POLISH ACADEMY OF SCIENCE
COMMITTEE FOR ELECTRONICS AND TELECOMMUNICATIONS

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY

KWARTALNIK ELEKTRONIKI I TELEKOMUNIKACJI

VOLUME 54 – No 3

WARSAW 2008

Address of Editorial Office
00-665 Warszawa, ul. Nowowiejska 15/19 Politechnika, pok. 484
Instytut Telekomunikacji, Gmach im. prof. JANUSZA GROSZKOWSKIEGO

Editor-on-duty
Mondays and Wednesday
From 2pm to 4pm
Phone number: (022) 234 77 37

Telephone numbers
Editor-in-chief : (022) 825 15 80; (022) 234 73 30
Responsible Secretary : 0500044131
www.tele.pw.edu.pl/keit

Ark. Wyd. 19,00 Ark. druk. 15,25	Podpisano do druku w październiku 2008 r.
Papier offset, kl. III 80 g. B-1	Druk ukończono w październiku 2008 r.

Publishing
Warszawska Drukarnia Naukowa PAN
00-656 Warszawa, ul. Śniadeckich 8
Tel./fsx 628-87-77

IMPORTANT MESSAGE FOR THE AUTHORS

The Editorial Board during their meeting on the 18th of January 2006 authorized the Editorial Office to introduce the following changes:

1. PUBLISHING THE ARTICLES IN ENGLISH LANGUAGE ONLY

Starting from No 1'2007 of E&T Quarterly, all the articles will be published in English only.

Each article prepared in English must be supplemented with a thorough summary in Polish (e.g. 2 pages), including the essential formulas, tables, diagrams etc. The Polish summary must be written on a separate page. The articles will be reviewed and their English correctness will be verified.

2. COVERING THE PUBLISHING EXPENSES BY AUTHORS

Starting from No'2007 of E&T Quarterly, a principle of publishing articles against payment is introduced, assuming non-profit making editorial office. According to the principle the authors or institutions employing them, will have to cover the expenses in amount of 760 PLN for each publishing sheet. The above amount will be used to supplement the limited financial means received from PAS for publishing; particularly to increase the capacity of next E&T Quaterly volumes and verify the English correctness of articles. It is neccessary to increase the capacity of E&T Quarterly volumes due to growing number of received articles, which delays their publishing.

In case of authors written request to accelerate the publishing of an article, the fee will amount to 1500 PLN for each publishing sheet.

In justifiable cases presented in writing, the editorial staff may decide to relieve authors from basic payment, either partially or fully. The payment must be made by bank transfer into account of Warsaw Science Publishers The account number: Bank Zachodni WBK S.A. Warszawa Nr 94 1090 1883 0000 0001 0588 2816 with additional note: "For Electronics and Telecommunications Quarterly".

Editors

El
Elekt
T
mitte
of PA
origi
dely
optoc
T
as y

criti
bran
then
and

cial
The
Sci

cor
Mc

dis
the
mc

au
pu
in

ec

Dear Authors,

Electronics and Telecommunications Quarterly continues tradition of the "Rozprawy Elektrotechniczne" quarterly established 54 years ago.

The E&T Quarterly is a periodical of Electronics and Telecommunications Committee of Polish Academy of Science. It is published by Warsaw Science Publishers of PAS. The Quarterly is a scientific periodical where articles presenting the results of original, theoretical, experimental and reviewed works are published. They consider widely recognised aspects of modern electronics, telecommunications, microelectronics, optoelectronics, radioelectronics and medical electronics.

The authors are outstanding scientists, well-known experienced specialists as well as young researchers — mainly candidates for a doctor's degree.

The articles present original approaches to problems, interesting research results, critical estimation of theories and methods, discuss current state or progress in a given branch of technology and describe development prospects. The manner of writing mathematical parts of articles complies with IEC (International Electronics Commision) and ISO (International Organization of Standardization) standards.

All the articles published in E&T Quarterly are reviewed by known, domestic specialists which ensures that the publications are recognized as author's scientific output. The publishing of research work results completed within the framework *of Ministry of Science and Higher Education* GRANTs meets one of the requirements for those works.

The periodical is distributed among all those who deal with electronics and telecommunications in national scientific centres, as well as in numeral foreign institutions. Moreover it is subscribed by many specialists and libraries.

Each author is entitled to free of charge 20 copies of article, which allows for easier distribution to persons and institutions domestic and abroad, individually chosen by the author. The fact that the articles are published in English makes the quarterly even more accessible.

The articles received are published within half a year if the cooperation between author and the editorial staff is efficient. Instructions for authors concerning the form of publications are included in every volume of the quarterly; they may also be obtained in editorial office.

The articles may be submitted to the editorial office personally or by post; the editorial office address is shown on editorial page in each volume.

Editors

R. Ro
R. S.
EuCA
M. Kv
tal im
P. Site
1.5cc
plexin
D. Po
manag
Ł. Sa
the bi
R. Ro
J. Og
K. Op
K. Op
J. For
lving
P. My
W. W
W. W
ment
A. M
of m
D. K
D. K
Infor

CONTENTS

scien
near
of dy
of bi
quen
resig
as fr
of th
the
publ

meth
[http
rese
[pho

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 233–238

PHOTONICS LETTERS.
A new, peer reviewed Internet publication

RYSZARD ROMANIUK

Warsaw university of Technology
e-mail: rrom@ise.pw.edu.pl

Received 2008.07.09
Authorized 2008.08.28

The global scientific publications market, especially in technical sciences, is subject, during the last years, to an essential reconfiguration. There are observed the biggest changes ever since several decades. These changes embrace: a considerable increase in the number of journals of the global impact, emerging of new research branches and accompanying, relevant journals (for example: biophotonics, THz technology, metamaterials, optics of negative refraction, etc.), a fast, evolutionary 'internetization' of scientific journals, with all consequences of this process (leading very probably in the near future to the confinement of the role of pure text and classical figures on behalf of dynamic multimedia presentations of the research results), and a complete change of bibliometric tools - measuring the objective value of the publications. The consequences of these changes are irreversible. The market of strictly scientific publications resigns irretrievably from the static evaluation of the value of published work, as well as from the paper versions of the journal issues. These processes go in dependently of the strength of the sentiments of classical, analog publishers and authors. To stop the remaining doubts, one can say that the costs will ultimately decide. The digital publications are around one order of magnitude cheaper than the analog ones.

Trying to keep the pace with these revolutionary changes of the publications methods and tools, ruling the new market, the Photonics Society of Poland – PSP, [http://photonics.pl] has launched, since the second half of 2008, a new, peer reviewed, research and engineering, Internet based, fast publishing journal "Photonics Letters" [photonics.pl/PL].

The journal Photonics Letters is a publication issued only in the Internet, without any paper copy. According to the PSP estimate, only such publication initiative for a journal, issued digitally, very versatile, and published extremely fast, has a real chance for success on a very dense, difficult, and competitive global publishing market. During the initial period, the Photonics Letters journal will be issued as quarterly, next bimonthly and monthly. Biweekly issue is also not excluded, as the ultimate aim. The publication time for an Author will be, however, much shorter. Two first issues of Photonics Letters will be published in 2008 embracing the third and the fourth quarters of the year. The journal aspires to the high class of peer reviewed, fast, short, letter type publications, functioning since not very long in the Internet, under a general name - rapid Internet publications. The pattern publications are issued in electronics and photonics by professional associations of global reach like SPIE, IEEE, OSA, for example Electronics Letters, Photonics Technology Letters, Optics Letters, Photonics Journal etc. The journal accepts for publication short work, not extending over four pages of text, prepared according to a two-column format analogous to the one published by IEEE Transactions. The submitted work should have the structure of a full journal paper.

The journal covers the topical areas of optics, optoelectronics and photonics in the following aspects: fundamental and applied research, physics and technical, materials, components and devices, circuits and systems, technological and design, construction and manufacturing, and metrological. In particular these are the following areas of the photonics: theory and experiment, quantum, photon physics, ultra-low power, atto-watt, single photon, ultrafast, femto- and atto-second, high power, high intensity, tera-watt and peta-watt, nonlinear, integrated and planar as well as volume, information, image, lighting, photovoltaic, semiconductor, laser, terahertz -THz, IR, UV, synchrotron, roentgen, free electron lasers, optical fiber, communications, sensors and instrumentation, biomedical, chemical and environmental, industrial and engineering, synthesis and processing of materials, and educational. The journal also embraces all kinds of experiments of interaction of a photon beam with matter, as well as the usage of photonics in high energy physics experiments and astro-particle physics and astronomy. A very current subject, embraced by the journal, is synthesis of metamaterials for photonics, and in particular research on meta-glasses and their applications. This research leads to a negative refraction optics.

The journal also covers neighboring regions to the above mentioned ones. Photonics Letters accepts all work concerning construction of research and application systems: metrological, automation and robotics, astronomical, chemical, biomedical, for food and agriculture industry, environmental monitoring, mechatronic, electronic, nano- and microsystems, and everywhere the photonics is a functional part of a bigger unity. The functions played by photonics may embrace sensors, metrology, data transmission channels, EM wave energy delivery, and other types of functional components and subassemblies, etc. One can mention the following exemplary systems: application specific 3D imaging devices, integrated processors of the type - lab on chip, lab on fiber (capillary), MOEMS, MEMS, THz imaging, transmission with quantum coding, optical gauges of time, frequency and distance, devices made of metamaterials and photonic crystals, etc.

The journal embraces also the integration problems of photonic components, devices and circuits with the analogous, accompanying, functional devices of electronics and mechatronics. There is a considerable interest in simultaneous coexistence of photonic, electrical and mechanical signals in a single integrated micro-system. It also concerns other kinds of signals like chemical, biological, including biophotonics, as well as signal conversion and processing, functional integration, packaging of components and devices, construction, small scale manufacturing, etc. The journal covers all aspects of the software accompanying the theoretical and design work as well as hardware solutions. There are two main aspects of software: calculation software for modeling of photonic effects and phenomena in materials and components as well as functional software for servicing the photonic devices and systems as well as mixed integrated systems consisting of photonics, electronics and mechatronics.

The journal accepts for publications only original, fast, i.e. very current, research and technical work reports, prepared in a form of a short, fully structured paper. There are accepted the following forms of work and topics: laboratory reports, engineering solutions, calculation results, modeling, laboratory demonstration and technical work descriptions, application examples, new components and devices, processing of optical signals, signal conversion, new and modified sensor solutions, technical descriptions, methods of demonstration of photonic effects, dynamic documentation of laboratory work – specific for new multimedia Internet tools, material description for photonic applications, photonic measurements, microscopy, digital holography, optical memories, photonic processors, maintenance and exploitation of photonic apparatus, reliability and ruggedness of apparatus, apparatus survivability in benign and adverse work environment, specificity of industrial apparatus, design notes, application notes for photonic components, research remarks and debates, authors responses, system integration between photonics – electronics and mechatronics, hardware – software integration, practical advances in the construction of micro-systems including moems, improrement of imaging techniques in different spectrum regions, software for photonic and mixed integrated devices and systems. In the above sense, the main area of the journal activity is applied science and technology, as well as developments of practical engineering.

Despite the initial organization of the journal as quarterly, the basic intention of the PSP as the journal publisher, and the editorial board of Photonics Letters, is to provide the Authors with a very fast path to the peer reviewed, high quality, Internet publication. The aim is to reach two weeks from the paper submission to the publication on-line. This would be possible only via a complete automation of the submission, reviewing and paper verification processes. After the acceptance, the work would be immediately, with no delay, published on-line on the webpage of the journal. Each paper would be checked obligatory by two reviewers. Apart from classical text with figures, the journal intends to publish accompanying auxiliary multimedia files in a form of additional still and animated graphics, audio comments, and video recordings of the experiment. Possibly, in the future, the text form would be only a formal addition to the full multimedia record of the work for the publication.

Launching a completely new publication, in a very attractive field, is an extremely challenging task now, and usually a result of long series of negotiations on the local, national level as well as international and global levels. Locally in Poland, the agreements concerned social, professional organizations, research institutions, universities, industry, government and business, as well as key persons, representing this very dynamically developing branch of science and technology, which is now photonics. The consulted national organizations, relevant to the publication of Photonics Letters, embraced among others: Polish Academy of Sciences (PAN), Division IV of Technical Sciences, Committee of Electronics and Telecommunications, Section of Optoelectronics; Association of Polish Electrical Engineers (SEP), Polish Committee of Optoelectronics; Polish Physical Society, Section of Optics; and a number of other professional organizations of topical relevance. The consultations embraced also the following schools and industrial and business organizations: governmental laboratories, faculties of physics, electronics, IT, electrical engineering, mechatronics of universities and some private businesses. The consultations are continued with other Committees of the Polish Academy of Sciences including: Astronomy, Physics, Metrology and Research Apparatus, Electrical Engineering, Bio-cybernetics and Biomedical Engineering, and others. Photonics Society of Poland counts seriously on close interaction in publishing domain with other professional research and technical societies, and in particular with the Association of Polish Electrical Engineers.

On the international level, the problem of consultation of such a big initiative, like launching a new journal of a global extent, is much more complex than locally. There are carried out, by the PSP, intense and advanced consultations with SPIE – The International Society for Optical Engineering [spie.org]. The reason of proposed cooperation concerning the journal are close, friendly, and profitable, and effective relations between SPIE Poland Chapter [spie.pl] and SPIE. The PSP is a direct continuator of SPIE Poland Chapter. A participation of SPIE in launching the journal gives to it immediately a global character.

The guarantee of success for such complex initiative in cooperation between PSP and SPIE, local and global photonics society, is a great product of this cooperation spanning over a quarter of the century. The results of this cooperation are: publishing of over than 200 volumes of Proc. SPIE from the international technical conferences on photonics organized in Poland, cooperation on publishing of next 100 volumes of Proc. SPIE from the international conferences organized elsewhere, common organization, with SPIE, in 2005, in Warsaw, of an International Congress on Optics and Optoelectronics, which gathered around 1000 professional participants, etc. Apart from close consultations with SPIE, the PSP which has quite a recognition of the European photonics market, consults launching of the new Internet publication with photonics associations in the neighboring countries.

The decision of launching of Photonics Letters is supported, on the international level, by a number of developmental processes concerning the photonics in the global scale. The photonics is subject to a further dynamic development, much faster than other related branches of the industry. This industrial and related research development trend is supported by a few powerful application pillars. These pillars are: optical communications, reborn with only marginal losses after a serious crisis a few years ago, precision metrology offered by lasers and interferometry, continuously extending applications of lasers, imaging techniques in medicine, sensors and integrated instrumental systems of moems type. A nother observed process is a maturing reconstruction and integration of the research and technical professional community of photonics in the global scale. A particular layer of this integration is a different view now on the cooperation, and its strongly synergetic role, between professional societies acting in local and global scales. A perfect example of this is profitable, long-lasting cooperation between PSP and SPIE.

The local aims of PSP concerning launching of Photonics Letters are: building of a strong, international societal journal; making more active the local engineering community of photonics; strengthening of the ties of this community with analogous local communities in Europe; globalization of the local activities; increasing the cooperation with industry and national economy; tracing and participation in the development of the current research direction of photonics; promotion of these directions locally; stimulation of applied research in photonics for the local economy.

One of the main motivations to launch own journal was the recent transformation from SPIE Poland Chapter to the Photonics Society of Poland. SPIE Poland Chapter acted for the last twenty years as a legally registered technical professional society. Transformation of the name was approved by the administrative court on 2007/08. The society is highly evaluated by the local professional community due to its intense activities on behalf of its members. Locally, the major aim of Photonics Letters journal is to open a possibility of fast, peer reviewed publications with a wide international indexing. A number of authors of valuable work tend to publish the results locally, assuming proper level of the journal and global indexing. Despite the global access to the journal it is treated as a local one. One of the ideas of globalization of the new journal uses this authors' psychology. The idea is to build a global journal, in the future, by launching a network of combined local editions.

PSP intends to introduce new journal, in a close cooperation with SPIE and national research and technical community of photonics, to be indexed by the ISI-Thomson and other key indexes and scientific publication databases like: Scopus, Scholar, Scirus, Scitopia, etc. The intentions of PSP as a main publisher of Photonics Letters is a close cooperation in Poland with the publishers of the other two important, strictly photonics journals, which are indexed by the ISI-Thomson. These are: Optica Applicata – published by the Faculty of Physics, Wrocław University of Technology and

Opto-Electronics Review published by the Military University of Technology, SEP and Elsevier. Cooperation and usage of synergy in the publishing activities of all of these journals may base on utilizing of the new specificity of the fast Internet publication to strengthen the archival journals publishing the work results in extenso. These journals are also subject to the evolution towards Internet based publications. They publish much longer papers, usually being a full documentation of concrete experiments, construction of new apparatus, thus, this kind of publication will never be so fast as a publication of 'letter' type on-line.

The aims of PSP concerning the European and global photonics communities, related to launching the Photonics Letters are: strengthening and extending the cooperation with SPIE, and considerable internationalization of this cooperation, building a strong international photonic journal of clearly European character and simultaneously very global impact; extension of European cooperation between local associations; trial to create around the journal a wider scientific and technical platform for international cooperation and debate; built in and active participation of PSP in the global photonics publications market, etc.

The invited editorial board of Photonics Letters has members from Poland and other countries. The Board consists of the leaders of large research and technical laboratories in photonics of a considerable scientific impact in the global scale. The individual persons declared active participation in launching and building of the new journal. Each person is responsible for a particular area of photonics, usually associated with the specialty of the represented laboratory.

Photonics Society of Poland and Editorial Board of Photonics Letters has recently made a number of initial steps towards practical launching of the journal. The initial make-up of the Editorial Board was decided. The first web page of the journal was launched containing the basic technical information. This page will be transformed in direction of a fully functional submission and on-line data base machine in the analogous style as MySPIE, eXplore, OpticsInfoBase or Versita. A call was announced for the first two issues of the journal. There are continued further consultations, domestic and international concerning the launching and maintenance of the journal, and building its unique identity.

Photonics Society of Poland encourages the prospective Authors to publish the results of their research and technical work in the new, peer reviewed, rapid Internet journal 'Photonics Letters'.

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 239–251

Development of accelerator technology in Poland Impact of European CARE and EuCARD programs

RYSZARD S. ROMANIUK

Warsaw University of Technology,
rrom@ise.pw.edu.pl

Received 2008.07.03
Authorized 2008.08.29

The development of accelerator technology in Poland is strictly combined with the cooperation with specialist accelerator centers of global character, where the relevant knowledge is generated, allowing to build big and modern machines. These are relatively costly undertakings of interdisciplinary character. Most of them are financed from the local resources. Only the biggest machines are financed commonly by many nations like: LHC in CERN, ILC in Fermi Lab, E-XFEL in DESY. A similar financing solution has to be implemented in Poland, where a scientific and political campaign is underway on behalf of building two big machines, a Polish Synchrotron in Kraków and a Polish FEL in Świerk. Around these two projects, there are realized a dozen or so smaller ones.

Keywords: accelerator technology, synchrotron, linac, free electron laser, FEL, radio frequency superconducting circuits, financing of research

EuCARD and CARE are examples of large European framework research and development programs to build common, enormous research and development infrastructure in the area of accelerator technology. CARE – Coordination of Accelerator Research and Development in Europe, was realized during 2004-2008, and EuCARD – European Coordination of Accelerator Research and Development, is its close continuation, and is planned for realization during 2009-2012. These programs consist of research networks, trans-national access to large research infrastructure and quite a number of particular joint research activities. The latter concern apparatus design methods, technology development, material research, components, sub-assemblies, and considerable modernization parts of existing large infrastructure. Polish research and development community utilize the participation in these programs very efficiently for: exchange and training of researchers, increased participation in large experiments attracting young researchers to exciting and demanding topics, building of auxiliary infrastructure in Poland, access to new technologies and knowledge having industrial applications. Participation in large, global and European research programs is a strong stimulation for undertaking in Poland, our own brave programs of building accelerator infrastructure. Towards the end of 2007, the National Council of Nuclear Research, cooperating with the National Atomic Energy Agency recommended for realization (apart from the basic program of building of nuclear energy infrastructure) three big accelerator projects: Center of Hadron Therapy, National Center of Synchrotron Radiation, Polish Free Electron Laser – Polfel.

1. ACCELERATOR TECHNOLOGY

Accelerator technology is a branch of research and technology which subject is acceleration of charged particles, research of the beams, interactions between beams and technical applications [1-19]. It embraces theoretical aspects and technical associated with the beams, design and optimization of accelerators, building, exploitation, applications for research and technology; properties of accelerating particles individually and in beams; particle detectors; points of particle interactions; application of particle beams; electron and positron beams, muon (heavy electron), hadron (three gasho) like protons and meson (two quarks) beams, and ion beams. The particles were accelerated to relativistic velocities, some time ago electrostatically (Van de Graaff), and now in EM RF accelerators. The accelerators can be warm, working in room temperatures, or cold, the latter working in super fluid helium, usually below 2K. The electrical field intensity in a resonant cavity made of copper and tungsten of a warm accelerator may reach 200MV/m (typically below 100MV/m). The electrical field intensity in a resonant cavity made of single crystal niobium of a cold accelerator may reach above 45MV/m (typically around 30MV/m).

The ultimate confinement of the field value in warm accelerators are electrical breakdown and pulling out atoms from the metal surface in the resonant cavity. The surface of warm copper cavity can be hardened locally, in the most critical places, by covering it with tungsten. The field in cold cavities may be increased by application of very pure, single crystal niobium Nb instead of polycrystalline (as usually metals are). Other direction of development of cold cavities is looking for materials which are warm superconductors. A promising material is Nb_3Sn. The confinement for the largest field value is multipacting, or avalanche multiplication and acceleration of electron cloud pulled out form the metal surface. These electrons build up a parasitic dark current of the accelerator. Other confinement is existence of residual higher order modes (apart from the fundamental mode - which is used for acceleration) in the resonant structure of the accelerator. At the above values of the electrical field intensity, the deposit of EM power, in a form of standing wave is very large and may reach a few hundred kW per a single cavity. The aim of efficient acceleration is transferring this large power, in the largest possible degree, directly to the electron or positron beam. The most popularly used cavities in the superconducting RF accelerators are of the TESLA type. They are made of pure niobium, work at the frequency of around 1,3 GHz, and in temperature below 2K. A single cavity consists of nine cells. Each cell is half-wavelength long, thus the whole structure is around 1m long.

Apart from application of large accelerating machines, there are developed alternative laser-plasma accelerating schemes for particles. An accelerating tunnel is created by a high power laser in dense plasma. The EM field intensity in plasma tunnel may be three orders of magnitude higher than in cold RF accelerators, and may reach tens of GV/m. This allows for generating of very energetic particle beams at the cost of considerable shortening of the accelerating structures and decreasing the overall needed energy, thus, increasing the machine efficiency. The problems to solve are: increasing the stability and length of the dynamic accelerating, plasma tunnel. Still another way to obtain ultra-high-energy relativistic particles is cosmic radiation. However, these particles are very rare events and are not controllable. The effects of their interactions may be havewer, observed.

The accelerators are also used for indirect generation of neutral particle beams, like neutrons and neutrinos, via the spallation of atom nuclei. The accelerator technology, via high energy of the particle beams, is closely related to nuclear technologies. It is also used, which is most interesting, for generation of photon beams of very wide spectrum of energies now from far IR, via visible to X rays and in the future to gamma rays. The machines to generate the most powerful photon beams are free electron lasers, using longer and longer linear accelerators, undulators and SASE mechanism. In this way, the accelerator technology is closely combined with photonics, optoelectronics, optics, electronics and mechatronics. Availability of more and more powerful user machines extracts a relatively new branch of research and technology called photon physics. These activities are one of the paths leading to the harnessing of a photon, in a similar way we have harnessed the electron. A photon is, however, much more difficult to be harnessed, because it continuously keeps escaping. There are a number of difficult problems, along this photonic path, to be solved, and among them: a considerable slowing down of the photon, if not stopping him completely for a while, gathering photons, controlled interaction of photons with the same and different energies, guiding of the photon along a complex optical path, optics of refraction equal to one, less than one, close to zero and negative. Especially the optics of negative refraction seems now very challenging in the visible spectrum.

An inverse free efectron laser, or I-FEL, is a machine where a laser beam accelerates and structures into short coherent bunches a co-propagating electron beam. Research on I-FEL may lead to laser – based, table-top, partide accelerators.

2. INSTITUTIONS OF ACCELERATOR TECHNOLOGY

The development of accelerator technology, due to the character of experimental work demanding very large machines, is concentrated in a few research centers around the globe. The biggest accelerator complex in the world is located in CERN [20]. Big accelerator center of national character are: Fermi Lab [21], SLAC [22], ORNL-SNS [23], Jefferson Lab-CEBAF [24], DESY [25], INFN [26], IN2P3 [27], KEK [28], etc. Research teams from Poland cooperate closely with these laboratories in Switzerland, Italy, France, England, Germany, USA and Japan. A considerable number of young scientists stays there permanently on shorter or longer research fellowships. A number of them keeps tight relations with the country by carrying common research projects. The accumulated experience, knowledge, cooperation between large teams, carried technical debates, participation in common programs, gives with time the birth to brave ideas to build large accelerator infrastructure in Poland. The ideas, worth hundred million zł, of this kind, are: Polish Synchrotron [29] in Kraków, and Polish Free Electron Laser – POLFEL in Świerk [30]. Realization of such big projects in this country is important be cause of many reasons – active participation in the process of generation of knowledge, breeding of new generation of scientists, generation of technical applications, etc.

European coordinator of accelerator research, on the administrative level, is ESGARD - European Steering Group for Accelerator Research and Development [31]. The Group consists of directors of international and governmental institutes of nuclear and accelerator technologies in the UE countries which have their own big accelerator infrastructure. Poland has not a representative in ESGARD. The natural candidates are directors of the IPJ in Świerk [32] or IFJ in Kraków [33]. Other organizations may also be taken into account, as governmental institutions like IEA or PAA – National Atomic Energy Agency. The condition, to have a representative in ESGARD is to possess in particular country active accelerator infrastructure. Poland has not such an infrastructure now of considerable extent. A cyclotron is in possession of Heavy Ion Laboratory by Warsaw University [34]. Some time ago there were active small electron linacs. This infrastructure will not be modernized, possibly only as part of bigger undertakings. Such plans are related to the mentioned national synchrotron and FEL laboratories. Participation in ESGARD activities is very important for the national research and technical communities. This influential organization takes actively part in crucial decisions, preparations, and submission of research programs in this area, covering the whole Europe.

European coordinator of accelerator research, on the scientific level, is ECFA – European Committee on Future Accelerators [35]. The members of ECFA are professors, representing research institutions from the EU member states, which carry out research in the accelerator technology. Poland is represented by a representative of IPJ. Poland, as a member of CERN, is also represented in ICFA – the International Committee for Future Accelerators [36]. A part of the coordination activities is done by the International Union of Pure and Applied Physics IUPAP [37]. Poland is represented in IUPAP by a representative of IF PAN. The ECFA cooperates with analogous regional organizations like Asian ACFA [38] under the umbrella of ICFA which has more global extent.

The most important developmental factor for accelerator technology (and for other branches of research and technology) is existence in Europe of a relevant, lobbying, strongly organized, research and technological community, representing the activity area on the governmental, organization, financial, business, scientific, industrial and social levels. Participation of Poland in such big, strictly organized and coordinated in the scale of whole Europe, research programs requires the existence of analogous communities – actively interacting with the European ones.

The listed factors of existence of centers of influence are fulfilled in relevance to the European community of accelerator technology. In Europe, and in most of the countries interested in the participation in the accelerator technology programs, there exist relevant institutions, organizations and governmental bodies, as well as professional and community groups. A similar situation is in Poland. The governmental bodies are PAA -National Atomic Energy Agency [39] and interested governmental ministries. The professional bodies are topical consortia, research networks, centers of excellence, industrial clusters, technological platforms – which are members of big experiments. Other relevant institutions are units of PAN [40] and professional associations, like: Committee of Physics, Department III of PAN, Department IV of PAN – Technical Sciences, Polish Physical Society [41], Polish Society of Synchrotron Radiation [42], Polish Nucleonic Society [43], etc. The professional work force originates from several PAN and JBR institutes like: IFPiLM, ICHiTJ, IPJ, IFJ, IEA, ITME, ITE, IF-PAN, MITR-PŁ, etc. Among the national universities, AGH, UW and PW feature the widest participation in the international accelerator experiments.

The usual place of meetings of the international accelerator communities are conferences: PAC, EPAC [44], LINAC, FEL, ICALEPCS [45], etc. Proceedings of these conferences are published on a common, global Internet web portal JACoW – Joint Accelerator Conference Web site [46] or CERN portal INDICO – International Digital Conference [47].

3. FINANCING OF ACCELERATOR TECHNOLOGY

Big accelerator structure is financed generally on the level of relevant countries. Similar situation is in the case of different infrastructural investments like: space research, nuclear or large power utilities. Despite the national character, in most cases, the big infrastructure is open and made widely available for international proposals from research teams, nearly without any confinements. A separate issue is financing of these proposed experiments by directed grants. Only the biggest world experiments are financed for infrastructure fully by international research consortia. The kinds of costs have to be differentiated: direct infrastructure investment, basic maintenance, experiment maintenance, excess costs of particular experiment, etc. Similar solutions have to be assumed in Poland. The two mentioned projects have to be realized mostly from own resources. Simultaneously the managers of the projects have to look for a reliable and the widest possible international partnership. The partners are absolutely necessary, because at such complex and technologically advanced undertakings, not all necessary knowledge is available in a single country. New national research teams are trained extensively around such projects. Despite the creation of a new infrastructure, the international cooperation assures effective usage of the available exploitation time of the machine. The large and expensive infrastructure has to be of a double usage: research and user oriented. Connecting these two functions is not an easy logistic task which has to be mastered by the machine managerial team. Financing of such big and unique infrastructure now in Poland uses an outdated method of SPUB type of grants. The method is not optimal for building large machines in wide international cooperation, where the endeavor extends the frames of a single institution, even the frames of several joined specialist national institutions. The infrastructure of this size has to be inbuilt into an international network of similar machines.

Since the access of Poland to the EU, a possible source of co-financing for large research infrastructure are the European structural and framework programs. It means a co-financing and not full financing, because a very good program definition and a declaration of own financial input are necessary. Structural programs of EU (POKL, POIG, POIS,...) and the successive FP7 are a natural source of co-financing for these two programs of the Synchrotron and the FEL. FP7 was planned for 2007-2013 as a continuation of the previous FPs. Technical information is available via Cordis [48]. Polish laboratories participated in FP5 during the pre-accession period and in FP6 (2003-2008) after the accession. The budget of FP7 is around 50Mld∈ which is around 5% of the accumulated research budget of EU and separate budgets of the member countries. At this budget, influence of the program is planned mainly in the regions of focusing and amplification of the European integration in the research area, building of the ERA, investment in the human resources, amplification of mobility schemes, searching for directions to increase Europe's competiveness in innovative industries – strongly based on research, evolutionary building of knowledge based information society. FP7 is divided to large sub-programs: People – embracing investment in human resources including young researchers; Ideas – investment in innovative projects by European teams and building of European Council for Research; Cooperation – (the largest part of FP7) building of modern methods of trans national cooperation combined with coordination of national efforts and building of European Technological Platforms; Capacities – development of research infrastructure to rise the European research and innovation potential. These programs are amplified by building a Joint Research Centre [49] to improve the common research policy; and creation of Euratom - European Community for Atomic Energy [50].

Capacities – FP7 part concerning the infrastructure, has a budget of over 4Mld∈, (600M∈ per annum) which is relatively small in the European scale. Its impact may be visible only when the financing is strictly directed to narrow and the most demanding aims. Maximum co-financing of a single grant is 10M∈. Topical definition of the program is strictly determined, on the political level, by the ESFRI – European Strategic Forum of Research Infrastructure [51]. Poland participates in the activities of ESFRI by the representatives of the NCBR – National Center for Research and Development [52] and the Ministry of Research and Higher Education. Capacities is divided to a few priorities, from which two the biggest are: research infrastructure (1,7Mld∈) and SME (1,3Mld∈). EU defines the research infrastructure as: devices, resources and services to carry valuable basic research, and to disseminate, exchange and store the generated knowledge as results of these experiments. The research infrastructure embraces: research machines, archives of information, telemetric networks (GRID, GEANT), unique research structures. Research infrastructure is a tool in building e-science. The main components of e-science are: a researcher, an experiment and storage, analysis and processing of information. The research infrastructure fills the area of a triangle of science, with the edges made of research, education and innovation. The general non-topical priorities (topical priorities are defined by ESFRI) of EU concerning the research infrastructure are: usage optimization, development of the best existing infrastructure, help in building the new needed infrastructure of pan-European extent, amplification of trans-national processes in creation and usage of infrastructure.

The scale of realization of an individual research program of global extent inside the FP7 program is not directly expressed by the accumulated financing. The components are: participation declaration from the industry, additionally generated industrial contracts, rejuvenated trans-national activities, synergetic effects, number and quality of institutions affiliated with the program, interest of young researchers, predicted results of the program realization, interest in these results by research, governmental, business and industrial communities, and many more. In certain cases, the best ones, with many industrial contracts associated with the program, the financial multiplication of the program investments, of big predicted impact for industry, may be estimated for 2 -3. It means that in the case of EuCARD program the accumulated, virtual investment is of the order of 100Mln∈. The chances to approve a program of this size are associated with several key factors. One of these factors is a declared level of own financial participation. Theoretically the obligatory level is 25%. The declaration from EuCARD was around 70%.

4. CONSORTIA, NETWORKS AND TECHNOLOGICAL PLATFORMS OF ACCELERATOR TECHNOLOGY

Examples of Networking Activities (NA) and consortia in accelerator technology are (FP6): ELAN – European Linear Accelerator Network [53], BENE (MEGLIO) – European Beams for Neutrino Experiments [54], HHH- high energy, high intensity hadron beams [55], EuDET – Detector for ILC [56], FITAL – Consortium of Physics and Technology for high energy Linear Accelerators [57], Femtophysics, CEAT – Center of Nuclear Technologies. Such consortia and networks in Europe and in Poland are opened for particular purposes. Some of them like EDET or FITAL, and Femtophysics are closely associated with global experiments like ILC – International Linear Collider [58]. Research teams from Poland participate in most of these consortia and networks.

A general assumption for NA is building of cooperation culture between participants and communities using the common infrastructure. It is realized by activities on behalf of: users (training), communities (conferences, working groups), good research practices (exchange of personnel, work standards), trans-national access (TA), intensifying of common research activities, building of virtual infrastructure (data and software). The aim of TA is to provide the possibility or a real and virtual training of personnel, and optimal usage of the infrastructure. The aim of joint research activities JRA is the development of prototypes, methods, protocols, standards, software, providing of samples for research.

TESLA Technology Collaboration (TTC) [59] is a continuation of TESLA Collaboration. The aim of the latter was building of a planned linac and electron – positron collider – Tera-ElectronVolt Superconducting Linear Accelerator, located in DESY. Now the project was shifted to other location and is continued by the Fermi Lab. The TTC was evolutionary transformed into an important consultative body of global community character. Participation of a numerous research teams from Poland in the research programs of CARE and EuCARD resulted in acceptance of Warsaw University of Technology and Łódź University of Technology into this influential body of global character. The members in TTC are also Polish atomic institutes.

Potential participation of Poland in the ILC – International Linear Collider project is quite wide, and may embrace even a few hundred of the personnel. Numerable research teams from Poland have recently declared their interest in the work on design, construction and usage of this enormous linac, detectors and accompanying infrastructure. Some of the teams proposed a number of innovative research experiments using this machine. The coordination of activities on ILC in the wide scale is done by the GDE - Global Design Effort. To facilitate the cooperation with the GDE, the ILC communities in Poland have recently organized themselves into several topical consortia like: EuDET, FITAL, Femtophysics and other like Warsaw ILC Group [60], Warsaw E-XFEL Group, Warsaw ELHEP Group - Electronics and Photonics for High Energy Physics Experiments [61]. The project of ILC will belong to the largest research undertakings besides ITER [62] and LHC [63], with the total budget over 10Mld€.

The biggest in the world, working now, free electron laser FLASH is located in DESY/Hamburg. The laser uses a two hundred meter electron linac based on superconducting TESLA technology [64]. FLASH laser is generating a beam of VUV photons, the most intense ever. It is mainly a user facility. In parallel, it serves as a machine study and development laboratory for the planned European Roentgen E-XFEL Laser [65]. The linac for E-XFEL will be around 2 km long, i.e. ten times bigger than in FLASH. A large number of research teams from Poland participate in the development of both lasers, and in the programs of beam usage.

A consortium XFEL-Poland was inaugurated in 2007 [66] to facilitate and coordinate the cooperation with E-XFEL. Construction of this large machine will be done on commercial basis by the E-XFEL-G.mB.H. The partners of this business will be nondependent daughter firms E-XFEL-country. Here in this country, the partner will be E-XFEL-Poland Sp.z.o.o. A partner of this firm will be NCBR. A large number of European countries have declared participation in E-XFEL. The share is declared in a percentage of the total costs, which are now evaluated for around 1,5Mld€. Poland has initially declared, on the governmental level, a share of 30Mln€. This amount defines also a percentage of the future machine ownership. The share may be declared in cash as well as in-kind. The In-kind contribution embraces personnel and materials.

In parallel to the activities of the E-XFEL-Poland, during the same time and basing on similar rules, there was established a consortium on Femtophysics [67]. The aim is to prepare in Poland a relevant research and experimental structures for efficient cooperation with the international accelerator center for anti-proton and ion research FAIR GSI in Darmstadt [68]. This center possesses one of the biggest accelerator infrastructures in Europe. It does research on the structure of matter and evolution of the universe. The participation of research teams from Poland is quite numerous. The financing is provided on the governmental level.

POLFEL project [30] is closely combined with FLASH and E-XFEL, research and user projects realized in DESY. Polfel laser is designed as a part of a big European network of FEL lasers, cooperating with the biggest E-XFEL machine. It will be realized in IPJ Świerk. The total cost is estimated now for around 100Mln zł. The realization time is estimated for a decade. Polfel will use superconducting cavities organized in cryo-units of TESLA type. Each unit has eight cavities. The source of electrons will not be pulsed (klystron) but tunable CW (klystrode) [69]. The construction and exploitation of the machine will engage a large number of research and technical teams from the whole country.

Polfel is a part of a planned Research-Technology Park by the National Center of Atomic Energy CEAT. This large project engages the biggest institutes of atomic technologies, local government, universities, PGE - Polish Electrical Power Group, public and private businesses. The project consists of the activities of Polish Technological Platform for Nuclear Technologies [71]. The CEAT was established as a research and development consortium and one of its tasks is training of personnel for atomic industries [72] embracing also accelerator technology. Further plans combined with the development of this large research and industrial center is to establish a new university – MAT – Mazovian Academy of Technology, teaching engineers in modern branches like accelerator technology, nuclear energy industry, atomic technologies for industry, medicine, environmental protection and safety.

Polish Synchrotron [29] is a very important project to built in Poland a strong, research and technical, as well as industrial center of synchrotron radiation. Till now, the strong national synchrotron radiation communities were realizing their projects abroad in: Cern, Desy (HasyLab), Triest, ESRF, Grenoble, Soleil, Orsay Lure, in France [70], or in USA. The most important aspects of this project are: teaching, technical and industrial.

Now, and during the closest years, the most important accelerator project in the global scale is CERN's LHC – large hadron collider and their detectors: CMS, ALICE,

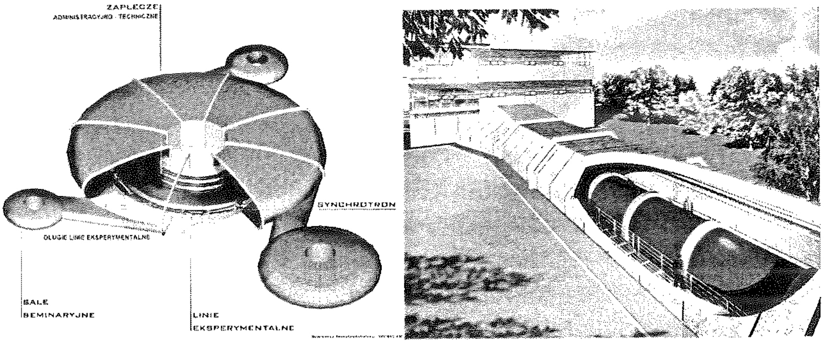


Fig. 1. Visualizations of synchrotron machine in Krakow National Center of Synchrotron Radiation (left) and POLFEL machine localized in the Świerk Park of Research and Technology (right)

ATLAS, LHCb [63]. The participation of research teams from Poland is quite big. A few hundred people take part in the realization of the accelerator, the detectors and accompanying infrastructure. These personnel includes physicists, engineers and technicians. Now there is continued intensive work on the first synchronization of the whole 27 km long cold accelerating loop and trials to launch the first beam. The tests will be done with low intensity beam. The full operability of the LHC (with the full power beam), which searches for the Higgs boson, is predicted in two years time. Next, the upgrade for increasing the luminosity twice should immediately start. The unique experiences gathered by these teams are forwarded to other planned immense international experiments like SLHC, ILC etc. A particularly large team of experts from Poland took part in the design and construction of the CMS - compact muon solenoid detector. A team of physicists and IT engineers takes part in building of a global numerical calculations Grid for the LHC experiment. The project called POLTIER engages a dozen of national IT centers. Indirectly, with these experiments are associated networks of Polish Nuclear Physics and PSAC- Polish Network of Astroparticles.

5. FP7 IA EuCARD PROGRAM

The debated FP7 EuCARD program was preceded by an analogous FP6 program CARE – Coordinated Accelerator Research in Europe [31]. CARE was realized in Europe during 2004-2008. The following research teams from Poland took part: Technical Universities form Warsaw and Łódź – working on the control system on the accelerator, Wrocław University of Technology – working on liquid helium systems, and IPJ – working on plasma sputtering of copper cavities with niobium. As a result of this participation, a modern ELHEP Laboratory of electronics and photonics systems for high energy physics experiments was established in the Institute of Electronic Systems, WUT. In these mentioned institutions a few Ph.Ds and D.Sc theses were accomplished from the research on accelerators. The same institutions, with the similar research teams and technical topics continue their participation in the EuCARD Program.

The EuCARD – European Coordination for Accelerator Research and Development program [73] was accepted by the EU in May 2008 for realization during the period of 2009-2012. It belongs to the FP7 Capacities group of programs and is associated with the rejuvenation and building of the community infrastructure in the area of accelerator technology. A very costly accelerator infrastructure, predicted mainly for fundamental research (as LHC and CLIC in CERN), as well as application research (as FLASH in DESY), in such branches as biology, medicine, material engineering, power, environment protection, safety, must be financed from the community, trans-national resources. The aggregated budget of the program (EU donation and own input) reaches around 35Mln€. The program belongs to a group of IA – Integrating Activities and consists of NA – network activities, TA – trans-national access and JRA – common joint research activities.

The EuCARD program of European accelerator technology is realized by a consortium of 40 research institutions: governmental, academia, and private ones, including the entities form Poland. A number of other associated institutions is combined with the program. The following country representatives participate in EuCARD: England, Austria, Finland, France, Spain, Malta, Germany, Poland, Russia, Switzerland, Sweden and Italy. The involved institutions are mainly universities or big governmental topical laboratories. The participants from Poland are Technical Universities from Warsaw, Wrocław and Łódź, and IFJ, IPJ.

The main aims of the program are: full usage of LHC accelerator potential, increase of LHC luminosity, development of CLIC accelerator, participation in ILC project development of high-intensity ion accelerators and accelerators for FEL lasers, etc. The main research subjects of EuCARD are: superconducting magnets of high field intensity, high temperature superconducting materials, materials and techniques for accelerator beam collimation, ultimate-gradient warm linear accelerators, superconducting RF technology, neutrino factory, alternative methods of particle acceleration and collimation. The realization of program topics is concentrated around the following European accelerator infrastructure: CERN in Geneva (LHC, SPS-CNGS, CTF3), INFN in Frascati (DAFNE, SPARCX), STFC/Cockroft in Daresbury (EMMA), GSI in Darmstadt (SIS, FAIR), DESY in Hamburg (FLASH, PETRA III, TTF), FZD in Drezden (ELBE), FZK in Karlsrue (ANKA), LOA in Paris (plasma accelerator), CEA in Saclay (SupraTech RF), LAL in Orsay (SupraTech RF), BESSY in Berlin (Hobicat), RAL in Oxfordzie (MICE).

6. SUMMARY

- The national, research and technical community of accelerator technology has recently organized themselves in a number of consortia and research networks and technological platforms. Such a social self-organization is now required as these bodies are partners to analogous European structures building together relevant and big research programs. These bodies enable absorption of program funds from operational and framework programs.

- Participation of the research and technical teams from Poland in large European and world wide programs in accelerator technology is absolutely needed for the development of this technology in the country.
- The programs for accelerator infrastructure development in Poland now embrace three big planned projects: Center for Hadron Therapy, National Center of Synchrotron Radiation, Polish Free Electron Laser POLFEL.
- The success in acceptance and then realization of a big European research and development program depends in a big degree on the interaction strength of the national, topical communities, its lobbying influences, the way the community is self-organized, the way it is bound to the international community, etc.

7. REFERENCES

1. D. A. Edwards, H. J. Syphers: *An introduction to the physics of high energy accelerators.* Wiley, 1993.
2. M. Reiser: *Theory and Design of Charged Particle Beams.* Wiley, 1994.
3. H. Padamsee, J. Knobloch, T. Hays: *RF superconductivity for accelerators.* Wiley, 1998.
4. A. W. Chao, M. Tigner (editors): *Handbook of Accelerator Physics and Engineering.* World Scientific, 1999.
5. G. Guignard (editor): *[The CLIC Study Team], A 3 TeV e^+/e^- Linear Collider Based on CLIC Technology.* CERN Report, July 2000.
6. E. J. N. Wilson: *An introduction to Particle Accelerators, Oxford.* 2001.
7. R. Brinkman, et al. (editors): *TESLA Technical Design Report.* DESY, March 2001.
8. K. Wille, J. McFall: *The Physics of Particle Accelerators: An Introduction.* Oxford, 2001.
9. R. C. Davidson: *Physics of Intense Charged Particle Beams in High Energy Accelerators.* World Scientific, 2001.
10. M. G. Minty, F. Zimmerman: *Measurements and Control of Charged Particle Beams.* Springer, 2003.
11. S. Y. Lee: *Accelerator Physics.* World Scientific, 2004.
12. S. I. Tzenow: *Contemporary Accelerator Physics.* World Scientific, 2004.
13. G. S. Nusinovich: *Modern Microwave and Millimeter-Wave Power Electronics.* Wiley, 2005.
14. S. E. Tsimring: *Electron Beams and Microwave Vacuum Electronics.* Wiley, 2006.
15. H. Widemann: Particle Accelerator Physics, Springer, 2007.
16. A. Sessler: *Engines of Discovery: A Century of Particle Accelerators.* World Scientific, 2007.
17. ILC Reference Design Report, CERN, Fermilab, KEK, DESY, August 2007.
18. T. P. Wangler: *RF linear accelerators.* Wiley, 2-nd edition, 2008.
19. A. Chao, W. Chou: *Reviews of Accelerator Science and Technology.* vol.1, World Scientific, 2008.
20. CERN [cern.ch];
21. Fermilab [fnal.gov];
22. SLAC [slac.stanford.edu];
23. ORNL [ornl.gov];
24. Jefferson Lab [jlab.org];

25. DESY [desy.de];
26. INFN [infn.it];
27. IN2P3 [in2p3.fr];
28. KEK [kek.jp];
29. PCPS [synchrotron.pl];
30. POLFEL [polfel.pl];
31. ESGARD [esgard.lal.in2p3.fr];
32. IPJ [ipj.gov.pl];
33. IFJ PAN [ifj.edu.pl];
34. SLCJ UW [slcj.uw.edu.pl];
35. ECFA – European Committee for Future Accelerators [committees2.web.cern.ch/ Committees2/ECFA];
36. ICFA – International Committee for Future Accelerators [fnal.gov/directorate/icfa];
37. IUPAP [iupap.org];
38. ACFA – Asian Committee for Future Accelerators [acfahep.kek.jp];
39. PAA [www.paa.gov.pl];
40. PAN [pan.pl];
41 PTF [ptf.fuw.edu.pl];
42. PTPS [synchrotron.org.pl];
43. PTN [ptn.nuclear.pl];
44. EPAC [epac.web.cern.ch/EPAC/Welcome.html], [epac08.org];
45. ICALEPCS [icalepcs.org];
46. JACOW [accelconf.web.cern.ch], [jacow.org];
47. INDICO [indico.cern.ch];
48. CORDIS-FP7 [cordis.europa.eu/fp7];
49. EU-JRC [www.jrc.ec.europa.eu];
50. EURATOM [cordis.europa.eu/fp7/euratom];
51. ESFRI [cordis.europa.eu/esfri];
52. NCBiR [ncbir.gov.pl];
53. ELAN [elan.desy.de];
54. BENE [bene.web.cern.ch/bene], [nu.to.infn.it/Neutrino_*super*_Beam];
55. HHH [care-hhh.web.cern.ch];
56. EuDET [eudet.org];
57. FITAL [fital.fuw.edu.pl];
58. ILC – International Linea Collider [linearcollider.org];
59. TESLA TC [tesla-new.desy.de];
60. ILC Warsaw [hep.fuw.edu.pl/ilc];
61. E-XFEL-ILC Warsaw Group ILC News Line, 5 July 2007 [linearcollider.org/newsline/archive/ 2007/20070705.html], Warsaw ELHEP Group [desy.de/ rrom];
62. ITER [iter.org];
63. LHC [cern.ch/LHC], [lhc.fuw.edu.pl], CMS [cms.cern.ch], ATLAS [atlas.cern.ch];
64. FLASH – Free Electron Laser in Hamburg [flash.desy.de];
65. E-XFEL – European X-Ray Free Electron Laser [xfel.desy.de], [xfel.eu];
66. XFEL Polska [xfel.pl];
67. PSFN, Konsorcjum Femtofizyka [slcj.uw.edu.pl/pnpn/en/25.454.html];
68. GSI Darmstadt, FAIR [gsi.de/fair];
69. J. Sekutowicz: *Multi-cell Superconducting Structures for High Energy $e^{+-}e^{-}$ Colliders and Free Electron Laser Linacs*. Publ.House of WUT, 2008.
70. ESRF – European Synchrotron Radiation Facility [esrf.eu];
71. PPTN [platforma.centrumatomistyki.pl];
72. CEAT [public.ipj.gov.pl/ceat]; [73.] EuCARD [eucard.web.cern.ch/EuCARD];

Dev

MA

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 253-274

Development of the digital image data acquisition system for astronomical applications

MACIEJ KWIATKOWSKI[1], GRZEGORZ KASPROWICZ[1], KRZYSZTOF POŹNIAK1, RYSZARD ROMANIUK[1], GRZEGORZ WROCHNA[2]

[1]*Institute of Electronic Systems*
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warszawa
[2]*2Soltan Institute for Nuclear Studies, Świerk*
e-mail: Maciej.Kwiatkowski@cern.ch
K.Pozniak@elka.pw.edu.pl; rrom@ise.pw.edu.pl

Received 2008.06.10
Authorized 2008.07.25

The paper describes a piece of a development process for a new generation of high performance, remote control, CCD camera system designed for astronomical applications. A com-pletely new image acquisition system residing on a highly integrated PCBs was designed, man-ufactured, tested and commissioned. The design process lead through largely redundant PCB design and tests with implemented such features like NIOS, CCD emulator, and a number of open source software. The CCD sensor was positioned in a different way than previously, resulting in better performance of the astronomical video data acquisition system. The camera was built using a low-noise, 4Mpixel CCD circuit by STA. The electronic circuit of the camera is highly parameterized and reconfigurable, as well as modular in comparison with the solution of first generation, due to application of open software solutions and FPGA circuit. New algorithms were implemented into the FPGA chip. There were used the following advanced electronic circuits in the camera design: High speed USB2.0 interface, dedicated image processor, GigEth interface, SDRAM and FLASH memories, ARM926EJ-S ARM processor. Software solutions for the camera and its remote control, as well as image data acquisition are based only on the open source platform. The camera supports both API V4L2 driver and INDI protocol. It will be replicated in 20 pieces and is designed for continuous on-line, wide angle observations of the sky in the Pi-of-the-Sky experiment. The full system, largely reduced in redundant functionalities, thus, with increased reliability, is now under production stage and will be soon implemented in real work conditions.

Keywords: optical astronomy, high performance CCD cameras, optical counterparts to GRB, FPGA based image data acquisition systems, ARM microprocessors, NIOS, ISI interface, AMBA standard, INDI interface

1. INTRODUCTION

In recent years astronomy has evolved from static data recording into observations of dynamic objects varying in short (from the human point of view) time scales. Optical counterparts of gamma ray bursts (GRB) are good examples of such phenomena. These events occur a few times a day and are placed in an allegedly isotropic way in the sky. This implies simultaneous observation of large areas of the sky. The time of gamma emission varies between milliseconds and minutes, hence it is important to observe the optical counterparts with a detector which has time resolution of the order of single seconds. High optical and time resolution of the detectors generates large amount of data at the storage. The CCD cameras are typical examples of such hejdetectors. Their data stream is processed on-line in the system in order to reduce the amount of data to be stored. That approach increases the overall efficiency of the detector. The CCD cameras are typical examples of such detectors. Their data stream is processed on-line in the system in order to reduce the amount of data to be stored. That approach increases the overall efficiency of the detector.

Major aim of the “Pi of the Sky” [1] Project is to observe the sky for the search of optical counterparts to the GRB [2,3,4]. Differently to other similar projects, the ”Pi of the Sky” assumes a continuous observation of the whole sky, without any need to move a large telescope [5] after the discovery of the burst by the satellite network GRB Coordinates Networks (CGN) [6]. In order to be able to do this task, a set of 32 cameras is manufactured. They will be organized in two separately mounted matrix sub-sets containing 16 pieces each [7]. Now a simple pilot sys-tem is functioning, consisting of two cameras, localized in Las Campanas Southern European Observatory in Chile [8]. The designed and manufactured cameras are based on CY7C68013a microcontroller [9], compatible with the 8051 core, in fast version with rich peripheries, including among others the full-speed USB 2.0 device. The camera is equipped in gigabit Ethernet inter-face realized in the RTL8169s controller [10] with PCI bus. Data transfer between SDRAM (32MB) memory and Ethernet controller is realized via the PCI controller implemented in FPGA chip Cyclone I (EP1C6) by Altera [11]. FPGA realizes also the control of charge shift in CCD and manages digital data acquisition and stores them in SDRAM memory. The video signal con-verted by the AD9826 image processor [12]. FPGA realizes also the control of charge shift in CCD and manages digital data acquisition and stores them in SDRAM memory. The video signal is converted by the AD9826 image prcessor [12].

The efficiency of the system of first generation (with two cameras) is satisfactory, but its development possibilities are rather confined. The system consists of many strictly dedicated solutions including complicated VHDL blocks (as for example con-trollers Master/Slave PCI, SDRAM) and of 8051 microcontroller. Software for the latter was written in assembler. The software includes, among others, implementations of network protocols like IP, UDP and dedicated NUDP protocol. The degree of system complication confines the possibilities of free system development but also its efficient debugging, to a narrow circle of experts, who have been working with the system from several years. The new generation of the camera assumes the usage of widely applied technologies and protocols based on *Open Source* standard. All the design changes were mainly possible due to application of AT91SAM9260 (ARM926EJ-S) CPU type microprocessor by Atmel [13] which can work in the Linux OS environment.

2. CCD EMULATOR

Most of the CCD sensors are not protected against the over-voltage conditions and ESD. They are also very fragile to the output buffer current overload – such condition causes amplifier damage. There are many occasions for such events during the development cycle. It is not re-quired to have full, low noise functionality of the sensor for all the time. For example, during timing adjustment or the digital interface debugging, most of the sensor's low noise properties are not used. Instead, the other source of the picture data could be used. The emulator allows full-frame image generation (2048 *times* 2048 × 16 bit) based on built in simple test patterns or user de-fined ones, maximum 2048×2048 pixels. It has two independent video channels that can simulate dual output sensors, e.g. STA0820 (STA) or TH7899 (Atmel). All the critical voltages measurement features check if the camera meets sensor requirements. Stand-alone or PC-controlled operation makes the emulator a very versatile device. There is possible emulation of the CCD sensors other than STA0820 and Fairchild CCD442A used in the "Pi of the Sky" project cameras. In such a case, simple header is required. Galvanic isolation of the device eliminates possible problems with ground loops. Fig.1. shows the block schematic of the emulator and Fig.2. manufactured PCB of the CCD emulator. Its heart is a 8 bit AVR processor. It is running the firmware which controls all the peripherals. In addition, the device contains the user interface – keyboard and LCD text display. Thanks to that, it is possible to operate it in the stand-alone work mode. It is useful specially while low noise measurements during which all the external connections should be reduced in order to avoid the ground loops. Two communication interfaces allow connection to the PC with control application which can be used to set device parameters and download the test patterns.

The emulator was successfully build and tested with a real camera hardware [14]. The firm-ware and PC GUI software is still under development, but since the device can work stand-alone, it is not critical part of the project. The emulator will be used during the final assembly and tests of the detector based on 2× 32 cameras. The tests have proven the emulator a very powerful tool outperforming the initial concepts. It makes possible to take measurements not anticipated in the specification, as of, for instance, frequency characteristics of the camera low noise am-plifier. The future enhancements of the project include: further reduction of the readout noise, implementation of another

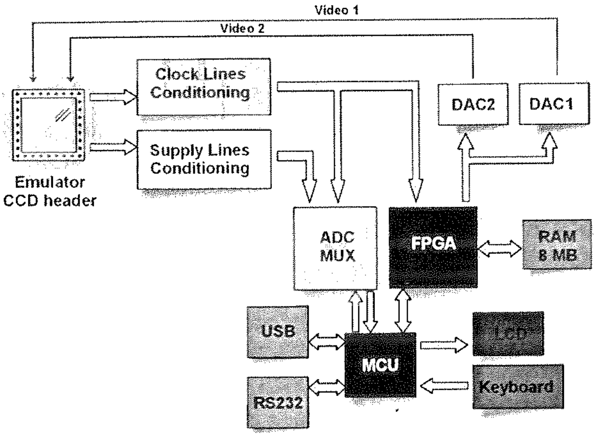


Fig. 1. Block diagram of CCD camera emulator

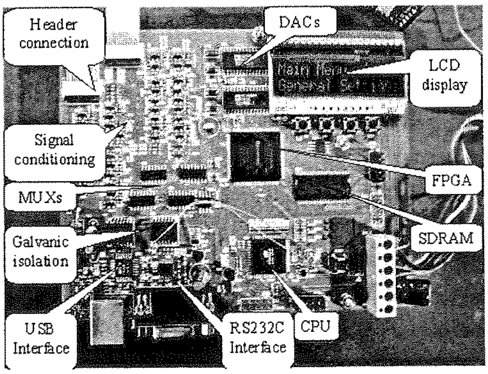


Fig. 2. PCB of CCD camera emulator

CCD emulation headers, increase of the maximum DAC and memory data throughput, reduction of the emulator latency. Universal character of the device, thanks to usage of the FPGA technology, makes it useful also in the other applications, not connected directly with the CCD cameras development. After software modification, the device can become a powerful and precision arbitrary signal generator with extremely long record time. This feature enables completely new applications.

3. NIOS II IN CCD CAMERA SYSTEM

The Nios II system implementation intended for "Pi of the Sky" cameras is shown in Fig. 3. [15]. All the blocks of the system, except for the CCD charge shift state machine and the custom PCI arbiter, are IP cores provided by Altera. Thanks to that, the time and effort of the hardware design could be significantly reduced. It is easy to notice that the external microcontroller is not necessary in this hardware implementation.

Embedded processor inside the FPGA can initiate the RTL8169s internal registers as well as interpret incoming device commands. In the case of exist-ing PCB boards, the external microcontroller can be used as a USB 2.0 bridge if required and if there is enough logic resource to implement support logic inside the FPGA. The CCD charge shift state machine can be adopted directly from the current version of the system. It is necessary only to equip it with the Avalon bus slave interface, so that the Nios could copy converted pixels (after analog to digital conversion in an external ADC) to the SDRAM memory.

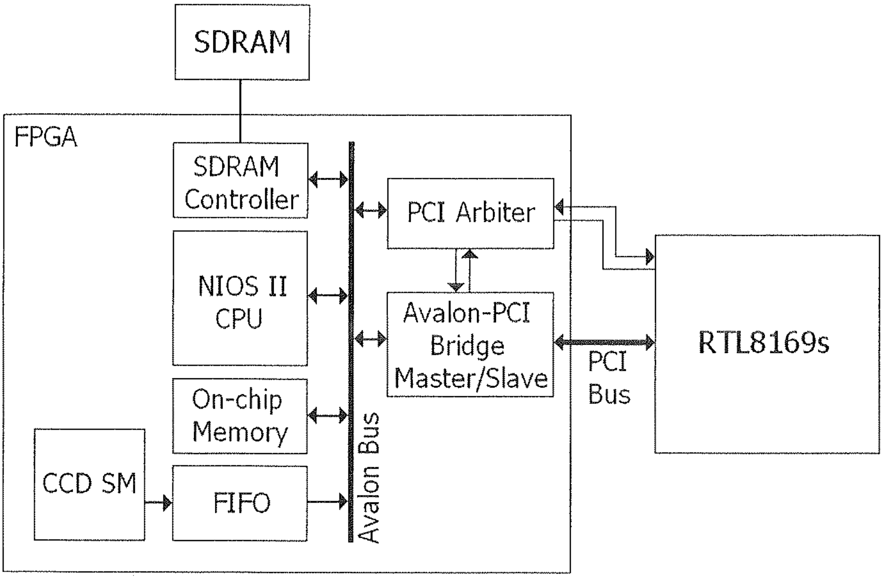


Fig. 3. NIOS II implementation in CCD camera

In case of sufficient amount of the FPGA logic resources, the CCD state machine could be equipped with the master Avalon bus interface. The Master Avalon bus interface together with DMA would allow direct data transfer without of the CPU intervention. Unfortunately, the FPGA used in current design is not big enough. That's why it is necessary to implement only quite confined version of the Nios II based system. It was obligatory to use the PCI Host Bridge because the processor resides on the Avalon bus side. Because in that system only two PCI devices exist (RTL8169s and Avalon-PCI Bridge), it is not necessary to implement bridge with multiple pending transactions that saves a little bit of logic resources. In spite of that, the Nios II processor was implemented in a minimal version without a support of the cache memory. On the Avalon bus there are two master devices (Nios II and Avalon-PCI Bridge) which can try to access the same memory in the same time so a small on-chip memory is required. In that memory there will be stored the fragment of a code which will be executed by the Nios II while the data transmission between Avalon-PCI Bridge

(RTL8169s) and main system memory (SDRAM). In the existing design there is only one memory available. This memory must be shared between the instructions of Nios II and the data. In that SDRAM memory all descriptor lists and buffers with data for the Ethernet controller will be placed.

Program fetching by the Nios II from the SDRAM will not be possible during the data transmission between the system main memory and the RTL8169s. Custom PCI arbiter is equipped with the Avalon bus slave interface which gives the Nios II ability to have the influence on PCI bus arbitration process. Using this feature, the Nios II can decide about granting or retaking access to the PCI bus. Additionally, the arbiter generates an interrupt to the processor when the RTL8169s receives the access to the PCI bus. The Interrupt handler should be placed in small on-chip memory. The implementation of a command parser as well as bootloader with the framework of interrupt handler is to be considered. The bootloader should then reload part of the on-chip memory with the handler of the received command.

All the command handlers can be placed in the main system memory and should be loaded to the on-chip memory only in case of the data transaction between the SDRAM and the RTL8169s. When there is no such data transaction, the instructions of the command handlers can be fetched from the main system memory. It is also possible to give up the custom arbiter on the PCI bus and to pass the arbitration process to the default Avalon bus arbiter. During the system design phase with the SOPC Builder it is possible to define arbitration priority levels for the Avalon master devices. For example defining priority level 4 for instruction master interface of the Nios II and priority level 1 for master interface of the Avalon-PCI Bridge will cause that on every 4 fetched instructions there will be one transaction between the SDRAM and the RTL8169s. Both arbitration conceptions will be tested during the software design stage and a better solution will be chosen. The complexity of the FPGA design, difficulties with debugging, licensed IP-blocks and non-standard solutions caused that this conception was finally resigned.

4. INDI, V4L2, ISI AND AMBA STANDARDS FOR CCD CAMERA SYSTEM

New implementation of the camera [16-18], including hardware and software parts, was carried on a specially designed development and redundant PCB. The PCB is presented in Fig.4. The board has Image Sensor Interface (ISI) outputs [19]. The ISI is compatible with the digital video signal transmission standard by ITU: ITU-R.BT601 [20]. The processor AT91SAM9260 is equipped with interface that supports this standard. ISI is designed for a direct cooperation with CMOS sensors. Usage of the ISI interface is possible after the analog video signal from CCD output is processed and managed by simple logical circuits in the FPGA. The simple possibility of ISI usage was not obvious from the beginning of the system design. There is only little data in the relevant web resources concerning the usage of ISI in this role and in a similar configuration. Due to these doubts, the development board has the ability to map the

FPGA circuit as a device visible in the address space of the CPU microprocessor, connected to the external bus. The FPGA chip was equipped with external SDRAM memory, serving as a buffer for image data storage.

In such a configuration, a state machine implemented in FPGA would carry image data ac-quisition in the SDRAM. The CPU would then initiate DMA transmission from the FPGA, visible as a device on the bus, to its own SDRAM memory.

After some encouraging initial tests with the usage of the ISI image interface, it is known that there is no need for a non-effective double data buffering and that the final camera PCB would be simplified. The tests have proven that there is no need for intermediate SDRAM memory buffer, and the FPGA circuit can be considerably smaller.

5. CAMERA SYSTEM CONSTRUCTION

The applied processor AT91SAM9260 by Atmel is based on the RM926EJS core. It has a rich set of peripheral devices. The core ARM926EJ-S, compatible with the architecture ARMv5TEJ, is characterised by the following features: hardware support for Java language, additional set of intructions for digital signal processig (DSP), unit of memory management MMU, internal bus compatible with Advanced Microcontroller Bus Architecture AMBA standard [21] and AHB – Advanced Highperformance Bus. The latter feature enables simultaneous data transfer between numerable devices of master and slave type. The kinds of master and slave devices and all possibilities of transmission between them are presented in Table 1. One of the AHB master type peripheries is image ISI interface. There are numerable advantages of this solution, which will be debated in the following chapter.

