

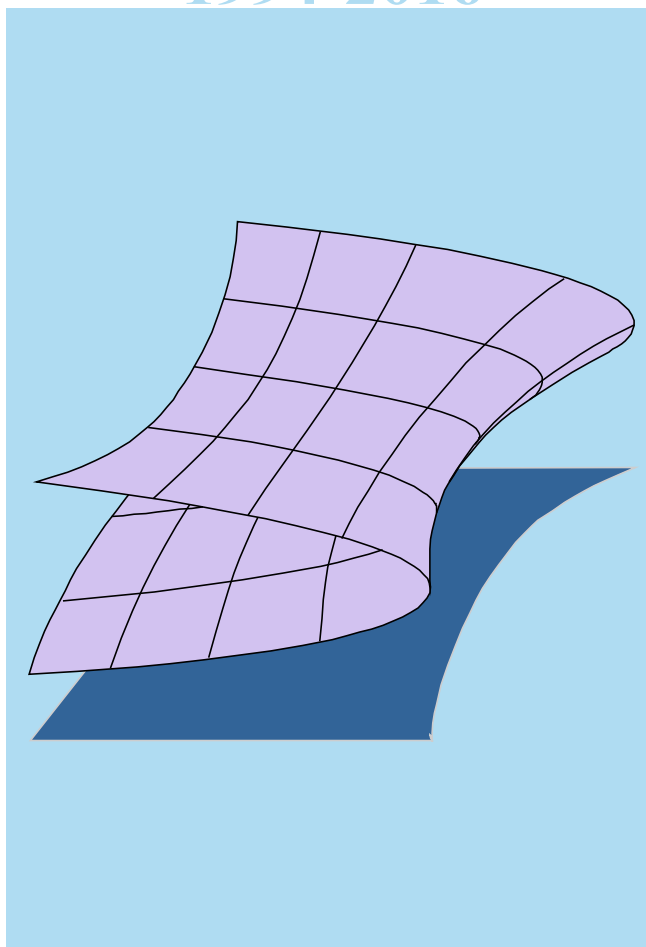
Проблемы Нелинейного Анализа в Инженерных Системах

Методы Подходы Гипотезы Решения

Международный журнал

Казанский национальный исследовательский технический университет им. А.Н. Туполева
(Казанский авиационный институт)

1994-2016



Losslessness in nonlinear Kirchhoff circuits and in relativity theory

Амплитудно-фазовая модуляция лазерного излучения в задачах генерации радиочастотных сигналов

Автоматизированное проектирование широкополосного диодного удвоителя частоты в диапазоне (26-40) ГГц

Моделирование нелинейного распространения мод в многомодовых волоконно-оптических линиях дальней связи

Об условиях трансверсальности в решении вариационных задач

Сравнение различных двигателей и методы повышения термического КПД

О конвертировании авиационных двигателей, отработавших ресурс в лётной эксплуатации

К проблеме жаропрочности рабочих лопаток турбины авиационных ГТД

To problem of science organization

О 70-летию исследований по механике в Казанском Научном Центре РАН

Advances in operational research for development

Казань

ПОЧЕТНЫЕ РЕДАКТОРЫ

В.Лакишмикантам, Президент МФНА, США

В.М.Матросов, Академик РАН, Президент АНН, Россия

И.Р.Пригожин, Лауреат Нобелевской Премии, Бельгия

РЕДАКТОРЫ

Г.Л.Дегтярев, Л.К.Кузьмина, КНИТУ-КАИ, РОССИЯ,

В.Ф.Журавлев, ИПМ РАН, РОССИЯ

РЕДАКЦИОННЫЙ КОМИТЕТ

В.В.Александров, МГУ, Москва, РОССИЯ

И.Антониу, Университет Аристотеля, Тессалоники, Греция

П.Борн, Главная Академия, Лилль, ФРАНЦИЯ

С.Н.Васильев, Институт Проблем Управления РАН, Москва, РОССИЯ

П.Дж.Вербос, Национальный научный Фонд, Вирджиния, США

И.Г.Горячева, Институт Проблем Механики РАН, Москва, РОССИЯ

Д.А.Губайдуллин, ИММ КазНЦ РАН, Казань, РОССИЯ

А.Л.Дончев, Американское математическое общество, Мичиган, США

Ю.Г.Евтушенко, Вычислительный Центр РАН, Москва, РОССИЯ

А.М.Елизаров, КФУ, Казань, РОССИЯ

М.С. Ел Нэйши, Университет Александрии, ЕГИПЕТ

В.Б.Колмановский, МГИЭМ (НИУ), Москва, РОССИЯ

П.С.Красильников, МАИ (НИУ), Москва, РОССИЯ

Джи Лин, Междунар. институт общей теории систем, Пенсильвания, США

А.Х.Найфе, Политехн.Институт, Гос.Университет, Вирджиния, США

Г.Николис, Университет, Брюссель, БЕЛЬГИЯ

Е. Онате, МЦВМК, Барселона, ИСПАНИЯ

В.Г.Пешехонов, “Концерн ЦНИИ Электроприбор”, Санкт-Петербург, РОССИЯ

Г.Г.Райкунов, ОРКК, Москва, РОССИЯ

Н.Х.Розов, МГУ, Москва, РОССИЯ

В.Ю.Рутковский, Институт Проблем Управления РАН, Москва, РОССИЯ

М.Х.Салахов, АН РТ, Казань, РОССИЯ

М.Самбандхам, МК, Атланта, США

А.Сидов, Национальный Исследовательский Центр, Берлин, ГЕРМАНИЯ

С.Я.Степанов, Вычислительный Центр РАН, Москва, РОССИЯ

А.Н.Тихонов, Гос.Инст.Информ.Технол. и Телекоммуникац., Москва, РОССИЯ

Ф.Л.Черноустько, Институт Проблем Механики РАН, Москва, РОССИЯ

Д.Д.Шилияк, Университет Санта Клара, Калифорния, США

О.А.Душина (ассист. редактора, переводы), КНИТУ им.А.Н.Туполева (КАИ), Казань, РОССИЯ

Главные цели Журнала -

- информирование специалистов соответствующих областей о состоянии в теории и приложениях в настоящий момент, о глобальных проблемах и актуальных задачах;
- содействие тесным рабочим контактам между учеными различных Университетов и Школ, теоретиками и прикладниками;
- "математизация" методов в решении задач, инициированных инженерной практикой;
- объединение усилий, синтез методов в решении проблем в различных сферах науки и образования.

В выпусках журнала публикуются обзоры и статьи, инженерные заметки, дискуссионные сообщения, постановки и решения задач во всех областях нелинейного анализа и их приложений в инженерных системах (в том числе, новые результаты, методы, подходы, решения, гипотезы, ...).

Авторы теоретических работ должны указать возможные области приложений в инженерной практике.

Языки публикаций - РУССКИЙ, АНГЛИЙСКИЙ, НЕМЕЦКИЙ, ФРАНЦУЗСКИЙ

Издание осуществляется в кооперации с Казанским Федеральным университетом (КФУ), с Московским авиационным институтом (национальным исследовательским университетом), с Международной научно-исследовательской лабораторией по нано- и биотехнологиям (INT).

HONORARY EDITORS

V.Lakshmikantham, *IFNA President, USA*

V.M.Matrossov, *RAS Academician, ANS President, Russia*

I.R.Prigogine, *Nobel Prize Laureate, Belgium*

EDITORS

G.L.Degtyarev, **L.K.Kuzmina**, *KNRTU-KAI, RUSSIA,*

V.F.Zhuravlev, *IPM, RAS, RUSSIA*

EDITORIAL BOARD

V.V.Alexandrov, *MSU, Moscow, RUSSIA*

I.Antoniou, *Aristotle University, Thessaloniki, GREECE*

P.Borne, *Lille Central Academy, FRANCE*

F.L.Chernousko, *Mechanics Problems Institute, RAS, Moscow, RUSSIA*

A.L.Dontchev, *American Mathematical Society, Michigan, USA*

A.M.Elizarov, *KFU, Kazan, RUSSIA*

M.S.El Naschie, *University of Alexandria, EGYPT*

Yu.G.Evtushenko, *RAS Computing Centre, Moscow, RUSSIA*

I.G.Goryacheva, *Mechanics Problems Institute, RAS, Moscow, RUSSIA*

D.F.Gubaidullin, *RAS KazSC IME, Kazan, RUSSIA*

V.B.Kolmanovsky, *MSIEM (NRU), Moscow, RUSSIA*

P.S.Krasilnikov, *MAI (NRU), Moscow, RUSSIA*

Yi Lin, *International Institute for General Systems Studies (IIGSS), PA, USA*

A.H.Nayfeh, *Virginia Polytechnic Institute, State University, USA*

G.Nicolis, *Free University, Brussels, BELGIUM*

V.G.Peshekhonov, *Concern CSRI Elektropribor, JSC, St. Petersburg, RUSSIA*

G.G.Raykunov, *ORCC, Moscow, RUSSIA*

N.Kh.Rozov, *MSU, Moscow, RUSSIA*

V.Yu.Rutkovsky, *Control Problems Institute, RAS, Moscow, RUSSIA*

M.Kh.Salakhov, *RT AS, Kazan, RUSSIA*

M.Sambandham, *MC, Atlanta, USA*

T.K.Serazetdinov, *KNRTU of A.N.Tupolev name (KAI), Kazan, RUSSIA*

D.D.Siljak, *Santa Clara University, California, USA*

S.Ya.Stepanov, *RAS Computing Centre, Moscow, RUSSIA*

A.Sydow, *GMD, Berlin, GERMANY*

A.N.Tikhonov, *State Inst.of Inform.Technol. and Telecomm., Moscow, RUSSIA*

S.N.Vasiliyev, *Control Problems Institute, RAS, Moscow, RUSSIA*

P.J.Werbos, *NSF, Virginia, USA*

O.A.Dushina (Assistant of Editor, translation), *KNRTU of A.N.Tupolev name (KAI), Kazan, RUSSIA*

Main goals of this Journal -

- to inform the specialists of appropriate fields about recent state in theory and applications; about global problems, and actual directions;
- to promote close working contacts between scientists of various Universities and Schools; between theorists and application oriented scientists;
- to mathematize the methods for solving the problems generated by engineering practice;
- to unite the efforts, to synthesize the methods in different areas of science and education.

In Journal the articles and reviews; the engineering notes; the discussional communications; the statements and solutions of problems in all areas of nonlinear analysis and their applications in engineering systems are published (including new results, methods, approaches, hypotheses,...). Authors of theoretical works should indicate the possible areas of applications in engineering practice.

The languages of publications are RUSSIAN, ENGLISH, GERMAN, FRENCH.

Edition is carried out in co-operation with Kazan Federal University (KFU), with Moscow Aviation Institute (National Research University), with International Nanobiological Testbed Ltd (INT).

От Редакционного Комитета

Единство – в разнообразии

Международный Журнал «Проблемы нелинейного анализа в инженерных системах» – *междисциплинарное двуязычное* научное периодическое Издание, представляющее исследования по нелинейным проблемам в целом, во всем разнообразии фундаментальных и прикладных наук, включая дисциплины естественного и гуманитарного циклов (математика, механика, физика, химия; инженерные, биологические, медицинские, социальные, политические науки; экология, космология; экономика и финансовая математика; также - нанонаука и нанотехнология, устойчивость и поддерживающее развитие, проблемы риска и защиты информации, проблемы исследования операций, ...).

В этом выпуске журнала «Проблемы нелинейного анализа в инженерных системах» (№2(46), т.22, 2016) опубликованы статьи, аналитические обзоры, научно-информационные материалы, отражающие видение специалистов, выделяющие некоторые актуальные проблемы настоящего и будущего *междисциплинарного* характера. Выделены два специальных раздела.

Один из них связан с результатами, представленными профессором О.Г.Морозовым, специалистом в области радиофотоники, представителем известных отечественных научных Школ в области квантовой электроники (Ю.Е. Польский и Г.И. Ильин) и в области микроволновых технологий (Г.А. Морозов и Ю.Е. Седельников), Приглашенным Редактором по этому разделу. Тематика представленных им статей связана с проблемами нелинейного анализа и синтеза радиофотонных систем и устройств. Она является очень важной как с теоретической, так и с прикладной точек зрения. Здесь опубликована часть работ этой серии:

О.Г. Морозов, Г.И. Ильин, Г.А. Морозов. Амплитудно-фазовая модуляция лазерного излучения в задачах генерации радиочастотных сигналов.

В.В. Березин, А.М. Щитов, А.С. Нечаев. Автоматизированное проектирование широкополосного диодного удвоителя частоты в диапазоне (26÷40) ГГц.

В.А. Бурдин, А.В. Бурдин. Моделирование нелинейного распространения мод в многомодовых волоконно-оптических линиях дальней связи.

Другая часть статей готовится к публикации в следующих выпусках.

Второй специальный раздел связан с результатами, представленными профессором Б.Г.Мингазовым, специалистом в области теории горения и воздушно-реактивных двигателей, представителем известных отечественных научных Школ по физико-техническим проблемам двигателей и энергетических установок летательных аппаратов (С.В.Румянцев, А.В.Талантов, Б.С.Виноградов, В.Е. Алемасов, А.Ф. Дрегалин), Приглашенным Редактором по этому разделу. Тематика представленных им статей связана с проблемами моделирования, анализа и синтеза для процессов и газодинамических течений в двигателях и энергетических установках, порожденных потребностями науки и инженерных приложений. В этом выпуске опубликована часть работ этой серии:

В.И. Панченко, Ли Цзывань. Сравнение различных двигателей и методы повышения термического КПД.

П.Г. Великанов, Чжан Бин. О конвертировании авиационных двигателей, отработавших ресурс в лётной эксплуатации.

Н.П. Великанова, Фань Вэйвэй. К проблеме жаропрочности рабочих лопаток турбины авиационных ГТД.

Другая часть статей готовится к публикации в следующих выпусках.

Здесь в целом: изучаемые объекты – *междисциплинарные* системы, требующие знаний из различных научных областей. Именно на стыке различных дисциплин происходит зарождение новых гипотез, обеспечивающих глубокое познание окружающего Мира, с *пониманием* происходящих процессов.

В этом отношении следует подчеркнуть, что *фундаментальной наукой о моделировании* сложных инженерных систем *является теоретическая механика*, которой принадлежит первостепенная роль в развитии методологии моделирования.

«Теоретическая механика - фундаментальная наука; она является ключевым предметом в подготовке инженеров, математиков, прикладников и физиков--теоретиков,...Для инженеров она является основой их специальных наук; для математиков – дорогой к современным обобщениям; для физиков - «прелюдией» к теории относительности, к статистической и квантовой механике», – R.Xill, Principles of Dynamics,Oxford,1964.

«Механика - это сплав математики со здравым смыслом», – Проф. М.Т. Нужин, представитель Казанской Четаевской Школы механики и устойчивости, Ректор Казанского университета.

Более того, в моделировании, как говорил великий И.Ньютон:

«вся трудность состоит в том, чтобы по явлениям движения распознать силы природы; а потом по этим силам изъяснить остальные явления», – (И.Ньютон).

Именно Механика, *фундаментальная наука об Искусстве Моделирования*, позволяет овладеть этим **искусством**, этим «know- how»... (причем – для любой области Знания).

Представленные в этом выпуске материалы, несомненно, будут способствовать объединению усилий теоретиков и прикладников, синтезу методов в решении проблем в междисциплинарных сферах науки и образования по направлениям, порожденным потребностями науки и инженерных приложений, в полном соответствии с девизом Международной Федерации Нелинейных Аналитиков –

«Единство – в разнообразии» (Профессор В.Лакшмикантам).

From the Editorial Board

Unity – in Diversity

International IFNA-ANS scientific Journal “Problems of nonlinear Analysis in Engineering Systems” is interdisciplinary scientific periodic Edition, presenting the works on nonlinear problems in all areas of fundamental and applied Sciences, including both natural and humanities disciplines: mathematics, mechanics, physics, chemistry; engineering, biological, medical, social, political sciences; ecology, cosmology; economics and financial mathematics; nanoscience and nanotechnology; stability and sustainable development, problems of risk and information protection, problems of operational research,...

In this issue of Journal “Problems of nonlinear Analysis in Engineering Systems” (No.2(46), Vol.22, 2016) it is published articles, analytical surveys, scientific-information papers that are reflecting the views of specialists and highlighting some urgent problems of present and future of interdisciplinary character. The issue has two special sections.

The first scientific section is connected with results of researches represented by Professor O.G. Morozov, the specialist in microwave photonics, a representative of well-known scientific Schools in the area of quantum electronics (Y.E. Polskii and G.I. Il'in) and the area of microwave technologies (G.A. Morozov Y.E. Sedel'nikov), the Guest Editor for this section. Subjects of articles presented by him are related to the problems of nonlinear analysis and synthesis of microwave photonics systems and devices. It is very important both from the theoretical and applied points of view.

A part of the works from this section is published here:

O.G. Morozov, G.I. Il'in, G.A. Morozov. Amplitude-phase modulation of laser radiation in generation problems of radio-frequency signals.

V.V. Berezin, A.M. Schitov, A.S. Nechaev. Computer-aided design of a broadband diode frequency doubler in the range (26÷40) GHz.

V.A. Burdin, A.V. Bourdine. Modelling nonlinear propagation of modes in long-haul multimode fiber optic links.

Another part of the works is preparing to publication in the next issues.

The second special section is prepared on the results, presented by Professor B.G.Mingazov, the specialist in the area of combustion theory and jet engines, a representative of well-known scientific Schools on physical and technical problems of aircraft engines and power plants (S.V.Rumyantsev, A.V. Talantov, B.S.Vinogradov, V.E. Alemasov, A.F. Dregalin), the Guest Editor for this section. Subjects of works presented by him are related to problems of modeling, analysis and synthesis for processes and gas-dynamic flows in engines and power plants, generated by the needs of science and engineering applications.

A part of work from this section is published in this issue:

V.I. Panchenko, Li Ziwan. Comparison of different engines and the methods of increasing the thermal efficiency.

P.G. Velikanov, Zhang Bing. About conversion of aircraft engines after life cycle in the flight operation.

N.P. Velikanova, Fan Weiwei. To heat-temperature strength problem of aviation gas turbine engine blades.

Another part of the works is preparing to publication in the next issues.

Here in whole: the studied objects are *interdisciplinary* systems requiring Knowledge from various scientific fields. Exactly the crossing of different disciplines gives the emerging of new hypotheses that provide a deep Knowledge of the around World, with the *understanding* of processes dynamics.

In this regard we emphasize that the *fundamental science for modelling* engineering systems is *theoretical mechanics*, which plays the primary role in the methodology development of modelling.

"Theoretical mechanics is a fundamental science; it is a key subject in training engineers, mathematicians, physicist -applied oriented and physicist-theorist... For engineers it is a base of their special sciences; for mathematicians it is "way" to modern generalizations; for physicists – it is "prelude" to theories of relativity, to statistical and quantum mechanics ", - R.Xill, Principles of Dynamics, Oxford, 1964.

"Mechanics – it is "an alloy" of mathematics with good senses" ..., - Prof. M.T. Nuzhin, Representative of Kazan Chetayev School of Mechanics and Stability, Rector of Kazan University.

Moreover, in regard to modelling great Isaak Newton noted:

"...the all difficulty in whole is contained in what, that it is necessary on motion phenomena to understand the forces of Nature, and after on these forces – to explain the another phenomena", – Isaak Newton.

And exactly Mechanics, **a fundamental science on ART of Modelling for any area of the Knowledge**, allows to possess *this Art, this "know-how"*...

The works presented in this issue will undoubtedly contribute the uniting efforts of theorists and applied oriented scientists, with synthesis of methods for solving problems in the interdisciplinary fields of science and education, in full accordance with the motto of International Federation of nonlinear analysts:

«Unity – in Diversity» (Professor Lakshmikantham)

Представление работ

Статьи, предназначенные для публикации в журнале, должны быть поданы в трех экземплярах. Статьи направляются по указанному ниже адресу или тому члену редакционного комитета, который, по мнению автора, наиболее близок к теме работы.

Адрес: Л.К.Кузьмина, Казанский авиационный институт (КНИТУ им.А.Н.Туполева)
Адамюк, 4-6, Казань-15, 420015, РОССИЯ
Lyudmila.Kuzmina@kpfu.ru
Тел.: (7) (843) 236-16-48
<http://pnaes.kpfu.ru/science/journals/ansj/>
http://www.kcn.ru/tat_en/science/ans/journals/ansj.html

Информация о подписке

“Проблемы нелинейного анализа в инженерных системах”, 2017, т.23 (два выпуска), ISSN 1727-687X.

Стоимость годовой подписки - 6600 руб. (включая пересылку) за любой год с 1996г.

Банковские реквизиты для платежа:

ФГБОУ ВПО КНИТУ им.А.Н.Туполева - КАИ
УФК по РТ (КНИТУ-КАИ л/с 20116X02750)
ИНН 1654003114 КПП 165501001 БИК 049205001
р/с 40501810292052000002
Отделение – НБ Республика Татарстан (БИК 049205001)
(X – печатается латинская буква).

с указанием: Для МНЖ “Проблемы нелинейного анализа в инженерных системах”.

Пожалуйста, информируйте нас о перечислении и сообщите номер платежного поручения по электронной почте или другим способом

Manuscript Submission

Manuscripts for publication in this Journal should be submitted in triplicate to the Editorial Office or to an individual member of the Board of Associate Editors who, in the opinion of the authors, is more closely involved with the topic of the paper.

The address of the Editorial Office:

Dr.Lyudmila Kuzmina, Kazan Aviation Institute (KNRTU of A.N.Tupolev's name)
Adamuck, 4-6, Kazan-15, 420015, RUSSIA
Lyudmila.Kuzmina@kpfu.ru
Tel.: (7) (843) 236-16-48
<http://pnaes.kpfu.ru/science/journals/ansj/>
http://www.kcn.ru/tat_en/science/ans/journals/ansj.html

Subscription information:

“Problems of Nonlinear Analysis in Engineering Systems”, 2017, vol.23 (two issues), ISSN 1727-687X.

Annual subscription rate: US\$200 (subscription rates include postage/air speed delivery), for any year from 1996.

Please, send this payment to:

Kazan State Technical University
S.W.I.F.T. SABRRUMMNA1
SBERBANK
(VOLGO-VYATSKY HEAD OFFICE
NIZHNIY NOVGOROD)
ACCOUNT 40503840762020200019
FOR CREDIT TO KGTU ANTUPOLEVA

with indication: For ISE “Problems of Nonlinear Analysis in Engineering Systems”.
Please, inform us about this transfer and the wire number by e-mail.

Losslessness in nonlinear Kirchhoff circuits and in relativity theory

A. Fettweis

Department of Electrical Engineering and Information Science/Communications
Engineering, Ruhr-Universität Bochum,
D-44780 Bochum, Germany

Abstract. Kirchhoff circuits are of importance not only for studying electrical phenomena but are ideally suited to model a broad range of physical systems for purposes where conservation of power and energy and related concepts such as passivity and losslessness are essential. They consist of interconnections of a variety of elements, which nowhere have to be linear and constant. If an element such as an inductance is nonlinear and/or explicitly time dependent and is to be characterized as being passive or, more specifically, lossless, its defining relation must have a specific form, but the classical relation for a relativistic mass is not of this type. It is shown that, preserving classical relativistic kinematics, requiring relativistic dynamics to approach the Newtonian one for appropriate limits, and putting prime emphasis on work done, thus on energy rather than momentum, one is naturally led to an expression for force in terms of mass and velocity whose form is in full agreement with that referred to for a nonlinear inductance. This alternative way of modifying Newton's second law requires Newton's third law to be also modified. These two modifications combined produce the same conservation of momentum and the same dynamics of particles in fields as classical relativity. The expression for kinetic energy, however, is different. Logically consistent derivations are presented, and a theoretical and an experimental result are pointed out that tend to offer some support to the alternative theory, or at least do not contradict it, as implausible as that theory may a priori appear to be. The paper complements and updates earlier results on the subject and improves the presentation.

Keywords: Newton's laws, relativity theory, Ilya Prigogine, Kirchhoff circuits

1. Introduction

This paper is dedicated to the memory of the great physicist and chemist Ilya Prigogine. His outstanding contributions have touched my own activities in at least two important ways: On the one hand he is particularly known for his work on non-equilibrium thermodynamics, thus an area in which the theory of an entirely different discipline, i.e. that of Kirchhoff circuits [1, 2], can play a helpful role [3]. On the other hand, he is one of the outstanding Russian émigrés who came to Belgium as boys in the aftermath of the Bolshevik revolution and who have contributed so much to the scientific life of that country. Another one of those was Vitold Belevitch [4], who has been the dominant figure in the theory of Kirchhoff circuits during the second half of the twentieth century, certainly in Europe and, in the opinion of many, also worldwide, and who is the grand master to whom this author owes the privilege of having been introduced into the beauty, generality, and rigor of that theory. The present paper shows how the theory of Kirchhoff circuits can throw some new light even on certain aspects of relativity theory.

Kirchhoff circuits are of importance not only in the context of modelling electric phenomena but also for a variety of other reasons. As an example, they can serve as reference circuits for applying wave-digital concepts, in particular to filtering [5] and numerical integration of ordinary [6] and, especially, partial differential equations describing physical systems [7 - 9]. While for filtering one is usually interested in linear circuits, numerical integration largely addresses nonlinear phenomena.

Kirchhoff circuits are composed of elements that are interconnected by merging element terminals into nodes. The interconnections are described by the Kirchhoff voltage and current laws. A crucial consequence of this is that conservation of power and thus energy is strictly guaranteed by the validity of the Kirchhoff laws. Hence, properties such as passivity and losslessness (which is a special case of passivity) are in turn strictly ensured if the corresponding properties hold for all the elements. In order to guarantee in a simple way that an element such as a nonlinear inductance is lossless, however, it should be represented in a somewhat different way than what is usually done [7 - 9], and this alternative representation turns out to be of rather fundamental physical importance.

Mass is, in a sense, also a lossless (conservative) element. In Newtonian mechanics this element is linear, but in relativistic mechanics it is nonlinear. Surprisingly, however, the relationship between force and velocity in classical relativity theory [10 - 19] does not have the general form one might expect in view of what has just been said about a nonlinear inductance.

In previous papers [20, 21] it has been shown that by making certain assumptions that at least are not a priori unreasonable one is naturally led to a relativistic expression between force, mass, and velocity that differs somewhat from the classical one but has a form that is in full agreement with the general form just referred to. This discrepancy is essentially due to the fact that in classical relativity, momentum is the quantity placed at the beginning of the derivations for establishing the laws of dynamics, while the alternative approach puts prime emphasis on energy and work done. There is no change concerning the use of the Lorentz transformation and thus the validity of relativistic kinematics, which is indeed known to be the result of intellectually highly satisfying and strict logical deductions. We thus maintain the second Einstein postulate, i.e. the one concerning the universality of the speed of light. Similarly, we uphold also the first one, i.e. the one concerning the universality of the fundamental laws of nature. In a sense, we do however complement the two Einstein postulates by what we will refer to as the *principle of Newtonian limit*.

The alternative modification of Newton's 2nd law (Section 4) requires a corresponding modification of his 3rd law (Section 5), although the latter remains untouched in classical relativity. The two changes combined imply that conservation of momentum holds as in classical relativity (Section 5) and that the dynamics of particles in fields (e.g. in particle accelerators) is also the same (Section 6). An expression, obtainable by standard approaches, for the energy of an electrostatic field observed in a moving reference frame (Section 7) agrees with the alternative theory, and so does, if properly interpreted, an extension to general electromagnetic fields. The Bertozzi experiment (Section 8) does not, to say the least, disprove it either, although that experiment involves kinetic energy. The expression for that quantity is indeed different in the two cases so that there have to exist ways of differentiating.

The present paper pursues several prime purposes. Firstly, the main result, i.e. the derivation of the alternative relation between force, velocity, and mass, is carried out in a somewhat different way than previously done, thus offering an even firmer basis for justifying that expression. Secondly, the Bertozzi experiment is analyzed more precisely than has been done in [21] and is shown to actually have a tendency of favoring the alternative approach, although the accuracy of the experiment is not high enough to allow us to draw definite conclusions. Thirdly, several inaccuracies in [21] are corrected, but we keep aiming at the same generality. Fourthly, the present paper is not only an improved but also a substantially expanded version of [22] and should thus be more easily accessible. Of course, as in [21] but contrary to what had been done in [20], we do not impose any restrictions on the integration constant occurring in relation with the energy expression. All previous results that are essential for understanding the paper are included for ease of presentation and thus for the benefit of the reader.

Clearly, the entire issue remains puzzling. It is hoped that the results presented hereafter will stimulate further experimental and theoretical investigations that bring about the needed clarification.

2. Passivity and losslessness in nonlinear Kirchhoff circuits

A nonlinear inductance that depends exclusively on its current, i , can be described by either one of the equations

$$u = D(L_g i) \quad \text{or} \quad u = L_l D i \quad (2.1)$$

where u is the voltage across the inductance and

$$D = d/dt, \quad (2.2)$$

$L_g = L_g(i) = \phi/i$ being the *global inductance* and $L_l = L_l(i) = d\phi/di$ the *local inductance*, and ϕ the magnetic flux. It can be shown that $L_g(i) \geq 0 \quad \forall i$ is necessary but not sufficient for guaranteeing passivity (in fact losslessness) while $L_l(i) \geq 0 \quad \forall i$ is sufficient but not necessary. In order to obtain a satisfactory characterization Meerkötter [23, 24] had suggested to represent a nonlinear inductance by means of a nonlinear ideal transformer terminated in a linear inductance, but a simpler, equivalent way is to write (2.1) in the form [7 - 9]

$$u = \sqrt{L} D(\sqrt{L} i) = \frac{1}{2} (D(L i) + L D i). \quad (2.3)$$

The two expressions in (2.3) correspond, in a sense, to the geometric and the arithmetic mean of the definitions (2.1). For the power absorbed and the stored energy, W_L , one finds from (2.3),

$$ui = DW_L, \quad W_L = \frac{1}{2} L i^2, \quad (2.4)$$

and for the *inductance*, L , the inequality $L(i) \geq 0$ is now necessary and sufficient for passivity (in fact, losslessness).

Writing (2.3) in the form $x = Dy$, $x = u/\sqrt{L}$, $y = i\sqrt{L}$, approximating D by means of the trapezoidal rule with step size T , and designating the discretized time variable by t_n , the result can be written in the form

$$b(t_n) = -a(t_n - T) \quad (2.5)$$

where the so-called *waves* (wave quantities), a and b , are given by

$$a = \frac{u + iR}{2\sqrt{R}}, \quad b = \frac{u - iR}{2\sqrt{R}}, \quad R = \frac{2L}{T}. \quad (2.6)$$

We obviously also have

$$ui = a^2 - b^2. \quad (2.7)$$

All this holds more generally if L is a function not only of i but of any of the dependent variables in the circuit and/or of the independent variable, say t . It holds even for multidimensional Kirchhoff circuits [7 - 9], in which case D is a partial differential operator and t has to be replaced by a vector of independent variables.

