4	3,30	± 8,70	5,25	± 1,95	0,45

Таблица 16. Ъглова модулация (фазова модулация) – част II

B_{FM},kHz	Δf_{max},kHz	P_c,W	P_t,W	ω(t)
3,2	0,80	11,25	5,64	$50 .10^3.\pi + 1,6 .10^3.\pi .\cos(0,8 .10^3.\pi .t)$

Таблица 17. Въздействие на амплитудно-модулирани сигнали върху трептящи кръгове – тип I (настроен трептящ кръг) – част I

U_c,V	ω_c,rad/s	ω_m,rad/s	m_a	L,μH	C,pF	R,Ω	Q	Q.U_c,V	$2Q\omega_m/\omega_c$	φ,rad	φ,rad	φ_c,rad	m_a'
0,6	$4 .10^6$	$1 .10^4$	0,4	625	100	20	125	75	0,625	0,5586	0,178 π	0,5 π	0,339

Таблица 17. Въздействие на амплитудно-модулирани сигнали върху трептящи кръгове – тип I (настроен трептящ кръг) – част II

Съставки								
Честоти, rad/s			Амплитуди, V - ВХОД			Амплитуди, V - ИЗХОД		
ω_c	$\omega_c-\omega_m$	$\omega_c+\omega_m$	U_c	$0,5.m_aU_c$	$0,5.m_aU_c$	Q.U_c	$0,5.m_a'.Q.U_c$	$0,5.m_a'.Q.U_c$
$400 .10^4$	$399 .10^4$	$401 .10^4$	0,6	0,12	0,12	75,00	12,72	12,72

Таблица 17. Въздействие на амплитудно-модулирани сигнали върху трептящи кръгове – тип I (настроен трептящ кръг) – част III

$u_{AM}'(t)$,V
75 .[1+ 0,339 .cos(1 .10^4.t – 0,178 .π)] .cos(4 .10^6.t – 0,5π)

Таблица 18. Въздействие на амплитудно-модулирани сигнали върху трептящи кръгове – тип II (разстроен трептящ кръг) – част I

U_c,V	ω_c,rad/s	ω_m,rad/s	m_a	ω_0,rad/s	Q	$K(\omega_c)$	$K(\omega_c-\omega_m)$	$K(\omega_c+\omega_m)$
2	4 .10^6	8 .10^4	0,6	397 .10^4	50	39,89	31,09	16,97

Таблица 18. Въздействие на амплитудно-модулирани сигнали върху трептящи кръгове – тип II (разстроен трептящ кръг) – част II

Съставки								
Честоти, rad/s			Амплитуди, V - ВХОД			Амплитуди, V - ИЗХОД		
ω_c	$\omega_c-\omega_m$	$\omega_c+\omega_m$	U_c	$0,5.m_aU_c$	$0,5.m_aU_c$	$K(\omega_c).U_c$	$0,5.m_a.U_c.K(\omega_c-\omega_m)$	$0,5.m_a.U_c.K(\omega_c+\omega_m)$
400 .10^4	392 .10^4	408 .10^4	2,0	0,60	0,60	79,78	18,65	10,18

Таблица 19. Амплитудно-импулсна модулация

					1	2	3	Съставки: честоти, rad/s; амплитуди, V										
A_m,V	A_c,V	N	m_{pa}	a_0	a_1	a_2	a_3	0	ω_m	ω_c	$\omega_c-\omega_m$	$\omega_c+\omega_m$	$2\omega_c$	$2\omega_c-\omega_m$	$2\omega_c+\omega_m$	$3\omega_c$	$3\omega_c-\omega_m$	$3\omega_c+\omega_m$
5	10	5	0,5	0,2	0,4	0,3	0,2	2,00	1,00	4,00	1,00	1,00	3,00	0,75	0,75	2,00	0,50	0,50

Таблица 20. Импулсно-кодова модулация

$u(t) = a.(e^{b.t} - 1)$

a,V	b,s^{-1}	t,s	0	1	2	3	4	5	6	7
1,3	0,3	u_A,V	0	0,4548	1,0688	1,8975	3,0162	4,5262	6,5645	9,3160
		u_D,V	0	0	1	2	3	5	7	9
		КК	000	000	001	010	011	101	111	-

Таблица 21. Проектиране на активен филтър с един операционен усилвател по метода с едноконтурна отрицателна обратна връзка: C1C2R3R4C5

a	b_1	b_0	R_N,kΩ	f_0,Hz	ст.	C_1	C_2	G_3	R_3	G_4	R_4	C_5	ω_0,rad/s	f_0,Hz	Q	k_0
2	0,25	1	100	750	Н	2,0000	6,0000	1,0000	1,0000	1,0000	1,0000	0,1250	-	-	4,0000	-
					Д,nF/kΩ	4,2441	12,7324	-	100,0000	-	100,0000	0,2653	4712,3890	750,0000	4,0000	8,0000
					С,nF/kΩ	4,3000	13,0000	-	100,0000	-	100,0000	0,2700	4626,9528	736,4024	4,0023	7,9630

Таблица 22. Проектиране на активен филтър с един операционен усилвател по метода с едноконтурна отрицателна обратна връзка: C1R2R3R4C5

a	b_1	b_0	R_N,kΩ	f_0,Hz	ст.	C_1	G_2	R_2	G_3	R_3	G_4	R_4	C_5	ω_0,rad/s	f_0,Hz	Q	k_0
10	1,65	1	100	250	Н	10,0000	14,5000	0,0690	1,0000	1,0000	1,0000	1,0000	0,1000	-	-	0,6061	-
					Д,nF/kΩ	63,6620	-	6,8966	-	100,0000	-	100,0000	0,6366	1570,7963	250,0000	0,6061	6,0606
					С,nF/kΩ	62,0000	-	6,8000	-	100,0000	-	100,0000	0,6200	1612,9032	256,7015	0,5986	5,9859

Таблица 23. Проектиране на активен филтър с един операционен усилвател по метода с едноконтурна отрицателна обратна връзка: C1--R3R4C5

a	b_1	b_0	R_N,kΩ	f_0,Hz	ст.	C_1	G_3	R_3	G_4	R_4	C_5	ω_0,rad/s	f_0,Hz	Q	k_0
2	1,55	1	10	1500	Н	1,3801	1,4491	0,6901	0,6901	1,4491	0,7246	-	-	0,6452	-
					Д,nF/kΩ	14,6436	-	6,9007	-	14,4914	7,6879	9424,7780	1500,0000	0,6452	1,2903
					С,nF/kΩ	15,0000	-	6,8000	-	15,0000	7,5000	9335,2006	1485,7433	0,6552	1,3761

Таблица 24. Проектиране на активен филтър с един операционен усилвател по метода с едноконтурна отрицателна обратна връзка: R1C2C3C4R5

a	b_1	b_0	R_N,kΩ	f_0,Hz	ст.	G_1	R_1	C_2	C_3	C_4	G_5	R_5	ω_0,rad/s	f_0,Hz	Q	k_0
9	1,95	1	10	3000	H	9,0000	0,1111	15,5500	1,0000	1,0000	0,1111	9,0000	-	-	0,5128	-
					Д,nF/kΩ	-	1,1111	82,4953	5,3052	5,3052	-	90,0000	18849,5559	3000,0000	0,5128	4,6154
					C,nF/kΩ	-	1,1000	82,0000	5,1000	5,1000	-	91,0000	19598,0466	3119,1260	0,5031	4,5760

Таблица 25. Проектиране на активен филтър с един операционен усилвател по метода с едноконтурна отрицателна обратна връзка: R1R2C3C4R5

a	b_1	b_0	R_N,kΩ	f_0,Hz	ст.	G_1	R_1	G_2	R_2	C_3	C_4	G_5	R_5	ω_0,rad/s	f_0,Hz	Q	k_0
2	0,25	1	100	750	H	2,0000	0,5000	6,0000	0,1667	1,0000	1,0000	0,1250	8,0000	-	-	4,0000	-
					Д,nF/kΩ	-	50,0000	-	16,6667	2,1221	2,1221	-	800,0000	4712,3890	750,0000	4,0000	8,0000
					C,nF/kΩ	-	51,0000	-	16,0000	2,2000	2,2000	-	820,0000	4548,4426	723,9071	4,1027	8,0392

Таблица 26. Проектиране на активен филтър с един операционен усилвател по метода с едноконтурна отрицателна обратна връзка: R1--C3C4R5

a	b_1	b_0	R_N,kΩ	f_0,Hz	ст.	G_1	R_1	C_3	C_4	G_5	R_5	ω_0,rad/s	f_0,Hz	Q	k_0
9	1,45	1	100	2500	H	2,5927	0,3857	3,4713	0,2881	0,3857	2,5927	-	-	0,6897	-
					Д,nF/kΩ	-	38,5701	2,2099	0,1834	-	259,2680	15707,9633	2500,0000	0,6897	6,2069
					C,nF/kΩ	-	39,0000	2,2000	0,1800	-	270,0000	15485,9631	2464,6676	0,6957	6,3995

4. ИЗВОДИ

Разработената програмна система, базирана на MS EXCEL и описана в публикацията, се използва за автоматизиране на процеса на генериране на варианти на задания за курсови задачи по дисциплината „Комуникационни вериги”, включена като задължителна в действащия учебен план на специалността „Комуникационна техника и технологии” за образователно-квалификационна степен „Бакалавър”.

ЛИТЕРАТУРА

[1] Манукова, А., А. Бороджиева. Комуникационни вериги – ръководство за упражнения. 104 стр., Русе, 2002.

[2] Стоянов, Г. Теоретични основи на съобщителната техника. София, Техника, 1993.

[3] Huelsman, L.P., P.E. Allen. “Introduction to the Theory and Design of Active Filters”, Moscow, 1984 (in Russian).

[4] Stoyanova, T., A. Borodzhieva. Analysis and Synthesis of Band-Pass Single-Amplifier Multiple-Loop Negative Feedback Active Filters Using MATLAB and MicroCAP. Summer School, Advanced Aspects of Theoretical Electrical Engineering, Sozopol ’07, 22.09. – 25.09.2007, Sozopol, Bulgaria, Regular papers, Part II, pp. 64 – 69.

[5] Stoyanova, T., A. Borodzhieva. Design and Computer Simulation of Low-Pass and High-Pass Single-Amplifier Multiple-Loop Negative Feedback Active Filters Using MATLAB and MicroCAP. Summer School, Advanced Aspects of Theoretical Electrical Engineering, Sozopol ’07, 22.09. – 25.09.2007, Sozopol, Bulgaria, Regular papers, Part II, pp. 58 – 63.

ПРИЛОЖЕНИЕ НА ТЕОРИЯТА НА СИГНАЛНИТЕ ГРАФИ ПРИ АНАЛИЗА НА АКТИВНИ БИКВАДРАТНИ ФИЛТРОВИ ЗВЕНА

Таня Методиева Стоянова[1], Адриана Найденова Бороджиева[2]

[1]Катедра „Теоретична и измервателна електротехника”,
[2]Катедра „Комуникационна техника и технологии”, Русенски университет „Ангел Кънчев”,
България, 7017 Русе, ул. „Студентска” № 8, тел.: [1](00359 82) 888 502, [2](00359 82) 888 734,
e-mail: [1]tstoyanova@uni-ruse.bg, [2]aborodjieva@ecs.uni-ruse.bg

Резюме: *В тази публикация е показано приложението на теорията на сигналните графи при извеждане на предавателната функция на активни нискочестотни, високочестотни и заграждащи биквадратни филтрови звена. Резултатите ще бъдат използвани в процеса на обучение по дисциплината „Комуникационни вериги”, изучавана от студентите от специалност „Комуникационна техника и технологии” от образователно-квалификационната степен „Бакалавър”.*

Ключови думи: Сигнални графи, анализ, биквадратни активни филтри.

1. ВЪВЕДЕНИЕ

Обект на анализ в публикацията са биквадратните активни филтрови звена с един операционен усилвател. Използваният метод за анализ на разглежданата линейна електронна схема е методът на насочените графи. По-долу ще бъде изведен изразът за предавателната функция, използвайки теорията на графите, която предоставя практически методи за изчисляване на предавателните функции и на входните и изходните импеданси на схемата.

Топологичните методи се основават на графичното представяне и обработване на информацията за електронната схема, в резултат на което се определят свойствата й [1].

Сигналният граф представлява схема, която се състои от точки (възли), свързани с насочени линии (ребра), и изразява система алгебрични уравнения. Възлите на сигналния граф не съвпадат с възлите на схемата, а съответстват на променливите, например напрежения или токове. При сигналните графи, възлите и свързващите ги линии изразяват линейни алгебрични уравнения.

2. СЪСТОЯНИЕ НА ПРОБЛЕМА – ПРИЛОЖЕНИЕ НА ТЕОРИЯТА НА СИГНАЛНИТЕ ГРАФИ ПРИ ИЗВЕЖДАНЕ НА ПРЕДАВАТЕЛНАТА ФУНКЦИЯ НА АКТИВНИ НИСКОЧЕСТОТНИ, ВИСОКОЧЕСТОТНИ И ЗАГРАЖДАЩИ БИКВАДРАТНИ ФИЛТРИ

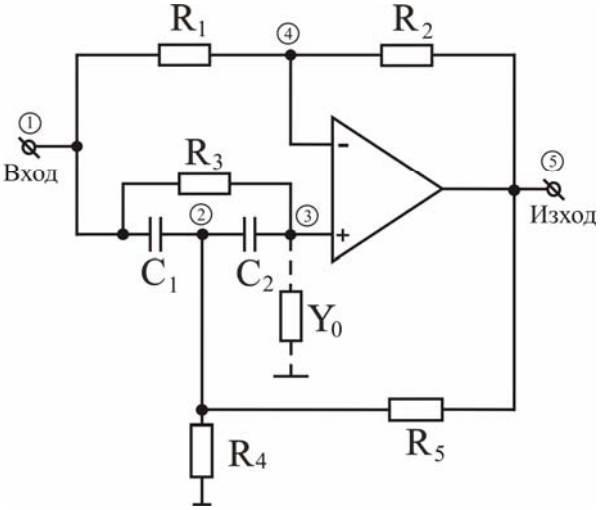
Ако два върха i и j са свързани с насочен граф, то тегло на графа се нарича коефициентът a_{ij}, който показва връзката между двете променливи x_i и x_j:

$x_i = a_{ij} x_j$. За определяне на тегловния коефициент между два произволни възела (най-често входящия и изходящия възел) се използва формулата на Мейсън, като първоначално се налага модифициране на графа, при което се елиминират всички клонове, които влизат във входящия възел – връх-корен.

За да се приложи класическата теория на насочените графи във веригите, съдържащи операционни усилватели, те се разглеждат в режим на малки сигнали, като се използват заместващи схеми, съдържащи елементи с едностранна проводимост − унистори. За да се различава от останалите сигнални графи, посоката му се означава с триъгълник, вместо със стрелка. Теглото на унисторния граф може да приема както положителни (връзка между изхода на операционния усилвател и неговия неинвертиращ вход), така и отрицателни стойности (връзка между изхода на операционния усилвател и неговия инвертиращ вход). Унисторният граф елиминира всички неунисторни клонове, които влизат във възела, към който е насочен.

На фиг. 1 е представена схемата на разглежданите биквадратни активни филтрови звена с един операционен усилвател [2, 3], с означени възли за построяване на сигналния граф, изобразен на фиг. 2, в който с пунктир са посочени всички ребра, които се елиминират.

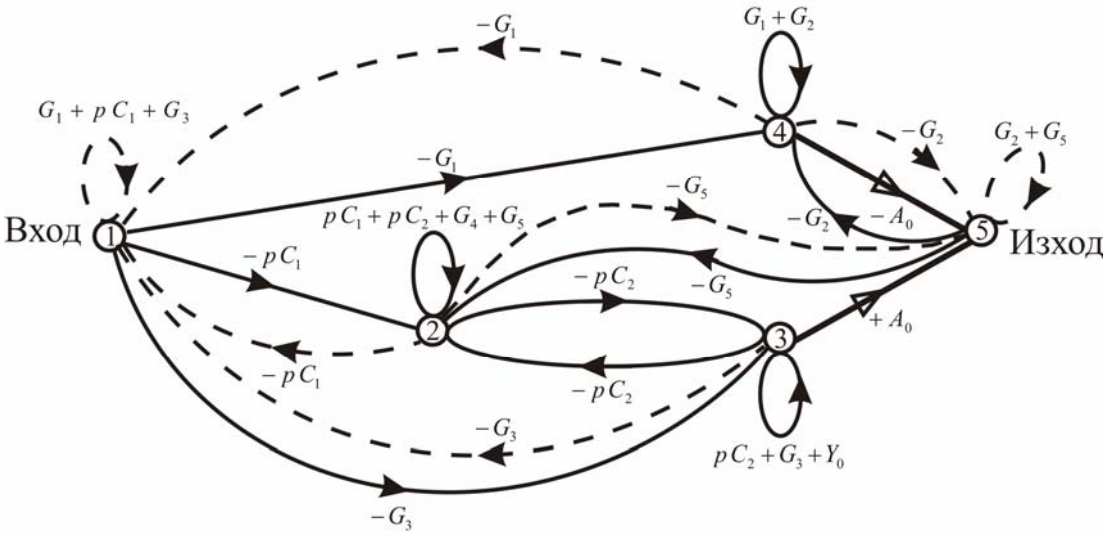
Предавателната функция на схемата може да се определи от относителния тегловен коефициент: $N_{ij} = \dfrac{\sum P_{ij}^k . \Delta^k}{\Delta}$, където i е входящият възел, j е изходящият възел, P_{ij}^k е директен (пряк) път, Δ^k е адюнгираното количество на този път, описващо частта от графа, недопираща се до пътя, а Δ е детерминантата на цялата схема [1].



Фиг. 1. Схема на биквадратни активни филтрови звена с един операционен усилвател с означени възли за построяване на сигналния граф

Контурите в сигналния граф на схемата са:

$$2 \to 2:\ pC_1 + pC_2 + G_4 + G_5, \tag{1}$$
$$3 \to 3:\ pC_2 + G_3 + Y_0, \tag{2}$$
$$4 \to 4:\ G_1 + G_2, \tag{3}$$
$$2 \to 3 \to 2:\ (-pC_2).(-pC_2), \tag{4}$$
$$2 \to 3 \to 5 \to 2:\ (-pC_2).(+A_0).(-G_5), \tag{5}$$
$$4 \to 5 \to 4:\ (-A_0).(-G_2). \tag{6}$$



Фиг. 2. Сигнален граф на схемата

Между възел „1” и възел „5” може да премине по следните пътища (последователности от възли): $1 \to 4 \to 5$ (P'_{15}), $1 \to 3 \to 5$ (P''_{15}) и $1 \to 2 \to 3 \to 5$ (P'''_{15}), за които се определят съответно P'_{15} и Δ', P''_{15} и Δ'', P'''_{15} и Δ''':

$$P'_{15} = (-G_1).(-A_0) = G_1.A_0, \tag{7}$$

$$\begin{aligned}\Delta'_k &= (-1)^2.(pC_1 + pC_2 + G_4 + G_5).(pC_2 + G_3 + Y_0) + (-1)^1.(-pC_2).(-pC_2) = \\ &= (pC_1 + pC_2 + G_4 + G_5).(G_3 + Y_0) + pC_2.(pC_1 + G_4 + G_5),\end{aligned} \tag{8}$$

$$P''_{15} = (-G_3).(+A_0) = -G_3.A_0, \tag{9}$$

$$\begin{aligned}\Delta''_k &= (-1)^2.(pC_1 + pC_2 + G_4 + G_5).(G_1 + G_2) = \\ &= (pC_1 + pC_2 + G_4 + G_5).(G_1 + G_2),\end{aligned} \tag{10}$$

$$P'''_{15} = (-pC_1).(-pC_2).(+A_0) = p^2C_1C_2.A_0, \tag{11}$$

$$\Delta'''_k = (-1)^1.(G_1 + G_2) = -(G_1 + G_2), \tag{12}$$

където: P_{ij}^k е произведението от проводимостите на клоните по пътя “k”, като проводимостта на изходния клон се приема за равна на единица, а Δ^k е адюнги-

раното количество, получено от изходната система при свързване на късо на всички клони по пътя "k" (ако пътят минава през всички възли $\Delta^k = 1$).

Предавателната функция (относителният тегловен коефициент на графа) може лесно да се намери без преобразувания, а направо от сигналния граф, като се използва формула на Мейсън [1]:

$$T(p) = \frac{N(p)}{D(p)} = \frac{\sum P_{15}^k \Delta^k}{\Delta}, \tag{13}$$

където числителят $N(p)$ има вида:

$$\begin{aligned} N(p) &= \sum P_{15}.\Delta_k = P'_{15}.\Delta'_k + P''_{15}.\Delta''_k + P'''_{15}.\Delta'''_k = \\ &= A_0.(pC_1 + pC_2 + G_4 + G_5).(G_1.Y_0 - G_2.G_3) + pC_2.(G_4 + G_5).G_1.A_0 - p^2 C_1 C_2.A_0.G_2, \end{aligned} \tag{14}$$

а знаменателят $D(p)$ има вида:

$$\begin{aligned} D(p) &= \Delta_{A_0} = (-1)^3.(pC_1 + pC_2 + G_4 + G_5).(pC_2 + G_3 + Y_0).(-A_0).(-G_2) + \\ &+ (-1)^2.(-pC_2).(-pC_2).(-A_0).(-G_2) + (-1)^2.(-pC_2).(+A_0).(-G_5).(G_1 + G_2) = \\ &= -G_2.(pC_1 + pC_2 + G_4 + G_5).(G_3 + Y_0).A_0 - pC_2.G_2.(pC_1 + G_4).A_0 + pC_2.G_5.G_1.A_0. \end{aligned} \tag{15}$$

Следователно предавателната функция на активни филтрови звена с един операционен усилвател може да се запише по следния начин:

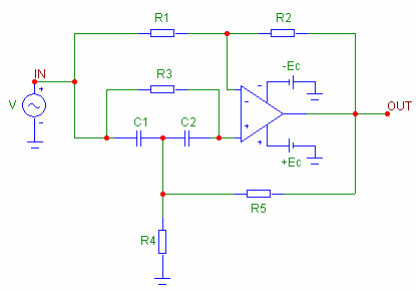
$$T(p) = \frac{A_0.(pC_1 + pC_2 + G_4 + G_5).(G_1.Y_0 - G_2.G_3) + pC_2.(G_4 + G_5).G_1.A_0 - p^2 C_1 C_2.A_0.G_2}{-G_2.(pC_1 + pC_2 + G_4 + G_5).(G_3 + Y_0).A_0 - pC_2.G_2.(pC_1 + G_4).A_0 + pC_2.G_5.G_1.A_0}. \tag{16}$$

В зависимост от вида на проводимостта Y_0 се различават следните случаи:

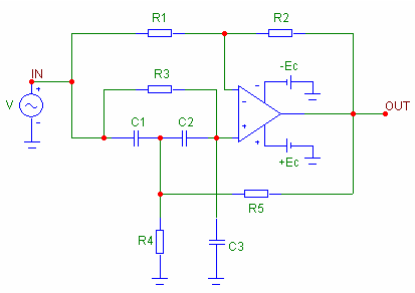
Случай 1: Заграждащ (режекторен) филтър – $Y_0 = 0$ (фиг. 3);

Случай 2: Нискочестотен полиномен филтър – $Y_0 = pC_3$ (фиг. 4);

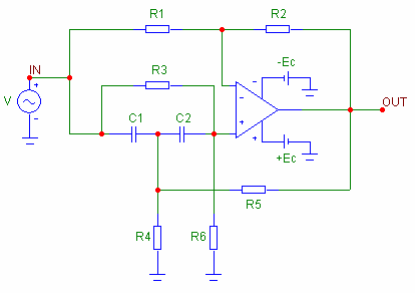
Случай 3: Високочестотен полиномен филтър – $Y_0 = G_6$ (фиг. 5).



Фиг. 3. Заграждащ (режекторен) филтър



Фиг. 4. Нискочестотен полиномен филтър



Фиг. 5. Високочестотен полиномен филтър

3. РЕЗУЛТАТИ

Случай 1: Заграждащ режекторен (режекторен) филтър – $Y_0 = 0$.

В този случай, за числителя $N(p)$, знаменателя $D(p)$ и предавателната функция $T(p)$ се получават съответно изразите:

$$N(p) = A_0.(pC_1 + pC_2 + G_4 + G_5).(-G_2.G_3) + pC_2.(G_4 + G_5).G_1.A_0 - p^2C_1C_2.A_0.G_2 = \\ = -p^2C_1C_2.G_2.A_0 - p.A_0.[G_2.G_3.(C_1 + C_2) - G_1.C_2.(G_4 + G_5)] - G_2.G_3.(G_4 + G_5).A_0, \quad (17)$$

$$D(p) = -G_2.(pC_1 + pC_2 + G_4 + G_5).G_3.A_0 - pC_2.G_2.(pC_1 + G_4).A_0 + pC_2.G_5.G_1.A_0 = \\ = -p^2C_1C_2.G_2.A_0 - p.A_0.[G_2.G_3.(C_1 + C_2) + C_2.G_2.G_4 - C_2.G_1.G_5] - G_2.G_3.(G_4 + G_5).A_0, \quad (18)$$

$$T(p) = \frac{p^2C_1C_2.G_2 + p.[G_2.G_3.(C_1 + C_2) - G_1.C_2.(G_4 + G_5)] + G_2.G_3.(G_4 + G_5)}{p^2C_1C_2.G_2 + p.[G_2.G_3.(C_1 + C_2) + C_2.G_2.G_4 - C_2.G_1.G_5] + G_2.G_3.(G_4 + G_5)}. \quad (19)$$

Условието за реализиране на предавателна функция на заграждащ филтър е коефициентът пред p^1 в числителя $N(p)$ да е равен на нула:

$$G_2.G_3.(C_1 + C_2) = G_1.C_2.(G_4 + G_5). \quad (20)$$

Тогава за предавателната функция на заграждащ филтър се записва:

$$T(p) = \frac{p^2C_1C_2.G_2 + G_2.G_3.(G_4 + G_5)}{p^2C_1C_2.G_2 + p.[G_2.G_3.(C_1 + C_2) + C_2.G_2.G_4 - C_2.G_1.G_5] + G_2.G_3.(G_4 + G_5)}. \quad (21)$$

След разделяне на числителя и знаменателя на предавателната функция с $C_1C_2.G_2$ се записва окончателно следния израз:

$$T(p) = \frac{p^2 + \dfrac{G_3.(G_4 + G_5)}{C_1C_2}}{p^2 + p.\left[\dfrac{G_3.(C_1 + C_2)}{C_1C_2} + \dfrac{G_4}{C_1} - \dfrac{G_1.G_5}{C_1.G_2}\right] + \dfrac{G_3.(G_4 + G_5)}{C_1C_2}}. \quad (22)$$

Случай 2: Нискочестотен полиномен филтър – $Y_0 = pC_3$.

В този случай, числителят и знаменателят се получават съответно:

$$
\begin{aligned}
&N(p) = A_0.(pC_1 + pC_2 + G_4 + G_5).(G_1.pC_3 - G_2.G_3) + \\
&+ pC_2.(G_4 + G_5).G_1.A_0 - p^2 C_1 C_2.A_0.G_2 = \\
&= -p^2.A_0.[C_1.C_2.G_2 - G_1.C_3.(C_1 + C_2)] - \\
&- p.A_0.[G_2.G_3.(C_1 + C_2) - G_1.C_2.(G_4 + G_5) - G_1.C_3.(G_4 + G_5)] - \\
&- G_2.G_3.A_0.(G_4 + G_5),
\end{aligned} \tag{23}
$$

$$
\begin{aligned}
&D(p) = -G_2.(pC_1 + pC_2 + G_4 + G_5).(G_3 + pC_3).A_0 - \\
&- pC_2.G_2.(pC_1 + G_4).A_0 + pC_2.G_5.G_1.A_0 = \\
&= -p^2.A_0.G_2.(C_1.C_2 + C_1.C_3 + C_2.C_3) - \\
&- p.A_0.[G_2.G_3.(C_1 + C_2) + G_2.G_4.(C_2 + C_3) + G_2.C_3.G_5 - C_2.G_1.G_5] - \\
&- G_2.G_3.(G_4 + G_5).A_0.
\end{aligned} \tag{24}
$$

Условието за реализиране на предавателна функция на нискочестотен полиномен филтър е коефициентът пред p^1 в числителя $N(p)$ да е равен на нула:

$$G_2.G_3.(C_1 + C_2) = G_1.(G_4 + G_5).(C_2 + C_3). \tag{25}$$

След отчитане на условието за реализиране на предавателна функция на нискочестотен полиномен филтър и след делене на числителя и знаменателя на предавателната функция с $-A_0$ се записва:

$$T(p) = \frac{p^2.[C_1.C_2.G_2 - G_1.C_3.(C_1 + C_2)] + G_2.G_3.(G_4 + G_5)}{p^2.G_2.(C_1.C_2 + C_1.C_3 + C_2.C_3) + p.[G_2.G_3.(C_1 + C_2) + G_2.G_4.(C_2 + C_3) + G_2.C_3.G_5 - C_2.G_1.G_5] + G_2.G_3.(G_4 + G_5)} \tag{26}$$

След разделяне на числителя и знаменателя на предавателната функция на филтъра с $C_1.C_2.G_2 + G_2.C_3.(C_1 + C_2)$ се записва:

$$T(p) = \frac{\dfrac{C_1.C_2.G_2 - G_1.C_3.(C_1 + C_2)}{G_2.(C_1.C_2 + C_1.C_3 + C_2.C_3)}.\left[p^2 + \dfrac{G_2.G_3.(G_4 + G_5)}{C_1.C_2.G_2 - G_1.C_3.(C_1 + C_2)}\right]}{p^2 + p.\left[\dfrac{G_3.(C_1 + C_2) + G_4.(C_2 + C_3) + C_3.G_5 - \dfrac{C_2.G_1.G_5}{G_2}}{C_1.C_2 + C_1.C_3 + C_2.C_3}\right] + \dfrac{G_3.(G_4 + G_5)}{C_1.C_2 + C_1.C_3 + C_2.C_3}} \tag{27}$$

Случай 3: Високочестотен полиномен филтър – $Y_0 = G_6$.

В този случай, за числителя $N(p)$ и знаменателя $D(p)$ на предавателната функция $T(p)$ се получават съответно изразите:

$$N(p)=A_0.(pC_1+pC_2+G_4+G_5).(G_1.G_6-G_2.G_3)+pC_2.(G_4+G_5).G_1.A_0-p^2C_1C_2.A_0.G_2=$$
$$=-p^2.A_0.C_1C_2.G_2-p.A_0.[(G_2.G_3-G_1.G_6).(C_1+C_2)-G_1.C_2.(G_4+G_5)]- \quad (28)$$
$$-A_0.(G_4+G_5).(G_2.G_3-G_1.G_6),$$

$$D(p)=-G_2.(pC_1+pC_2+G_4+G_5).(G_3+G_6).A_0-pC_2.G_2.(pC_1+G_4).A_0+pC_2.G_5.G_1.A_0=$$
$$=-p^2.A_0.C_1C_2.G_2-p.A_0.[C_2.G_2.G_4+G_2.(G_3+G_6).(C_1+C_2)-C_2.G_1.G_5]- \quad (29)$$
$$-G_2.(G_4+G_5).(G_3+G_6).A_0.$$

Условието за реализиране на предавателна функция на високочестотен полиномен филтър е коефициентът пред p^1 в числителя $N(p)$ да е равен на нула:

$$G_2.G_3.(C_1+C_2)=G_1.[C_2.(G_4+G_5)+G_6.(C_1+C_2)] \quad (30)$$

След отчитане на условието за реализиране на предавателна функция на високочестотен полиномен филтър и след делене на числителя и знаменателя на предавателната функция с $-A_0$:

$$T(p)=\frac{p^2.C_1C_2.G_2+(G_4+G_5).(G_2.G_3-G_1.G_6)}{p^2.C_1C_2.G_2+p.[C_2.G_2.G_4+G_2.(G_3+G_6).(C_1+C_2)-C_2.G_1.G_5]+G_2.(G_4+G_5).(G_3+G_6)}. \quad (31)$$

След разделяне на числителя и знаменателя на предавателната функция с $C_1C_2.G_2$ за предавателната функция на филтъра се записва:

$$T(p)=\frac{p^2+\frac{G_4+G_5}{C_1C_2}\left(G_3-\frac{G_1.G_6}{G_2}\right)}{p^2+p.\left[\frac{G_4}{C_1}+(G_3+G_6).\left(\frac{1}{C_1}+\frac{1}{C_2}\right)-\frac{G_1.G_5}{C_1.G_2}\right]+\frac{(G_4+G_5).(G_3+G_6)}{C_1C_2}}. \quad (32)$$

4. ИЗВОДИ

1. В публикацията е показано приложението на теорията на сигналните графи при извеждане на предавателната функция на активни нискочестотни, високочестотни и заграждащи биквадратни филтрови звена.

2. Резултатите ще бъдат използвани в процеса на обучение по дисциплината „Комуникационни вериги", изучавана от студентите от специалност „Комуникационна техника и технологии" от образователно-квалификационната степен „Бакалавър".

ЛИТЕРАТУРА

[1] Боянов, Й., Е. Шойкова. Теория на електронните схеми. София, Техника, 1995.
[2] Стоянов, Г. Теоретични основи на съобщителната техника. София, Техника, 1993.
[3] Huelsman, L.P., P.E. Allen. "Introduction to the Theory and Design of Active Filters", Moscow, 1984 (in Russian).

ХАРМОНИЗАЦИЯ НА УЧЕБНИЯ МАТЕРИАЛ ПО "ЕЛЕКТРИЧЕСКИ ВЕРИГИ" ("ОСНОВИ НА ЕЛЕКТРОТЕХНИКАТА") В СЪОТВЕТСТВИЕ С ЕВРОПЕЙСКИТЕ СТАНДАРТИ

Иван Стефанов Бозев[1], **Руси Александров Русев**[2]

[1]Катедра "Фундаментална подготовка", Висше държавно училище - "Колеж по телекомуникации и пощи", 1700, София, ул. "Академик Стефан Младенов" №1, България, phone: +359 2 8062 241, e-mail: IBozev@hctp.acad.bg

[2]Катедра "Фундаментална подготовка", Висше държавно училище - "Колеж по телекомуникации и пощи", 1700, София, ул. "Академик Стефан Младенов" №1, Българя, phone: +359 878 502 14 e-mail: rusi.russev@gmail.com

***Abstract**: В този доклад се разглеждат някой съществуващи разлики между установените в България традиции в обучението по електротехника и новоприетите европейски стандарти. Обърнато е внимание на по съществените разлики, който са направили впечатление на авторите. Направени са предложения за уточняване в национален мащаб на някои от елементите, имащи по няколко графични изображения. Поставен е въпросът за привеждането на учебната литература в съответствие с действащите стандарти.*

1. INTRODUCTION

Развитието на науката и технологиите изискват постоянно обновяване на стандартизираните термини и определения с цел по-лесно, по-ясно и разбира се еднозначно определяне на съответната информация. При разглеждане на източниците [1],…,[7] се вижда, че за периодът от 10 години (1969-1978) са настъпили някой изменения в термините и графичните изображения, използвани в електротехническа литература. В момента в Европа техническите стандарти се хармонизират (въвеждат се общи технически стандарти за всички европейски държави). Разбира се за страните от бившия "Съюз за икономическа взаимопомощ " (СИВ) изменението в техническите стандарти е по-голямо. От друга страна процеса на хармонизация е продължителен, и тогава е възможно сериозно разминаване между действащи стандарти и новоприети стандарти в областите, където се използват термини и определения на по общи стандарти. Така например в момента в България във областта на електротехниката има въведени европейски стандарти [10],…,[12] и все още действащи стандарти на СИВ [2],…,[7]. Между нововъведените стандарти и издадената до този момент електротехническа учебна литература се констатират някои съществени разлики.

2. PROBLEM STATEMENT

В стандарта БДС EN 80000-6:2008. Величини и единици. Част 6: Електромагнетизъм. (IEC 80000-6:2008 (EQV) Quantities and units – Part 6: Electromagne-

tism), както се вижда от заглавието, се определят допустимите за употреба величини и единици и техните означения в областта на електромагнетизма. В БДС EN 60375:2006. Споразумения, свързани с електрически и магнитни вериги (IEC 60375:2003. Conventions concerning electric and magnetic circuits) се разглеждат правилата за определяне на знаците, посоките и големините на отделните величини в електрическите схеми и магнитни вериги. Като български стандарти е необходимо те да бъдат спазвани, което, както беше казано по-горе ще доведе до еднозначност на информацията при съвременните методи на изчисление и обработка.

Нека да разгледаме последователно съществените разлики между тези стандарти и установени положения в учебната литература по електротехника.

В областта на термините и буквените означения има различни буквени означения при променлив ток: например при напрежението $\hat{U}$ е максималната стойност, а $\underline{U}$ е комплексната стойност. При импедансите и проводимостите при променлив ток за комплексните стойности съответно се използват подчертани главни букви ($\underline{Z}$,$\underline{Y}$), а за техните модули само главни букви (Z,Y).

В БДС EN 80000-6:2008 е дефиниран източник на напрежение, който се определя с неговото "напрежение на източника" u_s. Наименованието "електродвижеща сила ЕДС" (електродвижещо напрежение ЕДН) и символа E, използвани в литературата [1],[13],…,[19], се изключват от употреба в областта на електрическите вериги.

За източника на напрежение са стандартизирани две графични изображения, показани на Fig. 1 и Fig. 2, които се различават от използваното графично изображение в нашата литература (Fig. 3).

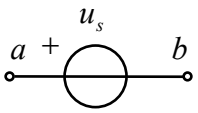


Fig.1

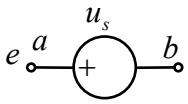


Fig.2

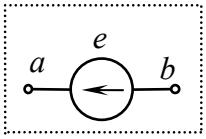


Fig.3

В стандартите на ISO и IEC понятието електродвижещ ток не е дефинирано като термин.

При това положение се преодоляват проблемите с различните посоки на електродвижещия ток при генераторите на ток ([13],[14],[18],[19]) и ([15],…,[17]), като се осигурява еднозначност между параметрите на дуалните елементи източник на напрежение с "напрежение на източника" u_s и източник на ток с "ток на източника" i_s.

В стандарта БДС EN 60375:2006 са определени графичните символи на източниците на ток (Fig. 4 и Fig. 5), които също се различават от използваното графично изображение в нашата литература (Fig. 6).

В отделните европейски държави обикновенно се е наложила една от двойките символи (Fig. 1 и Fig. 4 или Fig. 2 и Fig. 5). [20],…, [23]. Съществуващите

разлики във областта на електротехниката между отделните държави много добре се виждат, като се преглеждат WEB страниците на Уикипедия в тази област на различните езици [20],…,[23].

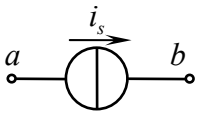


Fig.4

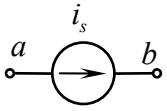


Fig.5

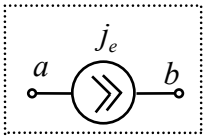


Fig.6

Авторите не са запознати с използването на определена двойка графични символи по новия стандарт в България. Като се имат предвид навиците, създадени при използваните до този момент графични символи, като по-ясно разграничаващи се са първата двойка символи. При втората двойка символи източникът на ток много прилича на източника на напрежение по стария стандарт и са възможни неволни грешки.

Дуалните вериги на източниците при постоянен ток са показани на Fig. 7 и Fig. 8.

Съответно дуалните параметри на източниците и на веригите са:

$$U_0 \Leftrightarrow I_K;\ R_S \Leftrightarrow G_S, \left(G_S = \frac{1}{R_S}\right);\ R_T \Leftrightarrow G_T, \left(G_T = \frac{1}{R_T}\right).$$

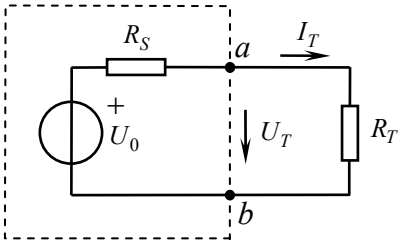


Fig.7

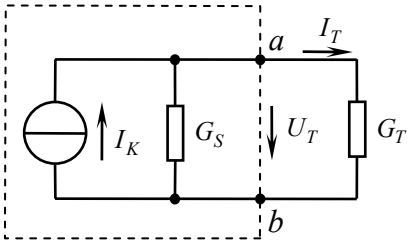


Fig.8

За управляемия източник на напрежение графичните символи са показани на Fig. 9 и Fig. 10.

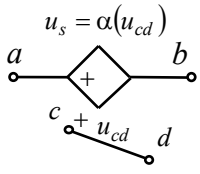


Fig.9

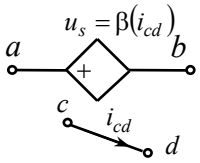


Fig.10

За управляемия източник на ток графичните символи са показани на Fig. 11 и Fig. 12.

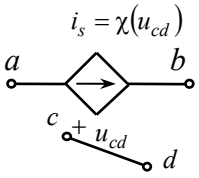


Fig.11

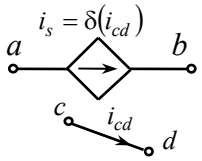


Fig.12

При определяне посоките на токовете в четириполюсниците е стандартизирана схемата на Fig.13, която се отличава от схемата, използвана в нашата учебната литература [13],…,[19] с посоката на тока i_2. Това няма значение при определянето на коефициентите на четириполюсника, но може би е от значение при други обстоятелства и след като е посочена като такава, би следвало в бъдещата литература да се използва именно тя.

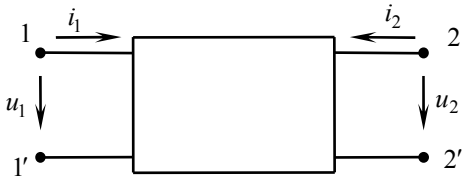


Fig.13

За взаимната индуктивност вместо символа M_{nm} е стандартизиран символа L_{nm}.

В същия стандарт законът на Кирхов за токовете е дефиниран по строго, като вместо условна положителна посока на токовете е зададена абсолютна посока, което може и да е от значение при по комплексни задачи. Преведена на български дефиницията на закона е: “Алгебричната сума от клоновите токовете към който и да е възел от една електрическа верига е нула. **БДС EN 60027-2: 2007**. (IEC60050-131:2002. / 131-15-09/). Този и всички преводи по-долу са на авторите и не са стандартизирани на български език. Цитираните по долу европейски стандарти са въведени като български, но не са преведени.

Законът на Кирхоф за напреженията се използва без да се въвеждат електродвижещите напрежения и формално разликата с досегашната литература е голяма.

Същото е положението с обобщения закон на Ом за активен участък от веригата (закон на Ом за цялата верига). За закона на Ом се приема само постоянното отношение между напрежението и тока. (Напрежението на изводите на идеален резистор е пропорционално на тока през резистора.).

Друга насока в хармонизацията на учебната литература по електротехника със европейските стандарти е дефинирането на отделните термини на български език. За сега посочените стандарти не са преведени и на практика нямаме стандартизирани термини на български език, което е много сериозен проблем при писането както на учебници, така и на друга техническа литература в областта на електротехниката. Авторите на доклада считат, че преподавателите в областта на електротехниката могат чрез участие в работата на съответните

технически комитети на “Български институт по стандартизация” да дадат своя принос за ясни преводи на стандартите.

При въвеждането на термините на български език в учебната литература може би е добре да се посочва и термина на английски език. Според авторите , в учебната литература е необходимо да се посочват стандартите, определящи стандартизираните символи, термини, определения и графични изображения.

При по подробно запознаване с новите стандарти е съвсем нормално да се открият още много изменения, но не това е целта на настоящия доклад. Основната цел е да се покаже, че за следването на бързите изменения в стандартите, следва да има и бързи изменения в учебната литература.

От друга страна възниква въпросът за действащите механизми за контрол по спазването на стандартите. Сега избора на рецензенти и научни редактори е на ниво Висше училище. На държавно ниво този проблем не се контролира пряко.

3. RESULTS

В статията е показано, че съществуват сериозни различия между нововъведените европейски стандарти и учебната литература по електротехника. Посочена е необходимостта от хармонизиране на учебния материал с новите изисквания.

4. CONCLUSIONS

С повдигането на настоящия въпрос авторите се надяват процесът на хармонизация на учебния материал по електрически вериги (основи на електротехниката) с новите европейски стандарти да бъде ускорен, което ще доведе до повишаване на качеството на учебния процес във областта на електротехниката в българските университети.

REFERENCES

[1] М. Клисаров и Г. Клисаров, Наръчник на електротехника. Държавно издателство “Техника”, София. 1969.

[2] БДС 2-701-78 (СТ СИВ 527-77). Единна система за конструкторска документация. Схеми електрически – класификация, термини и определения. (Действащ).