Table 1

Transmission configuration in AT915AM9260 AMBA Bus

	Master	0	1	2	3	4	5
Slave		Instructions ARM926	Data ARM926	Controller DMA	Host USB	Controller ISI	Ethernet
0	Internal SRAM0 4kB	+	+	+	+	+	+
1	Internal SRAM1 4kB	+	+	+	+	+	+
2	Internal ROM	+	+	+	+	-	-
3	External bus	+	+	+	+	+	+
4	Internal peripheries	+	+	+	+	-	-

Two other AHB devices of master type are: data bus and processor instruction bus. The data as well as the instructions may be taken from arbitrary memory: in-

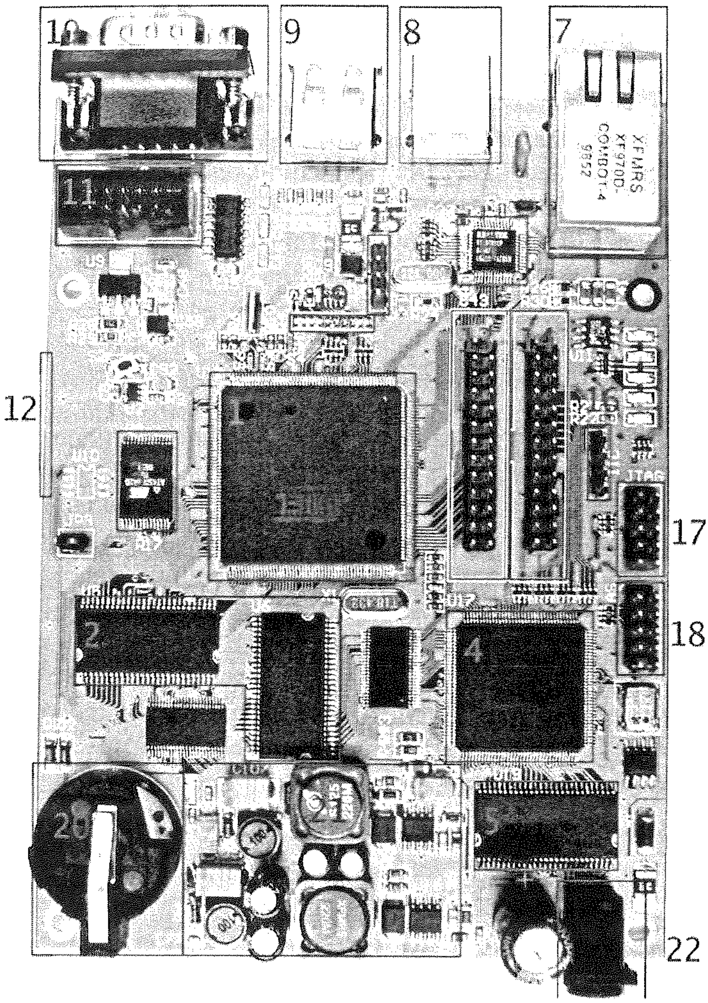


Fig. 4. Photo of a development board for astronomical camera 1. Processor AT91SAM9260; 2. SDRAM memory of processor 64MB 32bit (2×32MB 16bit); 3. DataFlash 8MB memories; 4. FPGA Cyclone EP1C3 circuit; 5. SDRAM memory of FPGA (32MB 16bit); 6. Circuit of the physical layer of Ethernet (PHY) RTL8201; 7. Ethernet socket 10/100; 8. USB socket Full-Speed Device; 9. USB Full-Speed Host socket; 10. RS232 socket; 11. Debug UART socket; 12. MMC socket; 13. ISI processor's connector; 14. ISI connector of FPGA chip; 15. UART connector; 16. I2C connector; 17. JTAG connector of FPGA chip; 18. Active Serial connector; 19. JTAG connector of processor; 20. Clock support battery socket; 21. Power supply section; 22. Power supply connector

---

ternal S
connect
a simpl
valid pr
in the f
develop
progran
bootloa
overcon
confine
nal SDI
U-boot

No
pending
possibil
comple
The U-
among
family.
mass m
fer Pro
debugg
The op
method
used to
the cor
kernel

Li
The ke
mation
is poss
also m
beacus
single
file sys
stored
may be

It
configu

ternal SRAM memory (2x4kB), internal ROM memory (32kB) or external memory connected to the external bus (like SDRAM). The nonvolatile ROM memory contains a simple bootloader (placed there by the processor manufacturer), which seeks for a valid program (beginning with instructions for jumps in the table of interrupt vectors) in the flash memory connected to the SPI bus, and loads it to SRAM memory. The development board features a serial flash memory AT45DB642 (64Mb) [22], with programmed bootloader, provided by the manufacturer (At91 bootstrap ver 1.6). The bootloader requires only a simple configuration and compilation. Its size may not overcome the size of the internal SRAM memory, which is 4kB. Therefore its role is confined only to initiate the clock of the main processor (PLL), to configure the external SDRAM memory (64MB) and to load from the serial flash memory the following U-boot bootloader.

Now after supplementing, the bootloader is much more advanced application. Depending on the configuration, it may take a few hundreds of kB, and thus, there is no possibility to reduce it to the required 4kB. It is loaded to the SDRAM. Due to its complexity it is necessary to configure the bootloader to the final system architecture. The U-boot was written based on the sources of Linux kernel. It may load and execute, among other functions, the Linux kernel. It may execute other programs outside the OS family. The bootloader has implemented a service for many devices and protocols like: mass memory – IDE, MMC, network devices and protocols -TFTP (trivial File Transfer Protocol), simple shell with many instructions which may be helpful at hardware debugging – view and modification arbitrary memory areas and processor registers. The operational system may be executed from the level of U-boot using numerable methods. To load Linux kernel, the TFTP protocol was used. It is a protocol frequently used to start stations without hard disks. After loading the kernel, U-boot gives out the control, via a jump to the loaded with the kernel, a decompression program. The kernel is decompressed and the system starts.

6. CAMERA SOFTWARE

Linux kernel is supporting its auto-configuration with BOOTP/DHCP protocol. The kernel, during the start-up, receives from the DHCP server the following information: own IP address, IP address of NFS server, access path to the file system. It is possible to mount the file system which is made accessible by the NFS server, or also mount it from the local MMC card. The network solution is more convenient beacuse provides a central management of many computers and their resources. A single server may store the following data for all 32 cameras: Linux kernel, separate file systems, including all configuration, start-up files, etc. Camera images may be stored on the same or separate NFS servers. The individual file and catalogue systems may be arbitrarily configured and structured.

It is possible to create arbitrary scripts for Linux shell, to manage the whole configuration and to automate its execution. The configuration space of the 32 camera system is rather wide. The embedded Linux in each camera enables compilation and execution of arbitrary ready application like streaming video, web server and other more dedicated application written by the user. A number of ready drivers may be immediately used, which have direct access to the hardware, like: I2C, SPI, GPIO. One may implement own kernel modules. A dedicated driver was written for the ISI image interface. The driver is compatible with the standard "Video for Linux v2" (V4L2 API) [23]. This module registers as a video device and it also reserves one of the SPI interfaces of the processor. The ISI interface serves for sending the video data stream, while the interface SPI serves for setting of the internal camera registers like: status, start/stop control, exposition time, gain, etc.

The interface I2C is used most frequently to communicate with the registers of integrated CMOS sensors. Also, a simpler SPI interface may be used for this purpose. The camera registers, the control state machine for image for image data transmission and acquisition from CCD sensor, all were implemented in the FPGA circuit residing on the camera digital board. The compatibility of the driver with V4L2 enables usage of an arbitrary video application, which makes the camera system very flexible and cooperative.

7. ISI VIDEO INTERFACE

The signals in the ISI interface are shown in Fig.5. The pixels data are sent serially via data bus which is 8-bit wide DATA[7..0] and is synchronized by a clock ISI_PCK (pixel clock). The signals ISI_VSYNC and ISI_HSYNC serve for synchronization of the beginning and end of the frame for each image line. Between the valid lines and frames, there are sent additional data (blanking intervals). All signals, together with the pixel clock are generated by the image sensor. The market offers a wide variety of CMOC image sensors, thus, the ISI interface may be configured differently. The configuration includes signal polarization, data format, etc. The accepted image formats are: RGB 8:8:8, RGB 5:6:5, YCbCr 4:2:2 and gray scale up to 12-bit resolution.

The ISI interface has two possible data paths: the first is codec path predicted for image data transfer for data storing in file, the second is preview path with frame scaling ability, predicted for visualization of the image on the LCD screen. Both paths have independent FIFO queues. The data are sent, via the master AHB interface of the AMBA bus, directly to the pointed memory area. The preview path has the ability to set a list of pointers for successive frames to be displayed. The preview path converts data to the RGB format which is accepted by LCD displays. The codec path converts image data to YCbCr format, where the luminance components may be under-sampled – loss compressed. The conversion coefficients are set in the internal registers of the ISI interface.

The K20 camera is equipped with a monochromatic CCD sensor by STA [24], of 2048×2048 (+14 dark columns) resolution. Very good noise parameters of the analog channel of the camera allowed for application of an image processor which samples

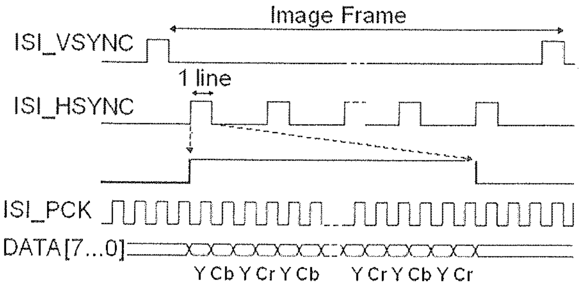


Fig. 5. Signals in ISI interface

the analog video signal with the resolution of 16-bits. It is not advisable to reduce the acquired image resolution to 12-bit, which is offered by the ISI interface in the grey scale mode. Therefore, to send data to the memory via the ISI interface, the YCbCr 4:2:2 mode was used. This format is as follows: 4 samples of luminance Y, 2 samples of chrominance Cb and 2 samples of chrominance Cr. Four samples of the image (pixels) in this format are stored via 8 bytes (samples are 8-bit). Exactly the same number of bytes is needed to store 4 pixels of a monochromatic image of the 16-bit resolution.

A drawback of such a solution is the lack of usage of the preview path with scaling ability. However, the ISI interface, in a monochromatic work mode, offers also a single data path, without the ability of hardware conversion. The maximum size of the transmitted image is 2048×2048, thus, the additional dark columns, which are accessible in the CCD sensor of STA, may not be transmitted. It is not a big drawback because they are used only during the device calibration procedures. Additionally, a test mode of the camera may be implemented, where a shift for the read out image may be activated, to uncover the dark columns, at the cost of loss of the equivalent number of active columns. A big advantage of the ISI interface is the ability to sent a big amount of data totally in the hardware domain, not engaging in any way the CPU. A diagram of data flow, starting from the analog signal, is presented in Fig.6. The video signal may be read out simultaneously from two outputs of the CCD chip, which increases two times the readout speed. The analog signal, after pre-amplification, is processed by the image processor AD9826. The charge shift clock of the CCD chip and other control signals of image processor are generated by the state machine implemented in FPGA chip. The same state machine receives processed image data and sends them via the ISI interface.

The external sensor, or as like in this case – the FPGA circuit, plays a role of a mater device, i.e. sends data on the bus DATA[7..0], generates pixel clock and synchronization signals H and V. The ISI interface is visible as a master device from the side of the ARM processor. The processor has to be fast enough to receive the sent data and forward them to the SDRAM memory. This process of fast data transmission is considerably facilitated by an inbuilt DMA controller with the ability of transmission initialization in the internal bus of the processor (AHB master device). It is necessary to set properly the priorities of arbitration in the AHB bus, to avoid conflicts with other devices of master type which may request the usage of SDRAM memory. The same memory serves for storing processor instructions and data. The most critical situation leads to the ISI internal buffer overflow (or a buffer of any other device of master type) and data loss. During commissioning tests of the camera, after proper initialization of the AHB bus, such situation was never observed. The inbuilt FIFO register queues and independent cache memories for the processor data and instructions contribute essentially to the increase of the system efficiency. The development board, in the case of not mounted CCD chip, was tested by using pattern images generated by the FPGA chip.

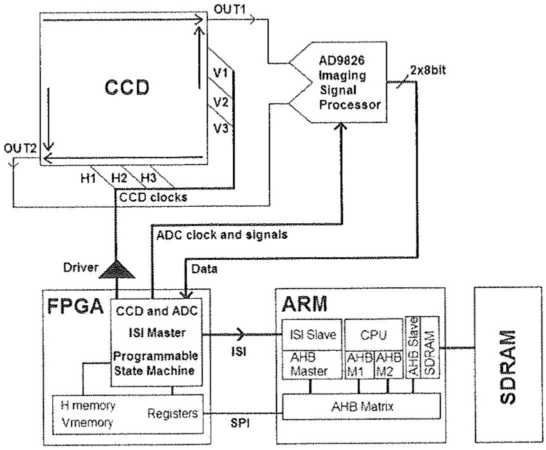


Fig. 6. Block diagram for data flow on the astronomical camera board

The development board has the ISI external connector. The old K20 camera, which was used for initial tests, has test pins, which were used to connect it with the ISI connector of the development board processor. Thus, it was possible to test the physical readout of the CCD chip. There were considerable problems with glitches and other interferences generated in the connecting cables. Test results were, however, favorable. The processor uses SPI bus for communication with FPGA chip registers. The registers are responsible for camera settings and control like: exposition time, start/stop of CCD chip readout, status checking. The FPGA chip has 2 small memories which store patterns of horizontal and vertical CCD signals. The programmable state machine is responsible for generation of control signals of charge shift in CCD chip. Thus, it is possible to generate, by a set of graphical tools, binary files describing the control signals and loading them during the work of the device.

8. V4L2 COMPATIBLE VIDEO DRIVER

Video device driver in Linux has a standardized structure *video_device*, which wraps, a well known structure *file_operations* from the area of character device drivers. Similarly to character devices, the video device must be registered in the system. This is done by the function *video_register_device*. This function requires, initiated previously, a *video_device* structure and an identifier determining a type of the device. A family of video devices is very broad. They may perform a lot of functions like: input or output, device type video streaming, data transmission during *blanking intervals*, teletext devices, *overlay* devices with direct access to the memory area of the imaging hardware, etc. CCD camera is registered as a device of *VFL_TYPE_GRABBER* type which is non overlay device type.

The structure *file – operations* contains pointers for methods typical for character devices like: *open, release* – these methods configures the device, and *write, read.* Read and write methods should be implemented respectively for the output and input devices, on condition that the device is not of the overlay type. In a device of the *overlay* type, the user applications do not call *read* and *write*, because the video stream is sent directly to the memory, exactly to the frame buffer. For the driver to be compatible with AP1 video ver.2, the ioctl method pointer of the *file_operations* structure should be set to the function *video_ioctl*. This function is an internal method of the V4L2 system. All calls of *ioctl* are redirected by it into separate functions of callback type. The V4L2 system, in 2.6.19 kernel, has 79 separatte functions of this type.

Video devices need initialization and setting of many parameters, to work properly. All these parameters, usually in quite a number, have to be accessible in the user area via the *ioctl* calls. The new version of video API for Linux splits a single driver method into numerable calls, because very frequently a code of the function was too long and unreadable. The pointers of all of these methods are included in the structure *video_device* and their implementation is left to the programmer. Not all of the methods, and usually most of them, need not to be implemented in a particular application. In this case it was 10 methods which were implemented for K20 astronomical camera driver. All the methods are called from the user level with the aid of known *ioctl* method with an appropriate parameter and data structure. The methods may be divided to types *get, set* and listing functions of available options for a particular parameter, like current image resolution, other resolutions availability for the device. This allows video application to list supported resolutions and set appropriate one. Resolution is of course only one example of camera parameters.

All CCD camera settings are included in the registers implemented in the FPGA circuit. The access to this registers is realized by the SPI bus. The camera driver uses one of the SPI ports of the processor. Apart from setting of image parameters it is possible to control the camera like: start/stop exposition of CCD chip, checking of camera status, start of ISI transmission, etc.

ISI transmission is of DMA type. Therefore, the role of the drivers *read* method is confined only to switching on the ISI transmission in the FPGA chip register and to wait for an interrupt after the transmission is over. In the next turn, the buffer is copied to the user area. The buffer for DMA transmission of ISI interface must be a continuous because this device does not support the transmission of *scatter-gather* type, to a non-continuous area of the memory. Linux OS operates on virtual memory, which means the processor physical memory is divided into frames (fragments), which are non-continuous in the physical memory. The virtual memory increases the system ability to operate in the wider address space. The address space for Linux is 4 GB, where is 1GB is intended for the kernel area and 3GB is intended for the user area. The address space does not depend on the available physical memory which is in the case of described here camera only 64MB. As a result of the allocation functions calls, like *malloc* in the user area, and *kmalloc* in the kernel area, a determined number of memory frames is assigned to the process. These frames are not continuous in the physical memory.

Some of the DMA devices accept a list of indicators – *scatterlist*. Then it is possible to initiate the DMA transmission to a non-continuous buffer in a physical memory. The ISI interface is unable to do this. Linux offers continuous allocation functions for memory areas, but for allocations of small buffers. The dimension of camera image frame is greater than 8MB. It is possible to enforce the allocation for such a large buffer, for example with the aid of *NOFAIL* flag. It is, however, a brute-force method which may cause system instability because it may force the release of virtual memory frames of important OS processes. In order to obtain a continuous and large memory area a simple method was applied. The *mem* parameter is forwarded to the kernel during the system start-up. The parameter determines the size of available operational memory. Despite the fact that the available memory on the development board is 64MB, the system is informed that only 55MB is available (*mem=55MB*). Then the upper 9MB of the memory is not occupied by the system and would not be divided to pages.

The access to the relevant reserved memory area may be obtained from the driver level. It is done by calling a function ioremap. The address of the gained buffer is forwarded to the DMA controller of ISI interface. After the transmission is over, the buffer is copied to the user area. It is a single drawback of this solution. However, at this size of the memory, only the buffer *scatter-gather* could be mapped directly to the user area. It is not a big drawback since the memory is copied very fast and does not cause any significant delay during the transmission process.

9. INDI PROTOCOL

INDI (Instrument Neutral Distributed Interface) [25] is a dissipated protocol serving for the control of astronomical equipment. It was designed by E.C.Downey from *ClearSky Institute* and is further developed as an *Open Source* project [26]. INDI is based on the architecture client-server-controller, what was presented in Fig.7.

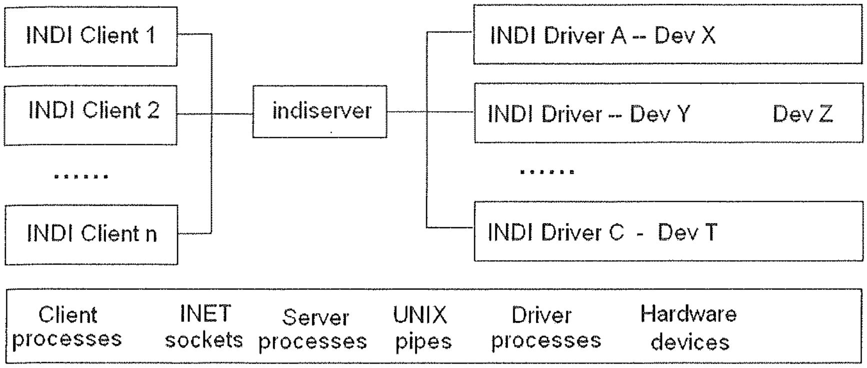


Fig. 7. Block diagram of INDI protocol of client-server-driver architecture type

The server and INDI drivers processes are executed directly on the camera board. Thus, the INDI drivers have direct access to the device, in practice via the driver of OS. The aim of the server is to execute a driver process or many processes like: camera, automatic concatenation, weather station, and directing the requests to these devices originating from a single or multiple user applications. The driver process includes a description of all device functions and is responsible for execution of requests from the clients. A list of functionalities is sent to the client, which generates a dynamic graphical user interface (GUI). In this way, INDI separates a low level device functionality from high level application of the user.

An arbitrary function added to the device must be only described and implemented in the driver, and the client application may turn it accessible for the user. Adding of a new device is confined only to implementation of its INDI driver. The driver consists usually of a single file c or CPP, which is compiled and linked with INDI framework libraries (indidrivermain.o, eventloop.o, liblilxml.a, libindicom.a). With the aid of API function of INDI protocol, all GUI controls are added to the driver source file (buttons, slides, radio-buttons, check boxes, edit fields, binary objects, etc.) Next, the functions of type ISxxx() has to be implemented for all of regivesed them. Those routines are called by main () functions of the INDI framework together with incoming communications messages transmitted from the clients. One of the functions family ISxxx() is a function ISGetProperties(), which sends a list of all mentioned GUI controls, and it is the first function regresed by the client.

The most popular client applications, which implement INDI protocol are *KStars* and *XEphem*. The latter is an extended program, also designed by the *ClearSky Institute* and available as *Open Source*. It may be used free of charge for research or private purposes. Before making connection to INDI server it is necessary to set a proper IP address and a port on which a connection is requested. A window called *INDI configuration* serves for this purpose.

This window enables setting of attribute names for automatic mount controller, in such a way that it is possible to trace a given object in the sky, as well as attribute

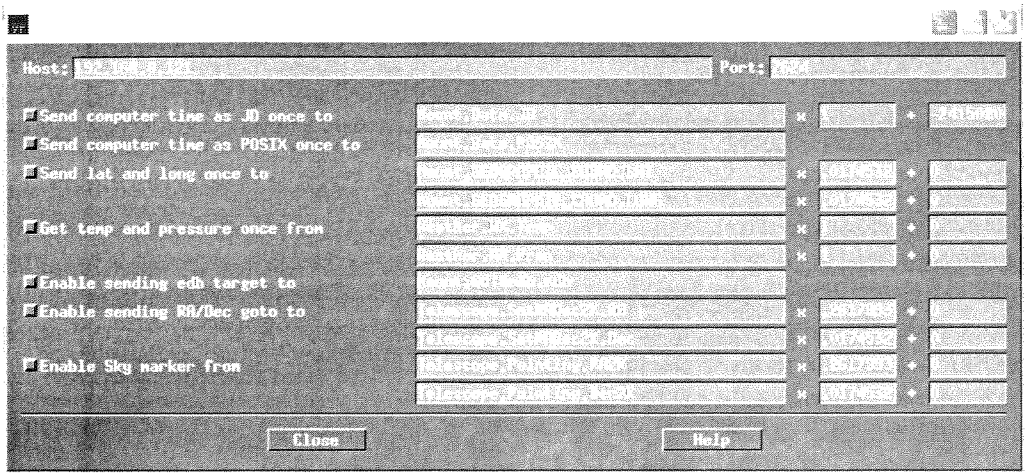


Fig. 8. Configuration panel of XEphems INDI interface

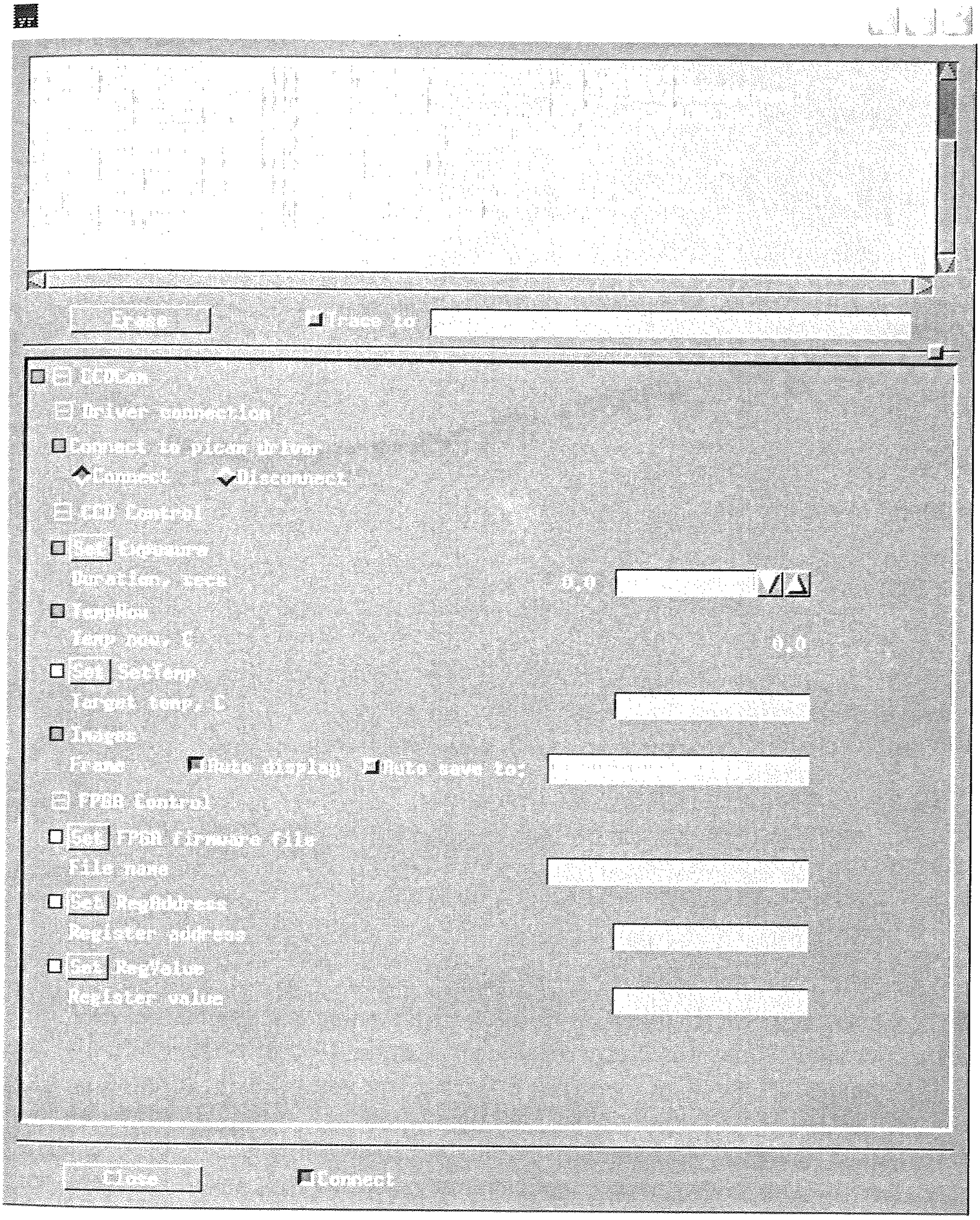
names for weather station. After setting the server parameters the operator goes to the window *INDI control panel* (Fig.9).

In the lower part of this windows, there is a button connect. After success connection to the INDI server dynamic GUI is displayed in the middle part of the window. All displayed controls were previously implemented in the INDI driver. In the upper part of the window, there are printed messages originating from the server and drivers. The control buttons of the first test implementation of the CCD camera INDI driver were divided to three groups: Driver connection, which enables connection with V4L2 driver, CCD control grouping functions combined with the camera servicing and management, and FPGA control group enabling loading of a new firmware for the FPGA chip or access to an arbitrary register of the camera in the test work mode. After the exposition is released *Set Exposure* in the window *Sky View*, there is presented a pattern generated in FPGA circuit embedded in the camera development board, Fig.10.

The INDI server may give access to more drivers for successive devices like: automatic mounting for camera positioning, the weather station monitoring telescope system safety, including protection form wind, rain or snow and excessive moisture, etc. Exemplary drivers controls for other devices are presented in Fig. 11.

10. ADVANTAGES OF ARCHITECTURE OF THE NEW CAMERA

Architecture of the new camera is based on ARM9 processor, which enables on-board running of the Linux OS. Such a solution opens a possibility to use widely accessible applications of *Open Source* type. All software components of the camera are based on *Open Source* standard. Thus, they are transparent and may be openly modified and dynamically fit to the current system needs. Flow of the measurement



Duration, secs 0.0

Temp now, C 0.0

Fig. 9. Control panel of XEphems INDI interface

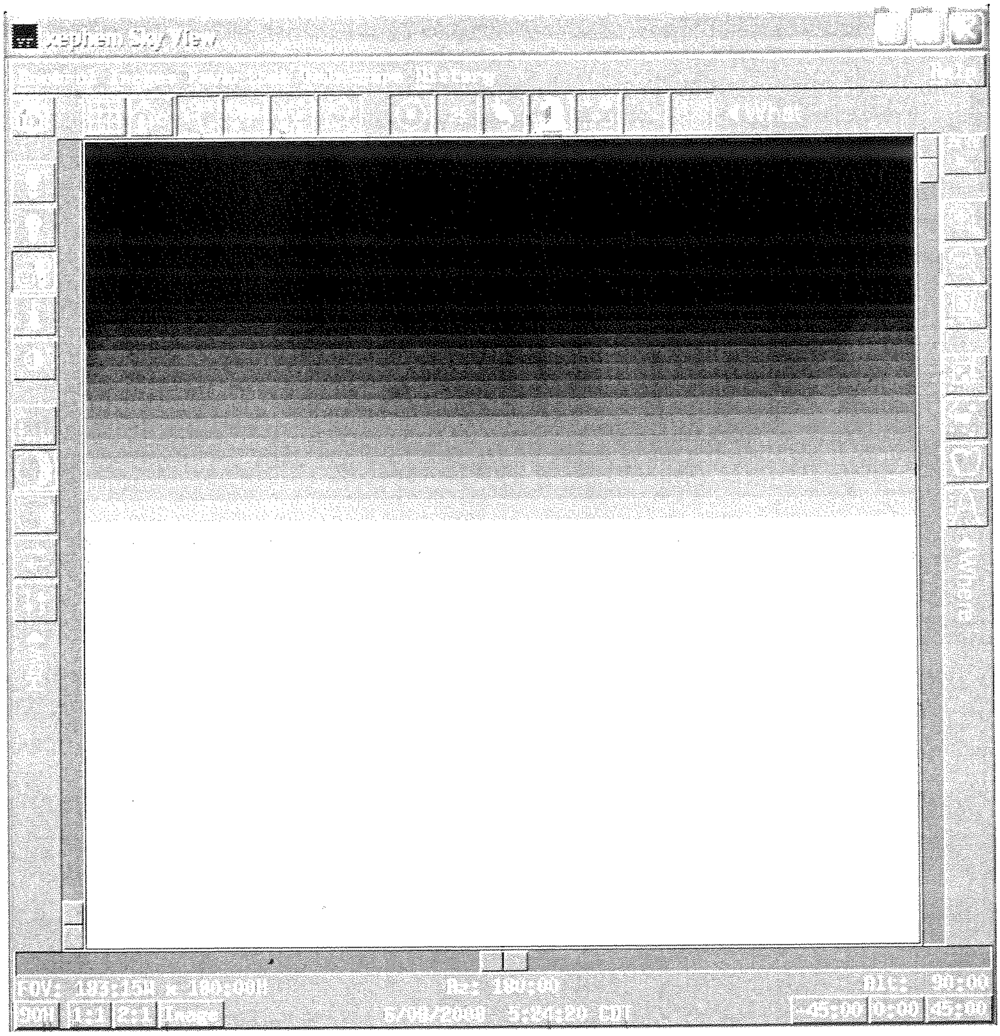
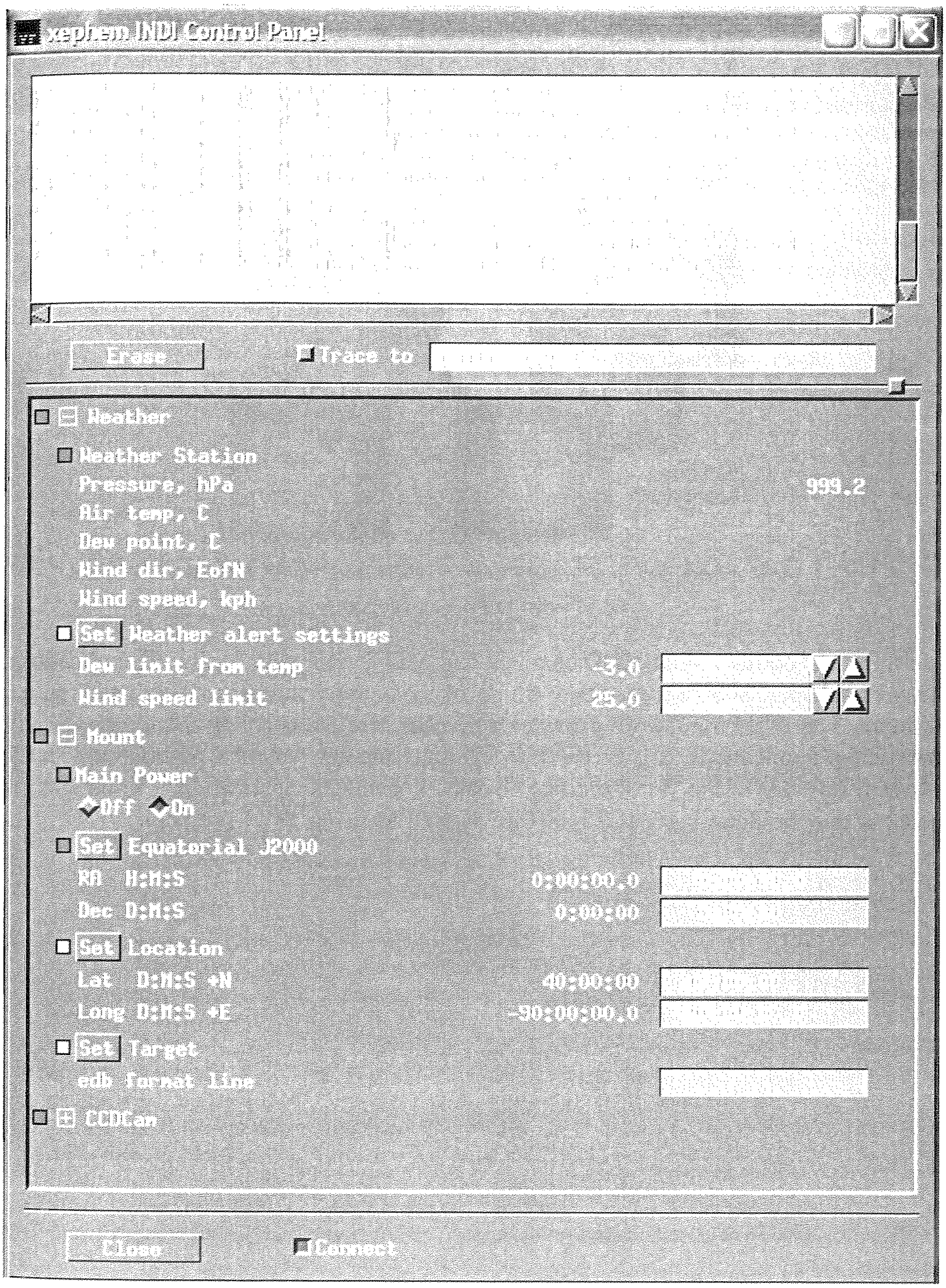


Fig. 10. Sky View window of the XEphem application

data (images) from the CCD chip, via all software components, Ethernet network up to the user application, was presented schematically in Fig.12.

Due to the fact that the camera has a video driver compatible with V4L2 standard, it is possible to run INDI server as well as an arbitrary server for video streaming (VSS) like FFmpeg which implements V4L2 standard. The latter server enables data acquisition in an arbitrary video application compatible with V4L2 like Mplayer. It is possible to run video application directly in the OS of the camera. In our case it is not necessary because the camera set works in a remotely controlled astronomical laboratory. The Mplayer and FFmpeg are able to service only standard controls of the camera, which are compatible with V4L2, such as exposition time, gain, etc. The user



☐ Trace to

☐ ⊟ Weather

☐ Weather Station

Pressure, hPa 999.2

Air temp, C

Dew point, C

Wind dir, EofN

Wind speed, kph

☐ Set Weather alert settings

Dew limit from temp -3.0

Wind speed limit 25.0

☐ ⊟ Mount

☐ Main Power

◆ Off ◆ On

☐ Set Equatorial J2000

RA H:M:S 0:00:00.0

Dec D:M:S 0:00:00

☐ Set Location

Lat D:M:S +N 40:00:00

Long D:M:S +E -90:00:00.0

☐ Set Target

edb format line

☐ ⊞ CCDCam

☐ Connect

Fig. 11. Exemplary INDI drivers

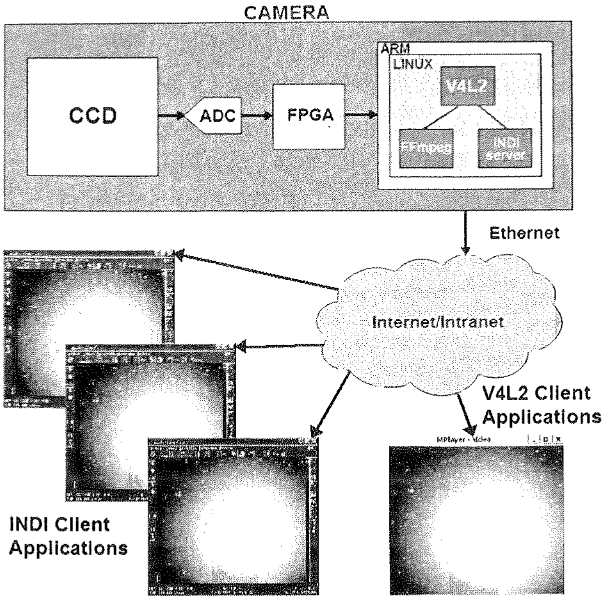


Fig. 12. Measurement data flow (image) from the CCD chip, via all software components, and Ethernet to the user application

controls, like cooling of the CCD chip, are compatible only with dedicated applications. Thanks to Open Source technology fitting of these applications to users needs in not difficult. The INDI framework provides a number of tools enabling implementation of own client applications. It is possible to integrate the new solution of camera with the existing software of the "Pi of the Sky" based on *test2K2K* application. The integration process may be done via V4L2 interface as well as using INDI protocol.

---

1. We
2. A.
M.
K
P i
S z
opt
3. M.
K
N a
n a
200
4. M.
l a
exp
5. A.
w i
wic
6. Str
7. M.
W
8. We
9. Co
10. Co
11. Alt
12. Im
13. Pro
14. G.
i w
Sky
15. M.
NI
Jan
16. G.
w i
U
200
17. M.
Pi
18. M.
n a
con
200
19. Vi
20. Vi
21. AN
22. Me
23. Int

11. REFERENCES

1. Web page of "Pi-of-the-Sky" [grb.fuw.edu.pl]
2. A. Burd, M. Ćwiok, H. Czyrkowski, R. Dąbrowski, W. Dominik, M. Grajda, M. Husejko, M. Jegier, A. Kalicki, G. Kasprowicz, K. Kierzkowski, K. Krupska, K. Kwiecińska, L. Mankiewicz, K. Nawrocki, B. Pilecki, L. W. Piotrowski, K. T. Poźniak, R. S. Romaniuk, R. Sałański, M. Sokołowski, D. Szczygieł, G. Wrochna, W. Zabołotny: *Pi of the Sky – all-sky, real-time search for fast optical transients, New Astronomy*, vol.10, no.5, April 2005, pp.409-416
3. M. Ćwiok, H. Czyrkowski, R. Dąbrowski, W. Dominik, G. Kasprowicz, K. Kwiecińska, K. Małek, L. Mankiewicz, M. Molak, J. Mrowca-Ciułacz, K. Nawrocki, L. W. Piotrowski, P. Sitek, M. Sokołowski, J. Użycki, G. Wrochna: *Search for optical counterparts of gamma ray bursts, Acta Physica Polonica*, vol.37, no.3, March 2006, pp.919-924
4. M. Ćwiok, J. Grygorczuk, J. Juchniewicz, L. Mankiewicz, J. Mrowca-Ciulacz, K. Nwrocki, M. Sokolowski, G. Wrochna: *Status of the full scale Pi of the Sky experiment (invited paper)*, Proc.SPIE, vol. 6347, November 2007, pp.6347.0K.1-4
5. A. Burd, H. Czyrkowski, R. Dąbrowski, W. Dominik, M. Grajda, G. Kasprowicz, L. Mankiewicz, S. Stankiewicz, G. Wrochna: *Low noise CCD cameras fro wide field astronomy*, Proc. SPIE, vol. 6159, February 2006, pp.160-166
6. Strona internetowa satelitarnej sieci koordynacyjnej rozbłysków gamma całego nieba [gcn.gsfc.nasa.gov]
7. M. Ćwiok, L. Mankiewicz, K. Nawrocki, M. Sokołowski, J. Użycki, G. Wrochna: *Full Pi of the Sky system and simulation*, Proc.SPIE, vol 6937, January 2008, pp.693702:1-9
8. Web page of European Southern Observatory [eso.org]
9. Controler CY7C68013A [cypress.com/products]
10. Controler RTL8169s [realtek.com.tw/products]
11. Altera Cyclone EP1C6 circuit [www.datasheetsite.com/datasheet/EP1C6]
12. Image processor AD9826 [analog.com/en/prod]
13. Processor ARM926EJ-S [arm.com/products/CPUs/ARM926EJ-S]
14. G. Kasprowicz, L. Mankiewicz, K. T. Poźniak, R. S. Romaniuk, S. Stankieiwcz, G. Wrochna: *Hardware emulator of the high-resolution CCD sensor for the "Pi of the Sky" experiment*, Proc.SPIE, vol. 6937, Jauuary 2008, pp.6937.08.1-7
15. M. Kwiatkowski, G. Kasprowicz, D. Rybka, R. S. Romaniuk, K. T. Poźniak: *NIOS II implementation in CCD camera for "Pi of the Sky" experiment*, Proc. SPIE, vol. 6937, January 2008, pp.6937.09.1-6
16. G. Kasprowicz, H. Czyrkowski, R. Dąbrowski, W. Dominik, L. Mankiewicz, K. T. Poźniak, R. S. Romaniuk, P. Sitek, M. Sokołowski, R. Sulej, J. Użycki: *New low noise CCD cameras for Pi of the Sky project*, Proc. SPIE, vol. 6347, November 2006, pp.6347.0N.1-9
17. M. Jegier, K. Nawrocki, K. T. Poźniak, M. Sokołowski: *Image acquisition in the Pi of the Sky project*, Proc.SPIE, vol.6347, November 2006, pp.6347.0O.1-5
18. M. Kwiatkowski, G. Kasprowicz, K. T. Poźniak, R. S. Romaniuk, G. Wrochna: *Advanced camera image data acquisition system for Pi of the Sky*, Proc. SPIE, WILGA 2008 conference on Photonics for Astronomy and High Energy Physics Experiments, 29 May – 01 June 2008
19. Video controller ISI [isi96.com/PHY-products/video-cores]
20. Video transmission standard ITU-R.BT601 [category.alldatasheet.com/ITU-R.BT601]
21. AMBA standard [javvin.com/networkingterms/AMBA]
22. Memory chip AT45DB642 [atmel.com]
23. Interface API V4L2 [v4l2spec.bytesex.org]

24. Web page Semiconductor Technology Associates [sta-inc.net]
25. Web page INDI transmission protocol [indi.sourceforge.net]
26. Web page of Clear Sky Institute [clearskyinstitute.com/]

E.T.Q.

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 275–285

Thermal-vacuum qualification test for space instrument

PIOTR SITEK

Institute of Electronic Systems, Warsaw University of Technology,
Nowowiejska 15/19, 00-665 Warszawa
Max-Planck Institute for Solar System Research
Max-Planck Str. 2, 37191 Katlenburg-Lindau

Received 2008.07.15
Authorized 2008.08.25

The paper presents preparation, proceeding and results of Thermal-Vacuum test of test the SIR-2[1] instrument for ISRO moon mission Chandrayaan-1. The TV test is the most test for space electronics, during temperature cycle thermal transfer and all electronic are checked. Temperature level, gradient and time of test is related to the target of mission and launch area. For optical measurement, to verify operability of instrument during part of test electroluminescence diodes, mounted in front of spectrometer, were used. After passing all tests instrument was delivered to India and mounted on panel of Chandrayaan-1 satellite. The TV test was performed at the Max-Planck Institute for Solar System Research (MPS) in Katlenburg-Lindau.

Keywords: thermal vacuum test, thermal cycles, thermal troubles

1. INTRODUCTION

The goal of Chandrayaan-1 mission [1] is survey the lunar surface to produce a complete map of its chemical characteristics and 3-dimensional topography. Mission duration is estimated on over 2 years, finally satellite will achieve lunar orbit of ~100 km.

SIR-2 is a redesigned, monolithic grating, near-infrared spectrometer which will fly on the Chandrayaan-1 mission, scheduled for launch in September 2008. With SIR-2, for the first time, an orbiting reflectance spectrometer will be flown in a scientific mission around the Moon with the intention to map the Moon in a two-year long orbit. The instrument is based on SIR, flown on ESA's SMART-1 technology mission,

---

[1] InfraRed Spectrometer

covering a wavelength range from 0.93 to 2.40 μm. SIR-2 has a spectral resolution of $\Delta\lambda$Pixel = 6 nm and an angular resolution of 2.2 millirad. SIR-2 is intended to accomplish a highly integrated study of the lunar mineralogy, together with two further optical instruments from a lunar polar circular orbit, extending thereby the wavelength region of CHANDRAYAAN-1's Hyper Spectral Imager into the near-infrared range, allowing for full coverage of the spectral range between 0.6 to 2.4 μm, which is important for the mineralogical identification.

The near-infrared spectrometer SIR-2 consists of three individual units:

- O-Box (front-end optics), located on an extension of the Anti-SS panel
- Sensor-Head/Radiator Unit (spectrometer), located on the Anti-SS panel
- E-Box (digital electronics and power converter), located inside the S/C cube on the Anti-SS panel.

For each of units independent thermal simulations were done. According to simulation [2], temperatures range for each of units was calculated, depended of mission phase and unit position. During TV test qualification levels for each unit must be reached, instrument can not be exposed to temperatures beyond the acceptance temperature range. At first each of units was tested independently, during test the following temperature was kept:

- for O-Box the acceptance temperature range to -65°C +35°C.
- for SHRU the satellite panel acceptance temperature range -5°C to +50°C and the Cooling Console acceptance temperature range -75oC to +50°C
- for E-Box the operation acceptance temperature range -5°C to +50°C and non-operation acceptance temperature range -30°C to +50°C.

2. PREPARATION TO TV TEST – DOCUMENTATION

Before starting operation in TV chamber we must be well prepared to tests. The first main part is preparing cycling procedure, test procedure and testing sequences for electronics.

The cycling procedure [3], [4], [5] should consist of:

- Ground stage. Warming up and heating instrument in maximum non-operation temperature.
- Hot start (for electronic/electrical systems). One of the most important tests, we are setting temperature to the maximum operation temperature and then we are switching on instrument. Instrument should be able to communicate with S/C, should be able to check all Housekeeping parameters, for most of experiments scientific data checking is not necessary.
- Lift off phase. This phase is simulating thermal conditions from launch of satellite until locating on orbit. The most important is temperature gradient visible by instrument.
- Dawn-Dusk phase. This phase is simulating transport of instrument in space from Earth orbit to destination object. During this phase instrument is exposed to the

lowest non-operation temperature. There should be simulated working of survival heaters and if the temperature is sufficient cold start for electronic/electrical systems.

- Cold start (of electronic/electrical systems). Setting instrument in the lowest operation temperature, waiting for temperature stabilization, turning on instrument and checking all parameters. During this test scientific data should be checked. After this test instrument should be continuously supplied and collect data.
- Thermal cycles. This phase is simulating main phase of flight, thermal environment during scientific measurement. The most important for electronic/electrical systems is checking scientific and Housekeeping data. There can be simulated sensitivity, noise, resolution and mechanical changes dependent on temperature changes. After last cycle electrical systems should be switched off.
- Warming up phase. The phase not occurred during flight. After all thermal tests, instrument temperature should be set to ambient temperature. The most important is setting right gradient of temperature.

Independent procedures should be implemented for each of instrument subunits. As an example TV-Test procedure for Electronic Box (PFM) for SIR-2 instrument.

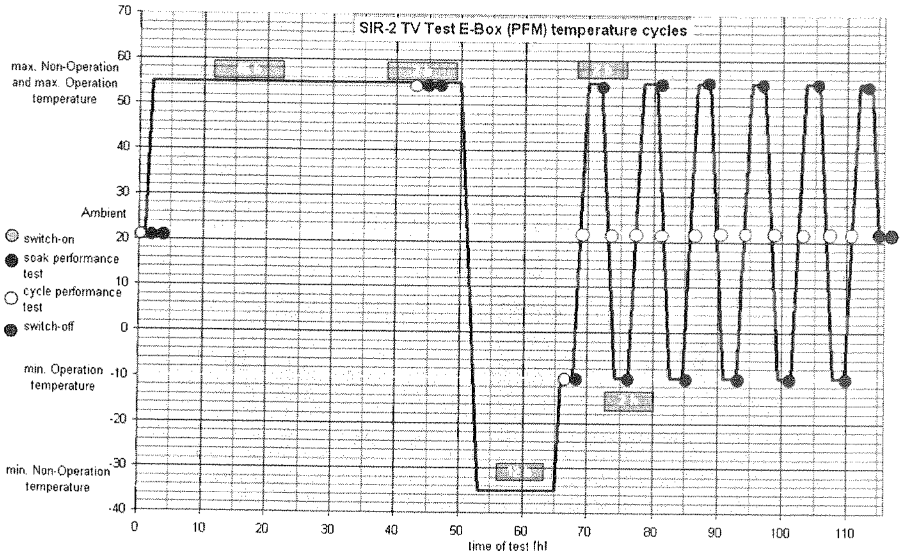


Fig. 1. Thermal-Vacuum cycles procedure for SIR-2 E-Box (PFM)

Before starting the warming up of instrument to maximum non-operation temperature, right level of vacuum should be obtained.

Testing procedure and testing sequence concern mainly electronic/electrical systems should contain following information:

- What kind of test should be proceeded (different parameters of transmission, different parameters of collecting data etc.).

- Which test should be done in which moment of cycle (minimum operation temperature, maximum operation temperature, rising slop, falling slope)

In figure 1 you can find four different tests, which were proceeded during testing E-Box of SIR-2:

- Switch on. At the beginning of test in ambient for checking connections through chamber, as a hot and cold start.
- Switch off. After checking connections thorough chamber, a hot start and finishing all cycles.
- Soak performance test. Short tests during each different frequencies, exposure times and sampling delays were tested. Proceeded at the end of each soak time.
- Cycle performance test. Long test with nominal settings. Proceed on slopes to monitor behavior of instrument during cycles.

For all thermal cycles always TRP[2] is defined, before starting test it is important to define all thermal points on instrument which should be monitored, to check right transmission of thermal energy. All external thermistors used during TV-test of E-Box [7] are shown on figure 2. The most critical points are:

- thermistor 2, the hottest point of PSU[3], close to transformer of DC/DC converter
- thermistor 5, the hottest point of ICU[4], close to ADC converter
- thermistor 7, the coldest point of ICU
- thermistor 1, point located close to FPGA

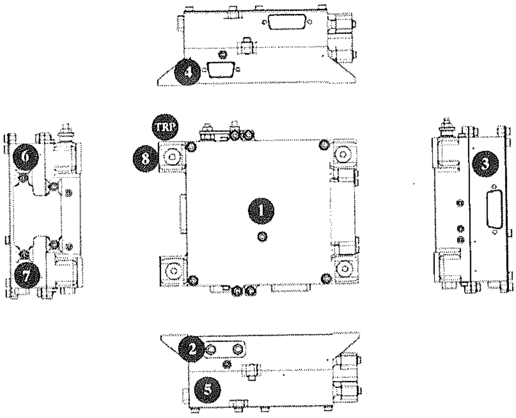


Fig. 2. Location on external thermistors during TV-test of E-Box for SIR-2 instrument

---

[2] Thermal Reference Point

[3] Power Supply Unit, lower part of E-Box

[4] Instrument Control Unit, upper part of E-Box

F. No.	Temperature sensor location	Notes
1	On top of the E-Box	used
2	On PSU box close to ground-stud	used
3	On ICU box close to connector J3	used
4	On PSU box close to connector J2	used
5	On ICU box (above ground-stud)	used
6	On ICU box close to dog-house	used
8	Attached to interface plate next to E-Box reference foot	used
10	Base plate (not shown in Figure 3)	used
11	Adapter plate for SIR-2 sensor-head board (not shown in Figure 3)	used
12	Close to SIR-2 sensor-head board (not shown in Figure 3)	used
TRP	Internal thermistor of the test facility that controls the temperature of the base plate.	used

3. PREPARATION TO TV TEST – HARDWARE

To execute tests we need testing environment and test object. The testing environment consist of TV-chamber, monitoring thermistors, connections between tested instrument and controlling instrument.

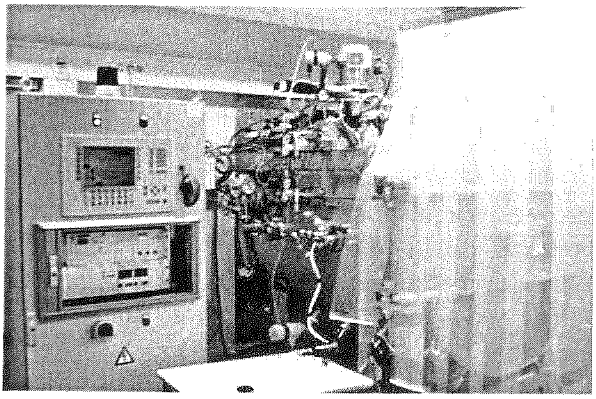


Fig. 3. One of two main TV-chambers used for TV tests in MPS MPG

It is important to prepare special set of cables which will connect our instrument through chamber connectors, to external power supply or monitoring systems. All cables and connectors inside chamber should be prepared for operating in vacuum and all should be tested before starting tests.

Table 1

Main parameters of TV-chambers used in MPS MPG

Max. test volume of the chamber:	600×600×800mm	
Typ. Vacuum:	min. 10E-5 mbar	
Temp. Range:	Shroud:	+150°C to -190°C
	Plate:	+150°C to -190°C
Temp. Gradient:	max. 2°C/minute	

To secure good thermal connection between instrument and chambers table special mounting plates should be prepared. Sometimes it is necessary to use additional material with highly thermally conductive between metal plates. For tests instruments with different thermal subunits MLI[5] should be prepared.

4. TV-TEST RESULTS

As a result of tests [6], [7], [8], we should obtain following data:

- thermal plots of instrument, with visible switch on and switch off time,
- thermal plots with gradients on cover of instrument, dependent of power dissipation in/on instrument
- monitored parameters plots showing right behavior of all parameters (voltages, currents etc.)

In figure 4 you can find thermal results of TV tests of E-Box (PFM) for SIR-2 instrument [6]. Please compare with figure 1 Thermal-Vacuum cycles procedure for SIR-2 E-Box (PFM) [4], "Ground stage" was simulated independently. For first 6 hours there were made simulations of warming up phase, hot start and thermal transfer proceeded. Between 6 and 18 hour minimum non-operation temperature was set. About 20 hour cold start was checked, there was maximum 4 degree increase of temperatures observed. Thermal cycles are shown between 24 and 78 hour. Additionally location and list of thermistors is shown.

As an electrical results of TV tests plot of FPGA core current is presented (one of 21 monitored parameters). Figure 5 presents current fluctuations related to changes of electronic temperature. There was voltage fluctuation observed during TV test, it should be interpreted as changes measuring system parameters (there is no temperature compensation circuit for housekeeping system in tested instrument). Measuring error

---

[5] Multi Layer Isolation

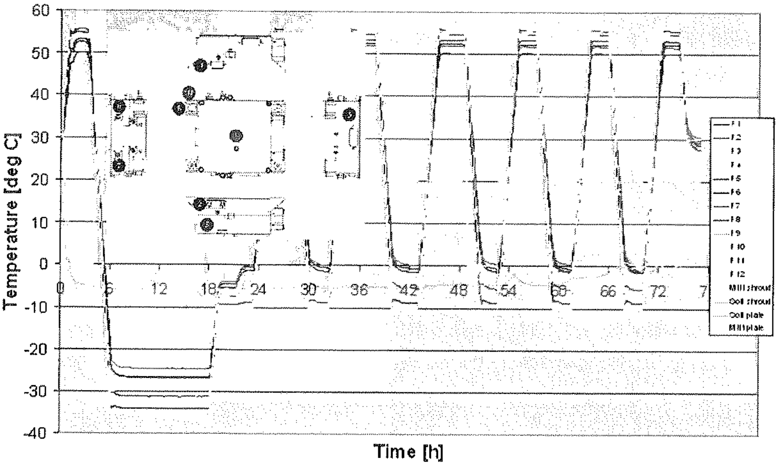


Fig. 4. TV test thermal results of E-Box (PFM) of SIR-2 instrument

related to fluctuation of housekeeping system parameters in acceptance range of temperatures is less than 0.4%. The plot shows current fluctuation during TV-test. Changes of power consumption are related to changes in acceptance range of temperature. Maximal change of power consumption during TV-test is 15%.

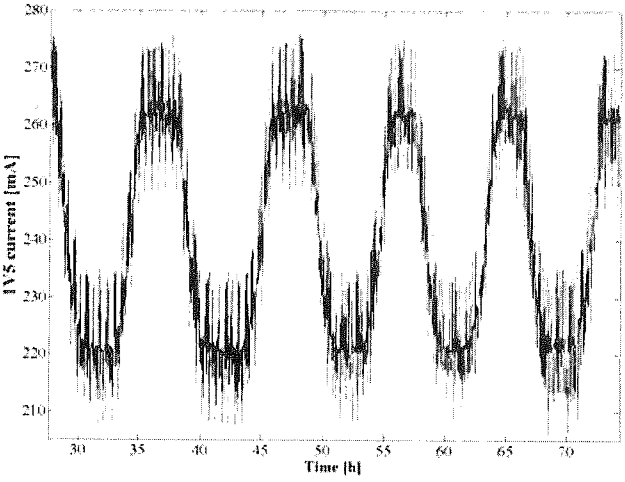


Fig. 5. Thermal fluctuation of measured parameters

During the TV test, a full instrument optical test of spectrometer was done. As a reference spectrum, light from three different electroluminescence diodes was used (wavelength of each LED marked on plot).

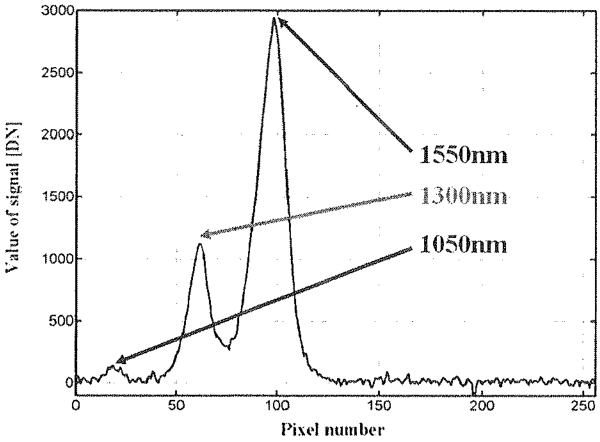


Fig. 6. Optical results from TV test

4.1. DETECTABLE PROBLEMS

Each of tests is made for checking instrument, belov list of problems which we can detect:

1. During pumping down we can observe drops of pressure in chamber, it means that probably there can be problem with outgassing. Wrong elements used during constructing instrument (some plastic elements), errors during coating.
2. Thermal problems in maximum operation temperature.
   a. problem with thermal energy transfer (related to wrong thermal connection, mistakes during thermal calculations)
   b. problem with communication (related mainly to overheating and temperature changes of dynamic parameters of electronics)
   c. problem with monitored parameters (related to overheating of elements)
   d. problem with power supply (overheating of DC/DC elements)
3. Thermal problems in minimum operation temperature.
   a. problem with "cold start" (elements parameters outside expected range)
   b. problem with monitored parameters (elements parameters outside expected range)
4. Thermal problems related to temperature gradient.
   a. problem with heat transfer
   b. mechanical problems with connection (differences between spill over of elements and PCB)
5. Electro compatibility problems.
   a. ground line problem (instruments and chambers GND)
   b. EM disturbance from chamber
   c. long cables problem

This is only main set of problems which can be detected, any of these problems need independent analyzes for finding solution.

4.2. PROBLEMS DETECTED ON TESTED INSTRUMENT

During TV test of E-Box (PFM) [7] we detected three independent kinds of problems, all were solved:

- instrument was reseted by high power disturbance inside/outside chamber (switching chamber heaters)
- instrument was randomly reset in temperatures close to maximum operation level
- at temperatures close to maximum operation level, Housekeeping data were unstable

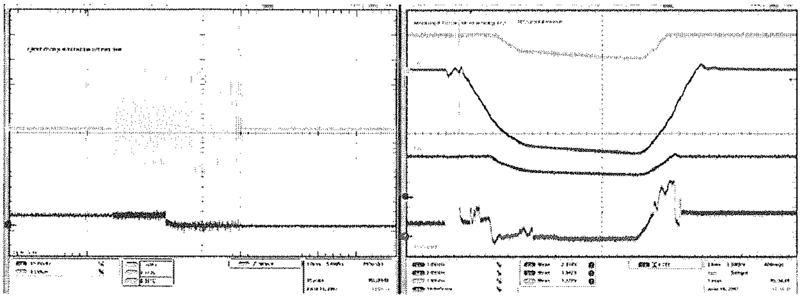


Fig. 7. The reset related to external disturbance (left), rests at high temperature (right)

The reason of resets related to high power disturbance inside and outside chamber was high sensitivity of reset line going through chamber. As a solution during tests filtering capacitor on reset line was added. Problem was seen only on testing harness, flight shielded harness secured of external disturbance. In figure 5 (left) upper line shows disturbance on reset line, lower line shows effect of reset, drop of communication signal.

Resets in high temperature were observed only in vacuum. The problem was non stability of DC/DC feedback line related to higher power consumption in high temperature. Problem was hard to solve, because instrument was not stable only close to maximum operation temperature in a vacuum, everything was stable in low temperatures and in air. As a solution increase of feedback loop capacitance was done.

Non stability of Houskeeping data was related to mounting error. Capacitor in reference voltage circuit of ADC converter for HK data had wrong polarity. As a solution new element with correct polarity was mounted.

5. PIECE OF ADVICE

- Before closing chamber, check all connection through chamber.
- Before closing chamber, check all thermistors.

- Before closing chamber, make basic test of instrument.
- Before closing chamber, check that “red tag items” were removed.
- Before closing chamber, make photo documentation of all connections.
- Before starting TV test, be sure that you have enough of nitrogen for cooling instrument during all cycles.
- Before starting TV test, be sure that you have enough of “dry air” for filling chamber after tests.
- Before starting TV test, check vacuum level.
- Prepare précise schedule of TV tests, with names of people responsible for each part of tests (there should be constantly someone checking main parameters of instrument).

  Note all names of data files with time, date and the most important chamber parameters value.

6. PERSPECTIVES

SIR was the first step in designing compact spectrometers. Successful launch and megabytes of collected date confirm idea of original design. Months of work and data analyses give us answer what and how can be improved to make instrument more flexible and data from instrument easier to analyzing.

SIR-2 confirm availability of our team to build compact spectrometers for space missions. Additionally new design help us understand and eliminate problems which appeared during mission of SMART-1. As all tests confirm from instrument all previous and detected problems were eliminated and data quality was improved. Now team is waiting for launch and first spectra of moon recorded by SIR-2.

Up to this moment we designed 1D spectrometers (1 dimension detector was used). As next step our team decided to design 2D infrared spectrometer. The new spectrometer will have better spectra and time resolution, will have memory for saving series of spectra, using more powerful FPGA gives possibility for picture analyzing on board.

7. CONCLUSIONS

Our specially designed TV test procedures of all subunits of SIR-2 instrument give us possibility to detect problems not-detectable in laboratory. TV test helped to check behavior of all electrical, optical and mechanical parameters of SIR-2 instrument which should be tested. During final test full spectrometer was located and tested in TV chamber. The new SIR-2 design eliminated problems which were detected during flight of SMART-1 mission. SIR-2 has additional space craft heaters which in case of necesity can be used for stabilization of temperature on instrument. As we predicted for calculated temperature range stabilization instrument temperature is not needed but can be helpful when simulation will not fit real environment. As a result of TV tests we obtained confirmation that instrument is able to work properly in simulated space environment. For all subunits qualification temperature levels have been obtained [9].

Each of final results was expected by team. Verg was difference between power consumption of space and commercial FPGA. In air and ambient temperature space logic needed almost 60% more power comparing to commercial logic. According influence of vacuum and temperature we observed differences on the level 85%, comparing to commercial FPGA in normal environment.

Work on SIR-2 instrument was finished. After finishing TV test instrument was successfully calibrated. In June 2008 SIR-2 Team delivered and mounted instrument on spacecraft panel in Bangalore in India, where satellite in assembled.

8. ACKNOWLEDGEMENTS

We are greatly indebted to K. Broenstad, S. Gadomski , K. Heerlein, W. Nowosielski, P. Orleanski, A. Solberg, K. Ullaland, U.Mall, A.Nathues, H.Perplies, A.Dannenberg, K.Heerlein and R.Orlik who were essential in the design and development phase of the SIR-2 electronics.

9. REFERENCES

1. http://www.isro.org/chandrayaan/htmls/home.htm
2. Chandrayaan-1/SIR-2. Thermal Design and Thermal Analysis. Instrument Thermal Mathematical Model. Ver 05, 03.10.2005.
3. A. Dannenberg: C1-SIR-TP-3012 "TV-Test Procedure of the SIR-2 STM Sensor-Head/Radiator Unit"
4. SIR-2 Team, C1-SIR-TP-3016 "TV-Test Procedure of the SIR-2 (P)FM E-Box".
5. SIR-2 Team, C1-SIR-TP-3020 "Qualification TV-Test Procedure of the SIR-2 (P)FM O-Box". item[6.] P. Sitek: C1-SIR-3018 V1.0 "Sensor Head Radiator Unit (SHRU) Thermal-Vacuum Test Report".
7. P. Sitek, K. Heerlein, C1-SIR-TR-3016 "Thermal Vacuum Test Report E-Box PFM".
8. P. Sitek: C1-SIR-TR-3032 V1.0 "FS Sensor Head Radiator Unit (FS-SHRU) Thermal-Vacuum Test Report".
9. SIR-2 Team, C1-SIR-ICD-3001 V2.2 "Interface Control and Instrument Description Document".

JERZY HE

sen
Th
Fo
rea
the

Ke

The
quantitie
external
or strain

Usi
is possib
ted mult
with slig
and spat
instrume
form of
ships fo

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 287–296

Fiber Bragg grating multiplexing system

JERZY HELSZTYŃSKI, LECH LEWANDOWSKI, WIESŁAW JASIEWICZ, KAZIMIERZ JĘDRZEJEWSKI

Institute of Electronic Systems
Warsaw University of Technology
15/19 Nowowiejska Street, 00-665 Warsaw, Poland
e-mail: llew@ise.pw.edu.pl

Received 2008.06.12
Authorized 2008.08.30

Fiber Bragg grating multiplexing system is presented. A set of many Bragg gratings sensors with very smooth side slopes were realized. They are tuned for different wavelengths. They form the distributed measurement system which can be applied for strain control. For signal detection special spectrophotometer with InGaAs linear photodiode array was realized. Fiber-optic switches were applied for flexible multiplexing of many sensors into the system.