The simplicity of expressions such as (2.5) and (2.7) and their combination with passivity and losslessness aspects are the essential reasons for the advantageous properties that can be obtained by making use of wave-digital principles for filtering [5] and numerical integration of ordinary and partial differential equations [6 - 9]. The inductance representation (2.3) can

be extended to coupled inductances, thus to matrix inductances \mathbf{L} , that matrix being symmetric and non-negative definite. With \mathbf{u} and \mathbf{i} being vector extensions of u and i one can then write,

$$\mathbf{u} = \frac{1}{2}(\mathbf{L}\mathbf{D}\mathbf{i} + \mathbf{D}(\mathbf{L}\mathbf{i})) \quad \text{or} \quad \mathbf{u} = \mathbf{L}^{T/2}\mathbf{D}(\mathbf{L}^{1/2}\mathbf{i}) \quad (2.8)$$

where the matrix $\mathbf{L}^{1/2}$ is such that $\mathbf{L}^{T/2}\mathbf{L}^{1/2} = \mathbf{L}$, the superscript T designating transposition and $\mathbf{L}^{T/2} = (\mathbf{L}^{1/2})^T$. The two expressions in (2.8) are equivalent if and only if $\mathbf{L}^{T/2}\mathbf{D}(\mathbf{L}^{1/2}) = (\mathbf{D}\mathbf{L}^{T/2})\mathbf{L}^{1/2}$. The general conditions under which the latter equality holds are not known [25], but it is sufficient that $\mathbf{L}^{1/2}$ is a diagonal matrix times a constant matrix (cf. the situation encountered later in relation with (3.13)).

The wide importance energy, wave, and scattering concepts have in physics suggests that relations such as (2.3) and (2.4) have some fundamental physical importance. This observation, however, is not honored by the classical theory of special relativity.

3. Relativistic mass

For a particle of rest mass m_0 , let

$$\mathbf{f} = (f_x, f_y, f_z)^T, \quad \mathbf{p} = (p_x, p_y, p_z)^T, \quad \mathbf{v} = (v_x, v_y, v_z)^T \quad (3.1)$$

be the force acting upon it, its momentum, and its velocity, respectively. Define \mathbf{v} , β , and α by

$$\mathbf{v}^T \mathbf{v} = v^2, \quad \beta = v/c, \quad \alpha = \sqrt{1 - \beta^2} \quad (3.2)$$

where c is the speed of light. According to the basic principles adopted in classical relativity theory the following holds (cf. (2.2)),

$$\mathbf{f} = \mathbf{D}\mathbf{p}, \quad \mathbf{p} = m_g \mathbf{v}, \quad m_g = m_0/\alpha \quad (3.3)$$

where we have written m_g instead of m in order to conform with the notation in (2.1). From (3.3) one finds for the power delivered to the particle, as is well known,

$$\mathbf{v}^T \mathbf{f} = \mathbf{D}(m_g c^2), \quad (3.4)$$

and this by making use of the identity

$$c^2 \mathbf{D} \frac{1}{\alpha} = \mathbf{v}^T \mathbf{D} \left(\frac{1}{\alpha} \mathbf{v} \right). \quad (3.5)$$

The quantity $m_g c^2$ is then interpreted as total energy while $(m_g - m_0)c^2$ is the kinetic energy.

Clearly, m_g as defined by (3.3) does indeed correspond to L_g in (2.1), not to L in (2.3). This is somewhat surprising and, as explained in [20 - 22], the reason for examining whether an expression for the *force* such as

$$\mathbf{f} = \sqrt{m} \mathbf{D}(\sqrt{m} \mathbf{v}) = \frac{1}{2}(\mathbf{D}(m \mathbf{v}) + m \mathbf{D} \mathbf{v}) \quad (3.6)$$

where the *mass* m is some function of v , say $m = m(\beta)$, could be of interest. In the affirmative, (3.4) would be replaced by

$$\mathbf{v}^T \mathbf{f} = \mathbf{D}E_k = \mathbf{D}E \quad (3.7)$$

where

$$E = E_i + E_k, \quad E_k = mv^2/2, \quad E_i = \text{const.} \quad (3.8)$$

Since $E_k = 0$ for $v = 0$, E_k would be the *kinetic energy*, thus E_i the *rest energy* (internal energy) and E the *total energy*. One could then assume m , not m_g , to be the quantity of prime interest. In particular, defining a rest mass $m_0 = m(0)$ one would want to preserve essential properties of classical relativity, although for m instead of m_g , and require that $m - m_0$ increases with E_k , preferably in a linear fashion, i.e., according to

$$E_k = K(m - m_0), \quad K = \text{const.},$$

and that $m = \infty$ for $v = c$. It then follows from the above expression for E_k (c.f. (3.8)) that $K = c^2/2$, that m must thus be of the form

$$m = \frac{m_0}{1 - \beta^2} = \frac{m_0}{\alpha^2}, \quad (3.9)$$

and that therefore $E_k = \frac{1}{2}(mc^2 - m_0c^2)$.

In view of $E_k = E - E_i$, that last result had suggested, in the earlier paper [20], to identify E with $mc^2/2$ and thus E_i with $m_0c^2/2$. This, however, leads to difficulties, while nothing contradicts the above equations if one does not impose any requirement on E_i , i.e., on the integration constant implied by (3.7). E and E_k are then still given by (3.8), thus E_k by

$$E_k = \frac{1}{2}mv^2 = E_0 \frac{\beta^2}{1 - \beta^2} = E_0 \frac{\beta^2}{\alpha^2}, \quad E_0 = \frac{1}{2}m_0c^2, \quad (3.10)$$

with m as defined in (3.9). The case discussed in [20] corresponds to choosing $E_i = E_0$. Adopting the present point of view, however, amounts to saying that a particle is, in general, characterized by two constants, E_i and E_0 . For $E_i = 2E_0$, the rest energy is $E_i = m_0c^2$, as in the classical case. In general, however, no simple relation has to exist between E_i and E_0 .

An attractive observation follows directly from (3.10). Using $v^2 = v_x^2 + v_y^2 + v_z^2$ the expression for E_k can indeed be decomposed into a sum of kinetic *energy components* as in Newtonian dynamics according to

$$E_k = E_{kx} + E_{ky} + E_{kz}$$

$$E_{kx} = \frac{1}{2}mv_x^2, \quad E_{ky} = \frac{1}{2}mv_y^2, \quad E_{kz} = \frac{1}{2}mv_z^2.$$

It is true that the triple $\mathbf{E}_k = (E_{kx}, E_{ky}, E_{kz})^T$ does not form a vector in the strict physical sense (it transforms in a more complicated way than coordinates do, contrary to the momentum), but this should not be of any concern. Nevertheless this observation may partly explain why classical relativity gives preference to momentum for establishing the foundations of dynamics theory. In fact, the emphasis placed on energy components (i.e., not only the three kinetic ones mentioned above) is also of fundamental importance for the wave-digital method of numerically integrating partial differential equations [7 - 9].

By means of \mathbf{f} one can define the quadruple

$$\begin{pmatrix} \mathbf{f} \\ \frac{1}{c} \mathbf{v}^T \mathbf{f} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \frac{1}{c} D \mathbf{E}_k \end{pmatrix}, \quad (3.11)$$

for which in (3.11) a second, equivalent expression in terms of \mathbf{E}_k (cf. (3.10)) is given and which can directly be compared to well-known results in relativity theory [10, 13 - 16, 19]. It can then easily be verified that this quadruple is identical to the four-vector (world vector, Minkowski force) originally introduced by Minkowski [27] and thus has a Lorentz invariant form. One immediate conclusion of this is that a Lorentz transformation (cf. Section 4) changes only that component of \mathbf{f} that is oriented in the direction of the relative movement of the two reference frames with respect to each other, i. e., the components of \mathbf{f} perpendicular to that direction remain unchanged, contrary to what holds in classical relativity theory.

Like for inductances one can go even a step further than what we have done so far and consider a particle-like object characterized, in addition to \mathbf{E}_i , by a constant nonnegative definite symmetric matrix \mathbf{m}_0 instead of simply a scalar mass m_0 . The scalar mass m has then to be replaced by the *matrix mass* \mathbf{m} given by (cf. (3.9)),

$$\mathbf{m} = \frac{1}{\alpha^2} \mathbf{m}_0. \quad (3.12)$$

Expression (3.6) for \mathbf{f} becomes

$$\mathbf{f} = \mathbf{m}_0 \frac{1}{\alpha} D \left(\frac{1}{\alpha} \mathbf{v} \right) = \frac{1}{\alpha} D \left(\mathbf{m}_0 \frac{1}{\alpha} \mathbf{v} \right) = \frac{1}{2} (D(\mathbf{m} \mathbf{v}) + \mathbf{m} D \mathbf{v}) = \frac{1}{\alpha} \mathbf{m}_0^{T/2} D \left(\frac{1}{\alpha} \mathbf{m}_0^{1/2} \mathbf{v} \right), \quad (3.13)$$

the adopted notation corresponding to that defined in relation with (2.8).

For the power delivered in the matrix case, (3.7) remains unchanged, but in (3.8) and (3.10) the expression for \mathbf{E}_k has to be replaced by

$$\mathbf{E}_k = \frac{1}{2} \mathbf{v}^T \mathbf{m} \mathbf{v} = \frac{1}{2\alpha^2} \mathbf{v}^T \mathbf{m}_0 \mathbf{v}. \quad (3.14)$$

The matrix case reduces to the scalar one if $\mathbf{m}_0 = \mathbf{1} m_0$ where $\mathbf{1}$ is the unit matrix. An irreducible example for the matrix case will be mentioned in Section 7.

4. Direct derivation of the alternative way of modifying Newton's second law

4.1. General relations

In view of the surprising results of Section 3 it appears opportune to examine whether there exists a way of deriving them, and possibly further related ones, by some direct approach. Clearly, since energy played an essential role in Section 3, such a direct approach would have to be centered on energy and work done rather than on momentum, as in classical relativity.

Consider thus two reference frames, \mathbf{S} and \mathbf{S}' , with \mathbf{S}' moving with constant velocity \mathbf{v}_0 , say, in the \mathbf{x} -direction of \mathbf{S} . For \mathbf{S} we use unprimed notation as in Section 3 (coordinates, velocities, forces, differential operators etc.) and for \mathbf{S}' the corresponding primed notation. Between \mathbf{S} and \mathbf{S}' we thus have the Lorentz transformation [10 -19]

$$\mathbf{x}' = \frac{1}{\alpha_0} (\mathbf{x} - \mathbf{v}_0 t), \quad y' = y, \quad z' = z, \quad t' = \frac{1}{\alpha_0} \left(t - \beta_0 \frac{\mathbf{x}}{c} \right) \quad (4.1)$$

where

$$\beta_0 = \mathbf{v}_0 / c, \quad \alpha_0 = \sqrt{1 - \beta_0^2}. \quad (4.2)$$

We consider a particle, \mathbf{P} , moving as discussed in Section 3. For \mathbf{P} it is appropriate to introduce vectors of position coordinates

$$\mathbf{r} = \mathbf{r}(t) = (x, y, z)^T, \quad (4.3)$$

$$\mathbf{r}' = \mathbf{r}'(t') = (x', y', z')^T,$$

while t' can be uniquely expressed in terms of t , say $t' = t'(t)$, and vice versa. While (3.1) and (3.2) refer to \mathbf{S} , we have, with respect to \mathbf{S}' ,

$$\mathbf{f}' = (f'_x, f'_y, f'_z)^T, \quad \mathbf{v}' = (v'_x, v'_y, v'_z)^T, \quad (4.4)$$

$$\mathbf{v}'^T \mathbf{v}' = v'^2, \quad \beta' = v'/c, \quad \alpha' = \sqrt{1 - \beta'^2}, \quad (4.5)$$

where (cf. (2.2))

$$\mathbf{v} = D\mathbf{r}, \quad \mathbf{v}' = D'\mathbf{r}' \quad (4.6)$$

$$D = d/dt, \quad D' = d/dt'. \quad (4.7)$$

As can be verified by means of (4.1), we have, as known [10 - 19],

$$\frac{dt}{dt'} = \frac{\alpha'}{\alpha} = \frac{\alpha'_x}{\alpha_x}, \quad \alpha'D = \alpha D' \quad (4.8)$$

where, in addition to quantities already defined,

$$\alpha_x = \sqrt{1 - \beta_x^2}, \quad \alpha'_x = \sqrt{1 - \beta'^2_x}, \quad \beta_x = \frac{v_x}{c}, \quad \beta'_x = \frac{v'_x}{c}, \quad (4.9)$$

and furthermore,

$$v'_x = \frac{v_x - v_0}{1 - \beta_x \beta_0}, \quad v'_y = \frac{\alpha'}{\alpha} v_y, \quad v'_z = \frac{\alpha'}{\alpha} v_z. \quad (4.10)$$

From these expressions, the following equalities can be derived:

$$\frac{1}{\alpha'^3} D'v'_x = \frac{1}{\alpha^3} Dv_x = D \frac{v_x}{\alpha} + \frac{v_y^2 + v_z^2}{\alpha^3 c^2} Dv_x - \frac{v_x}{\alpha^3 c^2} (v_y Dv_y + v_z Dv_z), \quad (4.11)$$

$$D'v'_y = \frac{\alpha'^2}{\alpha} D \frac{v_y}{\alpha} + v'_y \frac{1}{\alpha} D\alpha', \quad (4.12)$$

$$D'v'_z = \frac{\alpha'^2}{\alpha} D \frac{v_z}{\alpha} + v'_z \frac{1}{\alpha} D\alpha', \quad (4.13)$$

$$D(v_x/\alpha_x) = D'(v'_x/\alpha'_x). \quad (4.14)$$

4.2. Newton's second law

Newton's first law is purely qualitative and always valid; it thus is irrelevant for our purpose. His second law concerns the expression of the force. For examining it we consider two reference frames \mathbf{S} and \mathbf{S}' as discussed in Section 4.1 and a particle \mathbf{P} as so far. Let t_1 and t_2 be two time instants that are arbitrary close yet distinct, say $t_2 > t_1$. Quantities referring to t_1 and t_2 will be given subscripts 1 and 2, respectively, thus, e. g.,

$$\begin{aligned} t'_1 &= t'(t_1), & t'_2 &= t'(t_2), & t_1 &= t(t'_1), & t_2 &= t(t'_2), \\ \mathbf{r}_1 &= \mathbf{r}(t_1), & \mathbf{r}_2 &= \mathbf{r}(t_2), & \mathbf{r}'_1 &= \mathbf{r}'(t'_1), & \mathbf{r}'_2 &= \mathbf{r}'(t'_2), \\ \mathbf{v}_1 &= \mathbf{v}(t_1), & \mathbf{v}_2 &= \mathbf{v}(t_2), & \mathbf{v}'_1 &= \mathbf{v}'(t'_1), & \mathbf{v}'_2 &= \mathbf{v}'(t'_2), \\ x_1 &= x(t_1), & x_2 &= x(t_2), & v_{x1} &= v_x(t_1), & v_{x2} &= v_x(t_2) \end{aligned}$$

etc. We also define

$$\Delta \mathbf{t} = \mathbf{t}_2 - \mathbf{t}_1, \quad \Delta \mathbf{r} = (\Delta x, \Delta y, \Delta z)^T = \mathbf{r}_2 - \mathbf{r}_1, \quad (4.15)$$

$$\Delta \mathbf{t}' = \mathbf{t}'_2 - \mathbf{t}'_1, \quad \Delta \mathbf{r}' = (\Delta x', \Delta y', \Delta z')^T = \mathbf{r}'_2 - \mathbf{r}'_1 \quad (4.16)$$

etc. and use the simplified notation defined by

$$\mathcal{O}^n = \mathcal{O}\left((\Delta \mathbf{t})^n\right) = \text{order of } (\Delta \mathbf{t})^n \text{ for } n \geq 0.$$

Due to $\mathbf{t}' = \mathbf{t}'(\mathbf{t})$ and $\alpha d\mathbf{t} = \alpha' d\mathbf{t}'$ (cf. (4.8)) we have

$$\Delta \mathbf{t}' = \mathbf{t}'(\mathbf{t}_1 + \Delta \mathbf{t}) - \mathbf{t}'_1 = \left(\frac{d\mathbf{t}'}{d\mathbf{t}}\right)_1 \Delta \mathbf{t} + \mathcal{O}^2 = \frac{\alpha_1}{\alpha'_1} \Delta \mathbf{t} + \mathcal{O}^2$$

and therefore

$$\alpha'_1 \Delta \mathbf{t}' = \alpha_1 \Delta \mathbf{t} + \mathcal{O}^2, \quad \alpha_1'^2 (\Delta \mathbf{t}')^2 = \alpha_1^2 (\Delta \mathbf{t})^2 + \mathcal{O}^3, \quad (4.17)$$

and by Taylor series expansion,

$$\Delta \mathbf{r}' = \mathbf{v}_1 \Delta \mathbf{t}' + \frac{1}{2} (\Delta \mathbf{t}')^2 (D' \mathbf{v}')_1 + \mathcal{O}^3. \quad (4.18)$$

The position of \mathbf{S}' with respect to \mathbf{S} can be defined as $\mathbf{r}_0 = \mathbf{v}_0 \mathbf{t}$ where

$$\mathbf{v}_0 = (v_0, 0, 0)^T. \quad (4.19)$$

Consequently, the position, \mathbf{r}_r , of \mathbf{P} with respect to \mathbf{S}' but observed in \mathbf{S} , and the corresponding velocity, \mathbf{v}_r , are given by

$$\mathbf{r}_r = \mathbf{r} - \mathbf{r}_0, \quad \mathbf{v}_r = D\mathbf{r}_r = \mathbf{v} - \mathbf{v}_0. \quad (4.20)$$

Correspondingly, the displacement $\Delta \mathbf{r}'$ with respect to \mathbf{S}' is observed in \mathbf{S} as a *residual displacement*

$$\Delta \mathbf{r}_r = (\Delta x_r, \Delta y_r, \Delta z_r)^T = \Delta \mathbf{r} - \Delta \mathbf{r}_0, \quad (4.21)$$

where

$$\Delta \mathbf{r}_0 = \mathbf{v}_0 \Delta \mathbf{t}. \quad (4.22)$$

In Newtonian kinematics we have $\Delta \mathbf{r}_r = \Delta \mathbf{r}'$, but in relativistic kinematics we obtain from (4.1), (4.21), and (4.22)

$$\begin{aligned} \Delta x_r &= \Delta x - v_0 \Delta t = \alpha_0 \Delta x', \\ \Delta y_r &= \Delta y = \Delta y', \\ \Delta z_r &= \Delta z = \Delta z'. \end{aligned} \quad (4.23)$$

For the forces acting during $[\mathbf{t}_1, \mathbf{t}_2]$ we can write

$$\mathbf{f} = \mathbf{f}_1 + \mathcal{O}^1, \quad \mathbf{f}' = \mathbf{f}'_1 + \mathcal{O}^1,$$

thus for the corresponding work done in \mathbf{S}' (cf. (4.16)),

$$\Delta W' = \int_{\mathbf{t}'_1}^{\mathbf{t}'_2} \mathbf{f}'^T \mathbf{v}' d\mathbf{t}' = \int_{\mathbf{r}'_1}^{\mathbf{r}'_2} \mathbf{f}'^T d\mathbf{r}' = (\mathbf{f}'_1^T + \mathcal{O}^1) \Delta \mathbf{r}' \quad (4.24)$$

and that done in \mathbf{S} ,

$$\Delta W = \int_{\mathbf{t}_1}^{\mathbf{t}_2} \mathbf{f}^T \mathbf{v} d\mathbf{t} = \Delta W_0 + \Delta W_r \quad (4.25)$$

where (cf. (4.20))

$$\Delta W_0 = \int_{t_1}^{t_2} \mathbf{f}^T \mathbf{v}_0 dt, \quad (4.26)$$

$$\Delta W_r = \int_{t_1}^{t_2} \mathbf{f}^T \mathbf{v}_r dt = (\mathbf{f}_1^T + \mathbf{O}^1) \Delta \mathbf{r}_r. \quad (4.27)$$

Clearly, ΔW_0 can be considered to be the work done due to the displacement $\Delta \mathbf{r}_0$, and ΔW_r the *residual work* done. In Newtonian mechanics we have $\mathbf{v}_r = \mathbf{v}'$, $\mathbf{f} = \mathbf{f}'$, $\mathbf{t} = \mathbf{t}'$ and thus always $\Delta W_r = \Delta W'$, but the same is not true in the relativistic case.

If desired we may assume (as will be done later) that \mathbf{P} is *instantaneously motionless* in \mathbf{S}' at $\mathbf{t}' = \mathbf{t}'_1$, i.e., that

$$\mathbf{v}'_1 = 0, \quad \text{thus} \quad \mathbf{v}_1 = \mathbf{v}_0 = (\mathbf{v}_0, 0, 0)^T \quad (4.28)$$

(cf. (4.10)). This situation can indeed always be achieved by proper choice of \mathbf{S} and \mathbf{v}_0 . We then have (cf. (3.2)),

$$\alpha_1 = \alpha_0, \quad \alpha'_1 = 1, \quad \mathbf{v}_{x1} = \mathbf{v}_0, \quad \mathbf{v}_{y1} = \mathbf{v}_{z1} = 0, \quad (4.29)$$

and, due to (4.18), (4.23), (4.24) and (4.27),

$$\Delta \mathbf{r}' = \mathbf{O}^2, \quad \Delta \mathbf{r}_r = \mathbf{O}^2, \quad (4.30)$$

$$\Delta \mathbf{W}' = \mathbf{f}_1'^T \Delta \mathbf{r}' + \mathbf{O}^3, \quad \Delta \mathbf{W}_r = \mathbf{f}_1^T \Delta \mathbf{r}_r + \mathbf{O}^3. \quad (4.31)$$

The result expressed by (4.30) and (4.31) is in line with the notion of work done as a product of force and displacement.

If (4.28) holds, the situation in \mathbf{S}' becomes Newtonian at $\mathbf{t}' = \mathbf{t}'_1$, whence we may assume, as is also true in classical relativity, that

$$\mathbf{f}'_1 = m_0 (\mathbf{D}' \mathbf{v}')_1 = m_0 \lim_{\Delta t' \rightarrow 0} (\Delta \mathbf{v}' / \Delta t'), \quad (4.32)$$

m_0 being some positive constant (identical to the rest mass).

More generally, the assumption (4.32) may be considered to be a consequence of a broader requirement which we call the *principle of Newtonian limit*. We may express this principle in a simple way by saying that for appropriately defined limits relativistic dynamics (i.e., dynamics compatible with the Lorentz transformation) should give the same results as Newtonian dynamics. Expressed differently and more precisely, at time instants where \mathbf{P} is instantaneously motionless in the relevant reference frame, relativistic dynamics should, in the limit for $\Delta t' = 0$, give the same results as Newtonian dynamics.

Let us apply this principle to the ratio $\Delta W' / \Delta W_r$. According to what we have seen above (subsequently to (4.27)), we have in Newtonian dynamics

$$\lim_{\Delta t' \rightarrow 0} (\Delta W' / \Delta W_r) = 1, \quad (4.33)$$

and this even for any \mathbf{v}'_1 . Hence, it must hold in relativistic dynamics at least if $\mathbf{v}'_1 = \mathbf{0}$. Since $\Delta W'$ and ΔW_r are \mathbf{O}^2 (cf. (4.30) and (4.31)) we may thus equivalently say that for $\mathbf{v}'_1 = \mathbf{0}$ we have

$$\lim_{\Delta t' \rightarrow 0} \left(\frac{\Delta W'}{(\Delta t')^2} \right) = \lim_{\Delta t' \rightarrow 0} \left(\frac{\Delta W_r}{(\Delta t')^2} \right). \quad (4.34)$$

Since that last result must hold without restriction it must be true for any \mathbf{f}'_1 . Due to (4.18), (4.23), (4.31), and (4.32) an alternative way to express it is to require

$$\mathbf{f}'_1{}^T \mathbf{f}'_1 = (\alpha_0 \mathbf{f}_{x1}, \mathbf{f}_{y1}, \mathbf{f}_{z1}) \mathbf{f}'_1 \quad \forall \mathbf{f}'_1 \quad (4.35)$$

or else,

$$(\alpha_0 \mathbf{f}_{x1} - \mathbf{f}'_{x1}) \mathbf{f}'_{x1} + (\mathbf{f}_{y1} - \mathbf{f}'_{y1}) \mathbf{f}'_{y1} + (\mathbf{f}_{z1} - \mathbf{f}'_{z1}) \mathbf{f}'_{z1} = 0 \quad \forall \mathbf{f}'_1. \quad (4.36)$$

For this it is necessary and sufficient that

$$\alpha_0 \mathbf{f}_{x1} = \mathbf{f}'_{x1}, \quad \mathbf{f}_{y1} = \mathbf{f}'_{y1}, \quad \mathbf{f}_{z1} = \mathbf{f}'_{z1}. \quad (4.37)$$

Due to (4.28) we have $\mathbf{v}_{y1} = \mathbf{v}_{z1} = 0$ and thus, in view of (4.11) to (4.13) and (4.29),

$$\begin{aligned} (\mathbf{D}' \mathbf{v}'_x)_1 &= \left(\mathbf{D} \frac{\mathbf{v}_x}{\alpha} \right)_1, \\ (\mathbf{D}' \mathbf{v}'_y)_1 &= \frac{1}{\alpha_1} \left(\mathbf{D} \frac{\mathbf{v}_x}{\alpha} \right)_1, \quad (\mathbf{D}' \mathbf{v}'_z)_1 = \frac{1}{\alpha_1} \left(\mathbf{D} \frac{\mathbf{v}_z}{\alpha} \right)_1. \end{aligned}$$

Hence taking into account $\alpha_0 = \alpha_1$ (cf. (4.29)), (4.32) and (4.37) yield

$$\mathbf{f}_1 = \frac{m_0}{\alpha_1} \left(\mathbf{D} \frac{\mathbf{v}}{\alpha} \right)_1. \quad (4.38)$$

Clearly, (4.38) implies the same type of relations for all three components \mathbf{f}_{x1} , \mathbf{f}_{y1} , and \mathbf{f}_{z1} , and this despite the differences in (4.37). Thus the orientation originally adopted for \mathbf{v}_0 is totally irrelevant. But t_1 is arbitrary. Hence we can finally indeed replace (4.38) by

$$\mathbf{f} = \frac{m_0}{\alpha} \mathbf{D} \left(\frac{1}{\alpha} \mathbf{v} \right), \quad (4.39)$$

and therefore,

$$\mathbf{f} = \sqrt{m} \mathbf{D} (\sqrt{m} \mathbf{v}) = \frac{1}{2} (m \mathbf{D} \mathbf{v} + \mathbf{D} (m \mathbf{v})), \quad m = \frac{m_0}{\alpha^2}, \quad (4.40)$$

confirming (3.6) and (3.9).

In view of Einstein's second postulate, a relation of exactly the same type can of course be written for \mathbf{f}' , \mathbf{v}' , and α' , i.e.,

$$\mathbf{f}' = \frac{m_0}{\alpha'} \mathbf{D}' \left(\frac{1}{\alpha'} \mathbf{v}' \right).$$

In particular, if $\mathbf{v}_y = \mathbf{v}_z = 0$, i.e., if all phenomena take place exclusively along the x and x' axes, we have $\alpha_x = \alpha$, $\alpha'_x = \alpha'$, and we then obtain, using (4.14),

$$\alpha \mathbf{f} = \alpha' \mathbf{f}'. \quad (4.41)$$

More generally, we have

$$\mathbf{f}_x = \frac{1}{\alpha_0} \left(\frac{\beta_0}{c} \mathbf{v}'^T \mathbf{f}' + \mathbf{f}'_x \right), \quad \mathbf{f}_y = \mathbf{f}'_y, \quad \mathbf{f}_z = \mathbf{f}'_z, \quad (4.42)$$

$$\mathbf{v}^T \mathbf{f} = \frac{1}{\alpha_0} (\mathbf{v}'^T \mathbf{f}' + c \beta_0 \mathbf{f}'_x) \quad (4.43)$$

where all quantities are as defined before and used so far. This result can either be verified directly by means of (4.1) etc. or by recalling that a proper four-vector such as the Minkowski force (cf. (3.11)) transforms from \mathbf{S}' to \mathbf{S} by premultiplying the one in \mathbf{S}' with the matrix

$$\mathbf{M} = \begin{pmatrix} \frac{\beta_0}{\alpha_0} & \frac{1}{\alpha_0} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{1}{\alpha_0} & \frac{\beta_0}{\alpha_0} & 0 & 0 \end{pmatrix}, \quad (4.44)$$

(whose inverse \mathbf{M}^{-1} is obtained by simply changing the sign of β_0). The result (4.41) can of course be shown to be a special case of the present general one.

Observe that the entire derivation, from (4.32) to (4.38), involves \mathbf{v}' only via $(D'\mathbf{v}')_1$. This is of interest because it stresses the role of that quantity as an asymptotic measure. On the other hand, in order to obtain the classical relativistic expression for the force, one would have to replace the right-hand side of (4.33), as can be concluded by inspecting (4.34), by $1/\alpha_0$, which would amount to a somewhat surprising requirement.

So far we have taken it for granted that m_0 is constant. Nevertheless there may be some interest to allow also for rest masses m_0 that, although independent of \mathbf{v} , are dependent on t in some other way than via \mathbf{v} . Let us assume that in that case, (4.32) is of the form

$$\mathbf{f}'_1 = \sqrt{m_0} (D' \sqrt{m_0} \mathbf{v}')_1.$$

It is easily verified that under this assumption the above process for arriving at (4.40) would still be valid, but (4.39) would then in general no longer hold.

5. Modification of Newton's third law and conservation of momentum

In classical relativity, Newton's third law remains untouched. This is true to the point that it is common practice not even to mention that keeping it unchanged amounts to making a corresponding definite additional assumption. If we adopt the alternative viewpoint, however, such an assumption is nowhere justified, i.e., we have to modify not only Newton's second but also his third law.