[3] БДС 2-733-82 (СТ СИВ 869-78). Единна система за конструкторска документация. Схеми електрически. Общи изисквания за изпълнение. (Действащ).

[4] БДС 2- 721-80 (СТ СИВ 863-78). Единна система за конструкторска документация. Означения условни графични в електрическите схеми – резистори. (Действащ).

[5] БДС 2-724-80 (СТ СИВ 864-78)). Единна система за конструкторска документация. Означения условни графични в електрическите схеми – кондензатори. (Действащ).

[6] БДС 2-727-81 (СТ СИВ 869-78) Единна система за конструкторска документация. Означения условни графични в електрическите схеми – бобини индукционни, дросели, трансформатори, трансдуктори и магнитни усилватели. (Действащ).

[7] IEC 60050-121-1998 International electrotechnical vocabulary (IEV) - Chapter 121: Electromagnetism.

[8] IEC 60050-131-1998 International electrotechnical vocabulary (IEV) - Chapter 131: Circuit theory.

[9] БДС EN 80000-6:2008. Величини и единици. Част 6: Електромагнетизъм. (Действащ).

[10] БДС EN 60027-1: 2006. Буквени символи, използвани в електротехниката. Част 1: Общи положения (IEC 60027-1:1995 (препечатан) + A1:1997). (Действащ).

[11] БДС EN 60027-2: 2007. Буквени символи, използвани в електротехниката. Част 2: Телекомуникации и електроника (IEC 60027-2:2005). (Действащ).

[12] Л. Р. Нейман и К. С. Демирчян, Теоретические основы элекгротехники. “Энергия”. 1967.

[13] А. Е. Каплянский, А. П. Лысенко и Л. С. Полотовский, Теоретические основы электротехники. Москва, “Высшая школа, 1972.

[14] С. Л. Фархи, Основи на електротехниката. Част 1. Колеж по телекомуникации и пощи. София. 2002 г.

[15] С. Л. Фархи, Основи на електротехниката. Част 2. Колеж по телекомуникации и пощи. София. 2002 г.

[16] С. Фархи и С. Папазов, Теоретична електротехника. Част 1. Четвърто издание.

[17] К. Брандиски, Ж. Георгиев, В. Младенов и Р. Станчева. Теоретична електротехника. Част 1. ИК КИНГ, 2004.

[18] К. Брандиски, Ж. Георгиев, В. Младенов и Р. Станчева. Теоретична електротехника. Част 2. ИК КИНГ, 2005.

[19] http://de.wikipedia.org/wiki/Stromquelle

[20] http://en.wikipedia.org/wiki/Current_source

[21] http://fr.wikipedia.org/wiki/Source_de_courant

[22] http://ru.wikipedia.org/wiki/%D0%98%D1%81%D1%82%D0%BE%D1%87%D0%BD%D0%B8%D0%BA_%D1%82%D0%BE%D0%BA%D0%B0

SOFTWARE TOOL FOR SHORT TERM CONGESTION FORECASTING IN TRANSMISSION NETWORKS

Nenad Sijakovic, Miomir Kostic, Valeri Mladenov, Irina Bogatinova

Transmission System Planning and Analysis Department, Elektromreza Srbije, Vojvode Stepe 412, 11000 Belgrade, Serbia, +38111 3091113, nenad.sijakovic@ems.rs

Abstract: *Day Ahead Congestion Forecast (DACF) is a regular procedure for power flow forecasts in European interconnected network. It is performed on daily basis (today-for-tomorrow), whereby European TSOs participate in common activities such as producing the forecast network models of transmission network for day ahead, necessary for load-flow and contingency analyses for 24 hours of the next day. Exchanging those forecast models with all TSOs in Europe, collecting all the models of regional partners and merging them with their own model, a regional network model is made. AC Load Flow and Contingency Analyses are running on the regional (optionally on the whole merged Pan - European) interconnected transmission model. The goal is to prevent appearance of congestions in Pan European grid. Software tool for short term congestion forecasting, named Transmission System Planning Operational Tool, which is still under development, is described in the paper.*

Keywords: DACF – Day Ahead Congestion Forecast, congestions, contingency analysis, transmission network

1. INTRODUCTION

One of the main targets of the liberalization of the electricity supply sector in the European Union is the creation of a truly Internal Electricity Market (IEM). By introducing competition among generators and suppliers not only in their domestic markets, but also on an international scale, the economic efficiency of electricity supply shall be maximized for the benefit of the consumers end the entire economy. However, the transmission system operators (TSOs) have not historically designed their networks and interconnections between them with the primary objective of facilitating internal power trade. As a consequence, the integration of the national electricity markets is impeded by the limited amount of cross border transmission capacity at several borders.

To mitigate this problem, the EU supports on one hand measures to increase the transmission capacities, either by investment in new network facilities or by optimization and harmonization of operating standards that allow to better utilize the networks. On the other hand, the rules applied for managing the utilization of the existing transmission capacities, which can be summarized by the term “congestion management methods”, are of the particular importance for the efficiency of the IEM, and secure overall system operation in short and medium terms.

1.1. Congestion management

The main target of the Congestion management procedure is to recognize and prevent appearance of dangerous regimes in power system operation through process of operational planning and later on by system control measures in real time. Congestion management procedure can be divided into four separated steps:

Planning and coordination of the maintenance of the crucial network elements in the region,

Calculation and allocation of cross border transmission capacities on yearly, monthly, weekly, daily and intra daily levels,

DACF (Day Ahead Congestion Forecast) - day ahead contingency analysis in which every day TSOs are checking system operation security through “n-1” criteria on the merged regional network model, and

Real time dispatching actions.

1.2. DACF procedure

The third step of congestion management is, as mentioned above, based on DACF.

Each day every TSO is obliged to generate mathematical network models of its network (400 kV and 220 kV levels are modeled) for tomorrow’s characteristic hours. Those models in UCTE format are then being exchanged between all partners in the interconnection after which each TSO can perform security checks through contingency analysis on the merged regional network models for desired hours.

If the security criteria are not satisfied (in case of a problem concerning overload, voltage stability or any other issue regarding overall system operation security), TSO have the right to change cross border power exchange schedule or internal production schedule for that particular period of the day.

Software tool for short term congestion forecasting, named Transmission System Planning Operational Tool, or TSPOT, which is under development at the moment in the framework of FP7 project named SEETSOC is planned to deal with this issue.

2. TRANSMISSION SYSTEM PLANNING OPERATIONAL TOOL

2.1. TSPOT software features

Windows desktop application developed in Visual Studio under .net framework, data base based (DB on local server for local usage only). Application organization in the form of master form / child forms organization.

Master form: one, unique form, starting point for all applications

Child forms:

- cases (opens in new window, every case consists of four windows: Work Case Window, Notification/Results Window, Solution Explorer Window and Property & Settings Window)

- common Applications (opens in new window: Contingency Statistics Analyzer with Report Manager, Outage Scheduler, Format Converter, Model Compare Wizard, User Accounts & Logs Manager)
- MAST (many at the same time) applications (opens in the Case Window: Contingency Analyzer with Report Manager, Transmission Capacity Manager with Report Manager, Model Creator Wizard, Model Merger and Checker, Grid Data, Graphical System Presentation, Study Case with Report Manager, Auto Scheduler, Information Exchange)

2.2. TSPOT structure (modules)

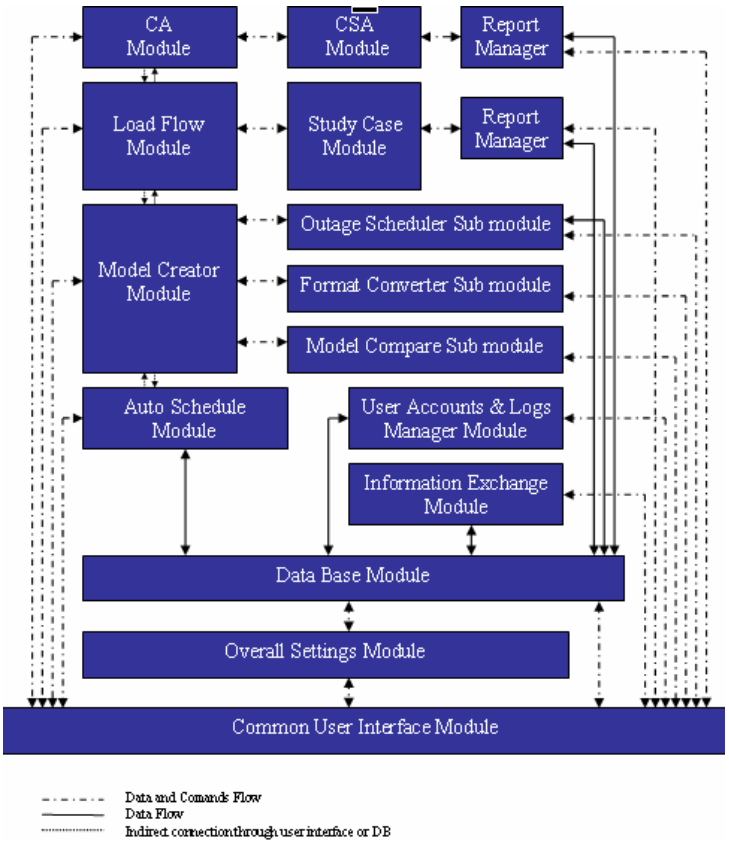


Image 1 - Transmission System Planning Operational Tool - Modules

1. Data Base Module
2. User Accounts & Logs Manager
3. Auto Schedule Module
4. Model Creator Module/Global functions/ Load Adjustment Sub module/Model Compare Sub module/Format Converter Sub module
5. Outage Scheduler Sub module
6. Load Flow Module
7. Contingency Analyzer Module
8. Contingency Statistics Analyzer Module
9. Report manager for Statistics Analyzer Module

10. Study Case Sub module
11. Report manager for Study Case Module
12. Information Exchange Module
13. Overall Settings Module

2.3. TSPOT functions

Network Models Managing

- usage of UCTE and new CIM XML data format for network modelling
- data base based network modelling
- function of format conversion from/into UCTE and new CIM XML data format
- function of model builder from snapshot reference model
- creation of models storage
- function of automated DACF model creator
- function of model merger
- function of model checker
- modular approach for inputs that will allow usage in different TSOs with different input data sources
- modular structure that can be easily connected to other relevant modules and sub modules of Congestion Forecasting Tool

Power Flow Calculation

- data base based AC and DC load flow machines
- possibility to work with common regional network model (SEE regional model consists of the following models: Italia, Austria, Slovenia, Croatia, Bosnia and Herzegovina, Hungary, Montenegro, Bulgaria, Romania, Serbia, FYROM, Greece, Albania, Ukraine West and Slovakia). All transmission network voltage levels starting with 110 kV should be taken into consideration.
- Load Flow settings, defined in DB, with huge spectra of freedom for users (for example, concerning AC load flow: tolerance, conversion from PV to PQ, number of iterations, flat start, possibility to select slack node, multiple slack node selection…
- Graphical System Presentation in the form of tree view (under Common User Interface Sub module

Contingency analysis (checking of n-1 criteria)

- primary function is to perform contingency analysis (n-1 criteria checking) on the existing network model. The ides is to make automatic machine that will change topology of the considered network for each case defined in Outage list. After topology modification each time full load flow calculation will take place in order to monitor load flows on the elements defined in Monitoring List and voltages in all nodes. If CA machine find overload on some of the moni-

tored elements or over/under voltage in any node, this will be recorded and put in the report. After calculation is done by CA machine, load flow module report is available through user interface (data grid view). User can insert that report into DB manually or that can be set as automatic function. CSA module uses this reports as a base for its statistical calculations.

- inputs:
  – List of contingencies (user defined selection of n-1 and n-2 contingencies)
  – List of outages
  – List of monitoring elements
- software should be able to import existing lists from txt file format, or to generate new lists.
- outputs: Contingency analysis (n-1 criteria checking) results which will be shown in CA window (after the calculation through user Interface) and after that optionally saved into DB in Contingency Statistics Analyzer Module & Report Manager Data Table

Contingency Statistics Analyzer

- editor (user interface) with tools for statistical analysis of "n-1" criteria checking results (CA outputs) with output in the form of data grids and diagrams. (Report manager for Contingency Statistics Analyzer)
  – results of "n-1" criteria checking performed by CA module presented in data grid view (table) with advanced filtering and export functions:
    - filter by contingency, outage or overload
    - filter by type of the outage or overload (is it connected to the domestic or foreign element?)
    - filter by date and hour of the concerned contingency
    - filter by overload level
    - presentation of the results of "n-1" criteria checking performed by CA module connected with network topology at that specific model
    - presentation of the results of "n-1" criteria checking performed by CA module connected with generation engagement at that specific model
    - export of the overall list or filtered list in the form of csv, txt, doc, xls or pdf file format
  – tabular presentation of the contingency appearance frequency in the selected time period:
    - filter by definable time period and selected hours
    - possibility to work with full or filtered contingency list
    - export of the table in the form of csv, txt, doc, xls or pdf file format
  – tabular presentation of the outage appearance frequency in the selected time period:
    - filter by definable time period and selected hours
    - possibility to work with full or filtered contingency list
    - export of the table in the form of csv, txt, doc, xls or pdf file format

– tabular presentation of the overload appearance frequency in the selected time period:
  - filter by definable time period and selected hours
  - possibility to work with full or filtered contingency list
  - export of the table in the form of csv, txt, doc, xls or pdf file format
– graphical presentation (diagrams) of the statistical analysis results with the folowing features:

- possibility to use three different types of diagrams (each type of diagrams contains three sub types)
  1. a) diagram that shows appearance of the contingencies during a year,
     b) diagram that shows appearance of the outages during a year,
     c) diagram that shows appearance of the overload during a year,
  2. a) diagram that shows contingency appearance frequency in the selected time period and selected hours,
     b) diagram that shows outage appearance frequency in the selected time period and selected hours
     c) diagram that shows overload appearance frequency in the selected time period and selected hours.
  3. diagram that shows overload levels of the contingencies in the selected time period and selected hours with filter by outages and filter by overloads.

3. CONCLUSION

TSPOT authors and developers hope that this software will be, together with other products of SEETSOC project which will be developed under FP7 framework, of great help for regional TSOs in the South-East European (SEE) region in the near future.

4. ACKNOWLEDGEMENT

The work was supported by FP7 project South–East European TSO Challenges (SEETSOC).

REFERENCES

[1] “Load Flow calculation and contingency analysis using PSA software tool” N Sijakovic, EMS, Belgrade 2006.
[2] “Calculation of cross border transmission capacities on tie lines of Serbia“ A. Kurcubic, B. Sumonja, N. Sijakovic, J. Petrovic, JUKO CIGRE Tara, 2006.
[3] “Possibility of Coordinated Auction appliance in Serbian network and influence of HPP Djerdap 1“ B. Sumonja, A. Kurcubic, N. Sijakovic, J. Petrovic, JUKO CIGRE Tara, 2006.
[4] “Critical outages of Transmission system elements in Serbian and neighbouring networks concerning security and voltage stability in Serbian network” N Sijakovic, J Vicovac, M Stancevic, A Kurcubic, JUKO CIGRE Vrnjacka Banja 2007.

APPLICATION OF MONITORING SYSTEM FOR TRANSFORMER SUBSTATIONS

Nikolina Petkova[1], Valeri Mladenov[1], Petar Nakov[2]

[1]Department of Theoretical Electrical Engineering, Technical University of Sofia,
8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria, phones: +3599652386, +3599652498,
e-mails: valerim@tu-sofia.bg, npetkova@tu-sofia.bg

[2]Department of Theoretical Electrical Engineering, Technical University of Sofia,
8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria, phone: +3599652107, e-mail: pnakov@tu-sofia.bg

Abstract: *This paper presents a monitoring system for transformers like an effective technique which preventive detection of possible failures, contributes to ensure continuity and reliability of the operation of such equipment, reducing costs and economic losses associated with them. The system consist of a set processes which correspond to filed measurement of operation and magnitude parameters, typical of a specific transformer. Online monitoring can warning and alarm in real time; it is possible to measured and registered data of loading; reactive power; earth current and power quality at the transformer substation.*

Keywords: Monitoring system, Transformers, Power distribution control

1. INTRODUCTION

The transformer is a critical item of equipment in power systems and its correct functioning is essential to the reliable operating of the system. Generally, the trend of transformer monitoring systems (TMS) is from data acquisition to data interpretation to give clear information to the operator. In this way the distribution network will be controlled automatically.

The reliability of operation of distribution networks can be increased by using automatic monitoring systems for transformers – not only for power transformers but also for distribution transformers. Remote monitoring can provide selective sharing of data among multiple sites in the most efficient and cost effective manner. An information centre of a utility acquires information on power plants and on HV/MV substations from a supervisory control and data acquisition (SCADA) system, and information on LV energy consumption of end users from automatic metering recorders. However, on-line data on the conditions of distribution transformer substations are currently not often available for remote diagnosis. For example, the information on the loading of distribution transformer substations is not in real time, but is limited and based only on technical checks performed few times a year by visiting maintenance electricians. Not only the conventional technical data, such as current, voltage, temperature etc., but also more advanced information about the status of transformers, such as expected transformer lifetime is needed by the operators to ensure reliable power delivery. Utilities find the implementation of communication between numerous monitoring systems too expensive and thus they have been ignored en-

tirely. However, the development of the infrastructure of wireless communication such as the mobile phone networks, gives new, cost effective possibilities to monitor distribution transformer substations. [1]

2. SUPERVISORY CONTROL AND DATA ACQUISITION (SCADA)

This kind of system is used to remotely monitor network infrastructure. A SCADA system allows users to monitor an entire plant or individual pieces of equipment and processes by collecting real-time data from various sensors throughout a network. It is very important for organizations to monitor these network activities, as doing so informs them of problems with their mission-critical processes. By monitoring these processes, organizations can quickly respond when there are problems within their network. A SCADA monitoring system contains several components that, when combined, allow operators to monitor their network. A SCADA system designed to monitor a network includes alarm sensors, Remote Terminal Units (RTUs), SCADA master units, and the communication network.

These parts interact to allow the operator to monitor the system. By deploying advanced RTUs, operators can more effectively monitor mission-critical processes through both discrete and analog alarms. "Derived Controls" can also be programmed into these advanced RTUs, providing for instant reactions of the monitoring system in the event that user-specified alarm combinations occur.

3. MONITORING OF POWER TRANSFORMERS

A multitude of different measurable variables can be collected for on-line monitoring of power transformers. However, it is very rarely useful to use the entire spectrum. Therefore, sensor technology must be adjusted to the specific requirements of a particular transformer or transformer bank, depending on their age and condition. From the experience of more than 120 monitoring systems the following general set-up of sensors for example is proposed for the use at a 400 kV power transformer:

- PT100 for measurement of top oil temperature;
- PT100 for measurement of ambient temperature;
- C.T. for measurement of load current (single phase);
- Measurement of voltage at measurement tap of bushing (three phase);
- Measurement of oil pressure of bushing;
- Sensor for measurement of oil humidity;
- Sensor for measurement of gas-in-oil content;
- Tap changer position;
- Power consumption of motor drive ;
- Digital inputs for switching status of fans and pumps.

For early error detection, the monitoring of the active part is of particular importance. It is fundamental to either measure the electrical variables load current and operating voltage directly at the transformer. Bushing-type current transformers are used for load current measurement. The load current is an important starting variable

for calculation of hot-spot temperature according to IEC 60354 [2,3] and thus presents ageing of active part insulation. This enables the evaluation not only of information regarding the temporary overload capacity of the transformer but also of the lifetime consumption. The voltage applied to the transformer is acquired at the measurement tap of the capacitor bushing by means of a voltage sensor. By this way a change of the capacitance of the bushing which is the start of deterioration of the bushing can be detected. Overvoltages represent an essential risk potential for the insulation of transformer windings. In connection with the volume of noxious gases, which are dissolved in oil, can draw deductions regarding possible damage to the insulation of the active part after the occurrence of voltage surges. For the gas-in-oil detection a Hydran sensor is used which reads a composite value of gases in ppm (H2 (100%), CO (18%), C2H2 (8%), C2H4 (1,5%)). As hydrogen is a key gas for problems in the active part, an increase in the output signal of the sensor is an indication for irregularities such as for example partial discharge or thermal overload. The evaluation of this measuring signal, together with the dependency on the temperature of the oil and the load current, provides a reliable basis for the continued operation of the transformer. In the event of an increase of gas-in-oil content, an immediate reaction can be effected via an off-line gas analysis to determine the concentration of the other components dissolved in the oil in order to limit the cause of the damage.

A capacitive thin film sensor is used for the detection of moisture in oil. There are several causes for an increase of water-in-oil content. Due to the fact that water is a result and also an origin of oil/paper insulation degradation the water in oil content is an important indicator for the condition of winding insulation, in particular for already aged transformers. In the cooling unit, not only the switching state of the oil pumps and the ventilators is monitored but also the temperature values. For this purpose, the input and output oil PT100 probes measure temperatures. The intention is to make selective statements regarding the state of the entire cooling plant from the measured values. Recordings of the tap changer position and the operating current help determine the number of switching operations of the tap changer and the total switched current, which gives information abut the burning of diverter switch contacts. If the limiting value, pre-set in accordance with the maintenance instructions, is exceeded a message is generated.

Due to the fact that serious damage to the transformer can be expected in the event of the failure of the tap changer, the monitoring of this mechanically and electrically highly stressed element is of great importance. In order to be able to obtain information regarding the mechanical state of the switch the power consumption of the tap changer drive mechanism is recorded.

The outputs of the above mentioned sensors are wired onto field bus terminals in the monitoring module installed at the transformer. Within these data acquisition units the analogue signals are digitised and send via a field bus to the monitoring server. By means of this industrial proven technology it is possible to monitor all transformers in one substation with a single system which is extremely cost effective. The erection of the server in an operating building offers the advantage that the ambi-

ent conditions (e.g. temperature, vibrations) are much more suitable for a PC. The connection to the protection and control system can be done either by dry relay contacts or a digital protocol according to IEC 60870-5-101.

During the last years the AREVA monitoring system MS 2000 was installed world-wide at power transformers of all major manufacturers. In the beginning utilities started to test the system with grid-coupling transformers of minor importance. Due to the good experience it is now operating at such strategic important points as nuclear power stations, pumped storage power stations and coal power stations. Most of these installations were retrofitted on-site at already aged transformers. Normally the installation of sensors requires no welding at the transformer and takes about two days. The transformer has to be taken out of operation only for half a day to install the voltage sensors and the tap changer monitoring module. An additional module installed on the MS 2000 monitoring server generates HTML-based web pages, which show the online and historical data, as in Fig. 1. [4]

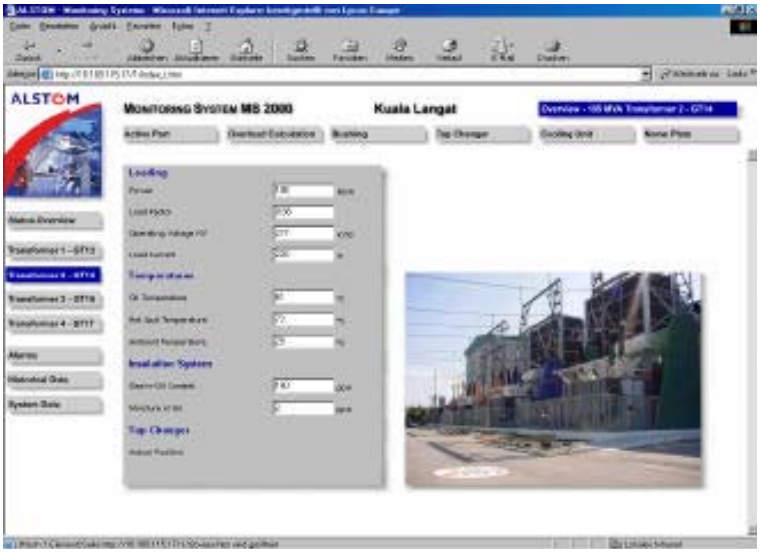


Fig. 1. Web based-based visualisation of transformer condition

4. THE WIMO SYSTEM

Wimo system is developed by Wimotec Ltd. [5] and it consists of a PC with proper software at the control centre, transformer monitoring units (TMU, i.e. the Wimo controllers) with a GSM module at the transformer substations and radio communication (GSM and sometimes also GPRS) between them, as shown in Fig. 2. A reliable server is also needed for data backup. The potential free inputs can be used to detect short circuits, smoke and humidity, to monitor transformer cabinet and for a temperature relay. The system is capable of communicating in both directions. The Wimo system not only sends messages (measuring data, warnings and alarms) from a distribution transformer substation to an operator via a mobile phone or a server, but also receives the instructions such as parameter settings of the operator via a GSM or GPRS network and Internet, as shown in Fig. 2.

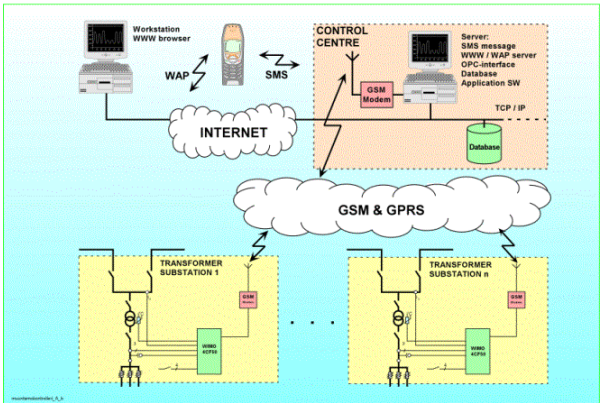


Fig. 2. The Wimo monitoring system for distribution transformers

5. CONCLUSION

Web-based technology have given monitoring systems a significant boost because they greatly increased access to information on operating condition of primary equipment. Furthers steps to enhance their potential to drastically shorten response times and larger diffusion of information are under way. Such tools are now able to provide utilities with a global and real time knowledge of their whole transmission system, complete expert system to diagnose and prevent failures, a global installed pool of all kind of equipment installed all over the country and direct links to manufacturers spare parts resources and maintenance teams. Besides these tools integrate all the information into a single system with the same interface for all users. They ideally support the drive for efficiency of utilities that need to increase the performances and availability of power grids.[4]

6. ACKNOWLEDGEMENT

The work was supported by FP7 project South–East European TSO Challenges (SEETSOC), R&DS of the Technical University - Sofia contract № 4149-M.

REFERENCES

[1] T. T. Vekara, S. Pettissalo, N. Rajkumar, “Remote Monitoring System for Transformer Substations”
[2] J.-.P. Dupraz, T..Jung&coll, “Digital Control and condition Monitoring: integration and application”, Cigré session 1998, paper 13-109
[3] S. Tenbohlen, G. Puldo, M. Lindners, G. Krost “Integration of Power Transformer monitoring and Overload Calculation into the Power System Control Surface“, CEPSI 2002, FUKUOKA
[4] T. Jung, S. Tenbohlen, J. Altwegg, P. Roussel, C. Harfouch, “Implementation of new monitoring tools and optimization of maintenance through the use of Web-based technology”, IEEE, 2004
[5] Wimotec Ltd., http://www.wimotec.com.
[6] S. Pettissalo, W. E. Kozlowski, H. Lyskawa, T. Kantecki, "WIMO – Remote Monitoring System for Transformer Substations," VI National Science and Technology Conference and Diagnostics of Industrial Processes PDD’03, Wladyslawowo, Poland, 15-17.9.2003.

ON THE DESCRIPTION OF THE SOLITARY WAVES IN THE PRESENCE OF DISTRIBUTED BANDWIDTH LIMITED OPTICAL AMPLIFICATION AND INTRAPULSE RAMAN SCATTERING

Ivan M. Uzunov

Department of Applied Physics, Technical University Sofia
8 Kl. Ohridski Blvd., Sofia 1000, Bulgaria, e-mail: ivan_uzunov@tu-sofia.bg

***Abstract**: A theoretical study of the possibility of existence of solitary waves in the presence of distributed bandwidth limited optical amplification and the intrapulse Raman scattering is proposed. Using the soliton perturbation theory, the stability of the fixed point for the soliton amplitude and velocity is analyzed. A relation that couples the soliton amplitude and velocity has been numerically verified. Next, a new perturbation approach is suggested that leads to the research of the equation of strongly nonlinear Duffing - Van der Pol oscillator. This equation has been examined by the recently developed hyperbolic perturbation methods of [6] and [7]. It turned out that the coupling of the equilibrium velocity and amplitude of stationary solution in both hyperbolic perturbation methods coincides. Even more, it goes out that this coupling coincides with the relation between the soliton amplitude and velocity derived by the adiabatic approximation of soliton perturbation theory.*

Keywords: (optical solitons, intrapulse Raman scattering, distributed bandwidth limited amplification, Duffing-Van der Pol oscillator)

1. INTRODUCTION

The intrapulse Raman scattering (IRS) plays important role among the higher-order nonlinear effects that occur in fibers. When the pulse spectrum becomes very broad its high-frequency components can transfer energy to its low-frequency components [1], which results in a continuous downshift of the soliton carrier frequency, phenomena known as the soliton self-frequency shift (SSFS) [1-3]. Erbium-doped fiber amplifiers (EDFAs) can be used to amplify optical pulses. When relatively short (about 1ps) pulses have to be amplified, IRS also should be taken into account. It was established [1-3], that the BLA can reduce the amount of spectral shift due to the SSFS and stabilize the soliton carrier frequency close to the gain peak.

The main analytical tool in this analysis of IRS and BLA is the soliton perturbation theory (SPT) [1-3]. Description, including changes in the shape of bright solitons in the presence of SSFS, has been suggested [4]. The adiabatic approximation of SPT has been employed in the study of the IRS and BLA and the fixed point for the soliton amplitude and velocity has been identified in [5].

In this paper, I intend to extend the results of the adiabatic approximation of SPT for the combined action of IRS and BLA [5], analyzing the stability of the fixed point for the soliton amplitude and velocity. Next, I suggest new perturbation approach,

which key idea is the usage of the equation of strongly nonlinear Duffing - Van der Pol oscillator. This equation has been studied by the hyperbolic perturbation method of [6], and the hyperbolic Lindstedt-Poincare perturbation method of [18]. Results of [7] are used directly, while the method [18] is applied here to the equation of strongly nonlinear Duffing - Van der Pol oscillator. The equilibrium velocity of perturbed homoclinic solution as a critical value of control parameter in hyperbolic perturbations methods will be evaluated and results compared.

BASIC EQUATION

Modified NLSE that describes pulse propagation in optical amplifiers is given by [1-2]:

$$i\frac{\partial U}{\partial x}+\frac{1}{2}\frac{\partial^2 U}{\partial t^2}+\left|U\right|^2 U+\gamma\frac{\partial}{\partial t}\left(\left|U\right|^2\right)U=i\delta U+i\beta\frac{\partial^2 U}{\partial t^2}. \qquad (1)$$

where the dimensionless variables (soliton units) are introduced as follows [1-2]:

$$x=z/L_D\,,t=T/T_0\,,U=\left(\gamma L_D\right)^{1/2}A.$$

Here, z and t are real longitudinal coordinate in the fiber, and time, $T=t-z/v_g=t-\beta_1 z$, v_g is the group velocity, $A(z,T)$ is the slowly varying envelope, $L_D=T_0^2\big/\left|\beta_2\right|$ is the dispersion length, T_0 is the width of the pulse, β_2 represents the dispersion of the group velocity. $\gamma=\dfrac{n_2\omega_0}{cA_{eff}}$ is the nonlinear parameter, n_2 is the nonlinear-index coefficient, ω_0 is the carrier wavelength of the optical pulse, A_{eff} is the effective core area, c is the speed of light in vacuum. The last term on the left-hand side of Eq. (1) is proportional to $\gamma=T_R/T_0$, where T_R is the first moment of the nonlinear response function. This term describes the IRS, and consequently is responsible for the SSFS. Next,

$$\delta=\left(g_0-\alpha\right)L_D/2\,,\beta=g_0L_D\left(T_2/T_0\right)^2/2$$

where g_0 is the gain, α is the fiber losses, and T_2 is the dipole relaxation time. The term proportional to β, represents the finite bandwidth of the fiber amplifier.

2. ADIABATIC PERTURBATION METHOD

For small perturbation, in the adiabatic approximation of SPT, the soliton solution may be written as follows [3]:

$$U(x,t)=\eta(x)\operatorname{sech}\left\{\eta(x)\left[t-\tau(x)\right]\right\}\exp\left[i\left(-k(x)t+\sigma(x)\right)\right] \tag{2}$$

where $\eta(x)$ and $k(x)$ are the soliton amplitude and velocity (frequency), correspondingly. The soliton position $\tau(x)$ and phase $\sigma(x)$ are defined by the equations: $d\tau(x)/dx=-k$, and $d\sigma(x)/dx=(\eta^2-k^2)/2$, respectively. Applying adiabatic perturbation method, the following system for the soliton's amplitude and velocity was obtained [3]:

$$\begin{aligned}\frac{d}{dx}\eta&=2\delta\eta-2\beta\left(k^2+\frac{1}{3}\eta^2\right)\eta\\ \frac{d}{dx}k&=-\frac{4}{3}\beta k\eta^2+\frac{8}{15}\gamma\eta^4\end{aligned} \tag{3}$$

In the case of IRS it was obtained [3]: $dk/dx=+8\gamma\eta^4/15$ i.e. the amplitude does not change, but the velocity changes with distance. When both BLA and IRS are present, following fixed point with positive amplitude [5]:

$$\eta_*=\sqrt{\frac{\left(5\sqrt{25\beta^4+144\delta\beta\gamma^2}-25\beta^2\right)}{24\gamma^2}}\ ,\ k_*=\frac{\left(\sqrt{25\beta^4+144\delta\beta\gamma^2}-5\beta^2\right)}{12\beta\gamma} \tag{4}$$

In addition to [5], I present here the eigenvalues $\lambda_{1,2}$ of linearized problem in the vicinity of the singular point given by Eq. (4). The eigenvalues $\lambda_{1,2}$ are solutions of the following quadratic equation in λ:

$\lambda^2+p\lambda+q=0$ with $p=\dfrac{5\beta}{9\gamma^2}\left(-5\beta^2+\sqrt{25\beta^4+144\delta\beta\gamma^2}\right)$,

$q=\dfrac{5\beta}{162\gamma^4}\left(-125\beta^5-720\gamma^2\beta^2\delta+\beta\left(25\beta^3+72\gamma^2\delta\right)\sqrt{\Lambda}\right)$, and $\Lambda=25\beta^4+144\beta\gamma^2\delta$.

Recalling the assumption regarding β,δ,γ, we see that p will be always positive $p>0$. The eigenvalues $\lambda_{1,2}$ are given by:

$$\begin{aligned}\lambda_{1,2}&=-\frac{p}{2}\pm\frac{1}{2}\sqrt{p^2-4q}=\\ &=\frac{1}{18\gamma^4}\left(5\beta\gamma^2\left(5\beta^2-\sqrt{\Lambda}\right)\pm2\sqrt{5}\sqrt{\beta\gamma^4\left(125\beta^5+540\beta^2\gamma^2\delta-\left(25\beta^3+36\gamma^2\delta\right)\sqrt{\Delta}\right)}\right).\end{aligned}$$

The case $q=0$ corresponds to higher order singular points and it requires $\beta=0$ or/and $\delta=0$, a situation which will be not discussed here. Let us first consider $q>0$ ($p>0$). For $p^2>4q$, the roots λ_1 and λ_2 are real and of the same sign, the sta-

ble nodal points appear for $p > 0$. For $p^2 = 4q$, the roots $\lambda_1 = \lambda_2 = -p/2$ are degenerate and the nodal point occurs. This situation appears in the case of the following system parameters: $\delta = 0{,}15; \beta = 0{,}45; \gamma = 5.10^{-4}$. Then $q = 0.36, p = 1.2 \Rightarrow p^2 = 4q = 1.44$, and degenerate roots are $\lambda_1 = \lambda_2 = -0.6$. For $p^2 < 4q$, the roots λ_1 and λ_2 are complex conjugate, the stable focal points appear for $p > 0$. Such an example case is illustrated for the following parameters: $\delta = 0{,}5; \beta = 0{,}01; \gamma = 5.10^{-4}$, where $q = 3.93, p = 3.46$, or $p^2 = 12 < 4q = 15.7$. The complex conjugate eigenvalues are $\lambda_{1,2} = -1.73 \mp 0.97i$, and they define the stable focal point. Using the obtained singular point given by Eq. (4), the following relation between the velocity k_* and the square of amplitude η_*^2 is obtained:

$$k_* = 2\gamma\eta_*^2 / (5\beta). \tag{5}$$

In order to confirm numerically the appearance of the singular points given by Eq. (4), as well as the relation Eq.(5), I solved numerically the system of Eq.(3) for the case $\delta = 0{,}5; \beta = 0{,}01; \gamma = 5.10^{-4}$, with the initial value for η: $\eta(0) = 0.1$, and initial values for k varying from -6 to 6 in steps of 2.5. The stable focal point occurs with the values of the amplitude $\eta_* = 11.39$ and the frequency $k_* = 2.59$, which satisfy Eq.(5).

3. NEW PERTURBATION APPROACH

An alternative perturbation approach to the adiabatic perturbation of SPT is proposed here, that comprises of two steps. First, analyzing the stationary solution of the basic equation, the equation of strongly nonlinear Duffing – Van der Pol oscillator is introduced. Next, methods of [6-7] are applied to the Eq. (11). In order to do this I will look for the stationary pulse solution of Eq. (1) into the form:

$$U(x,t) = u(\xi)\exp\left(i\left(f(\xi) + Kx\right)\right), \tag{6}$$

where $\xi = t - Mx$, and M and K are real numbers. M has a meaning of the unknown equilibrium velocity (more precisely the inverse velocity). Inserting Eq. (6) into Eq.(1), the following nonlinear system of ordinary differential equations for the functions $u(\xi)$ and $f(\xi)$ is obtained:

$$\begin{aligned} &-\delta u - \beta u'' + u f''/2 + f' u' - M u' + \beta u (f')^2 = 0 \\ &u''/2 + \beta u f'' + 2\beta u' f' + 2\gamma u^2 u' - u (f')^2/2 + Mu f' - Ku + u^3 = 0 \end{aligned} \tag{7a-7b}$$

We should mention here that the Eq.(7a-7b) in the absence of distributed optical amplification transform to Eq. (2.8a-2.8b) of [8]. Let us assume that due to the small-

ness of the parameters describing distributed BLA, δ and β, Eq. 7(a) can be approximated in the following way:

$$u\, f'' + 2\, f' u' - 2M\, u' = 0 \tag{8}$$

The applicability of this important approximation will be eventually justified by the validity of final results. Solution of Eq. (8) can be written as:

$$f = -M\xi + 2S_0 \int d\xi / u^2 \tag{9}$$

The inverse velocity M in the Eq. (9) and Eq. (6) corresponds to the soliton velocity k in the the SPT. Using Eq. (8), Eq. 7(b) can be transformed to:

$$u'' - \left(3M^2 + 2K\right)u + 2u^3 + \frac{8MS_0}{u} - \frac{4S_0^2}{u^3} + 4\left(\beta M + \gamma u^2\right)u' = 0 \tag{10}$$

In this equation, β and γ are related to distributed BLA, and γ connected to IRS. We consider the BLA as the primary perturbation, and IRS is considerably weaker than BLA. We expect that the inverse velocity M that is due to the effect of IRS will have a certain smallness $M \sim \gamma \sim \varepsilon$. So the term proportional to M^2 can be neglected, while the term proportional to βM will be kept in Eq. (10). I will assume here that $S_0 = 0$, therefore the phase function will be proportional to ξ with coefficient of proportionality determined by the inverse equilibrium velocity M. Using the small parameter $\varepsilon \sim \gamma_3$, Eq. (10) can be cast into the form:

$$u'' + c_1 u + c_3 u^3 = \varepsilon\left(\mu - \mu_1 u^2\right)u' = \varepsilon f\left(\mu, u, u'\right), \tag{11}$$

where: $c_1 = -2K$, $c_3 = 2$, $\mu = -4\beta M/\gamma$ and $\mu_1 = 4$. The coefficients c_1, c_3 on the left-hand side of Eq. (11) are not small. Neglecting the right-hand side of Eq. (11), we get the Duffing equation

$$u'' + c_1 u + c_3 u^3 = 0. \tag{12}$$

It will be called the generating equation. The solution of Eq. (12) (for $c_1 < 0$ and $c_3 > 0$) is:

$$u_0 = a_0 \operatorname{sech} \tau,\ \tau = \omega_0 \xi \tag{13}$$

where the amplitude a_0 and the frequency ω_0 are given by: $a_0^2 = -2c_1/c_3 = 2K$, and $\omega_0^2 = -c_1 = 2K\left(K > 0\right)$, respectively. The term on the right-hand side of Eq. (11)

can be considered as a Van der Pol perturbation to generating equation. Eq. (11) will be called here the equation of strongly nonlinear Duffing – Van der Pol oscillator. In the hyperbolic methods employed below, the parameter μ will be considered as control parameter, but as β/γ is given, our control parameter will be the inverse velocity M.

In accordance with [6], solution of Eq. (11) can be written as

$$u = \sum_{n=0}^{\infty} \varepsilon^n u_n = \sum_{n=0}^{\infty} \varepsilon^n a_n h(\tau) = \left(a_0 + \varepsilon a_1 + \varepsilon^2 a_2 + \ldots\right) h(\tau) \tag{14a}$$

where: $$d\tau/d\xi = \omega(\tau) = \sum_{n=0}^{\infty} \varepsilon^n \omega_n(\tau) = \omega_0 + \varepsilon\omega_1 + \varepsilon^2\omega_2 + \ldots . \tag{}$$

At each order the approximate solutions $u_n(\tau) = a_n h(\tau) = a_n \operatorname{sech}(\tau)$ have the shape of the solution of the generating equation, $\operatorname{sech}(\tau)$. The amplitudes a_n and τ depend on the small parameter ε. The control parameter μ is also expanded in the power of ε [6]:

$$\mu = \sum_{n=0}^{\infty} \varepsilon^n \mu_{Cn} = \mu_{C0} + \varepsilon\mu_{C1} + \varepsilon^2\mu_{C2} + \ldots . \tag{15}$$

We should identify the critical control parameter μ_{C0} under which a homoclinic solution forms. It has been found [6], that the value of μ_{C0} is given by:

$$\mu_{C0} = 2\mu_1 a_0^2/5 \tag{16}$$

Using Eq. (16), and the definition of $\mu_{C0} = -4\beta M/\gamma$, we obtain that

$$M_{C0} = -2\gamma a_0^2/(5\beta) . \tag{17}$$

Eq. (17) that relates the inverse velocity M_{C0} with amplitude a_0^2 of the stationary solution, compare well with the coupling between the soliton amplitude η_* and velocity k_* given by the Eq. (5). Using from [6], that: $\omega_1(\tau) = A_2 \tanh\tau = (-8K/5)\tanh\tau$, the perturbed solution of the Eq. (11) is given by:

$$\begin{aligned} u &= a_0 \operatorname{sech}\tau + O\left(\varepsilon^2\right) \\ u' &= -a_0\left[\omega_0 - (8\gamma K/5)\tanh\tau\right]\operatorname{sech}\tau\tanh\tau + O\left(\varepsilon^2\right) \end{aligned} \tag{18a - 18b}$$

where: a_0^2 and ω_0^2 are defined above. The correction to the frequency of the perturbed solution that describes the influence of the IRS on the stationary solution is: $\varepsilon\omega_1(\tau) = \left(-a_0^2\mu_1\gamma/5\right)\tanh\tau$.

Following [7], solution of Eq.(11) may be written as

$$u(\tau)=\sum_{n=0}^{\infty}\varepsilon^n u_n(\tau)=u_0(\tau)+\varepsilon u_1(\tau)+\varepsilon^2 u_2(\tau)+... \tag{19}$$

where: $u_0(\tau)=a_0\,\mathrm{sech}(\tau)$. The parameter μ is also expanded in the power of ε [7]:

$$\mu=\sum_{n=0}^{\infty}\varepsilon^n\mu_{Cn}=\mu_{C0}+\varepsilon\mu_{C1}+\varepsilon^2\mu_{C2}+... \tag{20}$$

I have found, that the values of μ_{C0} and M_{C0} are again given by Eq. (16), and Eq. (17), respectively. The expressions for M_{C0} obtained by the two hyperbolic perturbation methods [6], and [7] coincides. The relation between the amplitude and inverse equilibrium velocity derived by the proposed here perturbation approach, and the relation between the stationary soliton amplitude and it's frequency, obtained by the adiabatic approximation of the SPT compares vary well.

Using [7], I obtained the following perturbed solution of Eq. (11):

$$u(\tau)=u_0(\tau)+\varepsilon u_1(\tau)+O(\varepsilon^2)=a_0\left(1+(4\gamma a_0/5)\ln(\cosh(\tau))\tanh(\tau)\right)\mathrm{sech}(\tau)+O(\varepsilon^2) \tag{21a}$$

and

$$\mu=\mu_{C0}+\varepsilon\mu_{C1}+O(\varepsilon^2)=8a_0^2/5+O(\varepsilon^2)\,. \tag{21b}$$

The Eq.(21a) is similar to the functional shape of the Eq. (14) of the work [4].