Keywords: fiber Bragg grating, optical fiber sensor, fiber Bragg grating multiplexing system

1. INTRODUCTION

The fiber Bragg gratings are commonly used as sensing elements of many physical quantities. The geometrical and material properties variations under the influence of external fields make possible to construct fiber-optic sensors, e.g. temperature, pressure or strain. The information is directly contained in wavelength value of the grating.

Using easy obtainable fiber telecom devices multiplexing of measurement systems is possible. Each single sensor should be distinguished and recognized. Quasi-distributed multiplexed Bragg grating systems are reasonably easy to realize. Many gratings with slightly different parameters can be produced in one fiber. The measured parameter and spatial positioning of the grating can be simultaneously determined using only one instrument, e.g. optical spectrum analyzer. The multiplexed Bragg grating systems in form of "smart structures" can be integrally implemented into constructions, planes or ships for current control of their actual state: temperature or strain.

In the presented paper several switchable fibers can be used. In each fiber (measurement channel) identical set of Bragg gratings with different Bragg wavelengths has been written.

The specially designed spectrum analyzer is used for Bragg wavelength variations measurement and identification of individual sensors in each channel.

2. FIBER BRAGG GRATINGS

2.1. INTRODUCTION

The periodical refractive index profile in the fiber core is written by the interference fringes pattern formed by the UV laser irradiation. It is periodical in shape and forms a diffraction grating for the light guided by the fiber core. Because of refractive index differences along the core the propagated beam is partially reflected according to the Fresnel rules. The counter-propagating wave is build inside the Bragg grating. If the grating period is Λ the Bragg wavelength λ_B equals:

$$\lambda_B = 2n_{eff}\Lambda \tag{1}$$

where: n_{eff} – the effective refractive index profile. These reflections are constructively added forming so called Bragg reflection. Then the properly realized fiber Bragg grating forms very deep stop-band optical filter. Other wavelengths propagate with no influence and lossless through the grating. This behavior is shown schematically in Fig.1.

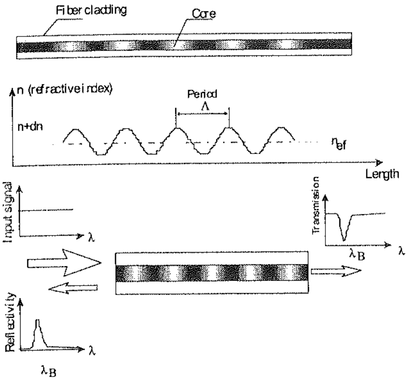


Fig. 1. The idea of Bragg grating operation

The modulation depth of the refraction index profile and grating length are decisive parameters of the transmission properties of the grating. The increase of length grating involves increase of reflectivity and decrease of FWHM. The full analysis is possible using the coupled modes theory [1].

The easy way to change the grating spectral characteristics during grating formation is the implementation of the interferometric method.

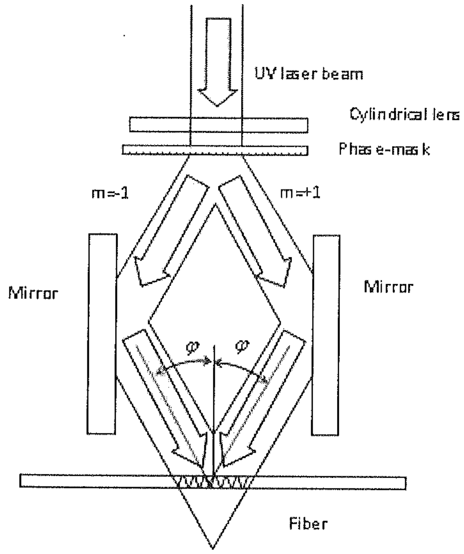


Fig. 2. The interferometr with spatial beam division

The fundamental advantage of interferometers with spatial beam division is the possibility of changing distances between fringes and therefore period of the grating. It implies the variation of Bragg wavelength λ_B. If the interferometer mirrors are mounted on the rotational and linear stages the angles of incidence of each beam against the fiber φ can be easily changed (Fig.3.). The formulas (2) and (3) concern this case:

$$\Lambda = \frac{\lambda_{uv}}{\sin\varphi} \tag{2}$$

$$\lambda_B = \frac{2n_{eff}\lambda_{uv}}{\sin\varphi} \tag{3}$$

where: λ_{uv} is the laser wavelength.
The relative Bragg wavelength change is given by:

$$\frac{\Delta\lambda}{\lambda_B} = -\Delta\varphi ctg\varphi \tag{4}$$

It is easy to calculate that the Bragg wavelength variation can be considerably high. For typical value of this angle in our case of 13^0 (it depends on the phase mask used for the beam splitting) the change of ±5° denote the Bragg wavelength changes in range 1000 – 1550nm. Such big sensitivity is advantageous from the point of view of theoretically high range of demanded device wavelength but is also very susceptible to any external influences of the environment such as temperature and vibrations caused by long time laser exposure during Bragg grating formation in the fiber core.

If the modulation of the refractive index profile n_{eff} is small, the approximate relation between the relative width of the spectral characteristic $\Delta\lambda_0/\lambda$, length of the grating L and a period Λ can be given in the form:

$$\frac{\Delta\lambda_0}{\lambda} \rightarrow \frac{\lambda_0}{n_{eff}L} = \frac{2}{N} \tag{5}$$

where: $N=L/\Lambda$

Severe requirements for applied optical equipment and stable environment conditions are necessary to realize Bragg gratings using amplitude division interferometers.

2.2. LABORATORY SETUP

Special Bragg gratings should be realized according to the method applied for presented multiplexed sensors system. We have assumed to make set of gratings with the Bragg wavelength differences of 9 nm and with the slope width of 1.5 nm. The amplitude division interferometer was realized. Its practical configuration is given in Fig.2. The optical system was mounted on the Coherent optical table (with pneumatic stability and leveling system). Special earth and building isolated concrete foundation cushion was also added. The environment was stabilized in form of air purification and conditioning system. The optical set-up is given schematically in Fig.3 and its photograph in Fig.4.

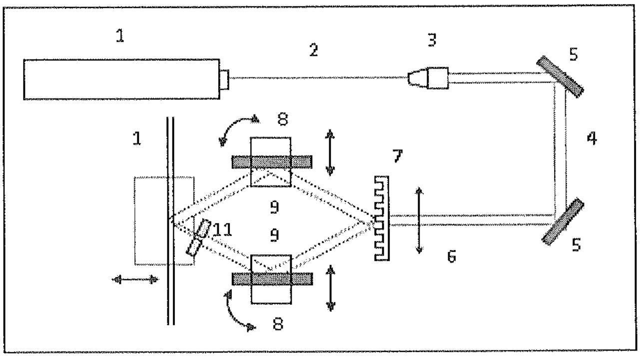


Fig. 3. Experiment setup for fabrication of fiber Bragg gratings; 1-frequency doubled argon ion laser (244 nm), 2-laser output beam, 3-beam expander, 4-collimated beam, 5-mirrors, 6-cylindrical lens, 7-phase mask, 8-inerferometer mirrors, 9-two-dimensional translation stages with rotators, 10-fiber, 11-shield with 0.3 mm gap

The double-frequency 100 mW, λ=244 nm CW Coherent argon-ion laser was used in the experiments. The laser beam diameter was increased to 8mm using 10×

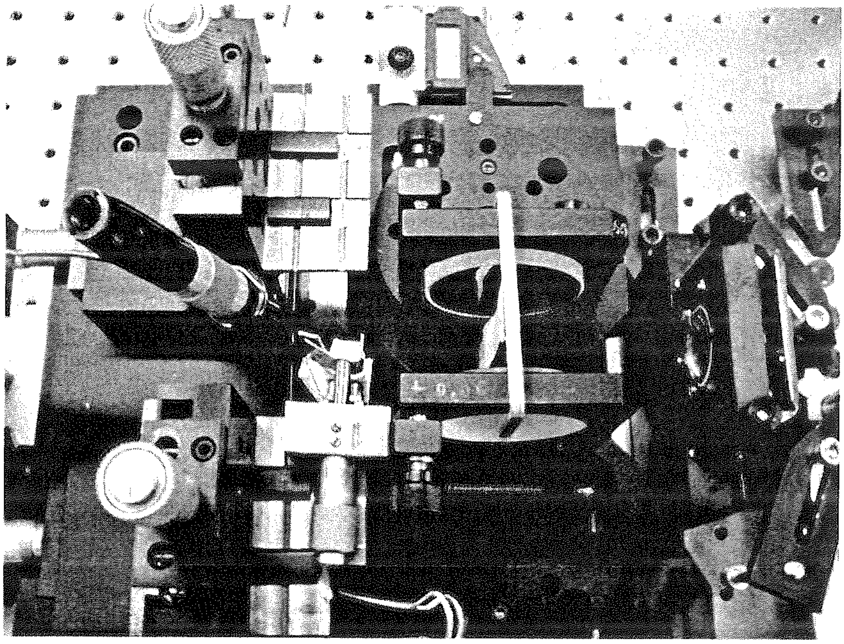


Fig. 4. Interferometer Setup

CVI beam collimator. The phase mask was applied for the amplitude beam splitting. Consequently the system symmetry was easily achieved. It was necessary for realization of series of identical in shape but with different Bragg wavelength gratings. The cylindrical lens applied in front of the phase mask increases the laser beam density. In consequence the grating writing time was significantly shortened.

To realize the Bragg grating of required parameters (according to (5)) it was necessary to reduce to 0.3 mm the width in one of the beams by slit diaphragm and consequently the grating length decrease.

The examples of spectral characteristics of Bragg gratings specially designed for presented multiplication system are given in Fig.5.

Grating slope can be modified significantly, from the fraction of nanometer to a few nanometers by changing the parameters of the interfering beams and the exposition time. Applying very narrow one of the beams we produced very broad gratings. But for significant refractive index profile changes the exposition time should be very long.

2.3. SENSOR SYSTEM

The information of the strain value is directly given by the Bragg wavelength shift. The fastest measurement can be performed with the use of optical spectrum analyzer. We can easily measure transmission or reflection characteristic of each grating. It is very expensive solution and difficult for application in real field conditions. Much cheaper

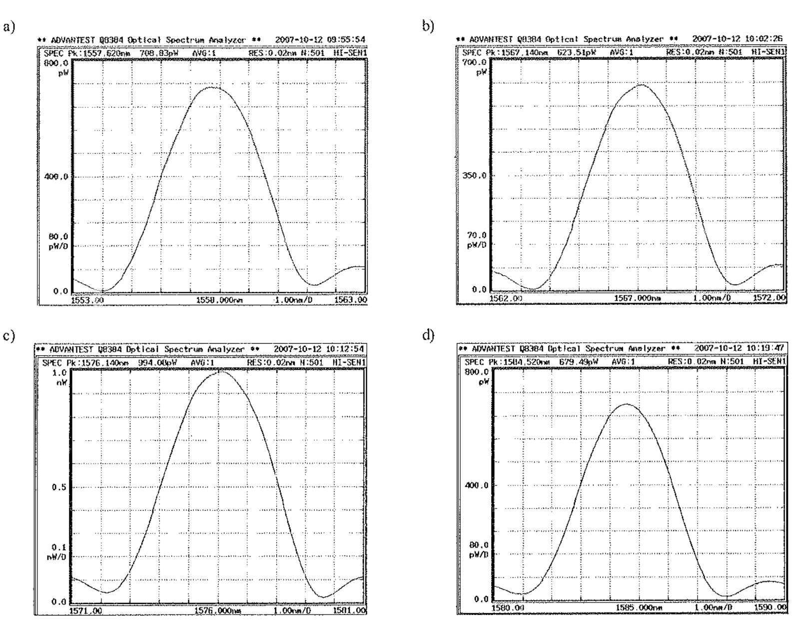


Fig. 5. Spectral characteristics of realized Bragg gratings for a) 1557,620 nm, b)1567,140 nm, c) 1576,140 nm, d) 1584,520 nm; exposition time for all gratings: 120 mins

solution is realized in presented paper. To analyze the spectrum changes we work on the grating slope and register the signal from individual elements of the photodiode array attached to the output of specially designed miniature spectrophotometer. In the case of zero strain the signal is taken from the lower point of spectral reflection characteristic. This point responds also to the spectral positioning of the photodiode elements (Fig.6). Application of such solution is possible if precise characteristics of Bragg gratings can be achieved in terms of Bragg wavelength and length of the slope. The spectral characteristics presented previously confirm the ability to produce precisely needed shapes. In Fig.7 variation of the output current versus stretch of the fiber is presented.

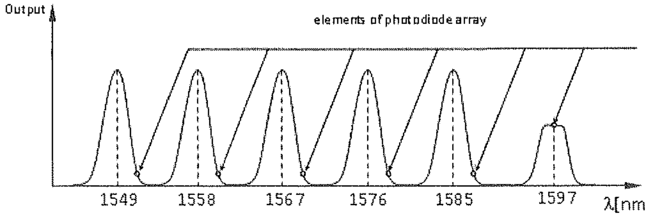


Fig. 6. Wavelength characteristics of Bragg gratings array for single sensors channel. Circles mean photodiode elements

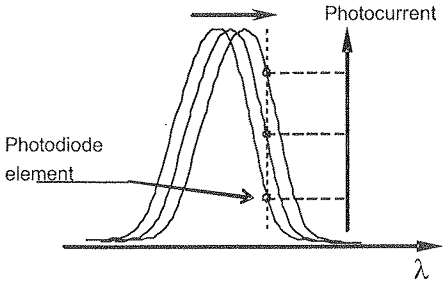


Fig. 7. The variation of the output current of individual diode under stretching conditions

The results of the experiment confirm the assumed idea of the system. The strain measurements were realized (stretch and compress). They were based on the signal level variation using specially designed miniature spectrophotometer assembled together with the photodiode array. The measurement range is wider if the slope of the Bragg grating characteristic is longer.

3. FUNCTIONAL OPTOELECTRONIC SETUP OF THE SENSOR SYSTEM

The designed Bragg grating multiplexed system is based on single-mode fibers working in the third window within the range 1550-1600nm. Set of six Bragg gratings of Bragg wavelengths equal to 1549 nm, 1558 nm, 1567 nm, 1576 nm, 1585 nm and 1597 nm was written in each fiber. The last grating is designed for reference signal. The distance between gratings can range from a few centimeters to many meters. Each grating forms local strain sensor of the object. They can be permanently bonded to the structure by the epoxy resin.

The functional optoelectronic diagram of the Bragg grating multiplexed sensor system is schematically presented in Fig.8.

The optical signal from the broadband optical source (1) is given to the input of each individual fiber (4) which is assembled of sets of sensors (Bragg gratings) through the fiberoptic switch (2) and optical circulator (3). The output circulator signal (Bragg wavelength reflected signal from individual sensors) is directed trough the second switch (2) onto the input of optical spectrophotometer (5) with the photodiode array (6). The photocurrent signal of the photodiodes is amplified by the input amplifiers and send to the data acquisition system (7) controlled by the computer (laptop).The whole sensor system: control, data acquisition and data processing is realized in the LabVIEW environment. The application of fiber switches allow to multiplex the number of fiber sensing arms. The number of channels depends on applied fiber switch. The 4, 8 or 16 channel switches are available on the market. In our case 4 channel switches were used and only two channels were put to use with 6 grating in each (5 sensing gratings and one reference grating).

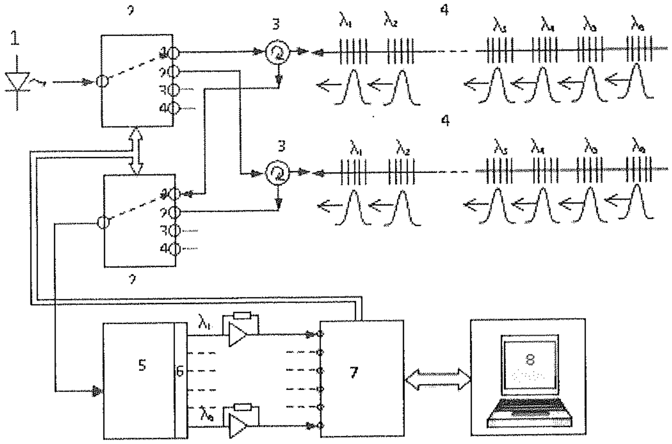


Fig. 8. The idea of the multiplexed system; 1-optical signal source (superluminescent diode), 2-fiber-optic switch, 3-optical circulator, 4-optical fiber with set of Bragg gratings, 5-miniature spectrophotometer, 6-photodetector array, 7-data acquisition system with input amplifiers, 8- computer

The specially designed optical detection for this particular application is constructed as a spectrophotometer with concave diffraction grating with 600 lines/mm and 100 mm radius (Jobin Yvon). Their spectral parameters were chosen for 1500-1600 nm operation and proper spatial sensitivity to cooperate with the physical dimensions of individual array segments. The photodiode array is installed at the spectrophotometer output. The spectral sensitivity of the spectrophotometer was estimated as 12 nm/mm. Photodiode module (Hamamatsu G8909-01) consist of 40 detection segments (photodiodes) 80 μm diameter each and spaced 250 μm. All elements are made on single InGaAs substrate. The total dimensions are: 22x2x2,5 mm. Only every fourth segment was used for detection of signals from consecutive sensing gratings. It can be calculated that for spectral sensitivity of the analyzing grating 12 nm/mm the distance between gratings is equal to 9 nm. The application of concave diffraction grating considerably simplified the optomechanical construction. No additional lenses or mirrors are necessary for effective coupling of fiber output signal with the receiver. The photograph of constructed spectrophotometer is given in Fig.9 and the calibration characteristic as a function of strain in Fig.10.

The scale constant was obtained experimentally. It is in total agreement with the theory. For germanium doped silica fiber in the range 1500-1600 nm it is 1.2 pm/$\mu\varepsilon$ [2].

The acquisition and signal processing module allows also to register and analyze static and dynamic (vibration) strains. In the vibrometer option the frequency can be as high as several hundred Hz in multichannel operation, for single sensor operation several kHz.

In Fig.11 view of the acquisition module screen is presented. The metal rod was periodically strained with the frequency 22 Hz and amplitude approx. 800 $\mu\varepsilon$

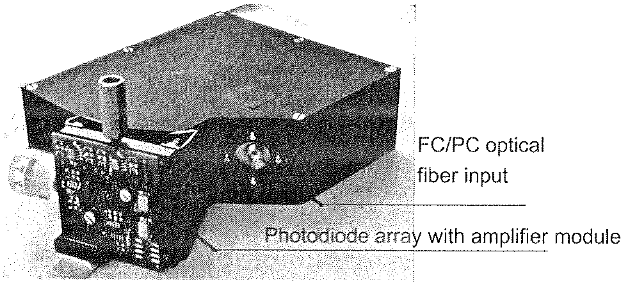


Fig. 9. The photograph of realized micro-spectrophotometer

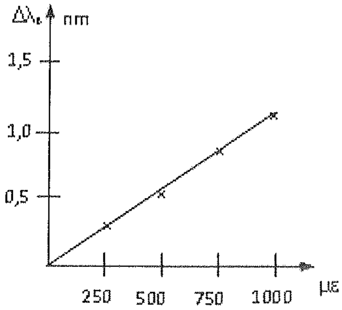


Fig. 10. Calibration characteristic of Bragg grating sensor

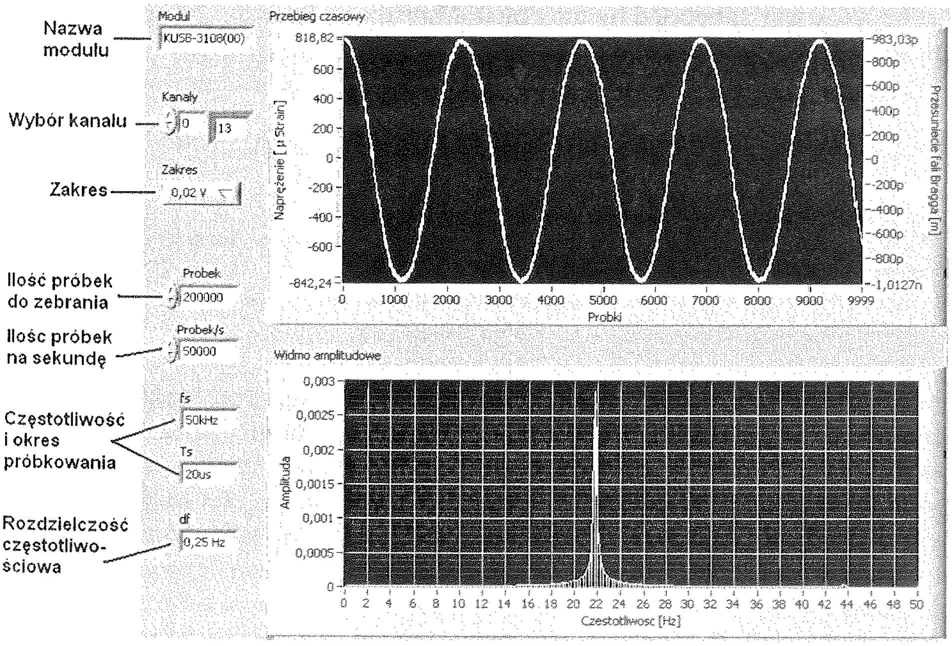

Modul: KUSB-3108(00)
Kanały: 0 13
Zakres: 0,02 V
Probek: 200000
Probek/s: 50000
fs: 50kHz
Ts: 20us
df: 0,25 Hz

Fig. 11. Front panel of the vibrometer – module screen

4. PARAMETERS

In the realized system the descending slope width of the grating is wider than 1.5 nm. The nonlinearity is smaller than 2%. It means that the total measurement range in stretching or compressing condition is 1250 $\mu\varepsilon$. The resolution is approx. 25 $\mu\varepsilon$. The total number of sensors is 24 which include 20 sensing and 4 reference gratings. In case of 16 arms fiber switch the number of gratins can increase to 96. The gratings can be physically distanced from couple of centimeters to several hundreds of meters. Each grating is individually calibrated and can be applied for static or dynamic measurements.

5. SUMMARY

The multiplexed fiberoptic Bragg grating system was presented. The flexible arrangement allows to apply tens of individual sensing elements using only one detecting system. The measured value and spatial positioning of the sensor can be simultaneously determined. These systems can be integrally implemented into e.g. constructions or planes for current control of their actual state.

6. ACKNOWLEGMENTS

The work was supported by the Polish Ministry of Science and Higher Education grant no. 1316/T10/2005/28.

7. REFERENCES

1. R. Kashyap: *Fiber Bragg gratings*, Academic Press, 1999.
2. A. Othonos, K. Kalli: *Fiber Bragg Gratings Fundamentals and Applications in Telecommunications and Sensing*, Artech House, 1999.

Th

syst
fibe
syst
real
data

Key

Fibe
Such a b
could be
growth o
of innova
years lar
sensing s
vibration
make the

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 297–309

The LabVIEW application for fiber Bragg grating sensor system management and data processing

DARIUSZ PODSIADŁY, LECH LEWANDOWSKI

Institute of Electronic Systems Warsaw University of Technology
15/19 Nowowiejska Street, 00-665 Warsaw, Poland
D.J.Podsiadly@elka.pw.edu.pl, llew@ise.pw.edu.pl

Received 2008.06.16
Authorized 2008.08.02

This paper describes design and realization of the fiber sensor system. The metrology system was developed for construction mechanical stress control. The system is based on fiber Bragg gratings as a physical quantity sensors. The major emphasis is placed on sensor system management which is running under the LabView environment. Specific tasks are realized by the programme such as acquiring, processing and visualizing of measurement data from the multiplexed sensor system.

Keywords: fiber Bragg grating, optical fiber sensor, strain sensor, LabVIEW

1. INTRODUCTION

Fiber optic telecommunication has made an unprecedented growth in recent years. Such a big progress is caused by constantly increasing data throughput needs, which could be satisfied by high data carrying capacity of an optical fiber. The dynamic growth of the optical telecommunication market was accompanied by the development of innovative technologies for optical sensing and fiber optic devices. During many years large numbers of fiber optic sensors were developed. They are usually used for sensing some physical quantities like temperature, strain, pressure, acceleration and vibration. Unique properties and large number of advantages of fiber optic sensors make them ideal candidates for use in sensing technologies.

2. FIBER BRAGG GRATINGS

Fiber optic sensors are very often based on Bragg gratings. The fiber Bragg grating (FBG) consist of refractive index modulation, created on the core of single mode fiber. When single mode fiber with inscribed Bragg grating is coupled with broadband light source of constant spectral density, some wavelengths in spectrum of transmitted light will be canceled out. It means that light guided along the core of optical fiber is scattered by each grating plane and if Bragg condition is satisfied scattered beams add constructively in the backward direction, which gives reflection. The spectrum of reflected light corresponds to spectrum of transmitted light – Fig.1. The center wavelength of reflected beam is defined by grating parameters and it is given by the formula: $\lambda_B = 2n_{eff}\Lambda$, where Λ is a period of the grating and n_{eff} is effective refractive index of the fiber core [1].

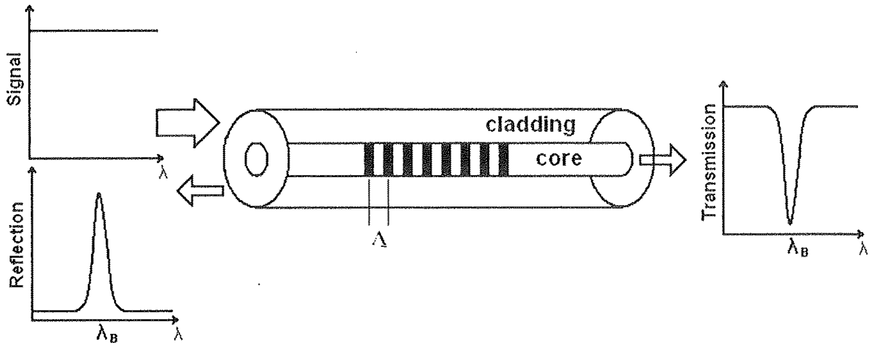


Fig. 1. Properties of fiber Bragg gratings

According to the Bragg condition the center wavelength of back reflected light depends on effective index of refraction of the core and the periodicity of the grating. The effective index of refraction of the core, as well as periodic spacing between the grating planes will be affected by changes in strain and temperature. Wavelength shift will be given by the formula:

$$\Delta\lambda = 2\left(\Lambda\frac{\partial n_{eff}}{\partial l} + n_{eff}\frac{\partial\Lambda}{\partial l}\right)\Delta l + 2\left(\Lambda\frac{\partial n_{eff}}{\partial T} + n_{eff}\frac{\partial\Lambda}{\partial T}\right)\Delta T \tag{1}$$

The first term in the equation (1) represents the strain effect on optical fibers which corresponds to a change in the grating spacing and strain-optic induced change in refractive index. The second term represents the effect of temperature on optical fiber.

3. THE IDEA OF FIBER BRAGG GRATING SENSOR SYSTEM

Information about strain applied to the grating is included in Bragg wavelength changes. The specially designed spectrometer has been built in order to measure each sensor Bragg wavelength variations. The diffracted light from spectrometer is directed onto the linear photodiode array next the output electrical signal is amplified, digitized and processed by computer.

An advantage of optical fiber sensors based on Bragg gratings is flexibility in distributed sensors network implementation. Due to very low insertion loss of the fiber, sensors can be serially multiplexed along a single mode optical fiber. The discussed fiber Bragg sensor system is built in this manner. Single sensors set consists of long fiber with six inscribed gratings placed in different regions of the fiber. Every grating has a different center Bragg wavelength which provides strain signals distinguishability. The signal from each grating sensor is detected by a single photodiode segment of linear photodiode array. In detection only decreasing slope of spectral reflectivity characteristic of single Bragg grating is used. Strain perturbations induces changes of single photodiode segment illumination corresponding to the applied tension, therefore photocurrent is increasing or decreasing according to Bragg wavelength change – Fig.2.

Output signals are dependent on input optical power inserted into fiber sensor. Measurement result must be immune to the level of source power, connector losses and non zero attenuation of every optical device used in the system. To achieve independence of fluctuating light levels reference grating must be used. The reference grating is the Bragg grating with flat top and its shielded from the stretching force applied to other sensors. It's used only to measure the light source output power. Output voltage level from each fiber sensor is divided by the value measured on reference grating. The division of these two quantities gives the relative values and solves the problem of unpredictable fluctuations of output signals, caused by source instability and losses.

4. THE FIBER BRAGG GRATING SENSOR SYSTEM REALIZATION

4.1. SENSOR SYSTEM CONSTRUCTION

The presented fiber Bragg grating sensor system is dedicated to smart structure technology for mechanical strain control. It consists of broadband light source which is coupled with 1x4 fiber optic switch. The switch connects optical channels by redirecting an incoming optical signal into selected output channel. Output signal is transmitted into optical circulator and next to selected fiber sensor set. [2], [3]. The light reflected from the gratings contains strain information and it is guided through circulator and optical switch to the spectrometer with linear photodiode array. Both optical switches are controlled by the same driver in parallel. Weak output signals from InGaAs detectors are amplified by transimpedance amplifiers block and digitized by

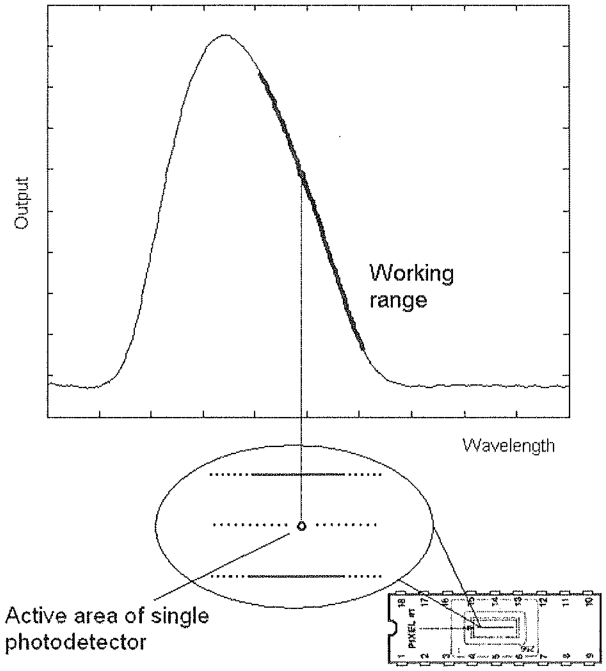


Fig. 2. Spectral characteristic of the Bragg grating sensor with marked working curve

KUSB-3108 data acquisition board. The last component in the system is the computer with LabVIEW environment and sensor system controlling application. The primary task of the computer is to read data from the system and to control external devices. Diagram of the system with all its components is presented below (Fig.3.)

4.2. MEASUREMENT ALGORITHM

Due to differences of spectral characteristic of every Bragg grating sensor scaling procedure must be performed. The aim of this procedure is to determine the slope of each grating characteristic (Fig.4).

The curve segment $U' - U''$ must be approximated with a linear function. To compute the slope the coordinates of U' and U'' points must be known. The voltage value on first point is measured by a single readout U' then the spectral characteristic is shifted in wavelength scale by $\Delta\lambda_{1nm}=1\text{nm}$ and the second measurement is performed U''. Every measured signal is divided by voltage value U''' from reference grating what gives relative output values.

Knowing the relative voltage difference: $\frac{\Delta U}{U'''} = \frac{U''}{U'''} - \frac{U'}{U'''}$ we are able to compute the slope:

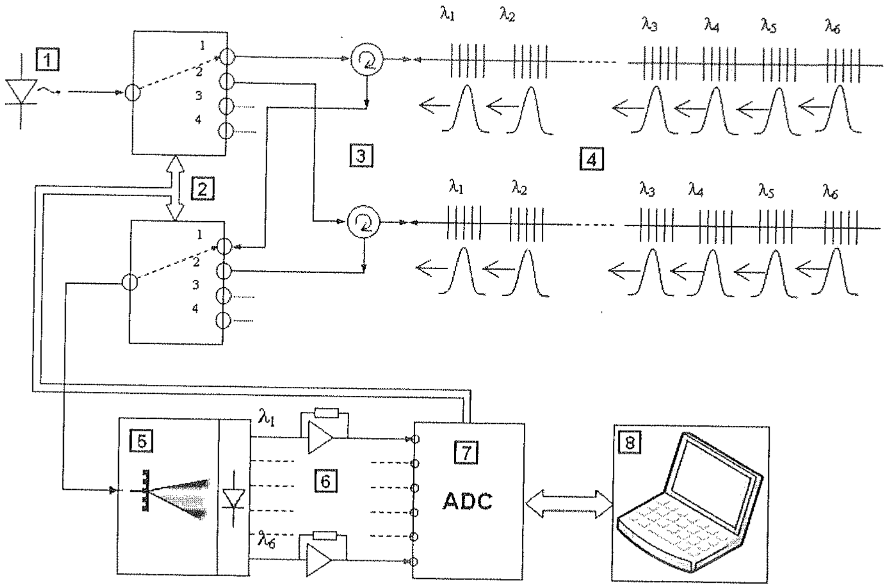


Fig. 3. The fiber Bragg grating sensor system diagram. Components: (1) broadband light source, (2) optical switches, (3) circulators, (4) set of fiber Bragg grating sensors, (5) spectrometer with linear photodiode array, (6) amplifiers block, (7) data acquisition board, (8) computer with LabVIEW environment

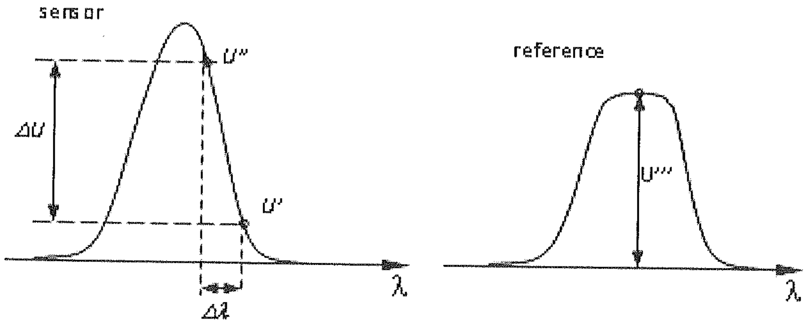


Fig. 4. Reflectivity of sensor grating and reference grating

$$S = \frac{\frac{\Delta U}{U'''}}{\Delta\lambda_{1nm}}\left[\frac{1}{nm}\right] \tag{2}$$

The scaling procedure described above has to be done only once and with no strain applied to the fiber sensor. This procedure is realized outside the LabView application and slope values are inserted into programme code.

Before measurement the calibrating procedure ought to done. The aim of calibration is to determine a relative voltage of static operating point $\frac{U'}{U'''_{op}}$ of sensor grating (U'''_{op} is the voltage from reference grating at the moment of calibration). Values of operating points are saved and stored in the disk file. When strain is applied to the sensor relative value $\frac{U''}{U'''}$ of the signal is measured. The Bragg wavelength shift can be computed with the formula:

$$\Delta\lambda = \frac{\Delta U_{rel}}{S} \tag{3}$$

where: $\Delta U_{rel} = \frac{U'}{U'''} - \frac{U'}{U'''_{op}}$

Knowing that strain sensitivity at the range1500-1600 nm for silica fibers grating is 1.2pm/1$\mu\varepsilon$ [1], the strain value can be computed using the formula:

$$x\,[\mu\varepsilon] = \frac{\Delta\lambda\,[nm]}{1.2}10^3 \tag{4}$$

5. APPLICATION DESCRIPTION

Above mentioned fiber Bragg grating sensor system diagram with description of measurement algorithm let to determine the operational requirements for the application.

The major task of the created programme is to control whole system, acquire and process measurement data from multiplexed sensor system, then present the results.

Functional requirements of the programme:

- Application has to determine the value of Bragg wavelength change and the value of strain which is causing wavelength change.
- Application has to manage the KUSB-3108 data acquisition board (DAQ) and optical switch module.
- Ability to perform calibration procedure on user's demand.
- Possibility of saving measurement results in spreadsheet files.

5.1. THE STRUCTURE OF THE APPLICATION

Computer application was divided into five main blocks which realize different types of functions (Fig.5).

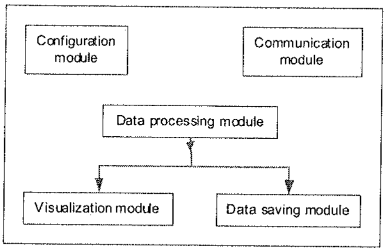


Fig. 5. Application structure

The application configuration module is responsible for performing all procedures which are needed to initialize variables and scaling factors. These variables are used by application to fulfill measurement algorithm. The crucial issue in calibration process is the measurement of signal level from each fiber sensor where no force is applied to the fiber. Electrical signals from photodetectors, including the signal from reference grating, are digitized by acquisition board in parallel. Readouts are saved to a file in application directory. So called static operating points are loaded to the programme when it is launched then are used for wavelength shift computations. Execution of calibration procedure is necessary condition to make all computations. Running application without configuration file will cause limitations in programme functionality and user will be prompted to perform calibration procedure.

The second important block in application structure is communication module which is used to control peripheral devices. External devices such as optical switch and data acquisition board are controlled by user through different interfaces. The switch board has embedded driver module and connection between computer and driver is realized by RS232 serial port. An application module implements custom controlling procedures using command set provided by Agiltron. The code was based on LabView built in Virtual Instruments and functions that communicate with devices connected to a serial port. Keithley Instrument's KUSB-3108 acquisition board is more complex device. It uses USB interface. Manufacturer supplies drivers and LabView libraries. The library is set of Virtual Instruments and it enables an access to data acquisition boards. Application communication module is initializing DAQ board, determines which channels will be used, sets channel types (single ended or differential), sets type of the conversion (in this case analog to digital conversion), executes autoranging procedure to set the most adequate input values range and input gain to minimize measuring error, starts data acquisition and gets the results to the LabView application. To provide the best measurement results, the configuration of KUSB-3108 module isn't fixed; it depends on application working mode which is actually chosen and on the level of input signal.

Data processing module is a part of application block diagram which performs mathematical operations on signal samples. Type of arithmetical operations and its sequence depends on measurement algorithm realization. When application is working in continuous measurement mode fast Fourier transformation on samples is applied. Additional signal processing procedure helps in interpretation of measurement data especially when sensors are put on periodical strain. The results of computations are sent to the visualization module. Final results are displayed on numeric indicators and plotted on graphs. More detailed description of visualization module will be placed in the next section where graphics user interface will be presented.

Storing measurements on hard disk drive was one of the application functionality requirements. This ability is granted by last application module – data saving module. This part of application collects all values which were displayed on indicators and build result array. When measurement is finished programme is waiting for user's action. After switching save button application converts results array into spreadsheet string, opens or creates the result file and writes data into file. In single measure mode new results are appended to existing file. Every saved signal value has its own time stamp. In continuous mode new results are saved in separate file due to large number of samples. Data files have *.xls extension thus can be opened using Microsoft Excel for example.

5.2. APPLICATION BLOCK DIAGRAM

The programme diagram presented on Fig. 6 is simplified. It focuses only on the most important procedures realized by application. Real operation network is more complex and is a little bit different to each working mode.

5.3. USER INTERFACE DESCRIPTION

Fiber sensor system controlling application is divided into two main parts. The system operator can distinguish two different application working modes. The first is configuration mode where system is set up to perform measurements, and the latter is measurement mode which can be further divided into single and continuous measurement mode. Switching between different application modes is realized by adequate tab selection.

Front panel of application in configuration mode is presented on Fig. 7. Using the controls placed on the panel user is allowed to perform scaling procedure. One of the most important buttons is fiber Bragg grating sensor set selector. Changing the value displayed on drop-down list causes channel change in optical switch device, the name of used serial port is indicated next to the button. Apart from that, selected value induces operations on configuration files dedicated to chosen sensor set. Spectral characteristic

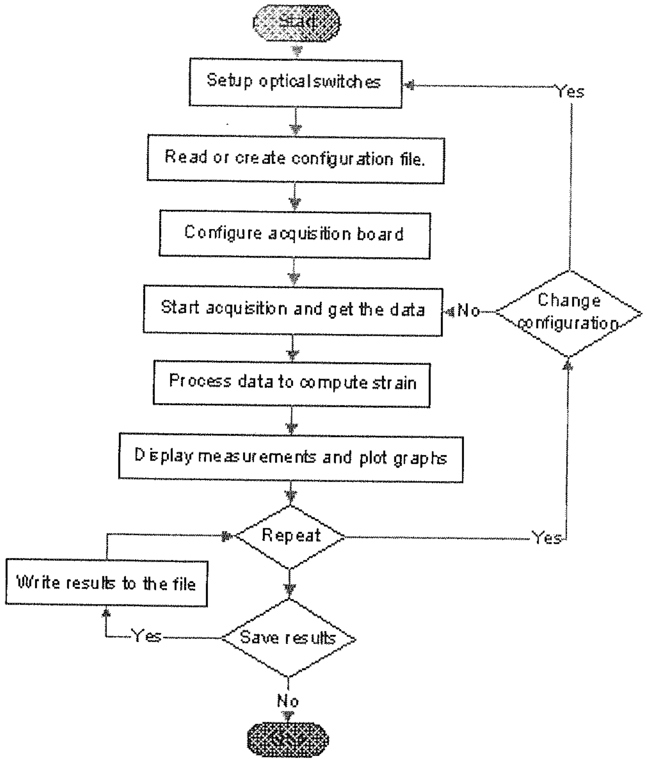
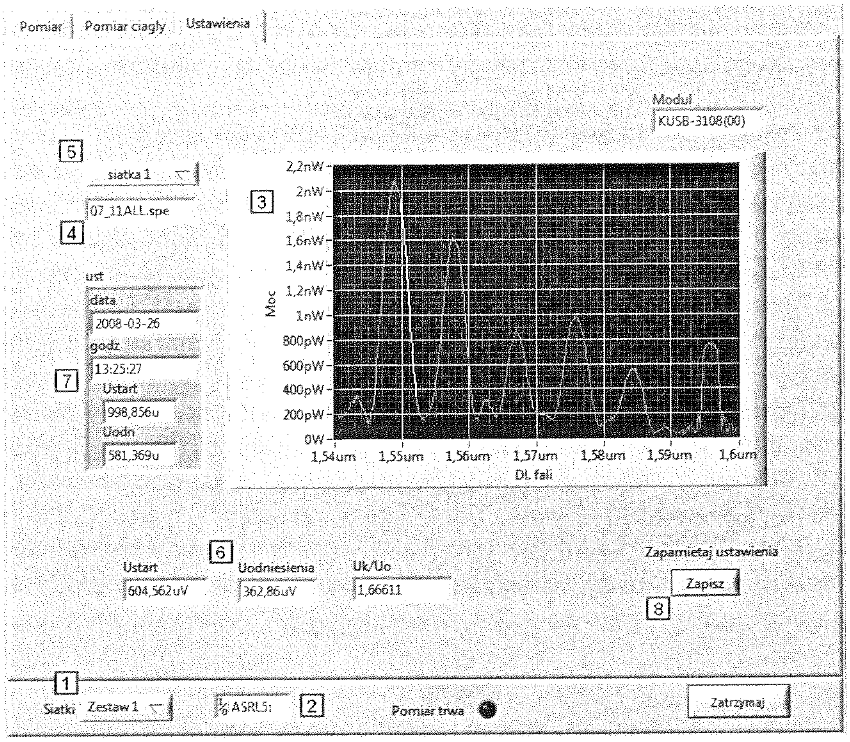


Fig. 6. Simplified application block diagram

of fiber Bragg grating sensor set is presented on the graph. Every sensor set was analyzed by optical spectrum analyzer on Bragg grating manufacturing stage. Application is reading spectrum analyzer data files and plot data points on characteristic. Name of file is also presented on the panel. Displaying sensors spectral characteristic isn't necessary but it might be useful when system is working with many Bragg grating sensors. The user may select a sensor to configuration from the drop-down list, actually selected Bragg grating is highlighted on spectral characteristic of sensors set. The input channel of data acquisition board corresponding to chosen grating is sampled in every second. The signal level from selected grating and reference grating is displayed on numeric indicators. The third indicator displays relative value of signal level. Previously stored configuration is displayed on the window placed on the left side of front panel. User gets the information about the time and date when calibrating procedure was taken. Displaying saved values of signal gives possibility to compare historical results to actual signal readouts. If selected sensor has not already been configured the window which is displaying saved results is empty. In such case user should press "Save (Zapisz) " button to complete calibration procedure. User must be aware of the fact that only configured sensors can be used in measurements.

Configuration mode.



Modul: KUSB-3108(00)

siatka 1

07_11ALL.spe

ust
data: 2008-03-26
godz: 13:25:27
Ustart: 998,856u
Uodn: 581,369u

Ustart: 604,562uV
Uodniesienia: 362,86uV
Uk/Uo: 1,66611

Siatki: Zestaw 1

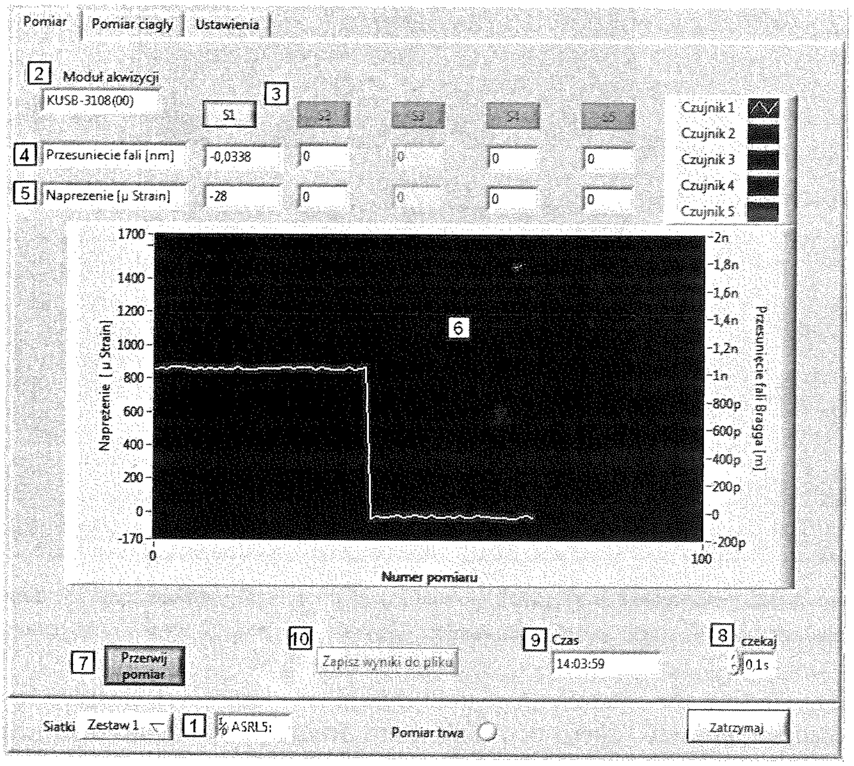
ASRL5:

Fig. 7. Configuration mode front panel. (1) Bragg grating set selector, (2) serial port indicator, (3) spectral characteristic of sensor set, (4) data file name indicator, (5) configured sensor selector, (6) measured signal indicators, (7) configuration file viewer, (8) save button

Single measurement front panel is presented on Fig.8. It consists of controls and indicators which allow user to perform measurements. In this mode user also can choose set of sensors to be monitored using sensor set selector. Changing the value displayed on drop-down list causes channel change in optical switch device, the name of used serial port is indicated next to the button. The name of data acquisition board is displayed on the indicator placed on the top of the front panel. If the name is displayed the communication between the device and the computer has been established without any errors. User is able to choose which strain sensor he wants to use in measurements. Each of five sensors marked from S1 to S5 has its own control button. If the button is in "ON" state which is indicated with green colour the sensor will be enabled and used in measurement. Only previously configured sensors can be selected, otherwise the button will be disabled and grayed out. The group of indicators is located below sensor selectors. In the first row values of center Bragg wavelength shift is located. The second row of indicators contains the value of computed strain. Numeric indicators show only last sampled value of strain. The 100 previously measured values are presented on the graph. In the range of elastic deformations of optical fiber the relation between center

Bragg wavelength shift and strain is linear and described by already known formula (4). This allows to display both quantities on the same graph with two different scales. Therefore the strain values axis is placed on the left side of the graph and wavelength shift scale is located on the other side. Measurement data points are plotted only when sensor is selected. If sensor will be disabled during measurement adequate plot will be cleared. The measurement sequence can be started and stopped by pressing the "Rozpocznij/Przerwij pomiar" button. Measurements are performed continuously with desired time intervals. The repetition rate can be set up using proper numeric control. Timestamp of single readout is displayed on the indicator. When measuring sequence is finished user may press save button to write all results into disc file.

Single measurement mode



Moduł akwizycji: KUSB-3108(00)

	S1	S2	S3	S4	S5
Przesuniecie fali [nm]	-0,0338	0	0	0	0
Naprezenie [µ Strain]	-28	0	0	0	0

Czas: 14:03:59

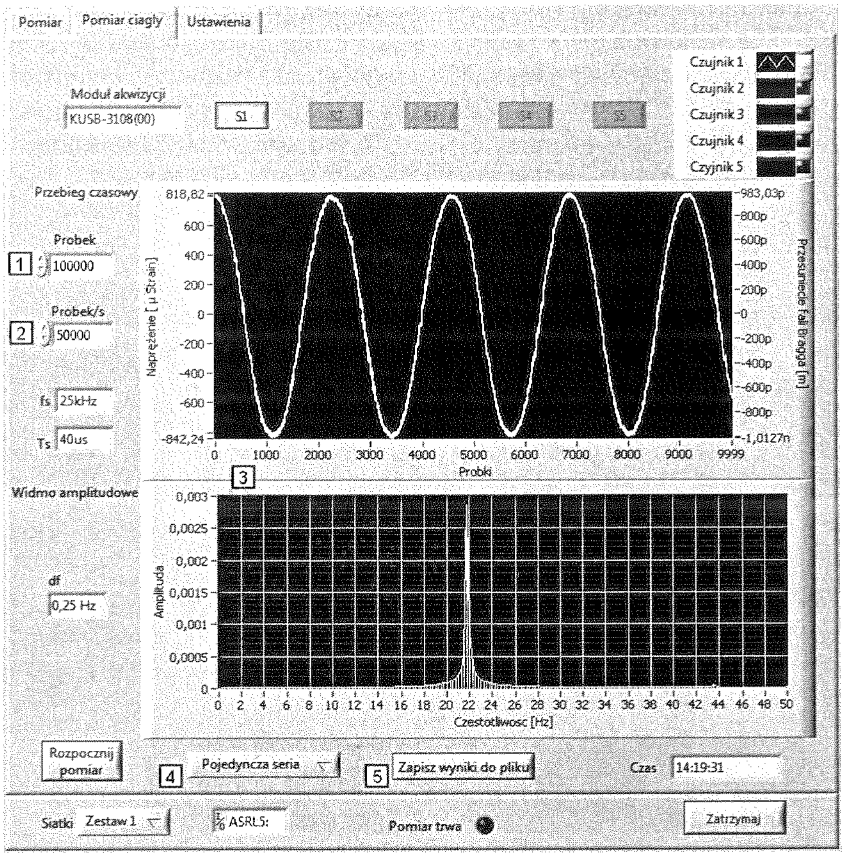
czekaj: 0,1s

Siatki: Zestaw 1 ASRL5:

Fig. 8. Single measurement front panel. (1) Bragg grating set selector with serial port indicator, (2) DAQ name indicator, (3) sensor selector, (4) wavelength shift indicator, (5) strain indicator, (6) measurement graph, (7) start/stop button, (8) time delay, (9) measurement timestamp, (10) save button

The last application functionality is ability to perform measurements in continuous mode. In this mode dynamic strain changes can be monitored and the system works as vibrometer. In this mode sampling period of every input channel can be accurately controlled and the results are buffered. Data acquisition device allows to acquire signals at wide frequency range up to 25kHz.

Continuous measurement mode



Moduł akwizycji: KUSB-3108(00)

Probek: 100000

Probek/s: 50000

fs: 25kHz

Ts: 40us

df: 0,25 Hz

Czas: 14:19:31

Siatki: Zestaw 1

ASRL5:

Fig. 9. Continuous measurement front panel. (1) number of samples, (2) samples rate, (3) measurement result graphs, (4) single sampling/continuous sampling selector, (5) write to file button

Before starting the measurement user should set number of samples to acquire and desired sample rate. These two values are needed to configure data acquisition board by setting up convenient buffer size and sampling frequency. Two indicators are displaying values of sampling period and frequency. Measurement results are presented graphically in time and frequency domain – Fig.9.

In continuous measurement mode application can work in two different manners. In single series submode application is configuring data acquisition board to get only specified number of samples and stop acquisition. In the second submode sampling is continuous. After getting a pack of samples next measurement is triggered. Results are displayed in a real time and acquisition must be stopped by the user. Type of acquisition scheme may be changed by control button. In continuous measurement mode set of recent measurement results can be saved to the spreadsheet file.

6. SUMMARY

This paper presents the project and the realization of fiber Bragg grating sensor system controlling application. Brief description of the whole system and its all components is also included. Created programme fulfills presented applications requirements. Application is able to process the signals from sensor system. Executed tests proved that the application gives correct results.

7. ACKNOWLEGMENTS

The work was supported by the Polish Ministry of Science and Higher Education grant no. 1316/T10/2005/28.

8. REFERENCES

1. A. Othonos, K. Kali: *Fiber Bragg Gratings*, Artech House 1999.
2. J. Helsztyński, L. Lewandowski: Fiber *Bragg gratings with very smooth side slopes*: Proc. of SPIE Vol. 7120, 71200R, (2008).
3. J. Helsztyński, L. Lewandowski, W. Jasiewicz, K. Jędrzejewski: *Fiber Bragg grating multiplexing system*, Electronics and Telecommunications Quarterly, vol. 54, 2008, no 3.

T
coeff

0%
da
mo
mo
ref

Ke

The
single-m
structure
[3]. Thi
polishin
low exc
In F
and tape
that the
the devi
through
tapering

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 311-321

The external refractive index dependence on coupling coefficient in the biconically fused single-mode fiber couplers

ŁUKASZ SAWICKI, KAZIMIERZ JĘDRZEJEWSKI

Institute of Electronic Systems
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warszawa, Poland
e-mail: kpj@ise.pw.edu.pl

Received 2008.06.12
Authorized 2008.08.27

The coupling coefficient of fused biconical taper single-mode coupler can vary from 0% to almost 100% by changing the refractive index of the surrounding media. From the data obtained the coupling mechanism cannot be attributed to power transfer between the modal fields guided by the fiber cores in the taper waist region. The results support the model of the coupling mechanism between modes guided by a glass core and an external refractive index.

Keywords: biconical fused tapers, couplers, sensors

1. INTRODUCTION

The technique of single-mode coupler fabrication made by fusing two identical single-mode fibers and consequently drawing to form of a twin-fiber biconical taper structure (BFT) has been proved as an effective method for coupler production [1] [2] [3]. This technique is relatively simple in comparison to mechanical [4] or chemical polishing [5] techniques. Fused biconical single-mode couplers can be made with a low excess loss and arbitrary coupling coefficient [1] [2] [3] [6] [7].

In BFT fabrication technique the fibers are brought into close proximity, heated and tapered [3] [8] [9] [10]. Experimentally it is observed during coupler fabrication that the light power transfer occurs between two fibers in the taper waist region of the device. In the twin identical fiber biconical taper structure the power transmitted through the coupler oscillates back and forth between two fiber output ports during tapering process. The monitoring of this transmitted light allows the number of full cycles in the power transfer to be recorded. The selection of coupling coefficient can be easily chosen. The excess loss of the coupler can be determined from a measurement of the power transmitted through one of the fibers before coupler is fabricated and the total power transmitted through both fiber arms after the coupler is made. The single-mode fiber commonly used in the coupler fabrication is the matched cladding fiber. For typical 1300nm ÷ 1600 nm range the fiber has a field diameter of 9-10 μm, cladding diameter 125 μm, and a cut-off wavelength of about 1200 nm. During coupler fabrication multi-mode propagation in the waist region can occur, because mode field is expanding to the external cladding surface. Signal transfers from weakly guided mode to the strong guided modes and the waist region of the coupler can propagate multi-mode. Therefore very important is to keep propagation in the taper waist region of fundamental mode only. To reduce occurrence of higher-order modes, cone of the taper has to be smooth [11]. It has been achieved by lengthening the taper cone regions.

2. THEORETICAL BACKGROUND

There are some models describing coupler mechanisms: theory of coupled modes [17], theory of evanescent wave coupling between the fields in the fiber cores at the taper waist region [16] and theory of the symmetric and antisymmetric modal fields interaction [13]. The coupler waist has been modeled as a composite waveguide with a glass core with the cross-section shape of a dumbbell and air cladding [9]. In this model it was assumed that the cross section of the cores of the fiber used in constructing the composite waveguide have become so small in the tapering process that the light fields extend to and are bound by the air-glass interface. Thus the medium surrounding the fused fibers, which is air during coupler fabrication, becomes the cladding of the coupler. A material with refraction index different from air can have a strong effect on power transfer in the waist region [8].

The phenomenon is described in terms of a superposition of a symmetric modal field $\varphi+$ and antisymmetric modal field $\varphi-$. The fields are bound by the residual cores of the two tapered fibers forming the coupler and the transverse spatial dependence of these modes can be approximated by [12]:

$$\varphi_{-} \approx \varphi_1 - \varphi_2 \qquad \varphi_{+} \approx \varphi_1 + \varphi_2 \tag{1}$$

where: φ_1 and φ_2 – the fundamental modes of the fibers.

In this model the taper waist section of the coupler is considered as two glass fiber cores of radius a separated by a distance ρ and surrounded by external medium. It has been shown that the antisymmetric mode φ_{-} of the coupler can be cut-off [12]. The state when the antisymmetric mode φ_{-} is cut-off makes the coupler significantly worse. Coupling mechanism, in which energy transfer in the composite waveguide is between modes bound by the glass and external media interface, was considered. Antisymmetric stimulation in the input elliptical taper caused to appear antisymmetric mode LP_{11} apart from transmission of the symmetrical fundamental mode LP_{01}. Both of them propagate with different phase constants along the waist of the taper. Total signal can occur in the first or second output port, depending on the waist length. As light propagates through the transition tapers to the composite waveguide structure at the coupler waist, the modal fields gradually change from being bounded by a waveguide with a glass core and glass cladding to one with a larger glass core and an external medium as a cladding. The reverse process of course occurs in the up tapers leading away from the coupler waist. To understand this process the cut-off normalized frequency for the antisymmetric mode should be defined and compared with normalized frequency.

Normalized frequency:

$$V = 2\pi a/\lambda(n_1^2 - n_2^2)^{1/2} \tag{2}$$

where : n_1, n_2 – refraction indices for core and cladding respectively
a – coreradius
λ – wavelength

Cut-off normalized frequency for the antisymmetric mode is [13]:

$$V_C = (8/(1 + 4\ln(\rho/a)))^{1/2} \tag{3}$$

where : a – coreradius
ρ – separation between the fiber cores

This equation applies to fibers with a step-index profile, waveguides with e.g. power-law profiles have larger cut-off frequencies than given above.

When normalized frequencies of the individual fibers forming the coupler do not fall below the cut-off normalized frequency value, formulas (2) and (3), antisymmetric mode is propagated in the composite waveguide. If the cross section of the composite waveguide is presented in form of two tangent circles, as a dumbbell cross section, the results of [12] can be used. For a composite waveguide with a dumbbell cross section the relevant fiber cores are in contact and the ratio $\rho/a = 2$, normalized cut-off frequency for the antisymmetric mode was estimated from (3) as $V_C = 1{,}45$. The exact cut-off frequency depends on the actual shape of the cross section of the composite waveguide. It was not possible to measure it during experiments. Since for a bare coupler the cladding material of the glass composite waveguide core is air, the normalized frequency of the individual glass cores without clad forming the composite waveguide is considerably larger than 1,45. In this way low loss propagation of the antisymmetric mode in the coupler is possible and the low loss coupling can be achieved [13]. On the other hand, when normalized frequency is smaller than $V_C = 1{,}45$ antisymmetric mode is not transmitted by the composite waveguide. It takes place only in cases for the refractive index of the immersion liquid in value of 1,45. The other configurations fulfil inequality: $V > V_C$, as shown in the Table 1.

With immersion of the coupler waist in a liquid with refraction index much less than silica the coupler loss should not be expected to change significantly. An increase

Table 1

The normalized frequency of the coupler for different wavelengths and refractive indices of used immersion liquids. The values that do not fulfil the inequality: $V > V_C$ marked bold. The values of refractive indices were calculated for demanded wavelengths – Tables 2 and 3

n_D λ [nm]	1,4	1,41	1,42	1,43	1,44	1,45
1200	5,28	4,80	4,27	3,66	2,91	1,88
1250	5,06	4,59	4,08	3,49	2,77	1,77
1300	4,84	4,40	3,91	3,33	2,63	1,65
1350	4,65	4,22	3,74	3,18	2,50	1,53
1400	4,46	4,05	3,58	3,04	2,38	**1,43**
1450	4,29	3,89	3,44	2,91	2,27	**1,32**
1500	4,13	3,75	3,30	2,79	2,16	**1,22**
1550	3,98	3,60	3,17	2,67	2,05	**1,13**
1600	3,84	3,47	3,05	2,56	1,95	**1,03**

in the refractive index of the liquid results in a corresponding decrease in normalized frequency of the coupler. The coupler loss should be expected to remain constant as the refractive index of the liquid increases until the normalized frequency of these liquid clad glass fiber cores becomes lower than the normalized cut-off frequency of the composite waveguide. Further increases in liquid refractive index ought to results in a rapid increase in coupler loss. Entire power transmitted by the coupler should be independent of surrounding medium for refractive indices of immersion liquid lower than silica.

Model of the coupler, where coupling coefficient is dependent on a degree of superposition of evanescent wave coupling between the fields in the waists fields, was also considered.

3. EXPERIMENTAL MEASUREMENTS

In this paper the dependence of the coupling ratio of fused biconical single-mode couplers on the refractive index of the surrounding medium is experimentally presented.

Fused biconical single-mode coupler are immersed in a immersion liquid (as shown in Fig.1). For light launched into port 1 measurements were made of the output powers P_3 and P_4 from ports 3 and 4.

The coupling coefficient (CR) was calculated, which determines relation of the output power transmitted in one fiber versus the output power transmitted in both:

$$CR_{14} = P_4/(P_3 + P_4)100\% \tag{4}$$

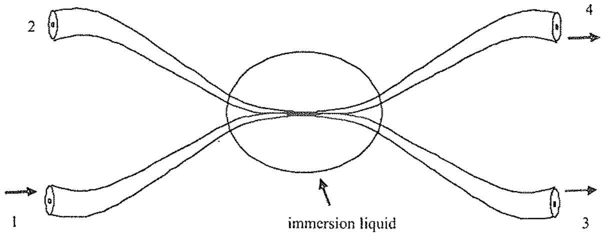


Fig. 1. Schematic diagram of a fused biconical single-mode coupler immersed in a immersion liquid

The excess loss (EL) was also calculated, which is specified as a value of power loss inside the coupler:

$$EL = 10\log(P_3 + P_4)/P_1\,[\mathrm{dB}] \tag{5}$$

A comparison of the total output power $P_3 + P_4$, measured for air and immersion liquids, with the input power P_1, measured for air, allows to plot coupling ratio (CR) and excess loss (EL) characteristics of the coupler versus refractive index of surrounding medium and the wavelength.

The couplers were mounted into aluminum housing to protect from damage. It caused handling problems while immersing into the liquid and while cleaning procedures. The measurement errors came out mainly of above difficulties and technical restrictions.

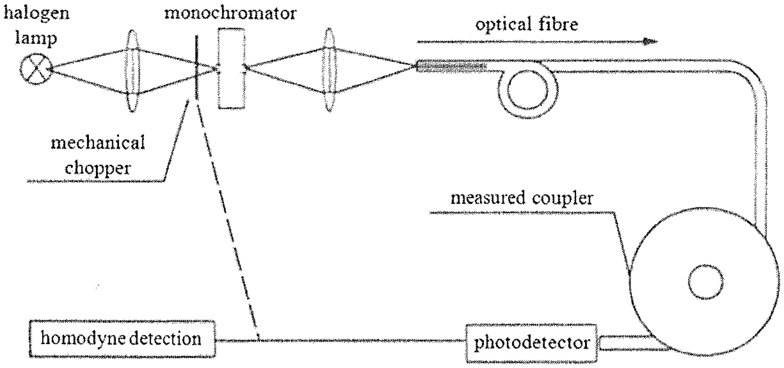


Fig. 2. Schematic diagram of experimental arrangement used to measure the characteristics of the single-mode coupler immersed in immersion liquid

The experimental setup is presented in Fig.2. The halogen lamp light passes through the mechanical chopper and monochromator to choose the appropriate wavelength in the range of 1000nm ÷ 1600nm and is focused by a microscope objective into the input end of fiber arms of the coupler. The light transmitted from the two output fiber arms is measured by photodetectors (InGaAs photodiodes). The homodyne detection system was used.

The total output power in air was chosen as a reference for measuring any increase in coupler excess loss after immersion in the liquids. The coupler is then characterised by measured coupling coefficients CR_{14} and coupler excess losses EL versus refractive indices of immersion liquids for appropriate wavelength. Cargille Labs produces immersion liquids and gives refractive index value n_D for 25°C and for wavelength 589 nm (sodium-D line). To match the results of the measurements to the used wavelength a table with appropriate calculations was made, according to the equation [14]:

$$n(\lambda) = C_A + (C_B/\lambda^2) + (C_C/\lambda^4) \tag{6}$$

where: C_A, C_B, C_C – Cauchy's coefficients. The results for the temperature of 25°C are shown in the Table 2.

Table 2

Refractive indices of immersion liquids for different wavelengths

λ [nm] \ n_D	1,40	1,41	1,42	1,43	1,44	1,45	1,46	tolerance adjustment
1000	1,3922	1,4022	1,4122	1,4222	1,4322	1,4422	1,4520	0,001
1050	1,3918	1,4018	1,4118	1,4218	1,4318	1,4418	1,4517	
1100	1,3914	1,4014	1,4115	1,4215	1,4315	1,4415	1,4513	
1150	1,3911	1,4011	1,4112	1,4212	1,4312	1,4412	1,4510	
1200	1,3909	1,4009	1,4109	1,4209	1,4309	1,4410	1,4508	0,005
1250	1,3906	1,4007	1,4107	1,4207	1,4307	1,4407	1,4505	
1310	1,3904	1,4004	1,4104	1,4205	1,4305	1,4405	1,4503	
1350	1,3902	1,4003	1,4103	1,4203	1,4303	1,4404	1,4502	
1400	1,3901	1,4001	1,4101	1,4202	1,4302	1,4402	1,4500	
1450	1,3899	1,4000	1,4100	1,4200	1,4300	1,4401	1,4499	
1500	1,3898	1,3998	1,4099	1,4199	1,4299	1,4400	1,4497	
1550	1,3897	1,3997	1,4097	1,4198	1,4298	1,4398	1,4496	
1600	1,3896	1,3996	1,4096	1,4197	1,4297	1,4397	1,4495	

Correct interpretation of the results demands the knowledge of refractive index dependence on light wavelength for fused quartz – Table 3.