Consider indeed action and reaction between two particles, P_1 and P_2 , the subscripts 1 and 2 being consistently used in this section in order to distinguish between quantities referring either to P_1 or P_2 , respectively. We thus assume P_1 and P_2 to be travelling with velocities \mathbf{v}_1 and \mathbf{v}_2 . Let there be an interaction between P_1 and P_2 and let \mathbf{f}_1 and \mathbf{f}_2 be the resulting forces acting upon P_1 and P_2 , respectively. We first assume \mathbf{f}_1 , \mathbf{f}_2 , \mathbf{v}_1 , and \mathbf{v}_2 to be all located on the same straight line, which we may assume to be the x axis of S , thus the x' axis of S' (defined as before); for S' we correspondingly have to consider \mathbf{f}'_1 , \mathbf{f}'_2 , \mathbf{v}'_1 , and \mathbf{v}'_2 (and similarly primed quantities as used hereafter). Defining

$$\beta_1 = \mathbf{v}_{x1}/c, \quad \beta'_1 = \mathbf{v}'_{x1}/c, \quad \beta_2 = \mathbf{v}_{x2}/c, \quad \beta'_2 = \mathbf{v}'_{x2}/c, \quad \alpha_i = \sqrt{1 - \beta_i^2}, \quad \alpha'_i = \sqrt{1 - \beta_i'^2}, \quad i = 1, 2, \quad (5.1)$$

we can write (cf.(4.10)),

$$\beta'_1 = \frac{\beta_1 - \beta_0}{1 - \beta_1 \beta_0}, \quad \beta'_2 = \frac{\beta_2 - \beta_0}{1 - \beta_2 \beta_0}, \quad \beta_0 = \mathbf{v}_0/c. \quad (5.2)$$

Assume P_1 and P_2 to be touching each other at $t = t_0$, which implies $\mathbf{x}_1(t_0) = \mathbf{x}_2(t_0)$ and thus (cf. (4.1)) $\mathbf{x}'_1(t'_0) = \mathbf{x}'_2(t'_0)$, $t'_0 = t'_1(t_0) = t'_2(t_0)$. To simplify the writing we add a subscript zero to specify evaluations at $t = t_0$, thus at $t' = t'_0$ (except that \mathbf{v}_0 and thus β_0 have the same meaning as so far). If $\beta_{10} = -\beta_{20}$ we adopt $\mathbf{S}' = \mathbf{S}$. If $\beta_{10} \neq -\beta_{20}$, choosing β_0 such that

$$\beta_0^2 - 2b\beta_0 + 1 = 0, \quad b = (1 + \beta_{10}\beta_{20})(\beta_{10} + \beta_{20}),$$

which, due to

$$b^2 = 1 + \left(\frac{\alpha_{10}\alpha_{20}}{\beta_{10} + \beta_{20}} \right)^2,$$

always yields real values for β_0 , we obtain, as can be shown (cf. (5.2)),

$$\beta'_{10} = -\beta'_{20}, \quad \text{thus} \quad \alpha'_{10} = \alpha'_{20}. \quad (5.3)$$

Note that the product of the two choices for β_0 is equal to 1, i.e., we must select that choice for which $|\beta_0| < 1$.

Due to the assumption concerning the orientation of the velocities and forces and the reference frames we may make use of (4.41) and thus write in particular,

$$\alpha_{10}\mathbf{f}_{10} = \alpha'_{10}\mathbf{f}'_{10}, \quad \alpha_{20}\mathbf{f}_{20} = \alpha'_{20}\mathbf{f}'_{20}. \quad (5.4)$$

But due to the symmetry of the velocities in \mathbf{S}' we must have $\mathbf{f}'_{10} = -\mathbf{f}'_{20}$. Hence, (5.3) and (5.4) yield

$$\alpha_{10}\mathbf{f}_{10} = -\alpha_{20}\mathbf{f}_{20}. \quad (5.5)$$

Furthermore if, while maintaining the assumption concerning the alignment of \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{f}_1 , \mathbf{f}_2 , we abandon the assumption concerning the coincidence of that common direction with the one of the \mathbf{x} and \mathbf{x}' axes, an orientation argument as used in Section 4.2 (following (4.38)) shows that (5.5) remains valid.

Consider now the general case of the relationship between action and reaction. In view of (5.5) it is clear that it could not be $\mathbf{f}_1 = -\mathbf{f}_2$ as in the classical theory. For modifying Newton's third law, however, the simplest way compatible with (5.5) is to write

$$\alpha_1\mathbf{f}_1 = -\alpha_2\mathbf{f}_2 \quad (5.6)$$

where α_1 and α_2 are defined according to the subscript convention mentioned at the beginning of this section (cf. (5.1)).

Next, let m_{10} and m_{20} be the rest masses of P_1 and P_2 . It follows from (4.39) and (5.6) that

$$D(\mathbf{p}_1 + \mathbf{p}_2) = 0$$

where

$$\mathbf{p}_1 = m_{10}\mathbf{v}_1/\alpha_1 \quad \text{and} \quad \mathbf{p}_2 = m_{20}\mathbf{v}_2/\alpha_2$$

are the momenta as mentioned in Section 3. More generally, if there are n particles P_1 to P_n and if forces $\mathbf{f}_{\nu 1}$ to $\mathbf{f}_{\nu k}$ are acting upon P_ν , which is moving with velocity \mathbf{v}_ν , we have

$$D\mathbf{p}_\nu = \alpha_\nu (\mathbf{f}_{\nu 1} + \dots + \mathbf{f}_{\nu k}) \quad (5.7)$$

where

$$\mathbf{p}_\nu = m_{\nu 0}\mathbf{v}_\nu / \alpha_\nu, \quad \alpha_\nu = \sqrt{1 - \beta_\nu^2}, \quad \beta_\nu = v_\nu / c, \quad v_\nu^2 = \mathbf{v}_\nu^T \mathbf{v}_\nu.$$

Summing (5.7) over all particles, there will be pairwise cancellations in the right-hand side due to relations of the type of (5.6), yielding

$$D(\mathbf{p}_1 + \cdots + \mathbf{p}_n) = 0. \quad (5.8)$$

Hence, conservation of momentum holds exactly as in classical relativity. Vice-versa, if in the alternative theory we require (5.8) to apply in all cases, we are forced to modify Newton's third law as shown in (5.6).

6. Moving particles in fields – forces and energy

Assume that a field (electromagnetic, gravitational etc.) is exerting a force \mathbf{f} upon a particle \mathbf{P} travelling with velocity \mathbf{v} , that the equipment, $\mathbf{E} \mathbf{q}$ (referred to hereafter by a subscript 1), producing the field is at rest ($\mathbf{v}_1 = 0$), and that the reaction upon $\mathbf{E} \mathbf{q}$ is \mathbf{f}_1 . Assume furthermore that all relevant distances are sufficiently small so that we can state (cf. (5.6)),

$$\alpha \mathbf{f} = -\alpha_1 \mathbf{f}_1 = -\mathbf{f}_1, \quad (6.1)$$

with α and α_1 as in (3.2) and as used in Section 5 (except that $\mathbf{E} \mathbf{q}$ assumes the former role of \mathbf{P}_1 , whence $\alpha_1 = 1$). If \mathbf{f}_1 is independent of the velocity of \mathbf{P} , it must be equal to the force existing for $\mathbf{v} = 0$. Hence, we then conclude from (6.1) that

$$\mathbf{f} = \mathbf{f}_0 / \alpha \quad (6.2)$$

where \mathbf{f}_0 is that force that the field exerts upon \mathbf{P} if $\mathbf{v} = \mathbf{0}$, and thus is equal to $-\mathbf{f}_1$, i. e., \mathbf{f}_0 is the force as used in all classical expressions. We have e. g. $\mathbf{f}_0 = q\mathbf{E}$ if q is a charge and \mathbf{E} the electric field, and correspondingly in the case of a gravitational field.

For an electromagnetic field this is confirmed and in fact extended to include the Lorentz force if one proceeds as is classically done (see e.g. Section 7.7 of [19]), i.e., if, at the instant considered, one derives \mathbf{f} by Lorentz transforming the force $q\mathbf{E}'$ observed in an appropriate other reference system \mathbf{S}' . For seeing this, let us first recall that the force \mathbf{f} considered here is identical to the corresponding part of the Minkowski four-vector (cf. Section 3) and that the electric and magnetic fields are known to transform between \mathbf{S}' and \mathbf{S} (both used as so far) according to

$$\mathbf{E}'_x = \mathbf{E}_x, \quad \mathbf{E}'_y = \frac{1}{\alpha_0} (\mathbf{E}_y - v_0 \mu \mathbf{H}_z), \quad \mathbf{E}'_z = \frac{1}{\alpha_0} (\mathbf{E}_z + v_0 \mu \mathbf{H}_y), \quad (6.3)$$

$$\mathbf{H}'_x = \mathbf{H}_x, \quad \mathbf{H}'_y = \frac{1}{\alpha_0} (\mathbf{H}_y + v_0 \varepsilon \mathbf{E}_z), \quad \mathbf{H}'_z = \frac{1}{\alpha_0} (\mathbf{H}_z - v_0 \varepsilon \mathbf{E}_y), \quad (6.4)$$

primed field quantities referring again to \mathbf{S}' and unprimed ones to \mathbf{S} .

Consider then a charge q that is instantaneously motionless in \mathbf{S}' , i.e., we assume the orientation of \mathbf{S} and \mathbf{S}' and the choice of \mathbf{v}_0 (cf. (4.19)) to be such that at the time instant considered we have $\mathbf{v}' = \mathbf{0}$, thus $\mathbf{v} = \mathbf{v}_0$ (cf. (4.10)) and therefore also $\alpha_0 = \alpha$. The force \mathbf{f}' acting on q is then given by

$$\mathbf{f}' = q\mathbf{E}', \quad \mathbf{E}' = (\mathbf{E}'_x, \mathbf{E}'_y, \mathbf{E}'_z)^T.$$

From (4.42) and (6.3) we thus obtain, making use of the magnetic induction defined by $\mathbf{B} = (\mathbf{B}_x, \mathbf{B}_y, \mathbf{B}_z)^T = \mu \mathbf{H}$,

$$f_x = q \frac{1}{\alpha} E_x, \quad f_y = q \frac{1}{\alpha} (E_y - v_0 B_z), \quad f_z = q \frac{1}{\alpha} (E_z - v_0 B_y).$$

This result can indeed be written in the form (6.2) with

$$\mathbf{f}_0 = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (6.5)$$

In the way the result is presented in (6.2) and (6.5) it is expressed by only using quantities defined in \mathbf{S} and general vector notation. Hence, it is valid without restriction.

On the other hand, a force \mathbf{f} acting on a particle of rest mass m_0 also satisfies (4.39). We thus derive from (6.2),

$$m_0 D\left(\frac{1}{\alpha} \mathbf{v}\right) = \mathbf{f}_0. \quad (6.6)$$

We conclude from (6.6) that the alternative theory leads to exactly the same dynamic behavior of a particle in a field as classical relativity theory.

This however does not include energy. Consider indeed a relation from classical electrodynamics such as

$$-\rho \mathbf{v}^T \mathbf{E} = \frac{1}{2} D(\varepsilon \mathbf{E}^T \mathbf{E} + \mu \mathbf{H}^T \mathbf{H}) + \text{div}(\mathbf{E} \times \mathbf{H})$$

(written in standard notation), where $\rho = \varepsilon \text{div} \mathbf{E}$ is the charge density and ε and μ are assumed constant (say $\varepsilon = \varepsilon_0$, $\mu = \mu_0$). In order to obtain a proper expression for the energy supplied, the alternative theory obviously requires, as has been shown in this section, to multiply the classical force density $\rho \mathbf{E}$ in the left-hand side by $1/\alpha$, and the right-hand side thus has to be multiplied in the same way. We consider here only the case of a constant α (see also Section 7). Clearly, the field energy density then turns out to be

$$w = \frac{1}{2\alpha} (\varepsilon \mathbf{E}^T \mathbf{E} + \mu \mathbf{H}^T \mathbf{H}) \quad (6.7)$$

and the Poynting vector similarly becomes $(\mathbf{E} \times \mathbf{H})/\alpha$. The classical expression $(\varepsilon \mathbf{E}^T \mathbf{E} + \mu \mathbf{H}^T \mathbf{H})/2$ is then something like an apparent energy density.

Note that in this and the subsequent sections we use the letters W and w for designating energies and energy densities, and this in order to facilitate distinguishing between energies and electric fields. Also note that if instead of a charge q we consider a charge density ρ travelling with velocity \mathbf{v} we can replace (6.5) by,

$$\hat{\mathbf{f}} = \hat{\mathbf{f}}_0/\alpha, \quad \hat{\mathbf{f}}_0 = \rho(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

where $\hat{\mathbf{f}}$ and $\hat{\mathbf{f}}_0$ are force densities, i.e., forces per unit volume.

7. Electromagnetic field moving with constant velocity

We define \mathbf{S} and \mathbf{S}' as so far. Let us first consider a field that in \mathbf{S}' is electrostatic, thus at rest, and given there by \mathbf{E}' . Clearly, \mathbf{E}' is independent of \mathbf{t}' . Although for the corresponding magnetic field we have $\mathbf{H}' = 0$, the field in \mathbf{S} comprises both \mathbf{E} and \mathbf{H} . Since we assume ε and μ constant we need not consider \mathbf{D} and \mathbf{B} . In \mathbf{S} , the field is moving with constant velocity \mathbf{v} . We have $\mathbf{v} = \mathbf{v}_0$ (cf. (4.10)) and therefore,

$$\mathbf{v} = \mathbf{v}_0, \quad \alpha = \alpha_0, \quad \beta = \beta_0, \quad (7.1)$$

the quantities thus involved being defined as so far (except that we assume \mathbf{v} to take the sign of \mathbf{v}_0). From (6.3), (6.4), and (7.1) we deduce, for any \mathbf{E}' and \mathbf{H}' ,

$$\begin{aligned} E_x &= E'_x, & E_y &= \frac{1}{\alpha}(E'_y + v\mu H'_z), & E_z &= \frac{1}{\alpha}(E'_z - v\mu H'_y), \\ H_x &= H'_x, & H_y &= \frac{1}{\alpha}(H'_y - v\varepsilon E'_z), & H_z &= \frac{1}{\alpha}(H'_z + v\varepsilon E'_y). \end{aligned} \quad (7.2)$$

For the field energy density defined by (6.7) we thus find for $\mathbf{H}' = \mathbf{0}$,

$$\mathbf{w} = \frac{1}{\alpha} \mathbf{w}' + \frac{\beta^2}{\alpha^3} \mathbf{w}'_0, \quad (7.3)$$

where

$$\mathbf{w}' = \frac{1}{2} \varepsilon (\mathbf{E}'^2_x + \mathbf{E}'^2_y + \mathbf{E}'^2_z), \quad \mathbf{w}'_0 = \varepsilon (\mathbf{E}'^2_y + \mathbf{E}'^2_z), \quad (7.4)$$

\mathbf{w}' being the field energy density in \mathbf{S}' . Due to the electrostatic hypothesis, the right-hand side of (7.3) is independent of \mathbf{t}' .

Let then

$$dV = dx \cdot dy \cdot dz, \quad dV' = dx' \cdot dy' \cdot dz'$$

be the elementary volumes in \mathbf{S} and \mathbf{S}' , respectively. As is known, it follows from (4.1) that for any fixed \mathbf{t} ,

$$dx = \alpha dx', \quad dy = dy', \quad dz = dz',$$

so that $dV = \alpha dV'$. We thus obtain from (7.3),

$$\mathbf{w} dV = \mathbf{w}' dV' + \frac{\beta^2}{\alpha^2} \mathbf{w}'_0 dV'. \quad (7.5)$$

This expression is precisely of the form determined by (3.8) and (3.10), with \mathbf{E} , \mathbf{E}_i , and \mathbf{E}_0 replaced by $\mathbf{w} dV$, $\mathbf{w}' dV'$, and $\mathbf{w}'_0 dV'$, respectively, and such that no strict relationship exists between \mathbf{w}' and \mathbf{w}'_0 . We may also integrate (7.5) over the entire volume, which leads to $\mathbf{W} = \mathbf{W}' + \mathbf{W}_k$ where

$$\mathbf{W} = \int_V \mathbf{w} dV, \quad \mathbf{W}' = \int_{V'} \mathbf{w}' dV', \quad \mathbf{W}_k = \frac{\beta^2}{\alpha^2} \int_{V'} \mathbf{w}'_0 dV'$$

Since \mathbf{w}' and \mathbf{w}'_0 are independent of \mathbf{t}' , \mathbf{W}' and \mathbf{W}_k may be considered to be evaluated at a constant \mathbf{t}' . Furthermore, V and V' may be the entire space.

In all cases we thus obviously obtain full agreement with the results developed in the preceding sections. The decomposition into rest energy and kinetic energy encountered in classical relativity theory does not offer a similarly elegant interpretation. This is remarkable since Maxwell's equations are known to be inherently compatible with the Lorentz transformation.

The second expression (7.4) can also be written as

$$\mathbf{w}'_0 = \varepsilon \mathbf{E}'^T \mathbf{E}' - \varepsilon \mathbf{E}'^2_x.$$

But due to $\mathbf{v} = \mathbf{v}_0$ and $\mathbf{v} = \mathbf{v}_0$ we have (cf. (7.1)) $\mathbf{v} \mathbf{E}'_x = \mathbf{v}^T \mathbf{E}'$. Hence, (7.5) can be expressed in the more general form

$$\mathbf{w} dV = \left(\mathbf{w}' + \frac{1}{2\alpha^2} \mathbf{v}^T \mathbf{m}'_0 \mathbf{v} \right) dV', \quad (7.6)$$

where

$$\mathbf{m}'_0 = \frac{2\varepsilon}{c^2} \left((\mathbf{E}'^T \mathbf{E}') \mathbf{1} - \mathbf{E}' \mathbf{E}'^T \right), \quad (7.7)$$

$\mathbf{1}$ being again the unit matrix of order 3. These expressions hold for any orientation of \mathbf{v} and have the form of the matrix case discussed in relation with (3.14).

Consider now a general electromagnetic field. Substitution of (7.2) in (6.7) yields, after some calculation, the more general expression

$$\mathbf{w}dV = \left(\mathbf{w}' + \frac{1}{2\alpha^2} \mathbf{v}^T \mathbf{m}'_0 \mathbf{v} + \frac{2}{c^2 \alpha^2} \mathbf{v}^T \mathbf{S}' \right) dV' \quad (7.8)$$

where

$$\mathbf{w}' = \frac{1}{2} (\varepsilon \mathbf{E}'^T \mathbf{E}' + \mu \mathbf{H}'^T \mathbf{H}'), \quad \mathbf{S}' = \mathbf{E}' \times \mathbf{H}', \quad (7.9)$$

$$\mathbf{m}'_0 = \frac{2}{c^2} \left((\varepsilon \mathbf{E}'^T \mathbf{E}' + \mu \mathbf{H}'^T \mathbf{H}') \mathbf{1} - (\varepsilon \mathbf{E}' \mathbf{E}'^T + \mu \mathbf{H}' \mathbf{H}'^T) \right). \quad (7.10)$$

Compared to (7.6), (7.7) and the first equation (7.4), the changes in \mathbf{w}' and \mathbf{m}'_0 are as immediately expected. There is however an additional term that is proportional to the component of the Poynting vector in \mathbf{S}' in the direction of \mathbf{v} , thus of the corresponding momentum density in \mathbf{S}' of the electromagnetic field. This additional term vanishes in particular if at the point \mathbf{P}' (thus at the position and the time instant) under consideration in \mathbf{S}' the field has zero momentum, i.e., in a sense, is there instantaneously motionless. If we then consider (6.7) but apply it to \mathbf{S}' instead of \mathbf{S} , it reduces to the first equality in (7.4). Hence, invoking again a principle akin to that of Newtonian limit (but now applied in an electromagnetic sense) we can state that \mathbf{w}' is, under the present assumptions, indeed equal to the field energy density at \mathbf{P}' .

8. The Bertozzi experiment

In 1964 Bertozzi [26] has published a few results of experiments aimed at determining the kinetic energy of fast electrons. Expressed in terms of a notation more suitable for our purpose, he had determined values of $\beta_b^2 = 1 - \alpha_b^2$, thus of v_b^2/c^2 , in terms of what we designate here by

$$\gamma = qu_0/m_0 c^2,$$

q being the electron charge, u_0 the voltage traversed by the electrons in the accelerator, m_0 the electron rest mass, v_b the velocity reached by the electron just before hitting the intended target, and c the speed of light. According to classical relativity theory α_b is simply related to the other quantities by

$$\frac{m_0 c^2}{\alpha_b} - m_0 c^2 = qu_0, \quad (8.1)$$

yielding

$$\alpha_b^2 = 1 - \beta_b^2 = 1/(1 + \gamma)^2, \quad \text{with} \quad \beta_b = v_b/c. \quad (8.2)$$

Since according to the alternative theory the dynamic behavior of particles in fields is exactly the same as that predicted by classical relativity, the result expressed by (8.2) is valid in exactly the same way in both theories (contrary to what had been stated in [21]).

Let E_k be the kinetic energy of the electron travelling at v_b . According to classical relativity we have $E_k = qu_0$, whence in [26] γ had indeed been set equal to $E_k/m_0 c^2$. According to the alternative theory, however, $E_k = m_0 v_b^2/2\alpha_b^2$, which in view of (8.2) yields

$$E_k = qu_0(1 + \gamma/2) = m_0 c^2 \gamma(1 + \gamma/2), \quad (8.3)$$

and this discrepancy was at the origin of the confusion in [21].

For testing the validity of equating E_k with qu_0 , Bertozzi had also, in two cases, determined the heat, E_h , generated by an electron hitting the target. Define $\delta = E_h/qu_0$. Since there are some unavoidable losses (e.g. by X-ray generation) we have $E_h < E_k$, thus $\delta < E_k/qu_0$. Hence, $\delta < 1$ if we had indeed $E_k = qu_0$. In both tests (1.5 MeV and 4.5 MeV), the measured result was $\delta = 1.067$, which is compatible with $\delta < 1$ only if one makes allowance for the low accuracy (10 % as stated in [26]) of the experiment.

For examining the situation according to the alternative theory, let f be the braking force acting upon the electron in the target and let f_f be the corresponding friction force acting on the target, which we may assume to be at rest. Due to the alternative form of Newton's third law we have

$$\alpha f = f_f, \quad \alpha = \sqrt{1 - \beta^2}, \quad \beta = v/c,$$

i.e., in view of (4.39), $f_f = m_0 D(v/\alpha)$, v being the electron velocity at the considered time instant. Hence (cf. (3.5)), the heat power generated by friction is $vf_f = D(m_0 c^2/\alpha)$, and the heat generated by the braking process is therefore given by

$$\left[\frac{m_0 c^2}{\alpha} \right]_{v=0}^{v=v_b} = \frac{m_0 c^2}{\alpha_b} - m_0 c^2 = qu_0,$$

the second equality being indeed a consequence of (8.1). The excess energy $E_k - qu_0$ (cf. (8.3)) will be converted partly into radiation and partly into (practically unmeasurable) kinetic energy of the target and its supporting body, but, due to secondary processes, a small part will again turn up in form of heat, thus causing $E_h > qu_0$, i.e., $\delta > 1$. This agrees with the tendency exhibited by the outcome of the experiment, i.e., without having to invoke an insufficient experimental accuracy.

9. The matrix case

In Sections 3 and 7 we have briefly mentioned the possibility of having objects whose mass cannot simply be described by a scalar, as is usually assumed, but only by a matrix. Let us examine a few details about the properties of such a (particle-like) object P , which we thus assume to be characterized, at vanishing velocity, by a constant (i.e. time-independent) symmetric positive definite matrix \mathbf{m}_0 called *matrix rest mass*.

We first have to re-examine the derivation presented in Section 4.2, i.e., for m_0 replaced by \mathbf{m}_0 and thus (4.32) by

$$\mathbf{f}'_i = \mathbf{m}_0 (D' \mathbf{v}')_i. \quad (9.1)$$

However, if \mathbf{m}_0 is a fully occupied matrix we cannot take it for granted anymore that the value ΔW_r defined by (4.25) and (4.26) is a proper estimate of $\Delta W'$. Indeed if \mathbf{m}_0 is as just stated it is not possible to sufficiently separate the phenomena in the x , y , and z directions, while due to (4.19) the corrective term (4.26) singles out the x direction. We therefore first assume that the directions of x' , y' , and z' coincide with the principal axes of \mathbf{m}_0 , i.e., that \mathbf{m}_0 is diagonal, thus $\mathbf{m}_0 = \text{diag}(m_1, m_2, m_3)$, and we may then again claim (4.34) to hold.

However, in view of (9.1), $(D'\mathbf{v}')_1$ appearing in (4.18) is no longer equal to a scalar (formerly: $1/\mathbf{m}_0$, cf. (4.32)) times \mathbf{f}'_1 , so that the two right vector factors \mathbf{f}'_1 in (4.35) have to be replaced by the original factor $(D'\mathbf{v}')_1$. We have to proceed correspondingly for the three right scalar factors in (4.36) where, due to (9.1), $\forall \mathbf{f}'_1$ implies $\forall (D'\mathbf{v}')_1$ so that (4.37) is still valid. From this result, we can follow again the same steps as in Section 4.2, using for \mathbf{m}_0 the property of being diagonal. One finds this way that (4.38) has to be replaced by

$$\mathbf{f}_1 = \frac{1}{\alpha_1} \mathbf{m}_0 \left(D \frac{\mathbf{v}}{\alpha} \right)_1$$

and thus (4.39) and (4.40) by (3.13), with \mathbf{m} given by (3.12). Observe that for vanishing velocity, \mathbf{P} is again fully characterized by \mathbf{m}_0 ; this is in agreement with Einstein's first postulate but holds in the strict form just mentioned only because of the specific choices of x' , y' , and z' with respect to x , y , and z (contrary to what is the case for the earlier considered scalar \mathbf{m}_0).

We now drop the assumption for \mathbf{m}_0 to be diagonal. Let \mathbf{P} be travelling in \mathbf{S} with velocity \mathbf{v} and subjected there to a force \mathbf{f} , \mathbf{m}_0 being the matrix mass for vanishing velocity, i.e. the matrix rest mass. Let \mathbf{S}_0 be a further reference system that is in a fixed, yet rotated position with respect to \mathbf{S} . There thus exists a constant orthogonal matrix \mathbf{U} such that $\mathbf{U}\mathbf{f}$ and $\mathbf{U}\mathbf{v}$ are the force and the velocity in \mathbf{S}_0 . Since for vanishing velocity we have $\mathbf{f} = \mathbf{m}_0 D\mathbf{v}$, the matrix rest mass in \mathbf{S}_0 is $\mathbf{U}\mathbf{m}_0\mathbf{U}^T$.

On the other hand, since \mathbf{m}_0 is constant we may indeed choose the constant matrix \mathbf{U} such that $\mathbf{U}\mathbf{m}_0\mathbf{U}^T$ is diagonal. Hence, using the result obtained above we may write for \mathbf{S}_0 ,

$$\mathbf{U}\mathbf{f} = \frac{1}{\alpha} (\mathbf{U}\mathbf{m}_0\mathbf{U}^T) D \left(\frac{1}{\alpha} \mathbf{U}\mathbf{v} \right). \quad (9.2)$$

This implies again (3.13) to hold, with \mathbf{m} as given by (3.12).

This neat result confirms the appropriateness of extending the scalar to the matrix case in the simple way we have chosen in Section 3. There are, however, also discrepancies, and this concerns in particular the classical concept of transforming four-vectors between \mathbf{S} and \mathbf{S}' by premultiplying them with the matrix \mathbf{M} (cf. (4.44)) or its inverse. The classical proof for justifying this approach in the case of the Minkowski force is indeed not applicable because \mathbf{M} and \mathbf{m}_0 will in general not commute. Nevertheless, quite simple proofs can be obtained by making use of elegant expressions such as

$$1 - \beta_0\beta_x = \alpha_0\alpha/\alpha', \quad 1 + \beta_0\beta'_x = \alpha_0\alpha'/\alpha, \quad 1 - \beta_x\beta'_x = \alpha_x\alpha'_x/\alpha_0, \quad \alpha'/\alpha = \alpha'_x/\alpha_x \quad (9.3)$$

which can indeed be shown to follow from (4.10) (cf. also (4.8)) and where all notation used is as defined earlier.

Let us thus derive the equivalent of (4.42). We find, using (3.13), (4.8) (4.10), and (9.3),

$$\mathbf{f} = \mathbf{m}_0 \frac{1}{\alpha} D \left(\frac{\mathbf{v}}{\alpha} \right) = \mathbf{m}_0 \frac{1}{\alpha'} D' \left(\frac{1}{\alpha'} ((\mathbf{v}'_x + \mathbf{v}_0)\alpha_0, \mathbf{v}'_y, \mathbf{v}'_z)^T \right). \quad (9.4)$$

This yields, applying the equivalent of (3.5) for \mathbf{S}' ,

$$\mathbf{f} = \mathbf{m}_0 \frac{1}{\alpha'} D' \left(\frac{1}{\alpha'} \begin{pmatrix} \mathbf{v}'_x / \alpha_0 \\ \mathbf{v}'_y \\ \mathbf{v}'_z \end{pmatrix} \right) + \mathbf{m}_0 \frac{\beta_0}{2c\alpha_0} \begin{pmatrix} D'(\mathbf{v}'\alpha')^2 \\ 0 \\ 0 \end{pmatrix}.$$

That expression cannot usually be reduced to the form of (4.42), which is the one that would be obtained by applying the standard procedure based on (4.44). For \mathbf{m}_0 diagonal, say for $\mathbf{m}_0 = \text{diag}(m_1, m_2, m_3)$ one finds

$$f_x = \frac{1}{\alpha_0} \left(f'_x + m_1 \frac{\beta_0}{c\alpha'} \mathbf{v}'^T D' \left(\frac{\mathbf{v}'}{\alpha'} \right) \right), \quad f_y = f'_y, \quad f_z = f'_z,$$

but even that expression reduces to (4.42) only if $m_1 = m_2 = m_3$, which is equivalent to the scalar case.

A comment similar to that at the end of Section 4.2 can also be made in the present context. In particular, an object as we are considering here may be subject to influences other than forces (cf. Section 7), which can bring about an implicit time dependence.

10. Conclusions

Proper characterization of losslessness in Kirchhoff circuits requires the defining equations of elements such as inductances to have a very specific form, which in turn can be recognized to be of fundamental physical importance. Surprisingly, this form is not respected by the classical relativistic expression for the force acting on a particle. However, requiring the relativistic laws to become coincident with the Newtonian laws when approaching a time instant where the particle is instantaneously motionless, and applying this principle of Newtonian limit in particular to the force and to the work done, one finds an expression for the force, thus for the relativistic formulation of Newton's second law, that is in perfect agreement with that to be expected from the theory of nonlinear Kirchhoff circuits. This difference comes about essentially by putting prime emphasis on energy aspects (losslessness, work done, etc.) rather than on momentum, as in classical relativity. The alternative way of modifying Newton's second law is shown to require also his third law, which in the classical theory remains untouched, and as a consequence the expressions for forces in fields to be changed correspondingly. These changes combined imply that conservation of momentum and dynamics of particles remain exactly as in classical relativity, the latter aspect being confirmed and extended by appropriately interpreting the classical application of the Lorentz transformation to Maxwell's equations.

The rest energy of particles appears as an arbitrary integration constant, and its value thus remains fully compatible with the classical relativistic expression. Definite differences between the classical and the alternative theory become visible for the kinetic energy, which for the latter theory is rational in the velocity while it is known to be irrational for the former. For an electric field that is static with respect to a reference frame moving itself uniformly with respect to the observer, the kinetic term in the field energy agrees fully and directly with the alternative expression for kinetic energy, thus contrary to the classical expression. The few measured results published by Bertozzi about the kinetic energy of fast electrons favor the alternative theory, but the inaccuracy of 10% mentioned by the author is too large to allow definite conclusions to be drawn.

Some further-reaching ideas could be developed that concern the nature of particles and are compatible with the present results. A very brief outline of this can be found at the end of [21], although some aspects of that paper are revised or extended by the present one.

Acknowledgement

The author is greatly indebted to M. Bouten, M. Depenbrock, P. Fettweis, A. Hahn, J. Kummer, K.-A. Owenier, H. Rollnick, and R. Staufenbiel for many discussions on the subject of this paper and to F.-J. Reich for substantial help in preparing the manuscript.