4. CONCLUSION

Study of distributed BLA of optical pulses in the presence of IRS in single mode optical fibers is presented. Stability of the fixed point for the soliton amplitude and velocity identified earlier by the adiabatic approximation of SPT [5] is analyzed by numerical solution of linearized system in the neighborhood of fixed point. Obtained analytical expressions for the eigenvalues allow determination of the values of pulse and material parameters for which the fixed point is stable. Relation between the stationary amplitude and the soliton's velocity (Eq.(5)) is numerically verified.

An alternative perturbation approach has been suggested, based on the introduction of the equation of perturbed strongly nonlinear Duffing-Van der Pol oscillator. This equation has been analyzed through the hyperbolic perturbation methods of [6], and [7]. The equilibrium velocity of solution obtained in both methods coincides, and compares very well to one derived by the SPT. The obtained pulse form by the perturbation method of [7], is similar to the form of the Eq. (14) of [4].

5. ACKNOWLEDGMENTS

This research was supported by the Project 102 ни 122-20 with the TU-Sofia, Bulgaria.

REFERENCE

[1] G.P. Agrawal, “Nonlinear Fiber Optics”, Academic Press, third edition, (2001).

[2] G.P. Agrawal, “Applications of Nonlinear Fiber Optics”, Academic Press, (2001).

[3] A. Hasegawa and Y. Kodama, “Solitons in Optical Communications”, Clarendon Press, (1995).

[4] L. Gagnon and P.A. Belanger, ”Soliton self-frequency shift versus Galilean-like symmetry”, Opt. Lett.15, 466-469 (1990).

[5] K. J. Blow, N. J. Doran, and D. Wood, “Generation and stabilization of short soliton pulses in the amplified nonlinear Schrodinger equation”, J. Opt. Soc. Am. B 5, 381-390 (1988).

[6] Y.Y. Chen and S.H. Chen, “Homoclinic and heteroclinic solutions of cubic strongly nonlinear autonomous oscillators by the hyperbolic perturbation method”, Nonlinear Dynamics, 58, 417-429 (2009).

[7] Y.Y. Chen, S.H. Chen, and K.Y. Sze, “A Hyperbolic Linstedt-Poincare method for homoclinic motion of a kind of strongly nonlinear autonomous oscillators”, Acta Mech. Sin., 25, 721-729 (2009).

[8] M. Florjanczyk, and L. Gagnon, ”Exact solutions for a high-order nonlinear Schrodinger equation”, Phys. Rev. A41, pp. 4478-4485 (1990).

GENERATION OF SUB-50-FEMTOSECOND PULSES BY COMPRESSION THROUGH AMPLIFICATION FROM SOLITON BASED ERBIUM FIBER LASER

I. G. Koprinkov

Department of Applied Physics, Technical University of Sofia, 8 Kl. Ochridski Blvd, 1756 Sofia, Bulgaria, phone: 00359 2 965 3072, e-mail: igk@tu-sofia.bg

K. A. Stankov

APHALAS GmbH, Bertha-von Suttner-Str.5, D-37085 Göttingen, Germany, phone: 0049 551 7706147, e-mail: stankov@alphalas.com

Abstract: *A fully integrated all-fiber erbium oscillator-amplifier mode-locked laser system is presented. The oscillator operates in soliton regime using nonlinear polarization rotation mode locking mechanism. Two types of oscillators were built, ordinary and high-repetition rate ones, having about 50MHz and 200MHz fundamental repletion rate, respectively. Instead of the ordinary stretcher-compressor technique, the laser system exploits soliton compression during the amplification. Forward and backward erbium amplifiers were investigated. The pulse duration has been measured using a GRENOUILLE device. Compressed pulses after the amplifier as short as 48fs have been measured. The average power of the amplified emission reaches 30mW. To the best of our knowledge, this is the shortest pulse generated from a soliton based fiber laser system.*

Keywords: fiber lasers, mode locking, femtosecond pulses, solitons.

1. INTRODUCTION

The femtosecond fiber lasers (FSFL) have been subject to intensive research since the demonstration of their capability to generate ultrashort pulses [1]. Although the former generation of femtosecond lasers, *i.e.*, the titan sapphire lasers, remains superior in the generation of ultrashort and high-energy pulses, the FSFL are unsurpassed what concerns the stability and the reliability of the operation, accompanied by compactness and low cost. It makes the FSFL a preferable ultrashort light source for advanced scientific and technology applications in many areas as physics, chemistry, biology, high-precision metrology, medical imaging, micromachining, etc.

Four regimes of mode-locked operation of the fiber lasers can be distinguished: soliton regime [2-4], dispersion managed soliton regime (also known as stretched mode regime) [5, 6], self-similar regime [7, 8], and all-normal dispersion regime [9, 10]. All of the mode-locked regimes, but the soliton one, are capable to generate around or below 100 *fs* long pulses directly from the oscillator or after dechirping of the pulse in an external compressor. It is important to note that earlier recognized soliton lasers that generate $sub-100\ fs$ pulses use optical fiber as an external pulse

compression feedback loop in order to control the intracavity pulse in the otherwise color center bulk active medium [11, 12]. That is why, they cannot be considered as a real fiber lasers working in soliton regime inside the main laser cavity. One of the main advantages of the soliton oscillators is that they generate highly robust pulses in the mode-locking regime. The shortcoming of the soliton lasers is that the generated pulses are relatively long (not less than few hundred femtoseconds, in most of the cases) of relatively low energy, in agreement with the soliton area theorem. Recently, generation of 167 *fs* pulses at $194\,MHz$ repetition rate from a soliton oscillator has been reported [13]. These, according to authors, represent the shortest pulses obtained from soliton oscillator and the highest fundamental repetition rate achieved from all-fiber laser working in nonlinear polarization rotation (NPR) regime.

In this work, we report on a soliton based laser system consisting of soliton erbium ring oscillator and an erbium amplifier. The pulses from the oscillator are compressed by means of amplification without using stretcher-compressor techniques. This makes the entire system extremely simple and reliable in the operation. Sub $50\,fs$ pulses have been measured after the amplifier. To the best of our knowledge, these are the shortest pulses generated from a soliton based laser system. We have also built an oscillator having $200\,MHz$ fundamental repetition rate. This, to our knowledge, is the highest fundamental repetition rate of all-fiber oscillator operating in NPR mode locking regime.

2. EXPERIMENTAL SETUP

2.1. Femtosecond fiber oscillator with 50 MHz repetition rate

The principle scheme of our all-fiber soliton oscillator is shown in Fig. 1. In fact, it resembles the so-called “Tamura laser” [5] but instead of stretched mode regime, it operates in soliton regime. The laser cavity is built on single mode erbium doped gain fiber and Corning SMF-28 passive fiber. The total cavity fiber length is about $380\,cm$. The dispersion of the cavity of the oscillator is estimated to be $D = -0.08\,ps^2$, *i.e.*, it falls within the well expressed soliton regime. The cavity length sets the repetition rate of the laser at $52.6\,MHz$. The gain fiber is pumped by a laser diode (LD) having maximal pump power of $120\,mW$ centered at $976\,nm$. The pump laser emission is delivered into the oscillator ring by means of a wavelength division multiplexer (WDM) $980/1550\,nm$. The laser emission is taken out of the cavity by means of $10/90$ output coupler (OC). Unidirectional operation of the laser is provided by means of a polarization sensitive isolator (PSI). The mode locking operation of the laser is based on the NPR mechanism. It is achieved by means of combination of the PSI placed between two identical polarization controllers PC1 and PC2. The polarization controllers are fiber based and become equivalent to a sequence of effective $\lambda/4 - \lambda/2 - \lambda/4$ plates that allow reproducing any state of the polarization of the

light. The mode locking operation is initiated at suitable pump power after proper adjustment of PC1 and PC2. A number of mode locked states of the oscillator can be found at suitable settings of the PC1 and PC2 and the pump power.

2.2. Femtosecond fiber oscillator with 200 MHz repetition rate

We have also developed a high repetition rate version of our all-fiber ring soliton oscillator described in the preceding section. The laser cavity is generally same as that one shown in Fig. 1 but the total cavity length has been shortened down to about of $1m$. This has been achieved using a reduced configuration of the polarization controllers. In addition, a more powerful pump LD has been employed that is capable to deliver up to $600\,mW$ into the single mode erbium fiber.

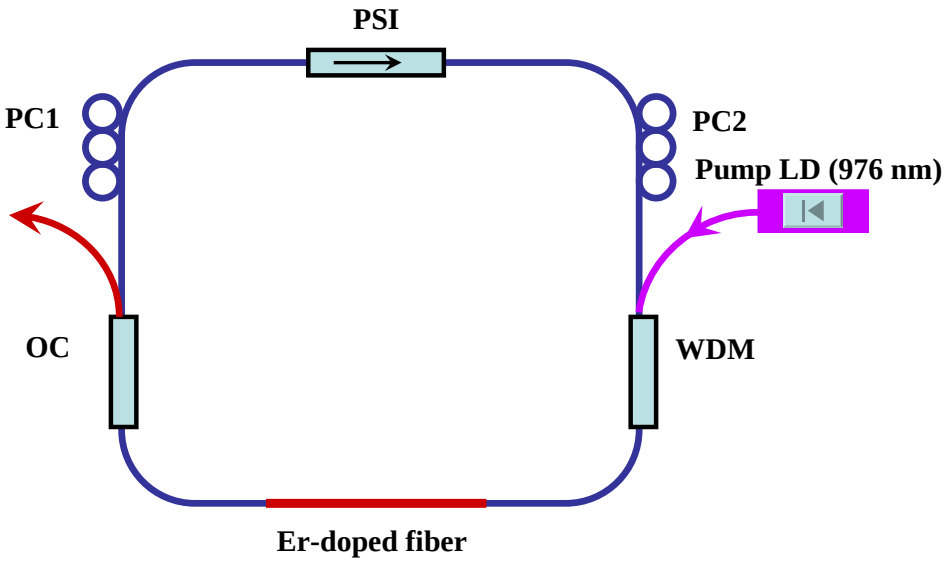


Fig. 1. Principle scheme of all fiber soliton oscillator

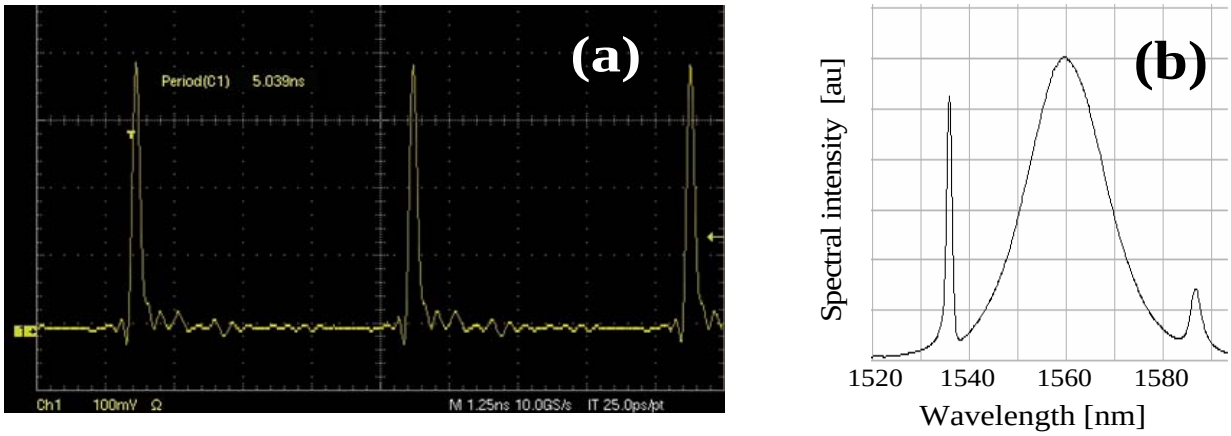


Fig. 2. Mode locking series (a) and spectrum (b) of all fiber high-repetition rate soliton oscillator

The pulse train of the high repetition rate oscillator, Fig. 2(a), has been monitored using a fast photodiode UPD-40-UR2-P (ALPHALAS) with 40 ps rise time and a

6 *GHz* digital oscilloscope Tektronix TDS 6604. The time separation between the pulses has been measured to be 5.039 *ns*, Fig. 2(a), which results in 200 *MHz* repetition rate. This, to our knowledge, is the highest fundamental repetition rate of all-fiber soliton laser using NPR mode-locking mechanism. The spectrum, Fig. 2(b) is substantially wider (20 *nm* FWHM) than that one of the ordinary repetition rate oscillator (see Fig. 4(a)). The average power of the oscillator is 9 *mW* at 570 *mW* pump power of the laser diode. It corresponds to 45 *pJ* pulse energy at the specified repetition rate.

2.3. Oscillator-amplifier femtosecond fiber laser system

The entire laser system, consisting of a soliton oscillator and a single amplifier, is shown schematically in Fig. 3. The active medium of the amplifier is the same single mode fiber as that one of the oscillator. Polarization insensitive isolator PII is placed after the oscillator in order to prevent interference between the oscillator and the amplifier. To prevent undesired feedback, the amplifier output is isolated by WDM-isolator (WDM-I). It also serves as an element for launching of the pump power from the LD into the erbium fiber of the amplifier. The pump power delivered by the LD into the gain fiber is about 230 mW at 1480 *nm*. The LD emission centered at 1480 *nm* ensures efficient pumping of the amplifier due to the low quantum defect.

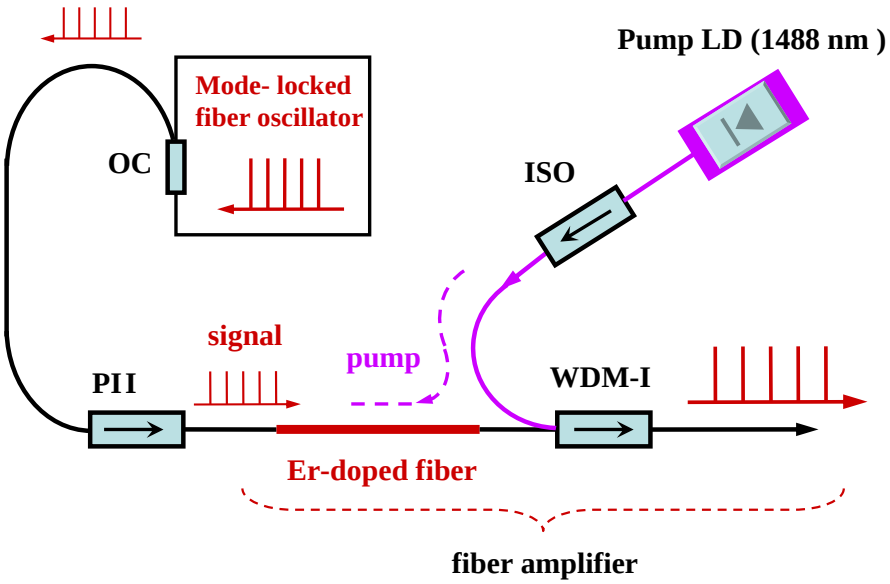


Fig. 3. All fiber oscillator-amplifier erbium laser system

3. RESULTS AND DISCUSSION

In this section, the properties of 50 *MHz* oscillator-amplifier laser system are presented. The evolution of the spectrum of the output pulse at different pump power in the amplifier is shown in Fig. 4: 80 mW (a), 120 mW (b), and 215 mW (c). The spec-

trum at low pump level, Fig. 4(a), almost repeats that one from the oscillator. Once the pump power grows above given level, Fig. 4(b), the spectrum at the output acquires an additional structure due to the self-phase modulation (SPM) in the amplifier fiber and the WDM-I fiber leads. The sharp lines from the both sides of the spectrum centered at 1550 *nm* and 1580 *nm* are the so-called Kelly sidebands. The latter originate from the periodic perturbations on the soliton in the oscillator due to the gain and loss. The SPM generate "red" frequency components on the leading edge of the pulse and "blue" frequency components on the trailing edge of the pulse. Such a positively chirped pulse undergoes compression while propagating along the fiber of the amplifier and the fiber leads of the WDM-I that have negative dispersion. The interplay of those two effects results in a compressed and amplified pulse at the output of the laser system.

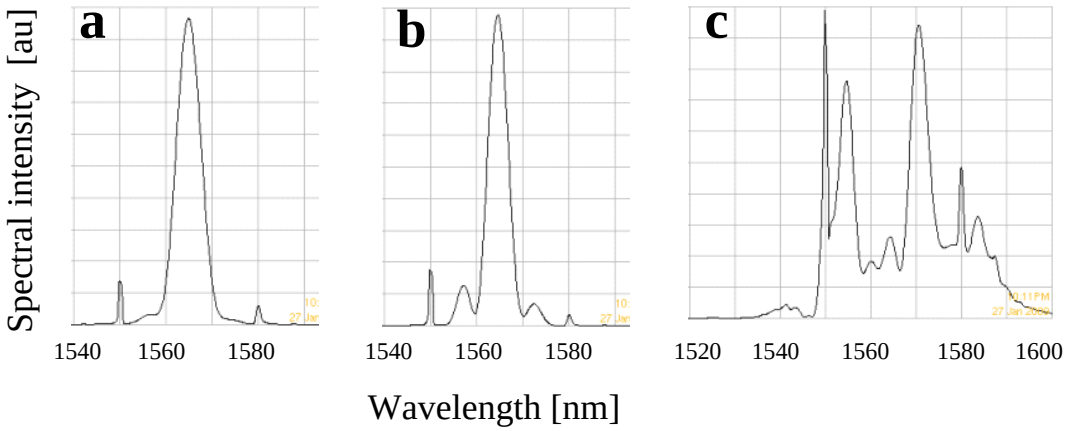


Fig. 4. Evolution of the laser spectrum at different pump power:
(a) 80 mW, (b) 120 mW, (c) 215 mW

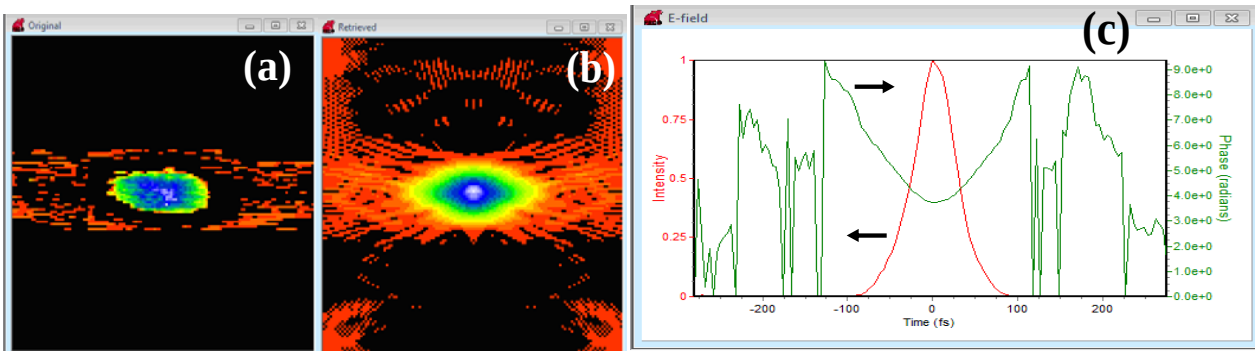


Fig. 5. Original (a), retrieved (b) FROG traces and the
temporal intensity and phase profiles (c) of the pulse.

The pulse duration has been measured using second harmonic GRENOUILLE [14]. A Proustite crystal has been used as the nonlinear crystal for the second harmonic generation. The original, the retrieved time-frequency (FROG) traces and the temporal intensity and phase profiles of the pulse are shown in Fig. 5 (a), (b) and (c), respectively. Inhomogeneities in the Proustite crystal used lead to some extra noise in the FROG traces and result in slight asymmetry in the time profile of the intensity.

The pulse duration from $48\ fs$ to $56\ fs$ (FWHM), depending on the proper optimization of the pump power of the amplifier, has been measured by GRENOUILLE. The average power of the amplified emission is about $30\,mW$. It gives $0.57\,nJ$ energy and $1.14x10^4\,W$ peak power of the pulse. To the best of our knowledge, the optical pulses so obtained are the shortest pulses generated from a soliton based fiber laser system.

4. CONCLUSIONS

An all-fiber erbium oscillator-amplifier mode-locked laser system is presented. The oscillator, having about $50\,MHz$ repetition rate, operates in soliton regime using nonlinear polarization rotation mode locking mechanism. In addition, high-repetition rate oscillator, having $200\,MHz$ fundamental repetition rate, has been also built. This is the highest fundamental repletion rate of all-fiber laser using nonlinear polarization rotation mode locking mechanism. The laser system exploits soliton compression during the amplification without using stretcher-compressor technique. Pulses shorter than $50\ fs$ and $30\ mW$ average power have been measured after the amplifier. To the best of our knowledge, these are the shortest pulses generated directly from a soliton based fiber laser system.

5. ACKNOWLEDGEMENTS

This work was done in the frame of the Contract for joint research activities between the Department of Applied Physics, Technical University of Sofia, and ALPHALAS GmbH. I. G. Koprinkov acknowledges with thanks the hospitality and the research support from ALPHALAS GmbH, where this work has been done.

REFERENCES

[1] M. E. Fermann, M. Hofer, F. Haberl, S. P. Craig-Ryan, Electron. Lett., **26**, 1737 (1990)
[2] J. D. Kafka, T. Baer, Opt. Lett. **14**, 1269 (1989)
[3] I. N. Duling III, Opt. Lett. **16**, 539 (1991)
[4] D. J. Richardson, R. I. Laming, D. N. Payne, M. W. Phillips, V. J. Matsas, Electron. Lett. **27**, 730 (1991).
[5] K. Tamura, E. P. Ippen, H. A. Haus, L. E. Nelson, Opt. Lett. **18**, 1080 (1993)
[6] K. Tamura, L. E. Nelson, H. A. Haus, E. P. Ippen, Appl. Phys. Lett., **64**, 149 (1994).
[7] M. E. Fermann, V. I. Kruglov, B. C. Thomsen, J. M. Dudley, J. D. Harvey, Phys. Rev. Lett., **84**, 6010 (2000)
[8] F. Ö. Ilday, J. R. Buckley, W.G. Clark, F. W. Wise, Phys. Rev. Lett. 94, 213902 (2004)
[9] A. Chong, J. Buckley, W. Renninger, F. Wise, Opt. Express **14**, 10095 (2006)
[10] A. Chong, W. H. Renninger, F. W. Wise, Opt. Lett. **32**, 2408 (2007)
[11] L. F. Mollenauer, R. H. Stolen, Opt. Lett., **9**, 13 (1984)
[12] F. M. Mitschke, L. F. Mollenauer, Opt. Lett., **12**, 407 (1987)
[13] J. Chen, J. W. Sickler, E. P. Ippen, F. X. Kärtner, Opt. Lett., **32**, 1566 (2007).
[14] S. Akturk, M. Kimmel, R. Trebino, Opt. Express, **19**, 4483 (2004).

OPTIMISATION OF THE BLOW-WOOD SPLIT-STEP FOURIER METHOD FOR INVESTIGATING OF THE N-SOLITON BOUND STATES IN OPTICAL FIBERS

Todor N. Arabadzhiev, Ivan M. Uzunov

Department of Applied Physics, Technical University-Sofia,
blv. Kliment Ohridski, 8, Sofia, Bulgaria

Abstract: *We propose optimized variants of the method of Blow and Wood by adopting adaptive propagation steps-size methodology for investigating of the N-soliton bound states. It is shown that the adaptive step-size methods of Blow and Wood have an excellent numerical accuracy and stability for long distance propagation calculations. To show the advantages of the new methods a comparison with the constant step-size Blow-Wood method was done for the 7-soliton bound state. The physical validation of the methods was done by investigating of the self-compression effect. It has been found the sample's number low-end in the self-compression points below which the investigated methods became unstable.*

1. INTRODUCTION

By using of the NSE we can compute and theoretically predict physical phenomena such as the modulation instability and propagation of N-soliton bound states in optical fiber. Historically, after the development of a wide variety of finite-difference methods for the NSE [1] and the development of the algorithm of the fast Fourier transformation, Hardin and Tapert for the first time investigated the NSE by using of the SSFM [2]. This method was developed further in many versions [3-7], but the question for the limits of applicability and effectivity of each variant and especially for the N-soliton bound states still remains open.

The investigating of the high order N-soliton bound states in optical fiber is interesting because of the fact that the nonlinear self-compression phenomenon dramatically limits performance of the standard constant step-size SSFM methods [1-7]. A solution of this problem may be using of some numerical techniques as the increasing of the sampling rate, decreasing of the propagation step-size or just decreasing of the sampling window. However, because of the fast computation time increasing, if decreasing of the propagation step-size method is chosen, an adaptive step-size technique must be applied. In ref. [7] are presented ideas and methods for selection of the propagation step-size of the SSFM, which have been used in this paper for development of new high-order accurate adaptive step-size version of the Blow-Wood method which as well-known has a fifth-order local accuracy [6]. In this paper the Satsuma-Yajima N-soliton bound states are investigated numerically with several new developed versions of the adaptive step-size Blow-Wood method in effort to find the best accuracy and correspondingly numerical stability. As well known, in a numerically stable algorithm, truncation error or errors in the input lessen significantly

as the algorithm proceeds, having little effect on the final output [1,4]. In our research, the results are presented in native way by comparing of the numerical and the well-known exact amplitude of the N-solitons [8].

2. DESCRIPTION OF THE METHODS

The developing of algorithm with adaptive propagation step-size is necessary for modeling of high order N-soliton bound sates, because of the fact that the nonlinear compression effect and efforts for avoiding of the aliasing in the frequency area make the ordinary [4] methods too slow and additional efforts for finding of the optimal propagation step-size are needed. To improve the quickness of the *Blow-Wood* method by finding of an optimal propagation step-size, we adopt several methodologies for constructing of the adaptive step-size methods presented in ref. [7]. As a result by adopting of two ideas for adapting of the propagation step for *i)*: limiting of the relative local error increasing (*Local Error Method*) and *ii):* limiting of the nonlinear phase increasing (*Rotating Phase Method*) several adaptive step-size versions of the *Blow-Wood* method were developed. We have found also that a slightly different (form ref. [7]) way for changing of the propagation step (presented below) ensure a good performance of the methods. We have tested the developed adaptive step-size methods in the cases of N-soliton bound cases up to N=7 and it was received that the numerical stability along the extremely long propagation distance was excellent. We will present below a numerical scheme of the adaptive step-size Blow-Wood methods for computing of the normalized NSE (ref. 3), where $\varphi(z,T)$ is normalized amplitude as a function of the normalized propagation distance z and time T:

The general part for the methods developed on the base of the idea for limiting of the relative error (BWA1 & BWA2 methods):

i) Coarse iteration:

$$\varphi\left(z+\Delta z,T\right)_{coarse}=\underbrace{\varphi\left(z+\frac{\Delta z}{j},T\right)_{coarse}}_{j\_times} \tag{1}$$

$$\varphi\left(z+\frac{\Delta z}{j},T\right)_{coarse}=\underbrace{\exp\left[\frac{\Delta z}{2j}\hat{D}\right]\exp\left[\frac{\Delta z}{j}\hat{N}\right]\exp\left[\frac{\Delta z}{2j}\hat{D}\right]}_{4times}$$

$$\exp\left[-\frac{\Delta z}{j}\hat{D}\right]\exp\left[-\frac{2}{j}\Delta z\hat{N}\right]\exp\left[-\frac{\Delta z}{j}\hat{D}\right]\underbrace{\exp\left[\frac{\Delta z}{2j}\hat{D}\right]\exp\left[\frac{\Delta z}{j}\hat{N}\right]\exp\left[\frac{\Delta z}{2j}\hat{D}\right]}_{4times}\varphi\left(z,T\right)$$

ii) Fine iteration:

$$\varphi\left(z+\Delta z,T\right)_{fine}=\underbrace{\varphi\left(z+\frac{\Delta z}{j+1},T\right)_{fine}}_{(j+1)\_times} \tag{2}$$

$$\varphi\left(z+\frac{\Delta z}{j+1},T\right)_{fine} =$$

$$\underbrace{\exp\left[\frac{\Delta z}{2(j+1)}\hat{D}\right]\exp\left[\frac{\Delta z}{j+1}\hat{N}\right]\exp\left[\frac{\Delta z}{2(j+1)}\hat{D}\right]}_{4times}\exp\left[-\frac{\Delta z}{j+1}\hat{D}\right]\exp\left[-\frac{2}{j+1}\Delta z\hat{N}\right]\exp\left[-\frac{\Delta z}{j+1}\hat{D}\right]$$

$$\underbrace{\exp\left[\frac{\Delta z}{2(j+1)}\hat{D}\right]\exp\left[\frac{\Delta z}{j+1}\hat{N}\right]\exp\left[\frac{\Delta z}{2(j+1)}\hat{D}\right]}_{4times}\varphi(z,T)$$

iii) Relative local error:

$$\varepsilon \le \left|\left(\varphi_{coarse}-\varphi_{fine}\right)/\varphi_{fine}\right| \tag{3}$$

iv) If the relative local error is greater than the admissible tolerance ε the following schemes for the adaptation of the propagation step-size are applied:

***Blow-Wood-Adaptive 1 method (BWA1)*:** We increase of the *j* value with *1* or two times (faster increasing of *j* lead to lower computational speed) and repeat the computations until the relative local error became smaller than admissible tolerance *ε*. However, if the computation speed becomes too low we continue the computation with the next propagation step. The amplitude φ_{fine} is used for the next propagation step. If the relative local error became much smaller than the admissible tolerance, we may decrease the value of *j* with 1:

***Blow-Wood-Aadaptive 2 method* (*BWA2)*:** We increase or decrease the *j* value following recipe given in ref. [7] in description of the *Local Error Method (LEM)*. However, if the computation speed becomes too low we continue the computation with the next propagation step. On each step the numerical amplitude is computed by using of an empirical formula proposed in ref. [7]:

$$\varphi = (4/3)\varphi_{fine} - (1/3)\varphi_{coarse} \tag{4}$$

From the presented above methods it can be shown that the numerical accuracy and stability depends from several components of the adaptive method: the splitting error of the basic *Blow-Wood* method, the sampling rate and the propagation step's size which are connected with the admissible tolerance ε. The method is so fast so the tolerance ε and initial step size Δz permits – the larger tolerance and propagation step mean the faster computations. One order smaller tolerance ε used in the adaptive method increases the numerical accuracy greatly but also increases the number of the necessary iterations. To decrease the variation of the amplitude and to achieve the optimal numerical stability we have to settle an appropriate maximal value of the propagation step size which in our algorithm (6-8) is equal to the initial step size. In this connection, if we start the computations with appropriate small propagation step but with very large tolerance, the propagation step size may have a constant value

equal to the initial step size, along the all computation distance. Then the adaptive BWA1,2 methods are equivalent of the constant Blow-Wood method with two times smaller propagation step.

Rotating Phase Blow-Wood method (RPM): Other variable step-size method is the nonlinear rotating phase method described in ref. [7]. Our version of this method is based on the using of the Blow-Wood method presented by formula (2). We impose an upper limit on the nonlinear phase increment during the propagation, which determine the propagation step-size (or increasing/decreasing of the parameter j in formula (2)). The used value of maximal nonlinear phase $\xi_{max,\ nl}$ was *0.02 rad*. The bound of the step-size satisfy the condition:

$$\Delta z \leq \frac{\phi_{\max,nl}}{\phi_{nl}} \tag{5}$$

For the practical reasons in the presented computations with the *RPM* the maximal propagation step-size was chosen to be equal of the initial step-size. If the nonlinear phase increment is higher from the mentioned phase limit, the step-size is decreased in the same way as it was described for the *BWA1* method ($j=j+1$ in (2)). Contrary is happened if the nonlinear phase increment is higher from the mentioned phase limit ($j=j-1$ in (2)).

3. RESULTS

First we can see in the figure 1 the performance of the constant step size Blow-Wood method for the 7-soliton bound state. Such a computation will be done with developed adaptive version of the Blow-Wood Method.

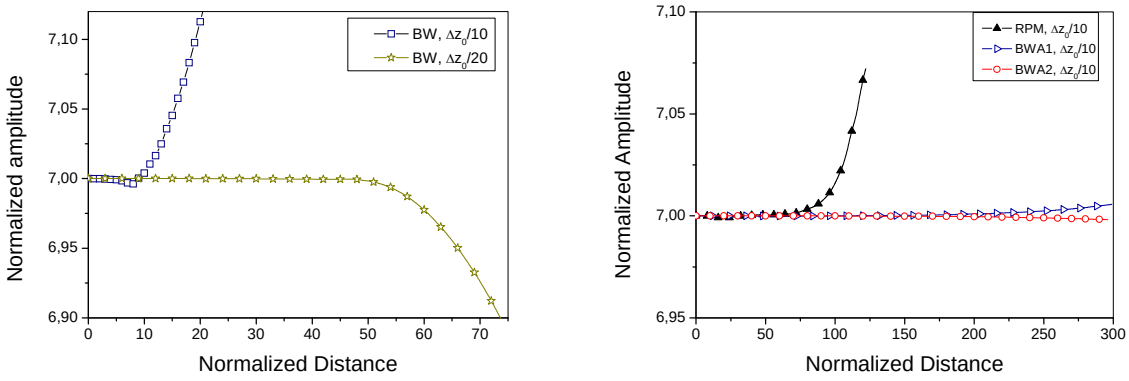


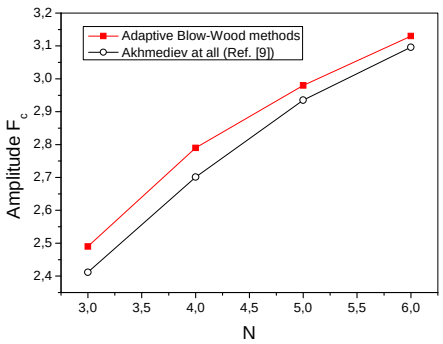
Figure 1 (left). 7-soliton bound state modeled with constant step-size Blow-Wood method (BW) with step-size: $\Delta z_0/10$ and $\Delta z_0/20$, the sampling rate: k=32768.

Figure 2 (right)**.** 7-soliton bound state and sampling rate *k=32768,* maximal propagation step-size $\Delta z_0/10$; *i)* the *BWA1* and *BWA2* methods for the cases $\varepsilon = 10^{-7}$; *ii) Rotating Phase Method (RPM)* with *maximal phase* $\xi_{max,\ nl}$ *=0.02.*

The basic parameters used in the computations were: the basic propagation step-size is $\Delta z_0 = 0.001$ as the initial and correspondingly the maximal propagation step-size was $\Delta z_0 / 10$, the distance is normalized in soliton periods. In all computation the time window (transversal sampling window) size L is connected with the sampling rate by relation $L = \sqrt{2\pi k}$. We consider such a case as the most difficult on the computational point of view.

Physical validation of the methods

The physical validation of the numerical methods by numerical evaluation of the self-compression effect at the point of maximal compression and comparison with some well-known results [10] has to be done. First we will evaluate and compare with ref. [10] the peak amplitude of the self-compressed pulse divided by the peak amplitude of the initial one: $F_c = \left|\varphi\left(z_{\max}, 0\right)\right| \Big/ \left|\varphi\left(0,0\right)\right|$, where z_{max} is the point of maximum self-compression. As we can see in fig. 3 a good coincidence of our results with the same in ref. [10] was observed.



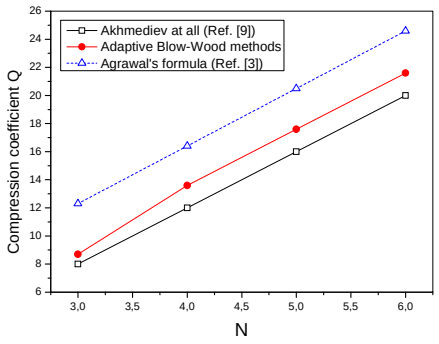


Figure 3 (left). The amplitude coefficient F_c as a function of N in the maximal self-compression point ref. [9]. The amplitude coefficients are identical for the presented adaptive Blow-Wood methods.

Figure 4 (right). The compression coefficient Q as a function of N for the maximal self-compression point, the compression coefficients are identical for the presented adaptive Blow-Wood methods.

The compression factor is computed by: $Q = T_{FWHM,0} \Big/ T_{FWHM,z_{opt}}$, where $T_{FWHM,0}$ is the initial pulse time length and $T_{FWHM,Zopt}$ is the pulse time length in the points with maximal self-compression. The received results (fig. 4) are in good agreement with the same results determined numerically in ref. [9] or by using of the empirical formula given in ref. [3].

From the figure 5 we can see that in the compression points the number of the samples decreases dramatically. As a result the aliasing effect and correspondingly the numerical artifacts are appeared. Such a decreasing can be delayed by increasing of the sampling rate (discretezation) but using of this method is very limited. We have found also that has a critical threshold or low-end from *six* samples in the self-

compression points below which all considered SSFM methods became unstable and the aliasing effect in the frequency area is observed even in a short propagation distance (first compression point). If the sampling rate is higher - the samples in the compression point are more. It may be concluded that the self-compression effect and the limited sampling rate put the boundaries in front of the SSFM methods for modeling of high order N-soliton bound states. Such a conclusion is valid especially if the sampling window size is connected with the sampling rate as it was done in our case. The used numerical parameters are: the sampling rate: 8192 and 16384 for N=6.

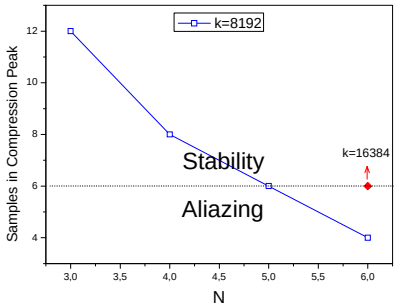


Figure 5. The number of the samples in the maximal self-compression point as a function of N – presented as a generalization on the base of the calculations with all considered adaptive methods (BWA1, BWA2, and RPM) - ref. [9].

4. CONCLUSION

It was developed three adaptive step-size versions of the Blow-Wood method. The methods were compared and tested on the case of the N-soliton bound case computation as an excellent long distance numerical stability was achieved. The physical validation of the methods was performed by investigation of the nonlinear self-compression effect in N-solitons evolution as the results are in good agreement with the same one presented in ref. [3] and ref. [9]. It is established the fact that for ensuring of a good numerical stability a critical number of six samples in the self-compression points is needed.

5. ACKNOWLEDGMENTS

The research was supported by the Project 102 ни 122-20 of the Technical University-Sofia.

REFERENCES

[1] T. Taha, M. Ablowitz, “Analytical and Numerical Aspects of Certain Nonlinear Evolution Equations. II. Numerical, Nonlinear Schrodinger Equation”, J. Comput. Physics 55, 203-230, (1984)

[2] R. Hardin, F. Tapert, “Application of the split step Fourier method to the numerical solution of nonlinear and variable coefficient wave equation.”, SIAM Rev. Cronicle 15,423, (1973)

[3] G. Agrawal, “Nonlinear Fiber Optics”, Acad. Press, (2001)

[4] T. Hohage, F. Schmidt, “On the Numerical Solution of Nonlinear Schrodinger Type Equations in Fiber Optics”, Konrad-Zuze-Zentrum fur Informationstechnik Berlin, ZIB-Report 02-04, (2002)

[5] M. Potasek, G. Agrawal S. Pinault, "Analytic and numerical study of pulse broadening in nonlinear dispersive fibers", JOSA B, 3 (2), 205-211, (1986)

[6] K. Blow, D. Wood, "Theoretical Description of Transient Stimulated Raman Scattering in Optical Fibers", 25 (12), 2665-2673, (1989)

[7] O. Sinkin, R. Holzlöhner, J. Zweck, C. Menyuk, “Optimization of the Split-step Fourier Method in Modeling Optical Fiber Communications Systems”, J. Lightwave Technologies 21, 61-68, (2003)

[8] J. Satsuma, N. Yajima, “Initial value problems of one dimensional self-modulation of nonlinear waves in dispersive media”, Progr. Theor. Phys. Suppl. 55, 286-306, (1974)

[9] N. Akhmediev, N. Mitzkevich, “Extremely high degree of N-soliton pulse compression in an optical fiber”, IEEE J. Quant. Electronics 27, 3, 849-857, (1991)

ALGORITMUS FOR FILTERING AND ADAPTIVE FILTER DESIGN

Antonio Andonov, Mariana Mihova

Department of Communication and Safety Equipment end Systems
Higher School of Tansport "T. Kableshkov"
158 G. Milev str., 1574 Sofia, andonov@vtu.bg

***Abstract**: Conventional signal processing systems for the extraction of information from a received or observed signal, operate in an open loop manner, i.e. the same processing operation is carried out in the previous the correct result in the previous time interval. In other words, the make the assumption that the signal degradation is a known and time-invariant quantity. This article examines adaptive processors are designet to approximate these estimators. They operate with a closed loop arrangement where the processor frequency response is controlled by the feedback algorithm. This permits them to compensate for time-varying distortions and still achieve performance, close to the optimal estimator function. Adaptive filters use a programmable filter whose frequency response or transfer function is altered, or adapted, to pass without degradation the desired components of a signal and to attenuate the undesired or interfering signals, i.e. to minimize the distortions on the input signal.*

Keywords: adaptive filters, adaptive filter algorithm

1. INTRODUCTION

The strategy by which the impulse response of the adaptive filter is altered is the adaptive filter algorithm. An adaptive filter is thus a time-varying filter whose impulse response at a particular time is dependent on the input sequence, the training sequence and adaptive filter algorithm. The time-varying nature of an adaptive filter gives rise to the concept of convergence. In a stationary environment, the convergence performance is a measure of how many data samples are required for the impulse response of the adaptive filter to approach that of the optimal filter. In a non-stationary environment, the convergence performance is also a measure of how closely the impulse response of the adaptive filter follows the now time-varying impulse response of the optimal filter. Adaptive filters may be both infinite impulse response IIR or finite impulse response FIR in nature. In this article, only adaptive FIR filters will be considered as FIR filters are inherently stable. Thus, provided the feedback loop is properly designed, then stability of the overall adaptive filter will be ensured.

2. ALGORITMUS FOR ADAPTIVE FITERING

In this article are proposed an adaptive FIR filter design with LMS gradient algorithm. The block diagram of the adaptive FIR filter is shown in Fig. 1. The input data

sequence {y(n)}, is convolved with the FIR sequence, $\{h_i(n)\}$, and the resulting filter output $\{\hat{x}(n)\}$, Equation (1), is:

$$\hat{x}(n) = \sum_{i=0}^{N-1} h_i(n-1)y(n-1) \tag{1}$$

This can also be written as a vector product:

$$\hat{x}(n) = h^T(n-1)w(n) \tag{2}$$

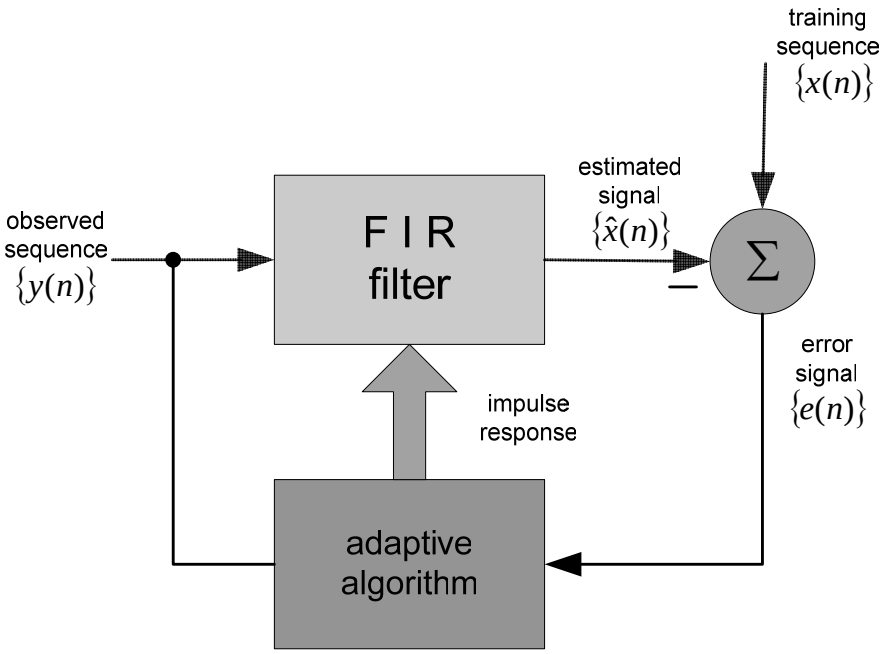


Fig. 1. General adaptive filter configuration

The difference here is that the weights of the FIR filter are adapted in a time varying manner: i.e. h(n-1) rather than simply h. The current output, $\hat{x}(n)$, is calculated using the previous set weight values, h(n-1). The training ssequence, x(n), is substracted from the output signal to yield a scalar error signal, e(n) [1].

$$e(n) = x(n) - \hat{x}(n) \tag{3}$$

The error is now used in conjunction with the input signal vector, y(n), to determine the next set of filter weight values, h(n). The notation h(n) is used because data up to and including y(n) and x(n) have been used to calculate h(n). The eventual objective of the adaptive algorithm is to estimate the optimum impulse response.