Table 3

Refractive indices of fused quartz in the temperature 24°C for various wavelengths [15]

λ[nm]	1200	1250	1300	1350	1400	1450	1500	1550	1600
n	1,44811	1,447545	1,44698	1,446412	1,445845	1,445266	1,444687	1,444089	1,443492

4. RESULTS

Measurement results of coupling coefficient (Fig. 3÷6 b and c) and total output power (Fig. 3÷6 a) characteristics for four couplers with different cross-section shape of the taper waist region as example are presented. During coupler fabrication light power transfers between two fibers in the taper waist region of the device. In the twin fiber biconical taper structure the power transmitted through the coupler oscillates back and forth between the two fiber output ports while tapering process. The power passing from one fiber to another and back has been defined as one full cycle. The monitoring of this transmitted light allows to record the number of full cycles in the power transfer and the selection of coupling coefficient. The couplers numbered from A to D vary mainly by cross-section shape of the taper waist region.

- Entire power transmitted by the coupler should be independent of surrounding medium for refractive index of immersion liquid lower than silica [9] (Fig. 3÷6 a).
- The coupler loss should be expected to remain constant as the refractive index of the liquid increase until the normalized frequency of these liquid clad glass fiber cores becomes lower than the normalized cut-off frequency of the composite waveguide [13] (Fig. 3÷6 a).
- Further increases in a liquid refractive index ought to result in an increase in coupler loss [13] (Fig. 3÷6 a).

Table 4

Main parameters of measured couplers

number of coupler		A	B	C	D
coupling coefficient CR_{14} [%] in the air	1310 [nm]	13,2	94,9	54,2	7,2
	1550 [nm]	12,3	88,2	56,3	92,6
coupling coefficient CR_{14} [%] in liquid of refractive index 1,44	1310 [nm]	92,1	60	91,3	86
	1550 [nm]	85,8	12,1	65,4	10,2
excess loss in the air EL [dB]	1310 [nm]	0,52	0,71	0,39	0,46

Diagram in Fig.3a shows the effect of the refractive index of the surrounding media on the total output power transmitted through coupler A. The results indicate that during immersion in liquid the excess loss of the coupler changes little. When refractive index of liquid reached value of 1,44 (after correction according to Tab.2) excess loss of the coupler starts to increase with refractive index of immersion liquid. The Fig.3a shows that for a refractive index of 1,45 loss of the coupler is increase significantly and for value of 1,46 the total output signal is almost fall to zero. The increase in loss for a refractive index of 1,45 or larger is attributed to the fact that light is gradually spread out to the surrounding medium, as the antisymmetric mode φ- is cut-off. The results of the experimental measurements show, that physical mechanism

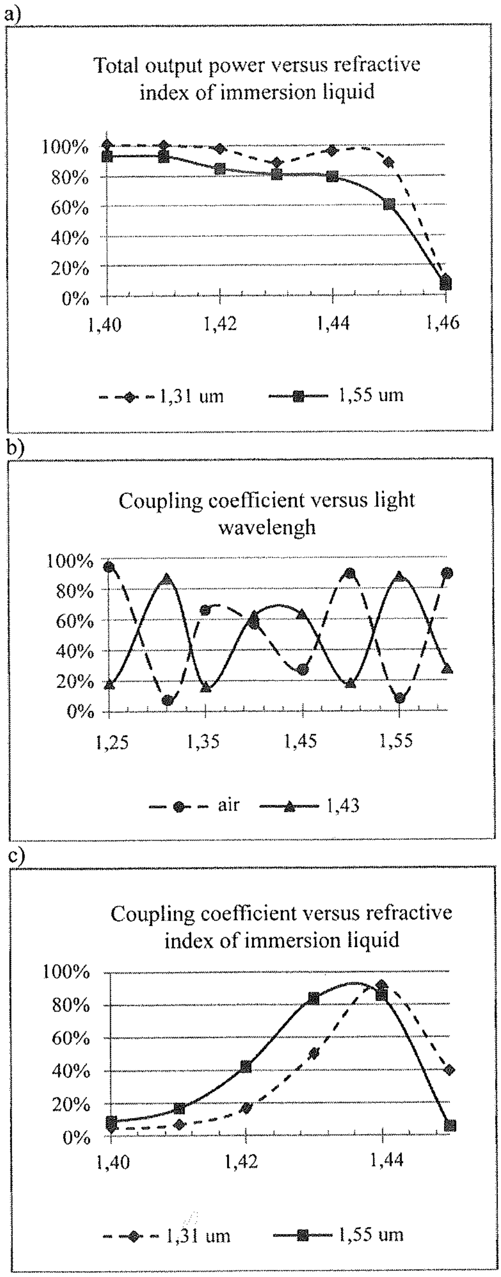


Fig. 3. Coupler A – 2 full cycles of signal transfer

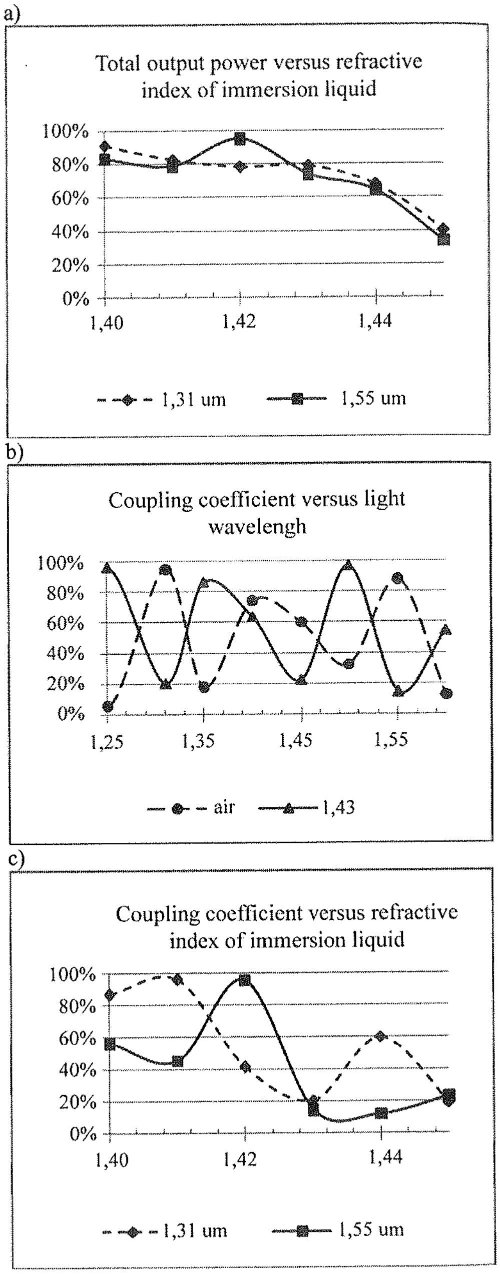


Fig. 4. Coupler B – 3 full cycles of signal transfer

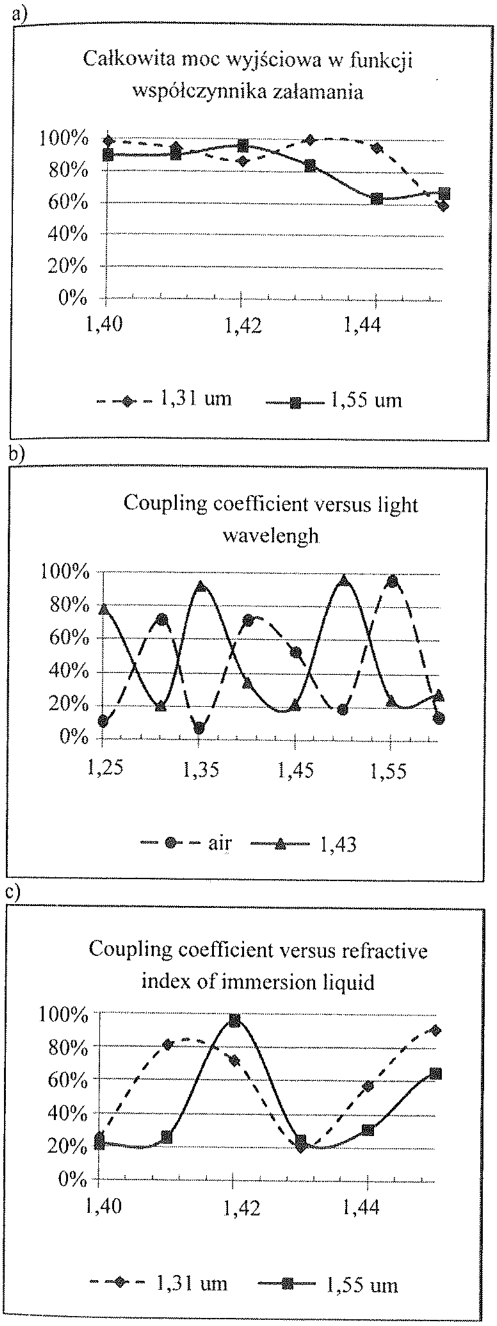


Fig. 5. Coupler C – 5 full cycles of signal transfer

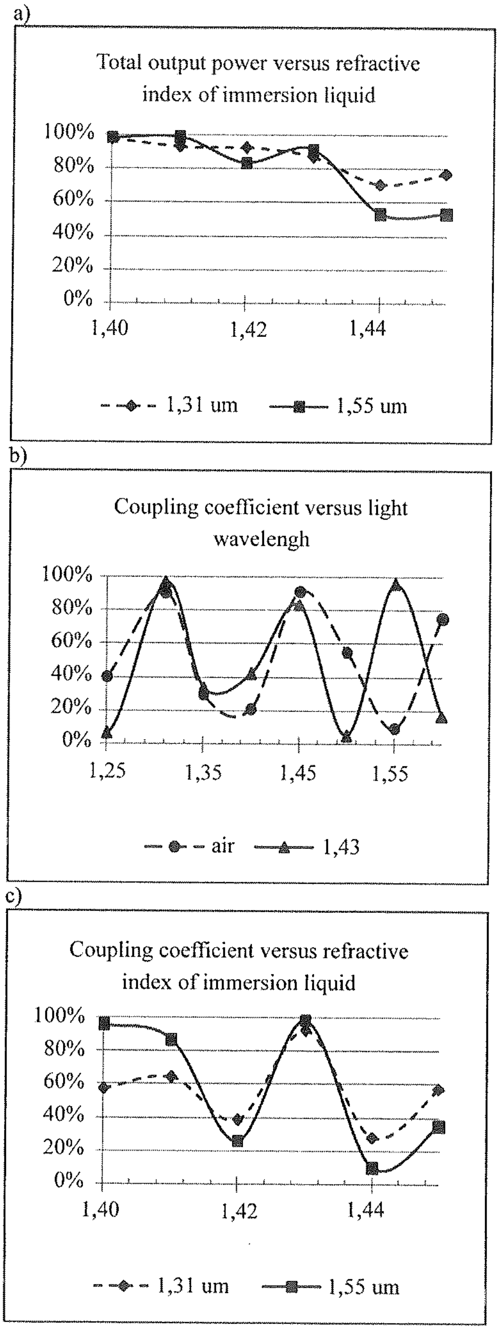


Fig. 6. Coupler D – 7 full cycles of signal transfer

responsible for coupling signals between the fibers is in good relation with the model of the symmetric and antisymmetric modal fields interaction [13]. This dependence of the refractive index of the surrounding medium on the total output power is typical for all the directional single-mode couplers that we have made.

Analysing diagram of coupling coefficient versus light wavelength (Fig.3b) we can see that for different refractive indices such coupler can be used as tunable multiplexer/demultiplexer. Similar results we can see on the others figures Figs.4b÷6b. The coupling coefficient CR_{14} of this coupler (Fig.3c) varies from 7,5% to 92% when immersed in liquids of indices between 1,40 and 1,44. In this way, by changing the refractive index of the immersion liquid, we can transfer optical power from one fiber to the other. For couplers with more cycles of signal transfers the switching can be obtained by changing the refractive index for example from 1,42 to 1,43. Similar results have been obtained for the all couplers we have made e.g. Figs.4c÷6c. It is worth to notice that complete 100% coupling was not achieved. According to the experiments it is noticeable (in Figs.3b÷6b) that when the number of full cycles of power transfer increase, the period of coupling coefficient versus wavelength doesn't change significantly and hasn't made device sufficiently sensitive, as we expected.

Experimental results agree qualitatively with previously analyzed model of coupler. A more rigorous comparison of this model with experiment requires a detailed description of changes in the modal fields in the taper transition region and the effect of cross sectional shape on the cut-off frequency of modes in the composite waveguide at the coupler waist.

5. CONCLUSIONS

The coupling coefficient CR_{14} of fused biconical single-mode coupler is shown to depend on the refractive index of the medium surrounding the taper waist. By controlling the external refractive index the coupling coefficient the signal can be switched between output fibers of the coupler. A better understanding of the coupling mechanism in fused biconical couplers should lead to improved fabrication technique, couplers performance, and couplers packaging technique. The capability to control the coupling coefficient via the external refractive index may also offer a means for constructing an optical coupler switch, fiber refractometer or tunable multiplexer / demultiplexer.

6. REFERENCES

1. B. S. Kawasaki, K. O. Hill, R. G. Lamont: *Biconical-Taper Single-Mode Fiber Coupler.* Optics Letters, 1981, Vol. 6 p. 327.
2. B. S. Kawasaki, M. Kawachi, K. O. Hill, D. C. Johnson: *A Single-Mode-Fiber Coupler With A Variable Coupling Ratio.* IEEE/OSA : J. Lightwave Technol., 1983, Vol. LT1 p. 176.

3. K. Jędrzejewski: *Biconical Fused Taper – A Universal Fibre Devices Technology*. Opto-Electronics Review, 2000, Vol. 8 pp. 153/159.
4. R. A. Bergh, G. Kotler, H. J. Shaw: *Single-Mode Fibre Optic Directional Coupler*. Electronics Letters, 1980, Vol. 16 pp. 260-261.
5. M. J. Digonnet, B. Y. Kim: *Single-Mode Fibre-Optic Components*. J. IEEE, 1988, Vol. 58(5) pp. S66-S78.
6. M. H. Slonecker, J. C. Williams: *Recent Advances In Single-Mode Fused Taper Coupler Technology*. Proc. Soc. Photo-Opt. Instrum. Eng., 1983, Vol. 412 p. 50.
7. T. Brichero, A. Fielding: *Stable Low-Loss Single-Mode Couplers*. Electronics Letters, 1984, Vol. 20 p. 230.
8. K. O. Hill, D. C. Johnson, R. C. Lamont: *Efficient Coupling-Ratio Control In Single-Mode-Fiber Biconical-Taper Couplers*. Washington : in Technical Digest, Conference on Optical Fiber Communication, Optical Society of America, 1984, paper WE2.
9. J. Bures, S. Lacroix, J. Lapierre: *Analyse d'un Coupleur Bidirectione A Fibres Optiques Monomodes Fusionnees*. Applied Optics: s.n., 1983, Vol. 22 p. 1918.
10. J. Bures, S. Lacroix, J. Lapierre: *Some Particular Properties Of Monomode Fused Fiber Couplers*. Applied Optics, 1984, Vol. 23 p. 968.
11. W. J. Stewart, J. D. Love: *Design Limitations On Tapers And Couplers In Single-Mode Fibres*. IOOC/ECOC Techn. Digest, 1985, Vol. 1 pp. 559-562.
12. J. D. Love, A. Ankiewicz: *Cutoff In Single-Mode Optical Fibre Couplers*. Electronics Letters, 1984, Vol. 20 p. 362.
13. Robert G. Lamont, Derwyn C. Johnson, Kenneth O. Hill: *Power Transfer In Fused Biconical-Taper Single-Mode Fiber Couplers: Dependence On External Refractive Index*. Applied Optics, 1985, Vol. 24, No.3, pp. 327-332
14. Cargille Laboratories: *Cargille Standard Series Liquids – Cauchy Equations*.
15. Współczynniki załamania szkła kwarcowego w temp. 24°C. Wikibooks. Wikipedia. http://pl.wikibooks.org/wiki/Aneks/Wsp%C3%B3%C5%82czynniki_za%C5%82amania.
16. M. Szustakowski: *Elementy Techniki Światłowodowej (in Polish)*. WNT, Warszawa, 1992.
17. Praca zbiorowa pod redakcją Jerzego Helsztyńskiego: *Laboratorium podstaw optoelektroiki I miernictwa optoelektronicznego (in Polish)*. Oficyna WPW, Warszawa 2003.

COF
the v
and
mod
The
Refra
glass
wide
capil
optic
Thus
guide
scatte
of li
Gaus
divid
two I
a cap
low-c
fields

Keyw

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 323–342

Guided modes in capillary optical fibers

RYSZARD S. ROMANIUK

Institute of Electronic Systems,
Warsaw University of Technology,
Nowowiejska 15/19, 00-665 Warszawa
rrom@ise.pw.edu.pl

Received 2008.07.25
Authorized 2008.09.03

A comparatively large group of capillary optical fibers, referred to in this paper as COF, consists of several families of optical filaments. The basic division line goes through the wave guidance mechanism. Two basic kinds of capillary optical fibers are of refractive and photonic mechanism of guided wave transmission. The work tries to compare wave modes in both kinds of optical fiber capillaries: refractive (RCOF) and photonic (PCOF). The differences are emphasized indicating prospective application areas of these fibers. Refractive COF carries most of the modal light in the ring-like, high-refraction, optical glass core encircling an empty capillary hole. Refractive capillary optical fibers are used widely for photonic instrumentation applications, due to the proximity of optical wave and capillary hole with the evanescent wave. The hole can be filled with a material subject to optical guided wave spectrometry. Photonic COF carries most of the light in air (or vacuum). Thus, photonic capillary optical fibers are considered for trunk optical communications, with guided wave travelling in vacuum rather than in glass - avoiding in this way the Rayleigh scattering. The fundamental mode in a refractive COF is LP_{01} or dark hollow beam (DHB) of light with zero intensity on fiber axis. The fundamental mode in a photonic COF is Gaussian beam with maximum intensity on fiber axis. The photonic COFs can be further divided to two basic groups: porous or holey/hollow and Bragg or OmniGuide fibers. These two kinds of PCOFs differ by the method of building a photonic band gap (PBG) around a capillary hole. The paper is a concise digest of fundamental kinds of singlemode (or low-order mode) COFs and their properties, with an emphasis on applications in two basic fields: instrumentation and telecommunications.

Keywords: capillary optical fibers, optical fiber capillaries, holey optical fibers, hollow optical fibers, photonic optical fibers, ring-core optical fibers, guided wave modes in optical fibers, refractive and diffractive-interference mechanism of wave guidance in optical fibers, photonic crystals

1. INTRODUCTION

Under the name of capillary optical fibers [1,2] one can understand a large family of optical, low-loss glass or polymer filaments with a single or multiple capillary spanning along the whole structure. The hole is positioned either axially in the fiber or off-axis (side hole OF). Here, we consider fibers with axial capillaries, surrounded either by refractive, or photonic (i.e. diffractive-interference) structures keeping the guided light wave, in close proximity, along the fiber axis. The COF, of a refractive structure analogous to a classical optical fiber (i.e. consisting of optical core and cladding), will be referred to as RCOF. The COF of a periodic refractive structure analogous to a photonic optical fiber will be referred to as PCOF. One can define a characteristic dimension d_{ch} of the internal refractive structures in COFs. This is a value with which are compared other dimensions in any singlemode fiber cross-section. Only for certain ranges of d_{ch} values, most frequently compared to the wavelength of transmitted light λ, one can speak of singlemode RCOF and PCOF.

A common feature of both methods of optical wave propagation in a COF is that the characteristic dimension of refractive and geometrical structures, influencing the wave, in all of these fibers, is of the order of $d_{ch} \approx \lambda/2$ [3,4]. When the characteristic dimension d_c in a fiber is much smaller than λ, what means in practice that the following condition is fulfilled $d_{ch} << \lambda$, then it is customary, in technical references, to speak about metaglass based optical fibers. Here, we speak of two kinds of materials, for refractive and photonic wave guidance, which in both cases are high quality optical glasses. The first material is called an optical fiber glass for the RCOF, while the second is called a photonic crystal for the PCOF. The quasi-crystal like structure, inside amorphous host, is created, in the latter, by the material structure rather than chemical composi-tion. Thus, the mechanism of optical beam guidance is fully refractive in singlemode RCOF, photonic, i.e. diffractive-interference in PCOF and resonant, plasmonic in metaglass COF.

Both kinds of COFs differ considerably in the construction [4,5-6]. A refractive capillary optical fiber consists of an annular, ring-like high index optical core around an air hole, a low-index optical cladding and a protective, high index jacket [1-4,10]. The fundamental mode of refractive capillary optical fiber is a dark hollow beam of light (DHB) of zero intensity on the fiber axis. A photonic capillary optical fiber consists of the capillary hole where the Gaussian light beam is guided, a cladding with two dimensional photonic bandgap keeping the light in air, a support cladding, and a jacket [5-9]. There are two kinds of structures of photonic bandgap material: porous glass of periodic refraction or circular Bragg reflector.

The major factor determining fiber construction and standards is communications. There are two basic parameters of importance in a classical communications optical fiber – losses and dispersion. Auxiliary parameters are: level of transmitted optical power at which nonlinear effects count, degrading digital transmission. The nonlinear effects are: self phase modulation, skew phase modulation, four photon mixing, stimulated

Rayleigh and Brillouin scattering and others. The total losses have two components – dissipation and absorption. A sum of these components, in a silica, single mode telecom fiber, for the wavelength of minimal losses $\lambda_{s(\min)} \sim 1,55\mu m$ is slightly below 0,15dB/km. Signal dispersion has two components: material and waveguide. The dispersion may be zeroed in a point and in several points for a particular band, usually, at a cost of some excess losses.

The capillary optical fiber, which is subject of consideration and comparison with classical optical fibers, possesses different properties concerning the above mentioned transmission parameters. This rises serious hopes for a new generation of optical fiber industrial trunk cable as well as instrumentation systems. The light wave can be guided inside an optical fiber in air rather than in glass.

A classical telecom oriented optical fiber has shifted dispersion characteristic to the area of the ultimate low-loss region. A telecommunications oriented capillary optical fiber is such a ultra low loss fiber, in which optical multimode or single mode wave, caused by arbitrary mechanism, refractive of diffractive-interference, is guided in a capillary hole (vacuum or air), which hole stretches along the whole filament (glass, polymer, composite, covered with noble metal, or metal and polymer). The propagation mechanism depends on fiber construction. In a photonic capillary optical fiber it is purely diffractive-interference. In a photonic porous fiber (with glass core) it may be also refractive, when the effective refraction of porous cladding is lower than the solid region of the core. The fundamental mode in a RCOF is HE_{11} in a form of a DHB, while in a PCOF it is HE_{11} in a form of Gaussian beam.

Refractive guidance in a capillary optical fiber is realized for small wavelengths in the region of VUV and X-rays, or for long wavelengths in the region of MIR and FIR, where the refraction of fiber materials is anomalous with $n<1$. This always lossy mechanism, the total external reflection, is effective only for grazing angle of the wave incidence. Thus, the fiber could not be bent too much. Above the critical angle, which is very small, the transmission losses increase abruptly.

The future of optical communications seems to be tight to transmission of light in the vacuum. Probably, the only capillary, single mode optical fibers, which are to be considered for telecommunications applications, will use the photonic mechanism of wave propagation. In a photonic capillary fiber, there exists in the cladding a two dimensional photonic band gap for particular wavelengths and directions of propagation (per analogism to a band gap in semiconductors). The band gap is tailored technologically (dimensions and number of holes in porous cladding, capillary dimension of core hole, value of air/glass factor), for the wavelength of minimum fiber losses $\lambda_{s(\min)}$. There is no need, in a capillary optical fiber, to shift the spectral dispersion characteristic minimum in the direction of minimal losses (as is in the classical optical fiber), because even 99% of optical wave power may be propagated in air (vacuum) of the capillary core.

The construction of studied capillary optical fibers for telecom industry are proposed in two solutions. These solutions are very similar, either to a classical fiber, or

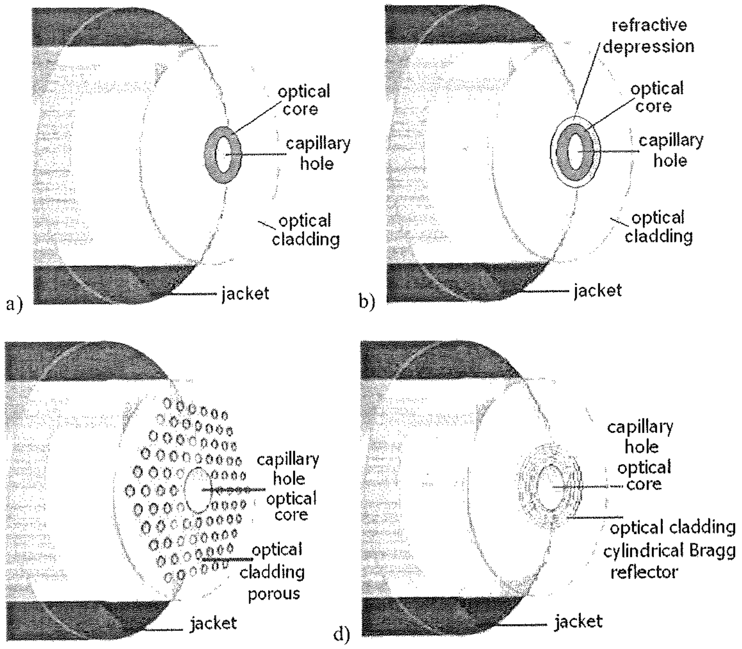


Fig. 1. Basic kinds of singlemode capillary optical fibers (COF), (a) refractive COF, referred to as RCOF, with ring-like, high-index optical core, (b) RCOF with refractive depression, referred to as DCOF, (c) photonic COF (PCOF), with porous optical cladding, referred to as PPCOF, (d) PCOF with Bragg optical cladding, referred to as BPCOF. The simplest solution of the RCOF, presented in (a) will be called throughout this work as COF. Characteristic dimension d_{ch} 23 M fig. 9c and d the distance between nanoholes in poroks cladding or periodic lengt of Bragg structure

to a photonic optical fiber, fig.1. In both cases, the core is a single mode capillary hole. In the first case, the optical cladding is a multilayer, cylindrical Bragg reflector. This solution is called OmniGuide, from the company which manufactured the first samples of this fiber [11]. The second solution, assumes porous optical cladding as a poly-capillary. Periodical distribution of nano-holes creates a photonic band gap around the optical core. This solution is called holey. The first solution of this type, which is close to commercialization, was proposed by the company BlazePhotonics [12]. Theoretically, there may also be considered an intermediate solution, of an ultralow loss, single mode COF for the HE_{11} mode, with a ring index core and photonic cladding. Such a fiber combines the features of RCOF and PCOF families.

2. DIFFERENCES BETWEEN COF AND CLASSICAL OPTICAL FIBER

There is a number of different factors which confine the transmission rate in a classical optical fiber transmission channel. These factors change considerably, when the transmission channel uses capillary optical fiber instead of classical, ultra low loss, or erbium doped, amplifying, high silica fiber.

Light velocity of an optical wave in a classical, silica telecom fiber is approximately 2/3 of the light velocity in vacuum, because the glass refraction is 1,45. Why not transmit the data faster? Taking into account the obvious fact that the data are transmitted at a group velocity rather than at a phase velocity.

The range of repeater-less, transmission line in classical system, with fiber non doped with erbium, is confined by losses s [dB/km], which are 0,2[dB/km], for the wavelength in the vicinity of zero material dispersion. The minimum losses are around 0,15dB/km. The wave is not subject to Raleigh scattering in vacuum. It is estimated, that the losses in capillary optical fiber can be at least ten times smaller and reach 0,01[dB/km]. In the capillary optical fiber, 99% of optical power may be guided in vacuum. Now the best capillaries have losses close to 1dB/km.

The range of repeater-less optical fiber line is also confined by the maximum allowed optical power that can be coupled to the fiber. Above a certain power, the material of fiber is optically nonlinear and such effects appear as: self phase modulation, skew phase modulation, four photon mixing. Nonlinear phenomena cause saturation of characteristic of the transmission rate as function of power in the channel. Transmission rate goes with power to the infinity in an ideal channel. In the vacuum, nonlinear effects are not caused by the fiber material.

The distance range of the transmission channel, assuming negligible losses of the fiber, is determined by signal dispersion. Dispersion stems from the dependence of fiber material refraction on wavelength. Broadband fiber transmission line carries monochromatic wave. This wave, however, has a finite spectral width. Thus, a residual material dispersion is present. The second component of dispersion is waveguide one. It is caused by different eigen-refractions for each mode propagated in fiber. The refractive index for all wavelengths in vacuum is 1.

Partial reflection of the wave at the boundaries between objects of different refractions is called the Fresnel reflection. In the capillary optical fiber, the Fresnel effect is present only at the fiber coupling with the source and detector and between the butt-coupled fibers.

The fiber glass has an amorphous, molecular structure. This structure features Raleigh scattering and (for shorter wavelengths) and multi-phonon absorption (for longer wavelengths). The scattering is a source of optical noise. Raleigh scattering has a residual nature in a capillary optical fiber.

To build a classical optical fibre, at least two materials (different glasses) are needed for core and cladding. In practice there are needed more than two different materials. In practice these are gradient index ones. All of them interact with the optical wave. These materials create core structure with gradient refractive index profile and optical cladding. A capillary optical fibre, in the porous solution may be manufactured from a single material – the best one, till now – ultrapure silica glass without any dopants.

3. REFRACTIVE KIND OF CAPILLARY OPTICAL FIBERS

An optical fiber, including COF, is fully described by a confined number of refractive and geometrical data. The input data of refraction and geometry in a COF for the modal analysis are:

- refractions: capillary $n_o = 1$, n_r core, n_p cladding, n_d depressed cladding
- differential refractions: Δn_{rd}, Δn_{rp}, Δn_{pd}
- numerical aperture: $NA_r = sqrt(n_r^2 - n_p^2)$, $NA_d = sqrt(n_p^2 - n_d^2)$,
- radiuses of areas: r_c capillary, r_r core, r_d refractive depression in cladding,
- relative radius of the capillary hole: $r_{rc}=r_r/r_c$,
- thickness of areas: $d_c = 2r_c, d_r = r_r - r_c$ core, $d_d = r_d - r_r$ refractive depression in cladding,
- refraction for analyzed COFs: high silica n_p=1,46, compound glass n_p=1,56, for $\lambda = 1\mu$m.
- numerical aperture: NA=0,1 for high silica fibers and NA=0,2 for compound glass fibers.
- dimensions: $r_c = 1 - 10\mu$m for all fibers.
- the additional assumptions are that the analyzed fiber has negligible optical losses.

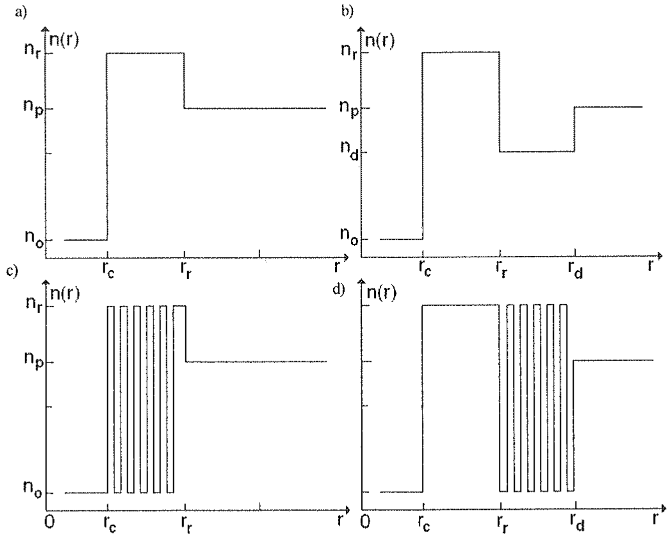


Fig. 2. Schematic diagrams of refractive index profiles n(r) of chosen fibers from a COF family; (a) and (b) refractive COFs, (c) and (d) photonic COFs; Not to scale. Left to right: (a) basic COF, (b) depressed cladding COF – DCOF, (c) Photonic COF – PCOF, (d) complex COF with ring index core, like basic COF, and photonic cladding. The characteristic dimension d_{ch} in figs. c and d is the period of refrachon chaupes

A basic geometrical solution of the refractive capillary fiber with a uniform cladding is called a COF, while a fiber with non-uniform cladding is called a DCOF (depressed cladding COF). A photonic capillary fiber is called a PCOF. Some of COF refractive indices in fiber cross-section are presented in fig.2. There are also possible other refractive index modifications. The common feature of the index profile is that, in any case of a COF, there is a capillary hole, filled with air, vacuum or other substance, liquid or gaseous, at the fiber axis.

The EM field is decomposed in a fiber into cylindrical rather than Cartesian components. The weak guidance assumption (which is a considerable simplification of a real situation) excludes hybrid field components and leaves only the transverse ones which is equivalent to the consideration of a fully linearly polarized plane wave. The electrical field components of a weakly guiding mode in a COF have the following form: $E(r,\phi,z) = e(r)\exp(-\ jm\phi)\exp(-\ j\beta x)$, where r,ϕ – radial and angular coordinates in a plane perpendicular to the fiber axis, z – length measured along fiber axis, $m = 0, \pm 1, \pm 2, ...$ – azimuthal mode number, β – propagation constant of a mode associated with effective refraction via the relation $n_{eff} = \beta/k_o$, $k_o = 2\pi/\lambda$ is a wavenumber in vacuum. Radial dependence of transverse field component $e(r)$ is an eigensolution of Helmholtz equation of the form: $e(r) = A_0 I_m(vr)$ for $r < r_c$; $A_1 J_m(ur) + A_2 Y_m(ur)$ for $r_c < r < r_r$; $A_3 K_m(wr)$ for $ar > r_r$, in a COF, and $e(r) = A_0 I_m(vr)$ for $r < r_c$. $A_1 J_m(ur) + A_2 Y_m(ur)$ for $r_c < r < r_r$; $A_3 I_m(wr) + A_4 Y_m(wr)$ for $r_r < r < r_d$; $A_5 K_m(sr)$ for $r > r_d$, in a DCOF, where $A_i (i = 0,1,2,3,4,5) = const$. J_m, Y_m are Bessel functions of the first and second kind, I_m, K_m are modified Bessel functions of the first and second kind, m-th order.

In certain fiber regions, like in the optical core, the EM field has a periodic distribution, i.e. is quided, while in the cladding and in the capillary hole, the EM field distribution is aperiodic, i.e. evanescent. Modal parameters, which are Bessel function arguments v, u, w, s for particular refraction areas of fiber cross section with periodic and aperiodic field distribution, are defined as follows: $v^2 = \beta^2 - k_o^2$, $u^2 = n_r^2 k_o^2 - \beta^2$, $w^2 = \beta^2 - n_d^2 k_o^2$, $s^2 = \beta^2 - n_p^2 k_o^2$. These parameters are combined with the propagation constant and, thus, have eigenvalues for individual modes.

The law of continuity condition preservation requires that field values and fields derivatives in the radial direction must be continuous on every boundary in the fiber cross-section: $e(r_c^-) = e(r_c^+)$, $de(r_c^-)dr = e(r_c^+)/dr$, $e(r_r^-) = e(r_r^+)$, $de(r_r^-)/dr = e(r_r^+)/dr$, $e(r_d^-) = e(r_d)$, $de(r_d^-)/dr = de(r_d)/dr$. A characteristic equation is formed from these conditions. The following matrices have to assume zero value, for COF, where the apostrophe means Bessel function derivative [4]:

$$\begin{vmatrix} I_m(vr_c) & -J_m(ur_c) & -Y_m(ur_c) & 0 \\ vI_m'(vr_c) & -uJ_m'(ur_c) & -uY_m'(ur_c) & 0 \\ 0 & J_m(ur_r) & Y_m(ur_r) & -K_m(wr_r) \\ 0 & uJ_m'(ur_r) & uY_m'(ur_r) & -wK_m'(wr_r) \end{vmatrix} = 0 \qquad (1)$$

An analogous equation, as above for a basic COF, is fulfilled for a DCOF, with this obvious difference that more refractive regions are to be considered. The field in the refractive depression is also evanescent but with a large geometrical extent due to low index contrast between the adjacent regions.

$$\begin{vmatrix} I_m(vr_c) & -J_m(ur_c) & -Y_m(ur_c) & 0 & 0 & 0 \\ vI_m^{'}(vr_c) & -uJ_m^{'}(ur_c) & -uY_m^{'}(ur_c) & 0 & 0 & 0 \\ 0 & J_m(ur_r) & Y_m(ur_r) & -I_m(wr_r) & -K_m(wr_r) & 0 \\ 0 & -uJ_m^{'}(ur_r) & -uY_m^{'}(ur_r) & -wK_m^{'}(wr_r) & -K_m^{'}(wr_d) & 0 \\ 0 & 0 & 0 & I_m(wr_d) & K_m(wr_d) & -K_m(sr_d) \\ 0 & 0 & 0 & wI_m^{'}(wr_d) & wK_m^{'}(wr_d) & -sK_m^{'}(sr_d) \end{vmatrix} = 0 \tag{2}$$

All the eigenvalues of wave modes in fiber, like the propagation constant β_{ml} and wave parameters v,u,w,s are obtained by numerical solution of the characteristic equation for each azimuthal m and radial $l=1,2,3...$ mode number. The modal cut-off is when the effective refraction is equal to cladding refraction. The transverse field distribution is obtained from equations $e(r)$ after determination of A_i constants from the β_{ml}. Figures below present calculation results for COF and DCOF fibers of given parameters. The parameters were chosen as realistic ones, available during the technological process of fiber pulling.

By applying a parameter of relative capillary hole dimension $r_{rc}=r_r/r_c$ (instead of absolute dimensions r_r and r_c), in the above expressions for the modal fields, in particular regions of the fiber, the eigenequation of the COF may be written in an analogous form like for a classical, weakly guiding optical fiber. This more general form of eigenequation (3) is reduced to a classical one when an obvious condition is fulfilled $r_{rc} \rightarrow \infty$. This form of the eigenequation for the COF allows for easier analysis of similarities and differences between both classes of optical fibers.

$$\begin{gathered} K_m +_1 (w)/uK_m(w) = \{uI_m(w/r_{rc})[J_m +_1 (u)Y_m +_1 (u/r_{rc})- \\ +J_m +_1 (u/r_{rc})Y_m +_1 (u)] + wI_m +_1 (w/r_{rc})[J_m +_1 (u)Y_m(u/r_{rc})- \\ +J_m(u/r_{rc})Y_m +_1 (u)]\}/uI_m(w/r_{rc})[J_m +_1 (u/r_{rc})Y_m(u) - J_m(u)Y_m +_1 (u/r_{rc})]+ \\ +wI_m +_1 (w/r_{rc})[J_m(u/r_{rc})Y_m(u) - J_m(u)Y_m(u/r_{rc})]. \end{gathered} \tag{3}$$

By analyzing the numerical solutions for COF eigenmodes for a range of realistic technological parameters, the set of wave-fiber behavior can be fully revealed. Some of the conclusions from the numerical solutions, related to the applications of COF, were gathered below. Basic differences between a classical, step-index, singlemode optical fiber and a COF were emphasized. The following properties of all fiber modes are completely determined by the effective refraction: field distribution, dispersion, overlap integral of modal field with the core, micro-bending losses, mode susceptibility to refraction and geometry changes in the fiber.

The effective refraction may change within the boundaries determined by maximal and minimal physical refractions in the fiber. The rate of change for effective refraction and mutual distances between its quantized values are determined by the local averaged value of physical refraction (refractive profile).

In the energetic interpretation, the squared and inversed refractive profile plays a role of a potential well. Comparison of a full vectorial analysis, done in MatLab environment, with the LP (linearly polarized) approximation yields an error not bigger than a fraction of %, for the most practical constructions of COF, further supporting the validity of weak guidance assumption.

The fundamental mode in both cases is LP_{01}. This notation 01 means that there is one field maximum in the radial cross section of the core and no field changes in the angular direction for a certain radius. The fundamental mode in a COF has zero field value at the fiber axis, forced by the capillary hole, while in a classical fiber it reaches the maximum there. A singlemode classical optical fiber of step-index and simple monotonic gradient profiles has always a zero value of the cut-off frequency for the fundamental mode.

The following condition is required for a fiber from the COF family to possess a nonzero cut-off frequency. A COF (as well as M and W fibers) has a nonzero cut-off frequency of the fundamental mode then, and only then, when the averaged value of the whole optical core refraction is lower than the relevant refraction in the cladding. A local depression of refraction in the cladding, height of refraction in the core and change in geometrical proportions of these areas lead, when the non-zero cut-off does exist, to the effective change of the cut-off value, what was presented in fig.3. Changes in the cut-off of fundamental mode turns the COF to an effective modal converter and filter, with the efficiency measured by the rate of the cut-off sweep against the normalized frequency V. The modal field selectivity is determined by the value of dn_{eff}/dV in the vicinity of cut-off. A number of other practical conditions have to be fulfilled to make an efficient modal filter out of a COF. The most important are: modes have to be clearly separated, i.e. possible for distinguishing basing on far field pattern in laboratory conditions, modal cut-off should be clearly defined in V space and not extend over a large range of V values, the modal cut-off should not depend too much on fiber disturbances like bending, vibration – called microphonics, heating, etc.

One of the essential questions of singlemode optical fibers, including COFs is the acceptable level of absolute power and power density in the core and the efficiency of power coupling from the source. An extended ring-like distribution of EM field in the fundamental mode facilitates pumping the COF with optical power nearly an order of magnitude higher than in a classical step-index singlemode optical fiber. The pumping is quite efficient over the whole area including the annular core, capillary hole and refractive depression around the core. Nearly all the power coupled to the capillary hole is accepted by the core. Due to this characteristic feature, single-mode and multimode COFs are extensively used in active optical fiber systems and optical fiber lasers as active medium and for a number of functional devices.

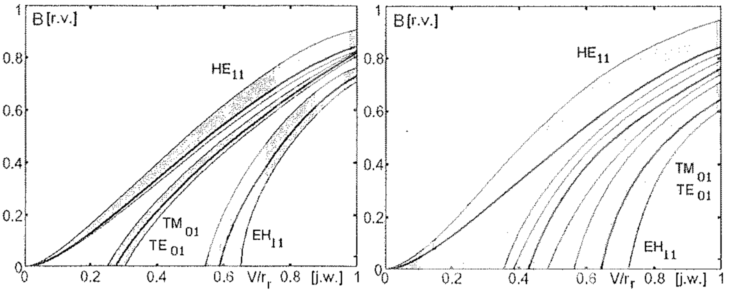


Fig. 3. Calculated dispersion characteristics and modal cut-off [10]; Normalized propagation constant $B=\beta_N=\mathrm{sqrt}[\beta^2/k^2-n_p^2)/NA^2]$ as a function of normalized frequency V/r_r for the lowest order modes in a refractive family of capillary optical fiber RCOF; Refractive and dimension data for the thick solid line characteristics: λ=800nm, d_c=3,5μm, $d_r=2r_r$=10μm, Δn_{rp}=0,2%, n−p=1,50. (a) left – COF, (b) right – DCOF. Darkened areas mean the extent of changes for practical range of available technological parameters of optical fiber; The characteristics were chosen in this way as not to overlap too much. In practice, they overlap considerably; [r.v.] - relative values. As opposite to a COF, the DCOF can have a wide extent of nonzero cut-off frequencies for the fundamental mode, for a certain range of V/r_r values

The COFs behave differently depending on the capillary hole diameter. The limits are very small holes, where the COF is similar to a classical singlemode fiber, and very large holes, where the fiber is similar to a planar optical waveguide. Between these two poles, two classes of COFs may be distinguished: small hole and large hole COFs, where the criterion is the value of the rate c/λ. Small hole COFs are the ones with this ratio around one, while big hole COFs are the ones with this ratio closer to 10. From practical point of view and potential applications both kinds of COFs: small hole and large hole, are very interesting.

A small hole COF is interesting because it is strictly singlemode, as opposite to the large hole COF, where the singlemode condition can be enforced by relatively easy removing of the higher order modes. A small diameter of the hole causes larger separation between the fundamental mode and the next one. The effective area where the COF is strictly singlemode can be even more extended in a DCOF. Simultaneously, the fundamental mode enters deeper into the optical cladding. The cut-off wavelength of LP_{01} increases with the hole diameter. The changes in cut-off for fundamental mode with the hole diameter are small in comparison with the relevant changes for higher order modes. The cut-off of a fundamental mode is relatively soft and slow. The values of derivatives dn_{eff}/dd_r and $dn_{eff}/d\lambda$, for $n_{eff} \approx n_p$ are small, which determines weak properties of such a fiber as a modal filter. When the fundamental mode is far from cut-off, or equivalently is strongly guided, the value of effective refraction is relatively high and sensitivity of the mode to fiber geometry (including microbending) and refraction is low. Such a fiber is not a source of modal noise. Coiling the COF (similarly to W and M fibers) causes shortening cut-off of the fundamental mode. Heating of COF causes slight lengthening of the cut-off.

A COF is really (in the meaning of laboratory practice) singlemode only for small enough diameters of the capillary hole. A confinement for small cores is when the modal field on the fiber axis is not zero, thus, interferes with the field from the opposite direction. Theoretically, for big holes it may be also singlemode. However, a large diameter of the hole in a COF (of the order 10 λ), causes nearly total overlapping of the cut-off characteristics for the fundamental mode and a group of the next ones. This means that the COF of large hole, even for thin core of high refraction (which acts contradictory) is susceptible to carrying many modes and the separation of these modes from the fundamental one turns impossible. However, separation of the modes LP_{02}, LP_{03},...from the fundamental one in a low mode COF is possible, while the separation of modes LP_{11}, LP_{21},... is nearly impossible. In a multimode COF of large dimensions of the refractive areas in comparison with the wavelength, there are propagated a few different kinds of surface modes - internal, with caustics positioned near the surface of the capillary hole, and external, (whispering gallery) with caustics positioned near the core-cladding boundary.

COF and DCOF fibers are in practice very similar to each other. In comparison with a COF, the DCOF has two more design parameters, which are refraction n_d and geometry d_d – the data of refractive depression. Thus, the DCOF is slightly more difficult to be designed, when compared with a basic COF. For a set core refraction, the fundamental mode cut-off is determined mainly by core thickness d_r and numerical aperture NA_r, and much weaker by other parameters like hole diameter d_c, width of refractive depression in the cladding d_d and numerical aperture of the depression NA_d. The cut-off wavelength of fundamental mode increases with the core thickness and with the capillary hole diameter. Refractive characteristic of the fundamental mode of DCOF as function of optical core thickness are analogous to the ones in COF with this difference that the values of derivatives dn_{eff}/dd_r and $dn_{eff}/d\lambda$, for $n_{eff} \approx n_p$ are bigger, which more precisely defines the very moment of the modal cut-off. The DCOF has much bigger effective core area than the COF and W fiber at similar values of refractive and dispersion parameters. Thus, the effective modal area is bigger and the fiber resembles more the ones with big effective areas A_{eff}.

One of the most fundamental characteristics of DCOF is the range of a singlemode area as a function of depression d_d and capillary hole d_c parameters. For example, for d_c=5μm, which is quite small, and for low-refraction, high-silica fiber, the DCOF is not guiding the fundamental mode, when the thickness of refractive depression in the cladding is bigger than d_d > 3μm. The core-cladding refractive depression is an effective tool in designing the single mode area of a DCOF. A diameter of the hole also influences the fundamental mode cut-off. Large diameter of the hole, up to a certain value, keeping the fiber single mode, compensates large width of refractive depression in the cladding. On the other hand, small diameter of the hole, requires compensation by stronger depression, for the modal cut-off not to be zeroed.

The characteristics in Fig.3 show, additionally, on the background of the fundamental mode range of characteristics, the areas of cut-off changeability of the higher order modes. Usually, for some practical range of characteristics, the modes are difficult for separation. The shaded areas were calculated for small, a few % changes in the COF and DCOF refraction and geometry. The fiber is unconditionally singlemode left to the HE_{11} characteristic. The widths of the areas are different for the fundamental and higher order modes. Because the relative value of effective index in the DCOF was around 0,3, which is not big, the fundamental mode is weakly confined by the core, in comparison with the COF, and its field deeply penetrates other refractive layers of the fiber. For the wavelength $\lambda \approx 1\mu m$ this particular DCOF fiber is singlemode for $d_c < 5\mu m$ and for a considerable refractive depression in the cladding. Effective refraction of the fundamental mode grows when d_c is smaller, making the fiber more resistant to microbending losses. Smaller diameter of the hole causes stronger propagation of the fundamental mode in DCOF, but also means smaller effective area of the core and lower efficiency of high power pumping. A compromise between the hole diameter and core thickness is a matter of fiber application. Strictly singlemode applications require either small hole COFs or larger hole, strong depression DCOFs.

Comparing COF and DCOF, the latter has a relatively broad possibility to shape the modal field distribution relative to its refractive and geometrical structure. In the active DCOF (and in other active fibers) an important parameter is an overlap factor of the modal field with the doped area of the fiber, or the active core. The pumping area is different in various fibers: in DCOF it is the whole area of the refractive core and the capillary hole, in double clad W and M fibers it is the inner cladding. The overlap factor in a COF, in the area close to the fundamental mode cut-off, abruptly decreases, due to fast increase of the modal field diameter. The efficiency of a COF laser is, thus, low. A different situation is in a DCOF, where the optical field is more confined by the refractive depression in the cladding. Additionally, the field is less attenuated which is equivalent to more distinctive marking of the cut-off point. This also could be seen in fig.3 from the slope of the particular characteristics.

There is a distinctive difference between the W, DCOF and COF fibers relative to the effective area of the core and the fundamental mode. The smallest effective area is in the W fiber for a set of equivalent refractive and geometrical parameters. Simultaneously, the W fiber has the broadest area where it is singlemode. The W and DCOF fibers are comparably resistant to micro-bending losses, while the COF is less resistive. Comparison of the analogous characteristics of COF and fibers from a similar family of refractive profiles gives necessary arguments for application of particular fibers in a variety of instrumental applications. The biggest similarity is between the W and DCOF refractive fibers, as these two families belong to the large effective area fibers. The other pair of similarity are COF and M (ring-index) fibers.

4. PHOTONIC KIND OF CAPILLARY OPTICAL FIBERS

Periodically porous optical cladding creates around a capillary hole, a two dimensional photonic band gap, efficiently confining optical field to the empty core, nondependent of fiber refraction. Field confinement in the capillary hole is effective for a narrow range of wavelengths, for example 1300-1800nm, or other range, depending on fiber construction. The basic technological problem is to obtain an ideal subtle structure of photonic capillary optical fiber, i.e. the smallest irregularities in cell shapes, the smoothest surface of the core hole and nano-capillaries in the cladding.

Building of micro-capillaries inside the filament is a thermodynamic process, thus certain degree of non-ideality for internal surfaces is non avoidable. These non-idealities [7-9] cause power losses from the fundamental mode, to non-propagated, lossy modes of higher order, or to surface modes propagated only at the air/glass boundary. These two loss mechanisms: optical surface modes (SM) and surface capillary wave (SCW) are now the fundamental confinement of broad applications of capillary optical fibers in telecommunications. Both loss mechanisms are associated with perturbation mode coupling. The perturbations in capillary optical fiber can be the following: internal surfaces (air/glass) roughness, axial changes of structure scale, tensions, bending, other causes leading to the change in effective refractive index along fiber axis. Figs.1 and 2. present a few kinds of construction solutions for photonic capillary optical fibers (besides refractive COFs) with different filling factor of air spaces in the glass, for porous solution.

Similarly to a classical optical fiber, the photonic optical fiber capillary possesses a finite number of well defined, guided, ultra low loss, core modes, and infinite number of leaky cladding modes and radiation modes. In the ideal case, the core modes carry almost all the optical power in air. The fundamental core modes in capillary optical fiber possess larger effective refractive index than the modes of higher order. Structure perturbation of silica glass only weakly couple core modes to the cladding ones via a small region of geometrical overlapping of the fields. Additional a special group of core, guided modes in periodic structures on the boundary of areas of different refraction are surface modes. Surface modes cannot be avoided, but their meaning may be reduced via optimized structure of single mode photonic capillary optical fiber (PCOF). In the initial solutions of the PCOF, the surface modes manifested their presence by strong resonant, quite wide (100nm) dissipation peak, present in the transmission band (a few hundred nm). The negative role of surface modes relies on being an effective liaison on the path of couplings between fundamental core modes and weakly lossy modes allowed by the fiber structure.

The surface modes possess the field maximum on the glass (periodic structure) and air boundary. The field decreases exponentially in both areas, glass and air, stronger in air , weaker in glass. The character of surface modes changes depending on distance of the areas boundary from periodic area. In a PCOF, the boundary and distance of periodic region is defined by the diameter of capillary core and air filling factor (i.e number of nanocapillaries). Internal surface of capillary enables propagation of surface modes of eigen-frequencies located inside the allowed photonic transmission band. Most of the optical power of the surface modes is localized in the glass on the boundary of optical core (for guided modes), or macro-capillary and optical cladding or nano-capillaries. If the amount of glass is reduced just around the optical core (capillary hole), the energies of surface modes increase, their field maximum shifts closer to the boundary, the penetration depth to the core increases, increasing the overlapping with core modes. From other side, the surface modes extend comparatively deep into the periodic structure of a PCOF, overlapping their own cladding evanescent field with lossy cladding modes.

Fig.4 presents schematically o diagram of modal structure for a PCOF. The difference between Fig.3 and Fig.4, which present nearly the same dispersive modal characteristics, can be explained as follows. Photonic crystal (or circular Bragg grating) fiber structure imposes additional propagation confinement on the modes. This confinement is called, per analogism to the semiconductor bandgap, a photonic bandgap. Inside the area of allowed propagation the modes behave similarly, with this difference that there are also surface modes guided refractively in the glass, besides capillary hole modes guided in the air. Low loss core modes have weak dispersion because they are guided nearly totally in the air (relative to other modes). Surface modes have much bigger dispersion. Fundamental modes have bigger effective eigen-refractions than the higher order modes. For certain wavelengths, the modal refractive indices equalize (respective modal refraction curves cross), and the character of individual modes may disappear. At some moment both modes, surface and core, would have the same propagation constant. That would be a very efficient means of optical power exchange between nondispersive core mode with dispersive surface mode and vice versa. In reality, in the transition period, there exist hybrid modes core-surface, out of which again separate core and surface modes appear, with the change in wavelength.

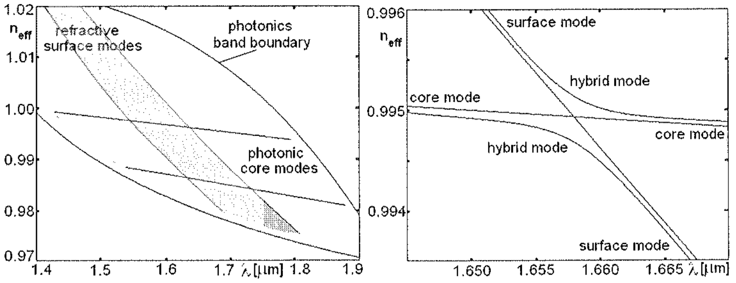


Fig. 4. Dispersion characteristics of fundamental core modes, surface modes and transient hybrid modes in a PCOF. Effective modal coefficient of refraction as a function of wavelength $n_{eff}(\lambda)$. Cladding modes propagate out of the photonic band but close to its boundary. (a) left – global picture showing the position of the individual modal characteristics and photonic bandgap edges; (b) right – detailed picture of guided-photonic and surface-refractive modal characteristics crossings and, thus, enabling power exchange between these modes [7-9]

The surface modes strongly depend on the internal diameter of capillary core. The possibilities to change this diameter are confined by the single mode condition. The surface modes strongly depend, in a discrete way on subtle structure of core surface. It is possible to choose such periodicity of this surface, that surface modes are eliminated or considerably reduced. Shape of the core is then not cylindrical but periodically waved.

Capillary Surface Wave in glass has to be considered in more deteil. The mechanism of creating periodic nano wave roughness on the internal glass surface of capillaries is analogous to the creation of wrinkles on flowing water surface or water excited by blowing wind. Building of wrinkles, called here capillary surface wave (CSW), may be a resonance process – out of this comes the periodicity, and is a dynamic balance between external force (optical fiber pulling) and surface tension. The bigger is the surface tension, the smaller is the surface perturbation amplitude and the space frequency is bigger. For the average technological conditions of high silica PCOF manufacturing (temperature, viscosity, surface tension, rate of fiber pulling and preform feeding) the amplitude of CSW is of the order of 0,1nm (for pure silica) and bigger than 0,5nm (for compound glasses, which have smaller surface tension). At first sight it may seem that CSW should not influence the optical wave propagation. However, in an ultra low loss telecom fiber, the interaction length is so big (tens of km), that the residual effects of modal transformations accumulate positively to relevant levels. It is estimated, that in a telecom grade PCOF, the nano CSW, of the dimension 0,1nm, distorting the surface of glass, are the cause of optical scattering of the order of 1dB/km. These results confirm recent achievements in PCOF losses reduction, for the best now available version of technology, just to this level. This level was achieved for $\lambda_{s(\min)} \sim 1,75\mu\mathrm{m}$. The attenuation of 1dB/km possesses almost only the scattering component, at estimated absorption component well below 0,1dB/km.

Assuming that nearly all losses from the best result of 1dB/km (which is a value non acceptable by telecom industry) come from the CSW effect, the only method of further development of telecom grade COF is precise investigation of SCW nature. It is necessary to answer a question which factors associated with SCW can be eliminated and which cannot, and if not, how to minimize them? In a classical telecom, pure silica optical fiber, the losses equal to 0,15dB/km stem from fundamental scattering and material absorption processes, and practically cannot be further minimized. The PCOF carries 99% of optical power in the air. Thus, the material factor of optical losses (at the same material for classical optical fiber and PCOF), calculating conservatively, may be reduced at least to 0,01dB/km. There is omitted direct radiation from the capillary fundamental mode, because it is caused solely by non-ideality of photonic bandgap, which is a consequence of non-idealities in fiber microstructure.

The thermodynamic equilibrium requires that each capillary surface wave (here present on the surface of molten glass, in the area of outfow meniscus of pulled fiber) has average energy proportional to: $E_K \sim 1/2k_BT$, where k_B Boltzmann constant, T-glass temperature. When the glass of pulled fiber cools down, the CSW gets frozen with the energy determining its amplitude and phase, for $T=T_g$, where T_g is a characteristic phase transition temperature for each glass. For pure silica glass $T_g \sim 1500K$, and for compound glasses this temperature is considerably lower. The average square spectral power density of surface roughness decides of optical power dissipation, for the space frequency spectrum in a confined bandwidth f+Δf:

$$S(f) = k_B T_g / 4\pi\gamma f, \tag{4}$$

where γ is surface tension (for silica glass $\gamma = 0,3\mathrm{J/m^2}$). The above dependence, valid for a nonconfined flat surface of liquid, is modified for capillary for diameter d:

$$S_k(f) = S(f)\coth(fd/2). \tag{5}$$

S_k=S for small and big space frequencies f. The spectrum of space frequencies f is confined from up and down by: gravitation, fiber dimensions, value of T_g, etc. Integrating the dependence on statistics of capillary fiber surface roughness, for sufficiently broad space frequency range, and small in comparison with dimension d, one obtains a root mean square roughness of glass depending on T_g and γ. For the pure silica glass it is s_{rms}=0,1nm, and for soft glasses it is several times more [8-9].

The roughness of glass surface (internal surface of optical core capillary hole) is the main cause of guided optical wave scattering from the fundamental mode to other modes. Scattering is effective if the space frequency of particular roughness component is equal to the differential value of propagation constants of potentially coupled modes – fundamental and coupled:

$$f = k(n_i - n_o), \tag{6}$$

where n_i and n_o are own, effective refraction indices of I-th coupled mode and fundamental mode o-th. Efficiency of coupling of both modes (measured by by rate of energy flow between them) is determined by S_k, which is an integral of modal fields overlapping on the glass/air boundary in the fiber cross section. Calculation of these overlapping integrals is associated with knowledge of precise distribution of modal fields and modal propagation constants. Total loss of capillary fiber is a statistical sum of power coupled to all modes, guided and cladding. In the capillary optical fiber, the modal field dimension, of al modes, is very strongly defined by the hole diameter, thus, in approximation, the field overlapping integrals are proportional to the part of optical power in fundamental mode carried in the hole, and to the differences of own effective refraction coefficients of fundamental mode and coupled mode. It is assumed, that the higher order modes preserve the proportions of penetrating the capillary glass, taking into account the difference in modal eigenrefractions. To obtain the total power losses (scattering) for the length unity of fiber, requires integration of S_k across the whole range of existing in fiber differences in effective modal refractions. The effective modal refractions (in a capillary fiber they are close to 1) are associated with space frequency of the mode, a subtle, discrete direction of propagation. Measuring the angular space distribution (against the axis of fiber) of scattered optical power from the PCOF, one may determine the spectrum of modal couplings.

One of the fundamental properties of a photonic optical fiber (the capillary considered here, with porous cladding is a POF) is a precise scaling of its properties in function of wavelength, for a constant rate of structure dimensions to the wavelength. For example, the $\lambda_{s(min)}$ – wavelength of minimum losses, is subject to proportional scaling. Other property of POF is dependence of scattering of propagated wave as $\lambda_{s(min)}^{-3}$ and not as λ^{-4}, as in a classical optical fiber, which is in agreement with Raleigh law. Differently from classical optical fiber, in POF the loss dependency $\alpha \sim \lambda_{s(min)}$ is fulfilled for all mechanism of scattering, not only for Raleigh scattering on the particles much smaller than the optical wavelength.

What are the possibilities to decrease the losses caused by the CSW effect in a PCOF? The changes of standard technological parameters (temperatures and pulling rate) do not influence the CSW.

Diminishing of the CSW amplitude is possible by reduction of glass phase transition temperature T_g of fiber material. Doped glasses have smaller values of T_g but simultaneously the value of surface tension γ decreases. Surface tension of pure silica, given in the glass catalogues is γ=0,3J/m^2. This parameter, which is given in catalogues, however, strongly depends on the state of glass surface, and is measured in normal conditions. Inside the capillary, the surface tension is bigger than on the free surface and equals γ=0,7-1J/m^2. Theoretical value for quartz, calculated from bounding energy is γ=5,2J/m^2. Practically measured values are smaller, because of covering the surface with silanol groups Si-OH. During capillary fiber production, the contamination of surface with silanol groups is to be avoided. The groups are source of strong absorption losses. Lack of Si-OH groups increases the value of surface tension in glass, during the fiber production, even to the value γ=1,5J/m^2, favorably decreasing the effect of CSW. Surface tension in capillary during fiber pulling may be measured via the pressure opposing the collapse. Certain possibilities of surface tension increase give new methods of optical fiber glass processing.

Decrease in CSW may be reached via decrease of the fundamental mode field penetration in the glass. The smaller field penetration decreases intermodal coupling to lossy, unguided modes and surface modes of higher order. The field penetration may be decreased by increasing the capillary hole (optical core) diameter, for a constant wavelength of guided fundamental mode. However, modes of higher order may be also guided. Additionally, for bigger optical cores (capillary hole diameters) the fiber bending losses increase. Also, the spectral density of surface modes increase.

Increasing the thickness of intermediate area between hole and periodic photonic structure is of importance. This thickness may be increased relative to the usual dimensions of grain boundaries. This gives this layer a sort of anti-resonant properties for surface modes. The evanescent field of fundamental mode is reduced.

Numerical analyses of the PCOF show, that the hexagonal geometry of mikro capillary (optical core) and increasing in the rate of air/glass leads to the reduction of field penetration of fundamental mode in the glass.

The choice of longer wavelength decreases the level of wave dissipation, up to the moment of balance with IR absorption component. The absorption and Raleigh scattering components depend on the proportion of light propagated in air p_p. Assuming p_p=99,9%, these components, in a PCOF are scaled by the value of 1-p_p. The additive CSW component remains. The estimates give total attenuation of the order of α ~0,1dB/km for $\lambda_{s(min)}$ ~1,90μm.

The above value of total attenuation is estimated for a typical catalogue value of surface tension in silica glass γ=0,3J/m^2. Assuming, still quite realistic value of surface tension for γ=1,5J/m^2 (which is still several times smaller than theoretical one) one obtains the attenuation of PCOF of the order of α ~0,02dB/km. A big value of γ is not very favorable for PCOF manufacturing, because of strong and increasing tendency to collapse of the core, and associated with this difficulty in obtaining large air/glass factors. Thus, is seems that glasses of big refraction, small γ, large level of CSW, do not fit for manufacturing of low-loss capillary optical fibers.

The scattering caused by CSW is reversely proportional to the differential own refractive index. The surface roughness components of small value of space frequency f possess the biggest input in the scattering. It is favorable to increase core capillary diameter and the nano-capillaries just adjacent to the core. This effect is a cause of appearing of exotic projects of PCOF with very complex of capillary holes for the optical core and photonic structure in the optical cladding. There remains the problem of practical realization possibilities of such structures.

The scattering caused by intermodal coupling can be reduced by elimination of surface modes in PCOF. This problem is associated with optimization of geometrical structure of PCOF. Current calculations give frequently structures which can not be realized practically with existing technologies (too complex geometrical shapes).

The level of CSW decreases with the increase in guided wavelength. An open question is, similarly to classical optical fiber communication, switching to different wave bandwidths. It is associated with different materials for fibers, which are transparent and have ultra low losses for longer wavelengths.

5. CONCLUSIONS

Capillary optical fibers – COF [10], have an axial hole along the whole length of the glass fiber. The light is guided using two mechanisms: refractive and diffractive-interference (called photonic). The characteristic dimension, for internal refractive and geometrical structures, in such fibers is $\lambda/2$. The manufacturing material is glass (fully amorphous substance) in a continuous form and in a structured form of a photonic crystal. The properties of photonic crystal (and thus photonic crystal fiber) depend primarily on the internal structure rather than on the chemical composition.