References

1. W. Cauer, "Theorie der linearen Wechselstromschaltungen", 2nd ed., Akademie-Verlag, Berlin, 1954. American edition: "Synthesis of Linear Communication Networks", McGraw Hill, New York, 1958
2. V. Belevitch, "Classical Network Theory", Holden-Day, San Francisco, 1968.
3. J. Meixner, "Beziehungen zwischen Netzwerktheorie und Thermodynamik", Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, Heft 181, Westdeutscher Verlag, Köln, 1967.
4. Fettweis, "In Memoriam - Vitold Belevitch", IEEE Transactions on Circuits and Systems - I, vol. 47, pp. 613 - 614, May 2000.
5. Fettweis, "Wave digital filters: theory and practice", Proc. IEEE, vol. 74, pp. 270-327, Feb. 1986, and vol. 75, p. 729, May 1987.
6. H. D. Fischer, "Wave digital filters for numerical integration", ntz-Archiv, vol. 6, pp. 37-40, 1984.
7. Fettweis, "Discrete passive modelling of physical systems described by partial differential equations", in "Multivariate Analysis: Future Directions"(C.R. Rao, editor), pp. 115-130, Elsevier, Amsterdam, 1993.
8. Fettweis, "Improved wave-digital approach to numerically integrating the PDEs of fluid dynamics", Proc. IEEE Int. Symp. Circuits and Systems, vol. III, pp. 361-364, Scottsdale, AZ, USA, 26-29 May 2002.
9. Fettweis, "Robust numerical integration using wave-digital concepts", Proceedings of the 5th DSPS Educators Conference, pp. 23 - 32, Tokyo Institute of Technology, Tokyo, Japan, 17 - 18 Sept. 2003.
10. P. G. Bergmann, "Introduction to the Theory of Relativity", Prentice Hall, 1942, and Dover, New York, 1976.
11. Sommerfeld, "Vorlesungen über Theoretische Physik", vol. 3, Elektrodynamik, Akademische Verlagsgesellschaft, Leipzig, 1949.
12. M. Born, "Die Relativitätstheorie Einsteins", 5th ed., Springer-Verlag, Berlin, 1969. American translation of an earlier German edition: "Einstein's Theory of Relativity", Dover, New York, 1965.
13. S. Weinberg, "Gravitation and Cosmology", Wiley, New York, 1972.
14. R. Becker and F. Sauter, "Theorie der Elektrizität", vol. 1, Teubner, Stuttgart, 1973.
15. J. D. Jackson, "Classical Electrodynamics", 2nd ed., J. Wiley, New York, 1975.
16. W. Panofsky and M. Phillips, "Classical Electricity and Magnetism", Addison-Wesley, Reading, MA, 1977.
17. J. van Bladel, "Relativity and Engineering", Springer-Verlag, Berlin, 1984.
18. Kacser, "Relativity, special theory", in Encyclopedia of Physics, 2nd ed., R. G. Lerner and G. L. Triggs, Eds., pp. 1052 - 1058, Wiley-VCH, New York, 1991.
19. H. and M. Ruder, "Die spezielle Relativitätstheorie", Vieweg, Braunschweig, Germany, 1993.

20. Fettweis, "The wave-digital method and some of its relativistic implications", IEEE Trans. Circuits and Systems I, vol. 49, No. 6, pp. 862- 868, June 2002, and No. 10, p. 1521, October 2002.
21. Fettweis, "Wave-digital concepts and relativity theory", in "Contemporary Issues in Systems Stability and Control with Applications"(Derong Liu and Panos Antsaklis, eds.), pp. 3 - 22, Birkhäuser, Boston, 2003.
22. Fettweis, "Nonlinear Kirchhoff circuits and relativity theory", AEÜ Int. J. Electronics (in print).
23. K. Meerkötter, Private communication, about 1991.
24. K. Meerkötter and T. Felderhoff, "Simulation of nonlinear transmission lines by wave digital filter principles", Proc. IEEE Int. Symp. Circuits and Systems, pp. 875 - 877, San Diego, CA, USA, May 1992.
25. N. K. Bose and A. Fettweis, "Isomorphism between skew-symmetric and orthogonal matrices via generalized Cayley Transform", Proc. Eur. Conf. Circuit Theory and Design, vol. I, pp. 133 - 136, Cracow, Poland, 1 - 4 Sept. 2003.
26. W. Bertozzi, "Speed and energy of relativistic electrons ", American J. of Physics, vol. 32, pp. 551 - 555, 1964.
27. H. Minkowski, "Raum und Zeit", published in "Das Relativitätsprinzip", Teubner, Leipzig, 1913; English version: "Space and Time", in "The Principle of Relativity", Methuen, 1923, and Dover, New York, 1952.

Alfred Fettweis Dr., Prof., (Belgium). He studied electrical engineering at Université Catholique de Louvain (1946 - 51) and received from there the degree "Docteur en sciences appliquées" in 1963. He worked for International Telephone and Telegraph Corporation in Belgium (1951 - 54, 1956 - 63) and the USA (1954 - 56), was a professor of theoretical electricity at Technical University of Eindhoven, Netherlands (1963 - 67), has been professor of communications engineering at Ruhr-Universität Bochum, Germany (since 1967, emeritus professor since 1992) and was a distinguished visiting professor at Notre Dame University, USA (1994 - 96). He received honorary doctorates from Linköping 1986, Mons 1988, Leuven 1988, Budapest 1995, and several awards, among which: Prix Acta Technica Belgica 1963, Prix Montefiore 1980, VDE-Ehrenring 1984, ITG Karl-Küpfmüller-Preis 1988, Eduard Rhein Basic Research Award 1993, IEEE Centennial (1984) and Millenium (1999) Medal, and Darlington Award 1980, Technical Achievement Award 1988, Van Valkenburg (Society) Award 2001, (first) Vitold Belevitch Award 2003 (all 4 from IEEE Circuits and Systems Society).

fettweis@nt.ruhr-uni-bochum.de.

Отсутствие потерь в нелинейных цепях Кирхгофа и в теории относительности

А. Феттвейс

Department of Electrical Engineering and Information Science/Communications
Engineering, Ruhr-Universität Bochum,
D-44780 Bochum, Germany

Аннотация. Цепи Кирхгофа имеют большое значение не только для исследования электрических явлений, но и для моделирования широкого ряда физических систем в тех случаях, когда важны сохранение мощности и энергии и связанные с ними понятия, например пассивность и отсутствие потерь. Они представляют собой взаимосвязи разнообразных элементов, которые не должны быть линейными и постоянными где-либо. Если некий элемент (например, индуктивность) является нелинейным и/или в явном виде зависящим от времени и его необходимо охарактеризовать как пассивный или, точнее, не имеющий потерь, то определяющее его выражение должно записываться в особой форме. Однако классическое понятие релятивистской массы к таким не относится. Показано, что если сохраняется классическая релятивистская кинематика, если релятивистская динамика приближается к ньютоновской при соответствующих предельных переходах, и если основное внимание уделяется совершенной работе, а не импульсу (количеству движения), а энергии, то естественным образом приходим к выражению для силы, записанному через массу и скорость, форма которого полностью согласуется с формой записи для нелинейной индуктивности. Этот альтернативный способ изменения второго закона Ньютона требует и изменения третьего закона Ньютона. Эти две модификации дают тот же закон сохранения импульса и ту же динамику частиц в поле, что и классическая теория относительности. Однако выражение для кинетической энергии отличается. Представлены логические выводы, теоретический и экспериментальный результат, которые в некоторой степени подтверждают альтернативную теорию или, по крайней мере, не противоречат ей, какой бы невероятной эта теория ни казалась. В работе дополняются и обновляются ранее полученные по данной тематике результаты, улучшена форма представления результатов.

Ключевые слова: законы Ньютона, теория относительности, Илья Пригожин, цепи Кирхгофа

1. Введение

Работа посвящается памяти великого физика и химика Ильи Пригожина. Его выдающийся вклад в науку коснулся автора по крайней мере по двум аспектам: с одной стороны, он особенно известен своей работой по неравновесной термодинамике – области, в которой может пригодиться теория цепей Кирхгофа [1, 2, 3], которая сама по себе никак не связана с неравновесной термодинамикой. С другой стороны, он был одним из выдающихся россиян, эмигрировавших в Бельгию в детстве после революции, который внес выдающийся вклад в научную жизнь Бельгии. Другим известным эмигрантом был Витольд Белевич [4] – известный в Европе и, по мнению многих, в мире специалист в области теории цепей Кирхгофа второй половины двадцатого века. Автор обязан этому выдающемуся ученому своей возможностью познакомиться с красотой, универсальностью и строгостью этой теории. В данной работе показано, как теория цепей Кирхгофа может пролить свет даже на некоторые аспекты теории относительности.

Цепи Кирхгофа имеют большое значение не только при моделировании электрических явлений, но и во многих других задачах. Например, они могут служить эталонными цепями для волновых цифровых подходов, в частности, к фильтрации [5] и к

численному интегрированию обыкновенных дифференциальных уравнений [6] и, особенно, дифференциальных уравнений в частных производных, описывающих физические системы [7-9]. При фильтрации особый интерес представляют линейные цепи, а при численном интегрировании – нелинейные явления.

В цепях Кирхгофа выводы элементов соединяются в узлы. Связи описываются законами Кирхгофа для токов и напряжений. Отсюда вытекает важное следствие, состоящее в том, что законы сохранения мощности и энергии строго гарантируются при соблюдении законов Кирхгофа. Значит, такие свойства как пассивность (passivity) и отсутствие потерь (частный случай пассивности), в свою очередь, строго гарантируются, если соответствующие свойства выполняются для всех элементов. Однако чтобы гарантировать, что элемент, такой как нелинейная индуктивность, не имеет потерь, его необходимо представить в виде, который немного отличается от обычного [7 - 9], и такой альтернативный вид элемента, оказывается, имеет фундаментальное физическое значение. Масса является в некоторой степени элементом без потерь (консервативным элементом). В механике Ньютона этот элемент линейный, а в релятивистской механике – нелинейный. Однако удивительно то, что взаимосвязь между силой и скоростью в классической теории относительности [10 - 19] не имеет той общей формы записи, которую можно было бы ожидать, исходя из сказанного выше о нелинейной индуктивности.

В предыдущих работах [20, 21] было показано, что при определенных допущениях (которые не являются априори необоснованными) мы естественным образом получаем релятивистское выражение, связывающее силу, массу и скорость, которое отличается от классического, но имеет вид, который полностью согласуется с только что упомянутой общей формой записи. Это несоответствие в основном связано с тем, что в классической теории относительности импульс (количество движения) – это величина, которая принимается в начале рассуждений для получения законов динамики, а в альтернативном подходе основное внимание уделяется энергии и проделанной работе. Изменений в отношении преобразования Лоренца, а следовательно, и справедливости релятивистской кинематики, которая, как известно, является результатом сложных и строгих логических выводов, нет. Таким образом, соблюдается второй постулат Эйнштейна (касающийся постоянства скорости света). Также соблюдается и первый постулат, касающийся универсальности физических законов. Однако в определенной степени мы дополняем два постулата Эйнштейна *принципом ньютоновского предела*.

Модификация второго закона Ньютона (раздел 4) требует соответствующего изменения третьего закона Ньютона (раздел 5), хотя в классической теории относительности последний остается неизменным. Сочетание этих двух модификаций указывает на то, что сохранение импульса соблюдается, как и в классической теории относительности (раздел 5), и что динамика частиц в полях (например, в ускорителях частиц) тоже не меняется (раздел 6). Выражение (которое можно получить стандартными методами) для энергии в электростатическом поле, наблюдаемом в движущейся системе координат (раздел 7), согласуется с альтернативной теорией. Распространение этого выражения на электромагнитные поля (при правильной его интерпретации) в общем случае также согласуется с альтернативной теорией. Эксперимент Бертоцци (раздел 8), хотя в нем и присутствует кинетическая энергия, также, по меньшей мере, не опровергает его. Выражение для этой величины, на самом деле, различно для этих двух случаев, поэтому должны быть способы различать их.

Данная статья преследует несколько целей. Во-первых, вывод альтернативной взаимосвязи между силой, скоростью и массой выполняется несколько иначе, чем ранее, поэтому предлагается более надежное обоснование этого выражения. Во-вторых,

эксперимент Бертоцци анализируется точнее, чем в [21]; показано, что эксперимент, скорее, подтверждает альтернативный подход, хотя точность эксперимента недостаточно высока, чтобы получить точные выводы. В-третьих, исправлено несколько неточностей в [21]. В-четвертых, данная работа не просто улучшена, но и существенно расширена по сравнению с [22], поэтому она должна быть гораздо удобнее для восприятия читателями. Безусловно, как в [21], но в отличие от [20], мы не накладываем каких-либо ограничений на константу интегрирования, возникающих в связи с выражением для энергии. Все важные для понимания статьи предыдущие результаты, опубликованные ранее, включены в текст для облегчения восприятия. Ясно, что вопрос в целом остается открытым. Автор надеется, что представленные ниже результаты послужат стимулом к дальнейшим экспериментальным и теоретическим исследованиям, которые смогут пролить свет на данную проблему.

2. Пассивность и отсутствие потерь в нелинейных цепях Кирхгофа

Нелинейную индуктивность, которая зависит только от своего тока i , можно описать одним из следующих уравнений

$$u = D(L_g i) \quad \text{или} \quad u = L_l D i \quad (2.1)$$

где u – напряжение на индуктивности и

$$D = d/dt, \quad (2.2)$$

$L_g = L_g(i) = \phi/i$ – глобальная индуктивность, $L_l = L_l(i) = d\phi/di$ – локальная индуктивность, ϕ – магнитный поток. Можно показать, что условие $L_g(i) \geq 0 \quad \forall i$ необходимо но недостаточно, чтобы обеспечить пассивность (фактически, отсутствие потерь), а $L_l(i) \geq 0 \quad \forall i$ достаточно, но не необходимо. Чтобы получить удовлетворительную характеристику, Meerkötter [23, 24] предложил представить нелинейную индуктивность с помощью нелинейного идеального трансформатора, присоединенного к линейной индуктивности; однако более простой и равноценный способ записи (2.1) в форме [7 - 9]

$$u = \sqrt{L} D(\sqrt{L} i) = \frac{1}{2} (D(L i) + L D i). \quad (2.3)$$

Два выражения в (2.3) в определенной степени соответствуют геометрическому и арифметическому среднему определений (2.1). Для поглощенной мощности и запасенной энергии W_L из (2.3) найдем,

$$u i = D W_L, \quad W_L = \frac{1}{2} L i^2, \quad (2.4)$$

а для индуктивности L неравенство $L(i) \geq 0$ теперь является необходимым и достаточным для пассивности (фактически, – для отсутствия потерь).

Записывая (2.3) в виде $x = D y$, $x = u/\sqrt{L}$, $y = i\sqrt{L}$, аппроксимируя D по формуле трапеций с шагом T и обозначая дискретизированную переменную времени через t_n , запишем результат в виде

$$b(t_n) = -a(t_n - T) \quad (2.5)$$

где так называемые волны (волновые величины) a и b имеют вид

$$a = \frac{u + i R}{2\sqrt{R}}, \quad b = \frac{u - i R}{2\sqrt{R}}, \quad R = \frac{2L}{T}. \quad (2.6)$$

Очевидно, что мы имеем также

$$ui = a^2 - b^2. \quad (2.7)$$

Всё приведенное выше имеет более общий характер, если L – это функция не только i (тока), но любой из зависимых переменных схемы и/или независимой переменной, обозначим ее t . Это верно даже для многомерных цепей Кирхгофа [7-9], когда D – оператор в частных производных, а t необходимо заменить вектором независимых переменных.

Простота таких выражений, как (2.5) и (2.7) и их комбинации с пассивностью и отсутствием потерь – важнейшие причины, по которым применение волновых цифровых подходов к фильтрации [5] и численному интегрированию обыкновенных дифференциальных уравнений и дифференциальных уравнений в частных производных дает существенные преимущества [6 - 9]. Представление индуктивности в виде (2.3) можно распространить на связанные индуктивности, а следовательно, на индуктивности в матричном виде L , причем эта матрица является симметричной и неотрицательно определенной. Если u и i – векторная запись u и i , то можно записать

$$u = \frac{1}{2}(LDi + D(Li)) \quad \text{или} \quad u = L^{T/2}D(L^{1/2}i) \quad (2.8)$$

где матрица $L^{1/2}$ такая, что $L^{T/2}L^{1/2} = L$, верхний индекс T обозначает транспонирование, и $L^{T/2} = (L^{1/2})^T$. Два выражения в (2.8) эквивалентны тогда и только тогда, когда $L^{T/2}D(L^{1/2}) = (DL^{T/2})L^{1/2}$. Общие условия, при которых последнее равенство выполняется, неизвестны [25], но достаточно того, что $L^{1/2}$ – диагональная матрица, умноженная на постоянную матрицу (см. (3.13)).

Значимость понятий энергии, волны и рассеяния в физике предполагает, что такие выражения как (2.3) и (2.4) имеют фундаментальное физическое значение. Однако это наблюдение не признается классической теорией относительности в рамках специальной теории.

3. Релятивистская масса

Пусть для частицы с массой покоя m_0

$$f = (f_x, f_y, f_z)^T, \quad p = (p_x, p_y, p_z)^T, \quad v = (v_x, v_y, v_z)^T \quad (3.1)$$

представляют собой действующую на нее силу, ее импульс и ее скорость, соответственно. Будем определять v , β и α так:

$$v^T v = v^2, \quad \beta = v/c, \quad \alpha = \sqrt{1 - \beta^2} \quad (3.2)$$

где c – скорость света. В соответствии с базовыми принципами, принятыми в классической теории относительности, выполняется следующее (см. (2.2)),

$$f = Dp, \quad p = m_g v, \quad m_g = m_0/\alpha \quad (3.3)$$

где m_g записано вместо m для соответствия обозначениям в (2.1). Из (3.3) для энергии, сообщаемой частице, можно найти известное соотношение

$$v^T f = D(m_g c^2), \quad (3.4)$$

а с учетом равенства –

$$c^2 D \frac{1}{\alpha} = \mathbf{v}^T D \left(\frac{1}{\alpha} \mathbf{v} \right). \quad (3.5)$$

Величину $m_g c^2$ тогда можно назвать общей энергией, а $(m_g - m_0) c^2$ – кинетической энергией. Ясно, что m_g из (3.3) на самом деле соответствует L_g из (2.1), а не L из (2.3). Это несколько удивительно и, как объясняется в [20 - 22], именно поэтому интересно проверить, может ли представлять интерес выражение для *силы* вида

$$\mathbf{f} = \sqrt{m} D(\sqrt{m} \mathbf{v}) = \frac{1}{2} (D(m \mathbf{v}) + m D \mathbf{v}) \quad (3.6)$$

где *масса* m – некая функция от \mathbf{v} , скажем, $m = m(\beta)$. В подтверждение, (3.4) будет заменено на

$$\mathbf{v}^T \mathbf{f} = D E_k = D E \quad (3.7)$$

где

$$E = E_i + E_k, \quad E_k = m v^2 / 2, \quad E_i = \text{const.} \quad (3.8)$$

Т.к. $E_k = 0$ при $v = 0$, то E_k будет *кинетической энергией*, а E_i – *энергией покоя* (внутренней энергией), E – *общей энергией*. Тогда можно предположить, что главный интерес представляет величина m , а не m_g . В частности, определяя массу покоя $m_0 = m(0)$, желательно сохранить важные характеристики классической теории относительности, хотя и не для m_g , а для m , и требовать, чтобы $m - m_0$ увеличивалось с ростом E_k , желательно линейно, т.е. в соответствии с соотношением

$$E_k = K (m - m_0), \quad K = \text{const.},$$

и чтобы $m = \infty$ при $v = c$. Из приведенного выше выражения для E_k (см. (3.8)) следует, что $K = c^2 / 2$ и что m должна иметь вид

$$m = \frac{m_0}{1 - \beta^2} = \frac{m_0}{\alpha^2}, \quad (3.9)$$

и поэтому $E_k = \frac{1}{2} (m c^2 - m_0 c^2)$.

С учетом $E_k = E - E_i$ в работе [20] предлагалось определять E через $m c^2 / 2$, и, следовательно, E_i – через $m_0 c^2 / 2$. Однако это связано с некоторыми трудностями, хотя ничего не противоречит приведенным выше уравнениям, если не накладывать каких-либо требований на E_i , т.е. на константу интегрирования из (3.7). E и E_k тогда определяются по (3.8), следовательно, по E_k , при помощи

$$E_k = \frac{1}{2} m v^2 = E_0 \frac{\beta^2}{1 - \beta^2} = E_0 \frac{\beta^2}{\alpha^2}, \quad E_0 = \frac{1}{2} m_0 c^2, \quad (3.10)$$

где m определяется по (3.9). Рассмотренный в [20] случай соответствует $E_i = E_0$. Принятая точка зрения равносильна тому, чтобы считать, что частица, в общем случае, характеризуется двумя константами: E_i и E_0 . При $E_i = 2E_0$ энергия покоя $E_i = m_0 c^2$, как в классическом случае. Однако в общем случае простого соотношения, связывающего E_i и E_0 , нет.

Непосредственно из (3.10) следует интересное наблюдение. С использованием $v^2 = v_x^2 + v_y^2 + v_z^2$ выражение для E_k можно, на самом деле, разложить на сумму компонент кинетической энергии из ньютоновской динамики:

$$E_k = E_{kx} + E_{ky} + E_{kz}$$

$$E_{kx} = \frac{1}{2} m v_x^2, \quad E_{ky} = \frac{1}{2} m v_y^2, \quad E_{kz} = \frac{1}{2} m v_z^2.$$

В строго физическом смысле тройное $E_k = (E_{kx}, E_{ky}, E_{kz})^T$ не образует вектор (преобразование здесь более сложное, чем в случае координат, в отличие от импульса), но это не должно волновать. Тем не менее, это наблюдение может частично объяснять, почему в классической теории вероятности при формулировке основ теории динамики отдается предпочтение импульсу. Фактически, упор на компонентах энергии (т.е. не только на упомянутых выше трех кинетических компонентах) также имеет важнейшее фундаментальное значение для волнового цифрового (wave-digital) метода численного интегрирования дифференциальных уравнений в частных производных [7 - 9].

При помощи \mathbf{f} можно определить квадруполь

$$\begin{pmatrix} \mathbf{f} \\ \frac{1}{c} \mathbf{v}^T \mathbf{f} \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \frac{1}{c} D E_k \end{pmatrix}, \quad (3.11)$$

для которого в (3.11) приводится второе, эквивалентное, выражение, записанное через E_k (см. (3.10)), и которое можно непосредственно сравнивать с известными результатами из теории относительности [10, 13 - 16, 19]. Можно легко проверить, что этот квадруполь аналогичен 4-вектору, который был впервые предложен Минковским [27], а значит, он инвариантен относительно преобразований Лоренца. Отсюда непосредственно следует, что преобразование Лоренца (см. Раздел 4) меняет только ту компоненту \mathbf{f} , которая ориентирована в направлении относительного движения двух систем координат друг относительно друга, т.е. компоненты \mathbf{f} , которые перпендикулярны этому направлению, останутся неизменными. В классической теории относительности ситуация обратная.

Как и в случае с индуктивностями, можно пойти еще дальше и рассмотреть объект, подобный частице, который характеризуется, помимо E_i , постоянной неотрицательной определенной симметричной матрицей \mathbf{m}_0 вместо скалярной массы m_0 . Тогда скалярную массу m необходимо заменить на *матричную массу* \mathbf{m} , которая имеет вид (см. (3.9))

$$\mathbf{m} = \frac{1}{\alpha^2} \mathbf{m}_0. \quad (3.12)$$

Выражение (3.6) для \mathbf{f} становится

$$\mathbf{f} = \mathbf{m}_0 \frac{1}{\alpha} D \left(\frac{1}{\alpha} \mathbf{v} \right) = \frac{1}{\alpha} D \left(\mathbf{m}_0 \frac{1}{\alpha} \mathbf{v} \right) = \frac{1}{2} (D(\mathbf{m} \mathbf{v}) + \mathbf{m} D \mathbf{v}) = \frac{1}{\alpha} \mathbf{m}_0^{T/2} D \left(\frac{1}{\alpha} \mathbf{m}_0^{1/2} \mathbf{v} \right), \quad (3.13)$$

принятая запись соответствует обозначениям, принятым в отношении (2.8).

В матричном случае (3.7) остается неизменным, но в (3.8) и (3.10) выражение для E_k необходимо заменить на

$$E_k = \frac{1}{2} \mathbf{v}^T \mathbf{m} \mathbf{v} = \frac{1}{2\alpha^2} \mathbf{v}^T \mathbf{m}_0 \mathbf{v}. \quad (3.14)$$

Случай с матрицей сводится к скалярному, если $\mathbf{m}_0 = \mathbf{1}m_0$, где $\mathbf{1}$ – единичная матрица. Пример с матрицей, где приведение к такому виду невозможно, упоминается в Разделе 7.

4. Прямой вывод альтернативного способа изменения второго закона Ньютона

4.1. Общие соотношения

С учетом неожиданных результатов Раздела 3, представляется уместным исследовать, существует ли способ получить эти (и, возможно, связанные с ними) результаты напрямую. Ясно, что т.к. энергия играет важную роль в Разделе 3, такой прямой подход должен концентрироваться на энергии и совершенной работе, а не на импульсе (количестве движения), как в классической теории относительности.

Рассмотрим две системы координат: \mathbf{S} и \mathbf{S}' , причем \mathbf{S}' движется с постоянной скоростью v_0 , допустим, в направлении x системы \mathbf{S} . Для \mathbf{S} мы используем обозначения без «штриха», как в Разделе 3 (координаты, скорости, силы, дифференциальные операторы и т.д.), а для \mathbf{S}' – соответствующие обозначения со «штрихом». Тогда \mathbf{S} и \mathbf{S}' связаны преобразованием Лоренца [10 -19]

$$x' = \frac{1}{\alpha_0}(x - v_0 t), \quad y' = y, \quad z' = z, \quad t' = \frac{1}{\alpha_0}\left(t - \beta_0 \frac{x}{c}\right) \quad (4.1)$$

где

$$\beta_0 = v_0/c, \quad \alpha_0 = \sqrt{1 - \beta_0^2}. \quad (4.2)$$

Рассмотрим частицу \mathbf{P} , двигающуюся, как описано в Разделе 3. Для частицы \mathbf{P} можно ввести векторы координат ее положения

$$\begin{aligned} \mathbf{r} &= \mathbf{r}(t) = (x, y, z)^T, \\ \mathbf{r}' &= \mathbf{r}'(t') = (x', y', z')^T, \end{aligned} \quad (4.3)$$

при этом t' можно однозначно выразить через t , пусть $t' = t'(t)$, и наоборот. В то время как (3.1) и (3.2) относятся к \mathbf{S} , имеем в отношении \mathbf{S}' :

$$\mathbf{f}' = (f'_x, f'_y, f'_z)^T, \quad \mathbf{v}' = (v'_x, v'_y, v'_z)^T, \quad (4.4)$$

$$\mathbf{v}'^T \mathbf{v}' = v'^2, \quad \beta' = v'/c, \quad \alpha' = \sqrt{1 - \beta'^2}, \quad (4.5)$$

где (см. (2.2))

$$\mathbf{v} = D\mathbf{r}, \quad \mathbf{v}' = D'\mathbf{r}' \quad (4.6)$$

$$D = d/dt, \quad D' = d/dt'. \quad (4.7)$$

Как можно проверить с помощью (4.1), имеем, как известно [10 - 19]:

$$\frac{dt}{dt'} = \frac{\alpha'}{\alpha} = \frac{\alpha'_x}{\alpha_x}, \quad \alpha'D = \alpha D' \quad (4.8)$$

где в дополнение к ранее определенным величинам:

$$\alpha_x = \sqrt{1 - \beta_x^2}, \quad \alpha'_x = \sqrt{1 - \beta'^2_x}, \quad \beta_x = \frac{v_x}{c}, \quad \beta'_x = \frac{v'_x}{c}, \quad (4.9)$$

и далее:

$$v'_x = \frac{v_x - v_0}{1 - \beta_x \beta_0}, \quad v'_y = \frac{\alpha'}{\alpha} v_y, \quad v'_z = \frac{\alpha'}{\alpha} v_z. \quad (4.10)$$

Из этих выражений можно получить следующие равенства:

$$\frac{1}{\alpha'^3} D'v'_x = \frac{1}{\alpha^3} Dv_x = D \frac{v_x}{\alpha} + \frac{v_y^2 + v_z^2}{\alpha^3 c^2} Dv_x - \frac{v_x}{\alpha^3 c^2} (v_y Dv_y + v_z Dv_z), \quad (4.11)$$

$$D'v'_y = \frac{\alpha'^2}{\alpha} D \frac{v_y}{\alpha} + v'_y \frac{1}{\alpha} D\alpha', \quad (4.12)$$

$$D'v'_z = \frac{\alpha'^2}{\alpha} D \frac{v_z}{\alpha} + v'_z \frac{1}{\alpha} D\alpha', \quad (4.13)$$

$$D(v_x/\alpha_x) = D'(v'_x/\alpha'_x). \quad (4.14)$$

4.2. Второй закон Ньютона

Первый закон Ньютона чисто качественный и всегда справедлив, поэтому для нашего случая он не актуален. Второй закон Ньютона касается выражения для силы. Рассмотрим две системы координат \mathbf{S} и \mathbf{S}' , как в Разделе 4.1, и ту же частицу P . Пусть t_1 и t_2 – два момента времени, которые произвольно близки, но различимы, пусть $t_2 > t_1$. Величины, относящиеся к t_1 и t_2 , будут обозначаться индексами 1 и 2, соответственно, т.е.

$$\begin{aligned} t'_1 &= t'(t_1), & t'_2 &= t'(t_2), & t_1 &= t(t'_1), & t_2 &= t(t'_2), \\ \mathbf{r}_1 &= \mathbf{r}(t_1), & \mathbf{r}_2 &= \mathbf{r}(t_2), & \mathbf{r}'_1 &= \mathbf{r}'(t'_1), & \mathbf{r}'_2 &= \mathbf{r}'(t'_2), \\ \mathbf{v}_1 &= \mathbf{v}(t_1), & \mathbf{v}_2 &= \mathbf{v}(t_2), & \mathbf{v}'_1 &= \mathbf{v}'(t'_1), & \mathbf{v}'_2 &= \mathbf{v}'(t'_2), \\ x_1 &= x(t_1), & x_2 &= x(t_2), & v_{x1} &= v_x(t_1), & v_{x2} &= v_x(t_2) \end{aligned}$$

и т.д. Также определим

$$\Delta t = t_2 - t_1, \quad \Delta \mathbf{r} = (\Delta x, \Delta y, \Delta z)^T = \mathbf{r}_2 - \mathbf{r}_1, \quad (4.15)$$

$$\Delta t' = t'_2 - t'_1, \quad \Delta \mathbf{r}' = (\Delta x', \Delta y', \Delta z')^T = \mathbf{r}'_2 - \mathbf{r}'_1 \quad (4.16)$$

и т.д. и используем упрощенную запись

$$O^n = O((\Delta t)^n) = \text{порядок величины } (\Delta t)^n \text{ при } n \geq 0.$$