One method of defining the impulse response of an adaptieve filter is to choose the coefficients to minimize a sum of squared errors cost function:

$$\xi = \sum_{k=1}^{n} (x(k) - \hat{x}(k))^2 \tag{4}$$

When this cost function is minimized with respect to the impulse response vector, h(n), associated with the estimate, $\hat{x}$(n), we obtain a least squares (LS) estimate. The impulse response vector h(n), which minimizes the sum of squared errors cost function, is now a function of the available data samples rather than ensemble averages. While an expression for h(n) could be derived by a similar method to that employed to derive the Wiener filter, a more elegant approach is to note that the two optimization are similar and obtain the solution by analogy. The optimum impulse response h_{opt} which minimizes the MSE is thus the solution to a set of N simultaneous linear equdtions [1]. Thus:

$$R_{yy}(n)\, h(n) = r_{yx}(n) \tag{5}$$

where:

$$R_{yy}(n) = \sum_{k=0}^{n} y(k)\, y^{T}(k) \tag{6}$$

and

$$r_{yx}(n) = \sum_{k=0}^{n} y(k)\, x(k) \tag{7}$$

In many applications of adaptive filters it is necessary then the impulse response be modified or updated as new data samples appear. The most direct method of meeting this requirement is to increment the upper limits of the summations of equations (4), (6) and (7), and completely resolve equation (5). However, a computationally more efficient method is derive a time recursion for h(n) in terms of the previous least squares solution h(n-1) and new data, y(n) and x(n). The first step is to obtain recursions for $R_{yy}(n)$ and $r_{yx}(n)$. This is straightforward as they are both defined as summations (6) and (7).

$$R_{yy}(n) = R_{yy}(n-1) + y(n)\, y^{T}(n) \tag{8}$$

$$r_{yx}(n) = r_{yx}(n-1) + y(n)\, x(n) \tag{9}$$

Substitute for r_{yx} in equation (9) using (5):

$$R_{yy}(n)\, h(n) = R_{yy}(n-1)\, h(n-1) + y(n)\, x(n) \tag{10}$$

Then use equation (8) to replace $R_{yy}(n-1)$:

$$R_{yy}(n)\, h(n) = \left(R_{yy}(n) - y(n) y^{T}(n)\right) h(n-1) + y(n)\, x(n) \tag{11}$$

After premultiplying by $R_{yy}^{-1}(n)$, this yields:

$$h(n) = h(n-1) + R_{yy}^{-1}(n)\, e(n) \tag{12}$$

When the substitution:

$$e(n) = x(n) - \hat{x}(n) \tag{13}$$

Is also included

$$e(n) = x(n) - h^T(n-1)\ y(n) \tag{14}$$

A recursion for $R_{yy}^{-1}(n)$ may be obtained by application of Sherman- Morrison identity to equation (8):

$$R_{yy}^{-1}(n) = R_{yy}^{-1}(n-1) - \frac{R_{yy}^{-1}(n-1)\ y^T(n)\ y^T(n)\ R_{yy}^{-1}(n-1)}{\left(1 + y^T(n)\ R_{yy}^{-1}(n-1)\ y(n)\right)} \tag{15}$$

Together, equations (12), (14), and (15) are known as the recursive least squares (RLS) algorithm.

REFERENCES

[1] E. Ifeachor and B. W. Barrie, Digital Signal Processing, Pearson Education Limited, New York, 2003.

NUCLEOTIDE GENOMIC SIGNALS OF DNA SEQUENCES

Paul Dan Cristea, Rodica Aurora Tuduce

BioMedical Engineering Center, University "Politehnica" of Bucharest, Spl. Independentei 313, 060042 Bucharest, sect.6, Romania, phone: +40-316 95 69, e-mail: {pcristea, trodica}@dsp.pub.ro

Abstract. *The paper presents the Nucleotide Genomic Signal (NuGS) methodology. NuGSs result from the conversion of symbolic nucleotide sequences into digital genomic signals. Using an unbiased mapping for this conversion allows to successfully apply the method to a large variety of problems related to DNA analysis. The range includes the study of regularities and symmetries at the scale of genomes or chromosomes, but also the local analysis of nucleotide sequences, such as genes, motifs or markers. This is especially important in the study of pathogen variability for tracking the development of pathogen resistance to drugs. The approach reveals not only ostensive features of extant DNA sequences, but also ancestral features that existed in early DNA sequences, but disappeared during evolution, in the process of species separation.*

Keywords: Nucleotide symbolic sequences, Nucleotide genomic signals, Genome symmetries, Pathogen variability, Sequence prediction.

1. INTRODUCTION

The nucleotide genomic signal (NuGS) methodology uses the conversion of symbolic nucleotide sequences into digital genomic signals [1]-[4]. There have been many attempts to represent nucleotide sequences as digital signals [5-9], most using some specific property of the nitrogenous bases, *e.g.*, the electron-ion interaction potential [9]. But such approaches are biased, *i.e.*, are adapted only for those applications for which the chosen property is of central importance, but not for other. The representation we adopted is unbiased, i.e., it is adequate for a large variety of problems related to DNA analysis. To achieve such a versatility, the representation is not using the cardinality of numbers – their capacity to express and handle quantities, using instead the ordinality of numbers – their capacity to classify objects, to extract and handle the order of objects. We have used this type of NuGS methodology for a large variety of applications, going from the global to the local analysis of nucleotide sequence features. The large scale analysis of genomic sequences reveals not only ostensive features of extant DNA sequences, of the size of entire chromosomes, or even entire genomes [10], but also ancestral features, which existed in DNA sequences, but disappeared during evolution [11], under the pressure of species separation. The method reveals surprising regularities along a sequence, in both prokaryotes and eukaryotes, which would be difficult or impossible to find by using only symbolic genomic sequences and standard statistical and pattern matching methods. The local analysis of nucleotide sequences is also important for the study of gene dynamics [12], especially in the context of pathogen variability and tracking the development of pathogen resistance to treatment [13]-[15]. Such results can help the fast diagnosis

and early assessment of drug efficiency, allowing a systematic use of the recent advances in molecular medicine to support clinical decisions [13], [15]. The genome regularities also allow to predict the nucleotides in a DNA sequence, using a methodology similar to time series prediction, and to estimate the cell self repair potential in processes such as replication, transcription or crossover [2]-[3].

2. NUCLEOTIDE GENOMIC SIGNALS

A DNA molecule comprises two strands which wind one around the other to form a *double helix*. Each strand is a *heteropolymer* built of four monomers – the *nucleotides*. A nucleotide comprises a sugar (the *deoxyribose*), a *phosphate group* and one of the four *nitrogenous bases*: A–*adenine*, C–*cytosine*, G–*guanine*, and T–*thymine* (or U–*uracil*, for RNA sequences), which gives the identity of the nucleotide. The strands are complementary: only the pairs A–T and C-–G normally exist.

The four nucleotides can be arranged in classes according to the three main dichotomies in their biochemical properties:

(1) *molecular structure* – A and G are *purines* (R), C and T are *pyrimidines* (Y);

(2) *strength of link* –A and T are linked by two hydrogen bonds (W - *weak bond*), C and G are linked by three hydrogen bonds (S - *strong bond*);

(3) *radical content* – A and C contain the *amino* (NH_3) group (M class) in the major grove, T and G the *keto* (C=O) group (K class).

To fully conserve the symmetry of the nucleotides and to classify them according to the three mentioned dichotomies, we have proposed a vector nucleotide tetrahedral representation [1]-[2]. After a recursive process, it resulted that, without loosing any information, it is possible to give up the less important amino-keto separation, and use a simpler two-dimensional nucleotide representation:

$$a = 1 + j, \quad t = 1 - j, \quad g = -1 + j, \quad c = -1 - j. \tag{1}$$

This mapping conserves all information in the initial symbolic sequence, without introducing artifacts related to the representation.

Two nucleotide genomic signals (NuGSs) can be associated to the phases of the complex representations $C\{Nu(k)\}$ of the nucleotides $Nu(k)$ in a sequence of n_b bases, $k \in \{1,2,\ldots,n_b\}$.

– The cumulated phase:

$$\theta_c(h) = \sum_{k=1}^{h} \arg(C\{Nu(k)\}) = \frac{\pi}{4}\left[3\left(n_G(h) - n_C(h)\right) + \left(n_A(h) - n_T(h)\right)\right] = \frac{\pi}{4} N(h), \tag{2}$$

where $n_A(h)$, $n_C(h)$, $n_G(h)$ and $n_T(h)$ are the number of occurences of adenine, cytosine, guanine and thymine nucleotides, respectively, in the first h samples of the sequence. $N(h)$ is the *nucleotide imbalance* – a NuGS which is a signature of the distribution of nucleotides in the sequence;

– The *unwrapped phase* is the phase of the elements in the sequence, corrected by adding $2m\pi$, $m \in Z$, Z – the set of integers, so that the absolute value of the difference between the phase of an element in the sequence and the phase of its preceding element be smaller than π:

$$\theta_u(1) = \arg(C\{Nu(1)\}),\ \theta_u(h) = \arg(C\{Nu(k)\}) + 2m\pi,\ m \in Z,$$
$$\text{so that } \left|\theta_u(h) - \theta_u(h-1)\right| < \pi,\ h \in \{2,\ldots,n_b\}. \tag{3}$$

For the case of the mapping (1), it results [1]-[4]:

$$\theta_u(h) = \theta_u(1) + \frac{\pi}{2}[n_+(h) - n_-(h)] = \theta_u(1) + \frac{\pi}{2}P(h), \quad h \in \{2,\ldots,n_b\}, \tag{4}$$

where n_+ is the number of positive pairs (A→G, G→C, C→T, T→A) in the sequence and n_- is the number of negative pairs (A→T, T→C, C→G, G→A) existing in the first h samples of the sequence, $h \in (2,\ldots, n_b\}$. $P(h)$ is the *nucleotide pair imbalance* – a NuGS which is a signature of the distribution of pairs of nucleotides in the sequence. For long sequences $\theta_u(1)$ is negligible.

Because of their direct statistical significance, it is convenient to use the nucleotide imbalance (N) and the nucleotide pair imbalance (P), instead of the cumulated phase (θ_c) and the unwrapped phase (θ_u), respectively.

3. SYMMETRIES IN NUCLEOTIDE GENOMIC SIGNALS

Below we the phase analysis of NuGSs for several eukaryote and prokaryote genomes.

Figure 1 presents the nucleotide imbalance (N) and the nucleotide pair imbalance (P) along the DNA sequence of *Homo sapiens* chromosome 11, downloaded from GenBank [16]. N is close to zero at the scale of the figure, in accordance to Chargaff's law [17] which states that $n_G \approx n_C$ and $n_A \approx n_T$, so that they balance each other in (2). A distinctive feature of P is its almost linear variation along the chromosome, which corresponds to a uniform average excess of the n_+ pairs with respect to the n_- pairs. This is a feature of all the investigated genomes. We have shown [2], [11] that recombination conserves this regularity, while local random mutations, such as uncorrelated SNPs tend to destroy it. The non-regularity visible in the interval 15,172–15,376 Kbp in Fig. 1 is caused by an error. Better curated ("phase 3") releases have a better linearity, as can seen in Fig.2 that gives θ_c (N) and θ_u (P) for chromosome 1 of *H. sapiens* [16].

The NuGSs of archaea and bacteria show even more striking regularities, with typical features and parameters that define a "physiognomy' of each genome. Figure 3 gives the N and P signals for the complete genome of *E. Coli* K12 [16]. The nucleotide pair imbalance P remains linear, like for eukaryotes, but the slope is nega-

tive. The significant change appears for the nucleotide imbalance *N*, which is no longer zero, but piece-wise linear. The separation points have a biological meaning: the minimum of *N* at 3,923,225 bp corresponds to the origin of genome replication, whereas the maximum at 1,550,413 bp corresponds to the terminus of replication.

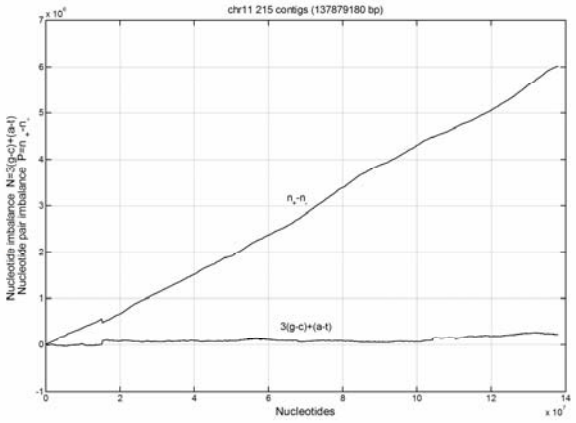


Fig. 1. *N* and *P* along the concatenated contigs of *Homo sapiens* chromosome 11 (length 137,879,180 bp) [16].

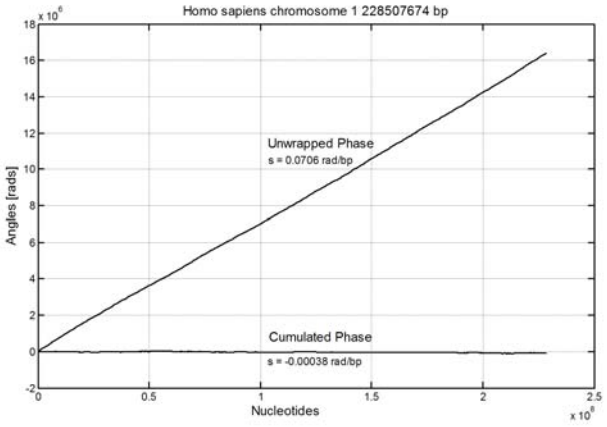


Fig. 2. θ_c and θ_u for phase-3 data on *H. sapiens* chromosome 1 (length 228,507,674 bp) [16].

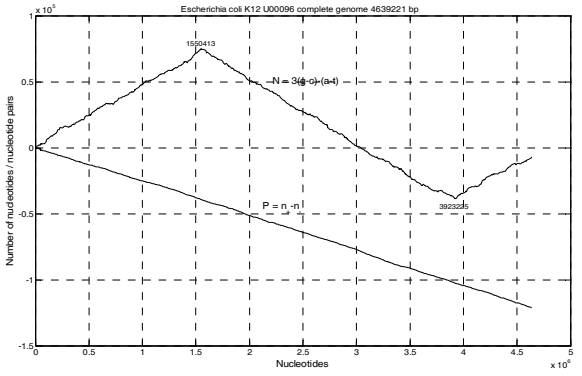


Fig. 3. *N* and *P* for *Escherichia Coli* strain K12 (length 4,639,221 bp) [16].

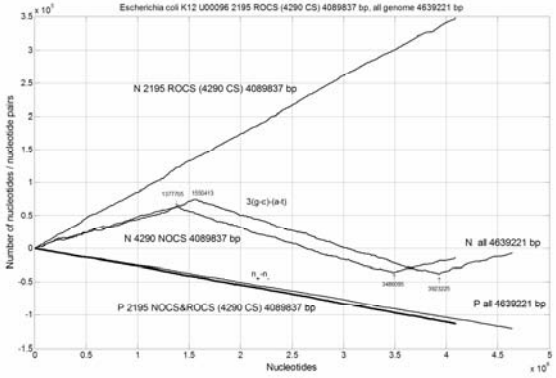


Fig. 4. *N* and *P* for *E. Coli* K12, before (4,639,221 bp) and after (4,089,837 bp) re-orienting its 2195 exons [16].

Figure 4 contains the NuGSs of *E.coli*, before and after re-orienting all its exons in the same (positive) direction:

- *N* and *P* signals for whole genome (length 4,639,221 bp, also shown in Fig. 3);
- *N* and *P* signals for the concatenated 4,290 exons (length 4,089,837 bp), after eliminating the non-coding segments, but maintaining the initial orientation of the exons in the genome.
- *N* and *P* signals for the concatenated exons, after re-orienting the 2195 exons that where in the negative direction in the genome given in GenBank [16].

The most remarkable result is the approximately linear shape of *N* after the re-orientation of the exons, suggesting a highly regular ancestral genomic structure. The current piece-wise linear *N* has evolved to allow only specific intermolecular interactions in processes such as replication, transcription and crossover [14]. *P* is invariant

because of the conservation of the n_+ and n_- numbers of nucleotide pairs in (4) when reversing a section of a DNA double helix, while simultaneously switching its strands, as occurs during crossover or re-orientation ([2], [11]).

4. DNA PREDICTION

A natural question related to processes such as DNA replication, DNA transcription to mRNA, and mRNA translation into polypeptide chains is whether there is enough redundancy in the genomic information to allow error correction. Despite the involvement of clearly distinct systems in these processes, they are functionally similar. We simulated the behavior of the human ribosome when translating mRNA strings. The current description of the ribosomal machine holds that a human ribosome, while processing a nucleotide, uses the information about the previous 35 nucleotides, but also about the next 3 nucleotides. In order to avoid obvious contradictions and conceptual difficulties, the model we adopted for checking the ability to predict nucleotide sequences has used only the information about the previous nucleotides.

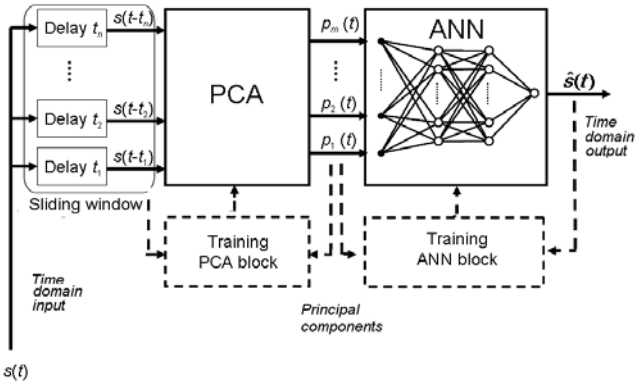


Fig. 5. Architecture of a prediction system using an ANN – as a learning unit, and a PCA block – to reduce the size and complexity of the ANN.

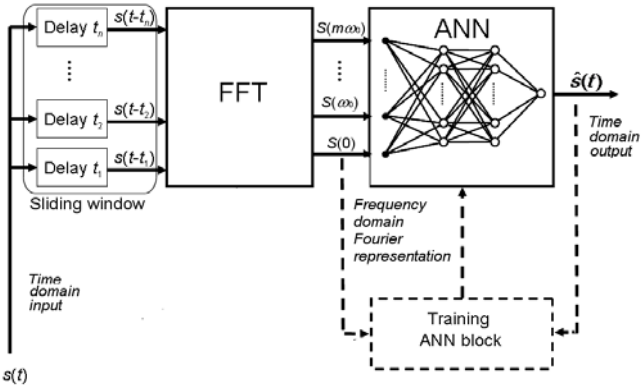


Fig. 7. Simplified architecture of a prediction system using an ANN and a predetermined FFT block.

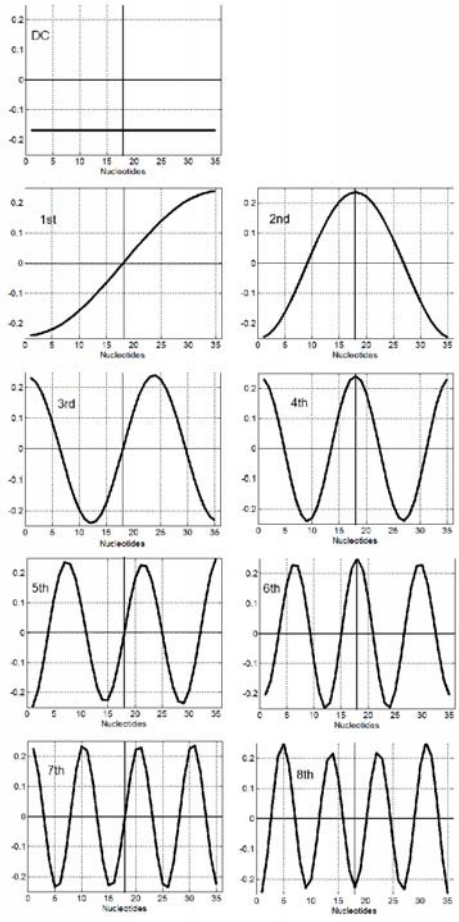


Fig. 6. Convolution functions (rows of the PCA block matrix in Fig.5) that allow to rebuild the sets of delayed signals with better than 10^{-5} accuracy.

Figure 5 shows the architecture of such a prediction system, based on a feed-forward artificial neural network (ANN) [18]-[19], similar to the classic *time series prediction*. Successive samples of the sequence $s(t)$ to be analyzed and predicted are applied at the input of the system, which is trained to produce an estimate of the current nucleotide. A principal component analysis (PCA) block is used to reduce the size and the complexity of the ANN, by replacing the samples of the input signal with its principal components.

The most interesting result of our experiments is that, after the training on NuGSs or other sequences satisfying certain mild statistical regularities, the rows of the resulting matrix $\boldsymbol{P}$ of the PCA block are closely related to the digital Fourier transform (DFT) operator, as shown in Fig. 6. It has been found that the set of $n = 35$ samples in the sliding window of the prediction system can be rebuilt from only $m = 9$ principal components, with a loss in the variance of the data set below 10^{-5}. The rows of the 9×35 matrix $\boldsymbol{P}$ of the trained PCA block are represented in Fig. 6. The resemblance to the DC and the first eight harmonics of the Fourier transform elements of a DFT operator is very good. This allows to reduce the complexity of the prediction system by replacing the PCA bloc in Fig. 5 with a predetermined FFT block, as shown in Fig. 7. Consequently, a training step is avoided and the prediction system becomes both simpler and more efficient. The architecture in Fig. 7 can be applied to increase the performance of almost any prediction system. and a principal component analysis (PCA) block.

5. CONCLUSIONS

The study of nucleotide sequences by using NuGSs reveals regularities in the structure of DNA and RNA molecules. This approach has been applied to describe the structure of nucleotide sequences, both in the current state and in a putative ancestral state, from which they have evolved.

The NuGSs method also provides a natural and efficient framework for the prediction of nucleotide sequences. The approach can help to locate areas in nucleotide molecules that have distinct statistical features and specific functionalities.

REFERENCES

[1] P. D. Cristea. Conversion of Nitrogenous Base Sequences into Genomic Signals. Journal of Cellular and Molecular Medicine, 6(2), pp. 279–303, 2002.

[2] P. D. Cristea. Representation and analysis of DNA sequences. In: E. Daugherty, I. Shmulevich, J. Chen, and Z.J. Wang, eds. Genomic Signal Processing and Statistics. Eurasip Book Series on Signal Processing and Communications, Hindawi Publ. Corp., pp. 15–65, 2005.

[3] P. D. Cristea, Rodica Tuduce, J. Cornelis, R. Deklerck, I. Nastac, M. Andrei. Signal Representation and Processing of Nucleotide Sequences, Proc. of the 7th IEEE Intl. Conf. on Bioinformatics and Bioengineering (IEEE BIBE 2007), pp.1214-1219, Harvard Medical School, Boston, USA, October 14-17, 2007.

[4] P. D. Cristea. Genetic signal representation and analysis, Proc. of Society of Photo-Optical Instrumentation Engineers (SPIE), 4623, pp. 77-84, 2002.

[5] J. A. Berger, S. K. Mitra, and J. Astola. Power spectrum analysis for DNA sequences, Proc. of the 7th Intl. Symp. on Signal Processing and its Applications, 2, pp. 29-32, 2003.

[6] P. Bernaola-Galvan, P. Carpena, R. Roman-Roldanet, J. L. Oliver. Study of statistical correlations in DNA sequences. Gene, 300, pp. 105-115, 2002.

[7] N. Chakravarthy, A. Spanias, L. D. Lasemidis, and K. Tsakalis. Autoregressive modeling and feature analysis of DNA sequences. EURASIP Journal of Genomic Signal Processing, 1, pp. 13-28, 2004.

[8] M. Akhtar, J. Epps, and E. Ambikairajah. On DNA numerical representations for period-3 based exon prediction, Proc. of IEEE Workshop on Genomic Signal Processing and Statistics (GENSIPS) 2007, pp. 1-4.

[9] I. Cosic. Macromolecular bioactivity: Is it Resonant Interaction between Molecules? - Theory and Applications. IEEE Trans. on BME, 41, pp. 1101-1114, 1994.

[10] P. D. Cristea. Large Scale Features in DNA Genomic Signals. Signal Processing [Special Issue on Genomic Signal Processing], Elsevier 2003, 83, pp. 871–888.

[11] P. D. Cristea. Genomic Signals of Re-Oriented ORFs. Eurasip – Journal on Applied Signal Processing [Special Issue on Genomic Signal Processing] 2004, 1: 132–137.

[12] P. D. Cristea. Genomic Signal Analysis of Pathogen Variability, Progress in Biomedical Optics and Imaging, Proc. of SPIE, 6088: P1-P12, 2006.

[13] P. D. Cristea. Genomic Signal Analysis of Mycobacterium tuberculosis. Progress in Biomedical Optics and Imaging, Proceedings of SPIE, 6447: C-1 – C8, 2007.

[14] S. T. Cole, R. Brosch, J. Parkhill, et al. Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. Nature, 393(6685), pp. 537-544, 1998.

[15] P. D. Cristea, D. Otelea, Rodica Tuduce. Genomic Signal analysis of HIV variability, Proc. of SPIE BIOS 2005, 6(14): pp. 362-372, 2005.

[16] GeneBank, NIH - National Institutes of Health, National Centre for Biotechnology Information, National Library of Medicine, (NCBI/GenBank), *http://www.ncbi.nlm.nih.gov/*, 2009

[17] E. Chargaff. Structure and function of nucleic acids as cell constituents. Federal Proceedings, 10, pp. 654–659, 1951.

[18] I. Năstac, P. D. Cristea. Neuro-Adaptive Forecasting for Nonstationary Sequences. SOFA 2005 –- IEEE Intl. Workshop on Soft Computing Applications, Szeged – Hungary & Arad – Romania, pp.179-186, 2005.

[19] P. D. Cristea, R. Deklerck, J. Cornelis, Rodica Tuduce, I. Năstac and M. Andrei. ANN Prediction of Nucleotide Sequences Link of Principal Component Analysis to Fourier Transform. IWSSIP 2007 – 14th Intl. Conf. on Systems, Signals and Image Processing, Maribor, Slovenia, 2007: 73-78.

FAULT DIAGNOSIS WITH OUTPUT OPTIMIZATION PROCEDURE

Alexandar Kamenov Ichtev

Faculty of Automatics, Technical University of Sofia, boulevard Kliment Ohridski 8, Sofia 1000, Bulgaria, phone: (+359 2) 965 2420, e-mail: ichtev@tu-sofia.bg

Abstract. *In this paper a method for fault diagnosis is presented. It is based on the optimization procedure and information obtained from bank of models. This is iterative method. It is capable of detecting single, as well as multiple, subsequent as well as simultaneously developing, modeled ones as well as partial faults in the system. The alternative optimization with output optimization is also disused. Experiments with a benchmark example – long bridge is carried out for demonstration purposes. The obtained results are discussed.*

Keywords: Fault diagnosis, optimization procedure, bridge benchmark

1. INTRODUCTION

In the last decades, there is increasing demand on performance for systems working in different environments. In order to satisfy this demand, more and more sophisticated systems with a larger number of sensors, actuators and other components are being built. As a result, the probability of a fault is increasing. On the other hand there are increasing safety demands. In order to satisfy those demands for automated systems reliable methods for fault detection and isolation are required.

This article proposes a method that continuously monitors the outputs of a plant. The objective is to detect fault as soon as possible after its occurrence as well as to determine its exact location and size. The results from fault detection might be used for fault tolerant control systems. The knowledge of the exact fault location will also be beneficial for rapid repair of the plant. This allows extension of the time between scheduled maintenances, which significantly reduces operational costs of the plant.

2. MULTIPLE MODEL APPROACH

The investigated systems are subject to abrupt as well as gradual developing faults. One way of describing such systems is by modeling them as hybrid dynamic system, whose state may jump as well as vary continuously. The jumps are used to model random abrupt faults. The dynamics between jumps is used to describe the investigated system, in faults free working regime or for gradually developing faults.

The model of the hybrid system is represented with

$$x(k+1) = \mathbf{F}(m(k+1))\mathbf{x}(k) + \mathbf{G}(m(k+1))\mathbf{u}(k) + \mathbf{T}(m(k+1))\boldsymbol{\eta}(k)$$

$$\mathbf{y}(k) = \mathbf{H}(m(k+1))\mathbf{x}(k) + \boldsymbol{\upsilon}(k)$$

where $m(k+1)$ is the model used in moment $k+1$.

It is assumed that the known model adequately describes the plant as well as that that jumps are described as first order Markov chain with transition probabilities from one model to another:

$$P\left(m_i(k+1)/m_j(k)\right)=\pi_{i,j}(k) \quad \text{with} \quad \sum_j \pi_{i,j}(k)=1$$

The transition probability matrix (π) is from great importance for the proper operation of the interactive multiple model (MM) algorithms [3]. However, the only way to determine this matrix is with the trial and error method.

The MM approach assumes that a set of N models can be set, which can approximate the hybrid system with the following N pairs of equations

$$\mathbf{x}(k+1)=\mathbf{F}_j\mathbf{x}(k)+\mathbf{G}_j\mathbf{u}(k)+\mathbf{T}_j\boldsymbol{\eta}(k)$$

$$\mathbf{y}(k)=\mathbf{H}_j\mathbf{x}(k)+\boldsymbol{\upsilon}(k) \qquad \text{for } j=1,2,\cdots,N$$

The set of all models will be referred to a model set M.

Each pair of equations corresponds to an operating mode of the system or presence of fault. If a mathematical description of the plant is known then the models with faults can be obtained directly from it. For example, *60%* partial fault of the sensor can be modeled by multiplying the corresponding row of matrix **H** with *0.4*. Total failure can be modeled if it is multiplied with *0*.

3. FAULT DETECTION AND ISOLATION

In the literature there are two types of methods for fault detection and isolation. The first one relies on information obtained form redundant sensors, actuators and component. This is known as hardware redundancy. The second type is based on additional information from computational methods. This is referred to as analytical redundancy [1], [2]. The first type requires use of redundant equipment. An important example for this approach is aerospace industry, where there are three (sometimes four) systems for the implementation of the same task. This leads to increased costs for implementation, maintenance and operation. Therefore, in this article the latter approach is adopted.

In the model based approach [1], [2], [3], [4], [5] for fault detection and isolation there are two main stages: (i) determination of the plant's models (can be performed off-line) and (ii) on-line residuals generation and evaluation. The residual is analytically calculated value that represents the difference between the measured values form the plant and analytically calculated ones. The residuals are determined from a model of the plant and actually measured values. They represent the difference between the current (real) situation with our expectation for the system's behavior.

4. PROBLEM FORMULATION

The problem for fault detection and isolation is usually solved by performing a hypothesis test in order to determine the current model [1], [2]. This is done by choosing one model from the model set M . Afterwards, it is assumed that this is the thru one. The hard decision can be serious drawback when a system is working with partial faults and fault tolerant control is used [4], [5]. One way to overcome this problem is to extend MM to linear differential inclusions. Let us consider the MM set M . Linear differential inclusions are defined as a set of all plants that are convex combination of the model in M :

$$M = \sum_{i=1}^{N} \mu_i M_i \quad \text{with} \quad \sum_{i=1}^{N} \mu_i = 1$$

The used criterion for convex combination is a weighted sum from the outputs:

$$\mathbf{y}(k) = \sum_{i=1}^{N} \mathbf{y}_i(k) \mu_i$$

where $\boldsymbol{\mu}$ is vector containing probabilities for each of the models. Each element of this vector represents the probability that the particular model is the true one at the given time instant. This vector will be referred to mode probability vector.

If the set M is known in advance, i.e. all possible faults are determined, then the fault detection and isolation task boils down to calculation of the mode probability vector for each time instant. If the probability that corresponds to the nominal model of the plant is close enough to one and the probabilities for the rest models are small enough (close to zero), then the system is fault free. Each significant difference from this situation indicates that there is a fault in the system, which is fault detection. The calculation of the mode probabilities also provides additional information. By tracking the most likely model it can be established which fault scenario is present in the system, which is fault isolation.

5. ALGORITHM

The base of the MM algorithms consists in the use of separate descriptions for each mode of the sustem. A model is set to represent this situation. Each model is evaluated for each time instant. The interactive MM algorithm consists from four main parts:

1) Mode dependant reinitialization. In this part the iteratively of the algorithm is given. It is from great importance, since in the no iterative algorithms it is assumed that the system mode did not changes, while in the fault detection this mode change is essential. The initial values for all models are chosen based on the estimated model form the previous time instant. This assessment is probabilistic sum from previous

estimates (see step 4 from this algorithm). This is how the interaction between the models is achieved. In this algorithm there is no need of knowing in advance the π matrix, in contrast to the proposed in [3] algorithm.

2) Model evaluation: All models are evaluated for each time instant and information regarding their outputs is stored.

3) Mode probability calculation and fault detection and isolation. The most significant difference between this approach and the one described in [3] is precisely in this step. Here it is proposed to be used an optimization procedure, as the presented results below are used with quadratic programming. Similar results are obtained by using other optimization algorithms that can solve the problem with limits.

During the optimization two limits are set in respect to the mode probabilities of the models.

$$\sum_{i=1}^{N} \mu_i(k) = 1 \quad \text{and} \quad 0 \le \mu_i(k) \le 1 \quad \text{for} \quad \text{i=1,2,}\cdots\text{,N}$$

The minimization at each step is

$$\min_{\boldsymbol{\mu}} \left[(\mathbf{Y}_m - \mathbf{Y}_s\boldsymbol{\mu})^T (\mathbf{Y}_m - \mathbf{Y}_s\boldsymbol{\mu}) \right] = \min_{\boldsymbol{\mu}} \left[\boldsymbol{\mu}^T \mathbf{Y}_s^T \mathbf{Y}_s \boldsymbol{\mu} - 2\mathbf{Y}_s^T \mathbf{Y}_m \boldsymbol{\mu}) \right]$$

where $\mathbf{Y}_m$ is the vector containing the measured outputs from the system $\mathbf{Y}_s$ is a vector containing the estimated output for each of the N models and $\boldsymbol{\mu}$ I s the vector containing the probabilities for each of the models.

$$\mathbf{Y}_m(k) = \begin{bmatrix} y_1(k) & y_2(k) & \cdots & y_r(k) \end{bmatrix}^T$$

$$\boldsymbol{\mu}_m(k) = \begin{bmatrix} \mu_1(k) & \mu_2(k) & \cdots & \mu_n(k) \end{bmatrix}^T$$

$$\mathbf{Y}_s(k) = \begin{bmatrix} y_{11}(k) & y_{12}(k) & \cdots & y_{1N}(k) \\ y_{21}(k) & y_{22}(k) & \cdots & y_{2N}(k) \\ \vdots & \vdots & \ddots & \vdots \\ y_{r1}(k) & y_{r2}(k) & \cdots & y_{rN}(k) \end{bmatrix}$$

The optimization is perfumed over the outputs of the system instead over the states. This aims to accelerate the algorithm, by simulating the output instead of perfuming the state estimation. It must be mentioned that, for the fault detection and isolation of component faults, it is better to use the optimization based on the states [3], while for fault detection of the actuators and especially sensor fault it is more convenient to use optimization over the outputs.

In some cases, especially when there is abrupt change in the operation point and when the system is subjected to strong noise, it is possible that some discrepancies

between the actual situation and calculated mode probabilities can occur. In such cases the optimization should be performed over a time window instead of just one time instant. This will slow down the fault detection procedure, but it will overcome the mentioned problems.

4) Estimate combination. In this last step of the algorithm the model and the initial conditions for the next time instant are determined. They are calculated as weighted sums from the states of all models. The probabilities from the mode probability vector are used.

6. EXAMPLE

As an illustrative example in this article a long bridge is considered. This plant would be good illustrative example, because for control purpose of such plant it is necessary to use a large number of sensors and actuators.

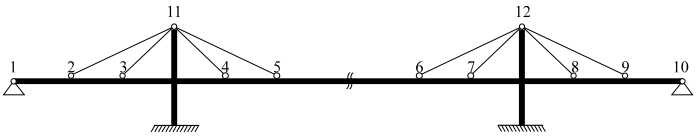


Fig. 1. Principle diagram of the plant

Mathematical models. As mathematical model of the plant a second order matrix differential equation is assumed:

$$\mathbf{M}\frac{d^2\mathbf{Y}}{dt^2}+\mathbf{C}\frac{d\mathbf{Y}}{dt}+\mathbf{K}\mathbf{Y}=\mathbf{D}\mathbf{V}+\mathbf{B}\mathbf{U}$$

where **Y** is n-dimensional vector of displacements in the basic points of the plant, **V** is a vector of external forces, **U** is a vector of the controls, and **M**, **C**, **K**, **D** and **B** are respectively matrices of masses, damping, stiffness inputs and controls. By introducing a *2n*-dimensional state vector $\mathbf{X}^T=\begin{bmatrix}\mathbf{Y}^T & \dot{\mathbf{Y}}^T\end{bmatrix}$, the model yields the form

$$\dot{\mathbf{X}}=\begin{bmatrix}\mathbf{0} & \mathbf{I}\\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{C}\end{bmatrix}\mathbf{X}+\begin{bmatrix}\mathbf{0}\\ \mathbf{M}^{-1}\mathbf{D}\end{bmatrix}\mathbf{V}+\begin{bmatrix}\mathbf{0}\\ \mathbf{M}^{-1}\mathbf{B}\end{bmatrix}\mathbf{U}$$

Experimental results. The nominal matrices in the plants model are

$$\mathbf{M}=diag\begin{bmatrix}m_1 & m_2 & m_3 & m_4 & m_5\end{bmatrix},$$

$$\mathbf{C}=\begin{bmatrix}c_{1,1} & -c_{1,2} & 0 & 0 & 0\\ -c_{1,2} & c_{2,2} & -c_{2,3} & 0 & 0\\ 0 & -c_{2,3} & c_{3,3} & -c_{3,4} & 0\\ 0 & 0 & -c_{3,4} & c_{4,4} & -c_{4,5}\\ 0 & 0 & 0 & -c_{4,5} & c_{5,5}\end{bmatrix},\ \mathbf{K}=\begin{bmatrix}k_{1,1} & -k_{1,2} & 0 & 0 & 0\\ -k_{1,2} & k_{2,2} & -k_{2,3} & 0 & 0\\ 0 & -k_{2,3} & k_{3,3} & -k_{3,4} & 0\\ 0 & 0 & -k_{3,4} & k_{4,4} & -k_{4,5}\\ 0 & 0 & 0 & -k_{4,5} & k_{5,5}\end{bmatrix}$$

$$D = \begin{bmatrix} d_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & d_2 \end{bmatrix}^T, \ \mathbf{B} = diag\begin{bmatrix} b_1 & b_2 & b_3 & b_4 & b_5 \end{bmatrix}$$

Their parameters have the following numerical values:

$$m_1 = m_2 = m_3 = m_4 = m_5 = 1,$$
$$k_{i,i} = 2, i = 1,2,3..., \ k_{i,i+1} = 1, i = 1,2,3..., \ c_{i,j} = 0.01k_{i,j},$$
$$d_1 = d_2 = b_1 = b_2 = b_3 = b_4 = b_5 = 1$$

The experiment involves sensor fault simulations. There are three total faults: in sensors *1*, *2* and *5*, and three partial faults, which are 40% error in the measurements. During the first *30* times samples (*3* seconds) of the experiment the bridge and sensors are with no fault. In the next *30* time samples the system is under the influence of fault 1. Further, alternately fault free and next fault modes with the same frequency are simulated. In the last two faults scenarios are simulated non modeled partial failures. Thus, the method's ability to cope with failures that are convex combination of modeling one is verified. This is achieved without introducing additional models in the model set.

As can be seen from Figure 2 the algorithm correctly detects the true model. Excluding the first interval, which is transitory, it can be seen that the probability of the correct model is more than 90% (0,9). If we take a limit of 50% (or even 70%), i.e. be considered the most likely model as a true one, it can be concluded that the algorithm works perfectly, even in the first (transitional) interval.

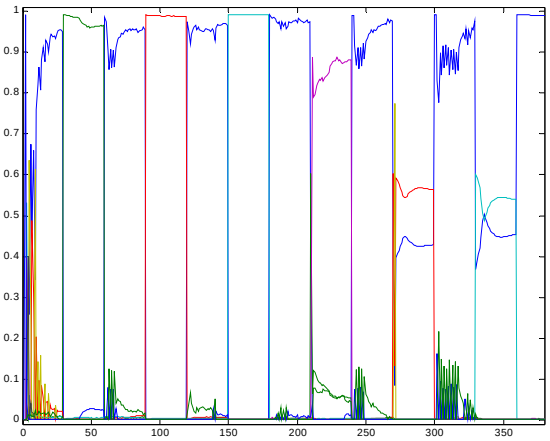


Fig. 2. Mode probabilities for each of the models

From great interest are the last two fault intervals. It can be seen that the partial non modeled faults can be successfully detected, isolated and identified as well. This shows that the proposed method can evaluate each partial failure without adding additional models. This is the main advantage of this method with respect to the other

methods in the literature. It is essential for control in the presence of partial faults. Faults tolerant systems are discussed in [4] [5].

7. CONCLUSION

This article proposes a new interactive method for fault detection and isolation. The output optimization allows faster evaluation then similar ones based on state optimization. Its main advantage is that it can detect and isolate partial faults, without introduction of additional faults. This advantage is demonstrated by the long bridge benchmark example. The methods advantages are important for fault tolerant control.

8. ACKNOWLEDGEMENT

The research results presented in this paper are partially funded by Technical University – Sofia research project No. 102ni046-08.

REFERENCES

[1] Gertler, Janos, Fault Detection and Diagnosis in engineering systems, Marcel Dekker, Inc., USA, 1998

[2] R. Isermann Fault-Diagnosis Systems, Springer 2006.

[3] Ichtev. A. “Fault tolerant control with the controlled bridge structures”. VSU’2002, v.1, pp. I – 137-142, 2002 (In Bulgarian)

[4] A. Ichtev, J. Hellendoorn, R. Babuska, S. Mollov, Fault Tolerant Model Based Predictive Control Using Multiple Takagi-Sugeno Fuzzy Models, FUZZ-IEEE 2002, Hawaii, USA, Volume: 1, 346 -351

[5] A. Ichtev, P. Petkov, Fault Tolerant Control by Compensating Partial Faults, 6th WSEAS conference, Crete, Greece, July 7-14, 2002, 259-264

SHORT TERM LOAD FORECASTING IN INTERCONNECTED POWER SYSTEM USING ANN: THE CONFIDENCE INTERVAL PROBLEM

Nikos E. Mastorakis, George J. Tsekouras

Department of Electrical & Computer Science, Hellenic Naval Academy, Terma Hatzikyriakou, ZIP:18539, Piraeus, GREECE, phone:302104581370, 30210458651,
e-mail: mastor@wseas.org, tsekouras_george_j@yahoo.gr

***Abstract**: One of the major issues during the application of Artificial Neural Networks for power system load prediction is the determination of the confidence intervals, because the respective estimation can not be applied directly, unlike to the classical forecasting methods. The most commonly used methods are: (a) the error output, (b) the re-sampling, (c) the multi-linear regression adapted to ANN. The scopus of this paper is (1) to present these basic methods, (2) to compare them proving the superiority of the re-sampling method and (3) to apply the last one for the short term forecasting of the next day load in interconnected Greek power system using as ANN training algorithm the scaled conjugate gradient one, which is properly optimized based on the evaluation set. Finally the next day load demand of the test set is estimated using the best ANN structure and the confidence intervals for the training, evaluation and test sets are compared.*

Keywords: Artificial neural networks (ANN), confidence interval, short-term load forecasting (STLF), power system

1. INTRODUCTION

In a deregulated electricity market, the load demand has to be predicted with the highest possible precision in different time periods: very short-term (for the next few minutes), short-term (for the next few hours to a week), midterm (for the next few weeks to few months) and long-term (for the next few months to years). Especially, short-term load forecasting is very significant problem, because its accuracy affects other operational issues of power systems, such as unit commitment, scheduling of spinning reserve, available transfer capability, system stability, application of load demand programs, etc. Higher reliability and lower operational costs for power systems are achieved by precise load forecasting. During last 15 years several forecasting methods have been implemented with different levels of success, such as ARMAX models [1], regression [2], ANNs [3], fuzzy logic [4], expert systems etc. Especially, in Greece, ANNs have been used successfully either for the interconnected power system [5-6], or autonomous big islands [7-9]. Some techniques belong to classical ANNs [5-7] or specialized ones [8] or they are based on ANNs combined with fuzzy logic algorithms [9].