The refractive capillary optical fibers possess a high-index, ring-like, optical core adjacent from one side to a capillary hole and optical cladding from the other. When the cladding has an index depression, then the fiber can have a nonzero cut-off frequency of the fundamental mode. The fiber fundamental mode is a dark hollow beam (DHB) of light. The DHB, for certain RCOF construction, can have a considerable gradient of the aperiodic, evanescent wave in the capillary hole. This gradient is used for sensing and other photonic, instrumental purposes, like guiding of nanoparticles inside the fiber. It is interesting, that not the absolute value of the field is important for sensing purposes but rather a large value of the decay gradient of the evanescent field. The application potential of refractive COFs is very large, especially in the field of such technologies like: MOEMS, lab on chip, lab on fiber, lab on capillary, etc.

The photonic capillary optical fibers [5] have also a considerable development potential, especially in the field of optical cable telecommunications industry. Current estimations, basing on intense research on such fibers, give the following data: loss for the wavelength of minimum losses – order of 0,1dB/km, broadband dispersion - above ten times smaller than in a classical optical fiber, level of nonlinear effects – for optical power above ten times higher, photonic transmission band – a few hundred of nm, etc. The basic confinement now for the application of these fibers (made of ultra pure silica glass), in the telecom systems, is the current value of losses, which are 1dB/km. These fibers also exhibit slightly longer wavelength of the minimum value of losses, now 1750nm, and in the future probably around 1900nm. Replacement of classical, single mode, dispersion shifted, 1550nm optimized optical fibers for communications with photonic capillaries would cause the next serious revolution in optical communications. The predictions say that such a revolution may happen within this decade.

6. ACKNOWLEDGMENT

This work was partly supported by Nat.Res.Prog. PBZ-MIN-009/T11/2003, Optoelectronic Components and Modules for Research and Industrial Applications coordinated by the Institute of Electronic Materials Technology, ITME, Warsaw and Center of Technology Transfer, CTT, Warsaw University of Technology.

7. REFERENCES

1. Polymicro Technologies, Capillary Optical Fibers: [www.polymicro.com] (2008.)
2. R. Romaniuk, J. Dorosz: *Design and fabrication of capillary optical fibers*, Technical Report (in Polish), Grant PBZ-MIN-009/T11/2003 (2003-2007), Warsaw University of Technology [www.ise.pw.edu.pl/ rrom/kapilary] (2007).
3. A. B. Sotsky, L. I. Sotskaya: *Modes of capillary optical fibers*, Optics Communications, vol.230, no. 1-3, pp.67-79 (2003).
4. V. Neves, F. S. C. Fernandes: *Modal characteristics of W-type and M-type dielectric profile fibres*, Microwave and Optics Technology Letters, vol.22, pp.355-357 (1998).
5. P. J. Roberts, F. Couny, H. Sabert, B. J. Mangan, D. P. Williams, L. Farr, M. W. Mason, A. Tomlinson, T. A. Birks, J. C. Knight, P. St. J. Russell; *Ultimate low-loss hollow core photonic crystal fibres*, Optics Express, 10 January 2005, vol.13, no.1, pp.236-244;

6. B. J. Mangan, L. Farr, A. Langford, P. J. Roberts, D. P. Williams, F. Couny, M. Lawman, M. Mason, S. Coupland, R. Flea, H. Sabert, T. A. Birks, J. C. Knight, P. St. J. Russell: "Low loss (1.7 dB/km) hollow core photonic bandgap fiber," in Proc. Opt. Fiber. Commun. Conf. (2004), paper PDP24.
7. J. A. West, Ch. M. Smith, N. F. Borelli, D. C. Allan, K. W. Koch: *Surface modes in air-core photonic band-gap fibers*, Optics Express, 19 April 2004, vol.12, no.8, pp.1485-1496.
8. K. Saitoh, N. A. Mortensen, M. Koshiba, "Air-core photonic band-gap fibers: the impact of surface modes," Opt. Express 12, 394-400 (2004),
9. D. C. Allan, N. F. Borrelli, M. T. Gallagher, D. Müller, C. M. Smith, N. Venkataraman, J. A. West, P. Zhang, K. W. Koch: "Surface modes and loss in air-core photonic bandgap fibers," in Photonic Crystal Materials and Devices, Ali Adibi, Axel Scherer, and Shawn Yu Lin;, eds. Proc. SPIE 5000, p. 161-174 (2003)
10. R. Romaniuk, *Capillary optical fiber – design, fabrication, characterization and application*, Bulletin of Polish Academy of Sciences, Technical Sciences, Vol.56, No.2., pp.87-102 (2008).
11. OmniGuide [omni-guide.com] (2008).
12. Blaze Photonics [blazephotonics.com] (2008).
13. Crystal Fibre A/S [crystal-fibre.com] (2008).

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 343–355

Physical and behavioural modelling of ion sensors*

JAN OGRODZKI

Institute of Electronic Systems,
Warsaw University of Technology,
Nowowiejska 15/19, 00-665 Warszawa
e-mail: J.Ogrodzki@ise.pw.edu.pl

Received 2008.06.04
Authorized 2008.08.30

This paper presents several approaches to modelling of potentiometric ion-selective sensors applicable in simulation of ion-sensor systems in environmental water measurements. Hardware-software co-design of systems like these requires models sufficiently simple and accurate for on-line simulation of a chemo-electrical measuring circuit which operates on the base of calibration-measurement or data-fusion methods. Sensor models proposed in this paper are dedicated to a combination of two-ions: one main and one interfering. To provide efficient data fusion for this two-ion case we propose only models limited to four parameters that shall be identified. The proposed models have been characterized by means of lab measurements and a complexity-accuracy trade-off has been discussed. The most accurate and yet moderately complex models are recommended for further application in the sensor-system design.

Keywords: water monitoring, potentiometric sensors, modelling, data fusion

1. INTRODUCTION

Monitoring of pollutants in human environment is vital for both safety as well as for quality of life. Due to progress of information technology and silicon-based micro-electronics it is basically possible to produce complex Microsystems that integrate both an array of chemical sensors as well as sophisticated raw information processing units. Such "smart sensors" can be superior to "raw sensors" in accuracy, linearity, lifetime and ease of use. When mass-produced – they can be also inexpensive. Yet the gain due to processing information from raw sensors can be envisioned only

---

[*] This paper was rapported by FP6 EU project WARMER

when there exist sufficiently accurate and reliable models of raw sensors. The model should express dependence of sensor outputs on important chemical factors. Chemical sensors are usually selective, i.e. sensitive only to some kinds of ions.

Ion selective electrodes known for over three decades [1- 6] are finding extensive practical applications especially in recent years. Sensors like these are built of polymeric ion-sensitive membranes operating as chemo-electrical transducers that convert ion concentration into an electrical signal. We may distinguish CHEMFET sensors, potentiometric sensors, and voltammetric sensors. CHEMFETs are chemical sensors based on Field Effect Transistors with ion-selective membranes. The potentiometric sensor is a zero current planar solid-state device operating as ions controlled voltage source. It has a ion-selective membrane similar to that of CHEMFET device though imposed on a passive substrate instead of an active FET device. This passive metallic electrode is made on a printed board or a silicone and then covered by an ion-selective membrane. The membrane contacts the extrinsic electrolyte subject to test. This paper is dedicated to electrical modeling of these sensors. Finally voltammetric sensors form a separate group of sensors that work in nonzero current conditions and their output signals may be of voltage, current or impedance type.

In this paper we present some results of modeling of potentiometric chemo-electrical transducers and discuss their practical verification. A goal of this modeling activity is to develop a simple but yet sufficiently accurate potentiometric sensor models that will be credible in further chemical data processing. Model simplicity means a small number of parameters to be determined on the data processing stage (calibration and measurement or data-fusion [7]) from fitting the model to sensor read-outs in a small number of steps and with a reduced consumption of measurement time and auxiliary liquids.

2. SO FAR KNOWN MODELS OF POTENTIOMETRIC SENSORS

In so far ion-sensors applications Nikolsky-Eisenmann (NE) model [8] has been ubiquitously used which is a simple semi-empirical model with known problems [9] yet useful especially for univalent ions. Chemists are used to apply this simple model despite of its several drawbacks. Its main drawback is that it is not invertible, i.e. for ions I_1 and I_2 NE model does not satisfy symmetry conditions consisting on equivalence of the formulae:

$$u = U_{012} + \frac{NV_T}{z_1} \ln(a_1 + K_{12} a_2^{z_1/z_2})$$

$$u = U_{021} + \frac{NV_T}{z_1} \ln(a_2 + K_{21} a_1^{z_2/z_1})$$

where parameters U_{021}, U_{012}, K_{21}, K_{12} satisfy:

$$U_{021} = U_{012} + \frac{NV_T}{z_1} \ln K_{12}, \ \ln K_{21} = -\frac{z_2}{z_1} \ln K_{12}$$

There are known other, more accurate models. The most complicated one was developed by Van den Berg [10] which is too difficult to be identified properly and its simplified variant called super-Nikolsky-Eisenmann (SNE) [11-16] which is equivalent to [18,19] of Nagele, Bakker, Pretsch though was elaborated quite independently. Important advantage of SNE model is that it satisfies the above mentioned symmetry conditions. For univalent ions of charge +1 or -1 models NE and SNE are identical. Both NE and SNE are unfortunately too inaccurate in the range of medium and small concentrations of main ions where influence of interfering ions can be observed. This range should be modelled more accurately.

NE and SNE models can be expressed by the following similar implicit formulae:

$$\sum_{i=1}^{N_{ions}} K_{1,i} a_i^{z_1/z_i} E^{z_1} = 1 \text{ (NE)} \tag{1}$$

$$\sum_{i=1}^{N_{ions}} K_{1,i}^{z_i/z_1} a_i E^{z_i} = 1 \text{ (SNE)} \tag{2}$$

where $E = \exp[-(u - U_0)/(NV_T)]$, $K_{11} = 1$, u is the sensor voltage response, V_T – a temperature dependent electro-thermal potential, N_{ions} – a number of ions involved which are all of one sign (all cations or all anions) and a_i, z_i ($i=1,\ldots,N_{ions}$) – activities and charges of the ions. Parameters of the models are as follows: U_0 – an offset voltage, N – a slope non-ideality coefficient and K_{1i} ($i=2,\ldots,N_{ions}$) – selectivity coefficients. Both NE and SNE have $N_{ions}+1$ parameters and are the simpliest and the least accurate among models which are discussed in this paper.

The models (1), (2) can be transformed into explicit forms. NE model for any ions gives:

$$u = U_0 + \frac{NV_T}{z_1}\ln\left(a_1 + \sum_{i=2}^{N_{ions}} K_{1i} a_i^{z_1/z_i}\right) \tag{3}$$

As for SNE model (2) it can be explicit analytically solved for u only in the case of valencies ±1, ±2, ±3 and ±4 [18]. For the most practical case of charges ±1 and ±2 equation (2) can be transformed into the form

$$u = U_0 + NV_T \operatorname{sgn}(z_1)\ln\left(A_1 + \sqrt{A_1^2 + 4A_2}\right) \tag{4}$$

where assuming N_1 ions of charges +1 (or -1 respectively) and N_2 ions of charges +2 (or -2 respectively), $N_1 + N_2 = N_{ions}$, we obtain:

$$A_1 = \begin{cases} \sum\limits_{i=1}^{N_1} K_{1,i} a_i, \; z_1 = \pm 1 \\ \sum\limits_{i=N_2+1}^{N_{ions}} K_{1,i}^{1/2} a_i, \; z_1 = \pm 2 \end{cases}, A_2 = \begin{cases} \sum\limits_{i=N_1+1}^{N_{ions}} K_{1,i}^{2} a_i, \; z_1 = \pm 1 \\ \sum\limits_{i=1}^{N_2} K_{1,i} a_i, \; z_1 = \pm 2 \end{cases}.$$

Asymptotically equivalent though slightly different from NE in a curvature of the knee is the model of Buck [1]

$$u = U_0 + N V_T \ln\left(a_1^{1/z_1} + \sum_{i=2}^{N_{ions}} K_{1i} a_i^{1/z_i}\right) \tag{5}$$

In this paper we propose several behavioural modifications of NE and SNE models. These modifications improve accuracy of modeling by some heuristic transformations. The modified models will be compared with NE and SNE references. The models to be compared will be all discussed in the case of $N_{ions} = 2$ because in WARMER we have only one pair of strongly interfering ions, typically K^+ and NH_4^+. The models will be of 4 parameters: U_0, N, K_{12} and one additional parameter improving accuracy of the model. In effect it will be possible to rank the models according to accuracy of fitting to sensors measurements obtained in lab and so to formulate some recommendation as to their use in CAD of chemo-electrical systems.

3. ADDITIONAL PHENOMENA TO BE TAKEN INTO ACCOUNT

From lab measurements of potentiometric sensors we can observe certain effects that have been neglected in physical description of Nernst, Nikolsky [8], Van den Berg [10], Nagele *et al* [18]. The first neglected effect consists of systematic inconsistence between ions concentration on the border of the membrane and in the tested electrolyte. This effect is caused by main ions collected by the ionophore and so taken from the electrolyte making a depletion layer on the phase border. In this layer ions activity is lower than in the solution bulk. Hence a measured concentration is slightly smaller than average concentration in the electrolyte. On the opposite side ligand complexes IL and free ions in the membrane are washed out from the solution and diluted in the electrolyte. These two dynamic phenomena are rather slow and so they are observed in practice as voltage drifts across the membrane. The second effect influencing the sensor selectivity is a not full dissociation of ion exchanging lipophyllic salt compounds. Hence concentration of lipophyllic salt ions is smaller than concentration of lipophyllic compounds themselves. The third effect observed is diffusion of ions inside the membrane. This diffusion decreases a voltage across the membrane. These effects are not taken into account in Van den Berg and Pretsch *et al* theory. As a result we observe vertical drift of characteristics, smaller than modeled curvature of the characteristics e of the

(5)

models.
nations.
els to be
we have
will be
y of the
f fitting
endation

,

n effects
en Berg
sistence
ctrolyte.
rom the
tivity is
smaller
mplexes
d in the
erved in
the sen-
pounds.
ophyllic
e mem-
are not
observe
cteristics

---

and an effect of slope of the characteristic in low concentration range. All mentioned effects are demonstrated in Fig. 1.

In this paper we propose certain behavioural modifications of known physical models that describe these second order effects with satisfactory accuracy.

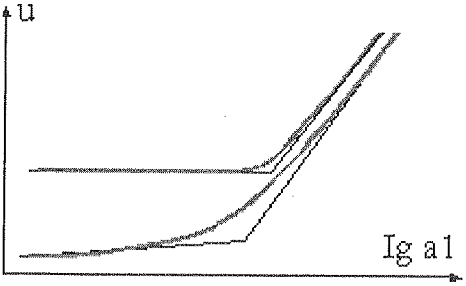


Fig. 1. Nonideal characteristics of chemical sensors. We observe three types of non-ideal effects: drift of the offset, a curvature weaker than in NE and SNE mods and a nonzero slope in saturation region

4. BEHAVIORAL MODELS FOR N_{IONS}=2

4.1. NIKOLSKY-EISENMANN MODEL WITH MODIFIED CURVATURE (NE2)

Measurements of sensor characteristics show that a true curvature of the knee of characteristic $u(a_1)$ is smaller than a curvature modeled by NE model. Hence to adjust the curvature an exponent w has been added:

$$u = U_0 + \frac{NV_T}{wz_1}\ln\left[a_1^w + \left(K_{12}a_2^{z_1/z_i}\right)^w\right] \tag{6}$$

Exponent w does not change left and right asymptotes of the model. We see that it has N_{ions}+2 = 4 parameters that have to be identified on the base of measurements: U_o, N, K_{12}, W

4.2. BUCK MODEL WITH MODIFIED CURVATURE (BUCK2)

Measurements of sensor characteristics show that a true curvature of the knee of characteristic $u(a_1)$ is smaller than a curvature modeled by BUCK model. Hence to adjust the curvature an exponent w has been added:

$$u = U_0 + \frac{NV_T}{w}\ln\left[a_1^{w/z_1} + \left(K_{12}a_2^{1/z_2}\right)^w\right] \tag{7}$$

Exponent w does not change left and right asymptotes of the model. We see that it has N_{ions}+2 = 4 parameters that have to be identified on the base of measurements.

4.3. SUPER NIKOLSKY-EISENMANN MODEL WITH A MODIFIED CURVATURE (SNE2)

Measurements of sensor characteristics show that a true curvature of the knee of characteristic $u(a_1)$ is smaller than a curvature modeled by NE model. Hence to adjust the curvature an exponent w has been added. For the case of $z_1 = 1$ and $z_2 = 2$ we obtain instead of (4) the following formula:

$$u = U_0 + NV_T \ln\left([a_1^{2w} + (4K_{12}^2 a_2)^w]^{1/2w} + a_1\right) \tag{8}$$

Complexity of model SNE2 (8) is increased by 1 parameter in comparison with SNE and so the model is of $N_{ions} + 2 = 4$ parameters which have to be identified.

4.4. NIKOLSKY-EISENMANN MODEL WITH A MODIFIED SATURATION SLOPE (NEZ)

In measurements of sensors we also observe an effect of positive slope of characteristic $u(a_1)$ in the range of small concentration of main ions saturation range (in logarithmic scale). This effect is not modeled by pure NE and SNE models. Moreover this slope decreases when concentration of interfering ions increases. An effect like this can be approximately modelled by making selectivity coefficient dependent on main and/or interfering ions. Hence the model (3) has to be enriched by a variable selectivity effect that yields two variants of models NEZ and NEZ 2.

$$u = U_o + \frac{NV_T}{z_1} \ln\left(a_1 + K_{12} a_1^{P_{12}} a_2^{z_1/z_2}\right) \tag{9}$$

Model NEZ has $2N_{\text{ions}} = 4$ parameters, its additional parameter P_{12} models a slope in the saturation region.

4.5. NIKOLSKY-EISENMANN MODEL WITH A MODIFIED INFLUENCE OF INTERFERING IONS (NEZ2)

In lab measurements we observe that discrepancy between characteristics in the range of low concentration of main ions for different levels of interfering ions is considerably different from characteristics of so far introduced models. Hence instead of a variable slope in (8) we may propose a variable influence of interfering ions not modelled neither by NE model nor by SNE model. An effect like this can be approximately modelled by an exponent in the interference describing term that yields the NEZ2 model:

$$u = U_o + \frac{NV_T}{z_1} \ln\left(a_1 + K_{12} a_2^{P_{12}}\right) \tag{10}$$

Model NEZ2 also has $2N_{\text{ions}} = 4$ parameters, but its additional parameter P_{12} models a strength of influence of interfering ions on the characteristics $u(a_1)$.

5. VERIFICATION OF THE PROPOSED MODELS BY LAB MEASUREMENTS

In this section the models introduced in Sections 2 and 4 are to be identified by means of an experiment where a parametric sensor of ammonia ions NH_4^+ worked in an electrolyte containing a fixed volume of an interfering solution including potassium ions K^+ or magnesium ions Mg^{+2} and subsequent additions of volumes of the main solution including ammonia ions. An experiment like this may give one branch of the sensor response $u(a_1 = a_{NH4}, a_2 = a_{INTERF})$ or two branches of the sensor response $u(a_1 = a_{NH4}, a_2 = a_{INTERF1,2})$. In these experiments a_{NH4} strongly increases and a_{INTERF} infinitesimally decreases due to a dilution effect. The measured response u is distorted by a systematic error caused by drift and a random error. In this experiment only ions of charge +1,+2 are taken into account. We discuss NE model (3), SNE model (4), BUCK model (5), NE2 model (6), BUCK2 model (7), SNE2 model (8), NEZ model (9), NEZ2 model (10). In the case of interfering ions I^{+1} models NE and SNE are equivalent and so one of them can be omitted.

In practical applications it is important to know: 1) how accurately the model approximates one branch of the characteristic $u(a_1, a_2\text{=const})$ and 2) how accurately is modelled influence of interfering ions, i.e. how accurately can be fitted two branches $u(a_1, a_2\text{=const1})$ and $u(a_1, a_2\text{=const2})$. In what follows these two cases will be discussed separately in chopters 5.1 and 5.2.

To verify the models a series of 10 measurements have been done per each branch of the characteristic. Measurements are performed in lab conditions using a semi-automatic measuring stand including a computer controlled micro-voltmeter and a precise manual biurete for dosage of necessary liquids. Measured results with interfering ions K^+ are shown in Fig. 2 (the one-branch case) and in Fig. 3 (the two-branch case) while with interfering ions Mg^{+2} are shown in Fig. 4 (the one-branch case) and in Fig. 5 (the two-branch case).

To fit the models MATLAB optimisation tools have been used to minimize absolute mean squared error $MSE = \frac{1}{n}\sum_{i=1}^{n}(u_i^{\text{measured}} - u_i^{\text{modeled}})^2$. Final accuracy of fitting is estimated by $RMSE = \sqrt{MSE}$ and a maximum error $ME = \max_i \left|u_i^{\text{measured}} - u_i^{\text{moodeled}}\right|$.

5.1. ONE-BRANCH VERIFICATION OF THE NH4 SENSOR WORKING WITH INTERFERING ION K$^+$

Numerical results of one-branch fitting for the discussed models are given in Table 1. Measured and fitted points of the characteristics have been plotted in Fig. 2.

Table 1

Absolute fitting accuracy in mV in the case of one branch and interfering ion K^+

	NE	NE2	BUCK2	NEZ	NEZ2
ME	0.86271	0.32114	0.32114	0.30745	0.86327
RMSE	0.4662	0.21322	0.21322	0.19961	0.46644

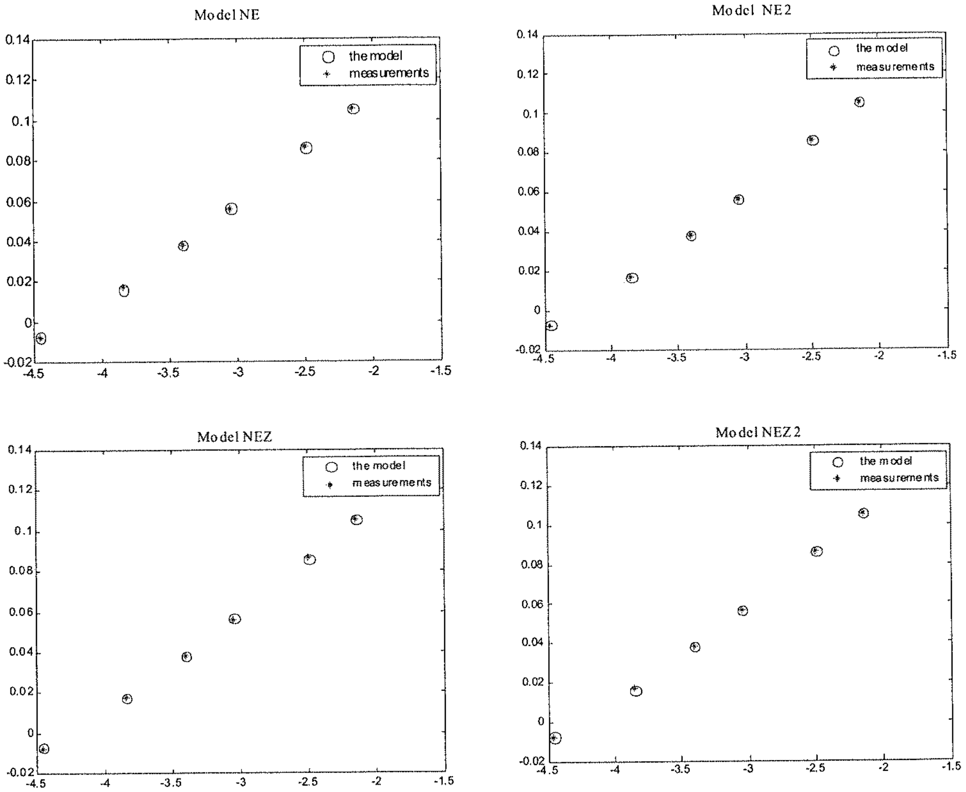


Fig. 2. Fitting of models NE, NE2, NEZ and NEZ2 (circles) to measurements (asterisks)

We can see that for all models involved ME and RMSE of fitting lay below 1mV. Models NE and NEZ2 are the worst, their RMSE appears to be about 0.46mV. The model NEZ is the most accurate with accuracy 0.19mV. The model NE2 is also acceptable with its accuracy 0.21mV. As we characterize an ammonia sensor its selectivity is very poor, even above 1, what means that the interfering potassium ion has even a little greater influence on the sensor than main ammonia ion.

5.2. TWO-BRANCH VERIFICATION OF THE NH4 SENSOR WORKING WITH INTERFERING ION K

Numerical results of two-branch fitting for the discussed models are given in Table 2. Measured and fitted points of the characteristics are plotted in Fig. 3.

Table 2

Absolute fitting accuracy in mV in the case of two-branches and interfering ion K^+

	NE	NE2	BUCK2	NEZ	NEZ2
ME	3.9279	3.9902	3.9902	3.8378	4.2198
RMSE	2.707	2.3317	2.3317	2.6091	2.6283

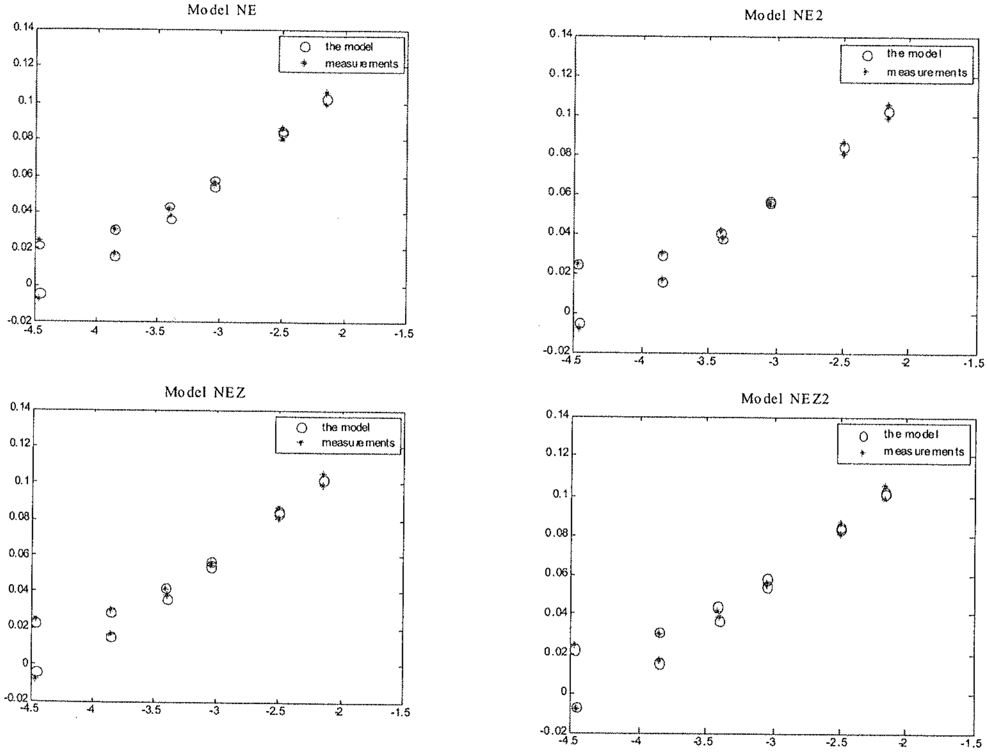


Fig. 3. Two-branch fitting of models NE, NE2, NEZ and NEZ2 (circles) to measurements (asterisks)

We can see that ME and RMSE of fitting lay below 4.2mV. Model NE is the worst, its RMSE appears to be about 2.7mV. The model NE2 is the most accurate with accuracy 2.3 mV. The model NEZ is of accuracy 2.6 mV. We observe that accuracy is not so good as for one-branch fitting.

5.3. ONE-BRANCH VERIFICATION OF THE NH4 SENSOR WORKING WITH INTERFERING ION MG^{+2}

Numerical results of one-branch fitting for the discussed models including SNE since the interfering ion is here of charge +2 are given in Table 3. Measured and fitted points of the characteristics have been plotted in Fig. 4.

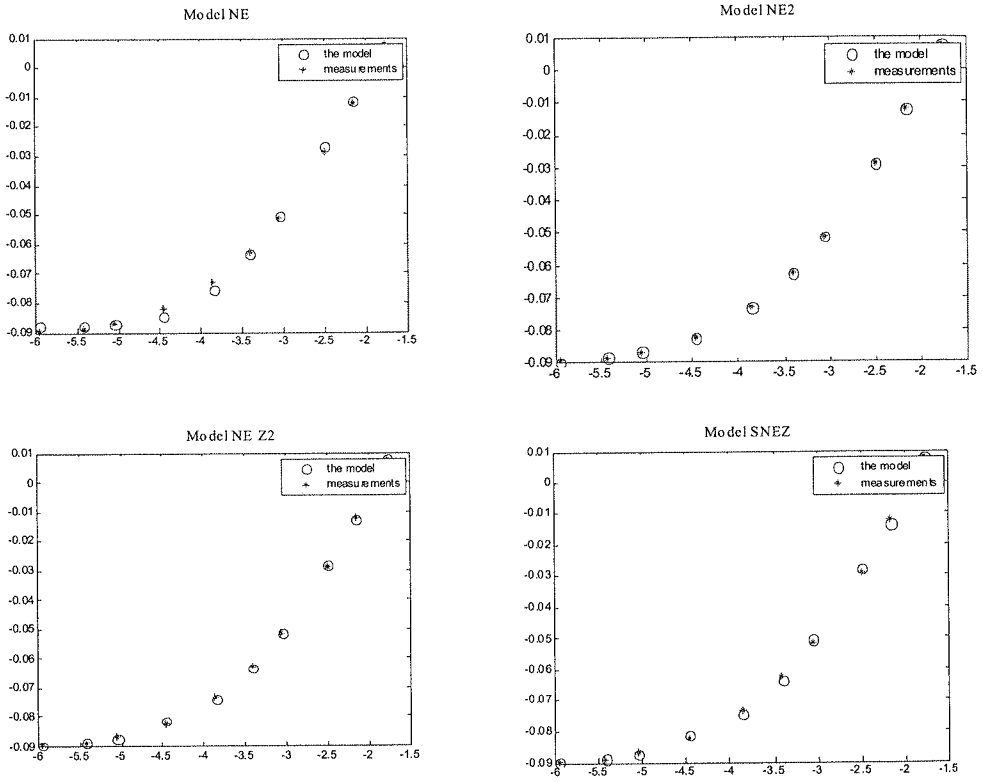
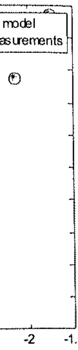


Fig. 4. One-branch fitting of models NE, NE2, NEZ2 and SNE2 (circles) to measurements (asterisks)

is the
te with
ccuracy

g SNE
d fitted



sterisks)

Table 3

Absolute fitting accuracy in mV in the case of one branch and interfering ion Mg^{+2}

	NE	NE2	BUCK2	NEZ	NEZ2	SNE	SNE2
ME	2.4631	0.37821	0.37821	1.1747	0.89645	3.876	1.6249
RMSE	1.16412	0.20528	0.20528	0.73843	0.51766	2.3538	0.91225

We can see that ME and RMSE of fitting lay below 1.6mV. Model SNE and NE are the worst, their RMSE appears to be about 2.3mV and 1.1mV respectively. However the model NE2 is the most accurate with accuracy 0.2mV. Models NEZ2, NEZ and SNE2 are also satisfactory with their accuracy 0.5mV, 0.7mV and 0.9mV respectively.

5.4. TWO-BRANCH VERIFICATION OF THE NH4 SENSOR WORKING WITH INTERFERING ION MG

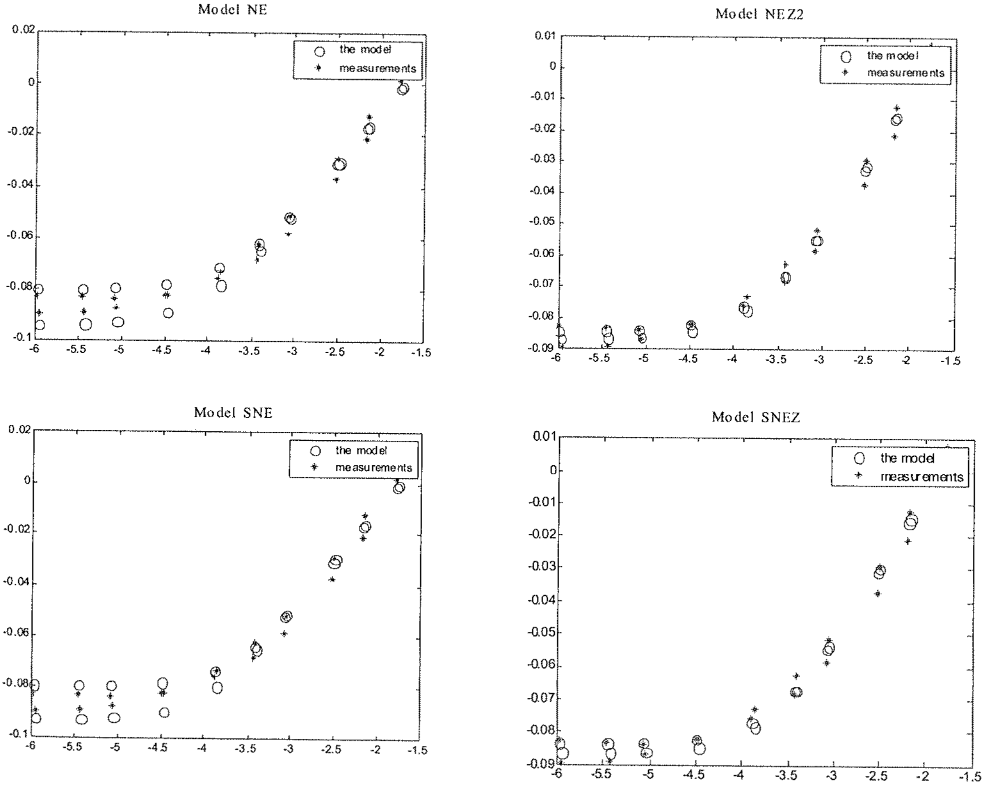


Fig. 5. Two-branch fitting of models NE, NEZ2, SNE and SNE2 (circles) to measurements (asterisks)

Numerical results of two-branch fitting for the discussed models including SNE and SNE2 are given in Table 4. Measured and fitted points of the characteristics have been plotted in Fig. 5.

Table 4

Absolute fitting accuracy in mV in the case of two-branch and interfering ion Mg^{+2}

	NE	NE2	BUCK2	NEZ	NEZ2	SNE	SNE2
ME	8.3156	8.3667	8.3667	8.3531	5.2692	8.3526	6.325
RMSE	4.7977	4.6951	4.6951	4.6852	2.8529	4.6322	3.1805

We can see that ME and RMSE of fitting lay below 8.3mV. Model SNE, NE, NE2 NEZ are the worst, their RMSE appears to be about 4.7mV. However the model NEZ2 is the most accurate with accuracy 2.8mV. Model SNE2 is also satisfactory.

6. CONCLUSIONS

In this section we state out several conclusions concerning the models discussed.

As for the two basic models NE and SNE the latter is theoretically more advantageous but in practice it appears less accurate though was developed for ions of charge ±2 to correct the model NE. We can state that so far known advanced description of chemical phenomena known from literature does not give us any constructive theory of operation of sensors like these. This fact is true because of existence of many parasitic effects like: a dynamic transfer of ions from and to the membrane and long time constants of all these mechanical and chemical phenomena. Because of this dynamics all second order effects are difficult to be modelled by means of algebraic formulae especially with a small number of parameters.

Measurements performed for electrolytes with ions of charge 1 have shown that RMSE of the models discussed was sufficiently small for chemical applications only in the case when a level of interfering ions is approximately constant (we work on one-branch of the characteristic). In this case RMSE is always smaller than 1mV and with models like NE2 or NEZ can be reduced to even 0.2mV.

In the case of variable level of interfering ions essential is modelling of interference. We state that there is no sufficient theory for modelling of this interference. If we work on two branches of characteristics, as in data fusion algorithms in woter probe design an error of the models increases in the best case to 2mV and can be even 3mV or 4mV. This considerably limits accuracy of testers making use of sensors like these and data fusion procedures.

Improvement of fitting accuracy may be obtained when random error of measurements is reduced by appropriate filtering of measured points. Unfortunately in data fusion this reduction is possible only by multiplication of a number of sensors.

Comparison of 8 models introduced in this paper: NE, NE2, BUCK, BUCK2, NEZ, NEZ2, SNE, SNE2 show that the most advantageous are models NEZ2 NEZ and NE2. They have one additional parameter and provide the smallest error in one-branch and two-branch cases. Among the parameters to be identified w of NE2, SNE2, NEZ2 are the best condifioned while P_{12} may be weaker conditioned.

7. REFERENCES

1. R. P. Buck, E. Lindner: *Tracing the History of Selective Ion Sensors*, Analytical Chemistry, vol. 73, no. 2, pp. 88-99, 2001.
2. E. Pretsch: *The New Wave of Potentiometri Ion Sensors*, Chimia, vol. 55, pp. 875-878, 2001.
3. H. Nam, G. S. Cha, T. D. Strong, J. Ha, J. H. Sim, R. W. Hower, et al: *Micropotentiometric Sensors*, Proceedings of the IEEE, vol. 91, no. 6, pp. 870-880, 2003.
4. F. Faridbod, M. R. Ganjali, R. Dinarvand, P. Norouzi: *The Fabrication of Potentiometric Membrane Sensors and their Applications*, Africal Journal of Biotechnology, vol. 6, pp. 2960-2987, 2007.
5. G. Horvai: *The Selectivity of Polymer Membrane Ion-selective Electrodes Obeying the Ion Exchanger Model*, Trends in Analytical Chemistry, vol. 16, no. 5, pp. 260-266, 1997.
6. Z. Brzózka, W. Wróblewski: *Chemical sensors*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, Poland, 1998 (in Polish).
7. L. J. Opalski: *Data Fusion for CHEMFET Arrays*, Bulletin of the Polish Academy of Sciences, no. 4, pp. 287-293, 2002.
8. P. Z. Nikolsky: *Physical Chemistry*, vol. 10, p. 495, 1937.
9. Y. Umezawa, K. Umezawa, H. Sato: *Selectivity Coefficients for Ion-selective Electrodes: Recommended Methods for Reporting KpotA,B Values*, Pure and Applied Chemistry, vol. 67, no. 3, pp. 507-518, 1995.
10. A. Van den Berg: *Ion Sensors Based on ISFETs with Synthetic Ionophors*, Ph. D. Thesis, University of Twente, The Netherlands, 1988.
11. J. Ogrodzki, L. J. Opalski: *Modeling of semiconductor ion sensors for CAD*, Proc. of MIXDES, Gdynia, Poland, 253-257, 2000.
12. J. Ogrodzki: *Chemfet Sensors Modeling for Water Monitoring Systems Design*, 9th Int. Workshop on Mixed Design of Integrated Circuits and Systems, Wrocław, Poland, pp. 151-156, June 2002.
13. J. Ogrodzki: *Physical Modeling of Ion-Selective Devices*, Int. Conf. on Signals and Electronic Systems, Świeradów Zdrój, Poland, pp. 269-274, Sept 2002.
14. J. Ogrodzki, L. J. Opalski, K. Zamłyński: *Modeling of CHEMFET sensors in SPICE*, Proc. of ECCTD, Kraków, Poland, Vol. II, 209-212, 2003.
15. J. Ogrodzki, W. Wróblewski: *Computer Modeling of CHEMFET Sensors for Data Fusion in Environmental Water Monitoring Systems*, Proc. AutoMoNet, Vien, Austria, 2004.
16. J. Ogrodzki: *Modeling of Potentiometric Sensors Sensitive to Ions of Valency One and Two for Possible Applications in WARMER Project*, SPIE Conference, Wilga, Poland, 2007.
17. P. Z. Wieczorek, L. J. Opalski, J. Ogrodzki: *Electrical properties of potentiomentric sensors – an empirical study*, Proc. SPIE Conference, Wilga, Poland, 2007.
18. M. Nagele, E. Bakker, E. Pretsch: *General Description of the Simultaneous Response of Potentiometric Ionophore-Based Sensors to Ions of Different Charge*, Analytical Chemistry, vol. 71, pp. 1041-48, 1999.
19. E. Bakker, E. Pretsch, P. Buhlmann: *Selectivity of Potentiometric Ion Sensors*, Analytical Chemistry, vol. 72, pp. 1127-1133, 2000.

c
p
li
fo
le

K

DS
transmi
while r
that the
to tran
the fina
Anothe
area of
a pulse
Analys
which r
reasons

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 357–366

TDR-based examination of DSL transmission line

KATARZYNA OPALSKA, ALEKSANDER BURD

Institute of Electronic Systems,
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warszawa, Poland
e-mail: K.Opalska@elka.pw.edu.pl

Received 2008.06.03
Authorized 2008.08.03

The paper presents results of TDR-based DSL line characterization. Measurement and computer analysis results are included for two distinct modes of TDR system: step and pulse. Measurement results emphasize the influence of a skin effect, which significantly limits measurement system resolution and accuracy. Results of numerous simulations allow for estimation of the dependence of reflected signal parameters on testing pulse width and/or length of line being examined.

Keywords: Time-domain reflectometry (TDR), DSL, transmission line, skin effect, time domain simulation, SPECTRE simulator

1. INTRODUCTION

DSL (Digital Subscriber Line) is a family of technologies that provide digital data transmission over the wires of a local telephone network. DSL uses high frequency, while regular telephone uses low frequency on the same telephone line. It is obvious, that the quality and the distance of telephone line (which was not originally intended to transmit anything else but sound) between the central office or the repeater and the final destination determine – so important nowadays – available transmission rate. Another important factor is the ease and time of faulty lines diagnosis – and this is the area of application of time domain reflectometry (TDR) technique [1]. In TDR system a pulse is applied to the tested line, its reflection (if any) is detected and registered. Analysis of the reflected signal brings information about variations in line impedance, which may occur due to open circuits, shorts, parasitic bridge tap, loading coil or other reasons.

The ‘last mile’ – part of the line delivering the signal directly to our houses – is most exposed to damage (due to environmental reasons, human factor and also the inevitable process of ageing – a lot of telephone lines is over 50 years old). This creates a demand on portable line analyzers suitable for field measurements.

TDR module for such an analyzer has been designed and constructed [2]. In this paper we present TDR-based testing of a typical, not so modern, telephone line. We concentrate on illustration of the importance of an appropriate selection of the testing pulse width, as this parameter – in combination with inevitable skin effect in the line – significantly limits measurement range and its accuracy. Numerous computer simulations – backing up measurements – allow for estimation of the dependence of reflected signal parameters on testing pulse width and/or length of line being examined and for pointing out the limited usefulness of copper wire for wide-band data transmission.

2. TDR-BASED MEASUREMENTS

There are two distinct operation principles of TDR units: step and pulse TDRs [3]. Both approaches have advantages and drawbacks. In – more popular nowadays – pulse devices (e.g. [4]), a narrow pulse (with variable width to be adjusted by the operator with respect to the tested line length) is transmitted into a line and only after that the receiver starts listening for reflections. This results in so called ‘dead zone’ or ‘blind spot’ corresponding to the width of transmitted pulse. Shorter pulse reduces dead zone, but unfortunately it also limits TDRs range because of line losses [5]. However this approach allows for relatively easy signal interpretation, as each distortion from zero level (neglecting noise) corresponds to line impedance variation causing the pulse to reflect (see Fig. 2 for an illustration). In the second mode (e.g. [6]) step pulse is sent to the line while the receiver simultaneously listens for returned signals. This not only eliminates dead zone, but also improves the signal to noise ratio due to higher energy transmitted. On the other hand it makes the signal more difficult for interpretation (see Fig. 3) which may be a serious obstacle for fault diagnosis. Figures 2, 3 illustrate the difference in output of both types of TDR devices for the case of an open-ended transmission line consisting of 3 sections, as depicted in Fig. 1.

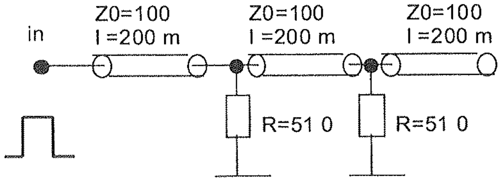


Fig. 1. Sample transmission line circuit with parastic resistances

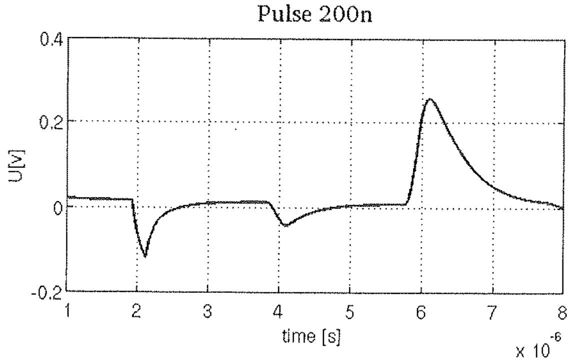


Fig. 2. Reflection of 200 ns pulse from 600 m long transmission line. Negative peaks correspond to parasitic resistances after each of 200 m section of the line. Final postive peak depicts reflection from the open end of the line

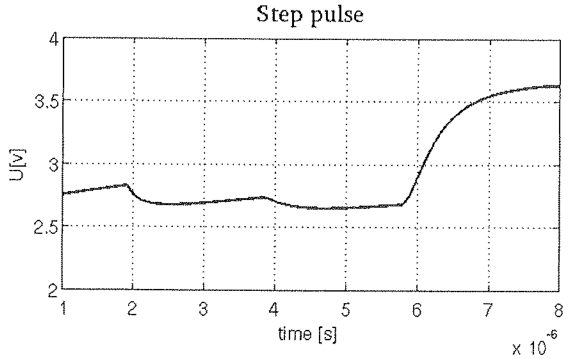


Fig. 3. Line from Fig. 1 with step pulse applied

3. CABLE CHARACTERIZATION

This section presents results of simulation/measurement-based characterization of a telephone cable. The line tested in laboratory consisted of 20 pairs of copper wires with 0.5 mm diameter, each of them 300 m length. Real-life telephone line consisting of the same type of wires was also examined.

First simple measurements in dedicated circuits brought basic results: DC resistance of copper wire was evaluated for 85 mΩ/m, wave impedance Z0 for 100 Ω, wave velocity for 0.21 m/ns.

TDR-based measurement relied on applying a symmetrical pulse (with ±5V amplitude) into a line under test, and visualization of the reflected signal on the generic-purpose oscilloscope or registration it by the designed TDR unit. Analysis of the signal (leading to line diagnosis) is left to system operator (usually – on experienced technician).

Numerous computer analysis of the transmission lines were performed using Spectre simulator [8] (obviously after verification of the simulation reliability). Simulations not only enhance user imagination in the domain of results interpretation, but also allow for estimation of some dependencies of TDRs output signal on parameters of the testing pulse.

Lines of several lengths were tested (in laboratory conditions) – starting from 300 m, up to 1.5 km, the longer lines obtained by sequential connection of several 300 m wires. Lines were tested with 2 kinds of pulse: step-like pulse (long pulse lasting approx. 200 μs) and narrow pulse of adjusted width (minimum 10 ns, useful range in order of hundreds nanoseconds/single microseconds). Rise time of testing pulse was in order of 2 ns. Figures 4 and 5 illustrate sample results obtained for step pulse reflected from open-ended 300 m line – measured and simulated (for simulation the waveforms obtained for longer lines are also depicted). Figures 6, 7 illustrate alike pair of the measured/simulated results obtained for the same line, but tested by a pulse with 1 μs width, while Figs. 8, 9 shorted 600 m line tested by a narrow pulse (20 ns width).

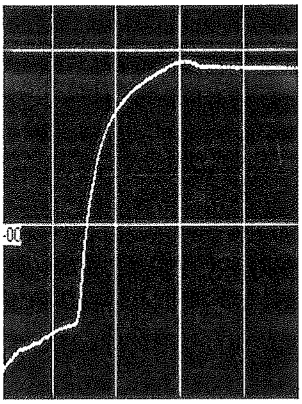


Fig. 4. Reflection of the step pulse from the open end of 300 m line (measured by the designed TDR unit). Time scale: 2 μ/div, vertical scale: 1V/div. Rise time of the respective signal part – approx 750 ns

Finally let us present measurement results for real line telephone line, opened at the end, with unknown (however not exceeding 1.5 km) length – see Figs. 10, 11. Two reflections were registered: one occurring at distance approximately 375 m from the source (local telephone office in the University), the second positioned at approximately 1300 m. The first reflection was probably caused by a loading coil inserted into a line, while the second one denotes lines (open) end.

Experiments (illustrated in Figs. 4, 7, 9, 10, 11 as wellas a great variety of others, not presented here) exposed the significant signal dispersion occurring in the copper wire. Rise time of testing pulse is not larger than 2 ns, while rise time of reflected signals slope counts in hundreds of ns (in 300 m line) or in microseconds (for lines exceeding ca 500 m). Unfortunately the dispersion cannot be avoided nor minimized,

Spec-
lations
ut also
ters of

from
several
pulse
useful
testing
or step
ulation
e alike
a pulse
(20 ns

d TDR
750 ns

ened at
1. Two
om the
imately
a line,

others,
copper
eflected
or lines
imized,

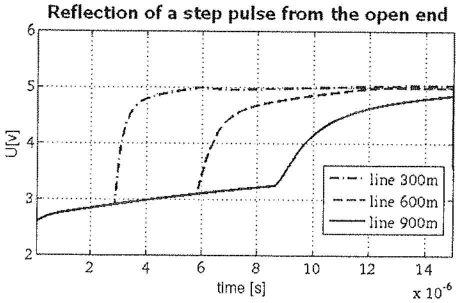


Fig. 5. Simulated reflection of the step pulse from the open end of line with variable length. Rise time for 300 m line approximately 800 ns

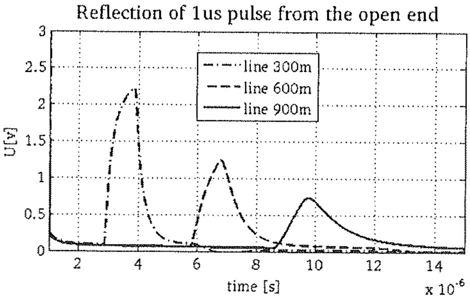


Fig. 6. Simulated reflection of 1 μs pulse from the open end of lines with variable length

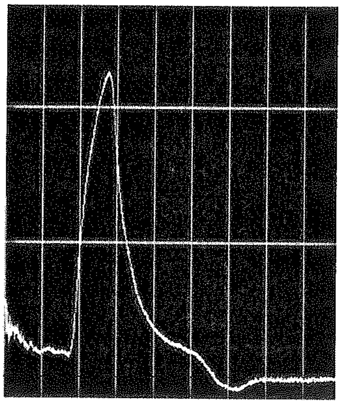


Fig. 7. Reflection of 1 μs pulse from the open end of 300 m line (measured by TDR module). Time scale: 1 μs/div, vertical scale: 1V/div. Peak voltage 2.2 V

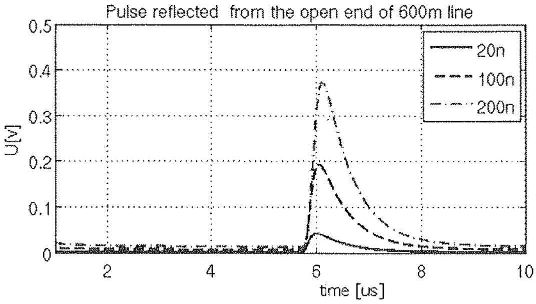


Fig. 8. Simulated reflections of narrow pulse with variable width (as described in the legend) from the open end of 600 m line

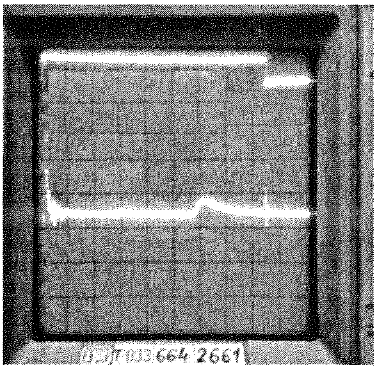


Fig. 9. Reflection of the positive narrow pulse (width 20 ns) from the open end of 600 m line (measured by the generic-purpose oscilloscope). Time scale: 1 μs/div, vertical scale: 50 mV/div

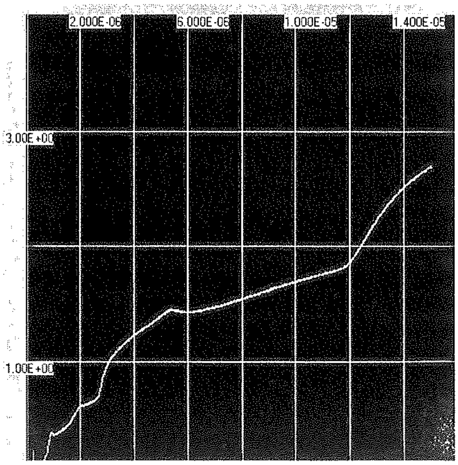


Fig. 10. Reflection of a step pulse from the open end of a telephone wire

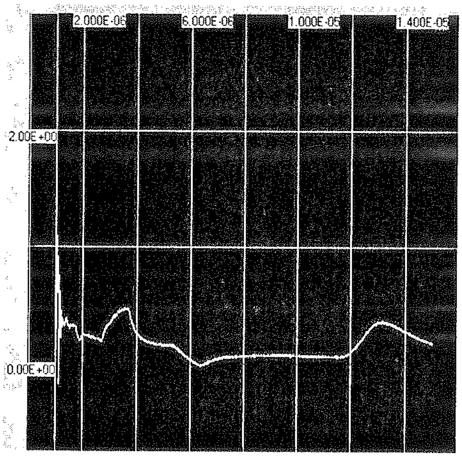


Fig. 11. As in Fig. 10, but testing pulse 1 μs width

as it is caused by the inherent feature of the cable itself – namely the skin effect [7] (frequency-dependent ohmic losses changing with cable geometry). In pulse TDR skin effect reduces amplitude of the reflected signal (obviously with accordance to the testing pulse width). In step TDR skin effect significantly deteriorates signal shape (slowering slopes) thus impeding signal interpretation. To make latler caseyet more complicated, the output of step TDR is influenced by DC cable resistance (responsible for linear rising of the signal between reflections).

Figure 12 illustrates the influence of two kinds of losses on step TDR output: serial resistance of the copper wire and skin-originated frequency dependent parasitic resistance. For comparison Fig.13 presents the corresponding signal measured by generic-purpose oscilloscope connected to copper wire with both types of losses evidently present, nuther of them negligible.

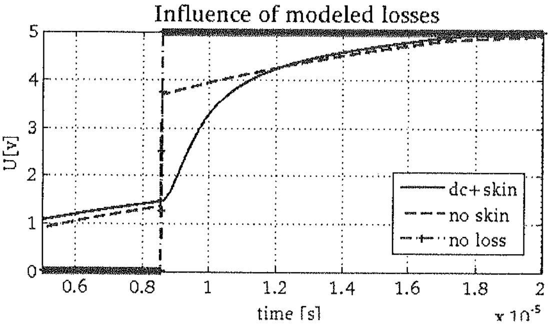


Fig. 12. Simulation with different level of transmission line model: reflection of the positive step pulse from the open end of 900 m long line

Analysis of the numerous computer simulated waveforms obtained for different variants of circuit configuration (variable line length and testing pulse width) enabled us

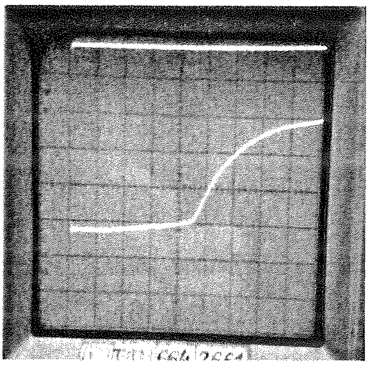


Fig. 13. Reflection of the step-like pulse from the open end of 900 m long line, time scale: 2 μs/div (slopes rise time in order of 5 μs). At the top of the screen – signal triggering the oscilloscope

to evaluate dependence of registered signal parameters (the peak value of the reflected pulse for pulse TDR and rise/fall time of the signal for step TDR) on transmission lines length. The characteristics – besides being help – ful for TDR system operator – illustrate limited usefulness of copper wire for wide – band data transmission, over relatively long distances. Sample results are presented in Figs. 14-16. Figures 14-15 ilustrate determination in output TDR signal with the increase of lines length.

It appeared, that – for interesting range of lines length ('last mile') – peak value of the reflection decreases exponentially (for fixed duration of the testing pulse), while rise time of the signal changes in a quasi-linear way.

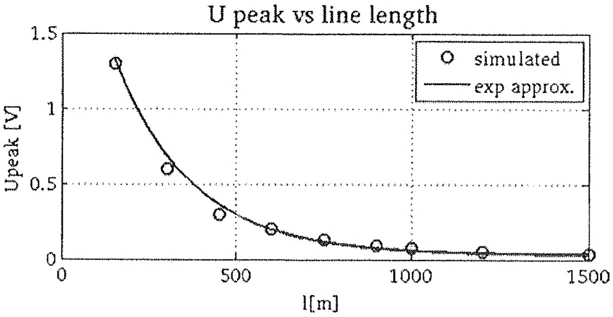


Fig. 14. Peak value of the reflected pulse versus lines lenght for tested pulse with 100 ns width (values for lines longer than 1500 m were unmeasurable). Solid line presents an exponential approximation providing the fit with error smaller than 20%

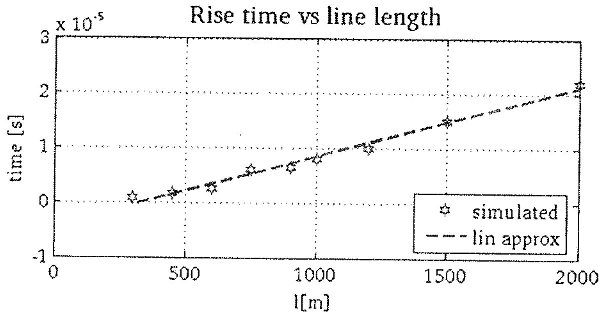


Fig. 15. Rise time of the respective part of signal for lines of variable length tested by a step pulse. Dashed line presents a linear approximation providing the fit with error smaller than 20%

Figure 16 illustrates the dependence of peak value of output (reflected) signal on testing pulse width for sample 600 m long line. This characteristic is very good approximated with an exponential function of a type $1-e\frac{-t}{\tau}$. Let us note also, that for testing pulses shorter than approximately 1 μs, linear approximation seems sufficiently accurate, as it produces an error less than 15%.

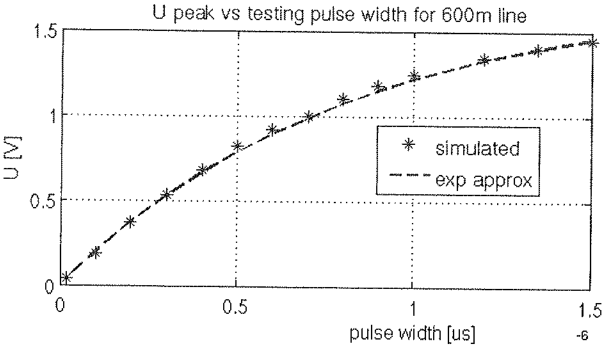


Fig. 16. Peak value of the reflected pulse versus width of a testing pulse. Dashed line presents an exponential approximation providing the fit with error smaller than 10%

4. CONCLUSIONS

We presented TDR-based testing of DSL transmission line backed by results of numerous computer simulation. We concentrated on – underestimated in our opinion – influence of the skin effect on the quality of the pulse transmitted along the relatively long transmission line. Skin effect depends on the material and geometry of the transmission line (telephone cable), so it cannot be avoided in any way. One can only learn how to deal with it and the paper aims in helping the reader to do it.

5. REFERENCES

1. I. J. Strickland: *Time-domain reflectometry measurements*, Textronix Inc. (1970)
2. A. Burd, T. Owczarek, K. Opalska: *TDR module for DSL line analyzer*, Technical report, ISE PW, 2008.
3. I. P. DeWinter, B. Ashley: *AN201 – Step vs Pulse TDR Technology*, AEA Technology Inc,www.aeatechnology.com, 2005.
4. xDSL Cable Qualifier CableSHARKR P3, operating manual
5. TDR Functionality on the ALT 2000, Trend Communications , ALT2000 Application Note No 2.
6. S. Misiaszek, A. Burd, K. Opalska, M. Radtke, M. Ramotowski, T. Starecki: *Time domain reflectometry educational tool kit for student laboratory*, Proc. ICSES, Wrocław, Poland, pp. 195-200 (2002).
7. M. Ramotowski, K. Opalska: *Evaluation of skin-effect distortions in delay lines with PSPICE models*, Proc. of SPIE, 6159, pp. 61592W-1 - 61592W-4 (2005).

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 367–375

Time-domain reflectometry module for DSL analyzer

KATARZYNA OPALSKA, ALEKSANDER BURD, TOMASZ OWCZAREK

Institute of Electronic Systems,
Warsaw University of Technology
Nowowiejska 15/19, 00-665 Warszawa, Poland
e-mail: K.Opalska@elka.pw.edu.pl

Received 2008.06.04
Authorized 2008.08.25

The paper presents a time domain relectometry module designed as a part of an analyzer of DSL transmission lines. Small size, programmable TDR module (to be included in the hand-held, battery-operated line analyzer) was designed, constructed and tested in the laboratory, proving its usefulness for detecting lines imperfections. The paper illustrates the principle of operation of the TDR unit, its basic building blocks and the software environment created for device testing.

Keywords: Time-domain reflectometry (TDR), DSL transmission line, pulse generation and registration, stroboscope sampling, microprocessor systems

1. INTRODUCTION

Time-Domain Reflectometry (TDR) is the well-known [1], simple and relatively inexpensive [2] method suitable for examination of transmission lines - especially for detecting discontinuities. The "radar-like" TDR technique relies on sending an electrical pulse into the tested line [3] and detecting, registration and analysis of the reflected pulse. TDR-based measurements are direct and prone to automation, the system is simple and relatively inexpensive – that are main advantages of the technique. On the other hand TDR has some drawbacks with the possibility of ambiguous results interpretations being the most serious one. But even if the analysis of the reflected waveform in not an easy task, it still allows for line diagnosis by pointing out the variations in line impedance, including shorts and open circuits. These - obviously unwanted – disturbances, affect the shape of the reflected signal which may be interpreted in terms of the location (marked by time position of the signal distortion) and kind (pointed out by the shape of the waveform) of line damage.

Presented in this paper TDR unit is designed as a part of a portable, battery-operated DSL analyzer equipped with dedicated GUI and controlled by a powerful microprocessor. The primary application of the analyzer is examination of the telephone wires – many of them being the simply copper, multi-pair coaxial cables – which are used nowadays also for transmission of digital signals with large - and still rising – bandwidth.

The assumptions for the designed TDR module have been formulated as follows:

1. the device allows for testing a line up to 10 km long with the resolution not larger than 5 m (the smaller the better),
2. line examination relies on delivering a symmetrical pulse to the line under test and registration of the reflected signal,
3. the results of the measurement are produced in the form of binary data containing digitized reflected signal,
4. device is controlled (via RS interface) by the main analyzer processor, which starts measurement, retrieves results and takes care for its interpretation.;
5. the last, but not least – the device is supposed to be a part of a portable line analyzer, so it should be small, spare in power consumption and – obviously – cheap.

In what follows we start with brief summary of basics of TDR measurements (sec. 2). Later on device structure and main technical specification are described (sec. 3). Section 4 presents software environment for device testing together with some measurement results.

2. BASICS OF TDR

At first let us remind basic idea of TDR measurements. Electrical pulse is applied to the examined line and reflected signal is registered. Time of arriving of reflected pulse determines distance of a discontinuity, which caused the reflection, from the source (so it might be convenient to scale CRT × axis in length units), while shape of the waveform depends on the type of a discontinuity. Discontinuities are clearly separated in time, so many of them may be located in a single measurement. Reflection (observed after $t=2l/v$, where l – line length, v – waves velocity) from the resistive line load results in a change of a pulse amplitude by the factor $\rho = (RL - Z0)/(RL + Z0)$, where RL – load resistance and Z0 – line characteristic impedance (see Fig. 1 for an illustration). Obviously for an uniform line terminated with a matched impedance no reflection occurs and so no change can be registered by TDR receiver. For short circuits RL=0 and applied pulse vanishes, while for an open end (RL=∞) pulses amplitude doubles.

Let us note, that in Fig.1 an unrealistic ideal pulse is assumed (with zero rise time). Finite rise time (determined by the maximum slope at the point of discontinuity) imposes significant - and unfortunately unavoidable – limitation on system resolution. Let us add also, that for reactive loads, the observed waveform has more complicated

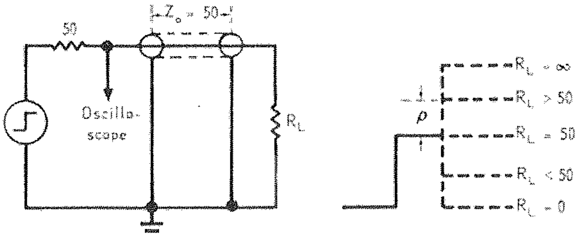


Fig. 1. Normalized (unity gain, zero rise time, positive going) pulse reflected from terminating resistance

shape depending upon character of the parasitic (capacitive or inductive) with slopes determined by the time constant formed by the load and Z0.

There are two distinct operation principles of TDR units: step and pulse TDRs [4], depending on the width of the testing pulse applied to the line. Both of them have their advantages and drawbacks. Pulse TDRs have a 'dead zone' ('blind spot') corresponding to the width of transmitted pulse. In the step mode (illustrated above) the receiver starts listening for returned signals with the starting slope of a testing pulse, which not only eliminates dead zone, but also improves the signal to noise ratio due to higher energy transmitted. On the other hand results obtained in pulse mode are easier for interpretation, which probably explains its frequent use in recent devices.

3. STRUCTURE OF TDR MODULE

Main computer system of the analyzer (referred in what follows as MASTER) is in charge of setting measurement conditions (start and final measurement time, time step, pulse length etc.), measurement triggering and processing its results (including signal visualization and its interpretation up to evaluation of tested line quality).

TDR module contains an internal microprocessor controlling the measurement process and data transmission to and from MASTER system via RS interface. A dedicated transmission protocol was adopted in order to ensure data consistence, minimum transmission time and error handling.

Measurement relies on applying a symmetrical pulse to the line under test, then sampling the line in appropriate intervals and saving the result in a memory (up to 128 kB, enough to store 64 ksamples, 2 bytes each). Minimum (and default) time interval between samples of registered signal is assumed for 5 ns (time needed for a waveform to travel to and fro along 0.5 m copper wire), longer intervals may be required by MASTER system. Sampling is performed by stroboscope technique (chosen mostly because of the demand of keeping the cost as low as possible), so an important task for modules microprocessor is time positioning of input signal and controlling the sequence of gathered samples.