Т.к. $t' = t'(t)$ и $\alpha dt = \alpha' dt'$ (см. (4.8)), имеем

$$\Delta t' = t'(t_1 + \Delta t) - t'_1 = (dt'/dt)_1 \Delta t + O^2 = \frac{\alpha_1}{\alpha'_1} \Delta t + O^2$$

и поэтому

$$\alpha'_1 \Delta t' = \alpha_1 \Delta t + O^2, \quad \alpha_1'^2 (\Delta t')^2 = \alpha_1^2 (\Delta t)^2 + O^3, \quad (4.17)$$

и из разложения в ряд Тейлора

$$\Delta \mathbf{r}' = \mathbf{v}_1 \Delta t' + \frac{1}{2} (\Delta t')^2 (D'\mathbf{v}')_1 + O^3. \quad (4.18)$$

Положение \mathbf{S}' относительно \mathbf{S} можно определить как $\mathbf{r}_0 = \mathbf{v}_0 t$, где

$$\mathbf{v}_0 = (v_0, 0, 0)^T. \quad (4.19)$$

Следовательно, положение \mathbf{r}_r частицы P относительно \mathbf{S}' , но наблюдаемое в \mathbf{S} , и соответствующая скорость \mathbf{v}_r записываются так:

$$\mathbf{r}_r = \mathbf{r} - \mathbf{r}_0, \quad \mathbf{v}_r = D\mathbf{r}_r = \mathbf{v} - \mathbf{v}_0. \quad (4.20)$$

Соответственно, смещение $\Delta\mathbf{r}'$ относительно \mathbf{S}' наблюдается в \mathbf{S} как *остаточное смещение*

$$\Delta\mathbf{r}_r = (\Delta x_r, \Delta y_r, \Delta z_r)^T = \Delta\mathbf{r} - \Delta\mathbf{r}_0, \quad (4.21)$$

где

$$\Delta\mathbf{r}_0 = \mathbf{v}_0 \Delta t. \quad (4.22)$$

В кинематике Ньютона имеем $\Delta\mathbf{r}_r = \Delta\mathbf{r}'$, но в релятивистской кинематике из (4.1), (4.21), и (4.22) получаем

$$\begin{aligned} \Delta x_r &= \Delta x - v_0 \Delta t = \alpha_0 \Delta x', \\ \Delta y_r &= \Delta y = \Delta y', \\ \Delta z_r &= \Delta z = \Delta z'. \end{aligned} \quad (4.23)$$

Для сил, действующих на промежутке времени $[t_1, t_2]$, можно записать

$$\mathbf{f} = \mathbf{f}_1 + \mathbf{O}^1, \quad \mathbf{f}' = \mathbf{f}'_1 + \mathbf{O}^1,$$

тогда для соответствующей работы, совершённой в \mathbf{S}' (см. (4.16)),

$$\Delta W' = \int_{t'_1}^{t'_2} \mathbf{f}'^T \mathbf{v}' dt' = \int_{\eta_1}^{\eta_2} \mathbf{f}'_1^T d\mathbf{r}' = (\mathbf{f}'_1^T + \mathbf{O}^1) \Delta\mathbf{r}' \quad (4.24)$$

для работы, совершённой в \mathbf{S} ,

$$\Delta W = \int_{t_1}^{t_2} \mathbf{f}^T \mathbf{v} dt = \Delta W_0 + \Delta W_r \quad (4.25)$$

где (см. (4.20))

$$\Delta W_0 = \int_{t_1}^{t_2} \mathbf{f}^T \mathbf{v}_0 dt, \quad (4.26)$$

$$\Delta W_r = \int_{t_1}^{t_2} \mathbf{f}^T \mathbf{v}_r dt = (\mathbf{f}_1^T + \mathbf{O}^1) \Delta\mathbf{r}_r. \quad (4.27)$$

Ясно, что ΔW_0 можно считать работой, совершённой на перемещении $\Delta\mathbf{r}_0$, а ΔW_r — совершённой *остаточной работой*. В ньютоновской механике имеем $\mathbf{v}_r = \mathbf{v}'$, $\mathbf{f} = \mathbf{f}'$, $t = t'$ и поэтому всегда $\Delta W_r = \Delta W'$, но то же самое неверно в релятивистской механике. Можно допустить (как будет сделано позже), что \mathbf{P} *мгновенно неподвижна* в \mathbf{S}' при $t' = t'_1$, т.е.

$$\mathbf{v}'_1 = 0, \quad \text{тогда} \quad \mathbf{v}_1 = \mathbf{v}_0 = (v_0, 0, 0)^T \quad (4.28)$$

(см. (4.10)). Этого всегда можно добиться правильным выбором \mathbf{S} и \mathbf{v}_0 . Тогда имеем (см. (3.2)),

$$\alpha_1 = \alpha_0, \quad \alpha'_1 = 1, \quad v_{x1} = v_0, \quad v_{y1} = v_{z1} = 0, \quad (4.29)$$

и с учетом (4.18), (4.23), (4.24) и (4.27)

$$\Delta\mathbf{r}' = \mathbf{O}^2, \quad \Delta\mathbf{r}_r = \mathbf{O}^2, \quad (4.30)$$

$$\Delta \mathbf{W}' = \mathbf{f}_1'^T \Delta \mathbf{r}' + \mathcal{O}^3, \quad \Delta \mathbf{W}_r = \mathbf{f}_1^T \Delta \mathbf{r}_r + \mathcal{O}^3. \quad (4.31)$$

Результат, который описывается выражениями (4.30) и (4.31), согласуется с работой, совершенной силой, в виде произведения силы на смещение.

Если (4.28) выполняется, то при $t' = t'_1$ система S' становится ньютоновской, откуда можно предположить (что также верно в классической теории относительности):

$$\mathbf{f}_1' = m_0 (D' \mathbf{v}')_1 = m_0 \lim_{\Delta t' \rightarrow 0} (\Delta \mathbf{v}' / \Delta t'), \quad (4.32)$$

m_0 – некая положительная константа (идентичная массе покоя).

В более общем смысле допущение (4.32) можно считать следствием более общего требования, которое мы называем *принципом ньютоновского предела*. Этот принцип можно сформулировать очень просто: при правильно определенных пределах релятивистская динамика (т.е. динамика, совместимая с преобразованием Лоренца) должна давать те же результаты, что и ньютоновская динамика. Точнее, другими словами, в моменты времени, когда P мгновенно неподвижна в соответствующей системе координат, релятивистская динамика в пределе для $\Delta t' = 0$ должна давать те же результаты, что и ньютоновская динамика.

Применим этот принцип к отношению $\Delta W' / \Delta W_r$. Согласно тому, что мы видели раньше (уравнение (4.27)), в ньютоновской динамике имеем:

$$\lim_{\Delta t' \rightarrow 0} (\Delta W' / \Delta W_r) = 1, \quad (4.33)$$

даже для любого \mathbf{v}'_1 . Следовательно, в релятивистской динамике это должно выполняться, по крайней мере, если $\mathbf{v}'_1 = \mathbf{0}$. Т.к. $\Delta W'$ и ΔW_r являются \mathcal{O}^2 (см. (4.30) и (4.31)), при $\mathbf{v}'_1 = \mathbf{0}$ аналогичным образом имеем:

$$\lim_{\Delta t' \rightarrow 0} \left(\frac{\Delta W'}{(\Delta t')^2} \right) = \lim_{\Delta t' \rightarrow 0} \left(\frac{\Delta W_r}{(\Delta t')^2} \right). \quad (4.34)$$

Т.к. последний результат должен выполняться без ограничений, он должен выполняться для любого \mathbf{f}'_1 . С учетом (4.18), (4.23), (4.31) и (4.32) альтернативный способ записать его – это потребовать, чтобы

$$\mathbf{f}_1'^T \mathbf{f}'_1 = (\alpha_0 \mathbf{f}_{x1}, \mathbf{f}_{y1}, \mathbf{f}_{z1}) \mathbf{f}'_1 \quad \forall \mathbf{f}'_1 \quad (4.35)$$

или

$$(\alpha_0 \mathbf{f}_{x1} - \mathbf{f}'_{x1}) \mathbf{f}'_{x1} + (\mathbf{f}_{y1} - \mathbf{f}'_{y1}) \mathbf{f}'_{y1} + (\mathbf{f}_{z1} - \mathbf{f}'_{z1}) \mathbf{f}'_{z1} = 0 \quad \forall \mathbf{f}'_1. \quad (4.36)$$

Для этого необходимо и достаточно, чтобы

$$\alpha_0 \mathbf{f}_{x1} = \mathbf{f}'_{x1}, \quad \mathbf{f}_{y1} = \mathbf{f}'_{y1}, \quad \mathbf{f}_{z1} = \mathbf{f}'_{z1}. \quad (4.37)$$

С учетом (4.28) имеем $\mathbf{v}_{y1} = \mathbf{v}_{z1} = 0$, откуда с учетом (4.11) – (4.13) и (4.29):

$$(D' \mathbf{v}'_x)_1 = \left(D \frac{\mathbf{v}_x}{\alpha} \right)_1, \\ (D' \mathbf{v}'_y)_1 = \frac{1}{\alpha_1} \left(D \frac{\mathbf{v}_x}{\alpha} \right)_1, \quad (D' \mathbf{v}'_z)_1 = \frac{1}{\alpha_1} \left(D \frac{\mathbf{v}_z}{\alpha} \right)_1.$$

Следовательно, учитывая $\alpha_0 = \alpha_1$ (см. (4.29)), (4.32) и (4.37), получаем

$$\mathbf{f}_1 = \frac{m_0}{\alpha_1} \left(D \frac{\mathbf{v}}{\alpha} \right)_1. \quad (4.38)$$

Ясно, что из (4.38) следует тот же тип взаимосвязи между всеми тремя компонентами f_{x1} , f_{y1} и f_{z1} , несмотря на различия в (4.37). Таким образом, изначально принятая ориентация \mathbf{v}_0 совершенно неуместна. Только t_1 является произвольной величиной. Следовательно, можно заменить (4.38) на

$$\mathbf{f} = \frac{m_0}{\alpha} D \left(\frac{1}{\alpha} \mathbf{v} \right), \quad (4.39)$$

и поэтому

$$\mathbf{f} = \sqrt{m} D (\sqrt{m} \mathbf{v}) = \frac{1}{2} (m D \mathbf{v} + D(m \mathbf{v})), \quad m = \frac{m_0}{\alpha^2}, \quad (4.40)$$

подтверждая (3.6) и (3.9).

С учетом второго постулата Эйнштейна для \mathbf{f}' , \mathbf{v}' и α' можно вывести выражение точно такого же типа:

$$\mathbf{f}' = \frac{m_0}{\alpha'} D' \left(\frac{1}{\alpha'} \mathbf{v}' \right).$$

В частности, если $v_y \equiv v_z \equiv 0$, т.е. если все события происходят только вдоль осей x и x' , имеем $\alpha_x = \alpha$, $\alpha'_x = \alpha'$, а затем с использованием (4.14) получаем

$$\alpha \mathbf{f} = \alpha' \mathbf{f}'. \quad (4.41)$$

В более общем смысле имеем

$$f_x = \frac{1}{\alpha_0} \left(\frac{\beta_0}{c} \mathbf{v}^T \mathbf{f}' + f'_x \right), \quad f_y = f'_y, \quad f_z = f'_z, \quad (4.42)$$

$$\mathbf{v}^T \mathbf{f} = \frac{1}{\alpha_0} (\mathbf{v}^T \mathbf{f}' + c \beta_0 f'_x) \quad (4.43)$$

где все величины уже были описаны выше. Этот результат можно проверить либо непосредственно при помощи (4.1) и т.д., либо с учетом того, что собственный 4-вектор, такой как сила Минковского (см. (3.11)), преобразуется из \mathbf{S}' в \mathbf{S} умножением этого вектора в \mathbf{S}' в обратном порядке на матрицу

$$\mathbf{M} = \begin{pmatrix} \frac{\beta_0}{\alpha_0} & \frac{1}{\alpha_0} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \frac{1}{\alpha_0} & \frac{\beta_0}{\alpha_0} & 0 & 0 \end{pmatrix}, \quad (4.44)$$

(обратная матрица \mathbf{M}^{-1} получается простым изменением знака β_0). Можно показать, что (4.41) – это частный случай данного общего случая.

Отметим, что полный вывод (от (4.32) до (4.38)) включает в себя \mathbf{v}' только в $(D' \mathbf{v}')_1$. Это интересно, поскольку так подчеркивается значимость этой величины как асимптотической меры. С другой стороны, чтобы получить классическое релятивистское выражение для силы, потребуется заменить правую часть (4.33), что видно из (4.34), на $1/\alpha_0$, а это достаточно неожиданное требование.

До сих пор m_0 считалась постоянной. Тем не менее, учет масс покоя m_0 может представлять интерес (m_0 не зависят от \mathbf{v} , но зависят от t). Допустим, что в этом случае (4.32) имеет вид

$$\mathbf{f}'_1 = \sqrt{m_0} (D' \sqrt{m_0} \mathbf{v})_1.$$

Легко проверить, что при этом допущении процесс получения (4.40) будет верным, но (4.39) в общем случае выполняться не будет.

5. Модификация третьего закона Ньютона и сохранение импульса

В классической теории относительности третий закон Ньютона остается неизменным. Обычно даже не упоминается, что это предполагает соответствующее дополнительное допущение. Если придерживаться другой точки зрения, то такое допущение необоснованно, т.е. необходимо модифицировать не только второй, но и третий закон Ньютона.

Рассмотрим действие и реакцию (противодействие) между двумя частицами P_1 и P_2 , индексы 1 и 2 в данном разделе обозначают величины, характеризующие P_1 и P_2 , соответственно. Считаем, что P_1 и P_2 движутся со скоростями \mathbf{v}_1 и \mathbf{v}_2 . Пусть P_1 и P_2 взаимодействуют, а \mathbf{f}_1 и \mathbf{f}_2 – результирующие силы, действующие на P_1 и P_2 , соответственно. Сначала допустим, что \mathbf{f}_1 , \mathbf{f}_2 , \mathbf{v}_1 и \mathbf{v}_2 расположены на одной и той же прямой линии, которую будем считать осью x системы S , (аналогично – осью x' системы S'); в отношении S' , соответственно, рассматриваем \mathbf{f}'_1 , \mathbf{f}'_2 , \mathbf{v}'_1 , \mathbf{v}'_2 . Запишем

$$\beta_1 = \mathbf{v}_{x1}/c, \quad \beta'_1 = \mathbf{v}'_{x1}/c, \quad \beta_2 = \mathbf{v}_{x2}/c, \quad \beta'_2 = \mathbf{v}'_{x2}/c, \quad \alpha_i = \sqrt{1 - \beta_i^2}, \quad \alpha'_i = \sqrt{1 - \beta_i'^2}, \quad i = 1, 2, \quad (5.1)$$

можно записать (см.(4.10)),

$$\beta'_1 = \frac{\beta_1 - \beta_0}{1 - \beta_1 \beta_0}, \quad \beta'_2 = \frac{\beta_2 - \beta_0}{1 - \beta_2 \beta_0}, \quad \beta_0 = \mathbf{v}_0/c. \quad (5.2)$$

Пусть P_1 и P_2 касаются друг друга при $t = t_0$, откуда $x_1(t_0) = x_2(t_0)$ и тогда (см. (4.1)) $x'_1(t'_0) = x'_2(t'_0)$, $t'_0 = t'_1(t_0) = t'_2(t_0)$. Для упрощения записи используем индекс «0» для величин при $t = t_0$, тогда $t' = t'_0$ (но при этом \mathbf{v}_0 и β_0 имеют то же значение, что и выше). Если $\beta_{10} = -\beta_{20}$, примем $S' = S$. Если $\beta_{10} \neq -\beta_{20}$, то, выбирая β_0 такие, что

$$\beta_0^2 - 2b\beta_0 + 1 = 0, \quad b = (1 + \beta_{10}\beta_{20})(\beta_{10} + \beta_{20}),$$

которые с учетом

$$b^2 = 1 + \left(\frac{\alpha_{10}\alpha_{20}}{\beta_{10} + \beta_{20}} \right)^2,$$

всегда дают действительные значения β_0 , получаем, как можно показать (см. (5.2)),

$$\beta'_{10} = -\beta'_{20}, \quad \text{тогда} \quad \alpha'_{10} = \alpha'_{20}. \quad (5.3)$$

Отметим, что произведение двух вариантов значений β_0 равняется 1, т.е. необходимо выбрать тот вариант, при котором $|\beta_0| < 1$.

Благодаря предположению об ориентации скоростей и сил в системах координат, мы можем применить (4.41) и записать, в частности

$$\alpha_{10}\mathbf{f}_{10} = \alpha'_{10}\mathbf{f}'_{10}, \quad \alpha_{20}\mathbf{f}_{20} = \alpha'_{20}\mathbf{f}'_{20}. \quad (5.4)$$

Но вследствие симметрии скоростей в S' необходимо иметь $\mathbf{f}'_{10} = -\mathbf{f}'_{20}$. Отсюда (5.3) и (5.4) дают

$$\alpha_{10}\mathbf{f}_{10} = -\alpha_{20}\mathbf{f}_{20}. \quad (5.5)$$

Далее, если при сохранении предположения об ориентации $\mathbf{v}_1, \mathbf{v}_2, \mathbf{f}_1, \mathbf{f}_2$, отвергнуть допущение о совпадении с направлением оси x , либо оси x' , то утверждение об ориентации, записанное аналогично записи в Разделе 4.2 (после (4.38)) показывает, что (5.5) остается верным.

Теперь рассмотрим общий случай соотношения между действием и реакцией (противодействием). С учетом (5.5) ясно, что равенство из классической теории $\mathbf{f}_1 = -\mathbf{f}_2$ здесь выполняться не может. Однако для модификации третьего закона Ньютона самой простой записью, которая не противоречит (5.5), будет

$$\alpha_1\mathbf{f}_1 = -\alpha_2\mathbf{f}_2 \quad (5.6)$$

где индексы величин α_1 и α_2 обозначают, как было упомянуто выше, принадлежность соответствующим частицам (см. (5.1)).

Далее, пусть m_{10} и m_{20} - массы покоя частиц P_1 и P_2 . Из (4.39) и (5.6) следует, что

$$D(\mathbf{p}_1 + \mathbf{p}_2) = 0$$

где

$$\mathbf{p}_1 = m_{10}\mathbf{v}_1 / \alpha_1 \quad \text{и} \quad \mathbf{p}_2 = m_{20}\mathbf{v}_2 / \alpha_2$$

– импульсы, как упоминается в разделе 3. В более общем смысле, если имеется n частиц от P_1 до P_n и если силы от $\mathbf{f}_{\nu 1}$ до $\mathbf{f}_{\nu k}$ действуют на P_ν , которая движется со скоростью \mathbf{v}_ν , имеем

$$D\mathbf{p}_\nu = \alpha_\nu (\mathbf{f}_{\nu 1} + \dots + \mathbf{f}_{\nu k}) \quad (5.7)$$

где

$$\mathbf{p}_\nu = m_{\nu 0}\mathbf{v}_\nu / \alpha_\nu, \quad \alpha_\nu = \sqrt{1 - \beta_\nu^2}, \quad \beta_\nu = v_\nu / c, \quad v_\nu^2 = \mathbf{v}_\nu^T \mathbf{v}_\nu.$$

Суммируя (5.7) по всем частицам, с учетом попарного сокращения в правой части благодаря соотношениям вида (5.6), получим

$$D(\mathbf{p}_1 + \dots + \mathbf{p}_n) = 0. \quad (5.8)$$

Следовательно, сохранение импульса выполняется точно так же, как в классической теории относительности. Наоборот, если в альтернативной теории мы требуем, чтобы (5.8) выполнялось во всех случаях, необходимо изменить третий закон Ньютона, как показано в (5.6).

6. Частицы, движущиеся в полях, – силы и энергия

Допустим, что поле (электромагнитное, гравитационное и т.д.) действует с силой \mathbf{f} на частицу P , которая движется со скоростью \mathbf{v} , что устройство (источник поля) $E q$ (далее обозначим его индексом 1), генерирующее поле, покоится ($\mathbf{v}_1 = 0$) и что реакция поля, создаваемого этим устройством $E q$ – это \mathbf{f}_1 . Кроме того, допустим, что все расстояния достаточно малы, откуда можно утверждать (см. (5.6)),

$$\alpha \mathbf{f} = -\alpha_1 \mathbf{f}_1 = -\mathbf{f}_1, \quad (6.1)$$

где α и α_1 аналогичны (3.2) и соответствующим данным из Раздела 5 (за исключением того, что $\mathbf{E}q$ выполняет роль P_1 , откуда $\alpha_1 = 1$). Если \mathbf{f}_1 не зависит от скорости \mathbf{P} , то она должна равняться силе, существующей при $\mathbf{v} = 0$. Значит, из (6.1) можно заключить, что

$$\mathbf{f} = \mathbf{f}_0 / \alpha \quad (6.2)$$

где \mathbf{f}_0 – сила, которая действует на \mathbf{P} , если $\mathbf{v} = \mathbf{0}$, и таким образом она равна $-\mathbf{f}_1$, т.е. \mathbf{f}_0 – сила в том понимании, как ее используют в классических выражениях. Например, $\mathbf{f}_0 = q\mathbf{E}$, если q – заряд, а \mathbf{E} – напряженность электрического поля (аналогично и в случае гравитационного поля).

В случае электромагнитного поля это подтверждается и, фактически, распространяется на силу Лоренца, если действовать как в классической теории (см. например Раздел 7.7 работы [19]), т.е. если в рассматриваемый момент времени получить \mathbf{f} посредством преобразования Лоренца силы $q\mathbf{E}'$, наблюдаемой в другой соответствующей системе координат \mathbf{S}' . Чтобы показать это, сначала вспомним, что рассматриваемая здесь сила \mathbf{f} тождественна соответствующей части 4-мерного вектора Минковского (см. Раздел 3) и что электрическое и магнитное поля можно преобразовывать для систем \mathbf{S}' и \mathbf{S} согласно соотношениям

$$\mathbf{E}'_x = \mathbf{E}_x, \quad \mathbf{E}'_y = \frac{1}{\alpha_0}(\mathbf{E}_y - v_0 \mu \mathbf{H}_z), \quad \mathbf{E}'_z = \frac{1}{\alpha_0}(\mathbf{E}_z + v_0 \mu \mathbf{H}_y), \quad (6.3)$$

$$\mathbf{H}'_x = \mathbf{H}_x, \quad \mathbf{H}'_y = \frac{1}{\alpha_0}(\mathbf{H}_y + v_0 \varepsilon \mathbf{E}_z), \quad \mathbf{H}'_z = \frac{1}{\alpha_0}(\mathbf{H}_z - v_0 \varepsilon \mathbf{E}_y), \quad (6.4)$$

Здесь величины с штрихом, опять же, относятся к \mathbf{S}' , без штриха – к \mathbf{S} .

Затем рассмотрим заряд q , который мгновенно неподвижен в системе \mathbf{S}' , т.е. допустим, что ориентация \mathbf{S} , \mathbf{S}' и выбор \mathbf{v}_0 (см. (4.19)) таковы, что в рассматриваемый момент времени $\mathbf{v}' = \mathbf{0}$, тогда $\mathbf{v} = \mathbf{v}_0$ (см. (4.10)) и поэтому $\alpha_0 = \alpha$. Сила \mathbf{f}' , действующая на q , тогда запишется так:

$$\mathbf{f}' = q\mathbf{E}', \quad \mathbf{E}' = (\mathbf{E}'_x, \mathbf{E}'_y, \mathbf{E}'_z)^T.$$

Из (4.42) и (6.3), таким образом, получим с использованием магнитной индукции $\mathbf{B} = (\mathbf{B}_x, \mathbf{B}_y, \mathbf{B}_z)^T = \mu \mathbf{H}$:

$$f_x = q \frac{1}{\alpha} E_x, \quad f_y = q \frac{1}{\alpha} (E_y - v_0 B_z), \quad f_z = q \frac{1}{\alpha} (E_z - v_0 B_y).$$

Этот результат можно переписать в виде (6.2) с

$$\mathbf{f}_0 = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (6.5)$$

В том виде, в каком результат представлен в (6.2) и (6.5), он выражается только через величины, определенные в \mathbf{S} , и в общей векторной записи. Поэтому он справедлив без ограничений.

С другой стороны, сила \mathbf{f} , действующая на частицу с массой покоя m_0 , также удовлетворяет (4.39). Поэтому получаем из (6.2):

$$m_0 D \left(\frac{1}{\alpha} \mathbf{v} \right) = \mathbf{f}_0. \quad (6.6)$$

Из (6.6) делаем вывод, что альтернативная теория дает точно такое же динамическое поведение частицы в поле, как и классическая теория относительности.

Однако сюда не входит энергия. Рассмотрим отношение из классической электродинамики:

$$-\rho \mathbf{v}^T \mathbf{E} = \frac{1}{2} D(\varepsilon \mathbf{E}^T \mathbf{E} + \mu \mathbf{H}^T \mathbf{H}) + \operatorname{div}(\mathbf{E} \times \mathbf{H})$$

(записанное в стандартном виде), где $\rho = \varepsilon \operatorname{div} \mathbf{E}$ – плотность заряда, ε и μ считаются постоянными (например, $\varepsilon = \varepsilon_0$, $\mu = \mu_0$). Чтобы получить правильное выражение для подводимой энергии, альтернативная теория требует (как было показано в этом разделе) умножить классическую плотность силы $\rho \mathbf{E}$ в левой части на $1/\alpha$. Следовательно, правую часть необходимо умножить таким же образом. Рассмотрим здесь только случай с постоянной α (см. также Раздел 7). Ясно, что тогда для плотности энергии поля имеем

$$w = \frac{1}{2\alpha} (\varepsilon \mathbf{E}^T \mathbf{E} + \mu \mathbf{H}^T \mathbf{H}), \quad (6.7)$$

и вектор Пойнтинга аналогичным образом становится $(\mathbf{E} \times \mathbf{H})/\alpha$. Классическое выражение $(\varepsilon \mathbf{E}^T \mathbf{E} + \mu \mathbf{H}^T \mathbf{H})/2$ тогда представляет собой нечто вроде плотности кажущейся энергии.

Отметим, что в этом и последующих разделах с помощью W и w мы обозначаем энергию и плотность энергии, чтобы различать энергии и электрические поля. Также отметим, что если вместо заряда q рассматривать плотность заряда ρ , движущегося со скоростью \mathbf{v} , то можно заменить (6.5) на

$$\hat{\mathbf{f}} = \hat{\mathbf{f}}_0/\alpha, \quad \hat{\mathbf{f}}_0 = \rho(\mathbf{E} + \mathbf{v} \times \mathbf{B})$$

где $\hat{\mathbf{f}}$ и $\hat{\mathbf{f}}_0$ – плотность силы, т.е. сила на единицу объема.

7. Электромагнитное поле,двигающееся с постоянной скоростью

Обозначения \mathbf{S} и \mathbf{S}' – те же самые. Сначала рассмотрим поле \mathbf{E}' , которое в системе \mathbf{S}' является электростатическим, т.е. покоящимся. Ясно, что \mathbf{E}' не зависит от t' . Несмотря на то, что для соответствующего магнитного поля имеем $\mathbf{H}' = 0$, поле в системе \mathbf{S} содержит и \mathbf{E} , и \mathbf{H} . Т.к. мы считаем ε и μ постоянными, нам не нужно учитывать \mathbf{D} и \mathbf{B} . В системе \mathbf{S} поле движется с постоянной скоростью \mathbf{v} . Имеем $\mathbf{v} = \mathbf{v}_0$ (см. (4.10)) и поэтому,

$$\mathbf{v} = \mathbf{v}_0, \quad \alpha = \alpha_0, \quad \beta = \beta_0, \quad (7.1)$$

Здесь мы используем те же обозначения, какие были введены (за исключением того, что считаем, что \mathbf{v} принимает знак \mathbf{v}_0). Из (6.3), (6.4) и (7.1) получаем для любого \mathbf{E}' и \mathbf{H}'

$$\begin{aligned} E_x &= E'_x, & E_y &= \frac{1}{\alpha} (E'_y + v \mu H'_z), & E_z &= \frac{1}{\alpha} (E'_z - v \mu H'_y), \\ H_x &= H'_x, & H_y &= \frac{1}{\alpha} (H'_y - v \varepsilon E'_z), & H_z &= \frac{1}{\alpha} (H'_z + v \varepsilon E'_y). \end{aligned} \quad (7.2)$$

Тогда для плотности энергии поля, определяемой выражением (6.7), найдем при $\mathbf{H}' = 0$,

$$w = \frac{1}{\alpha} w' + \frac{\beta^2}{\alpha^3} w'_0, \quad (7.3)$$

где

$$\mathbf{w}' = \frac{1}{2}\varepsilon(\mathbf{E}_x'^2 + \mathbf{E}_y'^2 + \mathbf{E}_z'^2), \quad \mathbf{w}'_0 = \varepsilon(\mathbf{E}_y'^2 + \mathbf{E}_z'^2), \quad (7.4)$$

\mathbf{w}' – плотность энергии поля в системе \mathbf{S}' . С учетом электростатической гипотезы правая часть (7.3) не зависит от t' .

Тогда пусть

$$dV = dx \cdot dy \cdot dz, \quad dV' = dx' \cdot dy' \cdot dz'$$

– элементарные объемы в системах \mathbf{S} и \mathbf{S}' , соответственно. Как известно, из (4.1) следует, что при любом фиксированном t

$$dx = \alpha dx', \quad dy = dy', \quad dz = dz',$$

так что $dV = \alpha dV'$. Тогда из (7.3) получаем

$$w dV = w' dV' + \frac{\beta^2}{\alpha^2} w'_0 dV'. \quad (7.5)$$

Это выражение записано в форме, определяемой (3.8) и (3.10), где \mathbf{E} , \mathbf{E}_i и \mathbf{E}_0 заменены на $w dV$, $w' dV'$ и $w'_0 dV'$, соответственно, так, чтобы не существовало строгой связи между \mathbf{w}' и \mathbf{w}'_0 . Можно также проинтегрировать (7.5) по всему объему, что приводит к $W = W' + W_k$ где

$$W = \int_V w dV, \quad W' = \int_{V'} w' dV', \quad W_k = \frac{\beta^2}{\alpha^2} \int_{V'} w'_0 dV'$$

Т.к. \mathbf{w}' и \mathbf{w}'_0 не зависят от t' , то можно считать, что W' и W_k определялись при постоянном t' . Кроме того, V и V' могут быть пространством (entire space).

Тогда во всех случаях, очевидно, получаем полное соответствие результатам из предыдущих разделов. Разложение (decomposition) на энергию покоя и кинетическую энергию, которое встречается в классической теории вероятности, здесь не дает такого же изящного объяснения. Это интересно, поскольку известно, что уравнения Максвелла заведомо совместимы с преобразованием Лоренца.