All modern models (such as ANNs, fuzzy, hybrid ones) lead to a prediction value which is rarely equal to the real one. The variance between the prediction and the real value should be quantified in advance. In classical models, such as regression models,

this is expressed by the confidence interval based on analytical calculations. In case of ANNs, Silva et al [10] presented three commonly used methods, which are: (a) the error output, (b) the re-sampling, (c) the multi-linear regression adapted to ANN. For fuzzy logic based methods, Elias et al have calculated the standard deviation analytically working out at the same time the respective problem [11].

In this paper the theoretical determination of the confidence intervals is presented for the application of Artificial Neural Networks to power system load prediction based on the aforementioned three commonly used methods. The superiority of the re-sampling method will be proved theoretically and it will be applied for the short-term forecasting of the next day load in interconnected Greek power system using as ANN training algorithm the scaled conjugate gradient one. This training method is properly optimized based on the evaluation set, as it has already presented in [12]. Meanwhile, the problem due to the mismatch between the confidence interval determined by the evaluation set and the one calculated by the test set is also arisen and is commented in results.

2. ANN SHORT-TERM LOAD FORECASTING METHODOLOGY

Before the mathematical base of confidence interval estimation is analyzed, the ANN methodology for short-term load forecasting is presented briefly (analytical description in [12]). Its flow chart is shown in Figure 1.

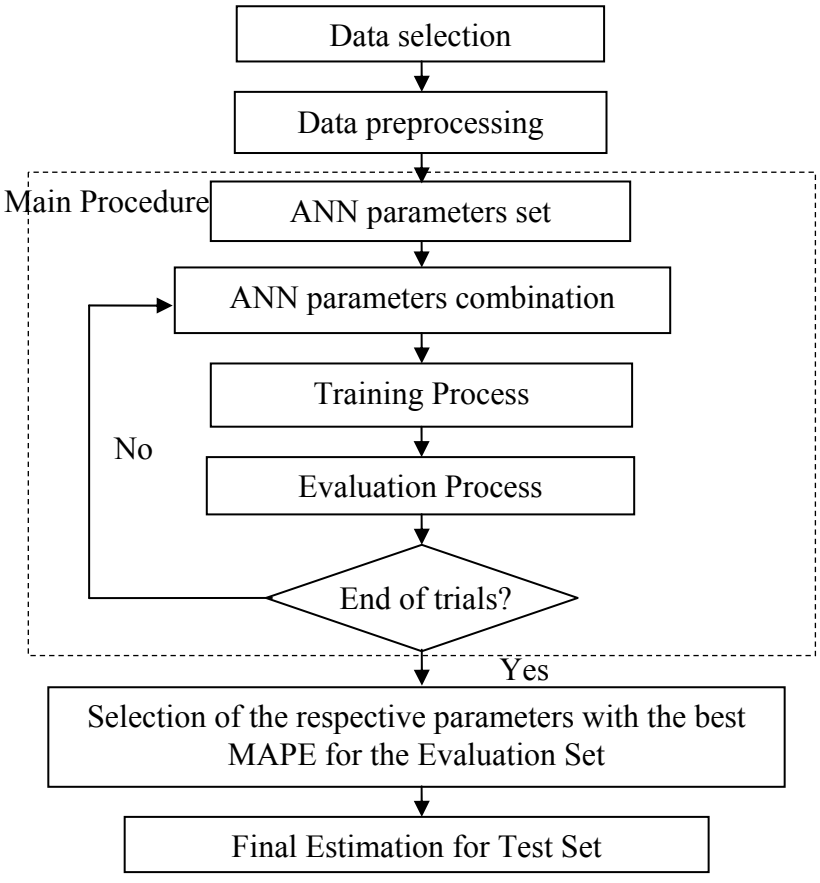


Fig. 1. Flowchart of the ANN methodology for the proper selection of ANN parameters for short-term load forecasting

The basic steps are the following:

(a) *Data selection*: In this step the input variables for load forecasting of *d*-th day are formed according to Kiartzis et al [5], Tsekouras et al [6, 12] including: (1) the hourly actual loads of the two previous days, (2) the maximum mean temperature per three hours and the minimum mean temperature per three hours for Athens and for Thessalonica, for the current and the previous day, (3) the temperature difference between the maximum mean temperature per three hours of the current day and the respective one of the last day for Athens and Thessalonica, (4) the temperature dispersion from comfortable living conditions temperature for Athens and for Thessalonica, for the current and the previous day, (5) seven digit binary numbers, which express the kind of the week day, (6) two sinusoidal functions $\cos(2\pi d/T)$ and $\sin(2\pi d/T)$), which express the seasonal behavior of the current day, where *T* is the number of the days of the current year. So, each input vector comprises 71 elements, while the output variables are the 24 hourly actual load demand of the current day.

(b) *Data preprocessing*: Data are examined, in order to modify or delete the values that are obviously wrong (noise suppression). In order to express the non linearity of the problem, non linear activation functions are preferably used, such as sigmoid activation functions. In order to avoid saturation problems, the input and the output values are normalized.

(c) *Main procedure*: The ANN is trained using the scaled conjugate gradient algorithm (SCGA), whose basic steps have been presented in [12-13]. The basic disadvantage of the SCGA algorithm is that its calculation complexity per iteration is twofold of the basic steepest descent method. Its basic advantage is that the error function decreases monotonically. The respective parameters of the neural network are selected through a set of trials. Specifically for each ANN parameter (such as the neurons of the hidden layer, the type of the activation functions (hyperbolic tangent, logistic, linear), the parameters of the activation functions, the maximum number of epochs, the SCGA algorithm parameters) the training algorithm is separately executed for the respective range of values (i.e. 20 to 70 neurons with step 1) based on the error function (sum of the square of errors for all neurons per epoch) for training set and the regions with satisfactory results (minimum Mean Absolute Percentage Error (*MAPE)* for evaluation set) are identified. Following, the training algorithm is repeatedly executed, while all parameters are simultaneously adjusted to their respective regions, so that the combination with the smallest *MAPE* for the evaluation set is selected. It is also noted that:

- The ANN is solved by using one hidden layer and calibrating the number of neurons properly according to Kolmogorov's theorem.
- During the training process for each ANN three stopping criterions are used: the weights stabilization, the respective error function not to be decreased or the excitation of the maximum number of epochs [6].

(d) *Final estimation for the test set*: The actual load demand (in MW) for the days of the test set is finally estimated by using the respective ANN parameters.

The best result for the *MAPE* index of the evaluation set is 1.487% for the case of Greek intercontinental power system with training years 1997-1999 (the training and the evaluation sets consist of the 90% and 10% of the normal days (no holidays) respectively), while the respective test set consists of the normal days of the year 2000. The *MAPE* indexes for the training and the test set are 1.294% and 1.781% respectively. These are obtained for an ANN with 52 neurons in the hidden layer using hyperbolic tangent activation function in both layers (tanh(0.5·x) for the hidden layer, tanh(0.25·x) for the output layer, where x is the respective sum of the properly weighted inputs of the neuron) (see [12, paragraph 4] for more details).

3. THE ESTIMATION OF THE CONFIDENCE INTERVAL FOR ANN STLF MODELS

The estimation of the confidence intervals for ANN models is not applied directly, unlike to the classical models. Three techniques have been mentioned in [10].

3.1. The error output technique

The ANN model has finally two outputs, the first one is the forecasting mean value and the second one the respective estimated absolute percentage error. After the training process a larger confidence interval is determined by multiplying the initial one by a proper factor, in order to obtain the required confidence degree, which can not be predefined.

3.2. The re-sampling technique

The prediction and the respective error are calculated for each set and for all available *m* input vectors. These errors are sorted in ascending order considering the signs and the cumulative sample distribution function of the prediction errors can be estimated by:

$$S_m(z) = \begin{cases} 0, & z < z_1 \\ r/m, & z_r \le z < z_{r+1} \\ 1, & z_m \le z \end{cases} \quad (1)$$

When *m* is large enough, $S_m(z)$ is a good approximation of the true cumulative probability distribution $F(z)$. The confidence interval is estimated by keeping the intermediate z_r and discarding the extreme values, according to the desired confidence degree. The intervals are computed in order to be symmetrical in probability (not necessarily symmetric in *z*). The number of cases to discard in each tail of the prediction error distribution is $n \cdot p$, where *p* is the probability in each tail. If $n \cdot p$ is a fractional number, the number of cases to discard in each tail is $\lfloor n \cdot p \rfloor$ for safeness reasons. If the cumulative probability distribution $F(Z_p)$ equals to *p*, then there is a probability *p*

that an error is less than or equal to Z_p , which indicates that Z_p is the lower confidence limit. Consequently, Z_{1-p} is the upper limit and there is a *(1-2p)* confidence interval for future errors.

3.3. The multi-linear regression adapted to ANN technique

This technique can be only applied, if linear activation functions are used at the output layer. In this case a multi-linear regression model can be implemented for each output neuron. The inputs of the regression model are taken from the outputs of the hidden neurons and the regression coefficients are taken as the connection weights of the output neuron. The computation of the confidence interval is accomplished through the estimation of the prediction error variance:

$$\sigma^2 = \sum_{i=1}^{m_1} \left(t_i - o_i\right)^2 \Big/ \left(m - p_c\right) \tag{2}$$

where p_c the number of coefficients of the regression model. The confidence interval for a point prediction τ can be obtained using the variance estimation (2), the ANN inputs and the desired confidence degree, which follows a *t*-Student's distribution with (m-p_c) degrees of freedom, as it is noted in [10].

3.4. The theoretical comparison of the techniques

The error output doubles the number of the ANN's outputs increasing the respective number of the ANN weights from 5016 to 6808 (if we assume that the best ANN is the same and only the number of the outputs changes) and the respective computational time by 84% approximately. The multi-linear regression adapted to ANN technique allows only the use of the linear activation function for the output layer, which deteriorates the *MAPE* results (from 1.509% to 1.539% for the substitution of the hyperbolic tangent from the linear activation function for the best ANN – (see [12 – Table 1])). So, the re-sampling technique is proposed as the most suitable for the confidence interval with high confidence degree, because it does not affect in computational time and the *MAPE* results.

4. APPLICATION OF THE RE-SAMPLING TECHNIQUE FOR SHORT-TERM LOAD FORECASTING IN GREEK INTERCONTINENTAL POWER BASED ON ANN STLF METHODOLOGY

After the selection of the best ANN (according to the respective methodology of §2) the 90% confidence interval is estimated using the re-sampling technique with the probability in each tail equal to 5%. In Fig. 2 the prediction errors of a typical summer day for Greek interconnected power system of the year 2000 (Thursday 8-6-

2000) are presented for the training, evaluation and test sets respectively, while in Fig. 3 the respective measured and estimated load values are presented with the 90% confidence intervals from the evaluation set and from the test set.

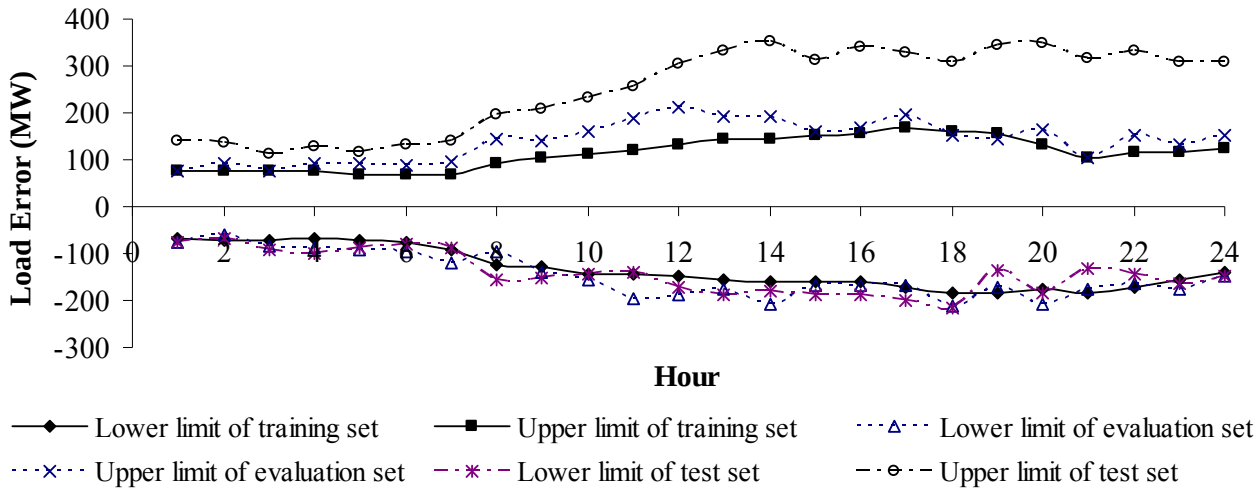


Fig. 2. 90% confidence interval limits with respect to the training, evaluation and test sets for the best ANN model for 8-6-2000 in Greek interconnected power system

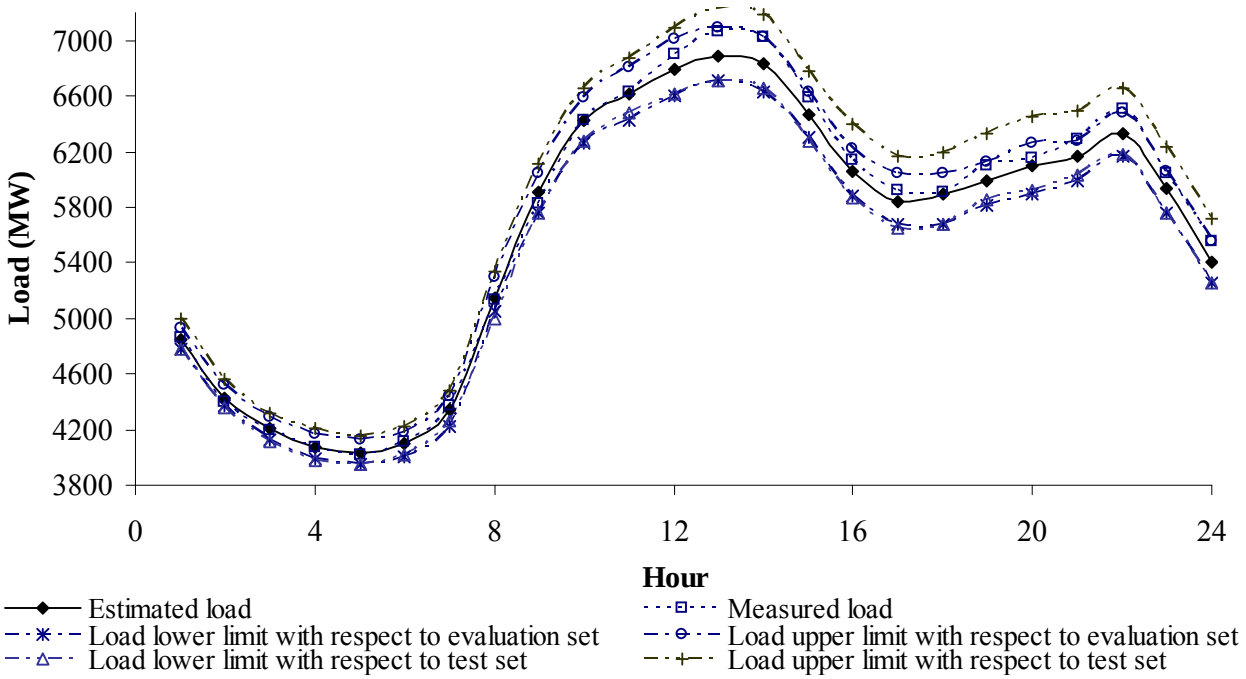


Fig. 3. Chronological active load curves of the measured load, the estimated load, the estimated load with the 5% lower limit with respect to evaluation set, the estimated load with the 5% upper limit with respect to evaluation set, the estimated load with the 5% lower limit with respect to test set, the estimated load with the 5% upper limit with respect to test set for the best ANN model for 8-6-2000 in Greek interconnected power system

From Fig. 2 it is observed that the lower limits of the confidence intervals for the three sets are quite similar. The ratio between the lower hourly error of the test set to the respective one of the evaluation set varies between 0.71 and 1.60, while the mean value is equal to 1.00. However, the upper limit of the confidence intervals for the test set is almost the double of the respective one of the evaluation set. The ratio between the upper hourly error of the test set to the respective one of the evaluation set varies between 1.22 and 3.02, while the mean value is equal to 1.78. From Fig. 3 it is observed that the confidence interval of the test set is broader than the respective one of the evaluation set. Similar behavior is presented for all days studied. In fact the limits of the test set are unknown in real applications, which means that they should be corrected, i.e. using the proper multiplying factor which is calculated by trial executions with historical data. The last issue is a problem that must be solved in the future.

5. CONCLUSIONS

One of the most significant problems in short-term power system load forecasting using ANN's methods is the determination of the confidence intervals. However, the respective estimation can not be applied directly, unlike to the classical forecasting methods. Here, three techniques have been presented: (a) the output-error, (b) the resampling, (c) the multi-linear regression adapted to ANN. The second one has been chosen, because it does not affect in computational time and *MAPE*, as it happens with other two techniques. However, another problem occurs as the confidence intervals between the training, evaluation and test sets differ for the same probability in tail. This has not been treated by Silva and Moulin [10] and certainly should be dealt with in future.

REFERENCES

[1] H.T. Yang, C.M. Huang, C.L. Huang, "Identification of ARMAX model for short term load forecasting: an evolutionary programming approach," IEEE Trans. on Power Systems, vol. 11, no. 1, 1996, pp. 403 – 408.

[2] T. Haida, S. Muto, "Regression based peak load forecasting using a transformation technique," IEEE Trans. on Power Systems, vol. 9, no. 4, 1994, pp. 1788 – 1794.

[3] H.S. Hippert, C.E. Pedreira, R.C. Souza, "Neural networks for short-term load forecasting: A review and evaluation," IEEE Trans. on Power Systems, vol. 16, no. 1, 2001, pp. 44-55.

[4] P.A.Mastorocostas, J.B.Theocharis, A.G. Bakirtzis, "Fuzzy modeling for short-term load forecasting using the orthogonal least squares method," IEEE Trans. on Power Systems, vol. 14, no. 1, 1999, pp. 29–36.

[5] A.G. Bakirtzis, V. Petridis, S.J. Kiartzis, M.C. Alexiadis, A.H. Maissis, "A neural network short term load forecasting model for the Greek power system," IEEE Trans. on Power Systems, vol. 11, no.2, 1996, pp. 858-863.

[6] G. J. Tsekouras, F.D. Kanellos, V.T. Kontargyri, C.D. Tsirekis, I.S. Karanasiou, Ch.N. Elias, A.D. Salis, N.E. Mastorakis, "A comparison of Artificial Neural Networks algorithms for short term load forecasting in Greek intercontinental power system," WSEAS International Conference on Circuits, Systems, Electronics, Control & Signal Processing, CSECS '08, Puerto De La Cruz, Canary Islands, Spain, 15-17 / 12 / 2008.

[7] S.J. Kiartzis, C.E. Zournas, J.M. Theocharis, A.G. Bakirtzis, V. Petridis, "Short term load forecasting in an autonomous power system using artificial neural networks", IEEE Trans. on Power Systems, Vol. 12, No.4, 1997, pp. 1591-1596.

[8] K. Kalaitzakis, G.S. Stavrakakis, E.M. Anagnostakis, "Short-term load forecasting based on artificial neural networks parallel implementation," Electric Power Systems Research, Vol. 63, 2002, pp. 185-196.

[9] A.G. Bakirtzis, J.B. Theocharis, S.J.Kiartzis, K.J.Satsios, "Short term load forecasting using fuzzy neural networks," IEEE Trans. on Power Systems, vol. 10, no. 3, 1995, pp. 1518-1524.

[10] A.P.A. Silva, L.S. Moulin, "Confidence intervals for neural network based short-term load forecasting", IEEE Trans. on Power Systems, vol. 15, no. 4, November 2000, pp.1191-1196.

[11] Ch. N. Elias, N. D. Hatziargyriou, "An annual midterm energy forecasting model using fuzzy logic", IEEE Trans. on Power Systems, vol. 24, no.1, January 2009, pp.469-478.

[12] G. J. Tsekouras, F.D. Kanellos, V.T. Kontargyri, C.D. Tsirekis, I.S. Karanasiou, Ch.N. Elias, A.D. Salis, N.E. Mastorakis, "Short term load forecasting in Greek intercontinental power system using ANN: A study for input variables," 10th WSEAS International Conference on Neural Networks, (NN '09), Prague, Czech Republic, March 23-25, 2009.

[13] M.F. Moller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks,* vol. 6, 1993, pp. 525-533.

METHOD OF DETERMINATION OF SHORT VIDEOIMPULSES PASSING PERIOD IN CODE SEQUENCES ON THE BASIS OF HEAVISIDE STEP FUNCTION APPROXIMATION

Volodymyr Lysechko, Yuliya Stepanenko

Ukrainian State Academy of Railway Transport – Kharkiv 61050, UKRAINE
e-mail lysechkov@ukr.net stepanenko7@live.ru

Abstract: *the task of determination of short video pulses passing periods in the sequences of video pulses with the improved cross-correlation properties with the use of rectangular function, Heaviside Step Function and their properties decides in the article. The use of the offered method of sequences forming allows to minimize interaction between signals in a time domain, and, as a result, substantially decrease intrasystem interference level. The offered mathematical apparatus allows to simplify forming of sequences ensembles from short video pulses with the improved cross-correlation properties.*

Keywords: video pulse, rectangular function, Heaviside Step Function , intrasystem interference

1. Problem formulation. The rise of possible number of active users in radio systems with the code – division access results in the rise of interference level, or intrasystem interferences, which appear due to the imperfect cross-correlation properties of signals, that are used. The decision of users number rise task in these systems is possible by the simple rising of interference threshold, but quality of communicztion here gets worse. Therefore the main directions of this task decision with saving of communication quality is signal forming with the improved cross-correlation properties and use of code with errors correction.

2. Analysis of literature. In modern technical literature of Ukrainian and foreign authors the following methods of fight with intrasystem interference are considered: frequency - division [3,4,5,6], time-division [5,6], spatial selection of signals, that uses aerials with the narrow diractional characteristics [4,5,6]. However, application of frequency and time division is limited by a bandwidth - time resources at a plenty of simultaneously working radio terminals [2,6], and application of aerials with the narrow directional characteristics results in complication of their construction and not always acceptable at construction of full-accessible radio networks [4,6]. In works [1,2] the method of short videoimpulses sequences ensembles forming is considered with the improved cross-correlation properties which meets a the intrasystem interference demand[3] levels for the systems of radio systems with the code – division access. But, the given method has not the detailed and clear mathematical ground.

3. Objective of the article is development of method of determination of short video pulses passing periods in the sequences of video pulses with the improved cross-correlation properties with the use of rectangular function, with the use of mathematical methods of analysis on the basis of rectangular function, *Heaviside Step Function* and their properties.

4. Basic material.

We will present the sequence of short videoimpulses in a kind [1, 3]

$$u_i(t) = \sum_{k=1}^{n_i} U_{k_i} \prod\left[t - \left(k \cdot Q_i - 1\right) \cdot \tau_u\right], \tag{1}$$

where $k = 1 \ldots, n_i$, – amount pulses in i - sequence, i=1,…,L;

$\prod[t - (k \cdot Q_i - 1) \cdot \tau_p]$ – rectangle function;

U_{ki} - amplitude of n – pulse of i – code sequence, equals [–1,1];

τ_p – duration of pulse;

$Q_i = T_i/\tau_u$ – on-off time ratio of i- pulse sequence;

T_i – period of pulses passing in i – sequence.

Function $\prod\left[t - \left(k \cdot Q_i - 1\right) \cdot \tau_p\right]$ in expression (1), has the appearance [1]

$$\prod\left[t - \left(k \cdot Q_i - 1\right) \cdot \tau_i\right] = \begin{cases} 1, \text{ if } (\text{kQi} - 1)\,\tau_p \\ 0, \text{ if } (\text{kQi} - 1)\,\tau_p \end{cases} \tag{2}$$

The cross-correlationt function of sequences is determined by expression [4]:

$$R_{ij}(\tau) = \frac{1}{2\sqrt{E_i E_j}} \int_{-T}^{T} u_i(t) u_j(t - \tau) dt, \tag{3}$$

where $u_i(t)$, $u_j(t)$ – i- and j- code sequences of video pulses, here $i \neq j$;

Ei and Ej – value of energies according to i- *and* j- sequences;

T – interval on which definite functions *of* $u_i(t)$ and $u_j(t)$.

Minimum similarity of two sequences [2], consists in that regardless of phase shift the coincidence is possible no more than on one pulse in every sequence. For implementation of this condition i and j sequences must not exceed the value [3]:

$$R_{ij}(\tau) \leq \frac{1}{n_i}, \tag{4}$$

where $n_i \geq n_j$ j – quantity of pulses in sequences.

If sequences have identical duration, but a different quantity of elements ($n_i \neq n_j$), for determination cross-correlation function we will use a next expression [4, 5]

$$R_{ij}(\tau) \leq \frac{1}{\sqrt{n_i n_j}}. \tag{5}$$

Code sequences of short video pulses, that meet condition of minimum similarity was called by sequences with the improved cross-correlation properties.

For implementation of terms (4), (5) it is suggested to use the sequences of video pulses which satisfy to such requirements [2]:

- quantity of short pulses (n_i, n_j) in the sequences *of* $u_j(t)$ must be different ($n_i > n_j$);
- duration of pulses in every sequence is equal ($\tau_{ui}=\tau_{uj}$);
- the periods of passing of impulses (T_i, T_j) in every sequence are assorted so that $n_iT_i \approx n_jT_j \approx T$, thus on-off ttime ratio of every sequence $Q_j > Q_i >> 1$, at $n_i > n_j$.
- On the basis of afore-mentioned limitations was built their system (6)which expressions (1, 3, 5) at the zero change of time shift to each other entered in the complement of ($R_{ij}(0)$, that is $\tau=0$).

From the system of limitations (6) the necessity of determination of pulses passing periods(T_i, T_j) *in every* sequence, the noted requirements must be executed here (3, 4, 5).

To work out this system (6) relative to the period of passing of pulses in a j sequence T_j it is necessary to represent a rectangular function $\Pi[t-(k\cdot Q_j-1)\cdot\tau_u]$ *thus*, that implementation of decision by the known methods was possible. A rectangular function is represented on fig. 1.

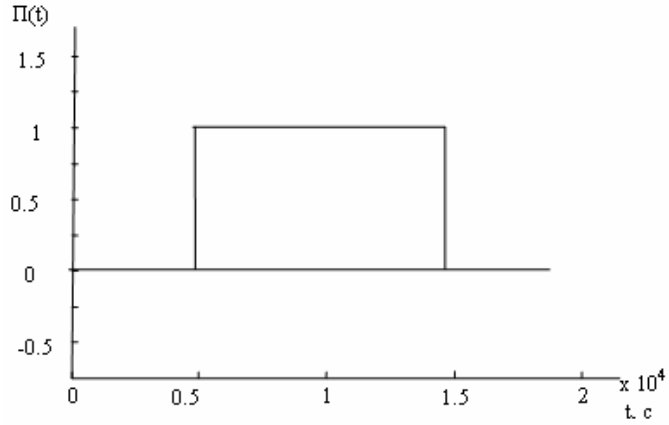


Fig. 1. Rectangular function

$$\begin{cases} u_i(t) = \sum_{k=1}^{n_i} U_{k_i} \Pi\left[t - (k \cdot Q_i - 1)\cdot \tau_u\right], \\ u_j(t) = \sum_{k=1}^{n_j} U_{k_j} \Pi\left[t - (k \cdot Q_j - 1)\cdot \tau_u\right], \\ Q_i = \dfrac{T_i}{\tau_u}, \\ Q_j = \dfrac{T_j}{\tau_u}, \\ R_{ij}(0) = \dfrac{1}{2\sqrt{E_i E_j}} \int_{-T}^{T} u_i(t) u_j(t) dt, \\ R_{ij}(0) \le \dfrac{1}{\sqrt{n_i n_j}}. \end{cases} \quad (6)$$

We approximate this function by means of the Heaviside Step Function [7, 8]

$$\prod[\frac{t}{\tau_u}-(k\cdot Q_j-1)]=H([t-(k\cdot Q_j-1)\cdot\tau_u]+\frac{\tau_u}{2})-H([t-k\cdot Q_j-1)\cdot\tau_u]-\frac{\tau_u}{2}) \quad (7)$$

$$\prod[\frac{t}{\tau_u}-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)]=H([t-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]+\frac{\tau_u}{2})-H([t-k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]-\frac{\tau_u}{2}). \quad (8)$$

The analytical form of approximation of Heaviside Step Function is of the form [8]

$$H(x)\approx\frac{1}{2}+\frac{1}{2}\cdot th(kx)=\frac{1}{1+e^{-2kx}}. \quad (9)$$

We will convert a rectangular function, using expressions (8) and (9).Then it can be written in a functional form [8]

$$\prod[\frac{t}{\tau_u}-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)]=\frac{1}{2}+\frac{1}{2}\cdot th([t-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]+\frac{\tau_u}{2})-\frac{1}{2}-\frac{1}{2}\cdot th([t-k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]-\frac{\tau_u}{2}),$$

$$\prod[\frac{t}{\tau_u}-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)]=\frac{1}{2}\cdot th([t-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]+\frac{\tau_u}{2})-th([t-k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]-\frac{\tau_u}{2}), \quad (10)$$

or in an exponential form [8]

$$\prod[\frac{t}{\tau_u}-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)]=\frac{1}{1+e^{-2k([t-(k\cdot\frac{n_j}{n_i}\cdot T_i-1\cdot\tau_u]+\frac{\tau_u}{2})}}-\frac{1}{1+e^{-2k([t-k\cdot\frac{n_j}{n_i}\cdot T_i-1\cdot\tau_u]+\frac{\tau_u}{2})}}. \quad (11)$$

We will put the got value of rectangular function in expression that describes the sequence of short video pulses u_j:

$$u_i=\sum_{i=1}^{n_i}\prod[\frac{t}{\tau_u}-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)]=\sum_{j=1}^{n_j}\frac{1}{2}th([t-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]+\frac{\tau_и}{2})-th([t-k\cdot\frac{n_j}{n_i}\cdot T_i-1)\cdot\tau_u]-\frac{\tau_u}{2}), \quad (12)$$

or:

$$u_i=\sum_{i=1}^{n_i}\prod[\frac{t}{\tau_u}-(k\cdot\frac{n_j}{n_i}\cdot T_i-1)]=\sum_{j=1}^{n_j}\frac{1}{1+e^{-2k([t-(k\cdot\frac{n_j}{n_i}\cdot T_i-1\cdot\tau_u]+\frac{\tau_u}{2})}}-\frac{1}{1+e^{-2k([t-k\cdot\frac{n_j}{n_i}\cdot T_i-1\cdot\tau_u]+\frac{\tau_u}{2})}}. \quad (13)$$

In *i* and *j* sequences a maximal value is taken to unit, and minimum to the zero, and here in expressions (12), (13) $Uk = 1$ in the definite moment of time for every pulse sequence.

Contrariwise it is necessary to provide the observance of condition (5). What we will oppose (5) and (3):

$$\frac{1}{2\sqrt{E_i E_j}} \int_{-T}^{T} u_i(t) u_j(t-\tau) \le \frac{1}{\sqrt{n_i n_j}}.$$

Substituting expressions (12) or (13) in (3) it is possible finally to define the period of passing of short video pulses in sequences T.

As a result of integration with the use of the module of character transformations MathCad, was got resulting value of period of passing of short videoimpulses in sequences with the improved cross-correlation properties in a *j* sequence:

$$T_j = \frac{T_i n_i}{n_j} = \frac{n_i \tau_u Q_i}{T}, \tag{14}$$

where: T – duration of ensemble sequences;

n_i – quantity of elements in a sequence $i=1...L, (n_i \neq n_j, i \neq j)$;

$Q_{\max} = \frac{T}{\tau_u \cdot n_{\min}}$, where Q_{max} – on-off time ratio of pulses in a sequence with the least of elements n_{min} [1, 3].

Thus, was developed mathematical apparatus of forming of ensembles of code sequences with small interaction, that allows to synthesize the sequences of short video pulses. Expression (14)lies in the basis of method, that possibility gives to define the period of passing of impulses in every sequence of the synthesized ensemble. All sequences from such ensembles have a low intrasystem interference level, that is determined by the values of the maximal outliers of lateral cross-correlation function petals of signals, that interact [2] .

The working capacity of expression is illustrated by an example in which are used four sequences *of s1(t)*, *s2(t)*, *s3(t)*, *s4(t)* with the quantity of elements $n_1=19$, $n_2=20$, $n_3=21$, $n_4=23$, signals duration is $T \approx 230$ mcs, pulses duration $\tau_u = 10$ ns. The pulses passing periods in every sequence were chosen accordingly: *T1=12.2* mcs, *T2=11.5* mcs, *T3=10.94* mcs and *T4=10* mcs. The type of sequences *of s1(t)* and *s2(t)* is presented on the Figure 2.

Setting the equal values of pulses amplitude, we will define energy of every sequence

$$E_i = n_i \cdot U_i^2 \cdot \tau_u, \tag{15}$$

where n_i – quantity of pulses in *i*- sequences;

τ – duration of pulse.

Obviously, energies of signals are different, as a result of a different quantity of pulses in every sequence. As sequences have different energies, for estimation cross-correlation in expression (3) they must be normalized [2]

$$s_{i\,norm}(t) = \frac{s_i(t)}{\sqrt{E_i}}. \tag{16}$$

The maximal values of sequences cross-correlation functions, calculated on a formula (3), with consideration of expression (14) are presented in table 1.

Table 1. Maximal values VCF

	s1(t)	*s2(t)*	*s3(t)*	*s4(t)*
s1(t)	*1*	*0,0513*	*0,0501*	*0,0488*
s2(t)	*0,0513*	*1*	*0.0478*	*0.0466*
s3(t)	*0,0501*	*0.0478*	*1*	*0.0455*
s4(t)	*0.0488*	*0.0466*	*0.0455*	*1*

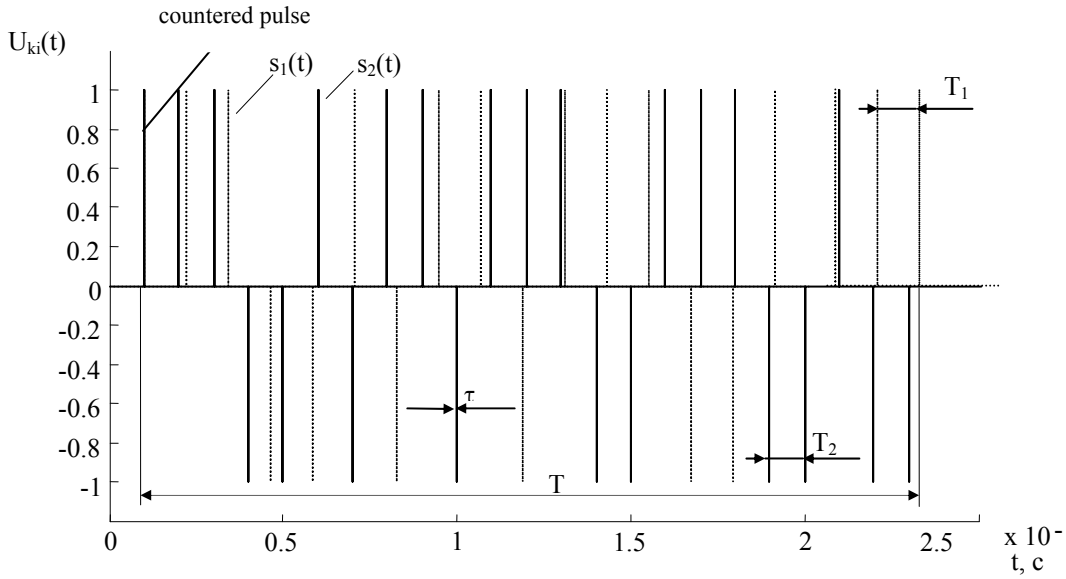


Fig. 3. Type of sequences $s_1(t)$ and $s_2(t)$ dependence on the quantity of elements

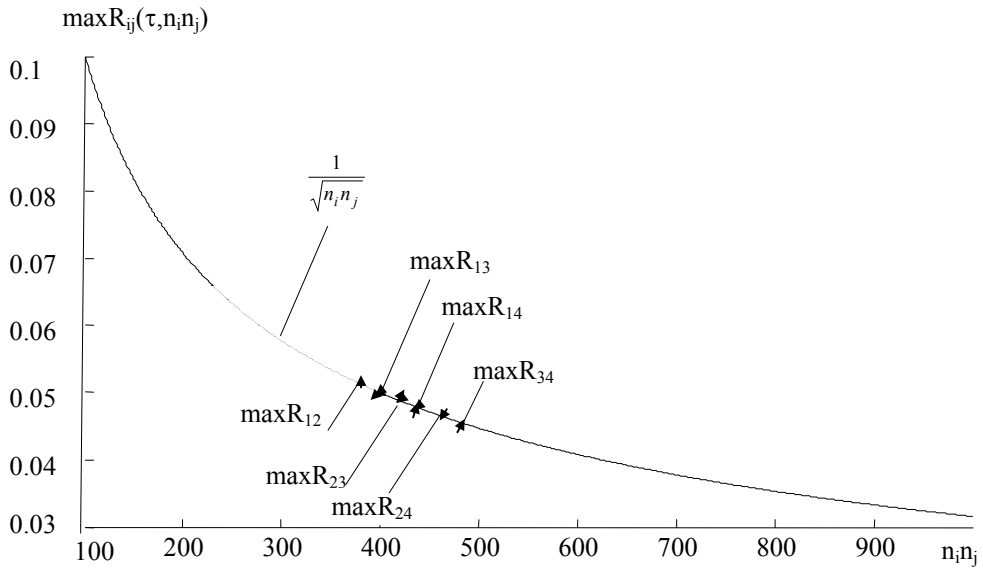


Fig. 4. Change of sequences cross-correlation coefficient of depending on the quantity of elements

On the Figure 4 the results of computations of maximums cross-correlation functions for the synthesized sequences with the small level of cross-correlation with the condition of minimum similarity are compared (5). Obviously, that the got sequences provide implementation of minimum similarity condition.

5. CONCLUSIONS

The use of the offered method of forming of sequences with the improved cross-correlation properties allows to minimize interaction between signals in a time domain, and, as a result, substantially decrease an intrasystem interference level. The offered mathematical apparatus allows to simplify forming of sequences ensembles from short video pulses with the improved cross-correlation properties. The radiation and reception of short video pulses is the separate scientific task [8], that in the given article was not considered. In the same time, there is foundation to consider that the use of synchronous transposition of time intervals in the sequences with the improved cross-correlation properties at definite terms insignificantly will worsen cross-correlation properties of signals, however considerably will rise the volume of signals ensembles that will provide possibility of their use in the existent systems of radio communication. It is foreseen to publish the results of such researches in next articles.

REFERENCES

[1] В. П. Лысечко, В.Н. Харченко. Метод борьбы с внутрисистемными радиопомехами // Системи обработки информации. – Х.: ХВУ –2004. – Вип. 2. – С.232-237.

[2] В. П. Лысечко. Метод определения параметров сложных сигналов // Системи обработки информации. – Х.: ХВУ –2004. – Вип. 7. – С.131-136.

[3] Помехозащищенность радиосистем со сложными сигналами, под ред. Г. И. Тузова. - М.: «Радио и связь», 1985 г. – 284 с.

[4] Т. А. Оганов, Помехоустойчивость инвариантного приема импульсных сигналов. – М.: Радио и связь, 1984. – 176 с.

[5] Л. Е. Варакин, Системы связи с шумоподобными сигналами. – М.:Радио и связь, 1985. – 384 с.

[6] Бернард Скляр. Цифровая связь. Теоретические основы и практическое применение. – М. Издательский дом «Вильямс», 2003. – 1104 с.

[7] Б. М. Болотовский, "Оливер Хевисайд", М. Изд. "Наука", 1985 г.

[8] Glover, I.; Grant, P. (2004). Digital Communications (2nd ed.). Pearson Education Ltd

HARDWARE IDENTIFICATION UNIT BASED ON GILLOU-QUISQUATER METHOD

Valentin S. Mollov, Yancho Z. Kolev

Department of Computer Systems, Technical University of Sofia, 6 Kl.Ohridsky blvd., 1000-Sofia, Bulgaria, phone: +359 (2) 9653523, e-mail: mollov@tu-sofia.bg

Abstract: *The paper presents implementation of a Hardware Protection Unit (HPU) for identification of software applications of any kind. The protection is against unauthorized usage of software and applies the Gillou-Quisquater cryptographic method which belongs to as called "zero-knowledge" challenge-response asymmetrical algorithms. It guarantees high level of security and arbitrary low level of complexity of the hardware. Implementation is prototyped over Xilinx FPGA Spartan3E 500K chip with limited number of dedicated multipliers. To avoid this inherent disadvantage, a separate Montgomery algorithm is applied instead of direct multiplication. Communication between the computer and the HPU uses EIA232 serial interface. Several examples with simulation waveforms are given to confirm the proper functionality of the prototyped device.*

Keywords: cryptography, zero-knowledge, hardware protection, FPGA

1. INTRODUCTION. ZERO-KNOWLEDGE IDENTIFICATION

The "zero-knowledge proof" concept has been first introduced by Shafi Goldwasser, Silvio Micali and Charles Rackoff during 80s, later extended and developed by Adi Shamir [1,2].

The "zero-knowledge" technique is applied inside interactive secret-key cryptosystem which uses some proof of knowledge, where the Prover (P) tries to convince the Verifier (V) that it knows a secret parameter (key). This proof itself could be accomplished through the ability to calculate complex arithmetic operations. For instance, if A has to prove that it knows two prime numbers p and q, so that p.q=n, it is possible as a proof to be accepted the ability of A to calculate square root of modulo n, which is count impossible in case of not-knowing p and q. This idea, basically inspired by the hash function modular arithmetic theory [1] could be applied inside an identification procedure, presuming that the knowledge of p and q (p.q=n) is associated with some identity I. So, the Prover could confirm its identity without revealing the secret values, and the verifier could accept or reject this proof by using only public parameters.

Here, we shall remind the following characteristics of zero-knowledge protocols:

(i) completeness: if both the prover and verifier are authentic, the Prover always accept the proof;

(ii) soundness: a threshold probability τ exists, so that for each probability $p>\tau$, there is a model extracting information (extractor), which effectively transforms every possible dishonest Prover, accepted by the verifier with probability p, into algorithm, calculating values, practically equivalent to the secret parameters of the honest

Prover [2]. We have to mention, that such a model presents an algorithm, extracting (inside a polynomial time limit) information from the interaction between the dishonest Prover and Verifier, which permits with a sufficient level of confidence the proper execution of the protocol. The soundness requirement means to have a guarantee that the protocol ensures and applies a proof of knowledge and to be successful it is necessary to use knowledge for the secret values or their equivalent, obtained with as called "extractor". Here it is important to clear, that soundness of such kind of protocols does not mean that it guarantees absolute security of the secret parameters. It prevents with a 1–p probability the successful performance of the protocol by dishonest Prover. The secret keys are kept secret due to specifics of the preliminary presumptions of applied mathematical functions during calculations.

(iii) zero-knowledge means that if the prover fulfills the protocol many times, even exchanging data with dishonest verifier, it does not reveal any information about the secret values.

2. GUILLOU-QUISQUATER METHOD TO PROVE IDENTITY

It is accepted, that the verifier stores an authentic copy of the public parameters to avoid application of any certificates or hash functions during the validity check [1].

2.1. Selection and calculation of the parameters. Protocol processing details

The involving parameters from both sides of the cryptographic protocol must satisfy some preliminary requirements. Also, some calculations over the initial prime number have to be done in advance. The procedure is as follows:

(i) two separate prime numbers p and q are chosen and the corresponding value $n=p.q$ is calculated;

(ii) $a<n$ is chosen and b is calculated, which satisfies the condition $a.b \bmod \varphi(n)=1$, i.e. $b=a^{-1} \bmod \varphi(n)$, where $\varphi(n)=(p-1)(q-1)$ is the Euler function;

(iii) u is a public user-identifier;

(iv) public values are n, a, and u; secret values are p, q, and b;

(v) $s=(u^{-1})^b \bmod n$;

(vi) the value of s is secret. It is known only by the dedicated person and could not be calculated by another person.

Fulfilling the protocol (method), the Prover P proves its identity, while the Verifier V checks (confirms) the proof. So, P and V operate over the following steps, presenting the communication exchange between the two sides of the protocol:

(1) P choose a random number r , $0 \le r \le n-1$, calculates $x=r^a \bmod n$ and sends x towards V;

(2) V choose a random number e („challenge") , $0 \le e \le a-1$, and sends e to P;

(3) P calculates $y=r.s^e \bmod n$ („response") and sends y to V;

(4) V checks if the condition $x'=x$ is true, where $x'=u^e.y^a \bmod n$;

(5) if the upper condition is true, so the protocol is successful.