Both TDR operation modes mentioned in sec. 2 (pulse and step) are available in the device. Mode selection (and pulse width if necessary) is left to system operator.

A structure of the designed TDR module is shown in Fig. 2. Further part of this section contains brief description of basic building blocks of the device.

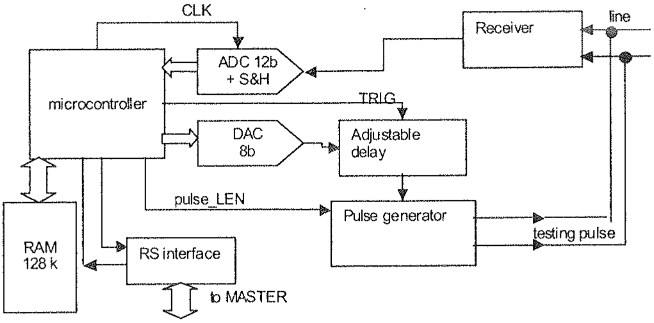


Fig. 2. Block structure of TDR system

Let us add here, that all circuits constituting TDR module are contained on the board 14 cm × 9 cm with some free space left for necessary connectors – which makes the device useful for hand-held DSL analyzer.

3.1. DIGITAL PART OF TDR SYSTEM

The main component of the digital part is the microprocessor controlling all phases of the measurements. Microprocessor C8051F066 (Silicon Laboratories) was chosen because of variety of its advantages, namely: rich functionality due to built-in components (including two DA and single AD converter and variety of counters); numerous pins providing direct access to controlling signals; fast transmission between processor and an external memory; low jitter and noise level of microprocessor clock (determining the accuracy of positioning samples on the time axis), and moderate price. Microprocessor controls both the exact time of applying test pulse into the line, as well as moments of line sampling. Time of pulse generation – very important because of an adopted stroboscope measurement technique – is adjusted in two ways: roughly by the moment of delivering TRIG signal (see Fig. 2) and smoothly by analogue block of delay generator (Fig. 5) controlled via built-in DA converter.

The second task of the processor – samples gathering – is realized by a 12-bit external AD converter (AD 9237 with 20 MS/s sampling rate, analogue bandwidth 500 MHz, including in its structure fast S&H circuit). The solution with the fast, external AD converter (neglecting the internal one) was chosen mostly because of S&H circuit, difficult to obtain in other way (an alternative is a home-made circuit, which however occupies more of the precious space on the printed board). Clock signal for ADC (specified for at least 1 MHz) is generated by PCA circuit built in the microprocessor structure.

Other functions of microprocessor system include control over storing the gathered samples (keeping their appropriate sequence!) in an external RAM memory and handling exchange of data between TDR unit and MASTER system.

3.2. ANALOGUE PART OF TDR SYSTEM

Analogue part of TDR device consists of the following blocks:

- generator of the symmetrical pulse with 5 V amplitude and adjustable width,
- receiver of the symmetrical line,
- digitally controlled delay generator,
- stabilizers delivering untypical supply voltages.

Block diagram of a symmetrical transmitter is presented in Fig. 3. Process of pulse generation starts after receiving **trig** signal from the delay generator (marked as '**out**' in Fig. 5), which in turn is triggered by a microprocessor. Basic transmitter function is realized by a pair of AC gates together with resistors setting output resistance for the required value of 100 Ω. Gates are controlled by a differential amplifier, which assures their backward symmetry. Input of the differential amplifier may be of two kinds: electrically conditioned pulse delivered from the microprocessor, or the shortest pulse available in the system (approximately 10 ns), triggered by a processor, but generated inside the transmitter block. Selection of pulse width depends on the **pulse_LEN** input (set by microprocessor), controlling the analogue key in the generator.

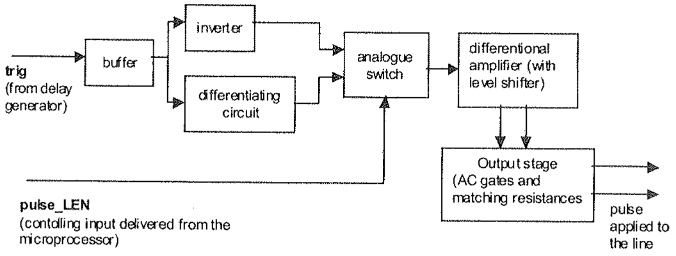


Fig. 3. Block diagram of the symmetrical transmiter

Structure of the receiver is shown in Fig. 4. The main function of the receiver is realized by a differential amplifier (BJTs with extra resistors in emitter to extend transition zone). Signal buffers (FET-based followers) and necessary level shifters make the circuit complete.

Delay generator (see Fig. 5) consists of two basic blocks: fast ramp generator and a comparator with hysteresis. **start** signal (delivered from the microprocessor) opens the switch (turns off initially saturated BJT key) and precise current source starts charging the forming capacitance. Voltage on the capacitance increases until it reaches the value – set by microprocessor via DA converter – corresponding to the required delay (with maximum delay limited by the electrical circuit to approximately 200 ns).

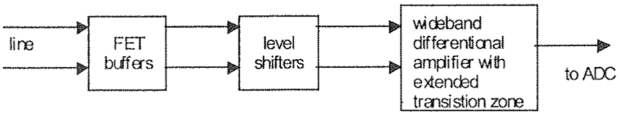


Fig. 4. Block diagram of line receiver

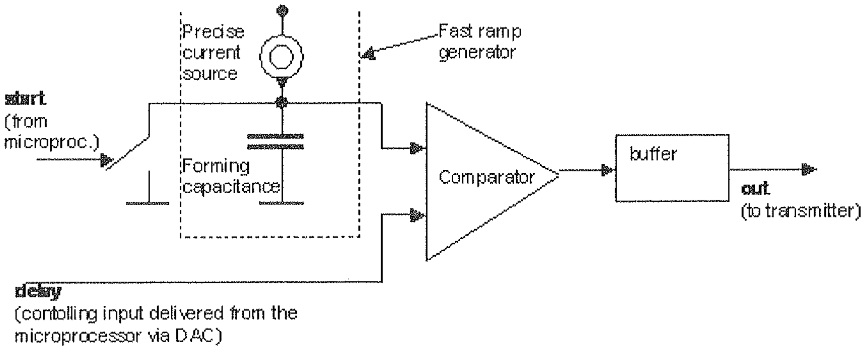


Fig. 5. Block diagram of delay generator

Comparator output (after necessary electrical conditioning) triggers the transmitter form Fig. 3 to start pulse generation.

All analogue blocks are placed on two mechanically tied boards with approximate dimensions 4.5×8 cm. Total parts cost does not exceed 15$.

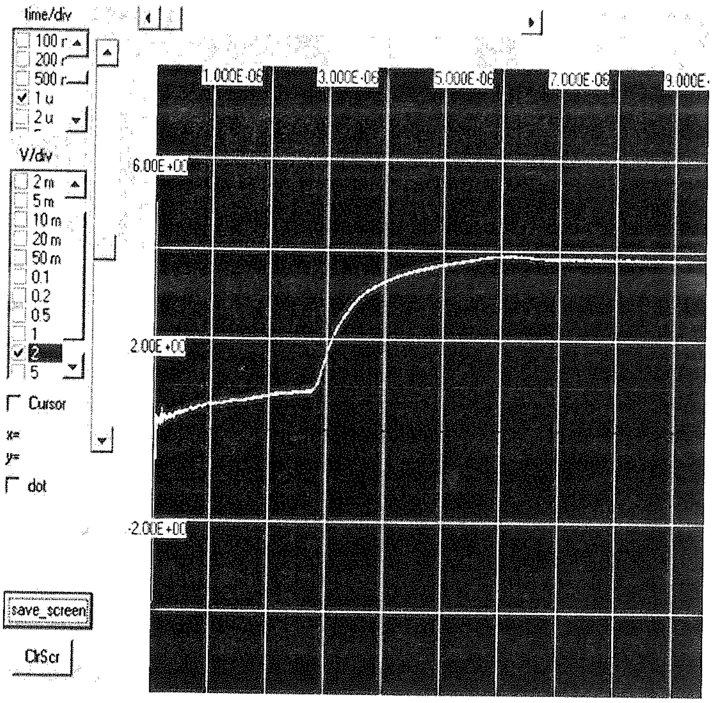
4. TESTING ENVIRONMENT

Software environment had to be designed in order to test the designed TDR module in the lack of a mainframe of DSL analyzer. Program, created in DELPHI environment, consist of 2 parts: emulation of a MASTER system (realizing data exchange with microprocessor of TDR system via RS interface) and graphical result visualization appropriate for PC screen. Graphical part was designed analogically to oscilloscopes screen – see Figs. 6, 7 for an illustration.

Controlling part of testing environment (see capture of the computer screen presented in Fig. 8) allows for:

- setting measurement parameters (delivered by the user from the computer keyboard),
- configuration of RS interface (selection of serial port number and transmission rate),
- system calibration,

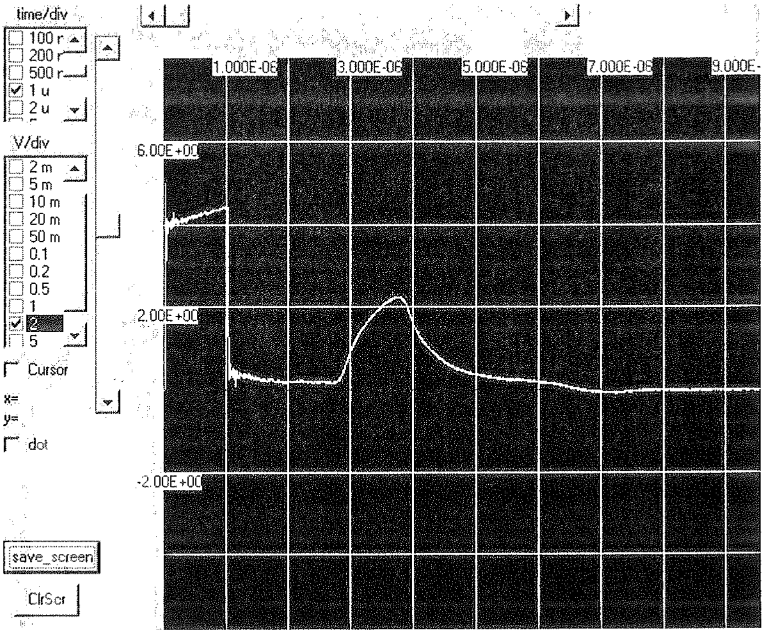
- measurements (single or continuous) with the possibility of averaging results of the subsequent experiments,
- phase inversion of the data received from TDR module,
- ‘soft pulse’ mode – processing of the measurement result obtained in response for step pulse, so as to evaluate line response to pulse of adjusted width (based on superposition of the adequately shifted opposite phase steps-induced waveforms - see Fig.7 for an illustration),
- breaking the measure in progress,
- restoring the initial state of the device,
- storing measurement results in the disc file and retrieving previously saved results.



time/div ☑ 1 u

V/div ☑ 2

Fig. 6. Visualization of a measured signal (reflection of the positive going step pulse from the open end of 300 m long line) in a testing environments

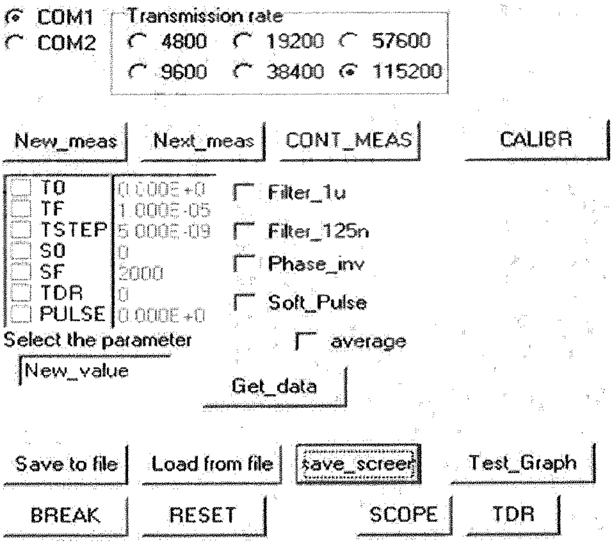


time/div
☐ 100 n
☐ 200 n
☐ 500 n
☑ 1 u
☐ 2 u

V/div
☐ 2 m
☐ 5 m
☐ 10 m
☐ 20 m
☐ 50 m
☐ 0.1
☐ 0.2
☐ 0.5
☐ 1
☑ 2
☐ 5

☐ Cursor
x=
y=
☐ dot

Fig. 7. Ilustration of the 'soft pulse' utility: reflection of 1 μs pulse from the open and of 300 m long line (obtained as the superposition of step-induced responses)



● COM1
○ COM2

Transmission rate
○ 4800 ○ 19200 ○ 57600
○ 9600 ○ 38400 ● 115200

☐ TO	0.000E+0
☐ TF	1.000E-05
☐ TSTEP	5.000E-09
☐ SO	0
☐ SF	2000
☐ TDR	0
☐ PULSE	0.000E+0

Select the parameter
New_value

☐ Filter_1u
☐ Filter_125n
☐ Phase_inv
☐ Soft_Pulse
☐ average

Fig. 8. Ilustration of a controlling part of software testing environment

5. CONCLUSIONS

A low-cost, small-size, programmable TDR unit designed for DSL line analyzer was presented. The device was tested with use of a dedicated software environment. Although the device is fully functional, further work is still needed. It concerns mainly the domain of signal processing so as to simplify and possibly automate measurement results interpretation.

6. REFERENCES

1. J. Strickland: *Time-domain reflectometry measurements*, Textronix Inc. (1970)
2. S. Misiaszek, A. Burd, K. Opalska, M. Radtke, M. Ramotowski, T. Starecki: *Time domain reflectometry educational tool kit for student laboratory*, Proc. ICSES, Wro-cław, Poland, pp. 195-200 (2002).
3. K. Opalska, S. Misiaszek: *A pulser with inverted microstrip line for time-domain reflectometry*, Proc. of SPIE, Vol/ 6937, pp. 693711-1 - 693711-6 (2007).
4. P. DeWinter, B. Ashley: *AN201 – Step vs Pulse TDR Technology*, AEA Technology Inc, www.aeatechnology.com, 2005
5. ×DSL Cable Qualifier CableSHARKR P3, operating manual.
6. A. Burd, T. Owczarek, K. Opalska: *TDR module for DSL line analyzer*, Technical report, ISE PW, 2008.

Dete

th
va
er
re
se
co

K

A s
sical, m
ordinar
method
usually
method
in a sh
ODEs s
equival
parallel

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 377–388

Determination of the initial conditions in the parallel method for the state equation solving

JAROSŁAW FORENC

Faculty of Electrical Engineering,
Białystok Technical University,
Wiejska 45D, 15-351 Białystok
e-mail: jarekf@pb.edu.pl

Received 2008.06.19
Authorized 2008.07.28

This paper presents the parallel method for state equation solving. The general idea of this method is based on the division of the integration interval into sub-intervals, in which the values of state variables are computed in parallel with the use of one of the well-known sequential numerical methods for state equation solving. Computations in particular sub-intervals require knowledge of initial conditions at the beginning of each sub-interval. In the proposed method the initial conditions are determined on the basis of an approximation of the convergence graph by the exponential function.

Keywords: state equation solving, ODEs, parallel computing

1. INTRODUCTION

A state space representation is often used in the analysis of the dynamics of physical, mechanical and electrical systems [1,2]. The state equation is a set of first-order ordinary differential equations (ODEs). Such equations are solved using numerical methods. Large and complicated physical systems modelling with a lot of parameters usually requires solving large sets of equations. The application of parallel, numerical methods of state equation solving can be helpful in achieving an accurate solution in a short time [3]. C.W. Gear [4,5] classified the means of achieving parallelism in ODEs solving into two main categories: parallelism across the system (problem) or equivalently parallelism across space and parallelism across the method or equivalently parallelism across time (steps). In some publications [3,6] parallelism across time is treated as a separate category. A good overview of parallel methods of ODEs solving can be found in some papers [4,5,7] and monographs [3,6].

In parallelism across the system and parallelism across the method each single processor has to communicate with all other processors in each integration step. It can be a problem, when computations are executed in parallel systems with a slow communication. In this paper we propose the parallel method for ODEs system solving, which belongs to the parallelism across time category and does not require communication in each integration step. This method originates from the approach towards the parallel analysis of transient states in electrical circuits published in [8,9]. The general idea of the method is based on a division of the integration interval into sub-intervals, in which the state variables values are computed in parallel with the use of one of the well-known sequential numerical methods of ODEs solving. The analysis of electrical circuit dynamics described by a nonlinear state equation is presented as a practical example of the application of the proposed method. Before the application of the parallel method, the nonlinear state equation was linearized with the use of the global linearization method [10-12]. A good approximation of the nonlinear state equation by a linear one requires computations with a small integration step size. In this case, the proposed method makes it possible to reduce the time of the computations.

2. THE METHOD OF PARALLEL STATE EQUATION SOLVING

The proposed method is designed for parallel solving of state equations, particularly when a state equation describes transient states in electrical circuits. The analysed state equation could be linear:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \mathbf{x}(t_0) = \mathbf{x}_0 \tag{1}$$

or nonlinear:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, t), \mathbf{x}(t_0) = x_0 \tag{2}$$

where: $\mathbf{x} \in \mathbf{R}^n$ – the vector of state variables, $\mathbf{u} \in \mathbf{R}^m$ – the vector of the input functions, $\mathbf{A}$ – the state matrix, $\mathbf{B}$ – the input matrix, $\mathbf{x}_0$ – the vector of initial conditions, $\mathbf{f}(\mathbf{x},\mathbf{u},t)$ – the vector of nonlinear functions.

The main aim of the application of the proposed method is to reduce the time of state equation solving. The general idea of this method is based on the division of the integration interval (t_0,t_N) into a given number of N equal sub-intervals (Fig. 1).

The number of sub-intervals should be equal to the number of processors used in computations. Computations of state variables values in particular sub-intervals are executed concurrently with the use of one of the well-known sequential numerical method for state equation solving e.g. Runge-Kutta or Fehlberg method. Parallel computations initiation requires the knowledge of state variable values (the initial conditions) at the beginning of each sub-interval, i.e. at time points $t_0,t_1,\ldots,t_{N-1}$. At

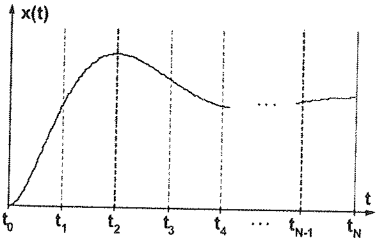


Fig. 1. Division of the integration interval into sub-intervals

time point t_0 the initial conditions are known from the assumption ($\mathbf{x}(t_0) = \mathbf{x}_0$). The remaining initial conditions are determined sequentially by one processor. From the convergence condition of the applied numerical method it results that the sequence of approximate solutions, obtained during repeated solving of the same state equation, but with decreased integration step size, is convergent to the exact solution (Fig. 2).

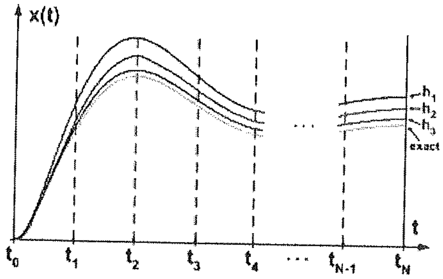


Fig. 2. Approximate solutions for different step sizes ($h_1 > h_2 > h_3$) and an exact solution

Let us check how the value of one state variable at a chosen time point during a decrease of the integration step size is changing. The obtained values are shown in the graph (Fig. 3), which contains integration step sizes (h_1, h_2, h_3; logarithmic scale) on the x-axis and the obtained values for these integration step sizes (x_{h1}, x_{h2}, x_{h3}) on the y-axis. Such a graph is called *the convergence graph.*

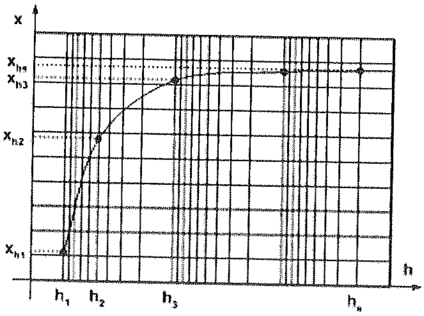


Fig. 3. The convergence graph in a logarithmic scale ($h_1 > h_2 > h_3 > h_s$)

The analysis of the above graph shows that the convergence graph of the sequential solution presented in a logarithmic scale could be approximated by the exponential function of the form:

$$x(t) = A(1 - e^{-\alpha t}) \tag{3}$$

The observation mentioned above was the basis for working out the determination method of state variable values at the beginning of each sub-interval. If the state variables values (x_{h1}, x_{h2}, x_{h3}) computed for large integration step sizes (h_1, h_2, h_3) are known, it is possible to compute A and α parameters of the exponential function (3) and then to determine an approximate value (x_{hs}) for a small h_s step size. In this method one processor (the master processor) solves state equation sequentially three times in the whole integration interval (t_0,t_N) with three integration step sizes: h_1, h_2 and h_3 ($h_1 > h_2 > h_3$). These computations are executed in a very short time because the h_1 step size is defined close to the limit of the stability of the applied sequential numerical method for state equation solving. Additionally the h_1, h_2 and h_3 step sizes should be chosen in such a way as to ensure the possibility of obtaining state variables values at time points $t_1, t_2, \ldots, t_{N-1}$.

The convergence graph could be an increasing function (Fig. 4a) or a decreasing function (Fig. 4b).

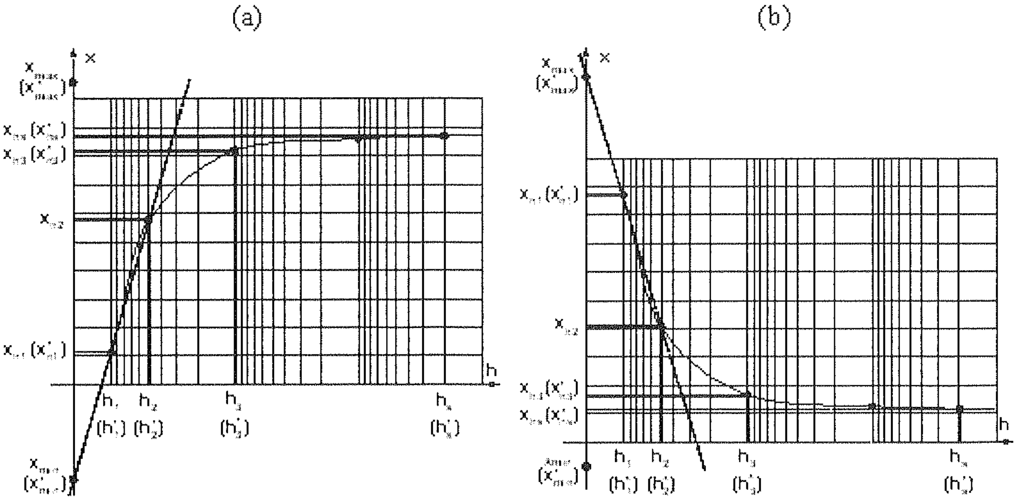


Fig. 4. The convergence graph as an increasing (a) and decreasing (b) function

In both cases it is necessary to use different formulas for the determination of initial conditions. Let us consider a situation when the convergence graph is an increasing function. The approximation of the convergence graph by an exponential function requires scaling the values on both axes of a coordinate system which causes passing of the graph through zero point of the coordinate system. The values on the axis of abscissa (h_1, h_2, h_3) are converted into distances from axis of ordinates (h_1^*, h_2^*, h_3^*). The scaling of values on the axis of ordinates requires the determination of lower (x_{min}) and upper (x_{max}) limits of the values for scaling. A lower limit is a point of crossing the convergence graph with the axis of ordinates, while an upper limit is determined on the basis of x_{h2} and x_{h3} values:

$$x_{\min} = \frac{x_{h2}h_1^* - x_{h1}h_2^*}{h_1^* - h_2^*}, \; x_{\max} = x_{h3} + (x_{h3} - x_{h2}) \tag{4}$$

The scaling is done on the basis of a simple proportion:

$$\frac{x - x_{\min}}{x_{\max} - x_{\min}} = \frac{x^* - x_{\min}^*}{x_{\max}^* - x_{\min}^*} \tag{5}$$

The transformation of (5) leads to a new value of x^*:

$$x^* = \frac{x_{\max}^* - x_{\min}^*}{x_{\max} - x_{\min}}(x - x_{\min}) + x_{\min}^* \tag{6}$$

Assuming, that $x_{min}^* = 0$ and $x_{max}^* = X$, where X is a positive constant, scaled values of x_{h1}^* and x_{h3}^* are obtained:

$$x_{h1}^* = \frac{X}{x_{\max} - x_{\min}}(x_{h1} - x_{\min}), \; x_{h3}^* = \frac{X}{x_{\max} - x_{\min}}(x_{h3} - x_{\min}) \tag{7}$$

Next, the parameters α and A of equation (3) are computed:

$$\alpha = \frac{\ln(x_{h3}^* - x_{h1}^*) + \ln x_{h1}^* - \ln x_{h3}^*}{h_3^* - h_1^*}, \; A = \frac{x_{h3}^*}{1 - e^{-\alpha h_3^*}} \tag{8}$$

As an approximate value of the initial condition, the value of A parameter of the exponential function (3), after reverse scaling, is assumed:

$$x_{hs} = \frac{A}{X}(x_{\max} - x_{\min}) + x_{\min} \tag{9}$$

When the convergence graph is a decreasing function (Fig. 4b) it is necessary to modify the (4-9) equations appropriately. In such case the lower (x_{min}) and the upper (x_{max}) limit of the values for scaling have the following form:

$$x_{\max} = \frac{x_{h2}h_1^* - x_{h1}h_2^*}{h_1^* - h_2^*}, \; x_{\min} = x_{h3} - (x_{h2} - x_{h3}) \tag{10}$$

After scaling values x_{h1}^* and x_{h3}^*:

$$x_{h1}^* = X - \frac{X}{x_{\max} - x_{\min}}(x_{h1} - x_{\min}), \; x_{h3}^* = X - \frac{X}{x_{\max} - x_{\min}}(x_{h3} - x_{\min}) \tag{11}$$

the parameters of exponential function are determined:

$$\alpha = \frac{\ln(x^*_{h1} - x^*_{h3}) + \ln x^*_{h1} - \ln x^*_{h3}}{h^*_3 - h^*_1}, \quad A = X - \frac{x^*_{h3}}{1 - e^{-\alpha h^*_3}} \tag{12}$$

and approximate value of initial condition is obtained:

$$x_{hs} = \frac{A}{X}(x_{\min} - x_{\max}) + x_{\max} \tag{13}$$

A monotonic convergence of the sequence of approximate solutions to the exact values was assumed in the described algorithm. It means that if $h_1 > h_2 > h_3$, than $x_{h1} \leq x_{h2} \leq x_{h3}$ (Fig. 4a) or $x_{h1} \geq x_{h2} \geq x_{h3}$ (Fig. 4b). In some cases the convergence graph has the forms presented in the Fig. 5a or 5b and the aforementioned conditions are not fulfilled. In such case it is necessary to execute additional computations with step size h_4 ($h_4 < h_3$), and to check the monotonicity of the convergence graph one more time but this time for the x_{h2}, x_{h3} and x_{h4} values.

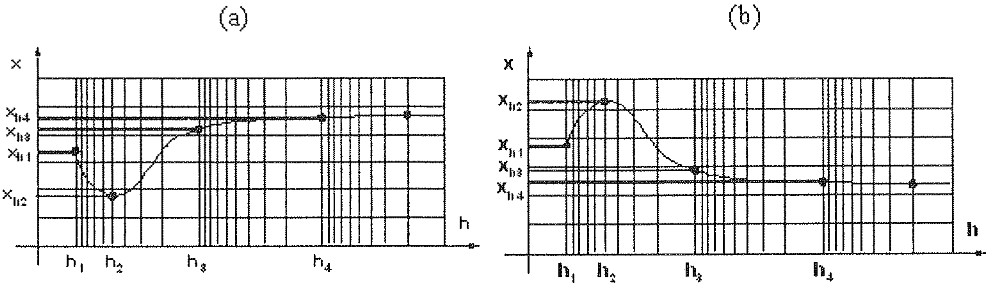


Fig. 5. The convergence graph (without the monotonicity) as an increasing function (a) and as a decreasing function (b)

If the monotonicity condition is not fulfilled again, the convergence graph has the form presented in Fig. 6. In such case the x_{h4} value, computed with h_4 step size, is assumed as the initial condition.

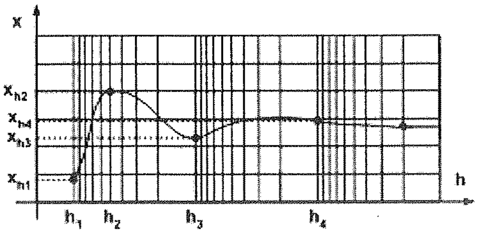


Fig. 6. The convergence graph without the monotonicity

Next, the master processor, on the basis of the obtained solutions, separately for each state variable and for each time point, computes parameters of equation (3), and then the initial conditions.

When all the initial conditions are determined, parallel computations of the state equation are started. The master processor sends the initial data to slave processors. The processors compute values of the state variables in subsequent sub-intervals: the master processor in the first sub-interval (t_0,t_1), the first slave processor in sub-interval (t_1,t_2), the second slave processor in sub-interval (t_2,t_3), etc. When the computations are finished, the slave processors send the obtained results to the master processor, which saves the complete solution on his local hard disk.

3. ANALYSIS OF THE TRANSIENT STATE IN ELECTRICAL CIRCUIT

The analysis of transient state in an electrical circuit, presented in Fig. 7, will be described as an application example for the proposed method.

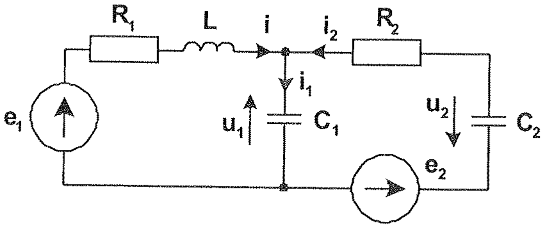


Fig. 7. The analysed nonlinear electrical circuit [13]

In the circuit above there is a non-linear resistance R_1 with the following characteristic:

$$R_1 = R_0(1 + ai^2) \tag{14}$$

The transient state in this circuit is described by the nonlinear state equation:

$$\begin{aligned} \dot{x}_1 &= -a_1x_1 - a_2x_1^3 - a_3x_2 + a_3e_1 & x_1(t_0) &= 0 \\ \dot{x}_2 &= a_4x_1 - a_5x_2 - a_5x_3 + a_5e_2 & x_2(t_0) &= 0 \\ \dot{x}_3 &= -a_6x_2 - a_6x_3 + a_6e_2 & x_3(t_0) &= 0 \end{aligned} \tag{15}$$

where $x_1 = i, x_2 = u_1$ and $x_3 = u_2$. The rest of coefficients on the right hand side of equation (15) results from the parameters of the circuit:

$$a_1 = \frac{R_0}{L},\ a_2 = \frac{aR_0}{L},\ a_3 = \frac{1}{L},\ a_4 = \frac{1}{C_1},\ a_5 = \frac{1}{R_2C_1},\ a_6 = \frac{1}{R_2C_2} \tag{16}$$

In order to examine the dynamics of the circuit, equation (15) have been linearized with the use of the global linearization method [10-12]. In this method the replacement of state variables is introduced:

$$z_1 = x_1,\ z_2 = -a_1 x_1 + x_2,\ z_3 = a_4 x_1 - a_5 x_2 + x_3 \tag{17}$$

Linearized equations have the following form [14]:

$$\begin{aligned} \dot{z}_1 &= z_2 + g_1(x, e) \\ \dot{z}_2 &= z_3 + \bar{g}_2(x, e) \\ \dot{z}_3 &= -k_1 z_1 - k_2 z_2 - k_3 z_2 + \bar{g}_3(x, e) \end{aligned} \tag{18}$$

where:

$$\begin{aligned} g_1(x, e) &= -a_2 x_1^3 - a_3 x_2 - x_2 + a_3 e_1 \\ \bar{g}_2(x, e) &= -a_1(-a_1 x_1 - a_2 x_1^3 - a_3 x_2 + a_3 e_1) - a_5 x_3 - x_3 + a_5 e_2 \\ \bar{g}_3(x, e) &= a_4(-a_1 x_1 - a_2 x_1^3 - a_3 x_2 + a_3 e_1) - a_6(x_2 + x_3 - e_2) \\ &-a_5(a_4 x_1 - a_5 x_2 - a_5 x_3 + a_5 e_2) + k_1 z_1 + k_2 z_2 + k_3 z_2 \end{aligned} \tag{19}$$

and

$$k_1 = a_3,\ k_2 = a_2,\ k_3 = a_1 \tag{20}$$

The final solution of state variables is determined by means of the inverse transformation:

$$\tilde{x}_1 = z_1,\ \tilde{x}_2 = a_1 z_1 + z_2,\ \tilde{x}_3 = (a_1 a_5 - a_4) z_1 + a_5 z_2 + z_3 \tag{21}$$

The computations were carried out with the use of a cluster of five workstations. Each node in the cluster contains an Intel Xeon 2.66 GHz processor and 1 GB RAM. Individual nodes are connected by a Gigabit Ethernet. The software environment is Ubuntu Linux and Open MPI as the message passing library.

The total time interval of transient analysis (t_0, t_N): $t_0 = 0$ $[s]$, $t_N = 0.001$ $[s]$ was divided into 5 equal sub-intervals $(t_0,t_1),\ldots,(t_4,t_5)$. The state variables values at t_1, t_2, t_3, t_4 time points (the initial conditions) were determined on the basis of computations executed with $h_1 = 10^{-9}$ $[s]$, $h_2 = 5 \cdot 10^{-10}$ $[s]$ and $h_3 = 10^{-10}$ $[s]$ step sizes. The main computations were carried out with the $h_s = 10^{-12}$ $[s]$ step size. The fourth-order Runge-Kutta method with a fixed integration step size was used as sequential numerical method for ODEs system solving.

In the proposed method, accuracy of the solution is mainly dependant on the de-termination accuracy of state variables values at the beginning of each sub-interval. In order to estimate the accuracy of determination of the initial conditions, the system of linear equations (18-20) was solved with the application of a sequential algorithm of the Runge-Kutta method with the same integration $h_s = 10^{-12}$ $[s]$ step size as in the parallel method. The obtained values of the sequential and parallel solution at t_1, t_2, t_3 and t_4 time points are presented in Table 1 and Table 2, respectively.

(17)

(18)

(19)

(20)

e trans-

(21)

tations.
RAM.
nent is

)01 [s]
lues at
compu-
es. The
h-order
merical

on the
nterval.
system
gorithm
s in the
t_1, t_2, t_3

Table 1

The values of the sequential solution at the beginning of sub-intervals

	$t_1 = 0.0002$ [s]	$t_2 = 0.0004$ [s]	$t_3 = 0.0006$ [s]	$t_4 = 0.0008$ [s]
$x_1(t)$	-0.003515	-0.000525	-0.000072	0.000014
$x_2(t)$	10.717674	10.227251	10.081765	10.029795
$x_3(t)$	-0.680025	-0.209873	-0.074366	-0.026698

Table 2

The values of the parallel solution at the beginning of sub-intervals

	$t_1 = 0.0002$ [s]	$t_2 = 0.0004$ [s]	$t_3 = 0.0006$ [s]	$t_4 = 0.0008$ [s]
$\tilde{x}_1(t)$	-0.003510	-0.000524	-0.000071	0.000014
$\tilde{x}_2(t)$	10.717939	10.227514	10.081934	10.029885
$\tilde{x}_3(t)$	-0.654188	-0.201514	-0.071219	-0.025504

Table 3 presents the relative errors of determination of the initial conditions. The sequential solution as an exact solution and the parallel solution as an approximate solution were assumed during the calculation of the relative errors. As we can see, the differences between the sequential solution and parallel solution at the beginning of sub-intervals have not exceeded 1.5% for the first state variable, 0.004% for the second state variable and 5% for the third state variable. In the case of the first and the third state variable the values are close to zero, so even a small difference causes large percentage errors which do not reflect the real results.

Table 3

The relative errors of the parallel solution

	$t_1 = 0.0002$ [s]	$t_2 = 0.0004$ [s]	$t_3 = 0.0006$ [s]	$t_4 = 0.0008$ [s]
$\delta_{\tilde{x}1(t)}$	0.151 %	0.348 %	0.878 %	1.221 %
$\delta_{\tilde{x}2(t)}$	0.003 %	0.003 %	0.002 %	0.001 %
$\delta_{\tilde{x}3(t)}$	3.949 %	4.148 %	4.406 %	4.685 %

Fig. 8-10 present the obtained parallel solution of all state variables. In these figures the limits of division of the total time interval of the transient analysis into 5 sub-intervals are also marked.

In order to determine the speedup, the total time of the transient analysis was divided into 2, 3, 4 and 5 sub-intervals. The time of the parallel method solution was referred to the time of the sequential solution (Runge-Kutta method with fixed

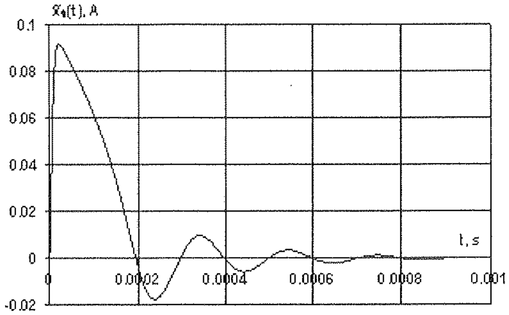


Fig. 8. Solution of the state variable $\tilde{x}_1(t)$

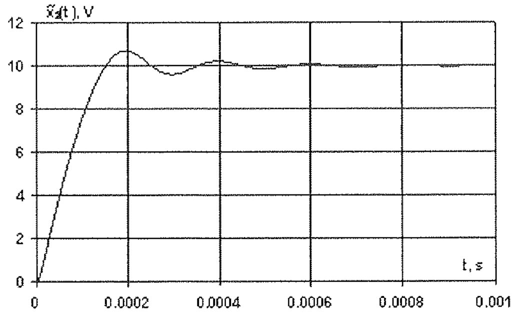


Fig. 9. Solution of the state variable $\tilde{x}_2(t)$

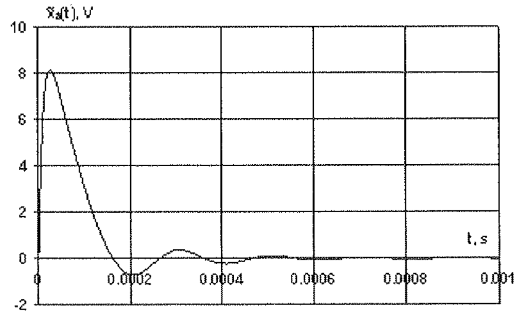


Fig. 10. Solution of the state variable $\tilde{x}_3(t)$

integration step size: $h_s = 10^{-12}$ [s]). The obtained values of the speedup are presented in Fig. 11.

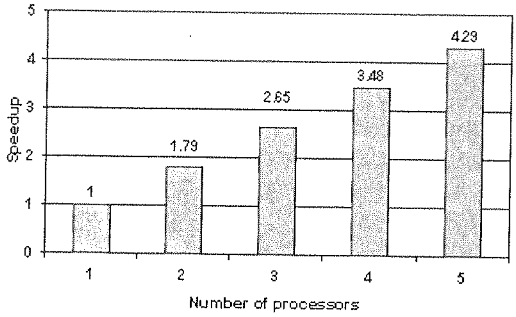


Fig. 11. Speedup in the proposed method

4. CONCLUSIONS

The state variables values at the beginning of each sub-interval determined on the basis of an approximation of the convergence graph solution by exponential function are close to the values of the sequential solution. It ensures a good accuracy of the parallel solution. The application of the presented algorithm also enables the reduction of the computation time. This is of particular importance in the case of systems of ODEs in which the function of the right-hand side of the state equation is expensive to evaluate, the number of equations in the system is large, the interval of integration is long, the system must be solved repeatedly or achieving high accuracy of computation results requires a very small integration step size. Good speedup and accuracy of the parallel solution were obtained in the presented example.

5. ACKNOWLEGMENTS

This work was carried out within the frame of Białystok Technical University Grant No: S/WE/3/08.

6. REFERENCES

1. N. S. Nise; *Control Systems Engineering. 4th Edition.* John Wiley & Sons, 2003.
2. A. Isidori: *Nonlinear Control Systems.* Springer-Verlag, Berlin Heidelberg New York, 1995.
3. K. Burrage: *Parallel and Sequential Methods for Ordinary Differential Equations.* Clarendon Press, Oxford, 1995.
4. C. W. Gear, X. Xuhai: *Parallelism Across Time in ODEs.* Applied Numerical Mathematics 11, pp. 45-68, 1993.

5. C. W. Gear: *Massive Parallelism Across Space in ODEs.* Applied Numerical Mathematics 11, pp. 27-43, 1993.
6. D. Petcu: *Parallelism in Solving Ordinary Differential Equations.* Mathematical Monographs 64, Timisoara University Press, 1998.
7. K. Burrage: *Parallel Methods for Initial Value Problems.* Applied Numerical Mathematics 11, pp. 5-25, 1993.
8. J. Forenc: *Speculative analysis of transient states in electrical circuits.* Doctoral thesis. Politechnika Białostocka, Białystok, 2006 (in Polish).
9. J. Forenc: *Time Domain Decomposition in Parallel Analysis of Transient States*, PARELEC'2006, International Symposium on Parallel Computing in Electrical Engineering, IEEE Computer Society, Los Alamitos, pp. 295-300, 2006.
10. T. Kaczorek: *The introduction to the Lie algebra.* The Seminary in the Technical University of Białystok, 2003 (not published).
11. A. Jordan, J. P. Nowacki: *Global Linearization of Non-Linear State Equations.* International Journal of Applied Electromagnetics and Mechanics, Vol. 19, Number 4, pp. 637-642, 2004
12. A. Jordan, T. Kaczorek, P. Myszkowski: *Linearization of nonlinear differential equations.* Wydawnictwa Politechniki Białostockiej, Białystok, 2007 (in Polish).
13. T. Kaczorek: *Application of polynomials and rational matrices in the theory of dynamic systems.* Wydawnictwa Politechniki Białostockiej, Białystok, 2004 (in Polish).
14. J. Forenc: *The Parallel Method for the State Equation Solving.* Computer Applications in Electrical Engineering, ed. by R. Nawrowski, Poznań University of Technology, Poznań, pp. 87-96, 2007.

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 389–395

The iterative approach to the optimal linearization method

PAWEŁ MYSZKOWSKI, JAROSŁAW FORENC

Faculty of Electrical Engineering,
Bialystok Technical University,
Wiejska 45D, 15-351 Białystok
e-mail: pmyszkowski@we.pb.edu.pl jarekf@pb.edu.pl

Received 2008.07.09
Authorized 2008.08.28

In presented paper the analysis of the non-linear models of physical systems will be considered. The analysis of the non-linear system dynamics will be performed by means of the linear model obtained by the optimal linearization method. The paper presents a new approach to this method, based on iterative process.

Keywords: linearization, non-linear state equation, optimal linearization, iterative approach

1. INTRODUCTION

In considerations, usually, the physical systems are treated as linear. It follows from accepted simplified assumptions that characteristics of the circuit elements are linear. However, in many cases these assumptions cannot be accepted and a non-linear differential system (the state equation) is used in analysis:

$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{u}, t), \ \mathbf{x}(0) = \mathbf{x}_0 \tag{1}$$

where $\mathbf{x} \in R^n$ is a state vector, $\mathbf{u} \in R^m$ is an input vector and $\mathbf{x}_0$ represents a vector of initial conditions.

The analysis of the system dynamics with the use of the non-linear equation (1) is a complicated and ambiguous problem in many cases [1,2]. Considerations based on determination and analysis of Lyapunov function lead to a complicated description, especially for the time-varying systems of many state variables. The method which is more useful for analysing the dynamics of non-linear systems is the analysis of their linear equivalents, obtained by the use of linearization methods: Taylor's series expansion around the equilibrium point [1], optimal linearization [3,4] and global linearization [5,6,7]. In this paper the optimal linearization method is used in considerations.

It should be noted that linear equivalents of non-linear models also enable the determination of the stability area of the models with the use of analysis of eigenvalues of the linear system matrix.

2. THEORETICAL BASIS OF THE OPTIMAL LINEARIZATION

The optimal linearization method is based on the analysis of mean square error. To linearize the non-linear state equation (1) its linear approximation is introduced:

$$\dot{\mathbf{x}} = \mathbf{A}^*\mathbf{x} + \mathbf{B}^*\mathbf{u} \tag{2}$$

where $\mathbf{A}^*$ and $\mathbf{B}^*$ are the optimal matrices

$$\begin{aligned} \mathbf{A}^* &= \left[a_{ij}^*\right],\ i = 1,...,n,\ j = 1,...,n \\ \mathbf{B}^* &= \left[b_{ij}^*\right],\ i = 1,...,n,\ j = 1,...,m \end{aligned} \tag{3}$$

Optimal elements a_{ij}^* and B_{ij}^* are chosen is such a way as to obtain the minimal value of mean square error between the non-linear function (1) and its linear approximation. First the small difference between the non-linear equation and the linear equation $f(\mathbf{x}, \mathbf{u}, t)$ is defined:

$$\varepsilon = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} - f(\mathbf{x}, \mathbf{u}, t) \tag{4}$$

Next, to obtain unknown elements a_{ij}^* and b_{ij}^* $(i=1,2,...,n;\ j=1,2,...,n)$ of matrices $\mathbf{A}^*$ and $\mathbf{B}^*$, the following functional is minimized:

$$I\left(a_{ij}, b_{ij}\right) = \int_0^{t_1} \varepsilon^T \varepsilon dt \tag{5}$$

The optimal elements a_{ij}^* and b_{ij}^* are calculated with the use of following formulas:

$$\frac{\partial I(a_{ij}, b_{ij})}{\partial a_{ij}}\Big|_{a_{ij} = a_{ij}^*} = 0 \frac{\partial I(a_{ij}, b_{ij})}{\partial b_{ij}}\Big|_{b_{ij} = b_{ij}^*} = 0 \tag{6}$$

To determine values of optimal elements a_{ij}^* and b_{ij}^* base functions in the form of tabulated solution of linear approximation are used. Predominantly the Taylor's series expansion around the equilibrium point is used for this purpose. Elements a_{ij}^* and b_{ij}^* are determined with the use of numerical computations, however, for more than three state variables this process becomes very complicated and the complexity increases in direct proportion to the order of non-linear system.

The procedure for determining of the optimal elements a_{ij}^* and b_{ij}^* was introduced in [7].

3. THE ITERATIVE APPROACH TO OPTIMAL LINEARIZATION METHOD

The analysis of linear approximations obtained by the optimal linearization with the use of different base functions shows that the choice of the base function influences on the accuracy of linear approximation of non-linear state equation. It is illustrated in Fig. 1, that shows a comparison of solution of non-linear equation (NLIN) to solutions of linear equations obtained by the optimal linearization for different base functions: Taylor's series expansion around the equilibrium point (Taylor SS), optimal line (OPT.LINE) and Taylor's series expansion with the use of transient component (Taylor SS, Tr.Comp.). As an example there was used the non-linear model of electrical circuit containing a DC-motor supplied by a solar generator [8]. In Fig 1. results for the first state variable are presented.

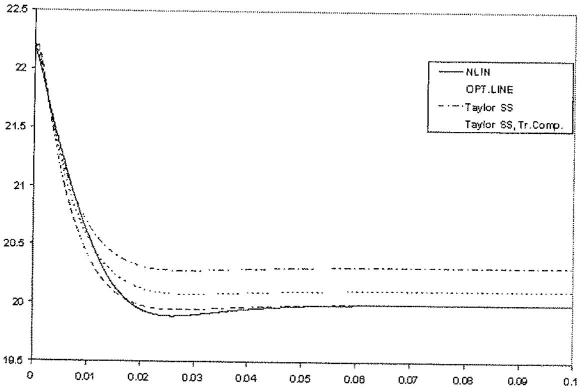


Fig. 1. The influence of the base function on the accuracy of linear approximation

There appears a question, if applying as a base function for the next iteration of this method the linear system obtained by the optimal linearization give as a result the linear system being a better approximation of analyzing non-linear system. This question is as important as determining different base functions is connected with a different amount of work, from the simplest Taylor's series expanstion through the optimal line to Taylor's series expansion with the use of transient component. If the iterative process for "the worst" base function will be convergent to the exact solution, it will be better way to apply it than to look for the better, although more complicated, base functions.

4. NUMERICAL EXAMPLE

As an example of the application of this method the analysis of an electrical circuit containing a diode with non-linear characteristic will be presented [9].

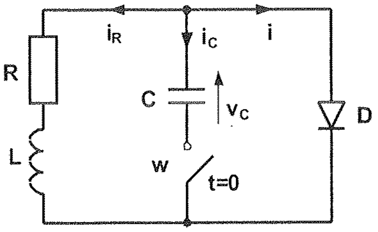


Fig. 2. Electrical circuit with a non-linear diode

The circuit is shown in Fig. 2. In this case the non-linear state equation takes the form $\dot{x} = f(x,t)$, $x(0) = x_0$. The diode has a non-linear characteristic $i = av_c + bv_c^2$. The state equation describing the system is as follows [9]:

$$\begin{aligned} \dot{x}_1 &= -a_1 x_1 - a_2 x_1^2 - a_3 x_2 \\ \dot{x}_2 &= a_4 x_1 - a_5 x_2 \\ x_1(0) &= v_{c,0},\ x_2(0) = 0 \end{aligned} \tag{7}$$

where $x_1 = v_c$ is the potential drop in the capacitor, $x_2 = i_R$ is the current passing through resistor (Fig. 2), and a and b are constants. Parameters of the system are given as follows:

$$a_1 = \frac{a}{C},\ a_2 = \frac{b}{C},\ a_3 = \frac{1}{C},\ a_4 = \frac{1}{L},\ a_5 = \frac{R}{L} \tag{8}$$

Following rated values of circuit parameters were used:

$$\begin{aligned} &R = 100\Omega,\ C = 5 \cdot 10^{-6}\mathrm{F},\ L = 0.5\mathrm{mH}, \\ &a = 3.5 \cdot 10^{-3}\mathrm{AV}^{-1},\ b = 10^{-2}\mathrm{AV}^{-2},\ v_{c,0} = 20\mathrm{V} \end{aligned} \tag{9}$$

Graphs of both state variables for rated values of circuit parameters are presented in Fig. 3. Additionally, there are presented graphs for $b = 0$, when the diode has a linear characteristic.

The non-linear equation (7) was linearized with the use of the optimal linearization method. The state equation (7) has no input, therefore the optimal linear equation is given as follows:

$$\dot{\mathbf{x}} = \mathbf{A}^*\mathbf{x} \tag{10}$$

As a base function a Taylor's series expansion around the equilibrium point was used. Optimal matrix $\mathbf{A}_1^*$ for the first approximation has the following values:

$$\mathbf{A}_1^* = \begin{bmatrix} -53221.4 & 2383740 \\ 2000 & -200000 \end{bmatrix} \tag{11}$$

Next the linear equation $\dot{\mathbf{x}} = \mathbf{A}_1^*\mathbf{x}$ was used for determining the second approximation. Obtained optimal matrix $\mathbf{A}_2^*$ is given as follows:

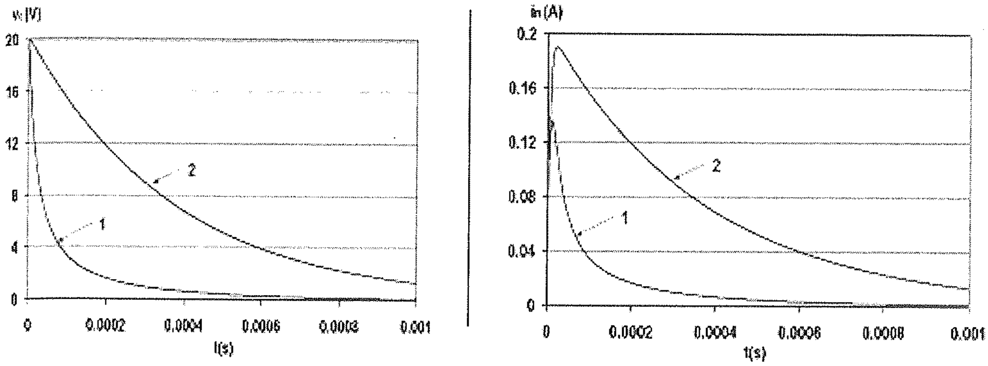


Fig. 3. Numerical solution of non-linear equation for rated values of circuit parameters (1) and for $b = 0_{(2)}$ for both state variables

$$\mathbf{A}_2^* = \begin{bmatrix} -46156.2 & 2183110 \\ 2000 & -200000 \end{bmatrix} \tag{12}$$

This process was continued and in third iteration the optimal matrix $\mathbf{A}_3^*$ has the following values

$$\mathbf{A}_3^* = \begin{bmatrix} -46970.8 & 2216770 \\ 2000 & -200000 \end{bmatrix} \tag{13}$$

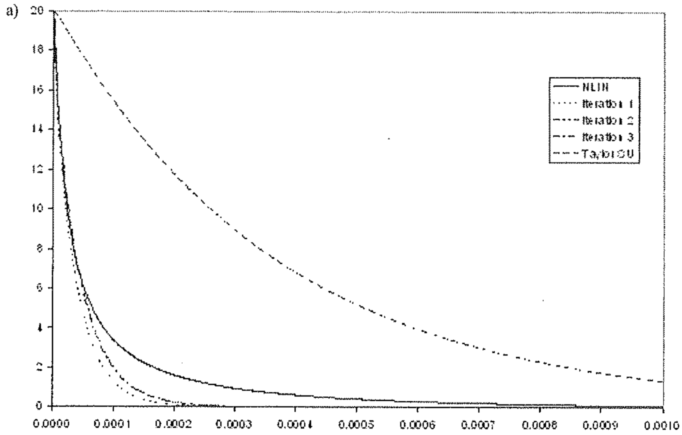
Results for both state variables are presented in Fig.4. The numerical solutions of obtained linear systems with matrices (11), (12) and (13) were compared with numerical solution of non-linear equation (7) and additionally with numerical solution of Taylor's series expansion of Eq. (7).

The difference between the numerical solution of equation (7) and that of equation (10) can be characterized by the norm $\|x\|_1 = \max\limits_{1 \le i \le n} |x_i|$ [10]

$$\|x_i(t) - \tilde{x}_i(t)\|_1 \le \varepsilon_i, i = 1, 2, ..., n \tag{14}$$

where $\tilde{x}_i(t)$ is a numerical solution of i-th state variable of non-linear equation and $x_i(t)$ is a numerical solution of i-th state variable of linear equation.

In Table 1 the values of norm for previously used approximations are presented.



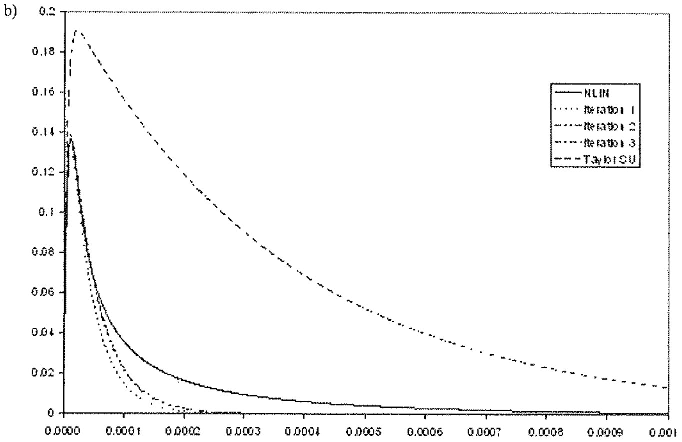


Fig. 4. The comparison between numerical solution of non-linear equation (7) and numerical solution of non-linear equations with matrices (11), (12), (13) – (a) – state veriable x_1, (b) – state veriable x_2

Table 1

Values of norms 14 for different linear approximations of non-linear equation

	ε_1	ε_2
Taylor's series expansion	12.11753	0.121073
Optimal linearizarion, iteration 1	2.064131	0.0206
Optimal linearizarion, iteration 2	1.551733	0.015494
Optimal linearizarion, iteration 3	1.598607	0.015962

Tl
— the
sta
ap
— the
res
exp
Du
lysed.

Th

1. T.
theo
2. T.
3. A.
Ana
4. M.
stati
Phys
5. A.
6. R.
199
7. A.
atio
8. A.
Jour
9. T. K
drive
PJIIT
10. T. K
Polis

5. CONCLUSIONS AND FUTURE REMARKS

The above considerations allow us to make the following conclusions:

- the optimal linearization method gives a good approximation of the non-linear state equation, however the base function has an influence on the accuracy of this approximation and linear model,
- the iterative approach to the optimal linearization method enables to obtain better results in the first iteration, however, next iterations do not give as good results as expected to be.

During the future works on this method more complicated examples will be analysed.

6. ACKNOWLEGMENTS

This work has been carried out within the Dean's project no S/WE/3/08.

7. REFERENCES

1. T. Kaczorek, A. Dzieliński, W. Dąbrowski, R. Łopatka: *The basis of control theory*, WNT Warszawa 2005 (*in Polish*).
2. T. Kaczorek: *System and control theory*, PWN Warszawa, 1999 (*in Polish*).
3. A. Jordan et al: *Optimal Linearization of Non-Linear State Equations*, Rairo-Automatique, System Analysis and Control, pp. 263-271, 1987.
4. M. Barland et al: *Commande optimale d'un systeme generateur photovoltaique convertisseur statique – recepteur*, Revue Phys. Appl. 19, pp. 905-915, Commision des Publications, Francaises de Physique, Paris, 1984.
5. A. Isidori: *Nonlinear Control Systems*, Springer Verlag, 1995.
6. R. Marino, P. Tomei: *Nonlinear Control Design – geometric, adaptive, robust*, Prentice Hall, 1995.
7. A. Jordan, T. Kaczorek, P. Myszkowski: *Linearization of non-linear differential equations*, Publishing of Bialystok Technical University, Białystok, 2007 (*in Polish*).
8. A. Jordan, J. Nowacki: *Global Linearization of Non-Linear State Equations*, International Journal of Applied Electromagnetics and Mechanics, pp. 637-642, vol. 19, Nr 1-4, 2004.
9. T. Kaczorek, A. Jordan, J. Forenc: *Global linearization of a non-linear model of a DC drive system*, 2nd Grant Conference "Numerical methods in modelling of electric devices", pp.7-16, PJIIT, Warsaw, 2002.
10. T. Kaczorek: *Vectors and matrices in automatics and electrotechnics*, WNT, Warszawa, 1998 (*in Polish*).

S
ment
tation
in cas
suppl
by sw
– **2x1**
bridg

C
provic

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 397–404

Experimental Verification of Double Current Strain Gauge Meter

WOJCIECH WALENDZIUK, ADAM IDŹKOWSKI

Faculty of Electrical Engineering,
Bialystok Technical University,
Wiejska 45D, 15-351 Bialystok
e-mail: walenw@pb. edu.pl adam_i@we.pb.edu.pl

Received 2008.06.03
Authorized 2008.08.29

A four-terminal (4T) bridge-circuit with a current source connected in parallel to opposite arms, is presented. This two-output circuit is named a double current bridge. The example of its application is presented. Signal conditioning formulas of two-parameter measurement of strain and temperature are discussed in detail. An advantage of this solution is that temperature compensation is done without any additional temperature sensor. Some results achieved with the use of this bridge-circuit are briefly described. The standard uncertainties of strain measurement and temperature measurement have been calculated.

Keywords: multichannel measurement systems, bridge circuits, strain and temperature measurement, uncertainty

1. INTRODUCTION

Since 1998 Warsza has been proposing two types of structures for such measurement and for the primary signal conditioning on the input analogue part of instrumentation channels. One of them is the circuit of two four-arm classic bridges connected in cascade [1,2]. The other one has an unconventional supplying: the 4T circuit is supplied by two equal current sources in parallel to opposite arms – **2J** or in practice by switching over two unequal sources between these arms – **2x2J** or even only one – **2x1J**. For all of these circuits Warsza proposed a common name: **double current bridges**. The circuits were described in [3,4] and more extensively in [5].

Currently, an improvement of strain, pressure, force, torque or other measurements provide many circuit designs. They differ in a realization of temperature compensation.

The sensor's thermal error (drift of sensor's offset and span) is compensated in the digital part of a conditioner by proper correction algorithms. For pressure measurement a piezoresistive sensor can be powered by an adjustable current source combined with a programmable-gain amplifier and external trimmable resistors (e.g. MAX1450) [7], or two amplifiers and two digitally controlled potentiometers [8], or four digital-to-analog converters (DAC) resulting in a temperature-depended bridge voltage (e.g. MAX1452) [9].

A circuit proposed in this paper assures the temperature reading and compensation in whole measuring span. It is done without any additional sensor such as thermistor or thermistor-resistor parallel networks [10,11].

To illustrate this concept of simultaneous two-parameter measurement (one-axis strain and temperature) using strain gauges plugged in a double current bridge, an experimental bridge-circuit (2x1J) was built. Two output signals of the bridge (representing temperature and strain) are interfaced to a microcontroller by ADCs (analog-to-digital converters).

2. DOUBLE CURRENT BRIDGE

2J bridge differs from the Wheatstone bridge in the way of supplying [1]. It is possible to arrange it for two variable measurements. The bridge has two outputs: A-B and D-C. Two equal current supply sources J are connected in parallel to opposite arms (R_1, R_3) of the bridge. However, it is difficult to implement practically. If the excitations are not equal, the output voltage equations have additional components which are dependent on the difference ΔJ [1,5].

The bridge can also be supplied by a single current source J switched over to the same arms – 2x1J (Fig. 1). Then each of the output voltages is held, summed up in two cycles (the superposition theorem) and measured, i.e.:

$$U_{DC} = U_{DC_1} + U_{DC_2} \tag{1}$$

$$U_{AB} = U_{AB_1} + U_{AB_2} \tag{2}$$

The bridge in Fig. 1. works as unbalanced and produces two output voltages (U_{DC} and U_{AB}):

$$U_{AB} = T_0(\varepsilon_1 - \varepsilon_2) \tag{3}$$

$$U_{DC} = T_0(\varepsilon_1 + \varepsilon_2) \tag{4}$$

where : $\mathrm{T}_0 = JR_0 / 4$ – the initial voltage sensitivity is equal for both outputs,
$\varepsilon_1, \varepsilon_2$ – the relative changes in resistances R_1, R_2.

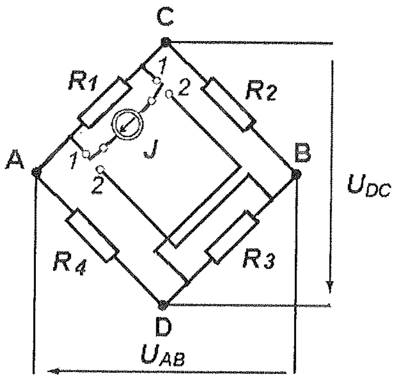


Fig. 1. The idea of double current bridge circuit for temperature and bending strain measurement (R_1 and R_2 are strain gauges) [6]

The formulas (3) and (4) are simplified and true if modules of the values $|\varepsilon_1|, |\varepsilon_2|$ are small enough.

3. PRACTICAL REALIZATION OF THE BRIDGE

The theory has been verified in the experimental circuit given in Fig. 2. The switched current source was constructed with the use of LM317 and four MOSFET switches (STP20NE06L), which has low on-resistance $R_{ON} = 0.06\ \Omega$. The current excitation can be manually adjustable from 9 mA to 38 mA. The transistors work in pairs – two switched on and two switched off at the same time. Their state of work is controlled by ATmega16 microcontroller ports (Fig. 2).

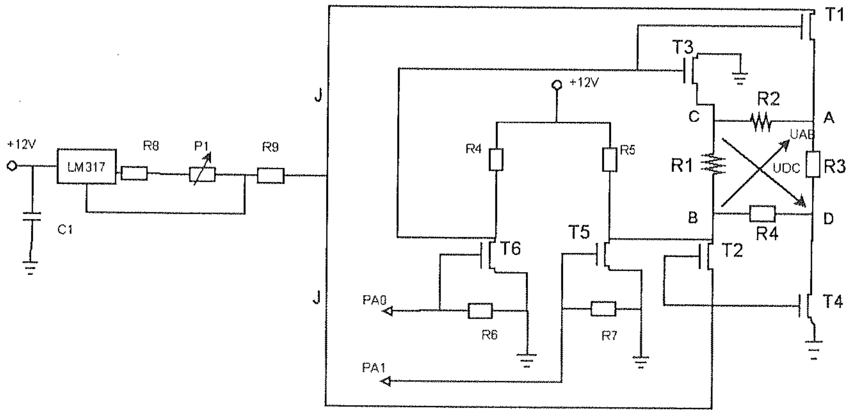


Fig. 2. Practical scheme of the bridge circuit

The output voltages U_{AB} and U_{DC} (two of them of positive sign and another two of negative sign) are connected to a 24-bit sigma-delta ADC (AD7718) via post-conditioning module (Fig. 3). It consists of instrumentation amplifiers (AD620AN) and ultra-precision voltage-dividers (MAX5491). The use of this module is necessary because AD7718 requires positive sign voltages of (0 - 2.56V).

The acquired voltage data from the circuit outputs are processed by the microcontroller.