Второе выражение (7.4) можно также записать в виде:

$$\mathbf{w}'_0 = \varepsilon \mathbf{E}'^T \mathbf{E}' - \varepsilon \mathbf{E}_x'^2.$$

Но в связи с $\mathbf{v} = \mathbf{v}_0$ и $V = V_0$ имеем (см. (7.1)) $\mathbf{v} \mathbf{E}'_x = \mathbf{v}^T \mathbf{E}'$. Отсюда (7.5) можно выразить в более общем виде

$$w dV = \left(\mathbf{w}' + \frac{1}{2\alpha^2} \mathbf{v}^T \mathbf{m}'_0 \mathbf{v} \right) dV', \quad (7.6)$$

где

$$\mathbf{m}'_0 = \frac{2\varepsilon}{c^2} \left((\mathbf{E}'^T \mathbf{E}') \mathbf{1} - \mathbf{E}' \mathbf{E}'^T \right), \quad (7.7)$$

$\mathbf{1}$ – единичная матрица порядка 3. Эти объяснения выполняются для любой ориентации \mathbf{v} и имеют вид как в случае с матрицей, который обсуждался в связи с (3.14).

Теперь рассмотрим в общем случае электромагнитное поле. Подстановка (7.2) в (6.7) дает, после некоторых вычислений, более общее выражение

$$w dV = \left(\mathbf{w}' + \frac{1}{2\alpha^2} \mathbf{v}^T \mathbf{m}'_0 \mathbf{v} + \frac{2}{c^2 \alpha^2} \mathbf{v}^T \mathbf{S}' \right) dV' \quad (7.8)$$

где

$$\mathbf{w}' = \frac{1}{2}(\varepsilon \mathbf{E}'^T \mathbf{E}' + \mu \mathbf{H}'^T \mathbf{H}'), \quad \mathbf{S}' = \mathbf{E}' \times \mathbf{H}', \quad (7.9)$$

$$\mathbf{m}'_0 = \frac{2}{c^2} \left((\varepsilon \mathbf{E}'^T \mathbf{E}' + \mu \mathbf{H}'^T \mathbf{H}') \mathbf{1} - (\varepsilon \mathbf{E}' \mathbf{E}'^T + \mu \mathbf{H}' \mathbf{H}'^T) \right). \quad (7.10)$$

Сравнивая с (7.6), (7.7) и первым уравнением (7.4), получаем, что изменения в \mathbf{w}' и \mathbf{m}'_0 такие, как ожидалось. Однако есть дополнительный член, который пропорционален компоненте вектора Пойнтинга в \mathbf{S}' в направлении \mathbf{v} , а значит, компоненте соответствующей плотности импульса в \mathbf{S}' электромагнитного поля. Этот дополнительный член исчезает, в частности, если в рассматриваемой точке \mathbf{P}' (т.е. в точке пространства/времени) в системе \mathbf{S}' поле обладает нулевым импульсом, т.е. в определенном смысле мгновенно неподвижно. Если затем применить (6.7) не к \mathbf{S} , а к \mathbf{S}' , то (6.7) сведется к первому уравнению (7.4). Отсюда опять возникает принцип, аналогичный ньютоновскому пределу (но теперь в электромагнитном смысле). Можно утверждать, что \mathbf{w}' при принятых допущениях на самом деле равняется плотности энергии поля в \mathbf{P}' .

8. Эксперимент Бертоцци

В 1964 Бертоцци [26] опубликовал ряд экспериментальных результатов, касающихся определения кинетической энергии быстрых электронов. Если использовать более подходящую для наших целей запись, то он определил величины $\beta_b^2 = 1 - \alpha_b^2$, а значит, v_b^2/c^2 , с точки зрения того, что мы обозначаем здесь через

$$\gamma = qu_0/m_0c^2,$$

q – заряд электрона, u_0 – напряжение, преодолеваемое электронами в ускорителе, m_0 – масса покоя электрона, v_b – скорость электрона непосредственно перед столкновением с целью, c – скорость света. В соответствии с классической теорией относительности связь α_b с другими величинами простая:

$$\frac{m_0c^2}{\alpha_b} - m_0c^2 = qu_0, \quad (8.1)$$

откуда

$$\alpha_b^2 = 1 - \beta_b^2 = 1/(1 + \gamma)^2, \quad \text{при} \quad \beta_b = v_b/c. \quad (8.2)$$

Т.к. в соответствии с альтернативной теорией динамическое поведение частиц в полях точно такое же, как описывает классическая теория, то результат (8.2) одинаково справедлив в обеих теориях (в отличие от того, что сказано в [21]).

Пусть E_k – кинетическая энергия электрона, обладающего скоростью v_b . Согласно классической теории относительности имеем $E_k = qu_0$, откуда в [26] было принято, что γ равна E_k/m_0c^2 . Согласно альтернативной теории, однако, $E_k = m_0v_b^2/2\alpha_b^2$, что с учетом (8.2) дает

$$E_k = qu_0(1 + \gamma/2) = m_0c^2\gamma(1 + \gamma/2), \quad (8.3)$$

и именно это расхождение привело к путанице в [21].

Чтобы проверить, справедливо ли приравнять E_k и qu_0 , Бертоцци также определил тепловую энергию (для двух случаев) E_n , сгенерированную при столкновении

электрона с целью. Пусть $\delta = E_h / qu_0$. Т.к. существуют неизбежные потери (например, на рентгеновское излучение), то имеем $E_h < E_k$, откуда $\delta < E_k / qu_0$. Следовательно, $\delta < 1$ если на самом деле $E_k = qu_0$. В обоих случаях (1.5 МэВ и 4.5 МэВ) результат измерения был $\delta = 1.067$, что согласуется с $\delta < 1$ только в том случае, если сделать поправку на низкую точность эксперимента (10 %, как указано в [26]).

Чтобы исследовать ситуацию с точки зрения альтернативной теории, пусть f – сила торможения, действующая на электрон со стороны цели, f_f – соответствующая сила трения, действующая на цель, которую мы считаем покоящейся. С учетом альтернативного вида третьего закона Ньютона имеем

$$\alpha f = f_f, \quad \alpha = \sqrt{1 - \beta^2}, \quad \beta = v/c,$$

т.е. с учетом (4.39) $f_f = m_0 D(v/\alpha)$, v – скорость электрона в рассматриваемый момент времени. Значит (см. (3.5)), тепловая энергия, сгенерированная трением, имеет вид $vf_f = D(m_0 c^2 / \alpha)$, а теплота, сгенерированная в процессе торможения, тогда записывается так:

$$\left[\frac{m_0 c^2}{\alpha} \right]_{v=0}^{v=v_b} = \frac{m_0 c^2}{\alpha_b} - m_0 c^2 = qu_0,$$

второе равенство, на самом деле, является следствием (8.1). Избыточная энергия $E_k - qu_0$ (см. (8.3)) частично превратится в энергию излучения, а частично – в кинетическую энергию цели и ее опоры (практически не поддается измерению), но, вследствие вторичных процессов, ее малая часть опять превратится в теплоту, вследствие чего $E_h > qu_0$, т.е. $\delta > 1$. Это согласуется с результатами эксперимента, т.е. допускать недостаточную точность эксперимента уже не нужно.

9. Матрица

В разделах 3 и 7 кратко упоминалось, что массу некоторых объектов следует описывать не скаляром, как обычно это делается, а матрицей. Рассмотрим некоторые свойства такого объекта \mathbf{P} (подобного частице). Будем считать, что при исчезающе малой скорости этот объект характеризуется постоянной (т.е. не зависящей от времени) симметричной положительной определенной матрицей \mathbf{m}_0 , которая называется *массой покоя в матричном виде*.

Вначале вернемся к выводу из Раздела 4.2, т.е. замене m_0 на \mathbf{m}_0 , т.е. замене (4.32) на

$$\mathbf{f}'_i = \mathbf{m}_0 (D' \mathbf{v}')_i. \quad (9.1)$$

Однако если \mathbf{m}_0 – «полностью занятая» (occupied) матрица, то нельзя больше принимать без доказательства, что величина ΔW_r , определяемая по (4.25) и (4.26), является правильной оценкой $\Delta W'$. В самом деле, если \mathbf{m}_0 на самом деле является такой матрицей, то невозможно удовлетворительно разграничить события в направлениях x , y , и z , в то же время в связи с (4.19) поправочный член (4.26) выделяет направление x . Поэтому вначале мы допускаем, что направления x' , y' и z' совпадают с главными осями \mathbf{m}_0 , т.е. что \mathbf{m}_0 является диагональной, таким образом, $\mathbf{m}_0 = \text{diag}(m_1, m_2, m_3)$, и опять можно утверждать, что (4.34) выполняется.

Однако с учетом (9.1) $(D'\mathbf{v})_1$, которое появляется в (4.18), более не равняется скаляру (ранее: $1/m_0$, см. (4.32)), умноженному на \mathbf{f}'_1 , поэтому два векторных множителя \mathbf{f}'_1 в правой части (4.35) необходимо заменить на исходный множитель $(D'\mathbf{v})_1$. Необходимо повторить эти действия в отношении трех скалярных множителей в правой части (4.36), где, в соответствии с (9.1), из $\forall \mathbf{f}'_1$ следует $\forall (D'\mathbf{v})_1$, поэтому (4.37) все еще справедливо. В соответствии с этим результатом, можно опять выполнить те же шаги, что в Разделе 4.2, используя свойство диагональности \mathbf{m}_0 . Так можно установить, что (4.38) следует заменить на

$$\mathbf{f}_1 = \frac{1}{\alpha_1} \mathbf{m}_0 \left(D \frac{\mathbf{v}}{\alpha} \right)_1$$

А значит, (4.39) и (4.40) необходимо заменить на (3.13), где \mathbf{m} имеет вид (3.12). Отметим, что при исчезающе малой скорости \mathbf{P} опять полностью характеризуется \mathbf{m}_0 ; это согласуется с первым постулатом Эйнштейна, но выполняется в только что упомянутом строгом виде только благодаря особому выбору x', y', z' относительно x, y, z (в отличие от случая с ранее рассмотренным скаляром m_0).

Теперь отбросим допущение о диагональности \mathbf{m}_0 . Пусть \mathbf{P} движется в \mathbf{S} со скоростью \mathbf{v} и испытывает там воздействие силы \mathbf{f} ; \mathbf{m}_0 – масса в матричном виде при исчезающе малой скорости, т.е. масса покоя в матричном виде. Пусть \mathbf{S}_0 – другая система координат, которая находится в зафиксированном, но повернутом положении относительно \mathbf{S} . Таким образом, существует постоянная ортогональная матрица \mathbf{U} , такая что $\mathbf{U}\mathbf{f}$ и $\mathbf{U}\mathbf{v}$ – сила и скорость в \mathbf{S}_0 . Т.к. при исчезающе малой скорости имеем $\mathbf{f} = \mathbf{m}_0 D \mathbf{v}$, масса покоя в матричном виде в \mathbf{S}_0 составляет $\mathbf{U}\mathbf{m}_0 \mathbf{U}^T$.

С другой стороны, т.к. \mathbf{m}_0 постоянна, то можно, на самом деле, выбрать постоянную матрицу \mathbf{U} , такую что $\mathbf{U}\mathbf{m}_0 \mathbf{U}^T$ диагональная. Значит, используя полученный выше результат, можно записать для \mathbf{S}_0 ,

$$\mathbf{U}\mathbf{f} = \frac{1}{\alpha} (\mathbf{U}\mathbf{m}_0 \mathbf{U}^T) D \left(\frac{1}{\alpha} \mathbf{U}\mathbf{v} \right). \quad (9.2)$$

Отсюда опять следует, что (3.13) выполняется; при этом \mathbf{m} имеет вид (3.12).

Этот результат подтверждает правильность распространения случая со скаляром на случай с матрицей в том простом виде, который был выбран в Разделе 3. Однако существуют и некоторые расхождения, которые касаются, в частности, классического понятия преобразования 4-мерных векторов между системами \mathbf{S} и \mathbf{S}' умножением их слева на матрицу \mathbf{M} (см. (4.44)) или обратную ей матрицу. Классическое доказательство этого метода в случае с силой Минковского не применимо, т.к. \mathbf{M} и \mathbf{m}_0 не коммутирующие в общем случае. Тем не менее, достаточно простые доказательства можно привести с использованием таких изящных выражений, как

$$1 - \beta_0 \beta_x = \alpha_0 \alpha' / \alpha', \quad 1 + \beta_0 \beta'_x = \alpha_0 \alpha' / \alpha \quad 1 - \beta_x \beta'_x = \alpha_x \alpha'_x / \alpha_0, \quad \alpha' / \alpha = \alpha'_x / \alpha_x \quad (9.3)$$

которые, как можно показать, следуют из (4.10) (см. также (4.8)), при этом сохраняются все приведенные выше обозначения.

Выведем выражение, эквивалентное (4.42). При помощи (3.13), (4.8) (4.10) и (9.3) получим

$$\mathbf{f} = \mathbf{m}_0 \frac{1}{\alpha} \mathbf{D} \left(\frac{\mathbf{v}}{\alpha} \right) = \mathbf{m}_0 \frac{1}{\alpha'} \mathbf{D}' \left(\frac{1}{\alpha'} ((\mathbf{v}'_x + v_0) \alpha_0, v'_y, v'_z)^T \right). \quad (9.4)$$

Отсюда, применяя эквивалент (3.5) для системы \mathbf{S}' , получим

$$\mathbf{f} = \mathbf{m}_0 \frac{1}{\alpha'} \mathbf{D}' \left(\frac{1}{\alpha'} \begin{pmatrix} \mathbf{v}'_x / \alpha_0 \\ v'_y \\ v'_z \end{pmatrix} \right) + \mathbf{m}_0 \frac{\beta_0}{2c\alpha_0} \begin{pmatrix} \mathbf{D}'(\mathbf{v}'\alpha')^2 \\ 0 \\ 0 \end{pmatrix}.$$

Обычно это выражение нельзя свести к виду (4.42), который бы мы получили при использовании стандартной методики на основе (4.44). Для диагональной \mathbf{m}_0 , пусть $\mathbf{m}_0 = \text{diag}(m_1, m_2, m_3)$, найдем

$$f_x = \frac{1}{\alpha_0} \left(f'_x + m_1 \frac{\beta_0}{c\alpha'} \mathbf{v}'^T \mathbf{D}' \left(\frac{\mathbf{v}'}{\alpha'} \right) \right), \quad f_y = f'_y, \quad f_z = f'_z,$$

но даже это выражение сводится к (4.42) только тогда, когда $m_1 = m_2 = m_3$, что эквивалентно случаю со скаляром.

В данном случае можно привести тот же комментарий, что и в конце Раздела 4.2. В частности, на рассматриваемый здесь объект могут влиять не только силы (см. Раздел 7), поэтому может возникнуть неявная зависимость от времени.

10. Выводы

Правильное определение отсутствия потерь в цепях Кирхгофа требует, чтобы определяющие уравнения для элементов, таких как индуктивность, были записаны в очень специфическом виде, который, в свою очередь, имеет фундаментальное физическое значение. Удивительно, что этот вид записи не учитывается в классическом релятивистском выражении для действующей на частицу силы. Однако требуя, чтобы релятивистские законы совпадали с ньютоновскими законами при приближении к моменту времени, когда частица является «мгновенно неподвижной», и применяя этот принцип ньютоновского предела, в частности, к силе и совершенной работе, получим выражение для силы, а значит, релятивистскую формулировку второго закона Ньютона, которая отлично согласуется с теорией нелинейных цепей Кирхгофа. Это отличие появляется в том случае, если основное внимание уделяется энергетическим аспектам (отсутствие потерь, совершенная работа и т.д.), а не импульсу, как в классической теории относительности. Альтернативный способ модификации второго закона Ньютона требует соответствующей модификации третьего закона, который в классической теории относительности остается неизменным, а также выражений для сил, действующих в полях. Из совокупности этих изменений следует, что сохранение импульса и динамика частиц остаются точно такими же, как в классической теории относительности, что подтверждено и детализировано применением преобразований Лоренца для уравнений Максвелла.

Энергия покоя частиц оказывается произвольной постоянной интегрирования, а ее значение полностью согласуется с выражением из классической теории относительности. Определенные различия между классической и альтернативной теориями становятся очевидными для кинетической энергии, которая в рамках альтернативной теории приемлема (rational) по скорости, а в рамках классической теории – нет. Для электрического поля, статического в отношении движущейся равномерно относительно наблюдателя системы координат, кинетический член энергии

поля полностью и напрямую согласуется с альтернативным выражением для кинетической энергии, т.е. отличается от классического выражения. Ряд опубликованных Бертоцци результатов измерения кинетической энергии быстрых электронов поддерживает альтернативную теорию, но упоминаемая автором погрешность в 10% слишком высока, чтобы делать какие-либо точные выводы. Можно выдвинуть более общие идеи, касающиеся природы частиц и согласующиеся с представленными результатами. Кратко эти идеи представлены в конце работы [21], хотя некоторые аспекты пересмотрены и раскрыты в настоящей статье.

Благодарности

Автор благодарит. Bouten, M. Depenbrock, P. Fettweis, A. Hahn, J. Kummer, K.-A. Owenier, H. Rollnick, R. Staufenbiel за обсуждение статьи и F.-J. Reich за помощь в подготовке рукописи.

Список литературы

- [1] W. Cauer, "Theorie der linearen Wechselstromschaltungen", 2nd ed., Akademie-Verlag, Berlin, 1954. American edition: "Synthesis of Linear Communication Networks", McGraw Hill, New York, 1958
- [2] V. Belevitch, "Classical Network Theory", Holden-Day, San Francisco, 1968.
- [3] J. Meixner, "Beziehungen zwischen Netzwerktheorie und Thermodynamik", Arbeitsgemeinschaft für Forschung des Landes Nordrhein-Westfalen, Heft 181, Westdeutscher Verlag, Köln, 1967.
- [4] Fettweis, "In Memoriam - Vitold Belevitch", IEEE Transactions on Circuits and Systems - I, vol. 47, pp. 613 - 614, May 2000.
- [5] Fettweis, "Wave digital filters: theory and practice", Proc. IEEE, vol. 74, pp. 270-327, Feb. 1986, and vol.75, p. 729, May 1987.
- [6] H. D. Fischer, "Wave digital filters for numerical integration", ntz-Archiv, vol. 6, pp. 37-40, 1984.
- [7] Fettweis, "Discrete passive modelling of physical systems described by partial differential equations", in "Multivariate Analysis: Future Directions"(C.R. Rao, editor), pp. 115-130, Elsevier, Amsterdam, 1993.
- [8] Fettweis, "Improved wave-digital approach to numerically integrating the PDEs of fluid dynamics", Proc. IEEE Int. Symp. Circuits and Systems, vol. III, pp. 361-364, Scottsdale, AZ, USA, 26-29 May 2002.
- [9] Fettweis, "Robust numerical integration using wave-digital concepts", Proceedings of the 5th DSPS Educators Conference, pp. 23 - 32, Tokyo Institute of Technology, Tokyo, Japan, 17 - 18 Sept. 2003.
- [10] P. G. Bergmann, "Introduction to the Theory of Relativity", Prentice Hall, 1942, and Dover, New York, 1976.
- [11] Sommerfeld, "Vorlesungen über Theoretische Physik", vol. 3, Elektrodynamik, Akademische Verlagsgesellschaft, Leipzig, 1949.
- [12] M. Born, "Die Relativitätstheorie Einsteins", 5th ed., Springer-Verlag, Berlin, 1969. American translation of an earlier German edition: "Einstein's Theory of Relativity", Dover, New York, 1965.
- [13] S. Weinberg, "Gravitation and Cosmology", Wiley, New York, 1972.
- [14] R. Becker and F. Sauter, "Theorie der Elektrizität", vol.1, Teubner, Stuttgart, 1973.
- [15] J. D. Jackson, "Classical Electrodynamics", 2nd ed., J. Wiley, New York, 1975.

- [16] W. Panofsky and M. Phillips, "Classical Electricity and Magnetism", Addison-Wesley, Reading, MA, 1977.
- [17] J. van Bladel, "Relativity and Engineering", Springer-Verlag, Berlin, 1984.
- [18] Kacser, "Relativity, special theory", in Encyclopaedia of Physics, 2nd ed., R. G. Lerner and G. L. Triggs, Eds., pp. 1052 - 1058, Wiley-VCH, New York, 1991.
- [19] H. and M. Ruder, "Die spezielle Relativitätstheorie", Vieweg, Braunschweig, Germany, 1993.
- [20] Fettweis, "The wave-digital method and some of its relativistic implications", IEEE Trans. Circuits and Systems I, vol. 49, No. 6, pp. 862- 868, June 2002, and No. 10, p. 1521, October 2002.
- [21] Fettweis, "Wave-digital concepts and relativity theory", in "Contemporary Issues in Systems Stability and Control with Applications"(Derong Liu and Panos Antsaklis, eds.), pp. 3 - 22, Birkhäuser, Boston, 2003.
- [22] Fettweis, "Nonlinear Kirchhoff circuits and relativity theory", AEÜ Int. J. Electronics (in print).
- [23] K. Meerkötter, Private communication, about 1991.
- [24] K. Meerkötter and T. Felderhoff, "Simulation of nonlinear transmission lines by wave digital filter principles", Proc. IEEE Int. Symp. Circuits and Systems, pp. 875 - 877, San Diego, CA, USA, May 1992.
- [25] N. K. Bose and A. Fettweis, "Isomorphism between skew-symmetric and orthogonal matrices via generalized Cayley Transform", Proc. Eur. Conf. Circuit Theory and Design, vol. I, pp. 133 - 136, Cracow, Poland, 1 - 4 Sept. 2003.
- [26] W. Bertozzi, "Speed and energy of relativistic electrons ", American J. of Physics, vol. 32, pp. 551 - 555, 1964.
- [27] H. Minkowski, "Raum und Zeit", published in "Das Relativitätsprinzip", Teubner, Leipzig, 1913; English version: "Space and Time", in "The Principle of Relativity", Methuen, 1923, and Dover, New York, 1952.

Альфред Феттвейс, др., проф. (Бельгия). Он – выпускник университета (специальность «Электротехника», Université Catholique de Louvain (1946 - 51)). Получил там степень доктора прикладных наук ("Docteur en sciences appliquées") в 1963. Работал в компании «International Telephone and Telegraph Corporation» в Бельгии (1951 - 54, 1956 - 63) и США (1954 - 56); был профессором (специальность «Теория электроэнергии») в университете Technical University of Eindhoven, Нидерланды (1963 - 67), был профессором (специальность «Техника связи») в университете Ruhr-Universität Bochum, Германия (с 1967 г., почетный профессор с 1992 г.), был заслуженным нештатным профессором в университете Notre Dame University, США (1994 - 96). Получил почетные докторские степени Linköping 1986, Mons 1988, Leuven 1988, Budapest 1995 и ряд наград, включая: Prix Acta Technica Belgica 1963, Prix Montefiore 1980, VDE-Ehrenring 1984, ITG Karl-Kupfmüller-Preis 1988, Eduard Rhein Basic Research Award 1993, IEEE Centennial (1984) and Millenium (1999) Medal, и Darlington Award 1980, Technical Achievement Award 1988, Van Valkenburg (Society) Award 2001, (первая) Vitold Belevitch Award 2003 (все 4 – награды общества IEEE Circuits and Systems Society).

Амплитудно-фазовая модуляция лазерного излучения в задачах генерации радиочастотных сигналов

О.Г. Морозов, Г.И. Ильин, Г.А. Морозов

Казанский национальный исследовательский технический университет
им. А.Н. Туполева-КАИ
420111, Российская Федерация, г. Казань, ул. К. Маркса, 10

Аннотация. В статье рассмотрены основные структуры, характеристики, элементная база и направления развития генераторных узлов радиопотонных систем. Это важно с точки зрения определения требований к параметрам внешней амплитудно-фазовой модуляции лазерного излучения как ключевому процессу формирования радиочастотных сигналов в оптическом диапазоне и полигармонического лазерного излучения с составляющими, разностные частоты которых лежат в радиочастотном диапазоне. Методы радиофотоники, рассмотренные в статье, можно классифицировать следующим образом:

фотосмещение двух радиочастотных сигналов (несущих частот) в оптическом диапазоне как основа оптической генерации несущих радиочастот;

внешняя амплитудная или фазовая модуляция лазерного излучения для генерации двух оптических поднесущих с разностной частотой, кратной двум или четырем частотам модулирующих сигналов;

внешняя амплитудно-фазовая модуляция лазерного излучения для генерации двух оптических поднесущих с разностной частотой равной модулирующей радиочастоте.

Дополнительно рассмотрена элементная база для реализации амплитудно-фазовой модуляции в оптическом диапазоне для решения задач генерации несущих радиочастотных сигналов.

Ключевые слова: радиофотоника; оптическое фотосмещение; генерация несущих радиочастот; модуляция лазерного излучения; фотонные радиочастотные генераторы.

1. Введение

Современные методы и средства обработки радиосигналов в оптическом диапазоне длин волн [1-6] применяются в различных радиотехнических и оптико-электронных системах оборонного и гражданского назначения: радиолокационные станции (РЛС) предупреждения, разведки наземных и бортовых средств, радиоэлектронной борьбы и связи по технологиям «радио-по-волокну», «широкополосные сигналы-по-волокну», создания специальных форматов передачи данных с использованием амплитудной и фазовой модуляции в волоконно-оптических магистралях. К задачам передачи, приема и обработки информации в системах радиофотоники относятся [7-9]: генерация несущих радиочастот; радиочастот гетеродинов; полигармонических и многочастотных излучений с разностными частотами, лежащими в радиодиапазоне; их передача по оптическому волокну и преобразование и т.д.

Ключевым элементом систем радиофотоники является подсистема формирования модулированного лазерного излучения [1-6], обеспечивающая как перенос радиосигнала в оптический диапазон (например, с целью измерения мгновенной частоты), так и формирование полигармонического оптического сигнала с разностной частотой, лежащей в радиодиапазоне (например, для генерации непрерывной радиочастоты, равной разностной, после фотосмещения). При этом, как правило, используются либо отдельный набор лазеров, разнесенных на частоту радиосигнала; либо лазеры с внутренней модуляцией или синхронизацией мод; либо лазеры с внешней модуляцией на основе электрооптических модуляторов с рабочим диапазоном

частот от 1 до 100 ГГц амплитудного, фазового, амплитудно-фазового и поляризационного типов для создания фотонных радиочастотных генераторов.

В силу природной симметрии модулированных сигналов и наибольшего достижимого коэффициента модуляционного преобразования особенно широкое применение в системах радиопотоники имеет амплитудно-фазовая модуляция с полным или частичным подавлением опорного одночастотного излучения лазера [10-13]. В статье с точки зрения определения требований к процессам амплитудно-фазовой модуляции и формирования полигармонического лазерного излучения с разностными частотами радиодиапазона рассмотрены основные структуры, характеристики, элементная база и направления развития фотонных радиочастотных генераторов.

С общей точки зрения методы радиопотоники, рассмотренные в статье, классифицированы следующим образом: фотосмещение двух оптических несущих как основа оптической генерации радиочастотных несущих (раздел 2), внешняя амплитудная или фазовая модуляция лазерного излучения для генерации двух оптических «поднесущих» с разностной частотой, кратной двум или четырем модулирующим радиочастотам (раздел 3); внешняя амплитудно-фазовая модуляция лазерного излучения для генерации двух оптических поднесущих с разностной частотой равной модулирующей радиочастоте (раздел 4). Дополнительно рассмотрена элементная база для реализации амплитудно-фазовой модуляции в оптическом диапазоне (раздел 5) для задач генерации радиочастотных сигналов. В заключении обсуждаются достоинства использования непрерывного одночастотного лазера и двухпортового модулятора для комплексной амплитудно-фазовой модуляции оптической несущей частоты и создания на их основе фотонных радиочастотных генераторов.

2. Оптическое фотосмещение как основная процедура генерации несущих радиочастотных сигналов

Перестраиваемые генераторы радиочастотных сигналов с низким уровнем фазовых шумов применяются в различных системах: радиолокационных станциях (РЛС), беспроводной связи, программно-определяемого радио, информационно-измерительных и т.д. [3, 14-15]. Обычно, радиочастотные несущие сигналы генерируются с помощью сложных радиоэлектронных схем с многоступенчатым умножением частоты до достижения необходимого значения. Кроме того, во многих случаях сгенерированный радиочастотный сигнал должен быть распределен к удаленным устройствам. Распределение таких сигналов в электронной области на практике затруднительно из-за больших потерь в линиях распределения, например, в коаксиальном кабеле. Благодаря широкой полосе пропускания и низким потерям современных оптических волокон, распределение радиочастотных сигналов по оптическому волокну является идеальным решением этой задачи и значительно упрощает требования к самим генераторам.

Радиочастотные сигналы могут быть сгенерированы в оптической области с использованием процедуры оптического (фото) смещения, когда две оптические волны с разной длиной волны смешиваются в фотодетекторе. На выходе фотодетектора генерируется сигнал биений с частотой, соответствующей интервалу между длинами двух оптических волн. Предположим, что у нас есть две оптические волны:

$$E_1(t) = E_1 \cos(\omega_1 t + \phi_1), \quad (1)$$

$$E_2(t) = E_2 \cos(\omega_2 t + \phi_2), \quad (2)$$

где E_1, E_2 – значения их амплитуд, ω_1, ω_2 – угловых частот, а ϕ_1, ϕ_2 – фаз.

Принимая ограниченную ширину пропускания фотодетектора, ток на его выходе описываем соотношением:

$$I_{RF}(t) \propto 2A \cos[(\omega_1 - \omega_2)t + (\phi_1 - \phi_2)], \quad (3)$$

где A – амплитуда, определяемая значениями E_1, E_2 и коэффициентом усиления фотодетектора.

Как следует из формулы (3), сгенерирован электрический сигнал с частотой, равной разности частот $(\omega_1 - \omega_2)$ двух оптических волн (1)-(2). Значения данной частоты могут достигать терагерцевого диапазона, ограничением может быть только ширина полосы пропускания фотодетектора. Однако использование двух несинхронизированных лазерных диодов приводит к тому, что полученный радиочастотный сигнал будет обладать высоким уровнем фазового шума, поскольку фазы двух волн ϕ_1 и ϕ_2 не коррелированы.

Несколько предложенных методик [16-24] позволяют генерировать из двух оптических волн, синхронизированных по фазе, радиочастотные сигналы с низким уровнем фазового шума. К ним относятся: методика синхронизации с захватом частоты [16, 17], методика фазовой автоподстройки частоты (ФАПЧ) [18, 19] и комбинированная методика [20]. В ряде работ представлены генераторы на основе двухчастотных лазеров [21, 22]. Особого рассмотрения заслуживают методики, основанные на внешней амплитудно-фазовой модуляции одночастотного лазерного излучения [23-24].

3. Генерация радиочастотных несущих с помощью внешней амплитудной и фазовой модуляции лазерного излучения

Радиочастотные несущие также можно генерировать посредством внешней модуляции [10-15, 23-26]. Представим структурные схемы генераторов при работе в «минимальной» (рис.1,а) и «максимальной» (рис.1,б) точках модуляционной характеристики модулятора Маха-Цендера (ММЦ) [14, 15]. Электрический сигнал с удвоенной частотой $2f_m$ был оптически сгенерирован с помощью электрического смещения ММЦ $V_B=0$ в «минимальную» рабочую точку [14]. Сигнал миллиметрового диапазона 36 ГГц был получен при модулирующей частоте $f_m = 18$ ГГц.