2.2. Security of Guillou-Quisquater method. Numerical example

The Guillou-Quisquater method is probabilistic [3,4], so the following attributes are valid about it: (1) probability of corruption (false check): here a defines the level of security. False Prover could break the protocol during attack with a probability of value 1/a, suggesting in advance the random number e. To avoid that, the recommended length of a depends on the environment where the attacks could happened;

(2) suggestion of security: it is necessary to extract a^{th} root modulo the composite number n to break the protocol. It is count practically impossible to calculate if the prime numbers p and q are unknown.

Example: Let us have the following parameters in decimal number system:
– p=569, q=739, n=420 491,
– u=34 579
– a=54 955, b=233 875, $\varphi(n)$=419 184
– s=403 154
The Prover (P) and Verifier (V) calculate and transmit the following data:

P: r = 65 446 , x = 89525	$P \xrightarrow{x} V$
V: e = 38 980	$V \xrightarrow{e} P$
P: y = 83 551	$P \xrightarrow{y} V$
V: x′ = 89 525	$V \xrightarrow{c} P$

V: The protocol is successful, as x′=x.

2.3. Advantages and drawbacks of Guillou-Quisquater approach

Guillou-Quisquater protocol is cryptographic identification algorithm using zero-knowledge proof and uses asymmetrical challenge-response technique. The method itself fits quite well to the concept of authentication of the protected application software through a separate hardware protection unit which stores itself the secret values of the algorithm. It gives mathematical background of such protection with a high level of security. It is also easy to adapt towards such task and in this sense corresponds to the preliminary requirements. The method also provides arbitrary easy procedure for calculation public and secret parameters in respect to other protocols of this kind. When using serial interface, the Guillou-Quisquater method is especially convenient as it needs only three exchange sessions to fulfill the entire protocol.

Among the practical issues that arise during hardware implementation, the most limiting one is the amount of hardware resources needed to execute all arithmetic calculations. This limitation suggests lower speed processing effectiveness. It could be successfully avoid by applying suitable calculation algorithms, as it is done here (e.g. Montgomery algorithm).

3. HARDWARE PROTECTION MODULE (HPU) PROTOTYPING

3.1. HPU and application software role into identification algorithm

The main objective of applying special algorithms is to protect specific application software against unauthorized usage which is allowed to only one dedicated user. Here, we apply a separate Hardware Protection Unit (HPU) which maps the Guillou-Quisquater cryptographic method, a zero-knowledge asymmetric type algorithm for identification.

The application software has not to be run successfully if there is no verified link according to the protocol between the computer system and the HPU (provided by the software owner). We mean verified link to describe that the HPU possesses the unique secret key, so that it confirms its identity. The application software also identifies the HPU by sending messages to it (one or several times during operation). There should be full correspondence between the HPU and the protected software. It means that during the application of algorithm, based on the specific cryptographic method, by the software there should have a confirmation of identity of the HPU.

It is suitable to apply asymmetric exchange techniques, as in our case. Also, it is sensible to store the public parameters into the main computer system where the application software is run, while the secret parameters to be stored inside the HPU. The opportunity not to store the secret parameters in the computer is one of the dominant advantages offered by the asymmetrical identification algorithms. So, it avoids the option of easily revealing the secret keys by the intruder from the computer side, where it is expected usually.

Despite the fact that the secret values are stored in HPU only, the identification is possible to be done due to direct usage of public parameters, which by definition, are calculated based on the secret ones, i.e. the verifier (application software) will not need the secret values of the Prover (HPU) during verification. Revealing of the secret values from the public parameters is practically impossible because of the specific nature of mathematical calculations to be done.

Finally, the application of identification procedure guarantees that the protected application software will not run successfully without having the correspondent (unique secret keys) HPU, so that the owner of the software is able to control its copyright by providing or not the hardware protection device to the customer.

3.2. Calculation of 32-bit parameters of Guillou-Quisquater into the prototyped HPU. Block diagram. Simulations

The implemented HPU, based on Guillou-Quisquater algorithms uses the values of s, a, and n, as it is mentioned earlier. The values n and a are public, while s is secret. The rest of parameters p, q, b, u, $\varphi(n)$ are not needed directly, as they participate in the preliminary calculations, also p, q, b, and $\varphi(n)$ are not publically announced. For our practical implementation, the following 32-bit parameters of the Guillou-Quisquater protocol are used (shown in hexagonal numbers here):

p=0000 FDB7, q=0000 BC77, n=BAC8 5411, φ(n)=BAC6 99E4
a=2038 86A1, b=568A 37E9
u=0739 755A, s=7348 07EE

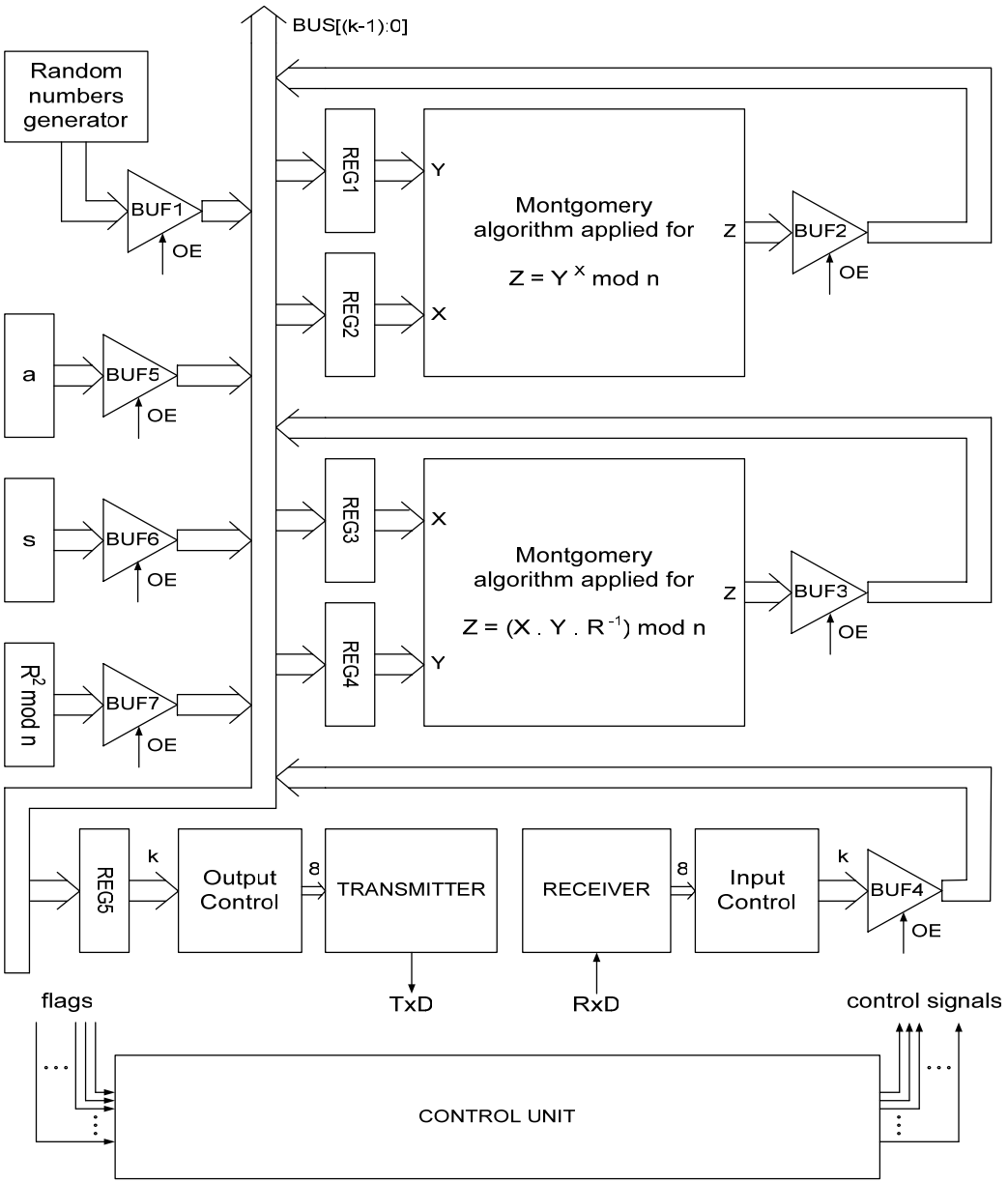


Fig. 1. Simplified block diagram of the prototyped HPU

The simplified block structure of the prototyped HPU is shown in Fig. 1. Montgomery algorithm is applied instead of direct multiplication for module n calculation. The overall algorithm is performed by the Control Unit with the correspondent input flags and output control signals. A separate block generates the random numbers (Random Number Generator). All participating values are physically buffered and enabled by the OE input signal. The REG1 to REG4 store the 32-bit values to drive the Montgomery algorithm which results are also buffered. Communication is accomplished through EIA232 interface, because of its simplicity and, as no other spe-

cial requirements are set at this stage. The block of Transmitter and Receiver sends and receives data via TxD and RxD buses accordingly.

Fig. 2 shows sample results of the operation of Montgomery modulo n multiplication block (the upper figure) and exchange data into receiver part of the EIA232 module (the lower figure).

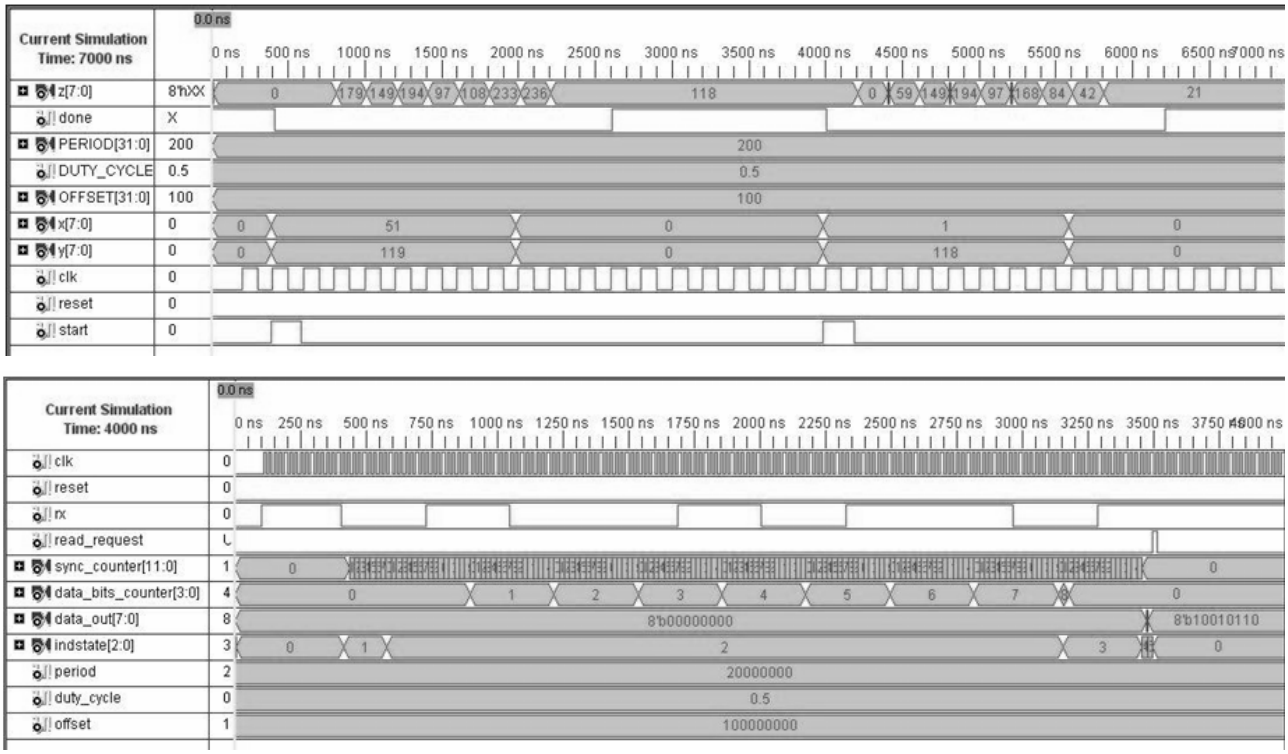


Fig. 2. Simulation waveforms from some of the HPU modules: Montgomery modulo n calculation block (the upper figure); Receiver part of the EIA232 module (the lower figure).

Implementation of the overall HPU is done over reconfigurable Field Programmable Gate Array (FPGA) chip with 500K gates from the Xilinx Spartan3E family. It gives sufficient amount of general-purposed resources (logical elements and dedicated memory), but does not provide enough on-board multipliers. To avoid this limitation, we apply Montgomery algorithms for modulo n calculation which includes several multiplications. Prototyping cryptographic algorithms over FPGAs is wide spreading approach [5,6]. It is very convenient and brings no cost for designer in case of frequent design adjustments.

4. CONCLUSIONS

This paper shows a very efficient approach to perform secure identification of any application software through as called zero-knowledge asymmetrical cryptographic algorithms with a separate hardware protection unit (HPU). Guillou-Quisquatter method is applied and the corresponding algorithm's steps are defined clearly in view of its implementation over dedicated FPGA. Specially tailored design approach to avoid chip limitation is proposed, as well. The results demonstrate a way to achieve a

high level of security in identification with storing the secret values of the algorithm only over the HPU. This approach could be also extended towards network security.

REFERENCES

[1] Menezes, A., van Oorshot, P., and Vanstone, S., Handbook of Applied Cryptography, CRC Press, 1997.

[2] S. Vaudenay, A Classical Introduction to Cryptography: Applications for Communications Security, Springer 2006.

[3] Guillou, Quisquater, A practical zero-knowledge protocol fitted to security microprocessor minimizing both transmission and memory, Advances in Cryptology EUROCRYPT '87, C. G. Guenther (ed.), Springer-Verlag, pp.123-128.

[4] Takeshi Okamoto, Design and Analysis of Praktical Digital Signature Schemes, Feb.2002, Ph.D. Thesis, JAIST, p.80.

[5] Quisquater, J.-J. and Couvreur, C., Fast decipherment algorithm for RSA public-key cryptosystem, Electronics Letters Volume: 18, Issue: 21, pp. 905-907, 1982

[6] Rouvroy, G., Standaert, F.-X., Quisquater, J.-J., and Legat, J.-D., Efficient uses of FPGAs for implementations of DES and its experimental linear cryptanalysis, IEEE Transactions on Computers, Volume: 52, Issue: 4, pp. 473-482, 2003.

HYDRAULIC ANALOGY FOR INDUCTIVE ELECTRIC ELEMENTS

George Popov[1)], Rumen Trifonov[2)]

[1)]Computer Science Department, Technical University of Sofia, Kl. Ohridsky Str, No.8, 1000, Bulgaria, phone: +3592 965-22-24, e-mail: popovg@tu-sofia.bg

[2)]Computer Science Department, Technical University of Sofia, Kl. Ohridsky Str,No.8, 1000, Bulgaria, phone: +3592 965-22-24, e-mail: r_trifonov@tu-sofia.bg

Abstract: *This document presents a new hydraulic model of coil and transformer. The method of hydraulic analogies is very useful to explain physical phenomena in area of electrical circuits, because hydraulic models have a lower level of abstraction.*

Keywords: electric circuits, electric chain, hydraulic model, coil, transformer

1. INTRODUCTION

Use of the hydraulic analogies to explain circuits and electronic circuits has many advantages: things are dealt with at a lower level of abstraction, as are associated with clear concepts of learner, to summarize the common phenomena in nature, it is possible to transfer knowledge and skills from one area to another, etc.

There are large known analogies between electric and hydraulic circuits:

- pressure – potential;
- difference in pressure – tension;
- flow – current;
- narrowing of the tube – resistance;
- pump – electric current generator;
- tower pressure – voltage generator;
- valve – diode;
- quantity of water – capacity.

It is interesting to note that there are hydraulic models of bipolar and field transistors, to whom are applicable some fundamental equations of electronics!

The abstract expression of electrical parameters by analogy has another useful side – it can be clarified Precedents factors and to derive appropriate mathematical formulas. Another useful part is the use of the developed mathematical apparatus for electrical circuits for the study of other phenomena.

It is obvious that, first and second Kirchhoff laws apply to hydraulic circuits. Also Ohm's law formula for the resistance $R = \rho \frac{l}{S}$, where ρ is resistivity, l – length, and S – the intersection of the tube.

At Fig. 1 is given a famous analogy of E. Aiseberg, where a capacitor is discharged through the resistance. Fig. 2 shows charging of capacitor by current generator.

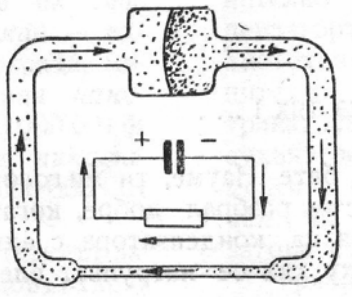


Fig. 1. Discharging of capacitor

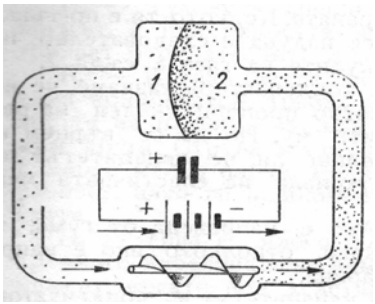


Fig. 2. Charging of capacitor

In Fig. 3 is given an analogy of a hydraulic vibrator circuit with its fluctuations. If there is parallel included a voltmeter to this resonant circuit we will be monitor a voltage's resonance.

Contrary, if there is connected serially an ampermeter in this circuit; we will see a resonance of currents.

Sometimes there is very difficult to explain these things to students in electrical engineering.

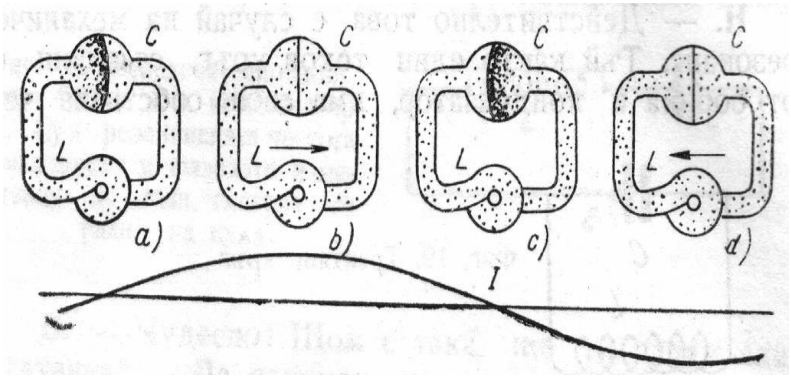


Fig. 3. A model LC circuit

In Fig. 4 shows the passage of alternating current in through a capacitor, here can be explained the role of the decoupling capacitor in electronic circuits. At Fig. 5 is shown that the two series connected capacitors have less capacity of each of them.

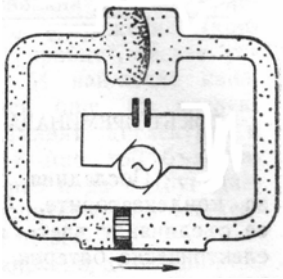


Fig. 4

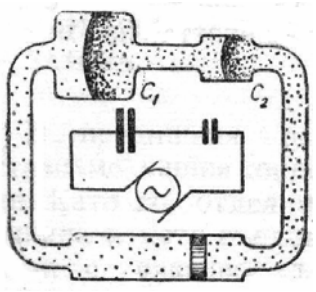


Fig. 5

From the foregoing it appears that the only model of the coil is not sufficiently adequate. At Fig. 3 phenomenon of inductance is modeled by the fluid inertia in the extended tube.

In this case cannot be modeled impact parameter L of the bobbin. Nor can realize a transformer, to model phenomena such as reaction of the current of the anchor and etc.

2. MODELING OF INDUCTANCE

If in the analogy (used in Fig. 3) it can replace an hydraulic coil with propeller with flywheel, there will have a more adequate model of the phenomenon inductance. It is clear that the resulting model has more adequate properties of these of real bobbin. The device resists any change to the current like real bobbin: when current increases the device make a resistance and when the current slows device supports it.

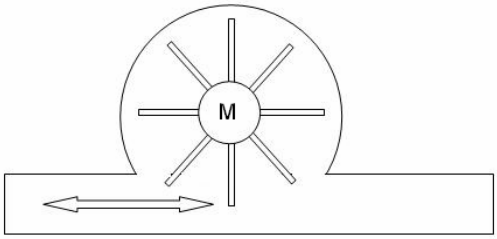


Fig. 6. A hydraulic model of the bobbin

The analogy is also another area – in the coils the electrical energy is converted into magnetic field, in the model – the energy of a moving fluid converts into kinetic energy.

The inductance of the bobbin depends on the actual number of coils and the parameter L, associated with ferromagnetic core, but here this is modeled with number of blades and mass flywheel M.

A parasitic resistance R of the bobbin is expressed through resistance tubes and bearings in the model.

If two hydraulic bobbins are contacted with a common axis, there is obtained a device with similar parameters of the real transformer:

- a means for galvanic decoupling (splitting of fluid – flows);
- transformer current and voltage (flow, pressure)
- transformer of impedances (momentum, force);
- a device with Inductive character.

The coefficient of transformation can be realized by different number and shape of blades or by mechanical gearbox system.

This model of the transformer has only one drawback – work with both AC and DC.

A better model of transformer is shown at Fig. 7, where two pipes (cylinders) have pistons connected together. They pass the AC fluid movements. To model the inductance and allowing to pass DC, parallel to them are connected two coils (like these at Fig. 6).

The model would be more appropriate if the relationship between the two pistons is made through first-generation lever. By this way it can model dephasing between output and input voltage.

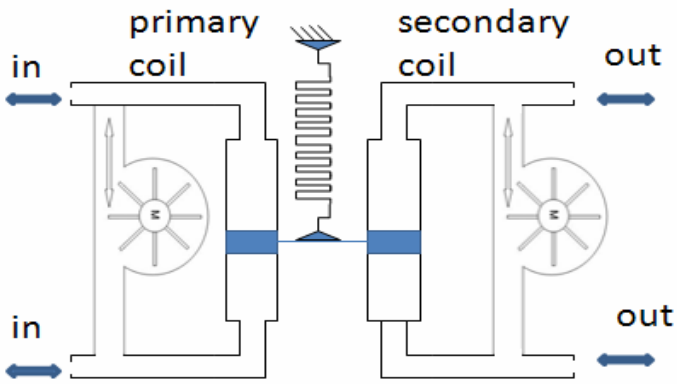


Fig. 7. A hydraulic model of transformer

Transformation ratio can be altered by varying the intersection of the cylinders and pistons.

3. CONCLUSIONS

Such devices are commonly used in hydraulic machines, but the idea here is to enrich the hydraulic circuits of analogies. This article may be written at the level of differential equations where there is such a similarity, but this should be the next step in the process of didactic teaching of the discipline, i.e. harmony in physics, respectively in the nature.

REFERENCES

[1] E. Aisberg, Le Transistor? ...Mais C Est Très Simple!, Societe Des Editions Radio – 1969

[2] E. Aisberg La Radio ? Mais C Est Très Simple, Dunod – 14/10/1998

[3] Popov G., Hydraulic Models of Inductive Elements, Symposium of metrology, Sozopol, 2006

[4] E. Laclais. Alarme? Pas de panique! Guide de l'installation reussie, PubliTronic, Pais-Bas, Avril, 1995.

[5] http://en.wikipedia.org/wiki/Hydraulic_analogy

A FUZZY LOGIC DSP PROCESSING FOR PET IMMUNITY PIR DETECTORS

Daniela Gotseva[1)], George Popov[2)]

[1)]Computer Science Department, Technical University of Sofia, Kl. Ohridsky Str, No.8, 1000, Bulgaria, phone: +3592 965-23-38, e-mail: dgoceva@tu-sofia.bg

[2)]Computer Science Department, Technical University of Sofia, Kl. Ohridsky Str, No.8, 1000, Bulgaria, phone: +3592 965-22-24, e-mail: popovg@tu-sofia.bg

Abstract: *This paper concerns fuzzy logic based DSP processing for high stable passive infrared detectors with increased pet immunity. The suggested algorithm analyzes the factors of movement and decide alarm situation using fuzzy logic model.*

Keywords: pet immunity PIR, motion detector, security systems, alarm, DSP PIR

1. INTRODUCTION

Passive Infrared motion detectors are most frequently used in the security systems. It is determined by the following factors: low cost, low false alarm ratio, high possibility of detection, no harmful emissions.

The sensor of PIR detector is called pyrosensor. It captures through facet lens the temperature of environment (fig. 1). Human body emits temperature waves with a length between 1 to 20 μm. When a human crosses the beam, the infrared energy is exposited on the pyrosensor. The change of this exposition is useful signal, which is a criterion of alarm situation. The signal of pyrosensor is shown in fig. 2.

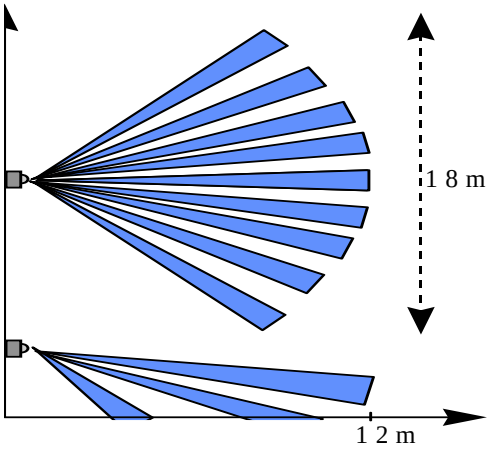


Fig. 1. Facet lens diagram

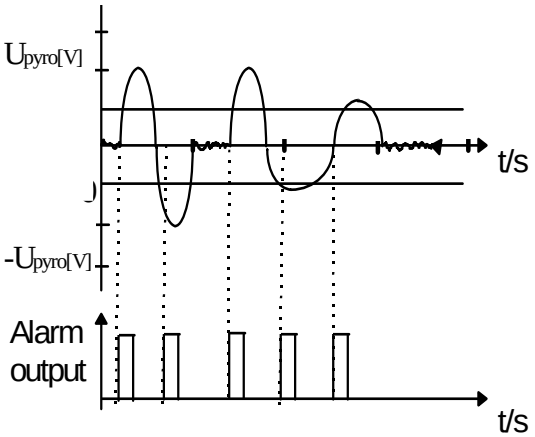


Fig. 2. The signal of pyrosensor

Most of detectors used analog circuits, as amplifiers, integrators and comparators. They are called threshold detectors. Generally one detector is composed by 4-8 operation amplifiers (fig. 3).

The disadvantages of analog PIR detector are: aging of components, low RFI immunity, low dynamic range, false alarms caused by environmental factors – air streams and vents, sunshine, small animals – mice and bugs, etc.

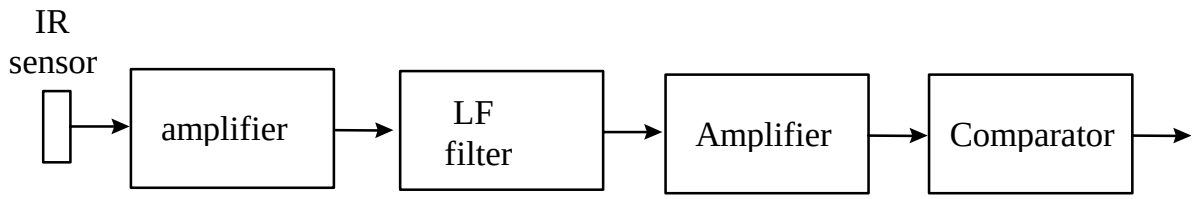


Fig. 3. Block diagram of analogue PIR detector

Some security detectors are designed to guard the premises in the presence of pets. These detectors consist of two sensors, each with its separate optical system implemented so as to monitor the various sectors of the protected area. If the signal of motion reaches both pyrosensors, that means a man is moving in protected area. Conversely, if the any pet is moving in the protected volume, it will receive alternate signals from both sensors. In Fig.4 sectors which are viewed by channel A are dark colored and these by channel B are light colored.

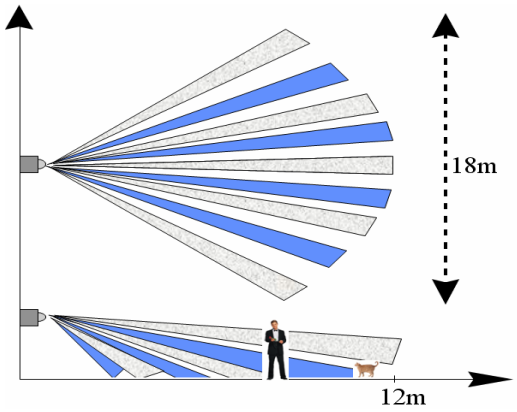


Fig. 4. Principle of pet immunity PIR detector

Block scheme of this type detector is given at Fig. 5.

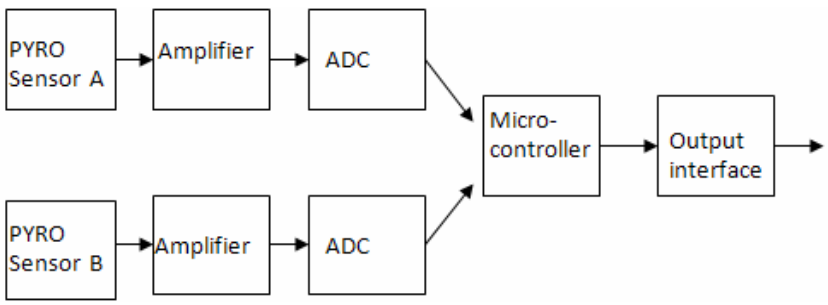


Fig. 5. Principle diagram of pet immunity PIR detector

2. FORMULATION OF RESEARCH TASK

If it is proceeded from the contradiction requirements for high probability of detection (Pd=1) and low false alarm ratio (FAR=0), it must use microcontroller solution of PIR detector.

We offer a special DSP processing for PIR detector. For a presence of alarm situation, we can try by symptoms of pyro's signal:

- amplitude – as much amplitude is high, the possibility of alarm situation is great;
- frequency – the frequency of human movement is from 0.8Hz up to 10Hz. (Intruder speed varies between 0.15m/s to 8m/s.). Low frequency means temperature fluctuation, sun spots and flows. High frequency means RFI interference, vibrations of basement etc.;
- signal form – if we eliminate constant component, the form of the signal must be sinusoidal. Every variation of this form, probably means false signal (false alarm);
- energy – as much energy is high, the possibility of alarm situation is great;

Every of symptom is described by the fuzzy function, receiving the result by table 1.

3. REALIZATION OF THE FUZZY LOGIC MODEL

Depending of fact, so the signal generated by pyrosensor is too weak, the signal is amplified by non inverting scale amplifier. The output signal of this circuit is within 0V-4.8V, because detectors power is 12V, stabilized by 78xx05 to 5.00V.

That is why the signal is defined by linguistic variables: amplitude, frequency and signal form.

For every one of these variables we assign the range and values of fuzzy constants.

Amplitude has changed into interval from 0 to 4,8V and can get the value:

- low – with diapason from 0 to 1,6V and function linear descending;
- medium – with diapason from 1,4V to 3,2V and function trapezoidal;
- high – with diapason from 3.00 to 4.8V and function linear ascending.

Frequency has changed into interval from 0Hz to 15Hz and can get the value:

- low – with diapason from 0Hz to 0.8Hz and function linear descending;
- medium – diapason from 0.6Hz to 8Hz and function trapezoidal;
- high – with diapason from 7.5Hz to 15Hz and function linear ascending;

The ideal form of the signal received by real intrusion is sinusoidal, symmetrical toward abscise. Any shift spoke about constant drive, result by temperature changes. Modulation with high frequency signs about RFI interference. If the form of the signal is not symmetrical, that means so this signal is generated by air flows and vents.

The form of the signal is presented in conditional units from 0 to 100. Depending his form, the signal is separated into 3 categories:

- ideal – with diapason from 80 to 100 and function linear ascending;

- similar – with diapason from 35 to 85 and function trapezoidal;
- random – with diapason from 0 to 40 and function linear ascending.

On the basis of signed input variables, we define one output linguistic variable, who displays possibility of alarm situation with diapason from 0% to 100%. She has next fuzzy constants:

- low possibility – with diapason from 0% to 35% and function linear descending;
- medium possibility – with diapason from 30% to 70% and function trapezoidal;
- high possibility – with diapason from 66% to 100% and function linear ascending.

Trough defined rules we make the basis of fuzzy rules, shown in Table 1 (shortened).

Table 1: Fuzzy rules

Channel A			Channel B			Possibility of alarm situation	
Amp1	Freq1	Form1	Amp2	Freq2	Form2		
low	low	similar	low	low	ideal	low	10
med	low	ideal	low	low	ideal	low	7
high	normal	ideal	low	low	ideal	medium	20
high	normal	ideal	med	norm	ideal	high	75
med	normal	ideal	med	norm	ideal	high	10
high	normal	ideal	high	normal	ideal	high	100
med	high	ideal	med	low	ideal	low	10
high	high	random	high	high	random	low	20
med	norm	ideal	high	norm	ideal	high	85
high	norm	similar	high	norm	similar	medium	50

4. CONCLUSION

The presented solution is simulated and tested using package MatLab. The obtained results proof that approach is very advantageous for the purposes of PIR detector design.

There is possibility through change to output function of detector (realize with DIP switch, potentiometer etc.) to realize PIR detector suitable for outdoor (noisy environment) or outdoor use.

Another advantage of this solution is a low cost realization – the necessary equipment is one small microcontroller with ADC and gain control (par example Microchip 16CXXX).

REFERENCES

[1] Gotseva D., Popov G., A fuzzi logic based DSP processing for infrared motion detectors, International Scientific Conference “Computer Science’2004”, Sofia, dec., 2004, pp.239-241.
[2] Gotseva, D. (2004) Embedding fuzzy variables into C++, CompSysTech’2004, II.8.1-II.8.5, Rousse, 2004.
[3] Capel V., (1993) Security Systems and Intruder Alarms, England. Oxford.
[4] Popov G., Pertova P., (1997) A model of Code Interface Control Panel for Security Systems with Time-out Nets. 11-th International Conference: Systems for Automation of Engineering and Research” and DECUS NUG Seminar’97, St. Konstantin Resort, Sept.,20-21th,1997, pp.64-67.
[5] Technical documentation Digigard 55,75 and 85 by Paradox Security Systems – Canada.

ELECTROACOUSTIC ANALYSIS OF MEMS FOR SOUND PRESSURE HARVESTING

Snejana Terzieva, Ivan Kralov, Ignat Ignatov

Department of Theoretical Electrical Engineering, Department of Mechanics,
Technical University of Sofia, 8, Kliment Ohridski Str., Sofia 1000, Bulgaria,
phone +0359965 23 94, +0359965 20 53, e-mail: ster19@tu-sofia.bg, kralov@tu-sofia.bg

Abstract: *A survey of noise energy harvesting methods is realized in this paper. Comparison between the basic functional parameters is carried out of the standard solutions for receiving, converting and storing the energy from the noise sources. The data collected could be used for future MEMS design.*

Keywords: Energy Harvesting, Micro-electro-mechanical systems, Noise, Piezoelectric.

1. INTRODUCTION

In the recent years, energy producing and harvesting from the surrounding environment becomes significant scientific area having practical application, and its development continues rapidly. The increased number of publications confirms it [1-7]. The developments in the micro-electro-mechanical systems (MEMS) have wide application, for example:

– Wireless receivers and life – supporting equipment supply;
– Battery recharging;
– Vehicles tire pressure monitoring used in the automotive industry;
– Remote control of security devices, etc.

Survey done by IDTechEx in 22 countries and 200 organizations in Europe, North America, East Asia and other countries indicates the between 2009 – 2019 the design and production of MEMS shall be speeded up.

The expanding publicity of problems, concerning the environment protection, could rise the energy harvesting technology (in this particular case sound pressure is used) using MEMS to a promising method for low voltage system supply. This will avoid the use of disposable batteries and the need of recycling.

The increase and development of portable electric devices having smaller and smaller mass, extends the necessity of highly effective energy sources to supply them. With the creation of technology and devices, which will be able to convert and store the energy from the surrounding environment, the question for excluding batteries as a main source supplying low power consuming devices arises.

A volume of $1\,cm^3$ lithium batteries consists of $2800\,J$, which is enough for supplying consumer with average power of $100\,\mu W$ nearly a year, but systems having higher energy consumption the battery replacement have to be made in a shorter periods. The most common systems convert solar, thermal energy and energy from me-

chanical oscillations in electrical. The main energy harvesting from mechanical oscillations approaches are three, namely: electromagnetic (inductive), electrostatic (capacitive) and piezoelectric [1, 3, 4, 7].

2. ANALYSIS OF THE SYSTEMS FOR SOUND PRESSURE HARVESTING

2.1. Electromagnetic sound pressure harvesters

The electromagnetic energy sources are widespread and the energy available could be harvested in inaccessible environment. The energy density of the electromagnetic sources is small (in order of 1 $\mu W/cm^2$). That requires the use of highly efficient amplifiers and transducers. In that case the control circuit is designed to harvest energy in longer periods, longer than 10 min, while in cycles shorter than 1s, by producing amplified output energy flow.

The structure of composite transducer, presented on fig. 1, consists of piezoelectric PZT fixed to Be-bronze horn – magnetostrictive Terfenol-D plate (produced by Gansu Tianxing Rare Earth Functional Materials Co., Ltd.) with dimensions 12 x 6 x 1 mm is longitudinal symmetrically located at the mechanical anchor of the ultrasonic horn where the displacement is zero [4]. Designed in this order the system aims to enhance the energy density and obtain convergence while transferring the wave trough larger cross sections to less cross sections. The experimental general design is shown on fig. 2. The three piezo-plates are longitudinal orientated and fixed, and their dimensions are 12 x 2 x 0.8 mm. They are glued to the ultrasonic horn by acrylic adhesive.

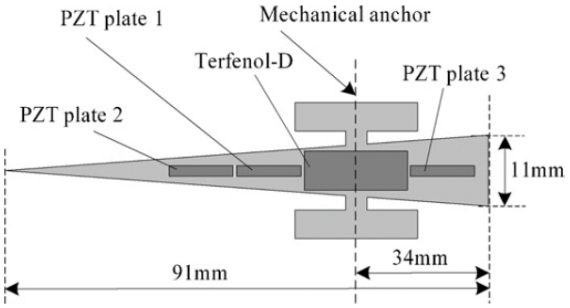


Fig. 1. Composite consisting of ultrasonic horn according to [4]

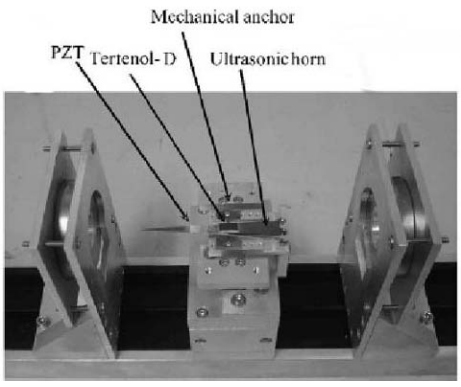


Fig. 2. Experimental apparatus photo of magnetoelectric transducer according to [4]

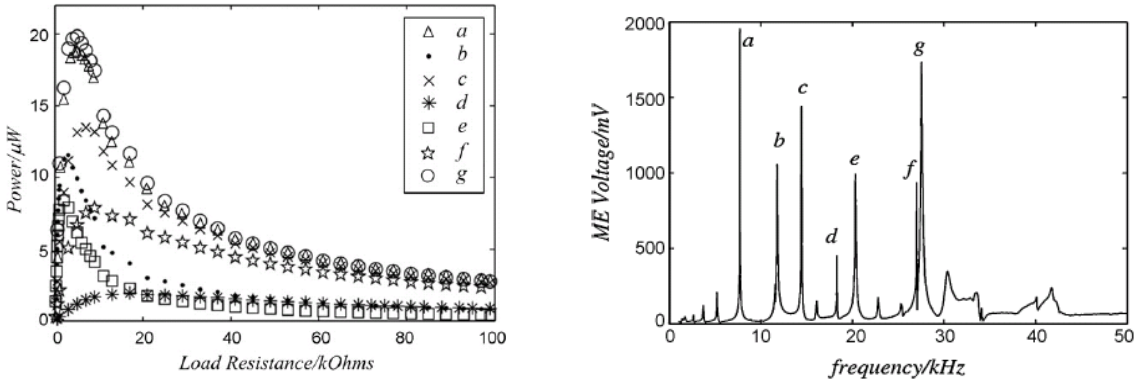


Fig. 3. Values of the voltage and power transduced by the system according to [4]

Fig. 3 shows the registered values of the transduced frequency and obtained power, as on the first graphic the registered multi peaks of the voltage (a to g) originates from the multiple vibrating modes at low frequencies, relative to the resonant frequencies of the ultrasonic horn. The maximum output power is obtained, while the output impedance matches the load impedance (in the case at resonant frequency 27 kHz).

The prototype of the composite Terfenol-D/PZT/Be–bronze, sound pressure harvesting system is designed to supply wireless receiver nodes, as during the tests is proved, that the system is able to support work mode of the sensor of $75\,mW$ consummation during transmitting and $1\,mW$ during receiving, in $600\,ms$ cycle of $60\,m$ radius.

2.2. Piezoelectric sound pressure harvesters

Most common and used piezo-materials in MEMS are ZnO (zinc oxide), AlN (aluminum nitride), PZT (lead zirconate titanate). The material is selected depending on various factors, some of them are: difficulties in producing it, ensuring certain parameters, compatibility with integrated circuits, etc.

The adoption of piezoelectric elements (plates, produced in specific way), for mechanical energy transducing (sound pressure in this particular case) is based on the direct piezoelectric effect – during mechanical deformation on the both surfaces of the plate opposite electrical charges occurred, i.e. voltage proportional to the deformation occurs. It could be assumed that in sinusoidal changing over time sound wave, on the plates output will be obtained sinusoidal voltage with the same frequency as the sound wave.

The most common piezo-transducer of sound pressure is the Helmholtz resonator with piezoelectric composite backplate. It consists of cavity, connected to the surrounding environment through narrow neck, as shown on fig. 4. With V is indicated the cavity's volume, $S = \pi a^2$ is the cross section area of the neck, l its length and a its radius, P_1 is the pressure of the surrounding environment, P_2 – the pressure of the cavity. The resonator is presented as acoustic oscillating system with lumped parame-

ters, which could be presented by equivalent mechanical oscillating system. The air transferred through the neck has finite mass and could induce dissipative losses. In the cavity, the air is compressible and it could be modeled as compliance. Moreover, based on electro acoustic analogy the model is transformed into equivalent electric circuit, which is presented by series RLC – circuit (fig. 5). There are three acoustic elements: L_A – mass, R_A – active resistance, C_A – compliance. They correspond to the electrical ones - coil with inductance, resistor with active resistance and capacitor respectively with certain capacity. The relevant equations according to [1] are:

$$L_A = \frac{\rho_0 V}{3(\pi a^2)^2} \quad \left[\frac{kg}{m^4}\right]; \quad R_A = \frac{8\pi\mu l}{S^2} \quad \left[\frac{kg}{m^4 s}\right]; \quad C_A = \frac{V}{\rho_0 c_0^2} \quad \left[\frac{m^3}{Pa}\right] \tag{1}$$

where ρ_0 is the mean density of the air, μ – dynamic viscosity of the air, c_0 – sound velocity.

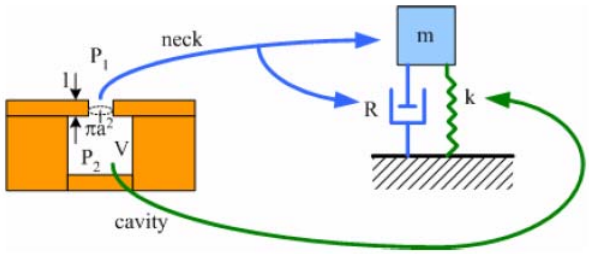


Fig. 4. Graphic showing the mechanical equivalent of the Helmholtz resonator

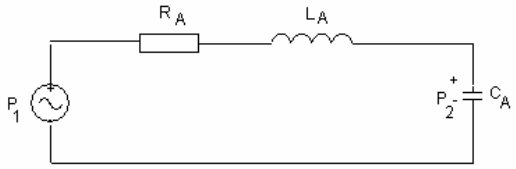


Fig. 5. Equivalent electrical circuit of the Helmholtz resonator

Using fig. 5 could be derived the dependencies for the complex coefficient $\dot{K}_{PA}$ of acoustic pressure intensification, the resonant frequency f_{res} and the pressure of that frequency PA_{res}. The equations are as follows:

$$\dot{K}_{PA} = \frac{P_2}{P_1} = \frac{\dfrac{1}{j\omega C_A}}{R_A + j\omega L_A + \dfrac{1}{j\omega C_A}}, \tag{2}$$

$$f_{res} = \frac{1}{2\pi\sqrt{L_A C_A}} \quad [Hz] \tag{3}$$

$$PA_{res} = \frac{L_A}{R_A} 2\pi f_{res} . \quad (4)$$

Equivalent electrical circuit is added to fig. 5 of peizo-crystal plate by transformer connection having relevant transducing coefficient *n* and capacity C_P – fig. 6.