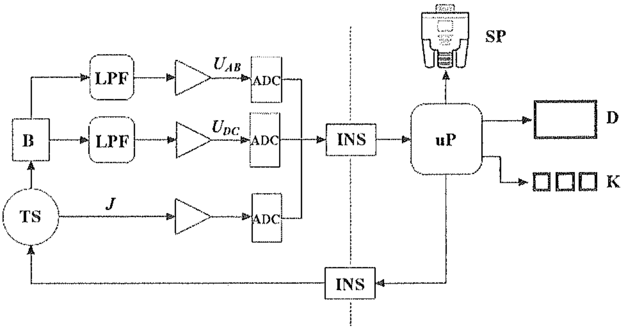


Fig. 3. Block diagram of measurement system (B – bridge, TS – MOSFET switches and current source, LPF – low pass filter, ADC – analog-to-digital converter, INS – insulation, uP – microprocessor, D – LCD display, K – keyboard, SP – serial port)

4. STRAIN AND TEMPERATURE MEASUREMENT

The increments in resistance of strain gauges consist of two parts ($\varepsilon_1 = \varepsilon' + \varepsilon''$, $\varepsilon_2 = \varepsilon' - \varepsilon''$, respectively). One part is increment due to temperature (5), the other one is the increment (or decrement) due to mechanical stress (6). If there are two strain gauges of the same type, the relative increments in temperature are of the same value and of the same sign. If strain gauges are glued to a beam such that the first gauge is stretched and the other compressed (Fig. 4), then the increments due to mechanical stress are of the opposite signs.

$$\varepsilon_1^{'}(\Delta T) = \varepsilon_2^{'}(\Delta T) = \varepsilon^{'} \tag{5}$$

$$\varepsilon_1^{''}(\varepsilon_b) = -\varepsilon_2^{''}(\varepsilon_b) = \varepsilon^{''} \tag{6}$$

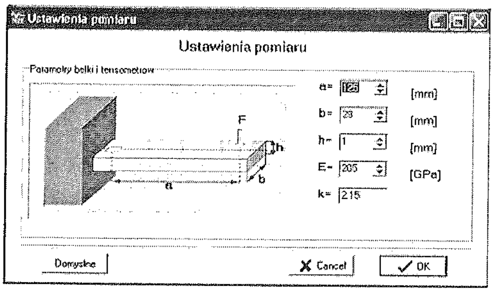
Increments could be considered as linear for both measured quantities, i.e. $\varepsilon'(\Delta T) = \alpha_T \Delta_T; \varepsilon''(\varepsilon_b) = k \cdot \varepsilon_b$ (where: α_T – the temperature coefficient of resistance, ΔT – change of temperature, $k = k_0(1 + \alpha_K \Delta T)$ – strain gauge factor, ε_b – bending strain).

After substitution, both functions can be presented as follows:

$$\varepsilon' = \alpha_T \cdot \Delta T = \frac{U_{DC}}{2T_0} = \frac{2}{JR_0}U_{DC} \Rightarrow \Delta T = \frac{2U_{DC}}{JR_0\alpha_T} \tag{7}$$

$$\varepsilon'' = k \cdot \varepsilon_b = \frac{U_{AB}}{2T_0} = \frac{2}{JR_0}U_{AB} \Rightarrow \varepsilon_b = \frac{2U_{AB}}{JR_0 k} \tag{8}$$

The both measured quantities depend linearly on the output voltages U_{AB} and U_{CD} (respectively), supplying current J and the parameters of the gauges.


Ustawienia pomiaru

Ustawienia pomiaru

Parametry belki i tensometrow

a= 125 [mm]
b= 23 [mm]
h= 1 [mm]
E= 205 [GPa]
k= 215

Fig. 4. Setting up the dimensions of cantilever beam Young's modulus and strain gauge factor to receive a value of stress, bending force or bending moment

Strain and temperature measurements were taken for four temperatures of a cantilever beam, while the beam was bent by a micrometer screw. The results of the experiment are shown in Fig. 5. It presents the charts of relative resistance increments (strain) of both strain gauges in the function of deflection X of the beam. The temperature has a significant influence on the value of strains and causes an offset (approximately 0.1% for temperature change ΔT = 40°C). If the value of this offset is known, the temperature effect can be easily corrected by microprocessor software.

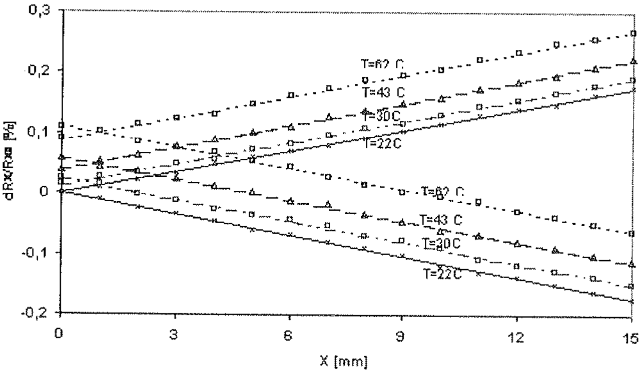


Fig. 5. Relative increments of resistance of both strain gauges in function of beam deflection X in different temperatures

5. UNCERTAINTY CALCULATIONS

Standard uncertainties (9) and (10) were calculated for the parameters in Table 1, after analysis of the circuit. The voltages are measured on 3 channels of ADC. Resolution of AD7718 converter is 24 bit. The measurement of current is done by voltage reading on the resistor 39 Ω (±0.1%). The variance (the method of type B) of current J ($u^2(J)$) is $7 \cdot 10^{-11}$ $(mA)^2$.

Table 1

Input data needed for the uncertainty calculations

Quantity and value	$U_{AB} = 2mV$	$U_{DC} = 1mV$	$R_{10} = 120\Omega$	$k = 2.15$	$\alpha_T = 0.0410^{-3}$ $l/°C$
Limiting error	$0.153\mu V$	$0.153\mu V$	0.24 Ω (±0.2%)	0.01075 (±0.5%)	0

$$u^2(\Delta T) = \left[\frac{\partial(\Delta T)}{\partial U_{DC}}\right]^2 \cdot u_B^2(U_{DC}) + \left[\frac{\partial(\Delta T)}{\partial J}\right]^2 \cdot u_B^2(J) + \left[\frac{\partial(\Delta T)}{\partial R_{10}}\right]^2 \cdot u_B^2(R_{10}) \tag{9}$$

$$u^2(\varepsilon_b) = \left[\frac{\partial(\varepsilon_b)}{\partial U_{AB}}\right]^2 \cdot u_B^2(U_{AB}) + \left[\frac{\partial(\varepsilon_b)}{\partial J}\right]^2 \cdot u_B^2(J) + \left[\frac{\partial(\varepsilon_b)}{\partial R_{10}}\right]^2 \cdot u_B^2(R_{10}) + \left[\frac{\partial(\varepsilon_b)}{\partial k}\right]^2 \cdot u_B^2(k) \tag{10}$$

The budgets of uncertainty of temperature measurement (first column) and of strain measurement (second column) are presented in Table 2.

Table 2

The budgets of uncertainty and the results of strain and temperature measurement

$\left[\frac{\partial(\Delta T)}{\partial U_{DC}}\right]^2 \cdot u_B^2(U_{DC}) = 6.36 \cdot 10^{-6}(K)^2$	$\left[\frac{\partial(\varepsilon_b)}{\partial U_{AB}}\right]^2 \cdot u_B^2(U_{AB}) = 0.0022 \cdot 10^{-12}$
$\left[\frac{\partial(\Delta T)}{\partial J}\right]^2 \cdot u_B^2(J) = 0.27 \cdot 10^{-9}(K)^2$	$\left[\frac{\partial(\varepsilon_b)}{\partial J}\right]^2 \cdot u_B^2(J) = 3.73 \cdot 10^{-19}$
$\left[\frac{\partial(\Delta T)}{\partial R_{10}}\right]^2 \cdot u_B^2(R_{10}) = 1.09 \cdot 10^{-3}(K)^2$	$\left[\frac{\partial(\varepsilon_b)}{\partial R_{10}}\right]^2 \cdot u_B^2(R_{10}) = 0.003 \cdot 10^{-12}$
	$\left[\frac{\partial(\varepsilon_b)}{\partial k}\right]^2 \cdot u_B^2(k) = 9.4 \cdot 10^{-12}$
$u^2(\Delta T) = 1.09 \cdot 10^{-3}(K)^2$ $u(\Delta T) = 0.03K$	$u^2(\varepsilon_b) = 0.0052 \cdot 10^{-12}$ $u(\varepsilon_b) = 0.072 \cdot 10^{-6}$
$\Delta T = (29.00 \pm 0.03)K$	$\varepsilon_b = \left(0.001 \pm 0.072 \cdot 10^{-6}\right)$

6. CONCLUSIONS

The experimental design of a double current bridge confirms that the Warsza's idea of a double current supply 4T circuits can be a supplement for well-known techniques of conditioning analogue signals. This way of supplying is technically more difficult than in a classic bridge. The advantage of this circuit is the way of temperature compensation, without additional temperature sensor or external resistors.

In case of force, stress, and pressure measurements, the mechanical elements are constructed as specially designed cantilever beams. The method of 2-parameter measurement with supplying from double current sources can be applied for semiconductor strain gauges and in Micro-Electro-Mechanical Systems (MEMS), where mechanical elements and electronics are integrated.

An alternative idea to 2J bridges is the method of Pallas-Areny and collaborators [12], based on detecting changes in resistance by measuring the time needed to discharge a given capacitor.

The procedure of uncertainty calculation was conducted in a classic way. Standard uncertainties of temperature and strain measurement are not higher than 0.1%.

7. ACKNOWLEGMENTS

This work was done in the frame of project W/WE/6/07.

8. REFERENCES

1. Z. L. Warsza: *Bridges Supplied by Two Current Sources-New Tool for Impedance Measurements and Signal Conditioning*, Proc. of IMEKO TC-7, 9th Symposium in Cracow (Poland), pp. 231-236, 2002.
2. Z. L. Warsza: *Backgrounds of two variable (2D) measurements of resistance increments by bridge cascade circuit*, Proc. of SPIE vol. 6347 part 2, 63472R, 2006.
3. Z. L. Warsza: *Two Parameter (2D) Measurements in Four Terminal (4T) Impedance Bridges as the New Tool for Signal Conditioning part 1 and 2*, Proc. of the 14th International Symposium on New Technologies in Measurement and Instrumentation and 10th Workshop of IMEKO TC-4, pp. 31-42, 2005.
4. Z. L. Warsza: *Two Parameter (2D) Measurements in Double – current Supply Four-terminal Resistance Circuits* Metrology and Measurement Systems vol. XIII, no 1 pp. 49-65, Polish Academy of Science, 2006.
5. Z. L. Warsza: *Four-Terminal (4-T) Immitance Circuits (4T) in Multivariable Measurements*, Industrial Institute of Control and Measurement PIAP, Warsaw 2004 – (in Polish).
6. A. Idźkowski, J. Makal, Z. L. Warsza: *Application of Double Current Bridge-Circuit for Simultaneous Measurements of Strain and Temperature*, IEEE Instrumentation, Measurement and Technology Conference, (IMTC 2007), Warsaw, Poland, pp. 1-4, 2007.
7. MAX1450, low cost 1%-accurate signal conditioner for piezorezistive sensors, Maxim product data sheet.
8. *Sensor circuits and digitally controlled potentiometers*, Intersil application note AN135, 2005.

9. *Sensor temperature compensation using the four DAC signal conditioning architecture*, Maxim application note 1839, 2002.
10. C. Swartz, C. Derrington, J. Gragg: *Temperature Compensation Methods For The Motorola X-ducer PressureSensor Element*, Motorola Semiconductor application notes.
11. *NPC-1210 series medium pressure sensor*, Product data sheet of GE NovaSensor Inc.
12. J. Jordana, R. Pallas-Areny: *A simple efficient interface circuit for piezoresistance pressure sensors*, Sensors and Actuators A-Physical vol. 127, Issue 1, 28, pp. 69-73, 2006.

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 405–414

Classification of the Space Domain Decomposition in the Parallel FDTD Implementation

WOJCIECH WALENDZIUK

Faculty of Electrical Engineering,
Bialystok Technical University,
Wiejska 45D, 15-351 Bialystok
e-mail: walenw@pb.edu.pl

Received 2008.06.04
Authorized 2008.08.03

The paper presents the approach to the space decomposition classification used in the Finite Differences Time Domain method. Proposed classification system simplifies the way of the identification of particular cases of the space decomposition in the Cartesian coordinate system. The work also contains the speedups results achieved with the use of a cluster system, for different type of a three – dimensional space decomposition.

Keywords: parallel computing, space decomposition, Finite Difference Time Domain method, electromagnetic field

1. INTRODUCTION

The method of finite differences in time domain is directly applied to the time form of Maxwell's equations. With the use of this method, curled equations of electromagnetic field are solved simultaneously in time and space domains. The basic system of Maxwell's equations in the vector form for isotropic environment can be presented as follows:

$$\nabla \times \vec{E} = -\mu \frac{\partial \vec{H}}{\partial t} \tag{1}$$

$$\nabla \times \vec{H} = \sigma \vec{E} + \varepsilon \frac{\partial \vec{E}}{\partial t} \tag{2}$$

The FDTD method is based on the algorithm created by K.S.Yee who formulated the differential form of Maxwell's curled equations in 1966. In his article [1] he presented an elementary three dimensional cell of cubic shape. Such cell contains nodes of a finite differences grid in which particular components of electric and magnetic field intensity vectors are placed. An area analyzed by the FDTD method is filled with those cells. Their number inside a numerical model is conditioned by the size of the analyzed area.

The bigger the number of Yee cells (Fig. 1.) filling the mentioned area, the greater the number of nodes of the finite differences grid, which results in growing correctness of analysis of the numerical model.

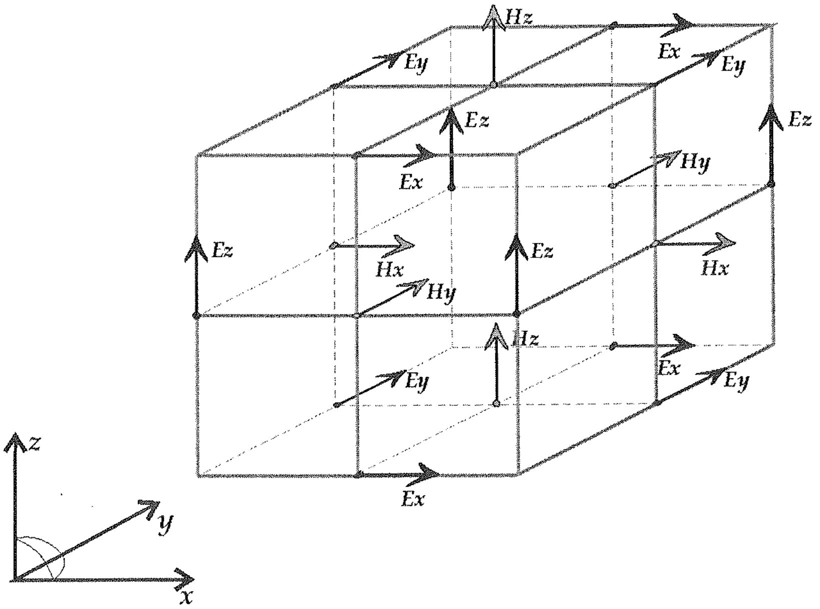


Fig. 1. Three – dimensional Yee cell

It can be easily noticed that all components of the electric field intensity $\vec{E}(E_x, E_y, E_z)$ are located in the centre of edges of particular surfaces of Yee cells. Components of magnetic field intensity $\vec{H}(H_x, H_y, H_z)$, however, are placed in the middle of surfaces forming it. It is worth stressing that the general idea of Maxwell's curled equations, (1) and (2), perfectly presents the fact that each component of the electric field vector is surrounded by proper components of magnetic field circulating around it. Moreover, components of the electric field vector surround a proper component of magnetic field. This structure helps to keep the unity of independent Maxwell's curled equations, describing electromagnetic wave propagation [2,3].

The FDTD is usually used as a sequential method where algorithm works on a single PC unit. Unfortunately, in the cases of numerically large areas, overcoming the capacity of memory of a single computational node, particular algorithms, enabling parallel computations (e.g. in a cluster system), should be used [5,6,7].

In this work we present a method of space decomposition enabling parallel transformation. It is based on decomposition of finite differences grid into smaller parts which are later transmitted to separate computational units. During computations working nodes are connected by a communicational network exchanging appropriate data necessary to keep the integrity of a global finite differences grid.

Algorithms composed according to this idea, usually base on the appropriate division of the computational area into subareas, as well as on distribution of interprocessor communicational tasks. The division of the computational area can take various forms, depending on the needs deriving from the type of computational task or the shape of the analyzed numerical model.

It is worth stressing that in all cases of research the Mur's absorbing boundary conditions were used on the outer surfaces of the analyzed area [4].

This work presents also a proposition of a way of signing particular types of space decomposition which makes their identification easier. Additionally, we present the results of speedups achieved for three – dimensional examples of space decomposition. The analyzed cases were done with the use of the "MORDOR" cluster system which is located at Bialystok Technical University. In computations were used nodes equipped with 64-bits Intel Xeon 3.2 GHz processor with 2 MB of cache memory. Particular nodes were also equipped with 2 GB operational memory [5].

2. SPACE DOMAIN DECOMPOSITION

Space decomposition, mentioned in the introduction, uses the division of a finite differences grid containing information about material data of the analyzed environment into smaller parts. They will later undergo computations in separate computational nodes. The choice of the division depends mostly on the way of observation of computations results. This can be described as follows: before the simulation is done, the type of decomposition can be chosen in such a way that the achieved data came from one computational node without the necessity of integrating it from a few nodes. Analyzed data collected from proper cross-sections of the examined object is then sent to the supervising node where it can undergo further transformation (post-processing). Because of a great number of examples of space decomposition in the Cartesian coordinates system, we propose to use the following system of signs:

$$R^{(2,3)}D^{(1,2,3)}_{(x,y,z)} \tag{3}$$

where:

R – symbol of the space type (e.g. two-dimensional or three-dimensional model),

D – type of decomposition (e.g. according to one, two or three axis of the Cartesian coordinate system), along O_x, O_y, O_z axis.

2.1. TWO DIMENSIONAL DECOMPOSITION – R^2

In the cases of the two-dimensional analysis (*TM-Transverse Magnetic, TE-Transverse Electric*), two types of decomposition can be mentioned:
- decomposition into stripes, marked as R^2D^1,
- decomposition into rectangles R^2D^2.

2.1.1. R^2D^1 DECOMPOSITION

R^2D^1 decomposition is based on the division along one of Cartesian axis. Three cases can be observed here:
- along axis ($R^2D^1_x$),
- along axis ($R^2D^1_y$),
- along axis ($R^2D^1_z$).

In this decomposition one edge has a constant length, and the other one is divided into fragments corresponding with the sizes of subareas. The width of the sub-areas depends on the type of decomposition. There can be two types: homogeneous [8] and heterogeneous [9] decompositions. In the homogeneous decomposition, for example in the case of decomposition along Ox axis, $x1, x2, x3$ have identical values ($x1 = x2 = x3$). Whereas in the heterogeneous decomposition they can take different values ($x1 \neq x2 \neq x3$). This type of decomposition can be presented as in Fig. 2.

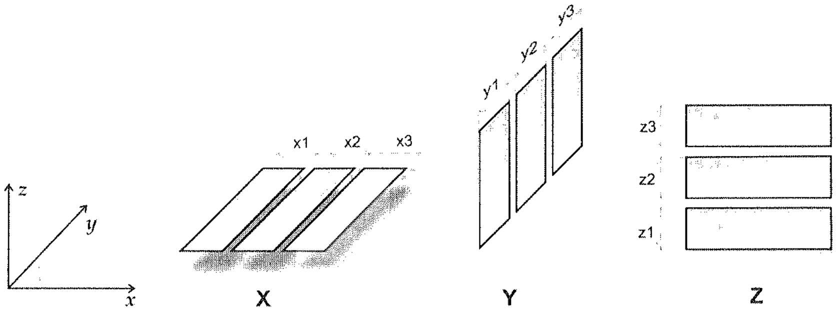


Fig. 2. R^2D^1 type of decomposition

It should be mentioned that the sub-areas marked by rectangles are represented in reality by a finite differences grid filling their areas.

2.1.2. R^2D^2 DECOMPOSITION

This type of decomposition differs from the previous one with the fact that in a two-dimensional area R^2 the decomposition is done along two chosen Cartesian axis (Fig. 3.). Similarly to the previous case (R^2D^1), the widths and the lengths can be regulated according to the type of decomposition (hetero- or homogeneous).

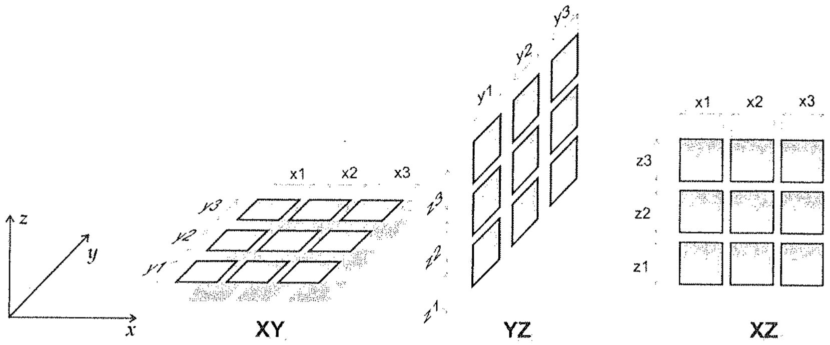


Fig. 3. R^2D^2 decomposition

2.2. THREE-DIMENSIONAL DECOMPOSITION R^3

The most correct visualization of a real object undergoing the numerical analysis can be achieved with the use of numerical simulation of a three-dimensional space. The decomposition of such area is far more complicated if compared with two-dimensional cases. The finite differences grid contains six components of electromagnetic field. They are most frequently recorded in computer's memory as separate tables of numerical variables. therefore, special attention should be paid to the places where space decomposition occurs. All components of electric and magnetic field intensity must be located inside appropriate sub-areas in such a way that the continuity of the finite differences grid is kept according to the structure of an elementary Yee cell. In the analysis of such three-dimensional area, two types of decomposition can be named:

— slice decomposition (in R^3D^1case),
— stripe decomposition (in R^3D^2 and R^3D^3cases).

2.2.1. R^3D^1 DECOMPOSITION

Space decomposition of the R^3D^1 ($R^3D^1_x$, $R^3D^1_y$, $R^3D^1_z$) type (Fig. 4.) is similar to the R^2D^1 type of decomposition of the computational area. The most important difference is the fact that the dividing is done on a finite differences grid of a three-dimensional space, not – in a two-dimensional one.

The algorithms decomposing the global area into sub-areas are then very similar, too. What differentiates them is the additional dimension according to one of Cartesian axis. The sizes of sub-areas are also conditioned by one of the axis of decomposition (e.g. when the decomposition is done along the *Ox* axis, the values $y = const$, $z = const$ are determined by the sizes of the finite differences grid, and $x = var$ is variable and it depends on the type of decomposition – homogeneous or heterogeneous).

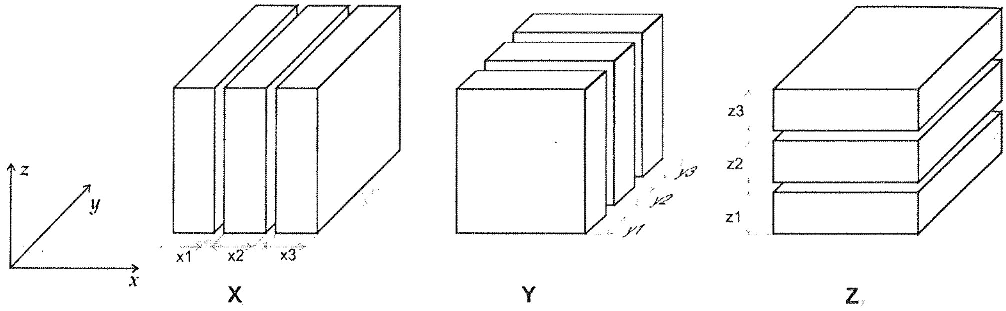


Fig. 4. R^3D^1 decomposition

The achieved speedups (Fig. 5.) for various densities of finite differences grids oscillated around the value of 3 for 6 computational nodes. Only for the smallest density (50×50×50) of Yee cells the speedup dropped below the value of 1. This means no benefit from using the parallel method in the computational process. This phenomenon was caused by an inconvenient coefficient of a long time devoted to communication to the time of computations.

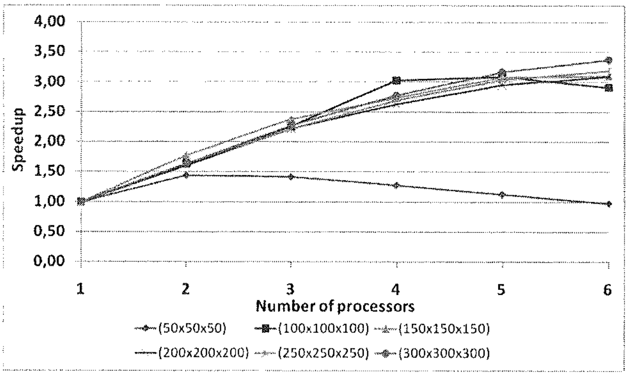


Fig. 5. The achieved speedups results for R^3D^1 decomposition

2.2.2. *R^3D^2 decomposition*

This type of decomposition (Fig. 6.), marked as R^3D^2 ($R^3D^2_{xy}$, $R^3D^2_{yz}$, $R^3D^2_{zx}$), takes into consideration two axis of coordinates, according to which a decomposition is done. Such case can appear convenient when the analyzed object, for example a dielectric, can be located in one segment of a sub-area. Then, there is no necessity to integrate the finite differences grid in order to analyze for example one of the components of the electric field. The supervising computer receives, in this case, only data from that computational node in which a given object is located.

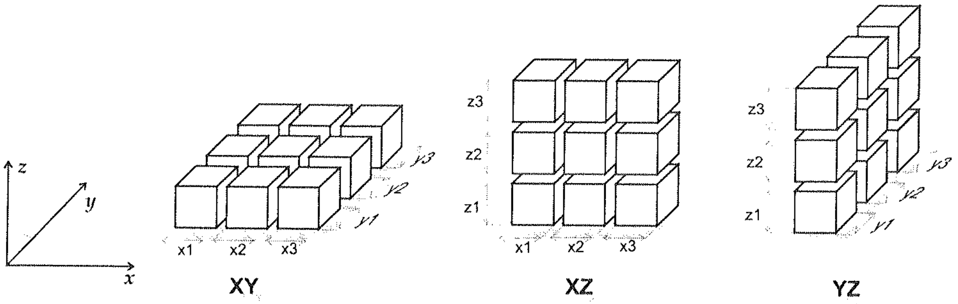


Fig. 6. R^3D^2 decomposition

Using this type of decomposition, even in spite of a bigger number of faces of communication, resulted in a bit better speedup of the algorithm for 6 computational processors, in relation to the R^3D^1 decomposition (Fig. 7.). This was caused by periodic data exchange in the 2×3 system of processors, according to the *Ox* axis of processors (1,3,5), and simultaneously – (2,4,6). Later, in the second cycle, the data exchange was fully parallel according to the *Oy* axis between processors (1,2), (3,4) and (5,6). This decomposition of data exchange in a fast Ethernet grid of 10 Gb/s caused only a sight speedup of the algorithm in relation to the previous case.

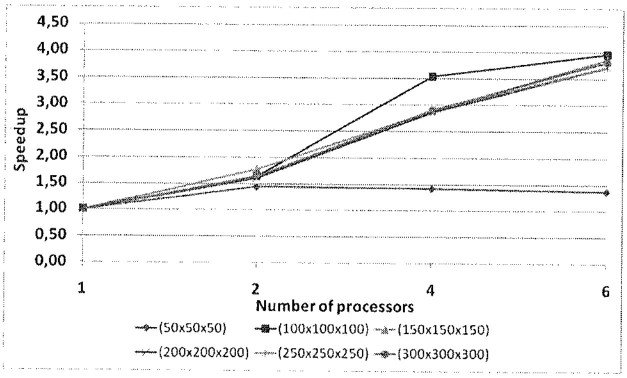


Fig. 7. The achieved speedups results for R^3D^2 decomposition

2.2.3. *R^3D^3 DECOMPOSITION*

The most complicated of all analyzed cases is decomposition of the R^3D^3 ($R^3D^3_{xyz}$) type (Fig. 8.). Surfaces dividing the analyzed area are located on all axis of a coordinate system. In the cases of the homogeneous decomposition, each area takes the shape of rectangles (all of them are identical). This type of decomposition appears to be efficient when the computational data assigned e.g. to visualization would come only from one computational node.

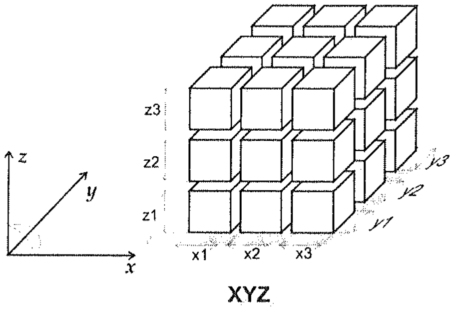


Fig. 8. R^3D^3 decomposition

Using this type of decomposition results in a significant decrease of speedup because of a big number of communication surfaces (Fig. 9.). In the case (200×200×200), (250×250×250) and (300×300×300) decrease of speedup below the value of 1 can be noticed. This prolongs the working time of the parallel program in comparison with its sequential version.

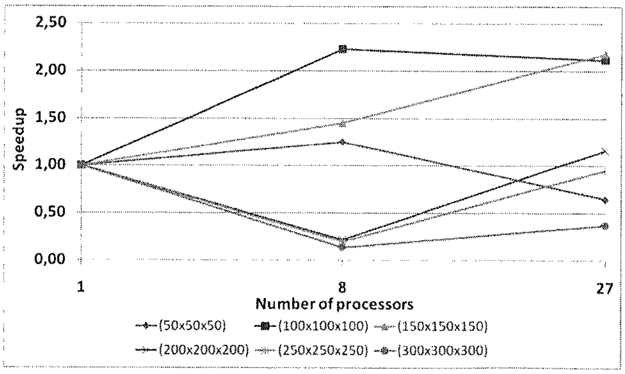


Fig. 9. The achieved speedups results for R^3D^2 decomposition

3. CONCLUSIONS

The presented way of identification of particular types of space decomposition in the Cartesian coordinates system can make an interesting proposition for common use. Applying this system of classification make discussing of proper cases easier, without unnecessary, descriptive discussion of a type of decomposition. It must be stressed that the proposed identification of a space decomposition can be applied not only in the FDTD method, but also in other numerical methods in which the analyzed area is decomposed in the Cartesian system.

This work presents also the results of the achieved speedups of parallel algorithms of the FDTD method. It appears that the value of speedup for particular cases of various densities of the finite differences grid and types of decomposition depends on the way of working out the algorithm of data exchange between computational nodes, algorithm implementation in a multiprocessor system, computational nodes efficiency and the Ethernet bandwidth.

The presented results achieved in the "MORDOR" cluster system can be considered as satisfactory, however not excellent. Space decompositions of R^3D^1 and R^3D^2 types resulted in similar speedups of the approximate value of 3. The R^3D^3 decomposition appeared the least efficient, but it could be applied in the cases where the visualization data would be located and computed in one the computational nodes. This solution would decrease the amount of data sent to visualization.

4. ACKNOWLEGMENTS

This work was supported by BTU grant S/WE/3/08.

5. REFERENCES

1. K. S. Yee: *Numerical Solution of Initial Boundary Value Problems Involving Maxwell's Equations in Isotropic Media*, IEEE Transactions on Antennas and Propagation, vol. AP-14, No 3, pp. 302-307, 1966.
2. A. Taflove: *Computational Electrodynamics. The Finite-Difference Time-Domain Method*, Artech House, Inc., 1995.
3. K. S. Kunz, R. J. Luebbers: *The Finite Difference Time Domain Method for Electromagnetics*, Boca Ra-ton, CRC Press, Inc. 1993.
4. G. Mur: *Absorbing Boundary Conditions for the Finite Difference Approximation of the Time-Domain Electromagnetic-Field Equations*, IEEE Transactions on Electromagnetic Compatibility, Vol. EMC-23, No. 4, pp. 377-382, 1981.
5. W. Walendziuk: *Parallel Computation of the SAR Distribution in a 3D Human Head Model*, Photonics applications in astronomy, communications, industry, and high-energy physics experiments, Proceedings of SPIE, Vol. 6937, 2007.

6. W. Walendziuk, J. Forenc, A. Jordan: *High Frequency Electromagnetic Fields Analysis With the Use of the Parallel FDTD Method*, Springer-Verlag, Berlin, Lecture Notes in Computer Science, vol.3019, pp. 528-535, 2004.
7. W. Walendziuk: *Parallel Techniques for Solving Maxwell's Curl Equations With the FDTD Method*, XVII International Conference: Electromagnetic disturbances EMD'2007, Bialystok, pp.91-94, 2007.
8. W. Walendziuk, J. Forenc: *Decomposition of the Computation Area in the Parallel Implementation of the FDTD Algorithm*, ISPDC'2002, Informatica Tomul XI, July 17-20, 2002, "Al.I.Cuza" University of Iasi, pp. 180-192, 2002.
9. W. Walendziuk, S. Kwieckowski: *The Use of the Dynamic Space Decomposition Algorithm of a Computational Area in Heterogeneous Cluster Computations*, PARELEC'2006, IEEE International Symposium on Parallel Computing in Electrical Engineering, IEEE. Los Alamitos, pp. 466-469, 2006.

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 415–427

Copper electrocrystallization in the presence of different kind of modeling electrodes

ANNA MARIA BIAŁOSTOCKA, ŁUKASZ ZANIEWSKI

Electrical Faculty, Bialystok Technical University,
ul. Wiejska 45D, PL 15-351 Bialystok, Poland
e-mail: bialan@pb.edu.pl

Received 2008.07.04
Authorized 2008.08.27

The electrochemical method of the metal setting plays important role in the intensive developing domain, which is materials engineering. The effects of electrochemical process working are the surface layers, which have the differences in the cristal stucture. Feature, which characterize this kind of layers is homogeneity of the structure and high level of adhesion to the base. Microstructure of the surface layers can be shaped with the help of electrical parameters selection during the electrochemical process, particularly current parametres as well chemical parametres. In the article is presented the method of electrical field shaping and setting layers influence. The researches present also the usability of steering of the electrical field distribution depend on the shape of the additional, modelling electrodes.

Keywords: electrochemistry, electrochemical process, Copper setting, Finite Elements Method, electric field shaping

1. INTRODUCTION

Metals produced electrochemically are applied specially as surface layers or negative forms. They are tools for elements creation by thermoplastic method processing. Dynamic development of technology and plastic industry is fensed by decreased plastic elements influence in the economic world. It has direct connection to necessity of replacing. Light elements made from artificial stuff replace heavy metallic elements. For this situation inquires into many branches of industry for example in moto industry or in the industry which produced tools and elements home load.

Plastics, like all constructive materials, are characterized by group of features which are advantages and disadvantages of the final product. The most important disadvantages of plastics, which made them different from other constructive materials, are low operation temperatures. These disadvantages do not get permission to eliminate plastics from applying in many branches of management, where they work in relative low temperatures. Plastics have many advantages, like easy forming, good endurance, strenght of the outer features and estetic features. Possible of getting high standard mould guaranted the creation of the good standard product, which will satisfy the needs of the purchasers. Many numbers of features have influence on the material's structure, which are made by the electrochemical method. These features are for example electrolyte contains, parameters, current type, environmental temperature and kind of electrodes setting. During electrodeposition we may control some numbers of features and by this control the quality of deposited layers. Very important is the fact that it is good to know which kind of features has influence on the final effect.

The aim of the researches presented in the article is demonstration of some kind of method. This method makes possible influence on deposited layers properties by the appriopriate electric field shaping near the covering element.

2. COPPER ELECTROCRYSTALLIZATION PROCESS

Basis for carrying out the process is current flow forced by energy source (Fig.1.), which is connected to the proper system of the electrodes sinking in the electrolyte. Electrolyte, it is multicomponent solution, which contains two-positive Copper ions – Cu^{2+}. During this process cathode is formed by the product on which surface layer is deposited, anode is made of metal which will be deposited (Copper).

In currentless coditions, ions in the electrolyte solution are under inner thermal motion state. Flow of the electric current is realiability thanks to both of signs, free ions presence in the electrolyte. But voltage applying makes possible ions motion in the electrolyte, then flow of the current and course coupled electrochemical reaction on the electrodes surfaces. Negative and positive ions in the electrolyte, similar to electrones in the metal, are charges carrier.

Inside the electric field negative ions (anions) go in the anode direction, which is connected to positive pole of the current source. Positive ions (cations) go in the cathode direction, which is connected to negative pole of the current source [1].

On the cathode surface follows reduction of positive copper ions to atomic forms with agreement to the reaction:

$$Cu^{2+} + 2e \rightarrow Cu^{0} \tag{1}$$

Simultaneously, on the anode's surface goes opposite process, i.e. oxidation of copper atoms to positive ions Cu^{2+}, and this process is described by the equation:

$$Cu^{0} - 2e \rightarrow Cu^{2+} \tag{2}$$

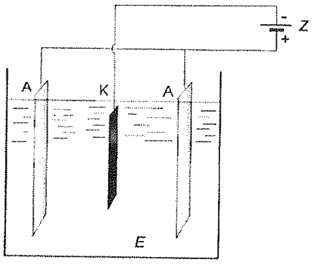


Fig. 1. Scheme of the electrochemical arangement for the electrodeposition: E – electrolyte solution, Z – source, A – anode, K – cathode

Processes on the two electrodes are coupled with each other. They take place with the same velocity and they are the reason of dynamic stability state in the electrochemical environment. The result of the oxidation and reduction reactions is copper fixing on the cathode and the anode dissolving [4].

Analyzed process of the copper layer deposition is limited by electric charge changing. During that kind of processes current density J becomes as a sum of anode current density and cathode current density and is determined by Butler-Volmer relation as follows [1]

$$J = J_0\left[\exp\left(\frac{(1-\beta)nFv}{RT}\right) - \exp\left(\frac{-\beta zFv}{RT}\right)\right] \tag{3}$$

where: v-overpotential; J_0 – changing current density between anode and cathode processes, in the state of stable dynamic equilibrium, when v=0; β-cathode charge transfer coefficient, which is measure of symmetry between anode and cathode processes, depends on kind of electrolyte; F-Faraday's constant [96500 Cmol^{-1}]; R-molar gas constant [8.314 Jmol^{-1}K^{-1}]; T-temperature [K]; n-number of ions which take place in oxidation and reduction processes.

Two fundamental processes have an influence on the effect of electrocrystallization. The shape of the crystalline setting layer to be precise velocity of the new embryos creation and crystal growth velocity depends on these two parallel processes. The course of processes determines metal ions transport from electrolyte depth to cathode surface and then ions reduction to atomic shape. The fundamental phase, which determines the crystalline structure and quality of setting metal is creation of crystal centres - crystal embryo and its growth.

On this stage of the electrocrystallization process important role play controls like: kind of crystal phase, kind of electrolyte, kind of ground material, its structure and parameters of realized process - current density.

Copper surface layer thickness determined the dependence [1]:

$$d = \alpha \cdot \frac{MIt}{zF\rho A} \tag{4}$$

where: M – mols mass of the fixing metal, z – number of electrones participating in the process of oxidation or reduction, F=96485 C/mol – Faradays constant, I – validity of current [A], t – time of fixing process, ρ – density of the fixing material, A – field of active cathode surface, α – cathode productiveness.

3. COPPER'S LAYER APPYING

Metals manufactured electrochemically are applied in the wide range as layers setted on final product's surfaces to modify them functional quantities. Present needs of microelectronics technology, electrooptics and electrotechnics materials engineering could be satisfjed as results of thin layer using, notones nanometric dimension. These needs permit to do the minimalization of elements, electronic components and then they get permission to achieve higher development in domain of integrated circuits applications. It visualizated also tendency to layers creation on final products surfaces. These layers allow to shape conservation in good way. It concerns not only surface layers deposited on metallic goods but also layers deposited on plastic goods.

Plastics are today very important group of constructive materials. Heavy metallic elements replacement of lightweight plastic elements in everyday practice happened common in moto industry, machine industry and home device. It was implemented plastic applying to auto's elements creation. Today, there are bumpers, carrying bodywork elements, supplying systems, electronic aid, electric fixtures elements and optics projector elements made from plastic. Negative forms are tools, which allow to achieve demanded shape and accuracy of elements created by thermoplastic method plastic precreation. Quality and accuracy of the form making decides in many cases not only about detail's quality, but about final product's quality too. Direct role during these negative form creation play electrochemical methods [3].

4. ELECTROCRYSTALLIZATION PROCESS MODIFICATION

During the researches in the measurement system (Fig.2.) there are carried the measurements of the potential distribution out by keeping constant current's value. The results are presented as a diagram of potential distribution on XZ cathode's surface (covered object) along Z axis, which is electrochemical bath's width (Fig.3.) [5].

There is visible violent increase of the potential cathode's edge. It was made use of dependance [6]:

$$E = -gradV \tag{5}$$

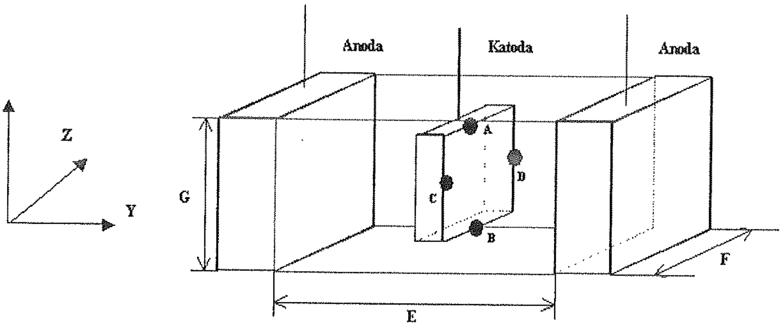


Fig. 2. Scheme of the electrodes's arrangement in the measurement system

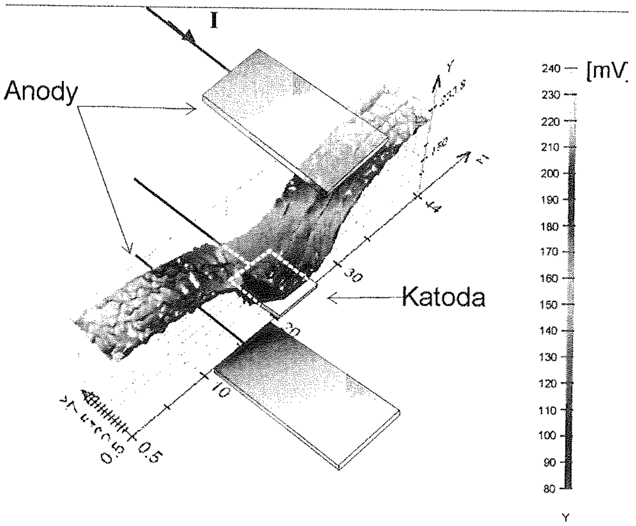


Fig. 3. The potential distribution on the XZ cathode's surface, made by computer aided measurement system coupled with microprocessor system of measurement probe travelling (measure along Z axis)

(the assumption was accepted, that vector *E* is the side of decreased potentials directed) and values of the electric field intensity at these points were received diversified. Higher values of the electric field intensity on the covered object edges decide about "edge effect" creation. "Edge effect" could be decreased after the conditions creation, which make uniform the electric field intensity around the covered object. There was applied the disturbance of the electric field distribution in order to set uniform metal layer on the cathode's surface. The disturbance was realized after the additional conductive elements applying in direct cathode's neighbourhood of the measurement system. These elements have the cathode potential. Based on the measurement results, actions were taken directed to modificated structure of the measurement system creating. Additional elements position will be changing (they will be called mobile elements) and connected to this geometry of the measurement system will be changing too.

Based on physical phenomenon, which proceed in electrochemical systems volume, the preliminary additional electrodes location in the area bewteen two anodes was decided (Fig.2.). The end effect this kind of modification of the electrochemical system should be the electric field intensity vectors, which values are similar on the whole cathode's surface and directed perpendicular to it. Connected to this the iam functions were taken under attension [6], [7]:

$$\| E_{xz} \|_1 = | E_{xz} | = \min \sqrt{E_x^2 + E_z^2} \qquad (6)$$

5. INVESTIGATION METHODS

In order to determination of the spacial additional electrodes location, which increase uniform of distribution of the electric field intensity around covered object, it was created numerical model of the electrolitical bath with preliminary location of the additional electrodes. Numerical simulation was achieved based on American Simulation Program E.M.R.C. NISA/EMAG. Author's program collaborated with upper licence program made possible automatic change of mobile elements according to aim functions values. Finite elements method applying is very common for simulation and analyze of electric field distribution. In order to acceleration and numer of results increasing the two-dimensional model was analyzed [8].

Model included additional structure simplifications. Deficiency of those simplifications will be making the time dilation incommensurable to gain benefits. It was for example: whole elements of system's geometry was cuboits, negligence of wires afferent voltage, acceptance of primary current's density division. Authors focus on these kind of model's quality, which values make view of the electric field distribution (electric field potential V, electric field intensity E) [8].

One of the input files included basic electrochemical system's configuration, which are two anodes's positions coordinates and cathode's position coordinates. Those coordinates wasn't relocated. Other input file included 50 random versions of additional electrodes's locations. Each of them contained 48 single elements, which were relocated. Examples of mobile elements's arrangement with electric field intensity distribution are shown as Fig.4 and Fig.5. Pictures presented program's possibilities, which concern mobile parts electrochemical system variations. Those locations appear as black elements in the net. Receiving electric field intensity distribution is uniform, then compatible to anticipation. Very problematic was complicated mobile elements's location and connected to this trouble with experiment realization. It was choosen optimal location's version when we taken electric field intensity distribution and mobile elements's locations under consideration. Results of numerical simulation as electric field intensity distribution with optimal additional electrodes's arrangement are showed as Fig.6 and Fig.7. It was taken six groups of points under consideration (Fig.8). Each group was one additional electrode and was eight single points fusion. Configuration was verified in real system for Copper electrodeposition [8].

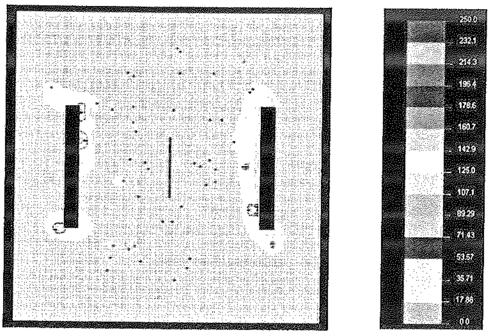


Fig. 4. The electric field distribution of the numerical model 2D with random arrengement of mobile elements

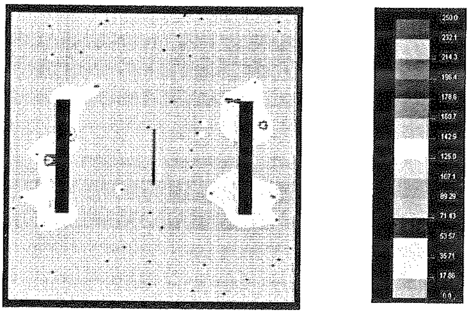


Fig. 5. The electric field distribution of the numerical model 2D with random arrengement of mobile elements

Results of numerical simulation were verified during the real time experiment of Copper electrochemical setting. Basis to conduct research was work place in Fig.9. Results of made computer-aided optimization was additional electrodes's introduction, which were situated in cathode's environmental. Additional electrodes with cathode's potential (0V) and dimensions 8 × 5 × 70 *mm* were situated in the real electrochemical system. Shapes of applied additional electrodes and samples are shown as Fig.10 and Fig.11 [8].

6. RESULTS OF THE EXPERIMENTS

The iam of carried researches out was modification's determination influence of the electric field distribution during electrodeposition process. It was achieving by making six additional elements witch cathode's potential to basis electrochemical system

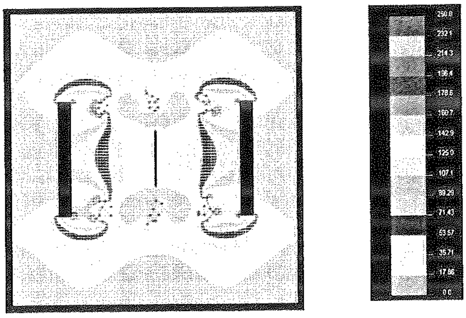


Fig. 6. The electric field distribution of the numerical model 2D with arrengement of mobile elements applied during real experiment

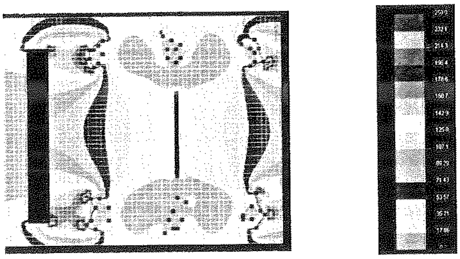


Fig. 7. The electric field distribution of the numerical model 2D with arrengement of mobile elements applied during real experiment (focus on the space near the cathode)

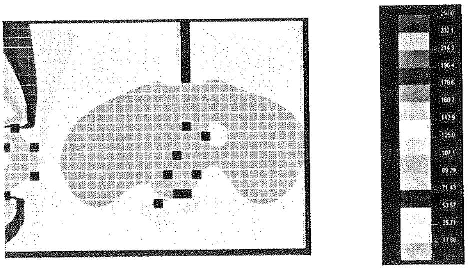


Fig. 8. The electric field distribution of the numerical model 2D with arrengement of mobile elements applied during real experiment (focus on the space near the edge of the cathode)

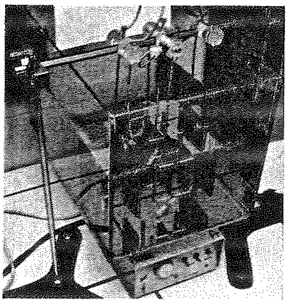


Fig. 9. Set-up for electrochemical experiments

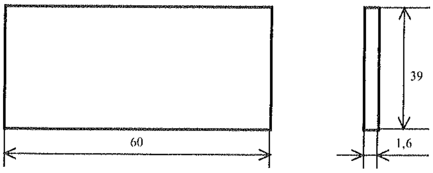


Fig. 10. Probes dimensions applied during electrodeposition process [mm]

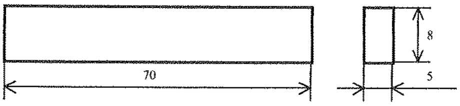


Fig. 11. Additional electrodes dimensions applied during electrodeposition process [mm]

(two anodes – cathode). Surface geometry, morphology and setting layers's structure were influenced by kind of element's cathodic "screnning" without negative influence of electric field force lines. Control of the additional electrodes arrangement during the electrodeposition process could makes layers with different graines's sizes. This arrangement decides about "edge effect" level of deposited Copper layers (Fig.12., Fig.13.) [8].

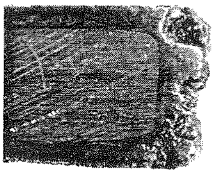


Fig. 12. Structure of deposited copper layers made by electrocrystallization method in basic system without additional electrodes

Results of surface's geometry analyze of deposited layers are showed as Fig.14. Fig.15. and Tab.1. Tab.2. Lower roughness irregularity of surface demonstrated Copper layers profile, which was deposited in system with additional electrodes. Results confirm roughness parameters's values. Derating of all roughness parameters is visible

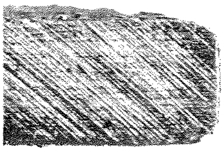


Fig. 13. Structure of deposited copper layers made by electrocrystallization method in modificated system with additional electrodes

(Tab.1, Tab.2) after additional electrodes arrangement applying in the electrochemical system. Single parameters's values: $R_a, R_g, R_v, R_t, R_p i R_z$ of Copper layers receiving in modificated system are five-times smaller than layers created in basis system. It is the activities's correctness confirmation, which consists of electrochemical system modification, additional elements applying with cathode's potential and located in precise position near cathode [8].

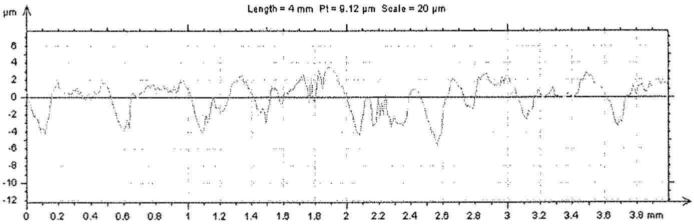


Fig. 14. Profile of deposited copper layers made by electrocrystallization method in basis system without additional electrodes

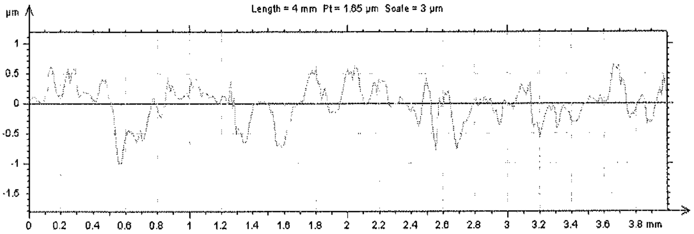
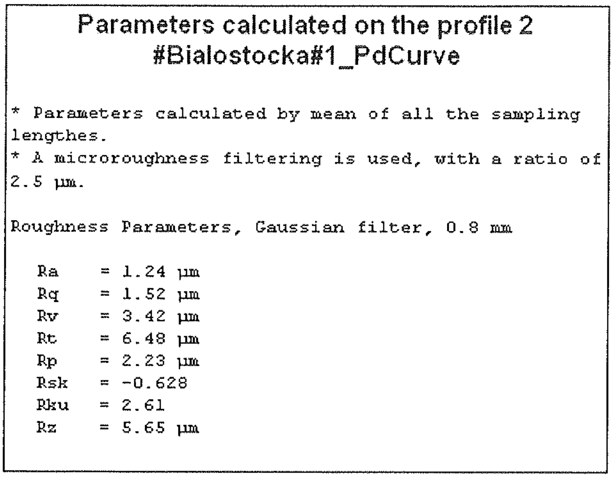


Fig. 15. Profile of deposited copper layers made by electrocrystallization method in modificated system with additional electrodes

Table 1

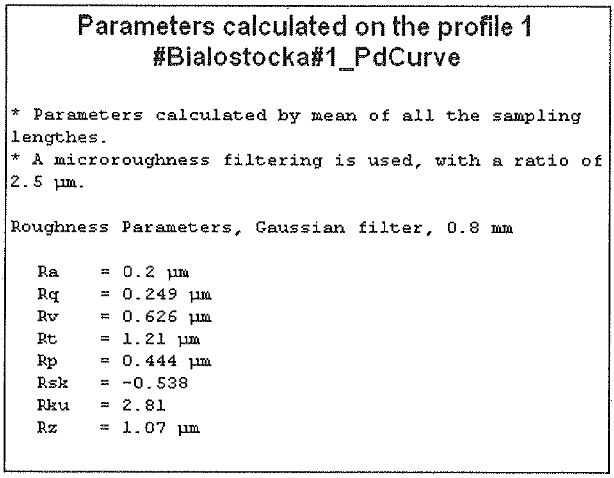
Roughness parameters of deposited copper layers made by electrocrystallization method in basis system without additional electrodes



```
Parameters calculated on the profile 2
#Bialostocka#1_PdCurve

* Parameters calculated by mean of all the sampling
lengthes.
* A microroughness filtering is used, with a ratio of
2.5 µm.

Roughness Parameters, Gaussian filter, 0.8 mm

  Ra    = 1.24 µm
  Rq    = 1.52 µm
  Rv    = 3.42 µm
  Rt    = 6.48 µm
  Rp    = 2.23 µm
  Rsk   = -0.628
  Rku   = 2.61
  Rz    = 5.65 µm
```

Table 2

Roughness parameters of deposited copper layers made by electrocrystallization method in modificated system with additional electrodes



```
Parameters calculated on the profile 1
#Bialostocka#1_PdCurve

* Parameters calculated by mean of all the sampling
lengthes.
* A microroughness filtering is used, with a ratio of
2.5 µm.

Roughness Parameters, Gaussian filter, 0.8 mm

  Ra    = 0.2 µm
  Rq    = 0.249 µm
  Rv    = 0.626 µm
  Rt    = 1.21 µm
  Rp    = 0.444 µm
  Rsk   = -0.538
  Rku   = 2.81
  Rz    = 1.07 µm
```

7. CONCLUSIONS

Researches presented in the article concern analyze of phenomenon proceed in electrochemical system during Copper layers electrodeposition, which could help in electric field optimization in the space between two anodes. Aim of realised research was condition creation for thickness of layers uniformity improvement after suitable aim function f_c applying. Planned aim achievement require fast optimize search method, getting electric field intensity distribution and additional electrodes's construction under consideration. There were analyzed many possible additional electrodes arrangements, the best – optimal – was laying out for experimental research. Layers deposited in system without additional electrodes are characterized by increased degree of surface development than layers deposited after additional electrodes applying. Optimal additional electrodes arrangement applying gets decreasing of Copper's surface roughness and deficiency of visible effects – called "edge effects" (Fig.13). Profilogram of surface irregularity demonstrated big change in values of single roughness perameters (Fig.14, Fig.15). Value of deposited layers's roughness in basis system is about $R_a = 1.24\mu m$ (Tab.1) and in modified system the same parameter is smaller – $R_a = 0.2\mu m$ (Tab.2). Differences are visible (5X) in all analyzed roughness parameters.

To sum up, researches's results presented in article give the answer, additional electrodes applying in electrochemical system for electrodeposition allow to conscious structure layers shaping and to control layers's properties. It is certificate – fixed electric field distribution and electric field shape have an effect on the Copper electrodeposition process and appear possibility of using all results in practice.

8. REFERENCES

1. M. Trzaska: *Elektrokrystalizacja metali.* Przegląd Elektrotechniczny, ss. 1166-1169, 11/2004 (in Polish).
2. M. Trzaska: *Możliwości kształtowania struktury i właściwości elektrolitycznych warstw niklowych.* Inżynieria Materiałowa, vol. 137, nr 6, ss. 581-584, 2003 (in Polish).
3. M. Trzaska: *Struktura i twardość warstw Ni, Cu, Co wytwarzanych elektrochemicznie,* Inżynieria Materiałowa, vol. 147, nr 5, ss. 698-700, 2005 (in Polish).
4. E. Jędrych: *Kształtowanie struktury i właściwości warstw miedzi metodą elektrokrystalizacji.* Praca dyplomowa magisterska. Politechnika Warszawska. Wydział Inżynierii Materiałowej. Warszawa 2006 (in Polish).
5. A. M. Białostocka, W. Walendziuk, S. Bolkowski: *Measurement of the distribution the electrical potential during the galvanotechnic process.* VIII Science – Technical Conference ZKwE'2003, Poznań/Kiekrz, pp. 193-195, volume I, 7-9.04.2003.
6. A. Białostocka, Ł. Zaniewski: *Optimization of the electrochemical method the cover's metallization,* VIII International Workshop for Candidates for a Doctor's Degree OWD'2006, Wisła-Kopydło, pp. 77-82, 21-24 October 2006.
7. M. Trzaska, A. Białostocka: *Electrodes modeling in the copper electrocrystllization process,* Computation Problems of Electrical Engineering, September 14-16, 2007 – Wilkasy, Poland, Przegląd Elektrotechniczny, nr 2, pp. 95-98, 2007.

8. A. Białostocka: *Metoda kształtowania pola elektrycznego w procesie elektrochemicznego osadzania miedzi*, Rozprawa Doktorska, Białystok 2008 (in Polish).

c
r
c
i

Sp
cascad
structio
to the s
tap, Fig
essenti
power
Splitter
Transm
describ
tenuatio

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 429–443

Analysis of the CaTV Splitterring Taps

DARIUSZ KRZEMIENIECKI, JAKUB JUCHNIEWICZ

Gdańskie Zakłady Teleelektroniczne TELKOM-TELMOR
80-425 Gdańsk ul. Mickiewicza 5/7
e-mail: dariusz.krzemieniecki@telmor.pl
jakub.juchniewicz@gmail.com

Received 2008.03.12
Authorized 2008.05.30

This article presents CaTV splitterring taps. Basic features of the splitterins taps and the influence of basic factor on crucial parameters were investigated in this article. Additionally, correction circuits for improvement of certain parameters were proposed. The research considered the two-way, three-way and eight-way splitterring taps. Some examples of applications of splitterring taps in passive devices used in CaTV network were presented in this paper.

Keywords: unidirectional transformer, basic configuration, inversed configuration, splitter, divider, matchingtransformer (autotransformer)

1. INTRODUCTION

Splitterring taps are common devices used in CaTV system as an alternative to cascade taps [11]. Both type of taps distribute CaTV signal to subscriber. Their construction is very simple: branched output of unidirectional transformer [5] is connected to the splitter [12] of relevant numbers of ways (Fig. 1 presents the two-way splitterring tap, Fig. 2 illustrates the three-way splitterring tap; dividers and autotransformers are essential elements of splitters [12]). This kind of implementation guarantees equal power division at branching outputs, what is difficult to achieve using cascade taps. Splitterring taps have typical for CaTV system working frequency band of 5-862 MHz. Transmission lines have the characteristic impedance of 75 Ω. Splitterring taps are described by following parameters: the main line attenuation IN-OUT, coupled line attenuation OUT-TAP, input reflectance IN, output reflectance OUT, branching (tapping) reflectance TAP, isolation between output and branch OUT-TAP, isolation between branches TAP A-B.

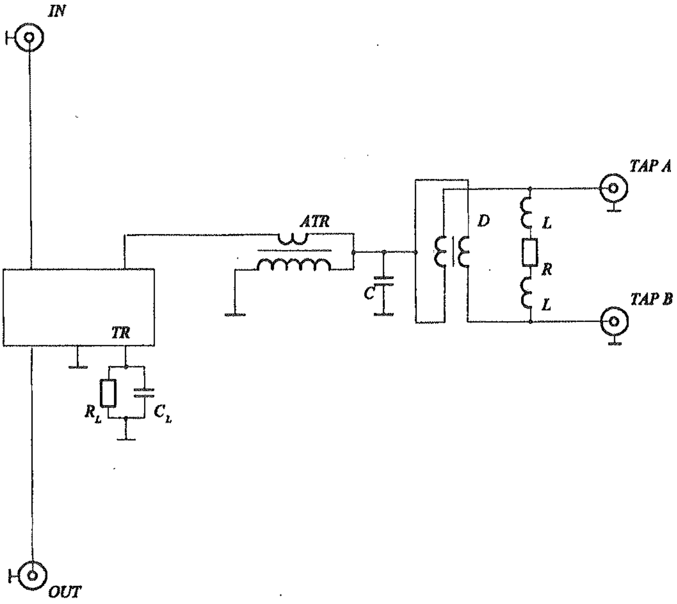


Fig. 1. The two-way splittering tap

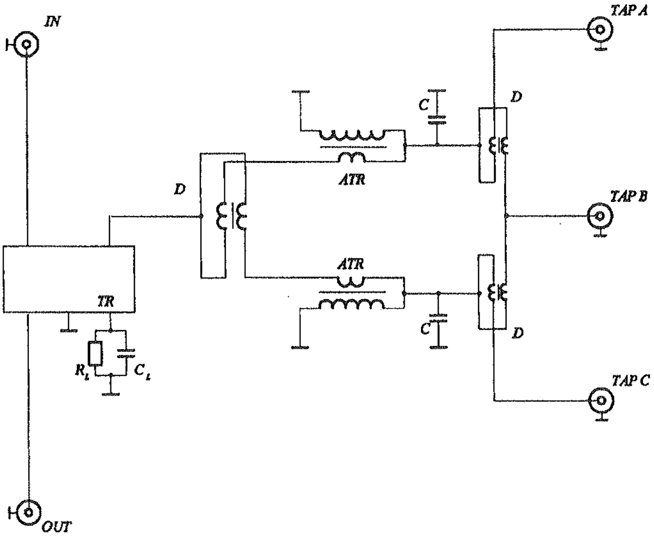


Fig. 2. The three-way splittering tap

Splitterring taps are band devices as the result of unidirectional transformers and autotransformers presence. Fig. 4, 5 illustrate IN-OUT and IN-TAP characteristics of the two-way splitterring tap versus wide band of frequencies.

2. THE CHARACTERIZATION OF THE SPLITTERING TAPS

The theoretical analysis relays upon synthesis of unidirectional transformers examined in [5], [9] and splitters described in [12]. The analysis is analogical to the analysis of mentioned elements and that is way it is omitted in this paper. Only the table with two-way splitterring tap admittance matrixes is given for illustrative purposes (see also Fig. 1).

Table 1

Admittance mattrix of the two-way splittering tap

node	IN	OUT	–	TAP A	TAP B
IN	y_{11T}	y_{13T}	y_{12T}	0	0
OUT	y_{31T}	y_{33T}	y_{23T}	0	0
-	y_{21T}	y_{32T}	$y_{22T} + y_{11S}$	y_{12S}	y_{13S}
TAP A	0	0	y_{21S}	y_{22S}	y_{23S}
TAP B	0	0	y_{31S}	y_{32S}	y_{33S}

The theoretical analysis were carried out for the two-way, three-way and four-way splitterring taps.

Assuming that theoretical analysis of splitterring taps can be treated as the synthesis of transformer analysis and splitter analysis, conclusions of transformers and splitters analysis results can be treated as conclusions of splitterring taps analysis as well. At the input and at the output splitterring taps has similar features to features of unidirectional transformer.

Depending on the chosen configuration of the transformer [5], [9] the reflectance at the input IN is better for basic configuration, and its worse for the inversed configuration. When evaluating the OUT-TAP parameter some similarity to the unidirectional transformer can be noticed: the isolation between output and branches is better for the inversed configuration of unidirectional transformer.

At the TAP output, splitterring taps are similar to splitters: worse TAP reflectance and isolation between TAP A-B branches.

It is necessary to be aware of the fact that for TAP A-B parameter, reflectance at the input of splitter is of high importance [12]. Transformer parameters should have influence on the isolation between TAP A-B branches.

Fig. 6, 7 present the influence of transformer permeability and ballast resistor on TAP A-B parameter, for basic configuration of transformer at 5 MHz for two-way and three-way splitterring tap respectively (the four-way splitterring tap has analogical to three-way splitterring tap characteristics). It is noticeable that elements of the transformer have influence on TAP A-B isolation.

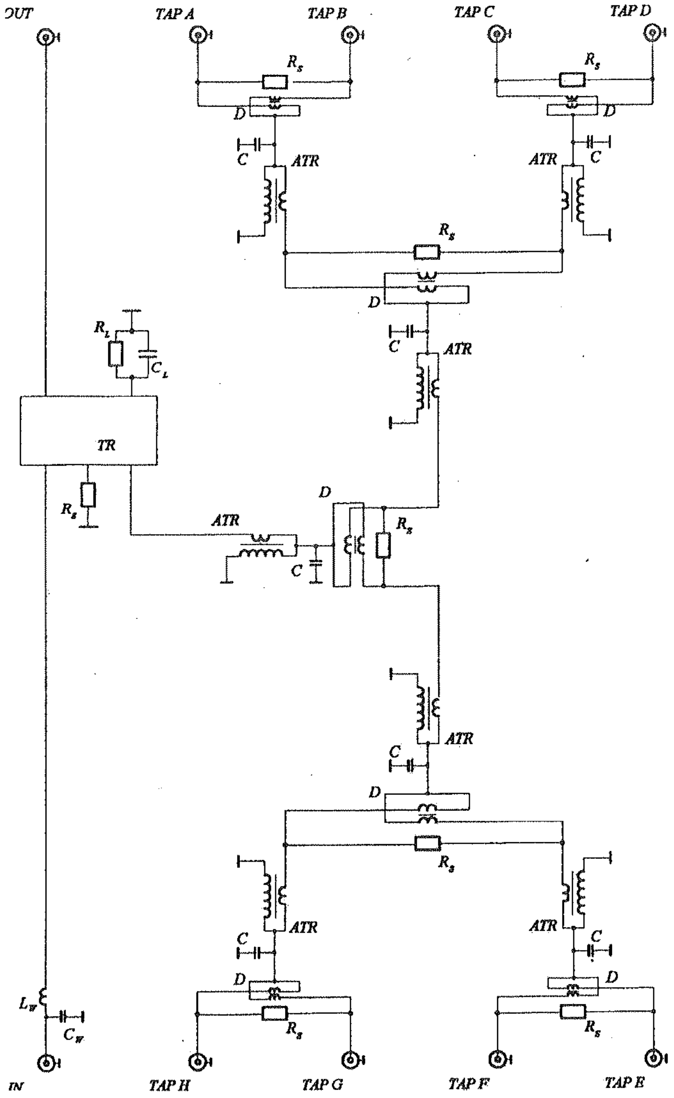


Fig. 3. The eight-way splittering tap

Moreover, splitterring taps have typical characteristics for every integral element; i.e. when taking into account the initial permeability of the transformer, the best performance can be met for permeability values in the range of 3000-6000 (for every unidirectional transformer), when considering permeability of inductive elements of the splitter – generally, lower permeability is better, similarly to splitter separately. Fig.