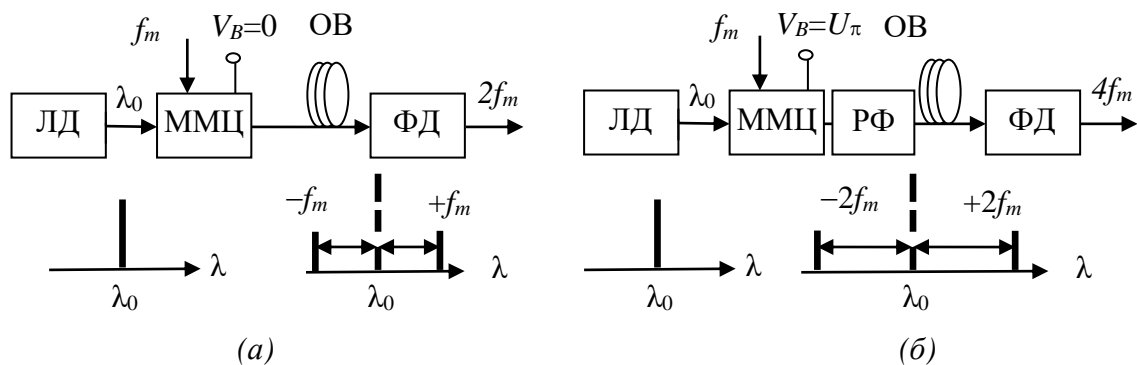


Рис. 1. Генератор радиочастотного сигнала с внешней модуляцией лазерного диода и смещением ММЦ в «минимальную» (а) и «максимальную» (б) рабочие точки: ЛД – лазерный диод; ФД – фотодетектор; РФ – режекторный фильтр

В [15] для генерации частоты в четыре раза большей, чем частота модулирующего сигнала, равная $4f_m$, использовался модулятор Маха-Цендера, смещенный в «максимальную» точку его модуляционной характеристики $V_B = U_\pi$, где U_π – полуволновое напряжение модулятора Маха-Цендера. Оптическая несущая с длиной волны λ_0 , а также гармонические составляющие 1-го и 3-го порядков были подавлены с помощью регулировки уровня модулирующего сигнала и режекторных фильтров. В качестве режекторных фильтров использовались перестраиваемые интерферометры Маха-Цендера и Фабри-Перо. Сигнал миллиметрового диапазона 60 ГГц был получен при модулирующей частоте $f_m = 15$ ГГц. Данные эффекты были получены в середине 90-х годов. Для сантиметрового диапазона радиочастот указанные эффекты получены нами в середине 80-х годов прошлого века. Сведения об этих работах приведены в [10-13].

Для широкополосных обзорных радиолокационных станций (РЛС), систем с растяжкой сигнала или программно-определяемых систем радиосвязи необходимо использовать радиочастотный сигнал с перестройкой частоты. В [23, 24] были продемонстрированы два подхода для генерации перестраиваемого по частоте радиочастотного сигнала с использованием оптического фильтра зафиксированного по длине волны: на основе амплитудного [23] и фазового [24] модулятора. Подавление оптической несущей осуществляется с помощью волоконной решетки Брэгга (ВРБ).

Основными причинами наличия высших составляющих в выходном спектре амплитудного модулятора являются его нелинейная характеристика, температурная нестабильность, отклонение параметров преобразования от оптимальных и погрешность юстировки поляризаторов преобразователя. Все они вызывают отклонение рабочей точки и быстрый рост амплитуд паразитных спектральных составляющих. Ключевым преимуществом использования оптического фазового модулятора является отсутствие необходимости включения постоянного смещения, что исключает проблему дрейфа положения рабочей точки [24]. В отличие от модулятора Маха-Цендера, которым можно управлять для выбора четных или нечетных гармонических составляющих, использование модулятора оптической фазы позволит генерировать их полный набор, включая оптическую несущую. Биения боковых гармонических составляющих, оставшихся на выходе фотодетектора после фильтрации в волоконной решетке Брэгга оптической несущей, позволят сгенерировать радиочастотный сигнал.

Необходимо отметить, что высшие оптические гармоники, переданные через одномодовое волокно, будут подвержены хроматической дисперсии, которая будет изменять соотношение фаз в боковых составляющих. Для сохранения соотношения одинаковых фаз необходима компенсация дисперсии. В условиях компенсации экспериментально было продемонстрировано, что при модулирующей частоте, лежащей в диапазоне (18,8÷25) ГГц, локально и дистанционно были получены две полосы сигналов миллиметрового диапазона (37,6÷50) ГГц и (75,2÷100) ГГц.

4. Генерация радиочастотных несущих с частотой, равной модулирующей (амплитудно-фазовая модуляция)

Рассмотрим переход от амплитудно-модулированного сигнала со стопроцентной модуляцией к сигналу биений. Это переход можно осуществить либо за счет подавления несущего колебания, либо за счет переключения фазы амплитудно-модулированного сигнала в момент, когда его огибающая достигает нулевого значения. Спектры исходного квазигармонического колебания со структурой амплитудно-модулированного сигнала (а) и спектры преобразованных колебаний со структурой

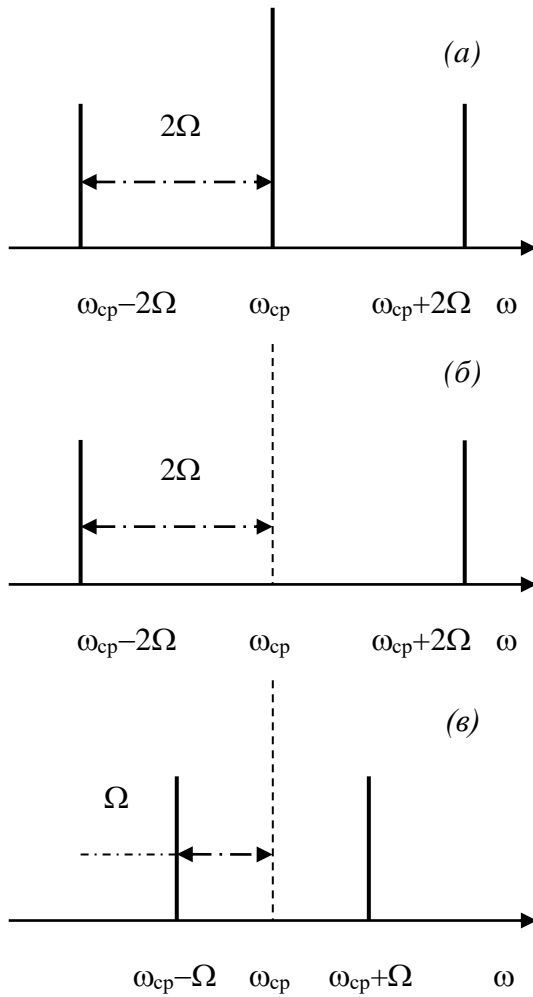


Рис. 2. Спектры амплитудно-модулированного сигнала до преобразования (а) и после преобразования по первому варианту (б) и по второму (в) варианту

получим $z=\pi/2$, $J_1(z)=0,64$, $J_3(z)=0,06$. Одночастотный входной и двухчастотный выходной спектры излучения генератора, а также радиочастотный сигнал на выходе фотодетектора показаны на рис. 3, а-в соответственно.

Максимальный коэффициент преобразования [15] равен 0,64. При коэффициенте преобразования 0,58 коэффициент нелинейных искажений не превышает 1%. Здесь еще два важных момента, которые отметим: это – равенство амплитуд спектральных составляющих и противоположность их фаз (что важно, например, для генерации солитонов [39] и построения фотонных фильтров радиочастотных сигналов [29, 40-41], вне зависимости от расположения рабочей точки) и простота перестройки частоты, что объясняется использованием одного модулирующего сигнала.

Возможность возникновения четных составляющих в спектре выходного излучения, показанная в [16], может быть полезна при генерации последовательностей импульсов с возвратом к нулю (RZ, return-to-zero) и с возвратом к нулю с подавлением постоянной составляющей (CSRZ, carrier-suppressed return-to-zero) [40]:

сигнала биений, полученных по первому (б) и по второму (в) вариантам, представлены на рис. 2.

И в этом случае полученные спектры существенно отличаются по разностным частотам 2Ω (рис. 2,б) и Ω (рис. 2,в) в зависимости от способа преобразования.

Рассмотрим первый вариант реализации. Генератор радиочастотного сигнала сантиметрового диапазона был построен нами на принципах амплитудно-фазовой модуляции в классическом амплитудном электрооптическом модуляторе (АЭОМ) на кристалле метаниобата лития LiNbO_3 , класс симметрии $3m$ [10÷13, 25÷26]. Соединение амплитудного электрооптического модулятора с лазерным диодом и фотодетектором осуществлялось с помощью волоконно-оптических шнуров (пигтейлов). При работе генератора в «нулевой» точке модуляционной характеристики выходной спектр описывается выражением:

$$E_{\text{вых}} = -jE_0 e^{j\omega_0 t} \left[2 \sum_{k=0}^{\infty} J_{2k+1}(z) \sin(2k+1)\Omega t \right], \quad (4)$$

где ω_0 – угловая частота оптического излучения, а Ω – угловая частота модулирующего излучения; индекс j в данном выражении указывает на то, что выходное излучение ортогонально вектору поляризации исходного излучения; $J_{2k+1}(z)$ – функция Бесселя $(2k+1)$ -го порядка. При полуволновом напряжении $U_m = U_{\lambda/2}$

$$e(t) = -j \frac{\sqrt{2}}{2} e^{j\omega_0 t} \left\{ \sum_{k=0}^{\infty} J_0(z) + 2 \sum_{k=1}^{\infty} J_{2k}(z) \cos 2k\Omega t + 2 \sum_{k=0}^{\infty} J_{2k+1}(z) \sin(2k+1)\Omega t \right\}. \quad (5)$$

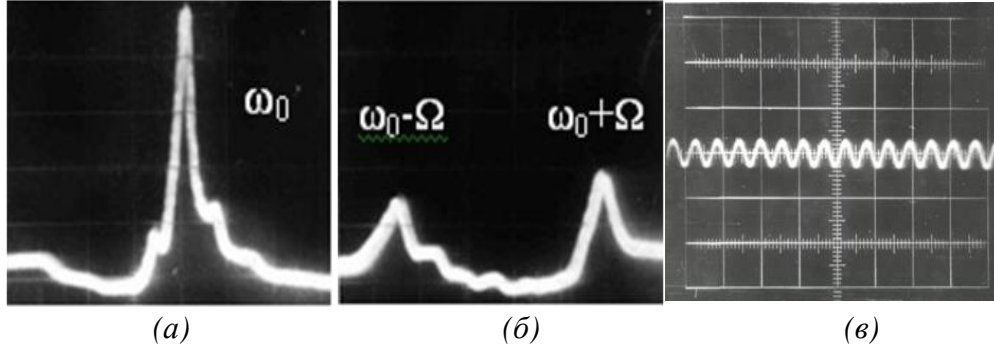


Рис. 3. Спектры входного (а) и выходного (б) излучения амплитудного электрооптического модулятора; осциллограмма радиочастотного сигнала (в) на выходе фотодетектора (рис. 1, а)

Теоретические результаты анализа спектра на выходе предлагаемого генератора приведены в табл. 1. Они получены в предположении, что модулятор смещен при нулевом приложенном напряжении $\Gamma_- = 0$, а поляризаторы либо ортогональны $\alpha = -\beta = 45^\circ$, либо параллельны $\alpha = \beta = 45^\circ$.

Таблица 1. Выходные спектральные характеристики АЭОМ

α, β, Γ_-	Спектр излучения на выходе АЭОМ
$\alpha = 45^\circ, \beta = -45^\circ, \Gamma_- = 0$	$E_{\text{вых}} = -j E e^{j\omega_0 t} \left[2 \sum_{k=0}^{\infty} J_{2k+1}(z) \sin(2k+1)\Omega t \right].$
$\alpha = 45^\circ, \beta = 45^\circ, \Gamma_- = 0$	$E_{\text{вых}}(t) = E e^{j\omega_0 t} \left\{ J_0(z) + 2 \sum_{k=1}^{\infty} J_{2k}(z) \cos 2k\Omega t \right\}.$

С выхода амплитудного электрооптического модулятора излучение поступает на поляризационный светоделитель, у которого одна из главных осей ориентирована под углом 45° к электрооптическому модулятору. На двух выходах поляризационного светоделителя получим два оптических сигнала по двум главным осям, причем один сигнал состоит из всех оптических частот четного порядка (первая строка табл. 1), а другой – из всех оптических частот нечетного (вторая строка табл. 1). Это позволит получить как RZ, так и CSRZ последовательности импульсов соответственно в диапазоне $(0,3 \div 1)$ ГГц, определяемом полосой пропускания амплитудного электрооптического модулятора [40]. Вся система эквивалентна генератору импульсов, который состоит из фазового модулятора и взаимодействующей пары гребенчатых фильтров для выборки четных и нечетных боковых полос. Главное преимущество предлагаемого метода заключается в том, что для разделения последовательностей не используется физическая пара взаимодействующих гребенчатых фильтров, что делает регулировку частоты повторения значительно проще.

Вторая версия источника двухчастотного излучения была смоделирована на базе двухпортового (ДП) модулятора Маха-Цендера [29]. При моделировании радиочастотный сигнал с частотой 10 ГГц подавался на оба порта двухпортового модулятора Маха-Цендера. Напряжение смещения изменяет фазовую задержку между

интерферирующими оптическими волнами на выходе двухпортового модулятора Маха-Цендера и таким образом меняет интенсивность выходного сигнала. Максимум передаточной функции соответствует случаю, когда разность фаз кратна 2π радиан (конструктивная интерференция), а минимум пропускания имеет место, когда разность фаз кратна π радиан (деструктивная интерференция).

Рассмотрим второй вариант реализации. Простейшим колебанием, удовлетворяющим этим требованиям, является колебание вида $S_1(t) = S_1 \cos(2\Omega t + \pi)$, где S_1 – его постоянная амплитуда, а π – начальная фаза. При этом фазу несущего колебания ω_0 будем коммутировать на π при прохождении модулирующего колебания минимума. В случае амплитудной модуляции сигналом $S_1(t)$ колебания коммутируемой несущей получим результирующее колебание со спектром следующего вида

$$e(t) = \frac{2E_0}{\pi} \sum_n \left\{ \left[\frac{1}{n} - \frac{m}{2} \left(\frac{1}{n-2} + \frac{1}{n+2} \right) \right] [\cos(\omega_0 + n\Omega)t - \cos(\omega_0 - n\Omega)t] \right\}. \quad (6)$$

где m – коэффициент амплитудной модуляции.

Из анализа (6) видно, что первое слагаемое в больших квадратных скобках определяет спектр коэффициентов ряда Фурье коммутируемой несущей, а второе и третье описывают подавляющее воздействие на его составляющие модулирующим колебанием. Степень подавления зависит от коэффициента модуляции m . Приняв $E_3 = 0$, получим, что оптимальный коэффициент модуляции $m_{\text{opt}} = 5/9$, при этом результирующее колебание практически двухчастотно ($E_1 = 0,76E_0$), так как амплитуда спектральных составляющих $E_n \leq E_1/15$ для $n \geq 5$. При изменении коэффициента модуляции в пределах $(0,85 - 1,15)$ m_{opt} коэффициент нелинейных искажений выходного колебания не будет превышать 1%.

Полного подавления боковых составляющих с $n \geq 3$ можно добиться при использовании для амплитудной модуляции сигнала вида $S_2(t) = S_0 |\sin \Omega t|$. Тогда результирующее колебание будет иметь следующий спектр

$$e(t) = \frac{2E_0}{\pi} (1-b) \sum_n \frac{1}{n} \{ \cos(\omega_0 + n\Omega)t - \cos(\omega_0 - n\Omega)t \} + \frac{\pi E_0 b}{4} \{ \cos(\omega_0 + \Omega)t - \cos(\omega_0 - \Omega)t \} \quad (7)$$

где b – коэффициент амплитудной модуляции.

Амплитуда спектральных составляющих будет определяться коэффициентами ряда Фурье и для $n=1$ $E_1 = [2E_0/\pi][1-b] + [\pi E_0 b/4]$, а для $n \geq 3$ $E_n = [2E_0/\pi n][1-b]$. При $b_{\text{opt}}=1$ спектр содержит две полезных составляющих на частотах $\omega_0 + \Omega$ и $\omega_0 - \Omega$, паразитные составляющие полностью подавлены. При изменении коэффициента модуляции в пределах $(0,7-1)b_{\text{opt}}$ коэффициент нелинейных искажений выходного колебания не будет превышать 1%.

Двухчастотный спектр может быть получен полной реализацией метода Ильина-Морозова [10-13] с помощью амплитудного модулятора Маха-Цендера (АММЦ), работающего на линейном участке рабочей характеристики, и фазового модулятора Маха-Цендера (ФММЦ). При этом для управления фазовым модулятором Маха-Цендера используется блок преобразования ГРЧ А – ГРЧ Ф (ГРЧ – генератор разностной частоты) синусоидального сигнала с частотой Ω в меандр с частотой $\Omega/2$ для фазовой коммутации с соответствующими амплитудами управляющих напряжений U , который представлен на рис. 4 и ранее нами не рассматривался. Волоконная решетка

Брэгга используется для дополнительного подавления оптической несущей источника лазерного излучения.

Спектр излучения для двух составляющих на выходе модуляторов в данном случае описывается выражением:

$$E_{\text{АФМ}}(t) = 0,49 E_0 \{ \sin(\omega + \Omega/2)t - \sin(\omega - \Omega/2)t \} + 0,007 E_0 \{ \sin(\omega + 3\Omega/2)t - \sin(\omega - 3\Omega/2)t \} + \dots \quad (8)$$

Как видно из (8) разностная частота между составляющими двухчастотного излучения Ω равна частоте модулирующего сигнала. Составляющие высших гармоник не учитываем в силу малости их амплитуд.

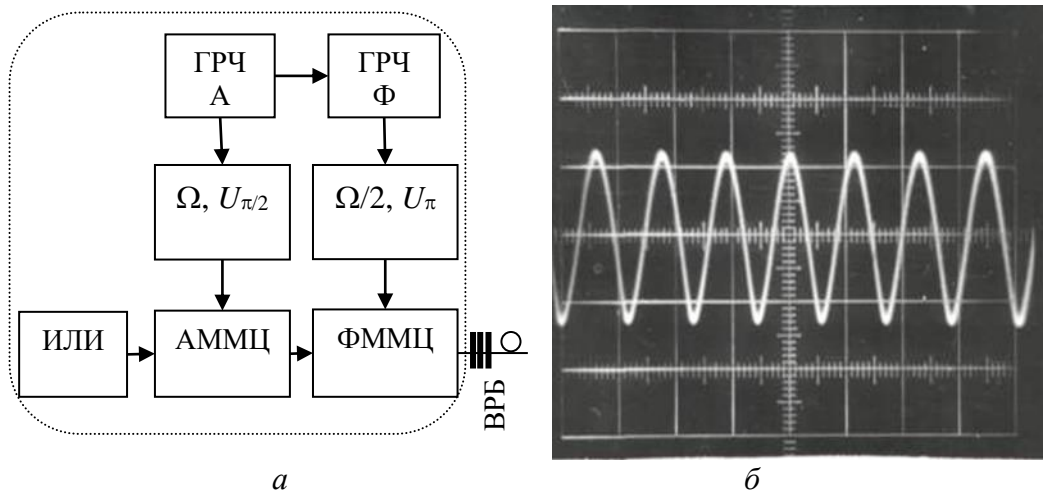


Рис. 4. Генератор двухчастотного симметричного излучения (а); колебание несущей на выходе фотодетектора (б) (здесь частота несущей в два раза меньше показанной на рис. 3, в)

Получено сужение разностной частоты в два раза по сравнению с классическими схемами ее удвоения, применяемыми в практике известных радиофотонных систем, например, при использовании одного амплитудного модулятора Маха-Цендера, работающего для подавления несущей в «нулевой» точке модуляционной характеристики, как указывалось в (4) и (6).

5. Электрооптические модуляторы для фотонных радиочастотных генераторов

В приложениях радиофотоники в основном используются двухканальные электрооптические модуляторы на основе интерферометра Маха-Цендера, реализующие амплитудную, фазовую и амплитудно-фазовую модуляцию. При приложении электрического поля к одному из двух параллельных электродов в первом оптическом канале происходит изменение коэффициента преломления и возникает разница фаз относительно оптической волны, распространяющейся во втором канале, которая определяет интенсивность и фазу сигнала на выходе модулятора.

В настоящее время активно используются электрооптические модуляторы как на основе ниобата лития LiNbO_3 , так и на основе арсенида галлия GaAs и фосфида индия InP , оптимизированные для работы в диапазонах длин волн 1,3–1,6 мкм. В случае использования подложек InP имеется возможность создания фотонных интегральных схем, включающих в себя лазеры, модуляторы, мультиплексоры по длине волны, волноводы и управляющую электронику. Для получения эффективной модуляции в

компактном интегрированном модуляторе необходимо обеспечить наибольшее изменение коэффициента преломления в модуляторной секции, при незначительном увеличении ее коэффициента поглощения и при приложении возможно меньшего напряжения к модулятору.

Основными характеристиками $LiNbO_3$ модуляторов являются полуволновое напряжение, рабочая полоса частот, вносимые оптические потери, стабильность по температуре и положению рабочей точки, оптическая чувствительность по коэффициенту преломления. Основное их назначение – работа в полосе частот до 40-60 ГГц. Технология модуляторов указанного класса наиболее приспособлена к работе с сигналами сверхвысокочастотного (СВЧ) и крайне высокочастотного (КВЧ) диапазонов. Однако существует и ряд недостатков, связанных с появлением фазовых искажений и неравномерности рабочей характеристики около верхней границы частотного диапазона. Основными фирмами – производителями $LiNbO_3$ модуляторов являются *JDSU*, *Avanex*, *Sumitomo*, *Photline* и *EOSPACE*.

Арсенид галлия $GaAs$ и фосфид индия InP в силу своих оптических свойств и высокого потенциала для интеграции в компактные многофункциональные устройства, а также в силу отсутствия симметрии в кристаллической структуре, играют важную роль в области расширения частотного диапазона обрабатываемых сигналов и создания перспективных систем радиофотоники. Условно их частотный диапазон можно оценить как до 100 ГГц для $GaAs$ и свыше 100 ГГц для InP . Несмотря на относительно малое значение электрооптического коэффициента ($r_{41}=1,4$ пм/В) и высокий коэффициент преломления (до 3,4), обеспечение 100%-го по объему кристалла взаимодействия оптического и электрического поля обеспечивает высокую электрооптическую эффективность полупроводников $GaAs$ и InP , сравнимой с достижимой в $LiNbO_3$. Однако полученная при этом характеристика «напряжение–длина» ($U_\pi L_\pi$) составляет в среднем 9 В·см, при предполагаемом 5,5 В·см, что позволяет делать электрооптические модуляторы меньшей длины и меньших размеров, чем модуляторы на основе $LiNbO_3$. Основным поставщиком электрооптических модуляторов на основе арсенида галлия $GaAs$ и фосфида индия InP является фирма *Bookham*.

Особое место в системах радиофотоники занимает поляризационный модулятор фирмы *Versawave Technologies* [42]. Поляризационный модулятор представляет собой специальный фазовый модулятор, который поддерживает обе моды ТМ и ТЕ, обеспечивая модуляцию фазы в каждой из них, но с противоположным знаком. Такая модуляция достигается при настройке контроллером поляризации вектора поляризации исходного излучения таким образом, чтобы его положение составляло 45° с одной из принципиальных осей поляризационного модулятора.

Нормированное оптическое поле на выходе поляризационного модулятора может быть описано следующим образом:

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} \propto \frac{\sqrt{2}}{2} \begin{bmatrix} \exp(j\omega_0 t + j\beta_1 \cos \omega_{RF} t) \\ \exp(j\omega_0 t - j\beta_1 \cos \omega_{RF} t) \end{bmatrix}, \quad (9)$$

где β_1 – индекс фазовой модуляции в поляризационном модуляторе, который определяется как $\beta_1 = \pi U_p / U_\pi$ при соответствующих значениях амплитуды модулирующего напряжения U_p и полуволнового напряжения U_π . Контроллер поляризации на выходе поляризационного модулятора позволяет либо объединить обе составляющие, либо выделить каждую из них.

Поляризационный модулятор изготавливается по технологии полупроводников $GaAs$, обладает коэффициентом контрастности 20, вносимыми оптическими потерями 3,5 дБ

и полуволновым напряжением до 3,5 В в диапазоне частот от 40 до 80 ГГц. При использовании модулятора указанного класса достаточно просто реализуются все виды модуляции, включая амплитудно-фазовую, полученную нами на кристаллах $LiNbO_3$ и представленную в работах [10–13, 43].

6. Заключение

Обзорные исследования, представленные в настоящей статье, показали, что развитие радиофотонных систем сдерживается сложностью формирования стабильных по частоте и синхронизированных по фазе излучений источника зондирования и гетеродина с возможностью их перестройки по частоте в общем случае по произвольному закону и с произвольной скоростью, что определяется необходимостью приема сигналов или измерений с требуемой чувствительностью и отношением сигнал/шум в широком диапазоне частот.

Указанные проблемы могут быть решены построением радиофотонных систем на основе одночастотных непрерывных лазеров и амплитудно-фазовых электрооптических модуляторов (возможно использование поляризационных модуляторов).

В основе решения поставленных задач лежит использование уникального метода [10–13], который получил название метода Ильина-Морозова, предложенного авторами в 1983 году и основанного на модуляции амплитуды и манипуляции фазы исходного одночастотного излучения для синтеза симметричных двухчастотных и полигармонических излучений. Выходное полигармоническое излучение при реализации предлагаемого метода Ильина-Морозова содержит в своей структуре симметричные составляющие равной амплитуды и альтернативной фазы с разностной частотой. Эти составляющие симметричны относительно исходной подавленной частоты. При этом разностная частота равна частоте модулирующего радиосигнала или ее удвоенному значению, и характеризуется высокой степенью спектральной чистоты при отклонениях параметров преобразования от оптимальных.

Разработка метода была осуществлена на основе детального исследования модуляционных характеристик амплитудного электрооптического модулятора по напряженности электрического поля, в отличие от широко использовавшегося в то время анализа модуляционных характеристик по его интенсивности. Было показано, что при работе амплитудного электрооптического модулятора в «нулевой (минимальной)» точке модуляционной характеристики выходное излучение является не амплитудно-модулированным с боковыми частотами на удвоенной частоте модуляции, а двухчастотным с подавленной несущей частотой и составляющими, разнесенными на удвоенную частоту модуляции [13]. Наиболее простой реализацией на уровне современных технологий является реализация на основе двухпортового модулятора Маха-Цендера.

В настоящей статье с использованием работ авторов [44–53] и их коллег [54–56] были показаны возможности генерации радиочастотных сигналов с низким уровнем фазовых шумов и генерации полигармонических излучений с радиочастотными компонентами. Кроме того, исследованы возможности применения технологий на основе ниобата лития, арсенида галлия и фосфида индия для создания элементов перспективных модуляционных систем радиофотоники в различных радиочастотных диапазонах.

Благодарности

Работа выполнена при финансовой поддержке Министерства образования и науки Российской Федерации в рамках государственного задания на оказание услуг

(выполнение работ) по организации научных исследований, выполняемых Казанским национальным исследовательским техническим университетом на кафедрах телевидения и мультимедийных систем, радиоэлектронных и квантовых устройств и в научно-исследовательском институте прикладной электродинамики, фотоники и живых систем (программы «Радиофотоника» 3.1962.2014/К и «Развитие» 3.2426.2014/К).

Список аббревиатур

РЛС – радиолокационная станция	АЭОМ – амплитудный электрооптический модулятор
ФАПЧ – фазовая автоподстройка частота	RZ (return-to-zero) – с возвратом к нулю
ММЦ – модулятора Маха-Цендера	CSRZ (carrier-suppressed return-to-zero) – с возвратом к нулю с подавлением постоянной составляющей
ЛД – лазерный диод	ГРЧ – генератор разностной частоты
ФД – фотодетектор	СВЧ – сверхвысокочастотный
РФ – режекторный фильтр	КВЧ – крайне высокочастотный
ВРБ – волоконная решетка Брэгга	
АММЦ – амплитудный модулятор Маха-Цендера	
ФММЦ – фазовый модулятор Маха-Цендера	

Литература

1. A. Seeds, "Microwave photonics," IEEE Trans. Microw. Theory Tech. 50(3), 877-887 (2002).
2. J. Capmany, D. Novak, "Microwave photonics combines two worlds," Nat. Photonics 1(6), 319-330 (2007).
3. J. Yao, "Microwave photonics," J. Lightwave Technol. 27(3), 314-335 (2009).
4. T. Berceli, P.R. Herzfeld, "Microwave photonics – a historical perspective," IEEE Trans. Microw. Theory Tech. 58 (11), 2992-3000 (2010).
5. J. Capmany, et al., "Microwave photonic signal processing," J. Lightwave Technol. 31(4), 571-586 (2013).
6. R.A. Minasian, E.H.W. Chan, X. Yi, "Microwave photonic signal processing," Opt. Express 21(19), 22918-22936 (2013).
7. L. Maleki, "The optoelectronic oscillator," Nat. Photonics 5(12), 728-730 (2011).
8. C.H. Lee, "Optical generation and control of microwave and millimeter-wave signals," IEEE MTT-S International Microwave Symposium Digest 2, 811-814 (1987).
9. P.R. Herzfeld, et al., "Indirect subharmonic optical injection locking of a millimeter-wave IMPATT oscillator," IEEE Trans. Microw. Theory Tech. 34(12), 1371-1376 (1986).
10. G.I. Il'in, O.G. Morozov, A.G. Il'in, "Theory of symmetrical two-frequency signals and key aspects of its application," Proc. of SPIE 9156, 91560M (2014).
11. O.G. Morozov, D.L. Aybatov, "Spectrum conversion investigation in lithium niobate Mach-Zehnder modulator," Proc. of SPIE 7523, 75230D (2010).
12. O.G. Morozov, et al., "Methodology of symmetric double frequency reflectometry for selective fiber optic structures," Proc. of SPIE 7026, 70260I (2008).
13. O.G. Morozov, et al., "Metrological aspects of symmetric double frequency and multi frequency reflectometry for fiber Bragg structures," Proc. of SPIE, 7026, 70260J (2008).
14. J.J. O'Reilly, et al., "Optical generation of very narrow line width millimeter wave signals," Electron. Lett. 28(25), 2309-2311 (1992).
15. J.J. O'Reilly, P.M. Lane, "Fiber-supported optical generation and delivery of 60 GHz signals," Electron. Lett. 30(16), 1329-1330 (1994).