The main difficulty is to produce piezo-materials with micro- and nano- dimensions having certain properties. Other challenge is of energy harvester of this type is the design and production of circuit, converting the harvested energy from the piezo – membranes to electric. Unlike the conventional power sources and batteries having lower internal impedance, in the piezoelectric sources the impedance is relatively higher. Numbers of works exist in which this problem is largely solved. Structural scheme of the sound pressure harvesting system is shown on fig. 7.

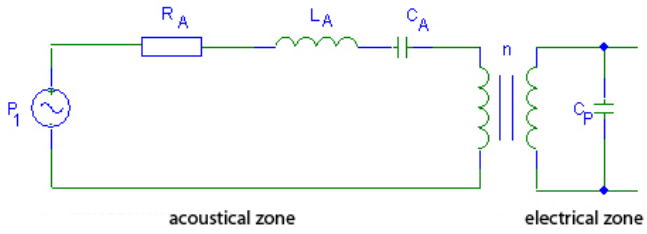


Fig. 6. Piezoelectric MEMS showing the acoustic and electric part of the system

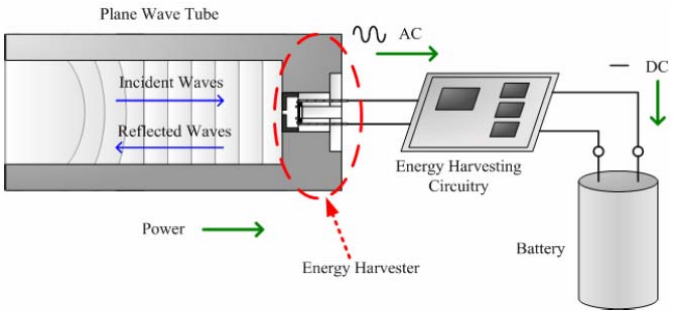


Fig. 7. General concept for sound pressure harvesting

In the resonant cavity, alternating sound pressure is obtained. It derives alternating by sign deformations in the membrane, applied steady to the cavity's bottom connected with a piezo-material. Through the conducting system, the generated alternating voltage is derived to a voltage transducer and amplifying unit. Then the voltage is applied to a consumer or accumulating device.

When designing MEMS with piezoelectric elements is derived from the concept of mechanical energy harvesting, as in this meaning, the building of the systems is divided in three components. The first component consists of composite deformable plate with piezo-element. It transduces the applied deformation in to a driving voltage. The second component is the design of the transducing device in order to achieve maximum sound pressure harvesting. The third component is electrical circuit, which receives the alternating electric signals from the piezoelectric diaphragm and transforms it in direct current. All components are subject to intensive development in order higher harvested energy density to be absorbed.

3. WORKING APPLICATIONS OF THE SOUND PRESSURE HARVESTING TECHNOLOGY

The harvesting devices of the British company Perpetuum, supplied by mechanical oscillations are applied to pumps, gas and oil installations. The energy harvested is used for wireless sensors supply determined periods in advance. With the assistance of Thechnology Partnership Cambridge UK cellular communicational platforms are build having supply without batteries, where the harvested mechanical oscillations supply low power GSM modules [5].

En Ocean has installed in a building 4200 light switches, light sensors and other low power receivers, which are supplied by sound pressure harvesters. That brought 40% reduction of the expenses for artificial light and light automation in the building [5].

Southampton University demonstrated piezoelectric sound pressure harvesters, which are able to ensure $118\,mW$ at a $4.1\,V$ and a load of $40\,k\Omega$ [5].

4. CONCLUSIONS

The main advantages of MEMS for sound pressure harvesting are:

– Environmentally friendly method;

– Application in objects like tunnels, metro stations underground facilities, where the harvesting from other alternative sources is not possible (wind, light, thermal sources, etc.);

– Possibility of using existing elements for modulating of the voltage.

Disadvantages of MEMS for sound pressure harvesting are:

– Low power and random nature of the sound waves;

– Expensive technology used to build the piezo-materials with micro- and nano-sizes.

In conclusion, it could be generalized that synthesis of new MEMS technologies for sound pressure harvesting is a recent challenge for research teams. In order to achieve prevalence will be necessary the development of alternative devices for energy storage.

REFERENCES

[1] Horovitz, St., Development of a MEMS-based Acoustic Energy Harvester, University of Florida, PhD Thesis, 2005.
[2] Li, P. et al., A Magnetoelectric Eenergy Harvester and Management Circuit for Wireless Sensor Network, Elsevier, vol. 157, pp. 100-106, 2009.
[3] Priya, Sh., D. Inman, Energy Harvesting Technologies, Springer science, pp. 7-9, 2007.
[4] Ramadass, Y., A. Chandrakasan, An Efficient Piezoelectric Energy Harvesting Interface Circuit Using a Bias-Flip Rectifier and Shared Inductor, IEEE Journal of Solid-State Circuits, Vol. 45, №. 1, pp. 189-192, January 2010.
[5] http://www.computescotland.com/energy-harvesting-devices-2103.php
[6] Beeby, S.P. et al., Macro and Micro Scale Electromagnetic Kinetic Energy Harvesting Generators, Proc. DTIP 2006, Stresa, Italy, April 26-28.
[7] Adhikari, S., M. I. Friswell and D. J. Inman, Piezoelectric Energy Harvesting from Broadband Random Vibrations, Smart Materials and Structures, Vol. 18, 115005, 2009.

AN ILLUSTRATION OF ELECTROMAGNETIC FIELD ANALOGY IN LABORATORY WORK ON THE BASIS OF NUMERICAL FIELD MODELING

Ilona Iatcheva, Rumena Stancheva, Kancho Ivanov

Department of Theoretical Electrical Engineering, Technical University of Sofia, Kl. Ohridski 8, 1000 Sofia, Bulgaria, phone: + 3592 965 3389, e-mail: iiach@tu-sofia.bg

Abstract. *The paper presents the possibilities and advantages of the application of electromagnetic field computation in the process of education in teaching theoretical electrical engineering. In particular, during the laboratory work, students investigate direct current field between two conductors in conducting medium on the basis of analogy to the electrostatic field of the same electrodes in dielectric. Simultaneously students have ability to use a demonstration file with results of numerical field modeling of the same laboratory experimental setup. The field distribution is investigated for different positions and shapes of the electrodes, using the finite element method and QuickField 5.6 software package. Using the field analogy, calculated conductance between these electrodes is transformed to the capacitance of the same electrodes in the air. The good visualization of the field distribution and the possibilities to find its parameters helps students to better understand the investigated processes and we hope it increases their interest in the field phenomena.*

Keywords: Education, Electromagnetic field, Modeling, Finite element method

1. INTRODUCTION

Important aim of the teaching electrical engineering is to give the students fundamental knowledge and understanding in electromagnetic field processes and phenomena. It is also very important to develop the ability to make proper formulation and intelligent solution of the electromagnetic field problems. But there are a lot of problems in educational process – electromagnetic field theory is not very attractive for the students, and very often obtaining of the analytical solution is too long process, which requires great amount of efforts. Often students see only abstract mathematics without any practical application. This fact decreases their interests and motivation for studying.

Authors' experience shows that courses and materials in which computer analysis of the problems and good visualization of the processes are included are well accepted by students and provoke their interest [1]. Modern numerical methods such as finite element method (FEM), boundary element method (BEM) and etc. are very suitable for modelling of the field problems and processes [2]. The obtained results for field distribution can be detailed visualised and analysed.

The described below laboratory work [3] deals with experimental investigation of direct current field between two electrodes in conducting medium. The experimentally obtained results are transformed on the basis of electrostatic analogy as results for the electrostatic field of the same electrodes in dielectric. For enriching the stu-

dents knowledge additional field computation as demonstration file is included. The file presents results of the FEM analysis of the same as experimentally investigated fields, but for a number of different positions and shapes of the electrodes. The detailed visualization and the possibilities for easy changing the "numerical experiments" can increase the student's interest to the investigated problem and to the process and subject of study.

2. DESCRIPTION OF THE INVESTIGATED PROBLEM IN LABORATORY WORK

The main goal of the considered laboratory work is experimental investigation of the electric field distribution between two rectangular electrodes in electrolyte and determination of some electromagnetic parameters. Additionally the values of these parameters are transformed for determining the pointed by analogy field parameters of the electrostatic field of a long two conductor line in air (Fig. 1).

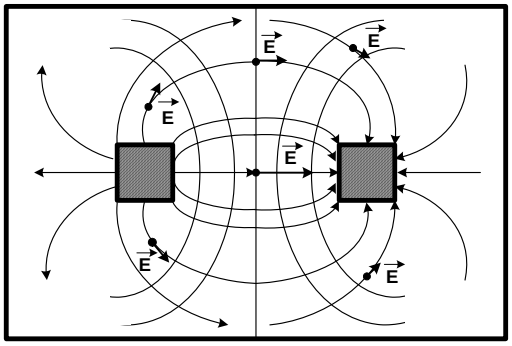


Fig.1. Electrostatic field of a long two conductor line in air

The investigation is based on the existing analogy between electrostatic field in dielectric and electric field of direct current in conductive medium. The analogical field's distribution can be obtained using mathematically similar equations:

- For the electrostatic field in dielectric

$$rot\vec{E} = 0, \quad \vec{D} = \varepsilon\vec{E}, \quad div\vec{D} = 0 \tag{1}$$

- For the electric field of direct current in conductive medium

$$rot\vec{E} = 0, \quad \vec{J} = \gamma\vec{E}, \quad div\vec{J} = 0 \tag{2}$$

where: $\vec{E}$ is the electric field strength;

$\vec{D}$ – the electric field displacement;

$\vec{J}$ – the current density in conductive medium;

ε – electric permittivity of the electrostatic field medium;

γ – conductivity of the electrolyte.

In the equations 1 and 2 the vector $\vec{D}$ from the first one corresponds to the vector $\vec{J}$ from the second. The correspondence also exist between parameters ε and γ, which define the physical properties of two medium. In case of equal shape of the boundaries the field distribution of the electrostatic field in the dielectric and electric field produced by direct current in low conductive medium are completely similar.

The described electrostatic analogy allows determination of the integral parameter, which characterizes one of the fields, using determination of the corresponding parameter from the other field. For example the capacitance *C* of a system of two conductive bodies in a dielectric can be obtained on the bases of measurements of the electrical conductance *G* between electrodes with the same shape, posed in electrolyte. The relation between parameters can be described with Equation 3:

$$C = \frac{\varepsilon}{\gamma} G . \tag{3}$$

So students investigate experimentally the electric field, produced in a tank with electrolyte in which two electrodes with different potentials have been posed (Fig.2). The applied voltage and the positions of the electrodes are the same as the ones of the conductors in air.

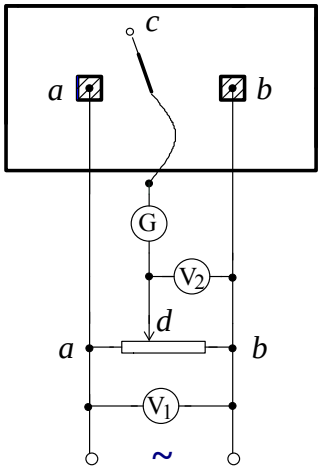


Fig. 2. Experimental setup – tank with electrolyte, in which two electrodes with different potentials have been posed

During the laboratory work the potentials in number of points with fixed coordinates (Fig. 3) are measured.

The electric field strength can be obtained using Equation 4:

$$\vec{E} = -\text{grad}V = \vec{x}_0 E_x + \vec{y}_0 E_y ,$$
$$E_x = -\frac{\partial V}{\partial x}, \quad E_y = -\frac{\partial V}{\partial y}. \tag{4}$$

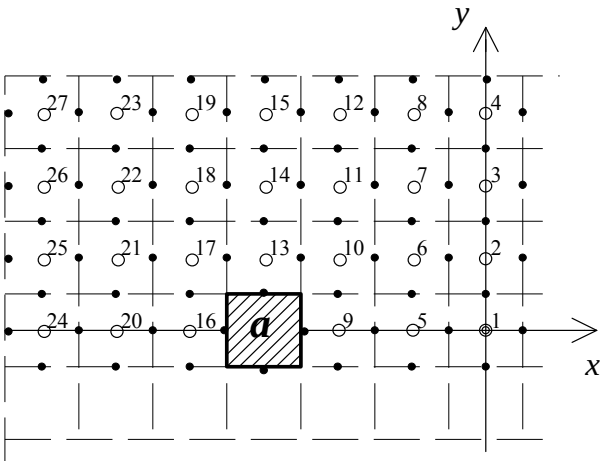
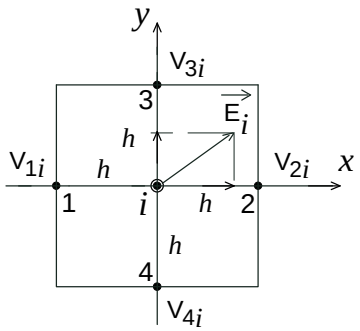


Fig. 3. The points with fixed coordinates, in which has been measured potential

On the basis of this theory students calculate electric field strength in points 1÷27, using finite difference method (Fig. 4)**.** Special attention is paid to the calculation of the electric conductance between the electrodes *G*.



$$E_{xi} = \frac{V_{1i} - V_{2i}}{2h},$$

$$E_{yi} = \frac{V_{4i} - V_{3i}}{2h}$$

Fig. 4. Electric field strength is obtained using finite difference method

3. NUMERICAL FIELD MODELLING AND DEMONSTRATION FILE

The electric field distribution in the experimentally studied region is modeled using Finite Element Method and QuickField 5.6 software package. During the laboratory work students can consider the results of the FEM analysis in a demonstration file. Some of the obtained results are presented in Figures 5,6,7,8 and 9. Figure 5 shows the investigated region, Figure 6 – the finite element mesh, Figure 7 – distribution of the electric potential in the electrolyte and the possibilities to obtain the field values in a point with given coordinates.

In Figure 8 is shown distribution of the electric field strength in case of circle instead of rectangular shape of the conductors in the air. The figure illustrates also the possibility to obtain capacitance of the system using energy or charge calculation.

In Figure 9 are shown color map of electric field strength in the electrolyte (the first field) and the chart of its changing along the fixed line.

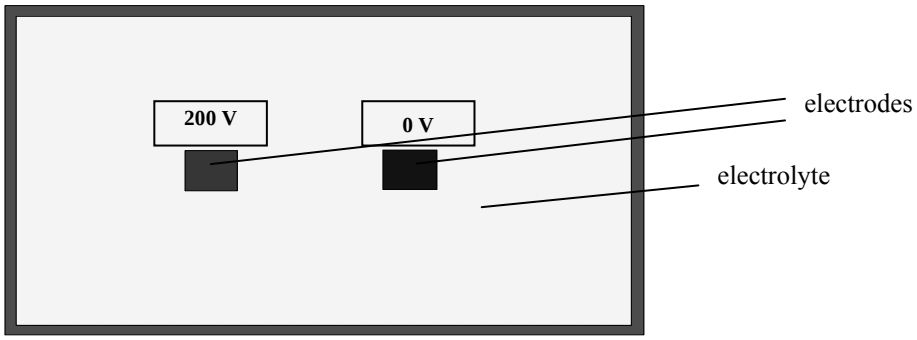


Fig. 5. The investigated region with electrodes

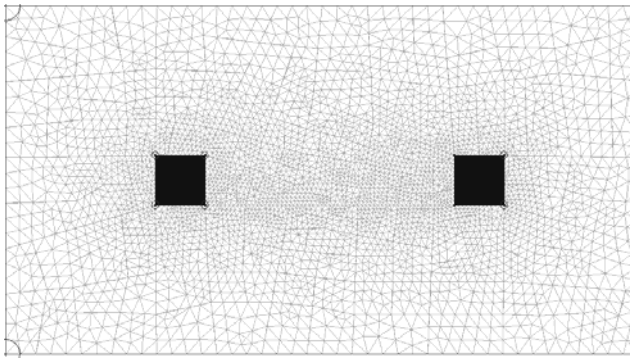


Fig. 6. The finite element mesh in the investigated region

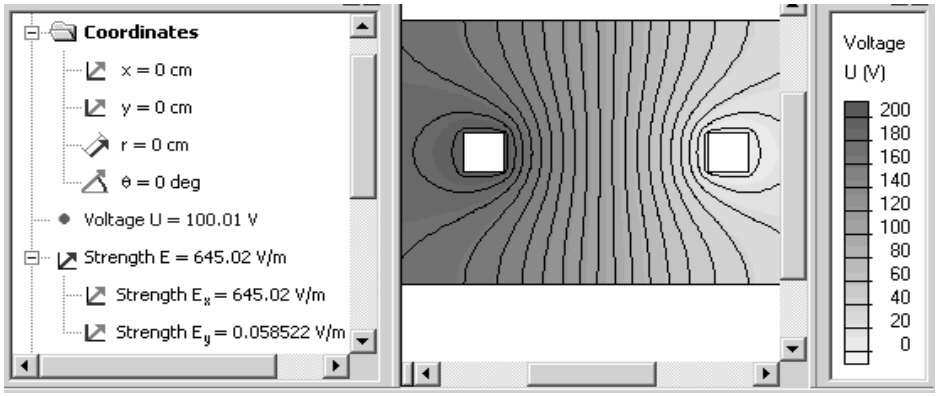


Fig. 7. Distribution of the electric potential in the electrolyte and field values in a point with coordinates x=0, y=0 (point N =1)

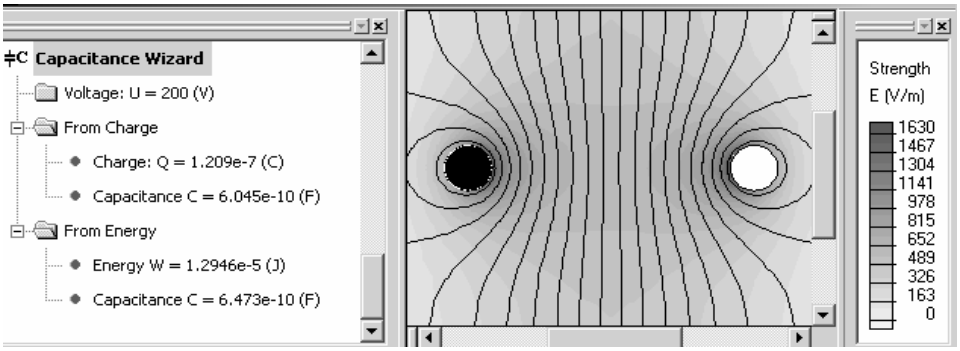


Fig. 8. Distribution of the electric field strength in case of circle shape of the conductors in the air and capacitance determination

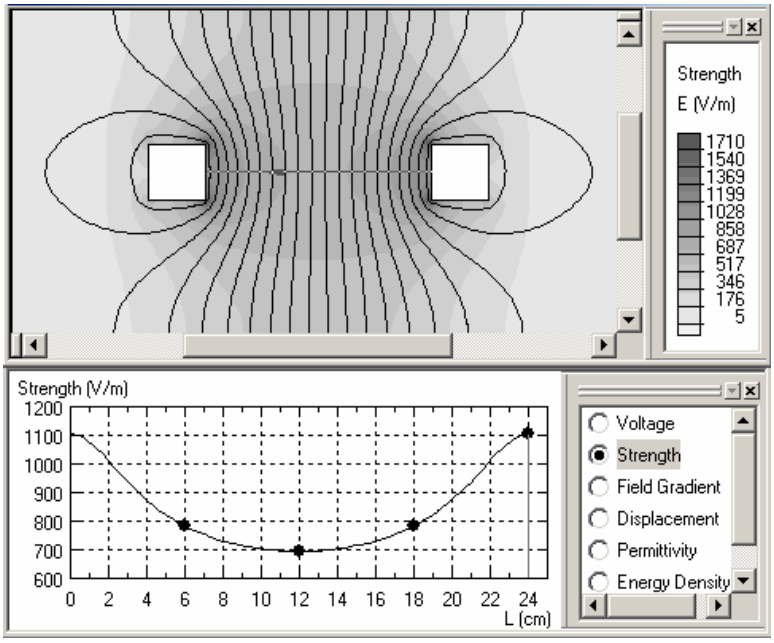


Fig. 9. The map of electric field strength in the electrolyte and the chart of its changing along the fixed line

Our example with described laboratory work shows that dealing with demonstration file including results of numerical field modeling students can realize the possibilities and advantages of numerical field investigation:

- The positions and shapes of the sources and boundaries of the investigated region, as well as boundary conditions can be described with high accuracy.
- The shapes and positions of the sources and boundaries and the material properties can be also easy changed. It means that it is very comfortable to produce a number of different "numerical experiments" and investigate different dependencies between dimensions, geometry, material properties of the investigated region and field distribution.
- As a result the field components can be obtained in arbitrary point and can be compared with measured ones.
- The good visualization of the field parameters helps students better to understand the investigated processes and fields.

4. CONCLUSION

The application of computer field modelling in the laboratory work is a good possibility to show the students that complete investigation of a given phenomenon has to be carried out not only experimentally, but also by numerical simulation. The students can realize the great possibilities and advantages of the numerical field investigation. The authors hope that such laboratory exercise improves the understanding level of the teaching material and increases the quality of education in this area. Students accumulate knowledge and experience and become much more interested in studying.

REFERENCES

[1] I. Iatcheva, R. Stancheva, “An Application of Electromagnetic Field Computation in Teaching Theoretical Electrical Engineering“. Proc. XIth Int. Symp. on Theoretical Electrical Engineering ISTET’01, Linz, Austria, 2001, pp 187-190.

[2] К. Брандиски, И. Ячева, “CAD-системи в електромагнетизма.” СИЕЛА, 2002.

[3] Брандиски, К. и др. Ръководство за лабораторни упражнения по ТЕ, ИК КИНГ, ISBN 954-9518-24-8, София, 2010.

[4] QuickField Finite Element Analysis System, Version 5.6, User’s Guide, Tera Analysis Ltd., 2009.

ИЗХОДЕН СИГНАЛ НА ПЛОСКИ ЕЛЕКТРОМАТНИТНИ ПРЕОБРАЗУВАТЕЛИ ПРИ КОНТРОЛА НА ФЕРОМАГНИТНИ ИЗДЕЛИЯ

Стефчо Георгиев Гунински, Стефан Стоянов Стайков, Калинка Миткова Тодорова

Кат. Обща електротехника, Технически университет – София,
бул. "Кл. Охридски" №8, 1000 София, тел.(+359 2) 965 36 65

Разгледани са плоски параметрични допиращи електромагнитни преобразуватели с формата на "правоъгълна" спирала [1]. Използван е получен в предишна работа аналитичен израз за изходния сигнал и са проведени поредица числени експерименти. Резултатите са използвани за построяване на серия ходографи на внесения импеданс при изменение на характерни величини.

Ключови думи: спирална намотка, параметричен преобразувател

Трайна тенденция в електромагнитния безразрушителен контрол е все по-честото използване на преобразуватели с некласически конструкции [2,3,5]. По-специалната форма и разположение на намотките им дава нови характеристики и редица предимства. Перспективна възможност са плоските спирални намотки. В работата е разгледан електромагнитен преобразувател (ЕМП) със спирална форма. Преобразувателят е параметричен и представлява плоска намотка от тънък проводник с пренебрежими размери на напречното сечение. Намотката има N на брой (незатворени) навивки и формата на "ъгловата спирала". На фиг. 1 са показани само две навивки, а пунктирите от точка М показват как токовият контур се затваря (отсечката МО) или как продължава третата навивка. Разстоянието между страните е $2a$. През ЕМП протича синусоидален възбудителен ток i с ефективна стойност I и ъглова честота $\omega = 2\pi f$.

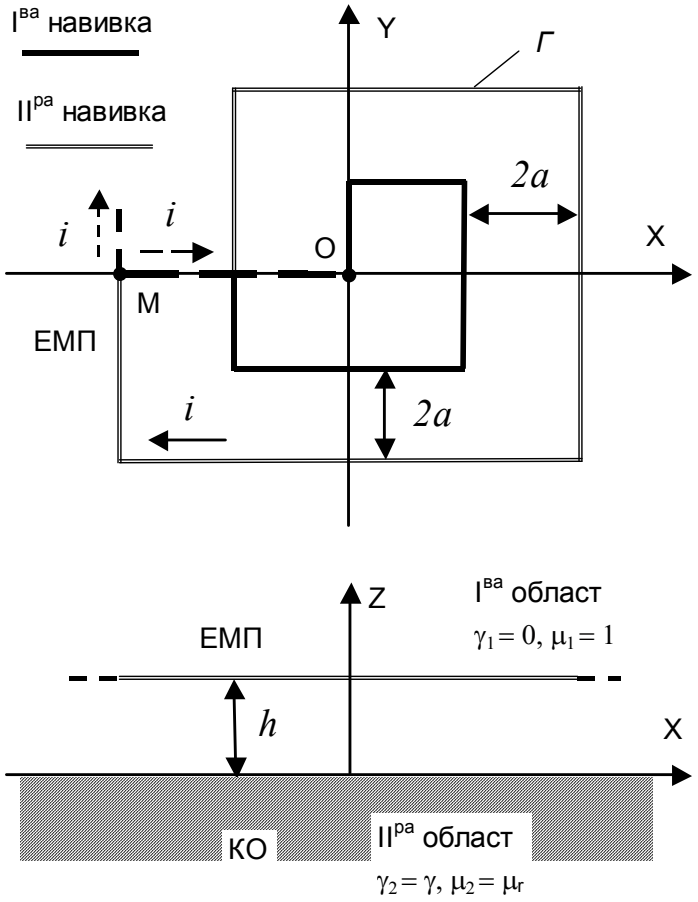
Преобразувателят е разположен на разстояние h над плосък контролиран обект (КО) с размери, позволяващи с достатъчна точност да се разглежда като линейно проводящо феромагнитно полупространство с относителна магнитна проницаемост μ_r и специфична електрическа проводимост γ (фиг. 1).

За да се намерят оптималните условия за контрол и подходящи размери на преобразователя, е необходимо да се проведат редица числени експерименти, като за целта е удобно да се използват получените в [1] резултати за изходния сигнал на преобразувателя.

Тъй като разглежданият ЕМП е параметричен, изходен сигнал е внесеният комплексен импеданс на спираловидната намотка [4]:

$$Z_{вн} = \frac{\dot{E}_{вн}}{\dot{I}} = j\frac{\omega}{\dot{I}}\oint_{\Gamma}\dot{\vec{A}}_{1вн}d\vec{l} \quad , \qquad (1)$$

където Γ е контурът на намотката на ЕМП (фиг. 1).



Фиг. 1. Допиращ параметричен ЕМП със спираловидна намотка, разположен над плосък контролиран обект

Удобно е да се премине към обичайните за вихровотоковия контрол относителни (безразмерни) величини:

$$u = a\lambda,\ v = a\nu$$

– Всички геометрични размери са нормирани спрямо половината от разстоянието между отделните страни на преобразувателя a (фиг. 1), т.е. $h^* = h/a$;

Обобщения параметър $\beta^2 = a\mu_0\gamma\omega$.

Тъй като спираловидната намотка се състои от отсечки, които са успоредни на координатните оси, интегрирането не е свързано с особени трудности, а резултата подлежи на опростяване. В крайна сметка се получава аналитичен израз от вида:

$$Z_{вн} = -2g\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F_f F_{ср} F_n du dv \ . \qquad (2)$$

където:

F_f е функция на *възбудителя*, която основно зависи от параметрите на ЕМП и тока през него:

$$F_f = \omega u^{-2} v^{-2} pQ\tilde{Q} \ , \tag{3}$$

където

$$\tilde{Q} = \cos(N+1)(u+v)\frac{\sin N(u+v)}{\sin(u+v)} - \\ - e^{ju}\cos\left[Nu-(N+1)v\right]\frac{\sin N(u-v)}{\sin(u-v)} + je^{-jNu}\sin Nu \ , \tag{4}$$

$$Q = \exp(-ju)\cos a\left[Nu-(N+1)v\right]\frac{\sin N(u-v)a}{\sin(u-v)a} - \cos(N+1)(u+v)\frac{\sin N(u+v)}{\sin(u+v)} + \\ + j\exp(jNu)\sin Nu \tag{5}$$

$F_{\text{ср}}$ е функция на *средата*, която основно зависи от параметрите на КО и има познатия за случая на полупространство вид [2,3];

F_n е функция на *положението*, която основно зависи от взаимното разположение на ЕМП и КО:

$$F_n = e^{-2ph^*} \ , \ p^2 = u^2 + v^2 \ , \ g = \mathrm{j}\mu_0/4\pi^2 \tag{6}$$

РЕЗУЛТАТИ ОТ ЧИСЛЕНИТЕ ЕКСПЕРИМЕНТИ

Като изходен сигнал се разглежда относителният внесен комплексен импеданс:

$$Z^* = Z_{вн} / X_0 \ , \tag{7}$$

като $Z_{вн}$ се изчислява от израза (2), а индуктивното съпротивление на преобразувателя X_0 при отсъствие на КО – по същата формула, като се положи $F_{ср} = F_n = 1$ и се вземе с обратен знак.

Резултати са онагледени графично чрез:

- ходографите на Z^* - графиките $\mathrm{Im}(Z^*)[\mathrm{Re}(Z^*)]$ при изменение на обобщения параметър β – фиг. 1,2,3.
- зависимостите на относителния внесен импеданс z^* и фазата му Ψ от обобщения параметър β – графиките $z^*(\beta)$ – фиг. 3,5,7 и $\Psi(\beta)$ – фиг. 4,6,8.

Изчисленията са проведени при различни стойности на броя на навивките N на преобразувателя (N = 1, 5, 30), различни разстояния между преобразувателя и обекта (h^* = 0,08; 0,2) и различна относителна магнитна проницаемост на КО (μ_r = 1, 10, 1000).

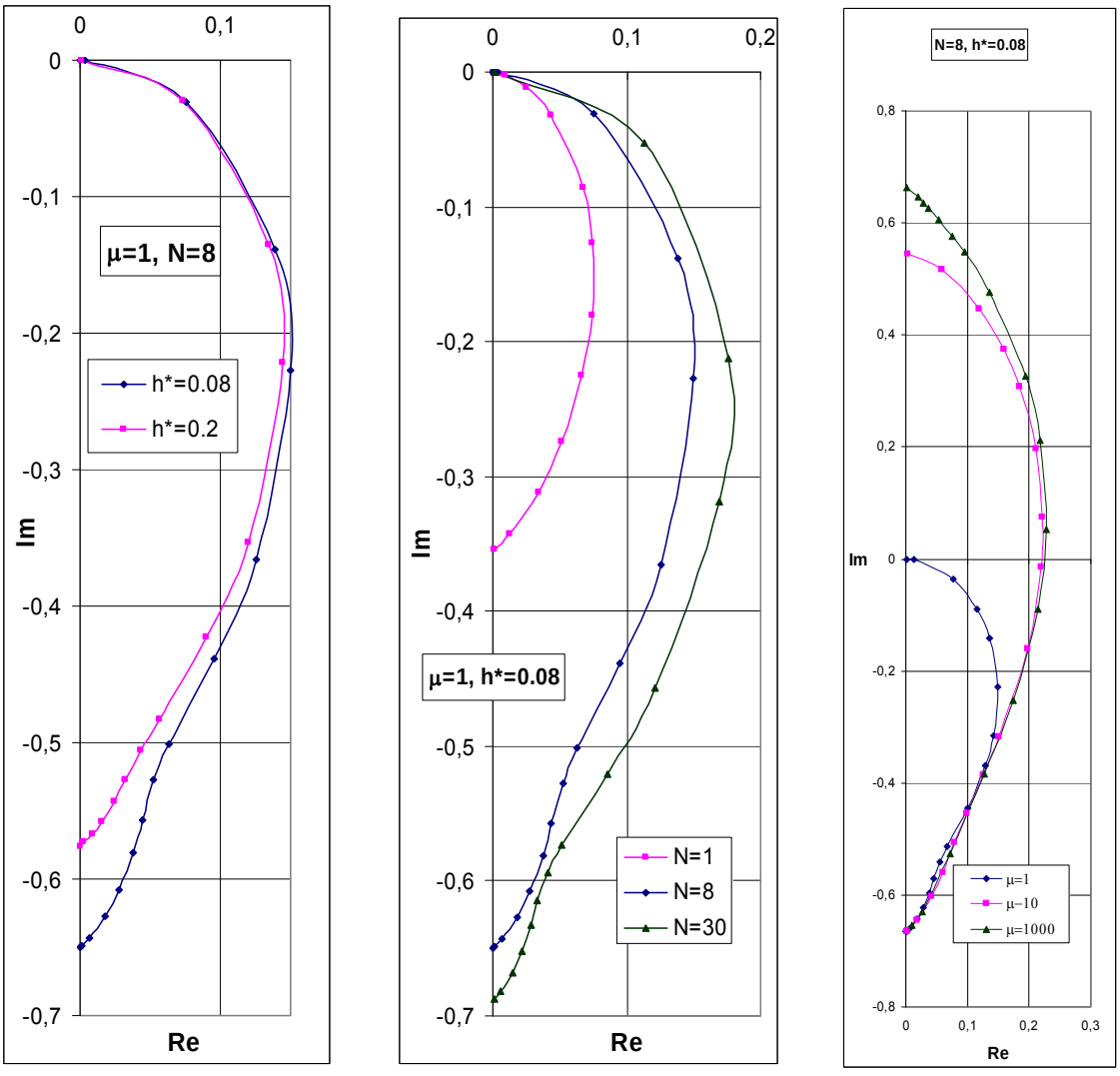
ЗАКЛЮЧЕНИЕ

От получените резултати може да се направят следните по-важни изводи:

- при малко разстояние h между преобразувателя и обекта и по-голям брой на навивките на преобразувателя за големи стойности на обобщения па-

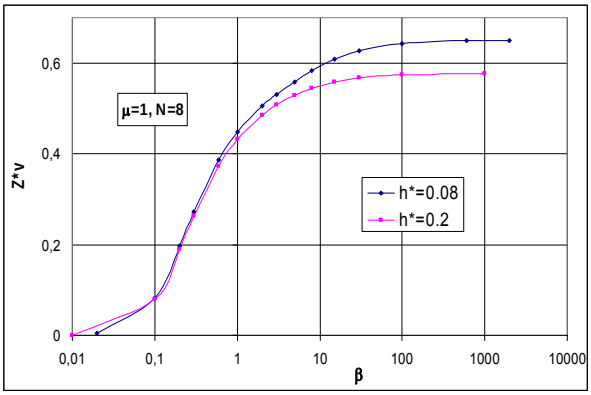
раметър β изменението на реактивното внесено съпротивление е много по-голямо от изменението на активното;

- в близката зона влиянието на отместването между преобразувателя и обекта е сравнително слабо, докато броя на навивките влияе по-значително, макар че при N > 10 това влияние намалява;
- максималната стойност на R^* практически не зависи от h^*, но се увеличава с увеличаването на броя на навивките N;
- с увеличаване на броя на навивките N фазата като функция на обобщения параметър се изменя по-рязко между две относително постоянни нива от 0 и – π/2 rad.

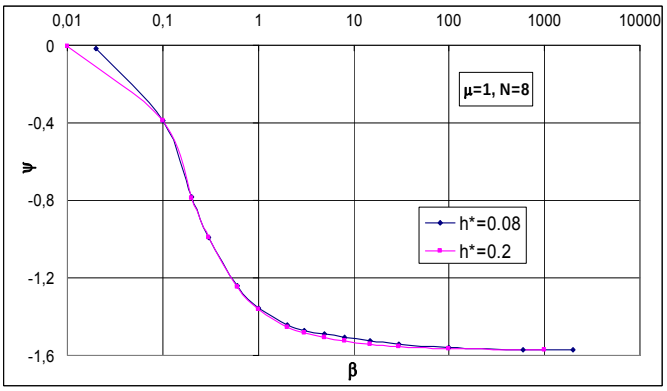


а) б) в)

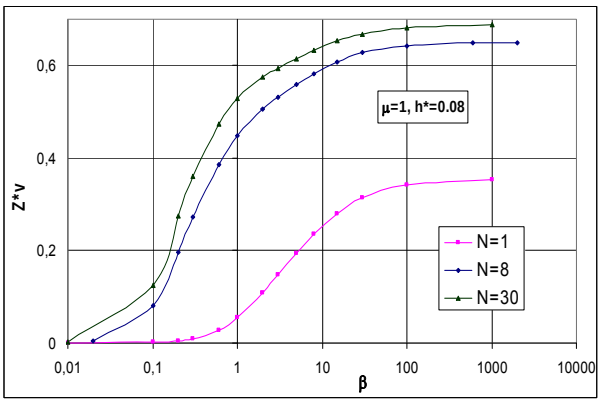
Фиг. 2. Ходограф на относителния внесен импеданс при изменение на *β:*
а) при различни разстояния **h** между преобразувателя и контролирания обект*;*
б) при различен брой на навивките **N** на преобразувателя*;*
в) при различна относителна магнитна проницаемост μ_r



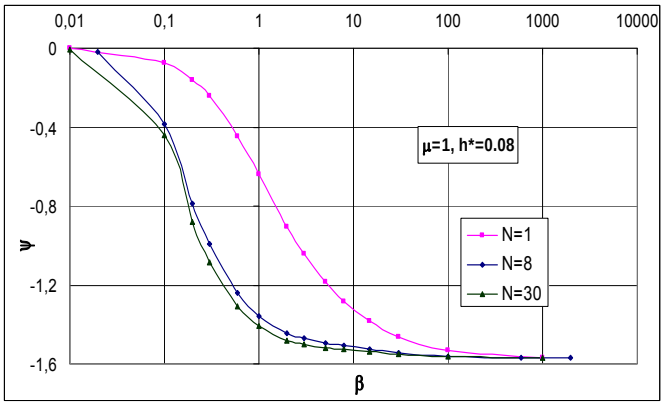
Фиг. 3. Зависимост на внесения импеданс от обобщения параметър β при различни разстояния **h** между преобразувателя и контролирания обект



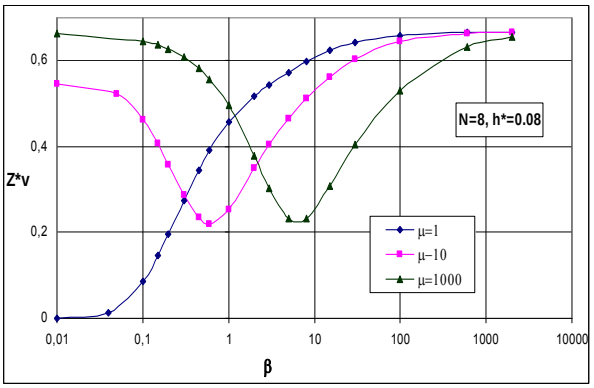
Фиг. 4. Зависимост на фазата ψ на внесения импеданс от обобщения параметър β при различни разстояния **h** между преобразувателя и контролирания обект



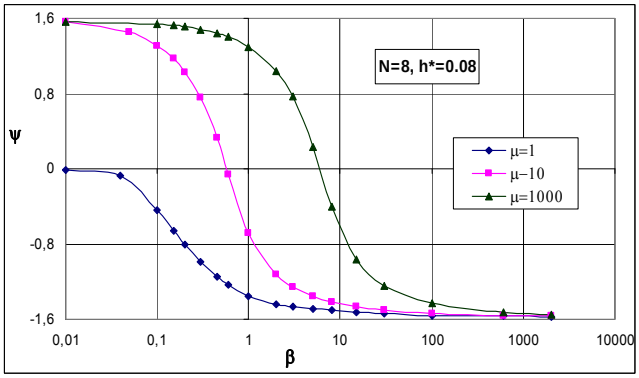
Фиг. 5. Зависимост на внесения импеданс от обобщения параметър β при различен брой на навивките **N** на преобразувателя



Фиг. 6. Зависимост на фазата ψ на внесения импеданс от обобщения параметър β при различен брой на навивките **N** на преобразувателя



Фиг. 7. Зависимост на внесения импеданс от обобщения параметър β при различна магнитна проницаемост μ_r



Фиг. 8. Зависимост на фазата ψ на внесения импеданс от обобщения параметър β при различна магнитна проницаемост μ_r

ЛИТЕРАТУРА

[1] Гунински С., И. Петров, Изходен сигнал на електромагнитен преобразувател със спирална форма, Научни известия на НТС по Машиностроене, год. XII, бр. 1 (79), юни 2005 год., Дефектоскопия‘2005. XX Национална конференция по безразрушителен контрол с международно участие, Созопол, 13-15 Юни 2005 г., с. 243-247.

[2] Гунински С., Вихровотоков преобразувател с триъгълна намотка при контрол на плоски изделия. VI национален научен симпозиум с международно участие "МЕТРОЛОГИЯ и надеждност '95", 19-23 септември 1995 г., Созопол, с. 140-143.

[3] Електромагнитна дефектометрия. Под ред. на Д. Цветков, Техника, София, 1981.

[4] Гунински С., Тримерни задачи на електромагнитния контрол при наличие на движение. Автореф. на докт. дисертация, София, 1991.

[5] W. Hurley, M. Duffy, S. O’Reilly and S. O’Mathuna. Impedance Formulas for Planar Magnetic Structures with Spiral Windings, IEEE Tranc. Of Industrial Electronics, vol 46, NO 2, April 1999, pp. 271-278.

АНАЛИЗ НА СИГНАЛА НА НЕКЛАСИЧЕСКИ ПАРАМЕТРИЧЕН ЕЛЕКТРОМАГНИТЕН ПРЕОБРАЗУВАТЕЛ ЗА ВИХРОВОТОКОВ КОНТРОЛ НА ПРОВОДЯЩА ПЛАСТИНА

Живко Асенов Даскалов, Калинка Миткова Тодорова, Стефчо Георгиев Гунински

Кат. Обща електротехника, Технически университет – София,
бул. "Кл. Охридски" №8, 1000 София, тел.(+359 2) 965 21 36
e-mail: jdaskalov@tu-sofia.bg, ktodorova@tu-sofia.bg, sgg@tu-sofia.bg

Резюме: *В доклада са представени резултатите от анализа на изходния сигнал на един некласически вихровотоков преобразувател, характеризиращ се с това, че възбудителното поле се създава от два съосни кръгови токови контура, през които протичат разместени по фаза синусоидални токове. Разгледан е случая, когато същият е разположен над плоска проводяща феромагнитна пластина с дебелина d.*

Изведени са аналитични изрази за внесените импеданси във всяка от двете намотки, които могат да се използват за определяне на влиянието на електрофизичните и геометрични параметри на изследвания обект при различни ефективни стойности и ъгли на дефазиране на възбудителните токове.

Ключови думи: електромагнитно поле, вихровотоков преобразувател, внесен импеданс

1. ВЪВЕДЕНИЕ

Електромагнитните вихровотокови методи намират широко приложение в практиката на безразрушителния контрол на качеството на проводящи обекти поради доказаните си предимства в сравнение с другите известни физически методи. Едно от възможните направления за усъвършенстване на възможностите им е включването на класическите (традиционните) параметрични вихровотокови преобразуватели (ПВТП) в подходящи измервателни схеми (мостови, резонансни и др.), в резултат на което ходографите на изменение на еквивалентния импеданс на така съставените вериги се получават различни по форма и по разположение в комплексната равнина, в сравнение с ходографите на изменение на внесения импеданс на самия ПВТП [1], [2].

Както е показано в [3], друг възможен подход за постигане на посочените цели е създаването на нови типове некласически вихровотокови преобразуватели (ВТП), които предоставят нови, по-благоприятни възможности за контрол.

2. ПОСТАНОВКА НА ЗАДАЧАТА

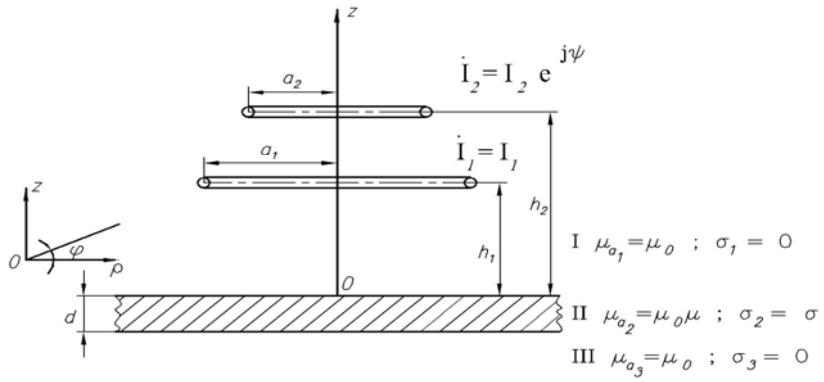
В настоящия доклад се разглежда един нов тип некласически допиращ ВТП от параметричен тип, предложен от авторите в [3], при който за разлика от до-

сега известните, възбудителното поле се създава от две намотки, през които протичат разместени по фаза синусоидални токове. Изследванията на сигнала на този ВТП, разположен над плоска проводяща среда, доказват възможностите му за получаване на различни по вид ходографи на изменение на внесения импеданс в една от намотките на този ВТП, при изменение на фазовата разлика между възбудителните токове в двете намотки, както и при изменение на геометричните им размери и взаимно разположение [3], [4], [5].