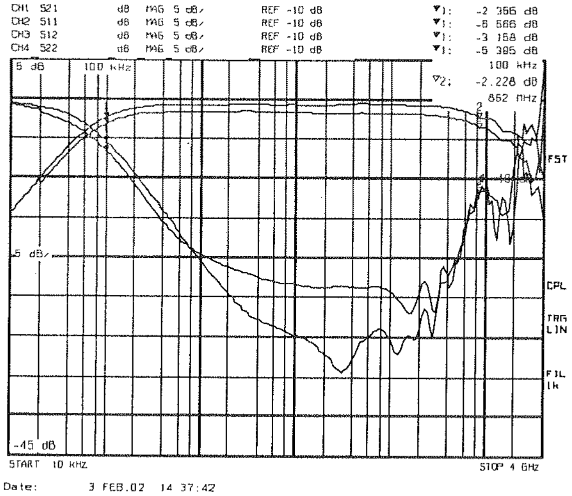


Fig. 4. The IN-OUT characteristic of the 15-dB dual splittering tap in wideband frequencies

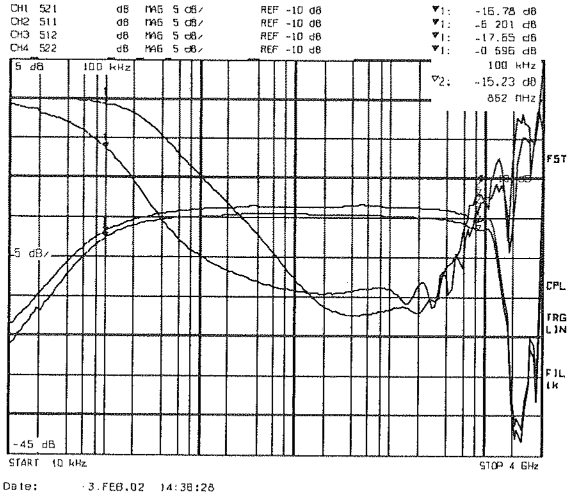


Fig. 5. The IN-TAP characteristic of the 15-dB dual splittering tap in wideband frequencies

8, 9 illustrate the initial permeability influence of unidirectional transformer and splitter respectively, for two-way splitterring tap at 5 MHz. Generally, permeability influence on device's parameters is noticeable at low frequencies, dependence of permeability at high frequencies is practically negligible [5], [9], [12].

There is no interaction between transformer and splitter (besides reflectance at the input influence on TAP A-B isolation, what was mentioned above). The example of no correlation are IN reflectance, OUT reflectance and OUT-TAP isolation, simulated characteristics versus the transformer permeability and separatory resistance of the divider, for two-way splitterring tap with transformer in basic configuration, is illustrated

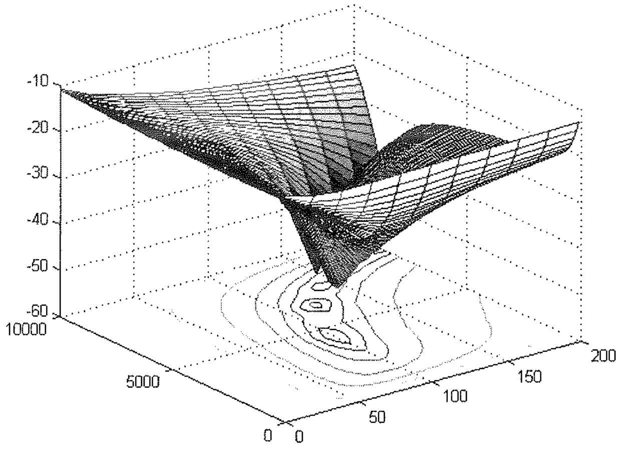


Fig. 6. The influrence of the initial permeabilitry and ballast resistance R_L of the transformer, on TAP A-B parameter of the two-wag splittering tap

by Fig. 10. There is no influence of separatory resistor R on parameters, which are determined by the unidirectional transformer.

The influence of capacitances and dissipation of the splittering taps, what is the result of placing the unidirectional transformer and autotransformers, is firmly negative [5], [9]. These factors limits working frequency band at high frequencies and have destructive influence on reflectances IN, OUT, TAP and isolations OUT-TAP and TAP A-B of the splittering taps. Special transformer winding techniques and core miniaturization can be done for preventive purposes.

For splitterring taps typical for unidirectional transformers correctors (low and high frequency) can be used, and for splitters - high frequency correction of attenuation characteristic can be done (C capacitance at Fig. 1, 2). For splitters, low frequency correction is ineffective [12].

Fig. 11 presents the two-way, three-way and eight-way (see also Fig. 3) basic characteristics of the splitterring taps. Fig. 12 illustrates circular charts of IN, OUT, TAP reflectances for two-way splitterring tap. The characteristic feature of the splitterring taps is poor, typical for splitters, isolation between TAP A-B branches. It is not getting better with higher number of section, neither with coupled lines attenuation. It is noticeable that for the eight-way splitterring tap inversed configuration with isolating resistor was utilized in order to improve OUT-TAP isolation. The inversed configuration with isolating resistor was introduced in papers [5], [9].

It is worth to compare splitterring tap with multidirectional transformers [6], [7], [9]. Similarly to splitterring taps, multidirectional transformers are characterized by equal level of coupled lines attenuation, but isolation between branches increases with higher number of section and with higher coupled lines attenuation level, keeping certain directivity. That differences are better for multidirectional transformers and stronger for higher number of sections, beside the fact that TAP A-B parameter is not the most advantageous. Another advantage is that multidirectional transformers are based only on one inductive element, and that is why it is easier to obtain steady IN-TAP characteristics. The main disadvantage of multidirectional transformers, is not the best IN and OUT reflectances at low frequencies. Considering that parameters, splitterring taps are better, which is the result of unidirectional transformer at the input.

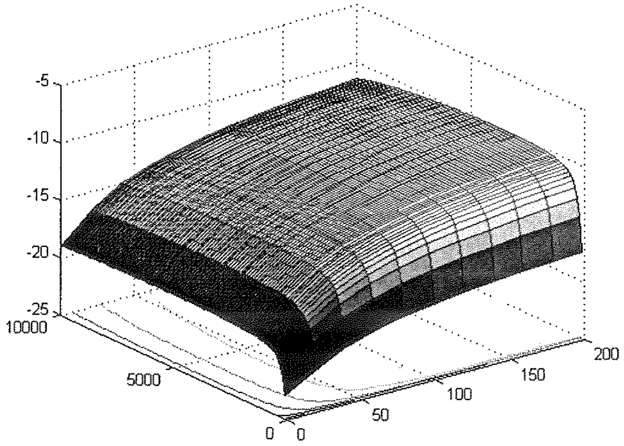


Fig. 7. The influrence of the initial perneability and ballast resistance R_L of the transformer, on TAP A-B parameter of the three-way splittering tap

3. OTHER APPLICATIONS

Splitterring taps can be used as autonomous devices and as parts of other devices as well. The example of utilization by different device is a branching-splitting advanced multimedia separator (its diagram is presented by Fig. 13, its photo is presented by Fig. 14). The advanced multimedia separator is not precisely a device of splitterring tap type, in fact, but it is a combination of transformer and splitter. Features of the splitters are noticeable at RTV outputs, whilst at the IN input and at the DATA output device is similar to the unidirectional transformer. It is necessary to remember that the input of the device, has significant influence on isolation between RTV outputs. The advanced multimedia separator is a synthesis of high pass filters HPF, low pass filter LPF and branching-splitting element.

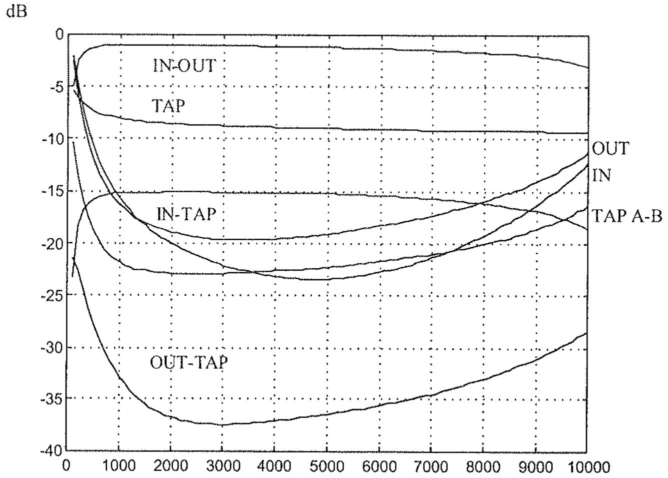


Fig. 8. The influence of the initial permeability of the transformer on the parameters of the two-way splitterring tap, at frequency 5 MHz

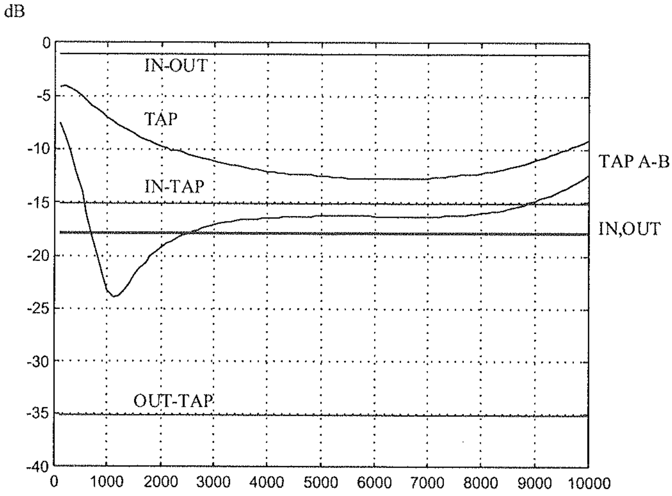


Fig. 9. The influence of the initial permeability of the divider on the parameters of the two-way splitterring tap, at frequency 5 MHz

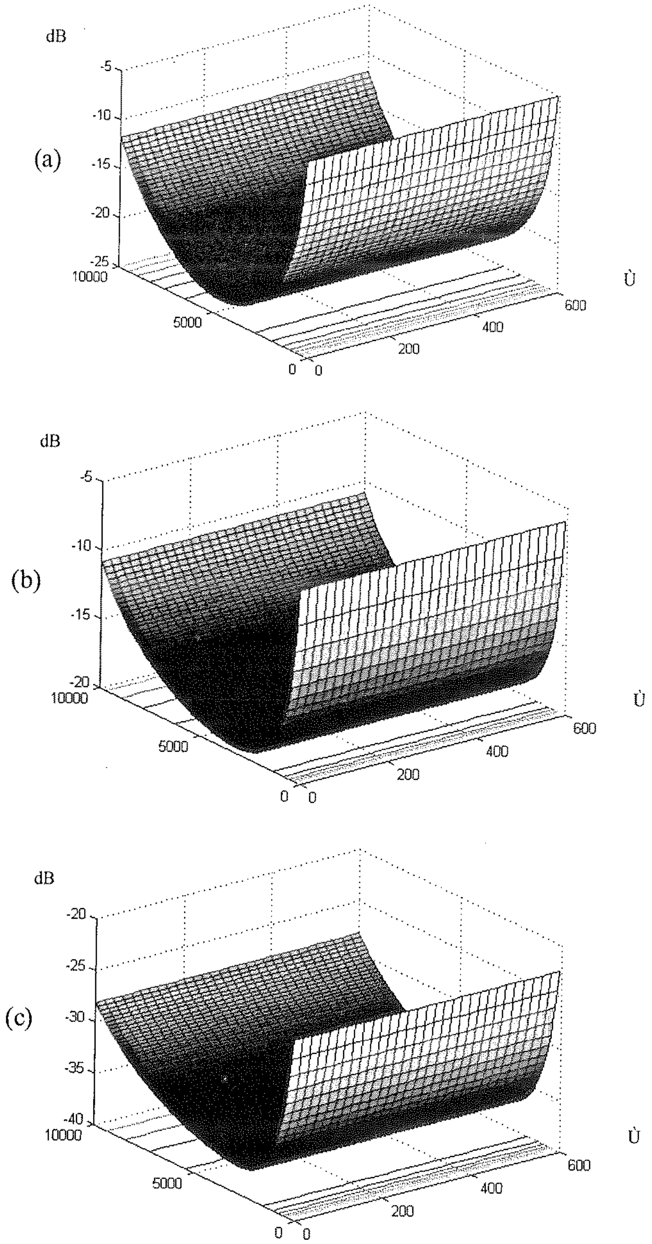


Fig. 10. Relations between the initial permeability of the unidirectional transformer and separatory resistor R of the splitter for parameters IN (a), OUT (b) and OUT-TAP (c)

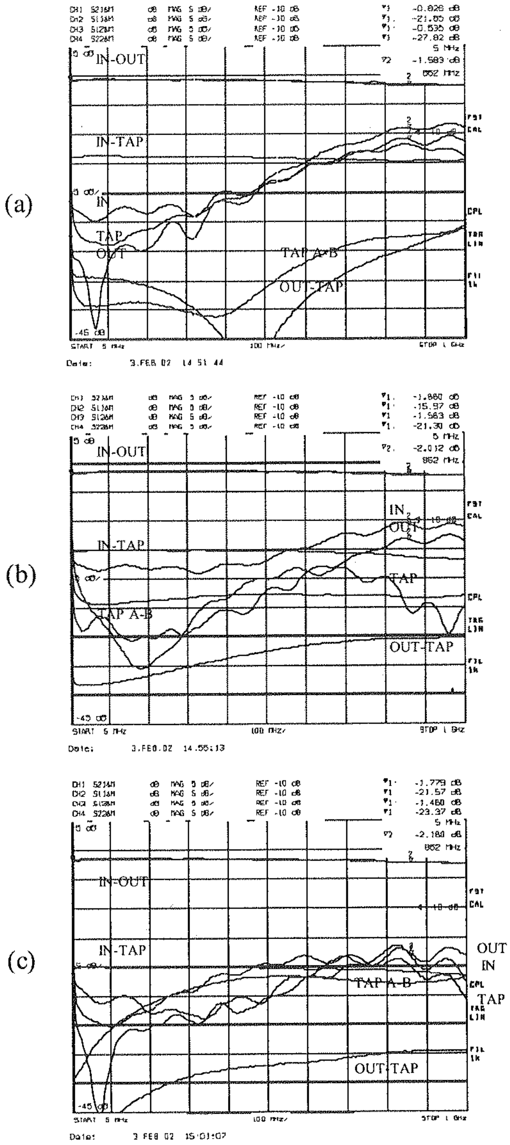


Fig. 11. Frequency characteristics of the splitterring taps: two-way (a), three-way (b) and eight-way (c)

The advanced multimedia separator uses high frequency correction at the "IN" input, at the "DATA" output and at splitter circuit. A "T" type circuit is a low frequency correction circuit, in this case it corrects low pass filter characteristic.

The advanced multimedia separator has isolation between input ground and device ground and suppresing line, which provides electromagnetic compatibility EMC [10]. Fig. 15 presents basic characteristics of the advanced separator.

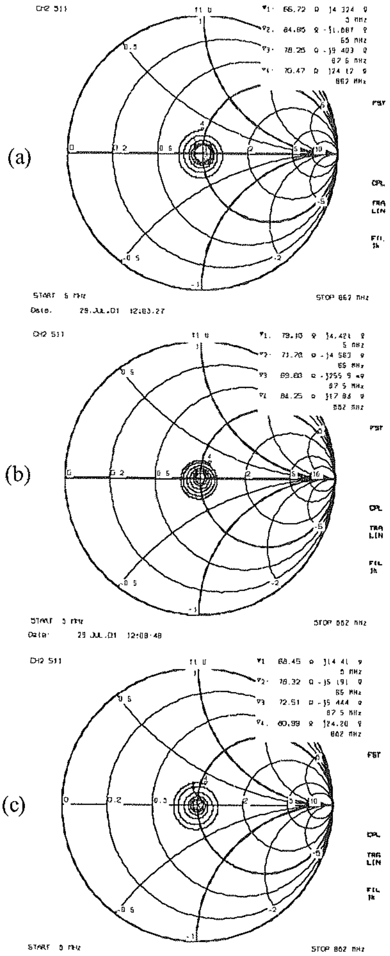


Fig. 12. Smith charts of the reflectances IN (a), OUT (b) and TAP (c) of the 14-dB two-way splitterring tap

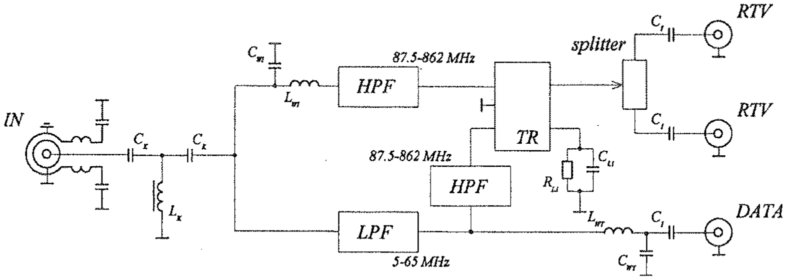


Fig. 13. The schematic diagram of the multimedia separator

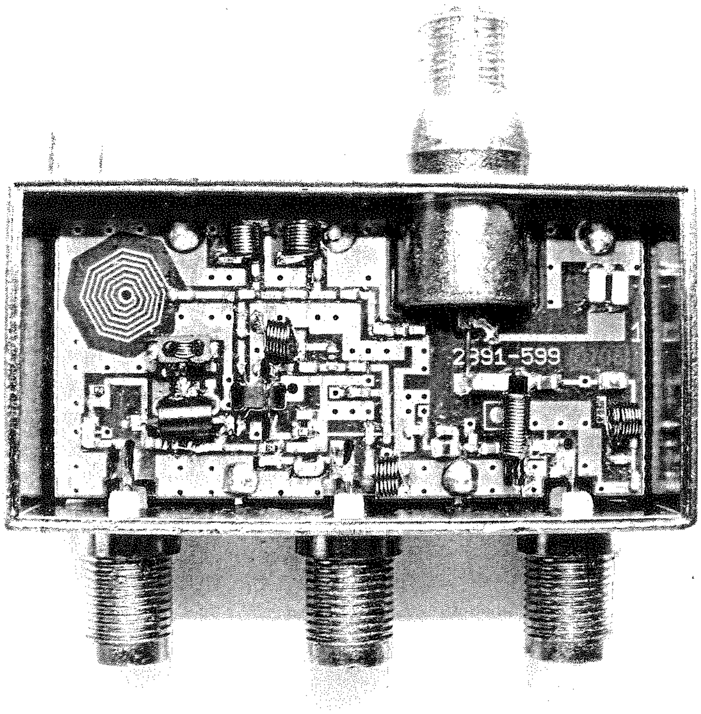


Fig. 14. The photo of the multimedia separator

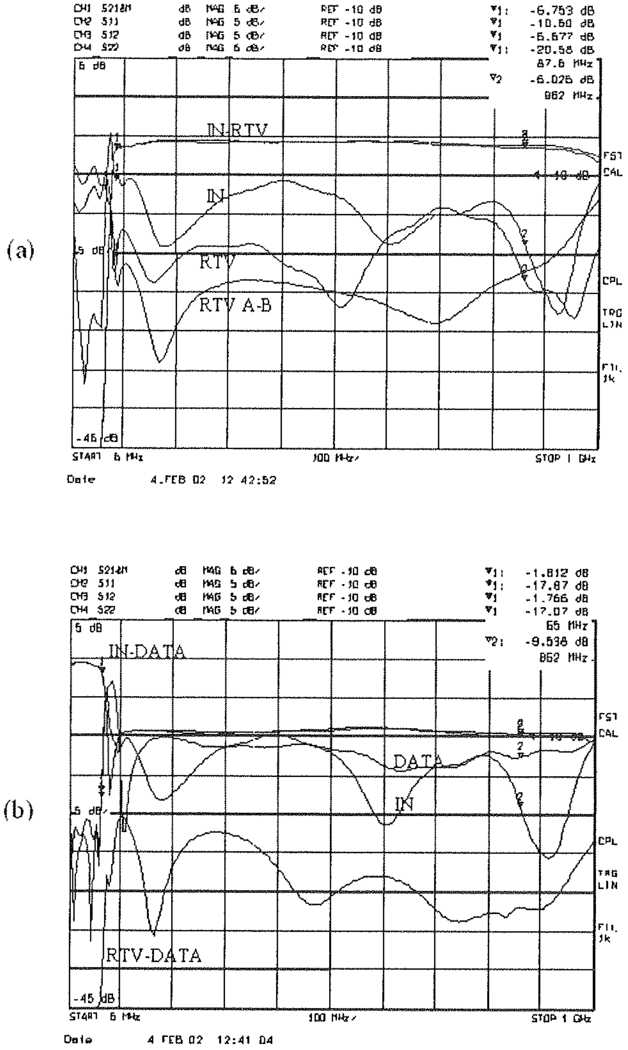


Fig. 15. Main characteristics of the advanced multimedia separator

4. CONCLUSIONS

Splitterring taps as the synthesis of the unidirectional transformers and splitters, have features of unidirectional transformers at the input and the output, and features of splitters at branched outputs. It is necessary to be aware of the fact that transformer features are significant, because of reflectance at the input influences on TAP A-B parameter. Due to majority of splitterring tap parameters it would be advisably using for the transformer ferrite cores of high permeability value. However, due to intermodulation distortions, cores of lower permeability values are utilized at the expense of worse reflectance at lower frequencies. There low frequency correctors typical for unidirectional transformers [5], [9] are helpful.

Moreover, it is necessary mention that developed mathematical model, because of causes given in [5], [9], is characterized by about 10 % precision.

Splitterring taps, which were described in this article, have CaTV working frequency band (5-862 MHz). It is worth to mention that some appropriate techniques, can widen the working frequency band of the splitterring taps. The appropriate constructions of transformers [5], [9] and splitters [12] and microstrip technique utilization, can widen the working frequency band of the splitterring taps even to 3 GHz.

The characteristic impedance of splitterring taps were assumed as 75 Ω (standard impedance for transmission lines in CaTV system). It is necessary to remember, that change in characteristic impedance value, comes with changes in parameters of the splittering taps. The increase in the characteristic impedance results in higher IN-OUT and IN-TAP attenuation. The reflectances IN, OUT, TAP also decreases its performance with characteristic impedance increase. Moreover, the fact that changes in characteristic impedance come with changes in value of optimal ballast resistor of the unidirectional transformer for TAP and OUT-TAP parameters, is of high importance. According to prior considerations, these changes have indirect influence on isolation between TAP A-B.

5. REFERENCES

1. D. I. Kim, M. Takahashi, K. Araki, Y. Naito: *Optimum design of the power dividers with ferrite toroids for CATV and/or MATV systems.* IEEE Trans. On Consumer Electronics, vol. CE-29, no 1, Feb. 1983, pp. 27-38.
2. Y. Naito: *Formulation of frequency dispersion of permeability.* Trans. IECE, vol. 59-c May 1976, pp. 297-304.
3. D. I. Kim, S. W. Jung, Y. Yun: *A high performance transformer-type Wilkinson power splitter with compensating circuit for CATV transmission system and its optimal design method.* IEEE Trans. On Consumer Electronics, vol. 50, no 3, Aug. 2004, pp. 934-940.
4. D. Krzemieniecki: *Advanced splittering multimedia separators.* National Telecommunication Symposium, Bydgoszcz 2001, vol. A, pp. 411-420 (in polish).
5. D. Krzemieniecki: *Investigation and optimization of the wideband directional transformer.* Electronics and Telecommunication Quarterly, 2005, 51, vol. 1, pp. 105-135 (in polish).
6. D. Krzemieniecki: *Investigation and optimization of the wideband two-directional transformer.* Electronics and Telecommunication Quarterly, 2005, 51, vol. 3, pp. 387-419 (in polish).
7. D. Krzemieniecki: *Investigation and optimization of the wideband four-directional transformer.* Electronics and Telecommunication Quarterly, 2005, 51, vol. 4, pp. 603-637 (in polish).
8. D. *Krzemieniecki*: *Current posibilities of projecting of the taps arrangements.* National Conference of Radiocommunication, Radio and Television, Poznań 2006, pp. 457-460 (in polish).
9. D. Krzemieniecki: *Modelling and designing of directional transformers*, PhD Thesis, Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, Poland (in Polish) 2006.
10. D. Krzemieniecki: *Advanced multimedia outlets and separators with full isolation..* National Conference of Radiocommunication, Radio and Television, Gdańsk 2007, pp. 415-418 (in polish).

11. D. Krzemieniecki: *Analysis of the CaTV cascade taps*. Electronics and Telecommunication Quarterly (in print).
12. D. Krzemieniecki, J. Juchniewicz: *Analysis of the CaTV splitters*. Electronics and Telecommunication Quarterly (in print).

A
used
deriv
by a
Fig.
by th

T
line i

T
line a
OUT,
The w
freque
Trans
ristics

ELECTRONICS AND TELECOMMUNICATIONS QUARTERLY, 2008, 54, no 3, pp. 445–459

Analysis of the CaTV Radio Coupler

DARIUSZ KRZEMIENIECKI, JAKUB JUCHNIEWICZ

Gdańskie Zakłady Teleelektroniczne TELKOM-TELMOR
80-425 Gdańsk ul. Mickiewicza 5/7
e-mail: dariusz.krzemieniecki@telmor.pl
jakub.juchniewicz@gmail.com

Received 2008.03.12
Authorized 2008.05.30

This article presents a radio coupler which is used as a part of some CaTV passive devices. Basic characteristics were investigated and a few examples of application were given in this content. Additionally correction circuits for improvement of important parameters were described.

Keywords: radio coupler

1. INTRODUCTION

A so-called radio coupler is the simplest coupling device of its kind and is widely used as a important part of devices designed for CaTV network. “Radio coupler” derives its name from its intended application – a select coupled line is mostly used by a radio band signal. Utilization as a part of wall outlets is the most common usage. Fig. 1 presents the outlook of the radio coupler, its usage in wall outlet is illustrated by the photo at Fig. 2.

The main line is determined by 0.5 coils upon a ferrite core while the coupled line is determined by 1.5 coils.

The radio coupler is described by following frequency dependent parameters: main line attenuation IN-OUT, coupled line attenuation IN-TAP, reflection coefficients IN, OUT, TAP and isolation between output and coupled line OUT-TAP.
The working frequency band of the radio coupler is 5 – 862 MHz, what is the working frequency band of CaTV system.
Transmission lines have characteristic impedance of 75 Ω. Fig. 4, 5 present characteristics of the radio coupler versus wide band of frequencies.

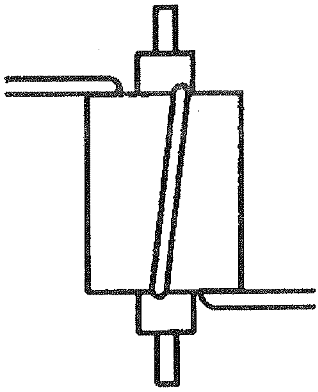


Fig. 1. The view of the radio coupler

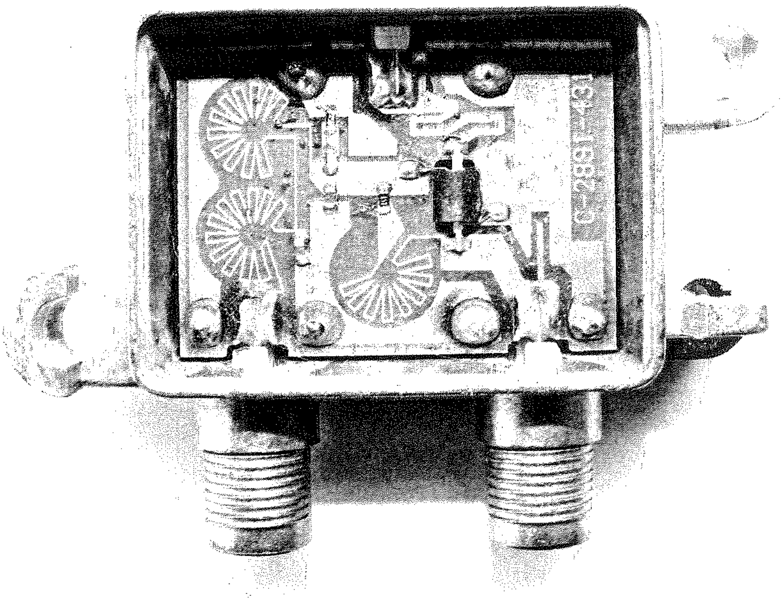


Fig. 2. The radio coupler in the wall outlet

2. THEORETICAL ANALYSIS

The analyzed radio coupler is a single three-way device. The node voltage method will be used in order to derive frequency dependent parameters of the radio coupler.

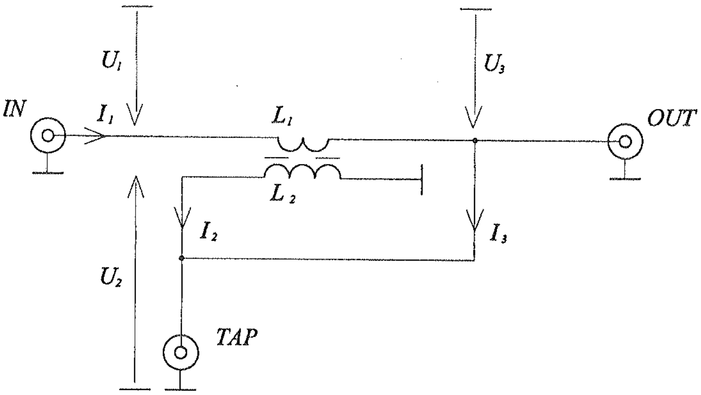


Fig. 3. Nodal model of the radio coupler

According to Fig. 3 we can write voltage equations, then current equations can be written.

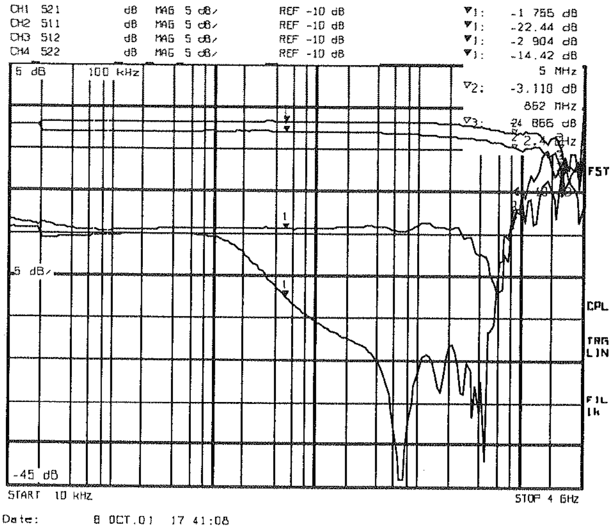


Fig. 4. The IN-OUT characteristic of the radio coupler in the wideband frequencies

$$U_1 - U_3 = j\omega L_1 I_1 - j\omega M_{12} I_2 \tag{1}$$

$$U_2 = -j\omega M_{21} I_1 + j\omega L_2 I_2 \tag{2}$$

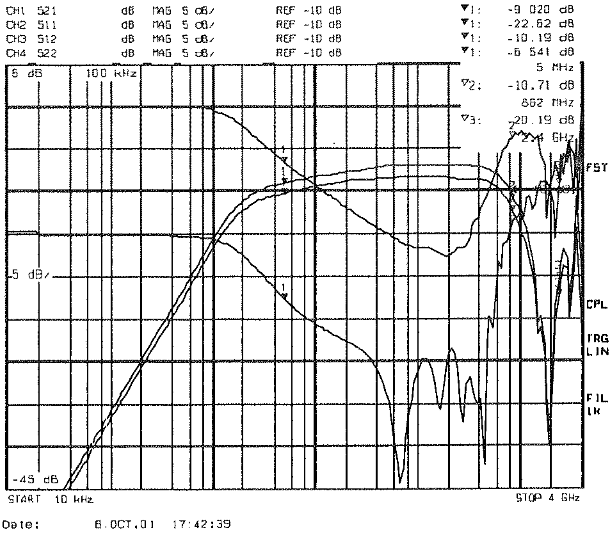


Fig. 5. The IN-TAP characteristic of the radio coupler in the wideband frequencies

$$I_1 = (U_1 - U_3)\zeta_{11} + U_2\zeta_{12} \tag{3}$$

$$I_2 = (U_1 - U_3)\zeta_{21} + U_2\zeta_{22} \tag{4}$$

A transconductance of the circuit is marked as a ζ_{ij} symbol.
According to equations above, node equations can be derived.
For node "IN":

$$U_1 y_0 = U_1\zeta_{11} - U_3\zeta_{11} + U_2\zeta_{12} \tag{5}$$

For node "OUT":

$$U_3 y_0 = -U_1\zeta_{11} + U_3\zeta_{11} - U_2\zeta_{12} \tag{6}$$

For node "TAP":

$$U_2 y_0 = U_1\zeta_{21} - U_2\zeta_{21} - U_3\zeta_{22} \tag{7}$$

The scattering matrix and inductances are derived analogically to [5].

Fig. 7, 8 present theoretical and measured characteristics respectively. The coupled line is used in limited frequency band because of poorer characteristics at high frequencies. The main line saves wide frequency band.

In order to improve parameters of the radio coupler, correction circuits are commonly utilized. Correction circuits are illustrated by Fig. 6.

It is necessary to emphasize that because of causes given in [5] the mathematical model has a about 10 % precision.

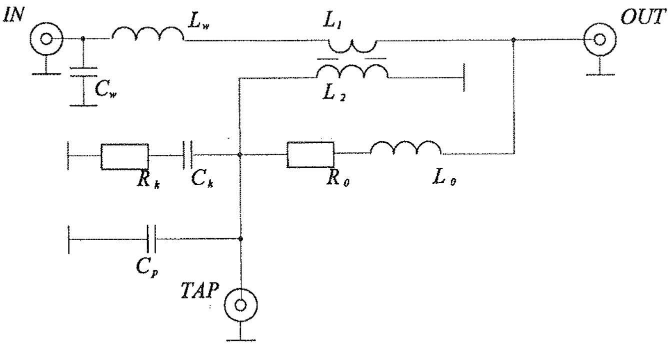


Fig. 6. The radio coupler with correcting circuits

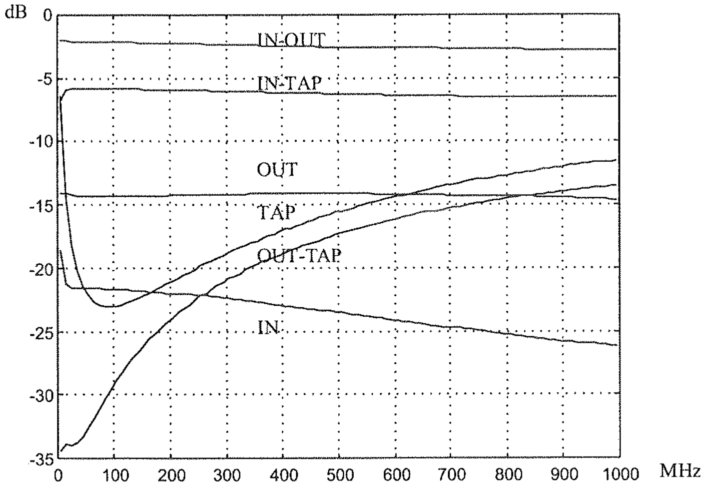


Fig. 7. Calculated frequency characteristics of the radio coupler

3. THE CHARACTERIZATION OF THE RADIO COUPLER

Fig. 10 illustrates how an initial permeability of a ferrite core influences on frequency dependent parameters of the radio coupler (at frequency of 5 MHz). Because of the IN reflectance, it is better to use ferrite cores of high permeability. The best performance for TAP parameter can be met for permeability values in the range of 3000-6000. Regarding the OUT-TAP isolation, permeability has no influence on this parameter. The influence of permeability is advantageous, because of the fact that attenuations of the IN-OUT and IN-TAP are slightly affected by the permeability value. The initial permeability has obviously influence on parameters of the coupler but only at low frequencies. With the increase in the frequency this influence decreases and no influence can be assumed at the frequency of 862 MHz.

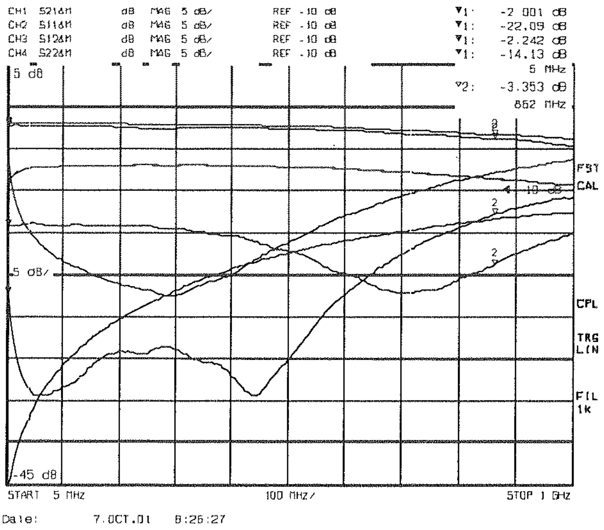


Fig. 8. Measured frequency characteristics of the radio coupler

An isolating resistor R_0 has mainly influence on the OUT-TAP isolation. This parameter has the optimum value of about 140 Ohm (Fig. 11).

The main purpose of R_kC_k circuit is correction of the IN-OUT characteristic at high frequencies. An increase in capacitance causes greater IN-TAP attenuation at high frequencies. It doesn't really matter if the limited band is utilized for the coupled line. Because of that fact, lower IN-TAP attenuation can be obtained. Good choice of R_k resistor can result in IN, TAP reflectances and OUT-TAP isolation improvement.

Fig. 12, 13 illustrate influence of capacitor and correction circuits on parameters of the radio coupler at 862 MHz.

Another improvement of the IN-OUT attenuation at high frequencies can be obtained by adding parallel capacitor in coupled line (Cp capacitor).

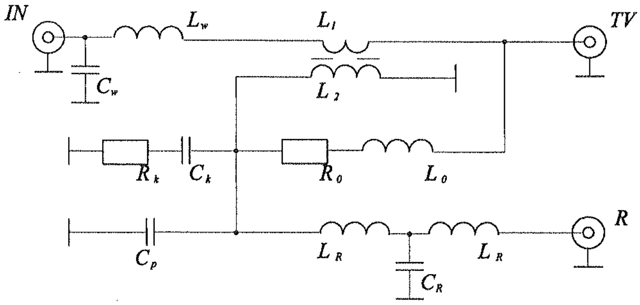


Fig. 9. The schematic diagram of the simple wall outlet

For low frequency correction of the radio coupler is not used because it is ineffective.

Moreover, at the input of the device, L_wC_w correction circuit can be used in order to improve IN, OUT reflectances at high frequencies (Fig. 14, 15).

In addition to our consideration, we can say that the coil configuration (0.5/1.5), which was mentioned in the introduction, is the optimum configuration. Every kind of change in the number of coils decreases the OUT-TAP isolation performance, that is way the research in this area wasn't carried out.

Correction circuits were used in the wall outlet presented in the introduction. This device uses low pass filter in coupled line for radio band separation (Fig. 9).

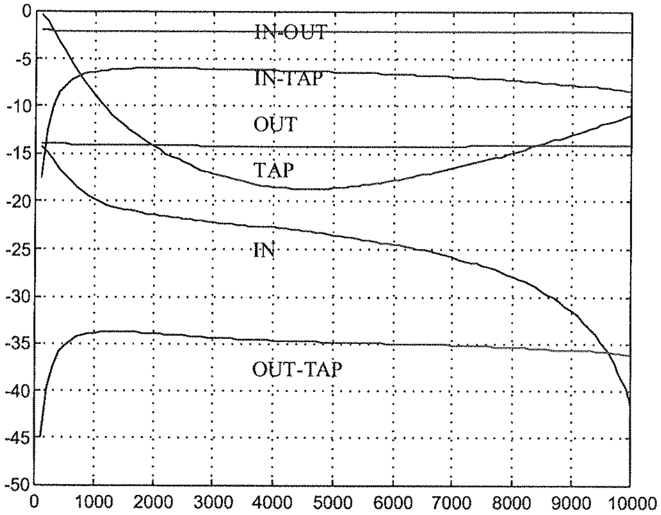


Fig. 10. The influence of the initial permeability on parameters of the radio coupler, at frequency 5 MHz

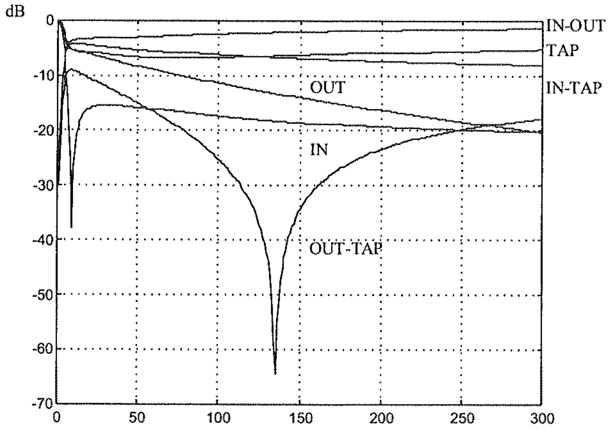


Fig. 11. The influence of the separatory resistor R_0 on parameters of the radio coupler, at frequency 5 MHz

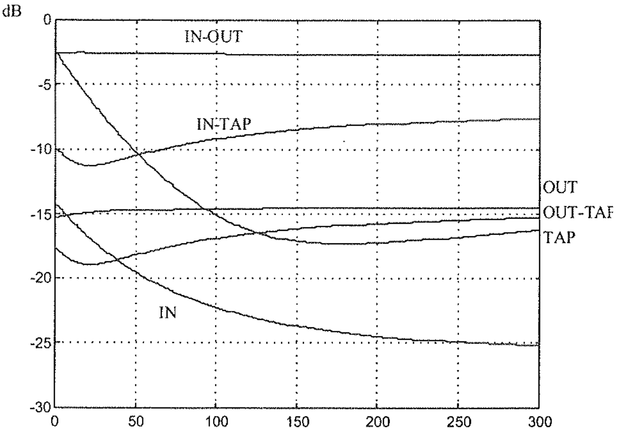


Fig. 12. The influence of the corrective resistance on parameters of the radio coupler, at frequency 862 MHz

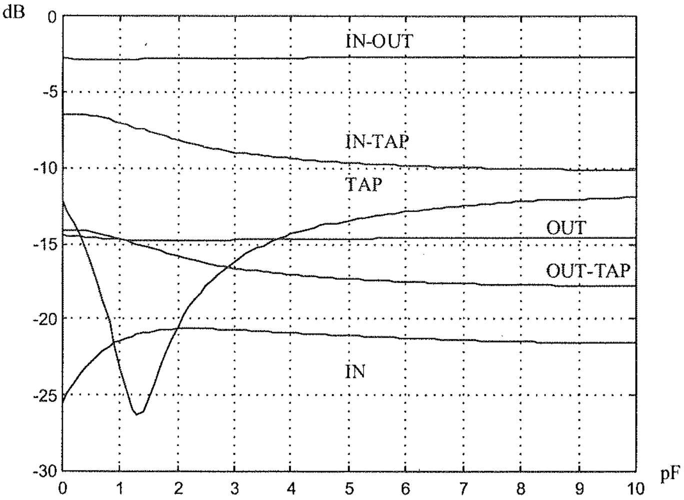


Fig. 13. The influence of the corrective capacitance on parameters of the radio coupler, at frequency 862 MHz

Basic characteristics of the wall outlet versus frequency are illustrated at included charts (Fig. 16). Moreover, circular charts of reflectances were presented (Fig. 17) in order to complete our considerations.

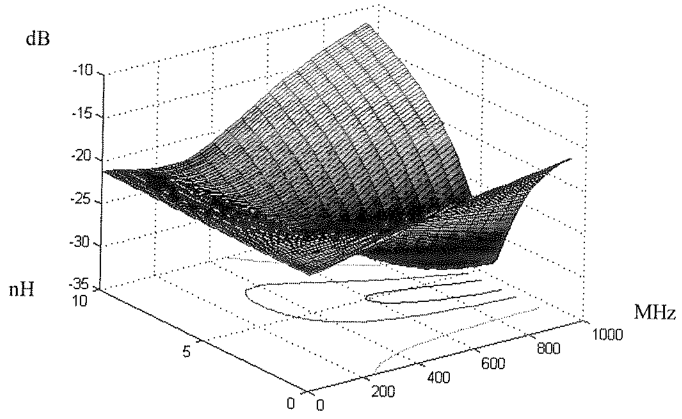


Fig. 14. The influence of the corrective inductance on the IN parameter

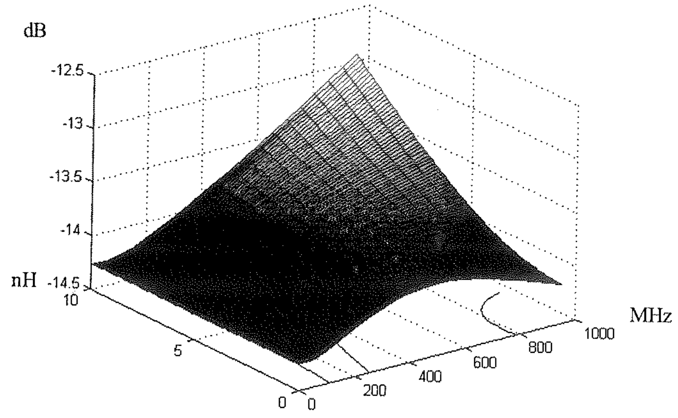


Fig. 15. The influence of the corective inductance on the OUT parameter

Other examples of the radio coupler are as following: an advanced multimedia wall outlet [4] and a Wi-Fi wall outlet [6], where the radio coupler is used for a radio signal separation. Fig. 18, 19, 20 and 21 present diagrams and photos of mentioned devices. Fig. 22 illustrate characteristics of the advanced multimedia wall outlet. The wall outlet is a synthesis of a directional transformer, coupler, high pass HPF and low pass LPF filters. This wall outlet uses typical correction circuits. The low frequency correction is provided by T-type filter, the high pass correction is provided by $L_w C_w$ circuit at the input and by correction circuits of the radio coupler.

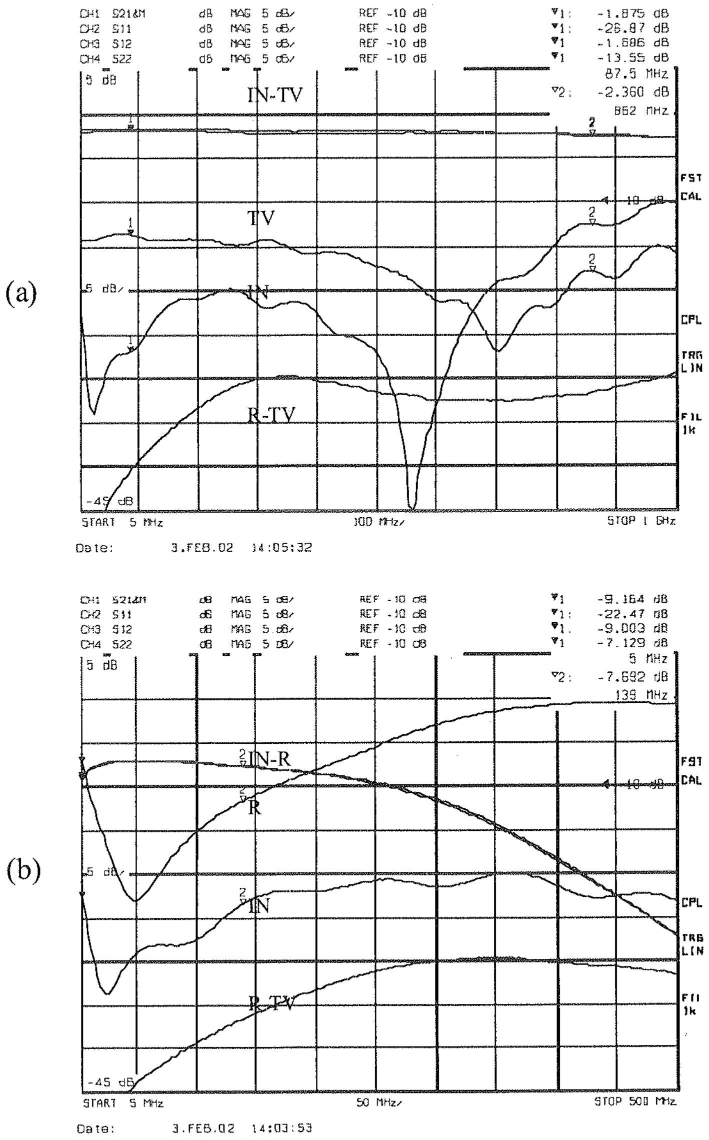


Fig. 16. Frequency characteristics of the wall outlet

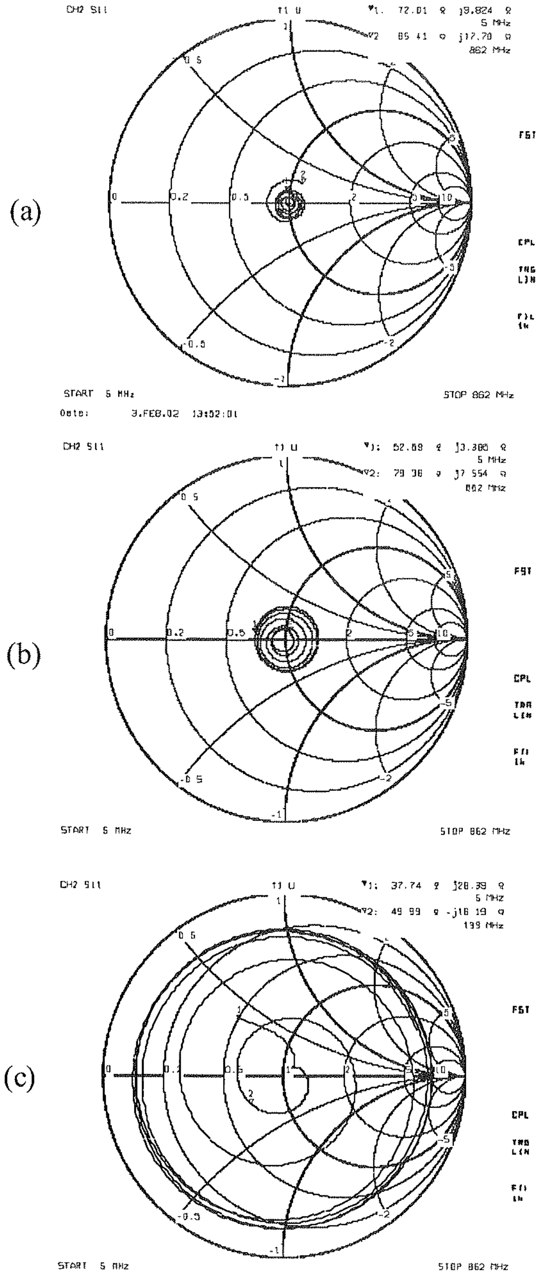


Fig. 17. Smith charts of the reflectances IN (a), TV (b), and R (c) of the wall outlet

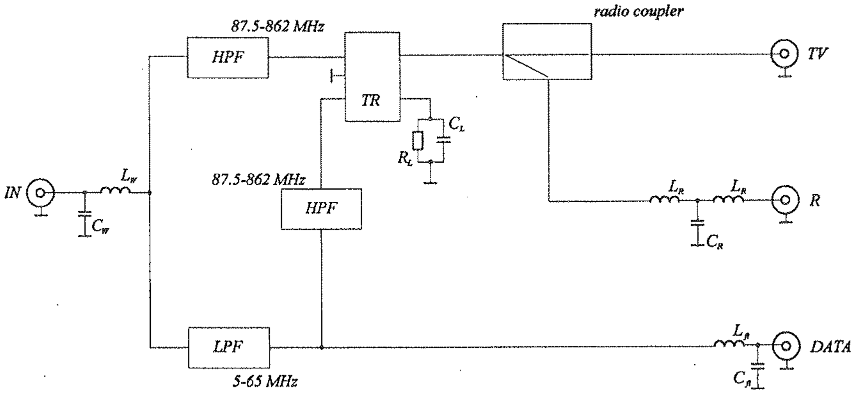


Fig. 18. The schematic diagram of the multimedia wall outlet

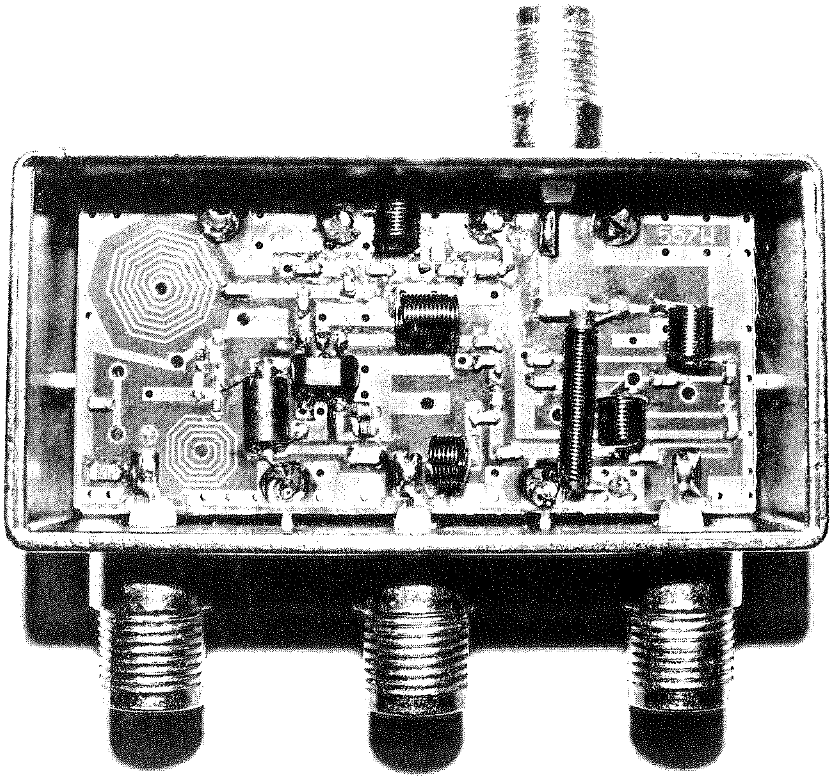


Fig. 19. The photo of the multimedia wall outlet

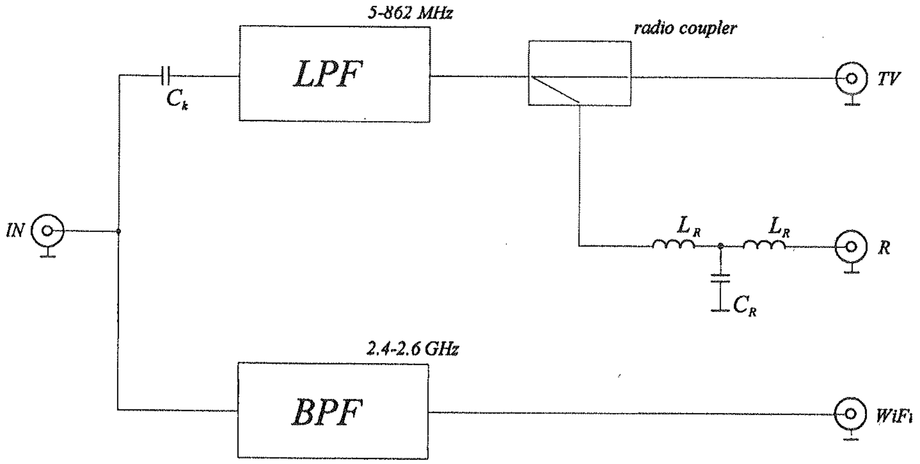


Fig. 20. The schematic diagram of the WiFi wall outlet

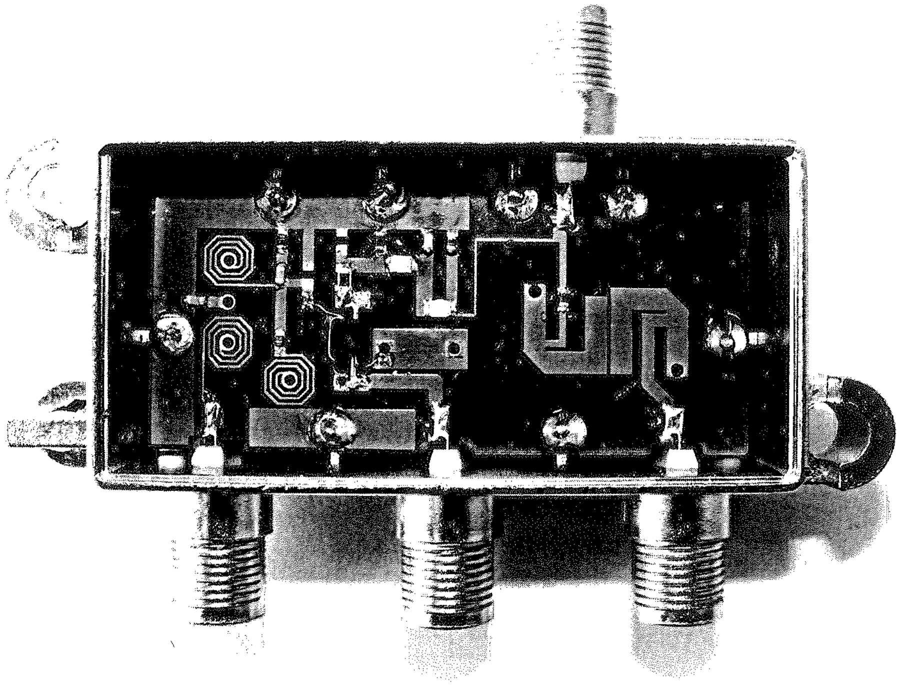


Fig. 21. The photo of the WiFi wall outlet

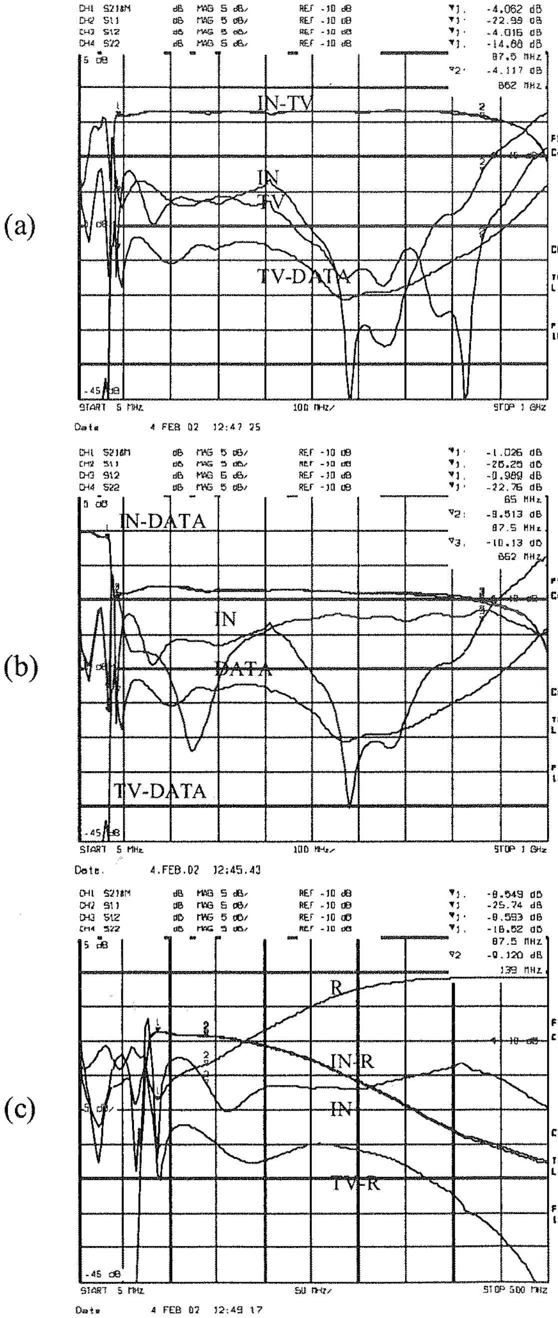


Fig. 22. Main characteristics of the advanced multimedia wall outlet

4. CONCLUSIONS

The main advantage of the radio coupler is its exceptional simplicity, which results in its low production cost. It is necessary to mention that a directional transformer is much more complicated solution in wall outlets. It is true that the second solution has better performance, but if the goal is the separation of radio signal band, the radio coupler is ideal solution then. Correction circuits, which where presented in this article, allow to obtain steady IN-OUT characteristics and can be used in various tapped type wall outlets design.

The coupler was investigated for the characteristic impedance of 75 Ohm. Some applications need different characteristic impedance and it is necessary to be aware of the fact, that with changes in its value, the changes in parameters of the radio coupler follow. With the rise in characteristic impedance values attenuations rise and reflectance go worse. The significant fact is that for assumed different characteristic impedance value, it is necessary to choose different resistance value R_0 in order to obtain optimal OUT-TAP isolation.

5. REFERENCES

1. D. I. Kim, M. Takahashi, K. Araki, Y. Naito: *Optimum design of the power dividers with ferrite toroids for CATV and/or MATV systems.* IEEE Trans. On Consumer Electronics, vol. CE-29, no 1, Feb. 1983, pp. 27-38.
2. Y. Naito: *Formulation of frequency dispersion of permeability.* Trans. IECE, vol. 59-c May 1976, pp. 297-304.
3. D. I. Kim, S. W. Jung, Y. Yun: *A high performance transformer-type Wilkinson power splitter with compensating circuit for CATV transmission system and its optimal design method.* IEEE Trans. On Consumer Electronics, vol. 50, no 3, Aug. 2004, pp. 934-940.
4. D. Krzemieniecki: *Advanced tapping multimedia outlets and separators.* National Telecommunication Symposium, Bydgoszcz 2001, vol. A, pp. 401-410 (in polish).
5. D. Krzemieniecki: *Investigation and optimization of the wideband directional transformer.* Electronics and Telecommunication Quarterly, 2005, 51, vol. 1, pp. 105-135 (in polish).
6. J. Juchniewicz, D. Krzemieniecki: *Distribution signals of the Wi-Fi system over CaTV network.* National Conference of Radiocommunication, Radio and Television, Wrocław 2008 (in polish, in print).

An ar
racter

Basic
The a
Pared
Layou
- Titl
- Aut
- Wo
- Con
- Mai

Pages

Main
Main
leters

Table
Table
th do
on th
cited

Math
Chara
the ba
mulas
shoul
sation

Refer
Refer
placed
1.
2.
3.

Figur
Figur
Corel
the ba
shoul

INFORMATION FOR AUTHORS

An article published in other magazines can not be submitted for publishing in E.T.Q. The size of an article 1800 characters each, including figures and tables.

Basic requirements
The article should be submitted to the editorial staff as a one side, clear, black and white computer print should be prePared in English. Floppy disc or a Cd with an electronic version of the article should be enclosed.
Layout of the article:

- Title
- Author (first name and surname of author/authors)
- Workplace (institution, address and e-mail)
- Concise summary
- Main text with following layout:
    - Introduction
    - Theory (if applicable)
    - Numerical results (if applicable)
    - Paragraph 1
    - Paragraph 2
    - ....
    - Conclusions
    - ....
    - Acknowledgements (if applicable)
    - References

Pages should have continuous numbering.

Main text
Main text cannot contain formatting such as spacing, underlining, words written in capital letters (except words in capital leters). Author can mark suggested formatting with pencil on the margin of the article using com monly marks.

Tables
Tables with their titles should be placed on a separate page at the end of the article. Titles of rows and columns letter with double line spacing. Annotations concerning tables should be placed directly below the table. Table Arabic numbers on the top each table. Table can contain algorithm and program listings. In such cases original preserved. Table should be cited in the text.

Mathematical formulas
Characters, numbers, letters and spacing of the formula should be adequate to layout of the article. Indexes reised above the basic line and clearly written. Special characters such as lines, arrows, dots should be place they are attributed to. Formulas should be numbered with Arabic numbers placed in brackets on the right side measure, letters and graphic symbols should be printed according to requirements of IEC (International Electronical) (International Organisation of Standardisation).

References
References should be placed at themain text with the subtitle "References", References should be numbered to references placed in the text. Examples of periodical [1], non-periodical [2] and book [3] references:

1. F. Valdoni: A new milimetre wave satellite. E.T.T. 1990, vol. 2, no 5, pp. 141– 148
2. K. Andersen: A resource allocation framework. XVI International Symposium, Stockholm (Sweden),
3. Y. P. Tvidis: Operation and modeling of the MOS transistors. New York, McGraw-Hill, 1987, p. 553

Figures
Figures should be clearly drawn on plain or milimetre paper in the format not smaller than 9×12 cm. Figures can (editor – CorelDRAW). Photos or diapositives will be accepted in black and white format not greater than 10×15 cm drawing and on the back of each photo author name and abbreviation of the article title should be placed. Figure's on a speared page. Figure should be cited in the text.

Additional information

On a separate page following information should be placed:

- mailing address (home or office)
- phone (home and/or office)
- e-mail

Authors in entitled to free of charge 20 copies of article. Additional copies or the whole magazine can be order expense. Author is obliged to perform the authors correction, which should be accomplished within 3 days starting from from the editorial staff. Corrected text should be returned to the editorial staff personally or by mail. Corrected the margin of copies received from the editorial staff or if needed on separate pages. In the case when the correction said time limit, correction will be performed by technical editorial staff of the publisher.

In case of changing of workplace or home address authors are asked to inform the editorial staff.