Показано е решението на задачата за аналитично определяне на внесените импеданси в двата възбудителни контура на описания по-долу некласически ВТП преобразувател при наличие на реакция от вихровите токове в плоска проводяща феромагнитна пластина при различни ефективни стойности и фазови отмествания между възбудителните токове в двата възбудителни токови контура.

3. ОПИСАНИЕ НА ТЕОРЕТИЧНИЯ МОДЕЛ

На фиг. 1 е показан теоретичният модел на описания по-горе некласически ПВТП от допиращ тип, разположен над проводяща феромагнитна пластина с дебелина ***d*** .



Фиг. 1. Некласически ВТП, разположен над проводяща феромагнитна пластина

Във въздушната среда (област I – за z > 0) с абсолютна магнитна проницаемост $\mu_{a1} = \mu_0$ и специфична електрическа проводимост $\sigma_1 = 0$) са разположени двете намотки на преобразувателя, представени като кръгови токови контури с пренебрежимо малко напречно сечение и с радиуси a_1 и a_2 , отстоящи на разстояния h_1 и h_2 от повърхността на проводяща феромагнитна пластина с дебелина d (област II – за - d < z < 0). Втората област се характеризира с електрофизични параметри : специфична електрическа проводимост $\sigma_2 = \sigma$ и относителна магнитна проницаемост $\mu_{r2} = \mu$. Третата област е също въздушна среда с абсолютна магнитна проницаемост $\mu_{a3} = \mu_0$ и специфична електрическа проводимост $\sigma_3 = 0$). Всички разглеждани области се считат за линейни, хомогенни и изотропни. През двата контура протичат синусоидални токове с кръгова честота $\omega = 2\pi f$ и комплексни ефективни стойности $\dot{I}_1$ и $\dot{I}_2$, дефазирани един от друг на ъгъл ψ .

За удобство при разглеждането се въвеждат следните означения:

$$m = \frac{I_2}{I_1}, \quad n = \frac{a_2}{a_1}, \quad p = \frac{h_2}{h_1} \quad \text{и} \quad \alpha = \frac{2h_1}{a_1} \tag{1}$$

4. ОПРЕДЕЛЯНЕ НА МАГНИТНОТО ПОЛЕ В ТРИТЕ ОБЛАСТИ

Векторът на магнитната индукция $\vec{B}$ се определя с помощта на магнитния вектор-потенциал $\vec{A}$ от известната зависимост:

$$\vec{B} = rot\,\vec{A} \tag{2}$$

При синусоидални възбудителни токове магнитният вектор-потенциал се получава като решение на векторното уравнение на Хелмхолц:

$$\nabla^2 \dot{\vec{A}}_i - k_i^2 \dot{\vec{A}}_i = -\mu_0 \dot{\vec{J}}_B \tag{3}$$

където: $i = 1, 2, 3$ – номер на съответната област;

$k_i^2 = j\omega\mu_0\mu_{r\ i}\sigma_i$, $\omega = 2\pi f$,

μ_{ri} и σ_i са съответно относителните магнитни проницаемости и специфичните електрически проводимости на отделните области.

В разглеждания случай:

$k_1^2 = k_3^2 = 0$, тъй като $\sigma_1 = \sigma_3 = 0$, а $k_2^2 = k^2 = j\omega\mu\mu_0\sigma$, тъй като $\sigma_2 = \sigma$ и $\mu_{r2} = \mu$.

В нашия случай външни възбудителни токове съществуват само в първата област и при въвеждане на цилиндрична координатна система (с оси ρ, φ и z) обемната им гъстота $\dot{J}_B$ се представя чрез обобщените функции на Дирак δ(x) по следния начин:

$$\dot{J}_B = \dot{J}_{B\varphi} = \dot{I}_1\left[\delta(\rho - a_1)\delta(z - h_1) + me^{j\psi}\,\delta(\rho - na_1)\delta(z - ph_1)\right] \tag{4}$$

Отчитайки горната зависимост, при решаване на векторното уравнение (3) комплексната ефективна стойност на магнитния вектор-потенциал в най-общ случай е функция на трите координати ρ, φ и z. Както се вижда от (4) и поради кръговата симетрия на разглеждания модел магнитният вектор-потенциал $\dot{\vec{A}}_i$ има компонента само по φ, зависеща от ρ и z и независеща от φ, т.е. $\dot{A}_{\varphi\ i}(\rho, z) = \dot{A}_i$, поради което уравнение (3) придобива скаларен вид:

$$\frac{1}{\rho}\frac{\partial}{\partial\rho}\left(\rho\frac{\partial\dot{A}_i}{\partial\rho}\right) + \frac{\partial^2\dot{A}_i}{\partial z^2} + \left(\frac{1}{\rho^2} - k_i^2\right)\dot{A}_i = -\mu_0\dot{J}_{ei} \tag{5}$$

Следователно уравнение (3) се свежда до две скаларни частни диференциални уравнения, решаването на които е удобно да се извърши чрез интегралната трансформация на Ханкел по координатата ρ .

Магнитният вектор-потенциал $\dot{A}_1$ във въздушната област I може да се представи във вида:

$$\dot{\vec{A}}_1 = \dot{\vec{A}}_{10} + \dot{\vec{A}}_{1BH}, \tag{6}$$

където: $\dot{\vec{A}}_{10}$ е магнитният вектор-потенциал на възбудителното поле при отсъствие на проводящата среда, а $\dot{\vec{A}}_{1BH}$ е съставка, отразяваща изменението му от индуктираните вихрови токове във внесена в полето проводяща среда.

Решението по-нататък се извършва по начина, описан в [6], с тази разлика, че след въвеждане на т.н. "обобщен параметър" $\beta_0 = a_1\sqrt{\omega\mu_0\sigma}$ и преминавайки към променлива x изразът за т.н. „функция на средата" F_{cp} за разглеждания случай на плоска проводяща феромагнитна пластина с относителна дебелина $d^* = \dfrac{2d}{a_1}$, както е показано в [7], се представя във вида:

$$F_{cp}(x) = \frac{\left[x^2\left(\mu^2-1\right) - j\mu\beta_0^2\right]\left(1 - e^{-\sqrt{x^2+j\mu\beta_0^2}.d^*}\right)}{\left[x^2\left(\mu^2+1\right) + j\mu\beta_0^2\right]\left(1 - e^{-\sqrt{x^2+j\mu\beta_0^2}.d^*}\right) + 2\mu x\sqrt{x^2+j\mu\beta_0^2}\left(1 + e^{-\sqrt{x^2+j\mu\beta_0^2}.d^*}\right)} \tag{7}$$

5. ОПРЕДЕЛЯНЕ НА ВНЕСЕНИЯ ИМПЕДАНС В НАМОТКИТЕ НА ВИХРОВОТОКОВИЯ ПРЕОБРАЗУВАТЕЛ

Използвайки израза за "внесения" магнитен вектор-потенциал в първата област, определен в [6], за внесените импеданси съответно в първата и втората намотка на разглеждания некласически вихровотоков преобразувател се получават следните изрази:

$$Z_{1BH} = \mathrm{Re}\left[Z_{1BH}\right] + j\,\mathrm{Im}\left[Z_{1BH}\right] = R_{1BH} + jX_{1BH} =$$
$$= \omega\pi a_1\mu_0\left\{\int_0^\infty F_{cp}(x)J_1(x)\left[e^{-\alpha x}J_1(x) + mne^{j\psi}e^{-0,5(p+1)\alpha x}J_1(nx)\right]dx\right\} \tag{8}$$

$$Z_{2BH} = \mathrm{Re}\left[Z_{2BH}\right] + j\,\mathrm{Im}\left[Z_{2BH}\right] = R_{2BH} + jX_{2BH} =$$
$$= \omega\pi a_1\mu_0\frac{n}{m}\left\{\int_0^\infty F_{cp}(x)J_1(x)\left[e^{-0,5(p+1)\alpha x}J_1(x) + mne^{j\psi}e^{-p\alpha x}J_1(nx)\right]dx\right\} \tag{9}$$

6. ЗАКЛЮЧЕНИЕ

От получените по-горе резултати се вижда, че аналитичните изрази, определящи изходния сигнал на преобразувателя (т.е. изразите за внесените импеданси в двата токови контура на преобразувателя) представляват достатъчно сложни по структура функционални зависимости от немалък брой влияещи параметри на възбудителя и на изследвания обект.

Това налага извършването на числени експерименти за изследване на влиянието на отделните параметри на възбудителя (в случая: α, m, n, p и ψ) и на изследвания обект (в случая : μ, σ и d) чрез изчисляване на сигнала на преобразувателя за достатъчен брой подходящи стойности на влияещите параметри при изменение на стойността на обобщения параметър β_0 .

ЛИТЕРАТУРА

[1] Цветков, Д. Н. и др. Електромагнитна дефектометрия, стр. 82 - 87. Техника, София, 1981.

[2] Даскалов, Ж. А. Изследване на взаимодействието на електромагнитно поле с многослойна феромагнитна среда. Докторска дисертация, стр. 101 - 115. Технически университет - София, 1984.

[3] Даскалов, Ж. А. Безразрушителен контрол на проводящи обекти с помощта на нов тип вихровотоков преобразувател. VIII Национална конференция с международно участие по безразрушителен контрол “Дефектоскопия′93”, стр. 31 - 33. София, май 1993.

[4] Даскалов, Ж. А., С. Г. Гунински и П. З. Костадинов. Анализ на сигнала на един некласически вихровотоков преобразувател за контрол на феромагнитни обекти. IX Национална конференция с международно участие по безразрушителен контрол “Дефектоскопия′94”, стр. 106 – 109, Созопол, 1994.

[5] Daskalov, J. A. Electromagnetic Control of Conductive Objects Using a New Type of Eddy Current Transducer. Non-Destructive Testing An Eastern-Western European Perspective. Edited by B. Bridge, A. Khalid and B. Yochev and published by The British Institute of Non-Destructive Testing. ISBN 0 903132 35 7, p. 202-206. September 1998.

[6] Даскалов, Ж. А. и С. Г. Гунински. Изследване на полето на вихровотоков преобразувател с две възбудителни намотки при наличие на реакция на проводяща неферомагнитна среда. XIV Национален Научен Симпозиум с международно участие “Метрология и метрологично осигуряване*2004”. Сборник доклади, стр. 157 ÷ 162. Созопол, септември 2004.

[7] Тодорова, К. М., Ж. А. Даскалов, С. Г. Гунински и П. З. Костадинов. Анализ на едно приближение на израза за сигнала на вихровотоков преобразувател, разположен над плоска многослойна среда. XVIII Национален Научен Симпозиум с международно участие “Метрология и метрологично осигуряване*2008”. Сборник доклади, стр.137÷ 141. Созопол, 10 – 14 септември 2008.

ВЛИЯНИЕ МЕШАЮЩИХ ФАКТОРОВ НА ВЫДЕЛЕНИЕ И ПАРАМЕТРИЗАЦИЮ СИГНАЛОВ ОТ ДЕФЕКТОВ ПРИ КОНТРОЛЕ ТОТ ПГ

Щукис Е., Лунин В., Жданов А.

МЭИ (Московский Энергетический институт), 111250, Москва, Красноказарменная ул., 14, 362-73-83, ShchukisEG@gmail.com, Valery.Lunin@mtu-net.ru, ZhdanovAG@mpei.ru

Резюме: *В статье отмечены трудности при анализе данных многочастотного вихретокового контроля теплообменных труб парогенераторов, обусловленные наличием множества мешающих факторов и шума. Предложен алгоритм выделения сигналов от дефектов на свободном от конструктивных элементов участке на основе дискретного вейвлет-преобразования.*

Abstract: *The article describes interfering factors influence on eddy current testing signals and problems during analysis this signals. The algorithm of diagnostically significant areas isolation was implemented via discrete wavelet transform.*

Keywords: discrete and continuous wavelet transform, eddy-current testing, signal processing.

1. ВВЕДЕНИЕ

Одной из важнейших составляющих безопасной эксплуатации энергоблока АЭС с водо-водяным энергетическим реактором (ВВЭР) является работоспособность парогенераторов (ПГ). Практика эксплуатации горизонтальных ПГ показала, что основным элементом, определяющим фактический срок службы этого объекта, являются теплообменные трубы (ТОТ).

Анализ повреждения теплообменных труб парогенераторов, изготовленных из аустенитной нержавеющей стали, показал, что наиболее вероятными причинами выхода их из строя являются: коррозионное растрескивание под напряжением (КРН), язвенная и питтинговая коррозия.

Для оценки состояния теплообменных труб ПГ в процессе эксплуатации проводится контроль их целостности. По результатам контроля, проводится превентивное глушение дефектных труб, что позволяет избежать возможного развития обнаруженного дефекта во время эксплуатации до сквозного и, соответственно, течи из перового контура во второй, что приводит к внеплановой остановке реактора и, следовательно, к экологическим и экономическим потерям.

Результаты контроля становятся основой для принятия решения об остаточном ресурсе ПГ и необходимости его замены. Опыт эксплуатации показал, что наиболее эффективным методом контроля является вихретоковый метод контроля (ВТК). Без ВТК своевременно обнаружить активизацию коррозионных процессов в ПГ не удается.

2. МНОГОЧАСТОТНЫЙ ВИХРЕТОКОВЫЙ КОНТРОЛЯ

Вихретоковый контроль основан на анализе взаимодействия внешнего электро-магнитного поля, которое создается индуктивной катушкой с переменным током, с электромагнитным полем вихревых токов, наводимых в объекте контроля (ОК) этой катушкой. Распределение и плотность вихревых токов определяется источником электромагнитного поля, геометрическими и электрофизическими параметрами ОК. В результате действия в ОК вихревых токов изменяется активное и индуктивное сопротивление катушки (вихретокового преобразователя ВТП). Таким образом, определяя активное и индуктивное изменение сопротивления преобразователя, можно судить о геометрических параметрах объекта контроля.

Метод многочастотного вихретокового контроля (ВТК) в общем случае позволяет зафиксировать наличие дефекта, его местоположение, глубину. Эти параметры определяются при использовании внутреннего проходного датчика, который является на данный момент наиболее производительным.

Следует отметить, что анализ диагностических сигналов с выхода датчика сопряжен со значительными трудностями, обусловленными наличием как множества мешающих факторов, таких как сигналы от конструктивных элементов, так и шумов наложенных на сигнал.

3. ПРИМЕНЕНИЕ ДИСКРЕТНОГО ВЕЙВЛЕТ-АНАЛИЗА ДЛЯ ВЫДЕЛЕНИЯ ДИАГНОСТИЧЕСКИ ЗНАЧИМЫХ ОБЛАСТЕЙ СИГНАЛА ПРИ КОНТРОЛЕ ТОТ

Существующие методы выделения в информационном сигнале диагностически значимых областей базируются на пороговом принципе: участок сигнала превышающий определенный уровень рассматривается как диагностически значимая область. Такой подход к определению диагностически значимых областей имеет существенный недостаток, заключающийся в том, что в этом случае к ним наряду с областями, соответствующими дефектам труб, относятся области, соответствующие конструктивным элементам (дистанционирующим решеткам, местам изгибов теплообменных труб), так как участки сигнала, соответствующие этим элементам, также превышают пороговый уровень.

В связи с этим для надежного обнаружения участков сигнала, соответствующих дефектам труб, необходимо сначала выделить участки сигнала, соответствующие конструктивным элементам, а затем по отдельности анализировать свободные участки и участки с конструктивными элементами на предмет обнаружения диагностически значимых областей, а именно дефектов.

Метод обнаружения сигналов от конструктивных элементов, а именно дистанционирующих решеток, был описан автором в публикации [1].

Методика выделения сигналов от дефектов на свободном от конструктивных элементов участке базируется на дискретном вейвлет-преобразование.

Для разработки методики и алгоритма надежного выделения сигналов от дефектов были выполнены следующие действия: смоделированы сигналы от дефектов; проанализированы спектры сигналов и определен диапазон изменения частоты диагностически значимых областей сигнала; проанализированы мешающие факторы; выполнен дискретный вейвлет-анализ исследуемых сигналов с последующей пороговой обработкой.

Модели сигналов, соответствующие дефектам, получены путем моделирования процедуры вихретокового контроля теплообменных трубок парогенератора с помощью программы конечно-элементного моделирования MagNum3D [2].

Анализ процесса на выходе вихретокового датчика показал, что основными видами помех, накладываемых на полезный сигнал являются: тепловой шум, обусловленный тепловым движением носителей заряда; внешние помехи от электромагнитных процессов, имеющих место в атмосфере, промышленных установках, трансформаторе и т.д.; пильгер-шум, обусловленный несовершенством технологии изготовления труб; тренд, возможность появления которого обусловлена смещением датчика относительно продольной оси, смещением нуля.

Модель диагностического сигнала от дефектов труб представляется набором сигналов от дефектов различного рода с аддитивно наложенными на них сигналами от помех. Аналитическое выражение, описывающее эту модель, имеет вид:

$$Sd(t) = D1(t1) + D2(t2) + D3(t3) + Trend(\mathrm{t}) + Noise(t) + StructureNoise(t), \qquad (1)$$

Где D1(t1), D2(t2), D3(t3) – сигналы от дефектов,

StructureNoise (t) – пильгер-шум,

Trend – тренд, обусловленный смещение датчика относительно продольной оси,

Noise (t) – белый гауссовский шум,

t1 = 200, t2 = 500, t3 = 1500 – координаты сигналов от дефектов,

t = 2000, длительность сигнала.

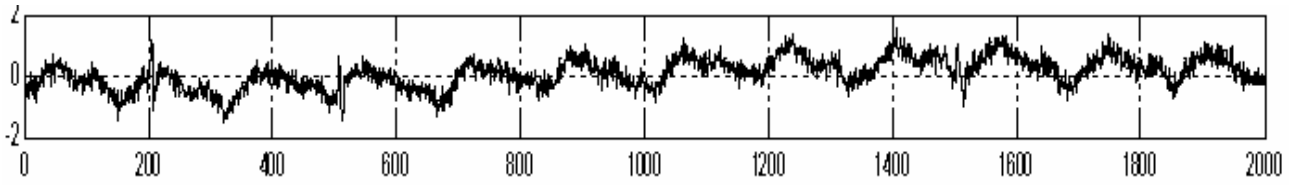


Рис. 1. Модель диагностического сигнала

Вейвлет-анализ позволяет анализировать исследуемый сигнал на различных уровнях разрешения, т.е. анализировать различные версии сигнала, каждой из которой соответствует определенный диапазон частот.

Учитывая, что сигналу от дефекта соответствует определенный диапазон частот, появляется возможность выполнять операцию выделения сигналов от

дефектов на базе версий сигнала, диапазон частот которых совпадает с диапазоном частот этих областей, что, несомненно, повысит эффективность операции выделения.

Для выделения сигналов от дефектов был выбран материнский биортогональный вейвлет bior4.4. В результате разложения исследуемого сигнала с помощью прямого алгоритма Малла были получены вейвлет-спектры, соответствующие аппроксимациям и детализациям исследуемого сигнала на различных уровнях разрешения. Анализ вейвлет-спектров различных уровней показал, что достаточно рассматривать первые 8 уровней сигнала, так как дальнейшее разложение не несет значимой информации о дефектах.

Далее по полученным вейвлет-спектрам аппроксимаций и детализаций соответствующих различным уровням разрешения, с помощью обратного алгоритма Малла выполнялось формирование версий аппроксимаций и детализаций сигнала, соответствующих различным уровням разрешения.

Для повышения отношения энергии сигнал/шум в восстановленных версиях в предлагаемом методе предусматривается пороговая обработка (трешолдинг), которая заключается в том, что приравниваются нулю вейвлет-коэффициенты, значения которых не превышают значения порога τ.

Значение порога с использование минимаксного критерия вычисляется для каждого уровня разложения отдельно [3]. Затем пороговое значение корректируется, на каждом уровне умножается на СКО коэффициентов вейвлет-спектра соответствующего уровня.

В качестве примера на рисунке 2 показана версия детализации для четвертого уровня разрешения после пороговой обработки.

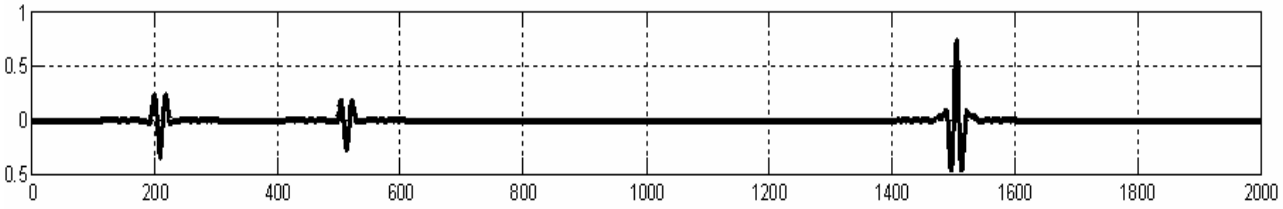


Рис. 2. Восстановленная детализации 5-ого уровня после пороговой обработки

Определение координат диагностически значимых областей выполняется по версиям детализаций, соответствующим частотным областям, в которых эти версии характеризуются наибольшим отношением энергии сигнал/шум.

Следует отметить, что координаты диагностически значимой области, полученные по вышеуказанной версии перекрывают координаты реальной области сигнала от дефекта. В связи с этим необходимо выполнить уточнение координат диагностически значимой области, определенных по принятой версии. Для этого определяют координату максимума восстановленного сигнала и приняв ее за центр диагностически значимой области уточняют координаты последней. На рисунке 3 выделены рамками участки диагностического сигнала, соответствующие дефектам.

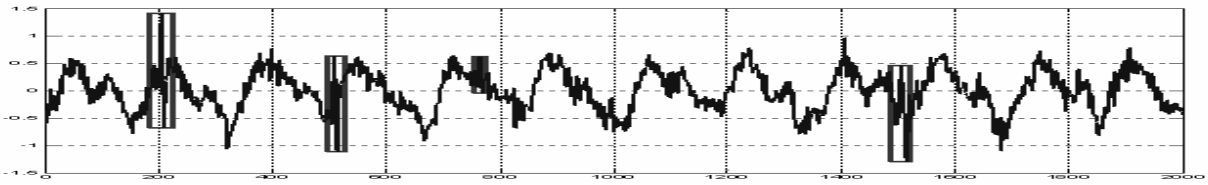


Рис. 3. Автоматическое выделение областей, относящихся к дефектам
на диагностическом сигнале Sd(t)

Проверка работы алгоритма на сигналах от трубок парогенератора показала хороший результат, рамки устанавливаются верно на сигналы от дефектов.

4. ПАРАМЕТРИЗАЦИЯ СИГНАЛОВ ОТ ДЕФЕКТОВ

После выделения сигналов от дефектов производится параметризация сигналов от дефектов, т.е. определяется глубина дефекта.

Получаемый с датчика вихретоковый сигнал комплексный, выше для выделения сигналов от дефектов были рассмотрены только действительные составляющие сигналов.

При многочастотном вихретоковом контроле теплообменных трубок парогенератора основным параметром для определения глубины дефекта является фаза сигнала от дефекта. Фаза сигнала определяется как угол между действительной полуосью на комплексной плоскости и вектором, проведенным из центра координат в точку с максимальным модулем сигнала. Пример показан на рисунке 4.

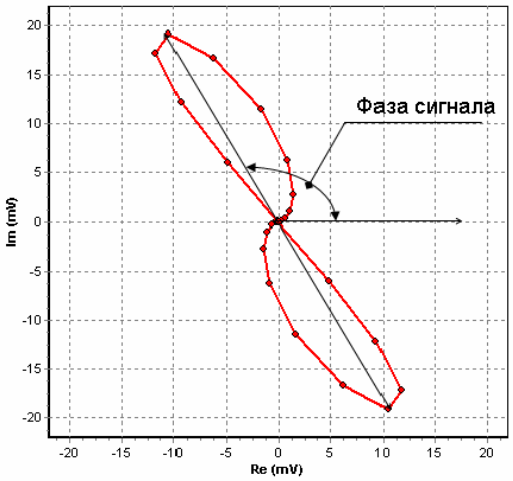


Рис. 4. Фаза сигнала от дефекта

Определение глубины дефекта производится по калибровочной зависимости. Для каждой частоты контроля своя калибровочная зависимость, она строится по фазам сигналов от дефектов с известными геометрическими параметрами. Вихретоковый контроль трубок ПГ проводится на трех частотах 280, 130 и 60 кГц. Пример калибровочной зависимости на частоте 280 кГц приведен на рисунке 5.

Фаза сигнала от дефекта может изменяться в зависимости от мешающих факторов. Проанализируем влияние шума на изменение фазы сигнала от дефекта.

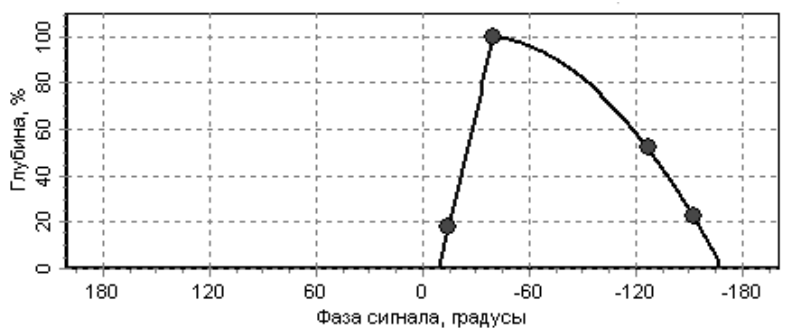


Рис. 5. Калибровочная зависимость

В предположении, что сигнал от дефекта с шумом имеет аддитивный характер, сложим шум и диагностический сигнал от дефекта и проанализируем изменение фазы сигнала от дефекта.

Рассмотрим сигналы от сквозных дефектов с различным максимальным значением: 1В, 0.5В, 0.3В, 0.1В. Сигнал от дефекта будем с равномерным шагом перемешать вдоль сигнала от шума. Рассчитаем на всех частотах максимальную, среднюю и минимальную абсолютную погрешность определения фазы сигнала от дефекта. Погрешность вычисляется как разность фаз не зашумленного сигнала от дефекта и сигнала от дефекта с шумом. Также приводится максимальная, минимальная и средняя вычисленная глубина дефекта, определяемая по калибровочной зависимости.

Если отношение энергии сигнал/шум меньше трех, погрешность определения фазы может превысить 40 градусов и глубину дефекта определять нельзя.

Если отношение энергии сигнал/шум больше трех, то максимальная погрешность определения фазы не превышает 6 градусов, а погрешность определения глубины 20% от толщины теплообменной трубки, результаты для всех частот приведены в Табл. 1.

Таблица 1. Абсолютная погрешность определения фазы и глубины

	F1 (280 кГц)	
	Фаза	Глуб.
Сред.	-0.85	96.69
Max	2.74	99.97
Мин	-5.28	82.39
	F2 (130 кГц)	
	Фаза	Глуб.
Сред.	-0.78	96.27
Max	0.60	99.88
Мин	-4.43	81.76
	F3 (60 кГц)	
	Фаза	Глуб.
Сред.	-1.24	89.46
Max	-0.38	96.80
Мин	-2.58	78.13

ЛИТЕРАТУРА

[1] Е. Щукис, В. Лунин. Вейвлет-анализ вихретоковых сигналов при контроле труб парогенераторов. Сборник докладов Национального Симпозиума с Международным участием “Метрология и Метрологическое Обеспечение 2008”, (Созополь), ТУ – София, Болгария 2008.

[2] Жданов А. Г., Лунин В. П. Банк вихретоковых сигналов от дефектов теплообменных труб. 9-я междун. науч.-техн. конф. студентов и аспирантов «Радиоэлектроника, электротехника и энергетика»: тез. докл.—Т. 1.— М.: Изд-во МЭИ, 2003.— С. 395.

[3] Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в MATLAB. – М.: ДМК Пресс, 2008.- 448с.

ПРИМЕНЕНИЕ ВИХРЕТОКОВЫХ ПРЕОБРАЗОВАТЕЛЕЙ В МЕТАЛЛУРГИЧЕСКОМ ПРОИЗВОДСТВЕ

Катанкин Р. А., Покровский А. Д., Сорокин А. Н.,
Терехин И. В., Шитиков В. С.

В настоящее время для целей технологического контроля в металлургическом производстве в основном применяются оптические и радиационные методы. Оптические методы неэффективны на станах горячей прокатки и при непрерывной разливке стали. Недостатками радиационных методов контроля являются радиационная опасность, сложность в эксплуатации, необходимость применения ряд эталонов и требование двухстороннего доступа. Другим существенным минусом является чувствительность радиоизотопных датчиков к шлакообразующей смеси, периодически добавляемой в кристаллизатор. В этой работе рассматриваются результаты разработки вихретокового контроля для измерения уровня жидкого металла в гильзовых, тонкослябовых и блюмовых кристаллизаторах,

Наиболее качественная сталь для изготовления металлопроката производится непрерывным способом с помощью гильзового кристаллизатора (сечением до 200 х 200 мм), конструкция которого показана на рис. 1. Жидкий металл непрерывно поступает сверху в охлаждаемую водой медную гильзу, где кристаллизуется. Качество отливаемого металла напрямую связано со стабильностью поддержания уровня жидкого металла в гильзе. Сегодня точность подержания уровня составляет ±5 мм.

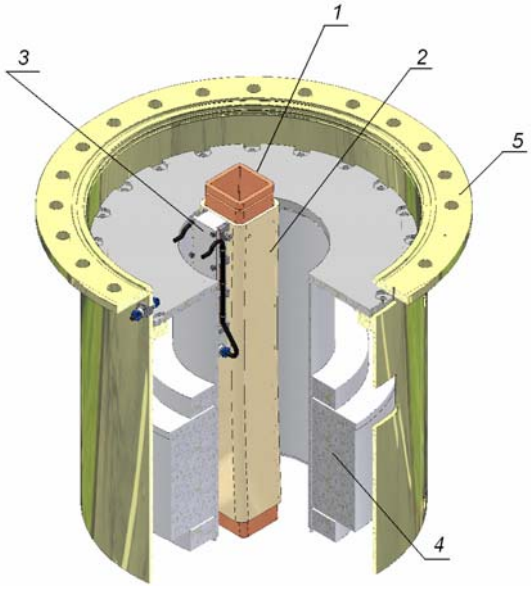


Рис. 1 – Гильзовый кристаллизатор:
1 – медная гильза, 2 – рубашка, 3 – вихретоковый датчик уровня,
4 – электромагнитный перемешиватель, 5 – корпус кристаллизатора

Для измерения уровня сейчас применяются только радиоизотопные датчики уровня. В качестве альтернативы радиоизотопному датчику уровня в НПХ «Техноап» и МЭИ разрабатывается вихретоковый преобразователь для измерения уровня металла. В гильзовом кристаллизатора вихретоковый преобразователь может быть установлен только сбоку от гильзы, как показано на рис. 1. Основными мешающими факторами при измерении уровня металла при таком расположении преобразователя являются помеха от электромагнитного перемешивателя стали, создающего вращающееся магнитное поле, и сигнал вихретокового преобразователя от медной стенки гильзы толщиной 10 – 15 мм. При разливке гильза нагревается, и сигнал вихретокового преобразователя от нагрева гильзы превосходит сигнал от наличия стали в кристаллизаторе более чем на порядок. Дополнительной сложностью является необходимость точной настройки вихретокового преобразователя в условиях многообразия типоразмеров гильз.

В НПХ «Техноап» было предложено использовать для измерения уровня металла вихретоковый преобразователь, содержащий набор измерительных катушек, расположенных внутри катушки возбуждения. Использование такой конструктивной схемы позволило перейти от измерения амплитуды и фазы вносимого напряжения к анализу формы распределения вносимого напряжения по измерительным катушкам. Пример такого распределения показан на рис. 2. Для приведенной конструкции преобразователя положение уровня металла определяется по положению максимума распределения амплитуды вносимой ЭДС. В результате появилась возможность уйти от необходимости калибровать вихретоковый преобразователь, а также возможность использовать при обработке алгоритмы пространственной фильтрации распределения.

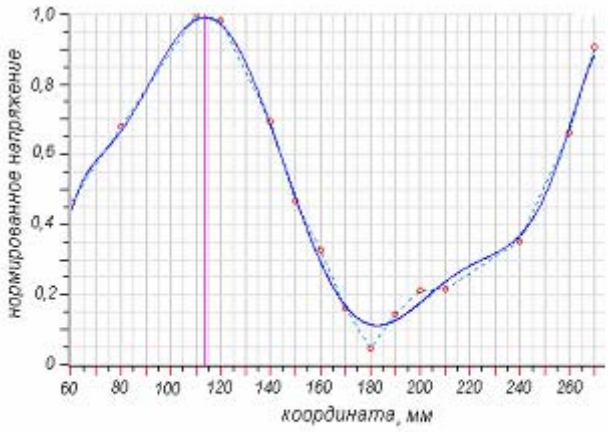


Рис. 2 – Рабочая характеристика вихретокового преобразователя

Было проведено математическое и лабораторное моделирование вихретокового преобразователя, опытный образец вихретокового преобразователя прошел испытания на металлургическом комбинате («Северсталь», г. Череповец).

В ходе промышленных испытаний синхронно фиксировались сигналы вихретокового и штатного радиоизотопного преобразователей. Анализ сигналов в отсутствие электромагнитного перемешивателя (рис. 3) показал, что по техническим параметрам вихретоковый преобразователь не уступает радиоизотопному преобразователю. Вихретоковый преобразователь имеет лучшие динамические свойства. Также подтвердилась нечувствительность вихретокового преобразователя к шлакообразующей смеси.

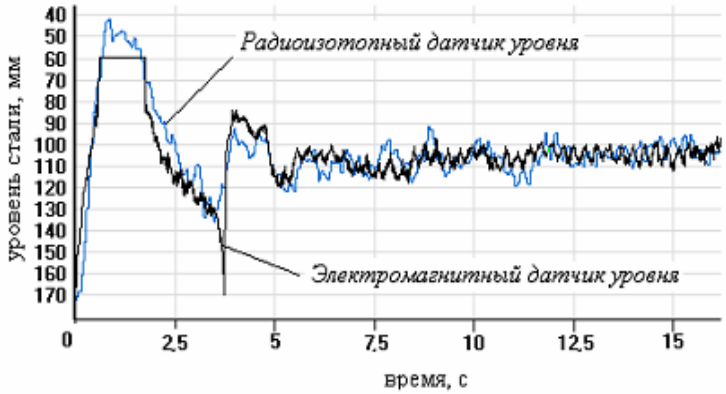


Рис. 3 - Временные диаграммы сигналов положения уровня стали, измеренного радиоизотопным и электромагнитным датчиками

При работающем электромагнитном перемешивателе неэкранированный вихретоковый преобразователь не работоспособен. Причина в том, что помеха электромагнитного перемешивателя содержит в своем спектре и частоту питания вихретокового преобразователя. Данные о помехе, полученные в ходе промышленных испытаний позволили создать математическую модель и лабораторный стенд для изучения влияния помехи перемешивателя. С их помощью удалось разработать экран для вихретокового преобразователя, который позволил снизить помеху в лабораторных условиях до необходимого уровня.

Планируются промышленные испытания экранированного вихретокового преобразователя. Если в ходе испытаний работоспособность вихретокового преобразователя в присутствии электромагнитного перемешивателя будет доказана, разрабатываемый вихретоковый преобразователь сможет в будущем заменить собой неэкологичные радиоизотопные датчики.

Сложности создания вихретокового преобразователя для использования в тонкослябовом кристаллизаторе средним сечением 250x1500 мм и блюмовом кристаллизаторе сечением 300x300 связаны с близостью противоположной к датчику медной стенки, близостью к преобразователю струи металла и наличием магнитопроводящего покрытия на медной гильзе. Из-за конструктивных особенностей кристаллизаторов, в которых планируется применение вихретокового преобразователя, располагать его возможно только на широкой стенке гильзы. Размещение преобразователя непосредственно над зеркалом поверхности металла не представляется возможным. Предложенный вихретоковый датчик располагается непосредственно над медной частью гильзы кристаллизатора, как показано на рис. 4.

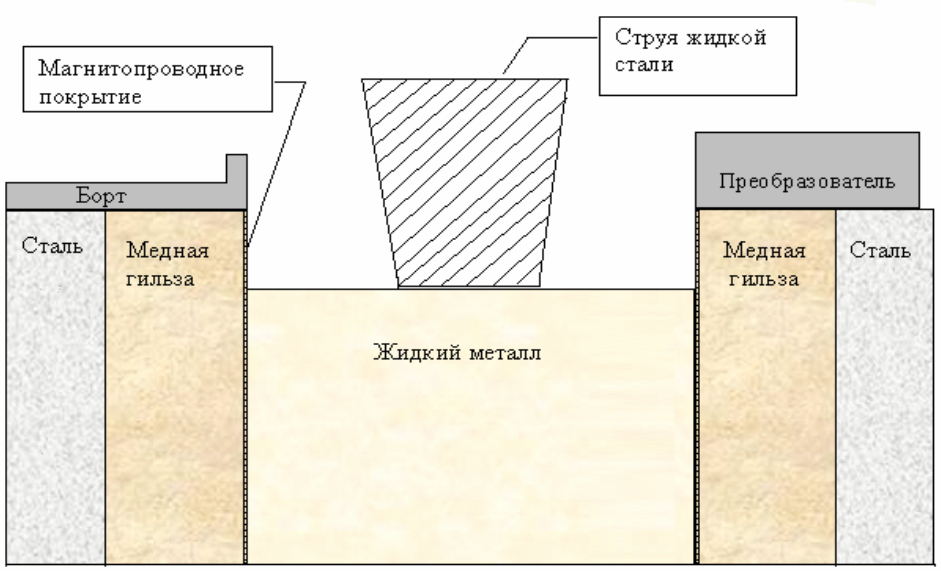


Рис. 4. Размещение преобразователя на кристаллизаторе

Максимальный сигнал от изменения уровня металла в кристаллизаторе получился для случая расположения ДУМ на узкой стенке кристаллизатора сечением 400х100, т.е. для случая максимального удаления противоположной стенки гильзы от датчика.

Рабочей частотой, на которой в меньшей степени будут сказываться температурные колебания в гильзе, а сигналы от жидкого металла будут значительными, была выбрана частота 2,5 кГц.

Для обеспечения работоспособности ВТП в условиях работающего ЭМП проанализировано применение нескольких приемов повышения отношения «сигнал-помеха»: экранирование измерительных катушек, дифференциальное включение измерительных обмоток по отношению к полю ЭМП, использование пассивного ФВЧ, подбор частоты возбуждения датчика уровня, установка экранирующих стальных пластин.

Вихретоковый прибор косвенного контроля был исследован в промышленных условиях. По результатам промышленных испытаний были определены погрешность вихретокового прибора контроля уровня в статическом режиме, определены его динамические свойства, а также была доказана нечувствительность вихретокового прибора к уровню шлакообразующей смеси. Временные диаграммы полученного после динамической коррекции сигнала вихретокового прибора на частоте 2,5 кГц и радиоизотопного показаны на рис. 5.

По результатам исследования характеристик вихретокового преобразователя предложены рекомендации по контролю уровня металла вихретоковыми преобразователями при непрерывной разливке стали, включающие:

– варианты конструкций преобразователей;

– диапазоны рабочих частот возбуждения преобразователей;

– соотношения между составляющими сигналов бортовых преобразователей;

– способы уменьшения уровней электромагнитных помех в тонкослябовом и блюмовом кристаллизаторах;

– алгоритм обработки сигналов бортовых преобразователей.

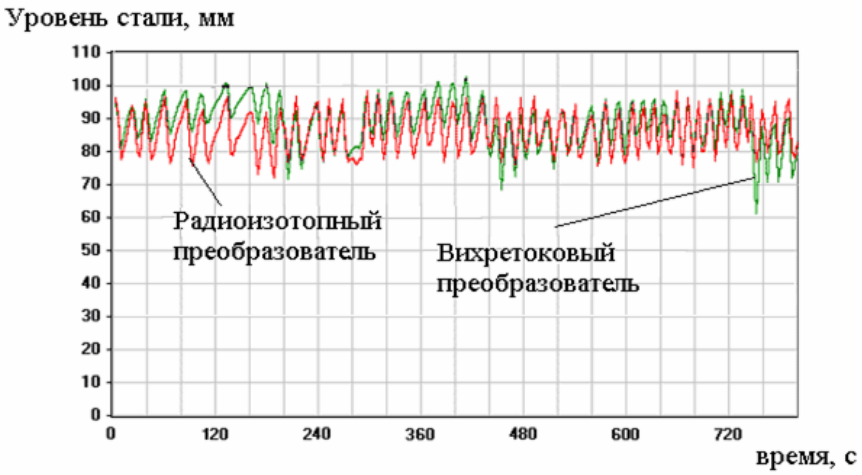


Рис. 5. Временная диаграмма сигналов уровня металла во время разливки

На основе сравнительных испытаний радиоизотопного преобразователя и вихретокового бортового преобразователя доказана применимость вихретокового метода для измерения уровня металла в гильзовом кристаллизаторе и подтверждена целесообразность замены радиоизотопного преобразователя на вихретоковый.

Мировая тенденция повышения доли тонкого проката на станах горячей прокатки и стремление получить высокоточную полосу на головном и хвостовом участках, прокатываемых в нестационарном режиме, поставили задачу измерять геометрию полосы не только на выходе стана, но и в межклетевых промежутках.

По существу, речь идет об обнаружении дефектов формы полосы (толщина, ширина, профиль и планшетность – изгиб полосы в направлении движения) не на уже готовой продукции на выходе из стана, а на стадии, когда дефекты только зарождаются. Поддержание эквивалентных профилей во всех клетях должно резко снизить появление непланшетности полосы. В настоящее время для целей технологического контроля геометрических параметров проката в основном применяются оптические и радиационные методы контроля.

Преимуществом вихретоковых средств контроля при работе с горячим прокатом является, прежде всего, то, что электромагнитное поле проникает через твердые непрозрачные стенки, а также через охлаждающую жидкость. Это допускает близкое расположение чувствительных элементов к полосе. Это делает более компактной и дешевой реализацию прибора для комплексных измерений. Измеритель может быть размещен полностью под полосой, что удобно с технологической точки зрения. Вихретоковые измерители толщины полосы не требуют применения большого количества эталонов.

Разработанный прибор для комплексных измерений представляет собой линейку вихретоковых измерительных преобразователей, которая располагается под полосой вблизи металла (рис. 6). Анализируя сигналы каждого датчика можно измерить толщину полосы над ним. Анализируя сигналы от всех датчиков можно проводить измерения ширины полосы. По измеренной толщине и ширине полосы в каждый момент времени можно судить о таких параметрах как планшетность и коробоватость, а также отслеживать поперечное смещение полосы от центра.

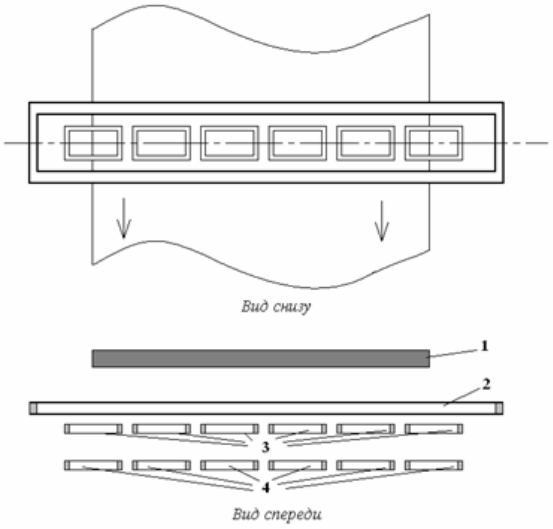


Рис.6 Расположение измерительной системы относительно полосы проката. Цифрами обозначены: 1 - полоса, 2 - обмотка возбуждения, 3 - измерительные катушки, 4 - компенсационные катушки.

Для измерения ширины с приемлемой точностью достаточно использовать одну частоту. Для измерения толщины используется две частоты, так как это дает дополнительные возможности отстройки от изменения проводимости. Анализ влияния различных факторов, таких как изменение зазора и проводимости позволил разработать эффективные алгоритмы выделения нужной информации и отстройки от мешающих факторов, а также позволил использовать измеренную ширину для коррекции измерений толщины в датчиках, расположенных над краем листа.