16. L. Goldberg, et al., "Microwave signal generation with injection locked laser diodes," *Electron. Lett.* 19(13), 491-493 (1983).
17. L. Goldberg, et al., "35 GHz microwave signal generation with injection locked laser diode," *Electron Lett.* 21(18), 714-715 (1985).
18. R.T. Ramos, A.J. Seeds, "Fast heterodyne optical phase-lock loop using double quantum well laser diodes," *Electron. Lett.* 28(1), 82-83 (1992).
19. H. Rideout, et al., "Discriminator-aided optical phase-lock loop incorporating a frequency down-conversion module," *IEEE Photon. Technol. Lett.* 18(22), 2344-2346 (2006).
20. A.C. Bordonalli, C. Walton, A.J. Seeds, "High-performance phase locking of wide line width semiconductor lasers by combined use of optical injection locking and optical phase-lock loop," *J. Lightw. Technol.* 17(2), 328-342 (1999).
21. X. Chen, Z. Deng, J.P. Yao, "Photonic generation of microwave signal using a dual-wavelength single-longitudinal-mode fiber ring laser," *IEEE Trans. Microw. Theory Tech.* 54(2), 804-809 (2006).
22. X. Chen, Z. Deng, J.P. Yao, "Ultranarrow dual-transmission-band fiber Bragg grating filter and its application in a dual-wavelength single-longitudinal-mode fiber ring laser," *Opt. Lett.* 30(16), 2068-2070 (2005).
23. G. Qi, et al., "Generation and distribution of a wide-band continuously tunable mm-wave signal with an optical external modulation technique," *IEEE Trans. Microw. Theory Tech.* 53(10), 3090-3097 (2005).
24. G. Qi, et al., "Optical generation and distribution of continuously tunable millimeter-wave signals using an optical phase modulator," *J. Lightw. Technol.* 23(9), 2687-2695 (2005).
25. O.G. Natanson, et al., "Development problems of frequency reflectometry for monitoring systems of optical fiber structures," *Proc. of SPIE* 5854, 215-223 (2005).
26. O. G. Natanson, et al., "Reflectometry in open and fiber mediums: technology transfer," *Proc. of SPIE* 5854, 205-214 (2005).
27. O.G. Morozov, et al., "Structural minimization of fiber optic sensor nets for monitoring of dangerous materials storage," *Proc. of SPIE* 7992, 79920E (2011).
28. O.G. Morozov, et al., "Characterization of stimulated Mandelstam-Brillouin scattering spectrum using a double-frequency probing radiation," *Proc. of SPIE* 8787, 878709 (2013).
29. O.G. Morozov, T.S. Sadeev, "All-optical microwave photonic filter based on two-frequency optical source," *Proc. of SPIE* 7992, 79920C (2011).
30. Diddams, S.A., "The evolving optical frequency comb," *J. Opt. Soc. Amer. B.* 27(11), B51-B62 (2010).
31. F.Z. Zhang, et al., "Flat optical frequency comb generation and its application for optical waveform generation," *Opt. Commun.* 290(1), 37-42 (2013).
32. C. He, et al., "Ultraflat optical frequency comb generated based on cascaded polarization modulators," *Opt. Lett.* 37(18), 3834-3836 (2012).
33. M. Fujiwara, et al., "Optical carrier supply module using flattened optical multicarrier generation based on sinusoidal amplitude and phase hybrid modulation," *J. Lightwave Technol.* 21(11), 2705 (2003).
34. Y.J. Dou, H.M. Zhang, M.Y. Yao, "Generation of flat optical-frequency comb using cascaded intensity and phase modulators," *Opt. Lett.* 36, 2749-2751 (2011).
35. R. Wu, et al., "Generation of very flat optical frequency combs from continuous-wave lasers using cascaded intensity and phase modulators driven by tailored radio frequency waveforms," *Opt. Lett.* 35(19), 3234-3236 (2010).
36. T. Yamamoto, et al., "Multicarrier light source with flattened spectrum using phase modulators and dispersion medium," *J. Lightwave Technol.* 27(19), 4297-4305 (2009).

37. C. Chen, et al., "Generation of a flat optical frequency comb based on a cascaded polarization modulator and phase modulator," *Opt. Lett.* 38(16), 3137-3140 (2013).
38. T.S. Sadeev, O.G. Morozov, "Investigation and analysis of electro-optical devices in implementation of microwave photonic filters," *Proc. of SPIE* 8410, 841007 (2012).
39. E.M. Dianov, et al., "Generation of a train of fundamental solitons at a high repetition rate in optical fibers," *Opt. Lett.* 14(18), 1008-1010 (1989).
40. O.G. Morozov, "RZ, CS-RZ and soliton generation for access networks applications: problems and variants of decisions," *Proc. of SPIE* 8410, 84100P (2012).
41. D.L. Aybatov, O.G. Morozov, T.S. Sadeev, "Dual port MZM based optical comb generator for all-optical microwave photonic devices," *Proc. of SPIE* 7992, 799202 (2011).
42. X. Zou, J. Yao, "Repetition-rate-tunable return-to-zero and carrier-suppressed return-to-zero optical pulse train generation using a polarization modulator," *Opt. Lett.* 34(3), 313-315 (2009).
43. O.G. Morozov, et al., "Synthesis of two-frequency symmetrical radiation and its application in fiber optical structures monitoring," *InTech*, 2012, <http://www.intechopen.com/books/fiber-optic-sensors/synthesis-of-two-frequency-symmetrical-radiation-and-its-application-in-fiber-optical-structures-monitoring> (10 December 2015).
44. O.G. Morozov, D.L. Aibatov, "Two-frequency scanning of FBG with arbitrary reflection spectrum," *Proc. of SPIE* 6605, 660506 (2007).
45. O.A. Stepustchenko, et al., "Optical refractometric FBG biosensors: problems of development and decision courses," *Proc. of SPIE* 7992, 79920D (2011).
46. O.G. Morozov, et al., "Instantaneous frequency measurement of microwave signals in optical range using "frequency-amplitude" conversion in the π -phase-shifted fiber Bragg grating," *Proc. of SPIE* 9136, 91361B (2014).
47. O.G. Morozov, et al., "Instantaneous microwave frequency measurement with monitoring of system temperature," *Proc. of SPIE* 9156, 91560N (2014).
48. O.G. Morozov, et al., "Instantaneous frequency measurement using double-frequency probing," *Proc. of SPIE* 8787, 878708 (2013).
49. O.G. Morozov, Y.E. Pol'ski, "Perspectives of fiber sensors based on optical reflectometry for nondestructive evaluation," *Proc. of SPIE* 2944, 178-183 (1996).
50. O.G. Morozov, et al., "Two-frequency analysis of fiber-optic structures," *Proc. of SPIE* 6277, 62770E (2006).
51. D.L. Aybatov, et al., "Distributed temperature fiber Bragg grating sensor," *Proc. of SPIE* 7374, 73740B (2009).
52. O.G. Morozov, et al., "Double mode system for FWM reducing," *Proc. of SPIE* 7026, 70260B (2008).
53. O.G. Morozov, et al., "Poly-harmonic Analysis of Raman and Mandelstam-Brillouin Scatterings and Bragg Reflection Spectra," *InTech*, 2015, <http://www.intechopen.com/books/advances-in-optical-fiber-technology-fundamental-optical-phenomena-and-applications/poly-harmonic-analysis-of-raman-and-mandelstam-brillouin-scatterings-and-bragg-reflection-spectra> (10 December 2015).
54. V.A. Andreev, et al., "Investigation of cross-phase modulation in multichannel dispersion managed soliton systems," *Proc. of SPIE* 7374, 737405-8 (2009).
55. A.Kh. Sultanov, et al., "Measuring of chromatic and polarization mode dispersion in optical fiber," *Proc. of SPIE* 4589, 192 (2011).
56. V.A. Burdin, M.V. Dashkov, "Polarization optical time domain reflectometer with linear extension of pulse width," *Proc. of SPIE* 7523, 75230K (2010).

Олег Геннадьевич Морозов, доктор технических наук, профессор (заведующий кафедрой радиофотоники и микроволновых технологий, директор научно-

исследовательского института прикладной электродинамики, фотоники и живых систем ФГБОУ ВО «Казанский национальный исследовательский технический университет им А.Н.Туполева-КАИ»). Область научных интересов – информационно-измерительные и телекоммуникационные системы оптического и микроволнового диапазонов, количество опубликованных работ – свыше 240.

E-mail: OGMorozov@kai.ru.

Герман Иванович Ильин, Заслуженный деятель науки Российской Федерации, Заслуженный деятель науки и техники Республики Татарстан, доктор технических наук, профессор (заведующий кафедрой радиоэлектронных и квантовых устройств ФГБОУ ВО «Казанский национальный исследовательский технический университет им А.Н.Туполева-КАИ»). Область научных интересов – лазерные и радиотехнические системы, теория потенциальной помехоустойчивости. Автор свыше 150 публикаций.

E-mail: GIILin@kai.ru

Геннадий Александрович Морозов, Заслуженный работник высшей школы Российской Федерации, Заслуженный деятель науки и техники Республики Татарстан, доктор технических наук, профессор (профессор кафедры радиоэлектронных и телекоммуникационных систем ФГБОУ ВО «Казанский национальный исследовательский технический университет им А.Н.Туполева-КАИ»). Область научных интересов – антенны и микроволновые технологии, биотехнологии. Автор свыше 150 публикаций.

E-mail: GAMorozov@kai.ru

Amplitude-phase modulation of laser radiation in generation problems of radio-frequency signals

O.G. Morozov, G.I. Il'in, G.A. Morozov

Kazan National Research Technical University n.a. A.N. Tupolev-KAI
10, K. Marx st., Kazan, Republic of Tatarstan, 420111, Russia

Abstract. From the point of view on requirements, defining the parameters of external amplitude-phase modulation, as key process for microwave frequency carrier's generation in the optical range, the basic structures, characteristics, element base and directions of development for microwave photonics generators are presented. Microwave photonics methods, discussed in the article, may be classified as follows: two optical carriers photo mixing as base method for generation of microwave frequency ones; external amplitude or phase electro-optic modulation for generation of two optical subcarriers with difference microwave frequency equal to two or four modulating ones; and external amplitude-phase electro-optic modulation for generation of two optical subcarriers with difference frequency, which is equal to microwave modulating one. Further, elements for realization of electro-optic amplitude-phase modulation in the optical range for the purposes of microwave frequency carrier's generation are discussed.

Keywords: microwave photonics; optical photo mixing generation of microwave frequency carriers; external electro-optic amplitude, phase and amplitude-phase modulation, element base of microwave photonic generators.

1. Introduction

Modern methods and means of microwave signals processing in the optical range [1-6] are applied in various radio and optical engineering systems for defense and civil purposes: radar warnings, ground and airborne reconnaissance assets, radio electronic warfare, RoF and UWBoF communications, amplitude and phase modulation for fiber optic communication lines, etc. The generation [7-9] of microwave frequencies (MF) carriers, the frequencies of local oscillators, the poly-harmonic radiations with difference microwave frequencies between components, etc. are related to the basic functions of transmission, reception and processing of information in microwave photonics systems.

A key element of microwave photonics systems is the subsystem of optical carrier modulation [1-6]. Its provides the transfer of microwave signal into an optical range, for example, to measure the instantaneous microwave frequencies, as well as the formation of poly-harmonic optical signal with a difference frequency, lying within the microwave frequencies range, for example, for microwave carrier generation with frequency equal to the difference one, achieved after photo mixing. Thus, as a rule, a separate set of lasers, spaced to the frequency of microwave signals, or lasers with internal modulation or modes synchronization, or lasers with external modulation, based on electro-optical modulators with operating frequency range up to 1-100 GHz of amplitude, phase, amplitude-phase, and polarization types are used for design of microwave photonics generators (MPG).

Due to the natural symmetry of modulated signals and the highest achievable ratio of the modulation conversions, amplitude-phase modulation with complete or partial suppression of the optical single-frequency laser radiation (carrier) has found a particularly wide application in the microwave photonics systems [10-13]. From the point of view on requirements, defining the parameters of external amplitude-phase modulation and formation of poly-harmonic laser radiation processes with a difference microwave band frequency, the basic structures, characteristics, components and directions of microwave photonics generators development are discussed.

From a general point of view, microwave photonics methods, shown in the article, may be classified as follows: two optical carriers photo mixing as base method for generation of microwave frequencies ones (section 2); external microwave amplitude or phase electro-optic modulation for generation of two optical subcarriers with difference frequency equal to two or four modulating ones (section 3); and external amplitude-phase electro-optic modulation for generation of two optical subcarriers with difference frequency, which is equal to microwave modulating one (section 4). Further, elements for realization of amplitude-phase modulation in the optical range for the purposes of microwave frequencies carrier's generation are presented (section 5). In conclusion we summarize advantages of one single frequency cw laser and one electro-optical dual-port modulator using for complex amplitude-phase modulation of optical carrier in order to design of microwave photonics generators with low phase noise.

2. Optical photo mixing as the basic procedure of microwave frequencies carrier generation

Tunable generators of microwave frequencies carriers with low level of phase noise are applied in various systems: radar, wireless communications, software defined radio, information and measurement [3, 14-15]. Typically, microwave frequencies carriers are generated by means of complex electronic circuits with multi-frequency multiplying to achieve the desired values. In addition, in many cases the generated microwave frequencies carrier must be distributed to remote devices. The distribution of such signals in the electronic field, in practice, is difficult because of large losses in distribution lines, for example, in coaxial cable. Thanks to the wide bandwidth and low loss of modern optical fibers, the distribution of microwave frequencies carriers over optical fiber is the ideal solution that greatly simplifies the requirements to the oscillators.

microwave frequencies carriers may be generated in the optical field with the use of optical photo mixing procedures, when two optical waves with different wavelengths are mixed in the photodetector (nonlinear quadrature element). At the photodetector output, a beat signal is generated with a frequency corresponding to the interval between the lengths of two optical waves. Suppose we have two optical waves:

$$E_1(t) = E_1 \cos(\omega_1 t + \phi_1), \quad (1)$$

$$E_2(t) = E_2 \cos(\omega_2 t + \phi_2), \quad (2)$$

where E_1, E_2 – its amplitude values, ω_1, ω_2 – angle frequencies, and ϕ_1, ϕ_2 – phase values.

Taking into account a restricted photodetector's bandpass its current may be described as follows:

$$I_{RF}(t) \propto 2A \cos[(\omega_1 - \omega_2)t + (\phi_1 - \phi_2)], \quad (3)$$

where A – amplitude, defined by values E_1, E_2 and the gain of the photodetector.

As follows from (3) an electrical signal is generated with a frequency equal to difference frequency between $\omega_1 - \omega_2$ of two optical waves (1)-(2). The value of this frequency can reach the terahertz range; the only restriction might be the photodetector bandwidth. However, the use of two unsynchronized laser diodes leads to the fact that the resulting radiofrequency signal will have a high level of phase noise since the phases of two wave's ϕ_1 and ϕ_2 are not correlated.

Several proposed methods [16-24] allow generation of microwave frequencies carriers with low phase noise level from two optical waves with synchronized phases. These methods

include synchronization methods with frequency capture [16, 17], phase-locked loop (PLL) [18, 19] and its combined version [20]. In a number of papers, generators based on dual-frequency lasers are presented [21, 22]. Methods based on external amplitude phase modulation of single-frequency laser radiation deserve special attention [23-24].

3. Microwave frequencies carrier generation by means of external amplitude or phase modulation of optical carrier

microwave frequencies carriers can also be generated by an external modulation [10-15, 23-26]. Here we present structural diagrams of generators while operating at "minimum" (fig.1,a) and "maximum" (fig.1,b) points of the modulation characteristics of Mach-Zehnder modulator (MZM) [14, 15]. Microwave signal with the doubled $2f_m$ frequency was optically generated with electrical bias Mach-Zehnder modulator $V_B = 0$ in the "minimum" operating point [14]. The signal of millimeter range of 36 GHz was obtained at modulation frequency $f_m = 18$ GHz. For frequency generation that exceeds in four times the frequency modulating signal $4f_m$ [15] a Mach-Zehnder modulator shifted to a "maximum" point of its modulation characteristic $V_B = U_\pi$ where U_π – is a Mach-Zehnder modulator half-wave voltage was used.

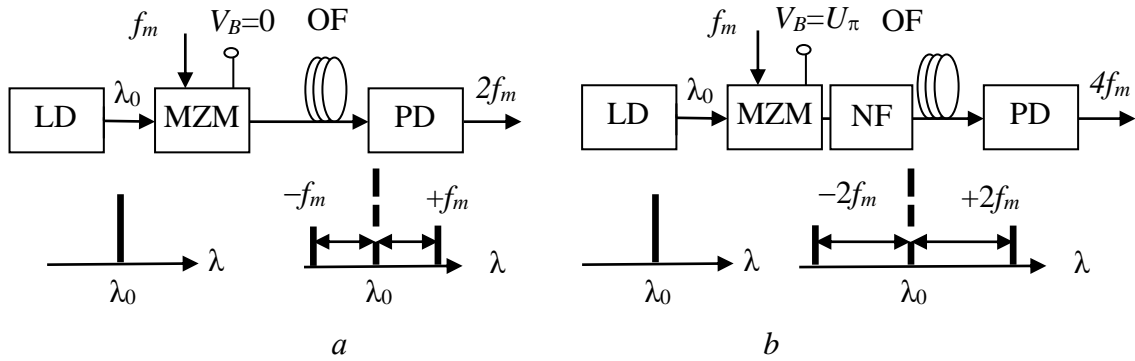


Fig. 1. Radiofrequency carrier generator with external modulation of the laser diode and MZM shift into «minimal» (a) and «maximal» (b) operating points: OF – optical fiber; LD – laser diode; PD – photodetector; NF – the notch filter

Optical carrier λ_0 , and harmonic components of the 1st and 3rd orders were suppressed by means of adjusting the level of the modulating signal and notch filters. As notch filters, Mach-Zehnder and Fabry-Perot tunable interferometers were used. The millimeter band signal at 60 GHz was achieved at a modulating frequency $f_m = 15$ GHz. For wideband review radar systems with a streamer signal or software-defined radio systems, it is necessary to use the microwave frequencies signal with a frequency tuning. Two approaches have been demonstrated in [23, 24] to generate frequency of tunable microwave frequencies signal, using an optical filter fixed on wavelength: based on the amplitude [23] and phase modulators [24]. The optical carrier suppression is carried out by the fiber Bragg gratings (FBG).

The main reasons for the presence of the higher components in the output spectrum of the amplitude modulator are nonlinear characteristics, temperature instability, and deviation of the transformation parameters from the optimal and the error of alignment of the polarizer's converter. They all cause the deviation of the operating point and the rapid growth of the amplitudes of spurious spectral components. A key advantage of optical phase modulator usage is that there is no need to switch on a constant shift, which eliminates the problem of drift of the operating point [24]. Unlike Mach-Zehnder modulator, which can be controlled for odd or even harmonic components selection, the optical phase modulator usage will

generate their full range including the optical carrier. The beats left after fiber Bragg gratings optical carrier filtering on side harmonic components at the photodetector output will allow generating the microwave frequencies signal.

It should be noted that higher optical harmonics transmitted through single-mode fiber will be subjected to chromatic dispersion, which will change the phase ratio in lateral components. To maintain the ratio of the same phases, there is necessity of dispersion compensation. Under conditions of compensation, it has been experimentally demonstrated that at modulating with a frequency lying in the range of $(18.8 \div 25)$ GHz, locally and remotely were obtained two bands of signals of millimeter range of $(37.6 \div 50)$ GHz and $(75.2 \div 100)$ GHz.

4. Microwave frequencies carrier generation with frequency equal to the modulating one (amplitude-phase modulation)

Let's consider the transition from amplitude-modulated signal with absolute modulation to the beat signal. This transition can be achieved either by suppressing the carrier oscillation (first option) or by switching the phase of the amplitude-modulated signal at the moment, when its envelope achieves a zero value (second option).

The spectrum of the initial quasi-harmonic oscillation with the amplitude-modulated signal structure (a) and spectrums of transformed oscillations with the beat signal structure received on the first (b) and second (c) options are presented in fig. 2.

In this case, the obtained spectra are significantly different at difference frequencies 2Ω (fig. 2,b) and Ω (fig. 2,c) depending on the conversion method. The first option implementation help us to get double sideband with suppressed carrier (DSB-SC) signal, and second double half sideband with suppressed carrier (DHSB-SC). Therefore, we can get in first option radiofrequency or microwave carrier with frequency 2Ω , and in second – with frequency Ω . It is the first time, when we show radiofrequency carrier generation with frequency equal to the modulating one.

Let us consider the first option implementation. The radiofrequency generator of the centimeter range was built on the principles of amplitude-phase modulation in a classical amplitude electro-optic modulator (AEOM) on lithium metaniobate LiNbO_3 crystal, the symmetry class 3m [10-13, 25-26].

Amplitude electro-optic modulator connection with laser diode and photodetector was carried out by means of fiber-optic pigtails. When operating the generator in a “zero” point of modulation characteristics output spectrum is described by the expression

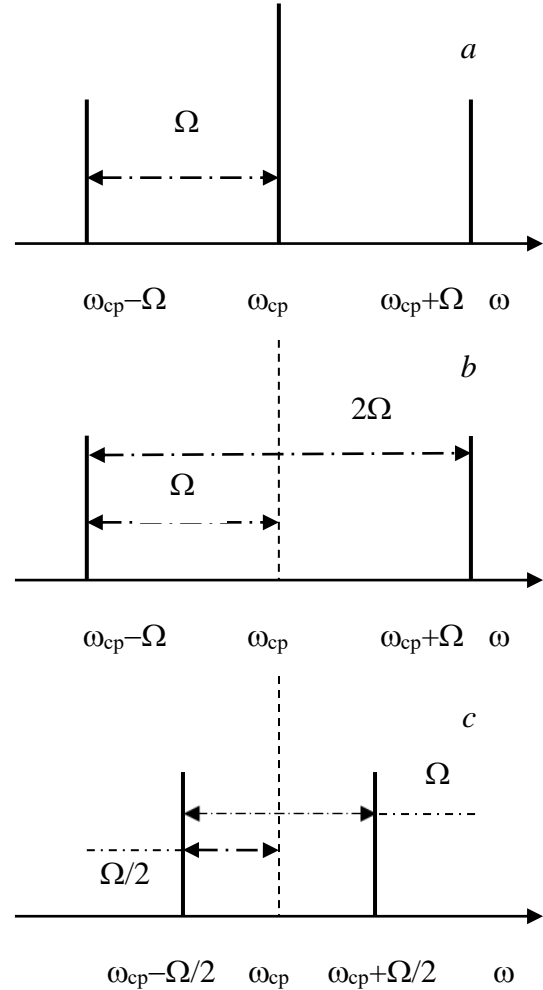


Fig. 2. Amplitude modulated spectrum signal before (a) and after modulation transformation by first (b) and second (c) options

$$E_{\text{BYX}} = -jE_0 e^{j\omega_0 t} \left[2 \sum_{k=0}^{\infty} \mathbf{J}_{2k+1}(z) \sin(2k+1)\Omega t \right], \quad (4)$$

where ω_0 is the angular frequency of optical carrier, and Ω – of modulating radiofrequency signal; index j in this expression indicates that the output radiation is orthogonal to the polarization vector of the source radiation; $\mathbf{J}_{2k+1}(z)$ is the Bessel function of the $(2k+1)$ order. At half-wave voltage $Um = U_{\lambda/2}$ we get $z=\pi/2$, $\mathbf{J}_1(z)=0,64$, $\mathbf{J}_3(z)=0,06$. The input single frequency and dual-frequency output spectra of the generator radiation, and the radiofrequency signal at the photodetector output are shown in Fig. 3, *a-c*, respectively.

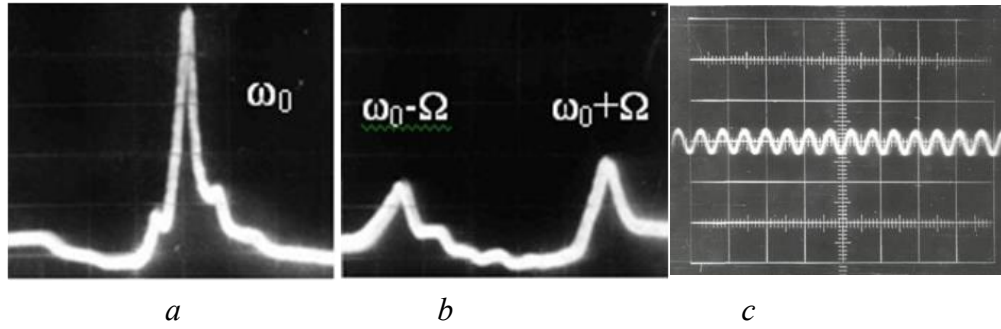


Fig. 3. Spectrum of input (*a*) and output (*b*) AEOM radiation, the waveform of the radiofrequency signal (*c*) at PD output (fig. 1, *a*)

A maximum coefficient of transformation (4) is 0.64. When the conversion coefficient is 0.58 the total harmonic distortion does not exceed 1%. There are two more important points worth noting, it is the equality of amplitudes of spectral components and opposite phases, which is important e.g. for the generation of solitons [39] and the construction of photonic filters of radio-frequency signals [29, 40-41], regardless of the location of the working point, and ease of frequency tuning, due to the use of a single modulating signal. The possibility of even-numbered components in the spectrum of the output radiation, is shown in (5), can be useful at generating RZ (return-to-zero) and CSRZ (carrier-suppressed return-to-zero) pulse sequences [40]:

$$e(t) = -j \frac{\sqrt{2}}{2} e^{j\omega_0 t} \times \left\{ \sum_{k=0}^{\infty} \mathbf{J}_0(z) + 2 \sum_{k=1}^{\infty} \mathbf{J}_{2k}(z) \cos 2k\Omega t + \right. \\ \left. + 2 \sum_{k=0}^{\infty} \mathbf{J}_{2k+1}(z) \sin(2k+1)\Omega t \right\}. \quad (5)$$

The theoretical results of the spectrum analysis on the output of the proposed generator are given in table 1, and are obtained under the assumption that the modulator is biased at zero applied voltage $\Gamma_- = 0$, and polarizers are either orthogonal $\alpha = -\beta = 45^\circ$ or parallel $\alpha = \beta = 45^\circ$.

Output radiation is fed from amplitude electro-optic modulator to a polarizing beam splitter; one of its principal axes is oriented at a 45° angle to amplitude electro-optic modulator. At two outputs of the polarization beam splitter will receive two optical signals along two principal axes, and one signal includes all of the optical frequencies of even order (first line of table 1) and the other consists of all the optical frequencies of the odd one (first line of table 1). This allows generating both return-to-zero and carrier-suppressed return-to-zero pulse consequences respectively in the range of $(0.3 \div 1)$ GHz, determined by the amplitude electro-optic modulator bandwidth⁴⁰. The whole system is equivalent to the pulse generator, which consists of a phase modulator and the interacting pair of comb filters for sampling odd

and even sidebands. The main advantage of the proposed method is that for the separation of the sequences is not used physical interacting pair of comb filters, which makes much easier the adjustment of the repetition frequency.

Table 1. AEOM output spectral characteristics

$\alpha, \beta, \Gamma_{\pm}$	AEOM output spectra
$\alpha=45^\circ, \beta=-45^\circ, \Gamma_{\pm} = 0$	$E_{\text{out}} = -jEe^{j\omega_0 t} \left[2 \sum_{k=0}^{\infty} \mathbf{J}_{2k+1}(z) \sin(2k+1)\Omega t \right]$
$\alpha=45^\circ, \beta=45^\circ, \Gamma_{\pm} = 0$	$E_{\text{out}}(t) = Ee^{j\omega_0 t} \left\{ \mathbf{J}_0(z) + 2 \sum_{k=1}^{\infty} \mathbf{J}_{2k}(z) \cos 2k\Omega t \right\}$

Let us consider the second option implementation. The simplest oscillation, satisfying modulation requirements is the oscillation of $S_1(t) = S_1 \cos(\Omega t + \pi)$ type, where S_1 is its constant amplitude, and π is an initial phase. In this case, the phase of the carrier wave ω_0 will be switched to π when passing the minimum of modulating oscillation. In the case of switched carrier amplitude modulation by $S_1(t)$ oscillation, the resulting spectrum has the following form:

$$e(t) = \frac{2E_0}{\pi} \sum_n \left\{ \left[\frac{1}{n} - \frac{m}{2} \left(\frac{1}{n-2} + \frac{1}{n+2} \right) \right] [\cos(\omega_0 + n\Omega/2)t - \cos(\omega_0 - n\Omega/2)t] \right\}. \quad (6)$$

where m is coefficient of amplitude modulation.

From the analysis of (6) we see that the first term in large square brackets defines the range of coefficients of Fourier series for switched carrier, and the second and third describe the vast influence on its components by modulating oscillation. The degree of suppression depends on the modulation coefficient m . Taking $E_3 = 0$, we get the optimum modulation index $m_{\text{opt}} = 5/9$, practically with the resulting dual-frequency oscillation ($E_{\pm 1} = 0,76E_0$), since the amplitude of spectral components $E_n \leq E_0/15$ for $n \geq 5$. When you change the modulation index in the range of $(0,85 - 1,15)$ m_{opt} coefficient of nonlinear distortion of the output oscillation does not exceed 1%. Complete suppression of the lateral components with $n \geq 3$ can be achieved using modulation oscillation of $S_2(t) = S_0 |\sin(\Omega/2)t|$ type for amplitude modulation. Then the resulting oscillation will have the following range

$$e(t) = \frac{2E_0}{\pi} (1-b) \sum_n \frac{1}{n} \{ \cos(\omega_0 + n\Omega/2)t - \cos(\omega_0 - n\Omega/2)t \} + \frac{\pi E_0 b}{4} \{ \cos(\omega_0 + \Omega/2)t - \cos(\omega_0 - \Omega/2)t \}. \quad (7)$$

where b is amplitude modulation coefficient.

The amplitude of the spectral components will be determined by the coefficients of Fourier series for $n=1$ $E_1 = [2E_0/\pi][1-b] + [\pi E_0 b/4]$, and for $n \geq 3$ $E_n = [2E_0/\pi n][1-b]$. When $b_{\text{opt}} = 1$ spectrum contains two useful components at $\omega_0 + \Omega/2$ and $\omega_0 - \Omega/2$ frequencies, spurious components are suppressed completely. When you change the modulation index in $(0,7 - 1)b_{\text{opt}}$ range the coefficient of nonlinear distortion of output oscillation does not exceed 1%. Dual-frequency radiation can be obtained by complete implementation of the Il'in-Morozov's method¹⁰⁻¹³ by using the amplitude Mach-Zehnder modulator (AMZM), operating in a linear range of the working characteristics, and phase Mach-Zehnder modulator (PMZM). At the same time to control phase Mach-Zehnder modulator we used a conversion unit from two radiofrequency generators (RFG A – RFG P) of a sinusoidal signal with a frequency Ω in a

meander with a frequency of $\Omega/2$ for the phase commutation with the corresponding amplitudes of the operating voltages $U_{\pi/2}$ and U_{π} , which is presented in fig. 4,a and was not considered before. Fiber Bragg gratings is used to further suppression of the laser source optical carrier. The spectrum for the two components at the modulator output in this case is described by the expression:

$$E_{\text{A}\Phi\text{M}}(t) = 0,49 E_0 \{ \sin(\omega + \Omega/2)t - \sin(\omega - \Omega/2)t \} + \\ + 0,007 E_0 \{ \sin(\omega + 3\Omega/2)t - \sin(\omega - 3\Omega/2)t \} + \dots \quad (8)$$

As can be seen from (8) the difference frequency between the components of dual-frequency radiation Ω is equal to the frequency of the modulating signal and its view after photodetector is shown on fig. 4,b.

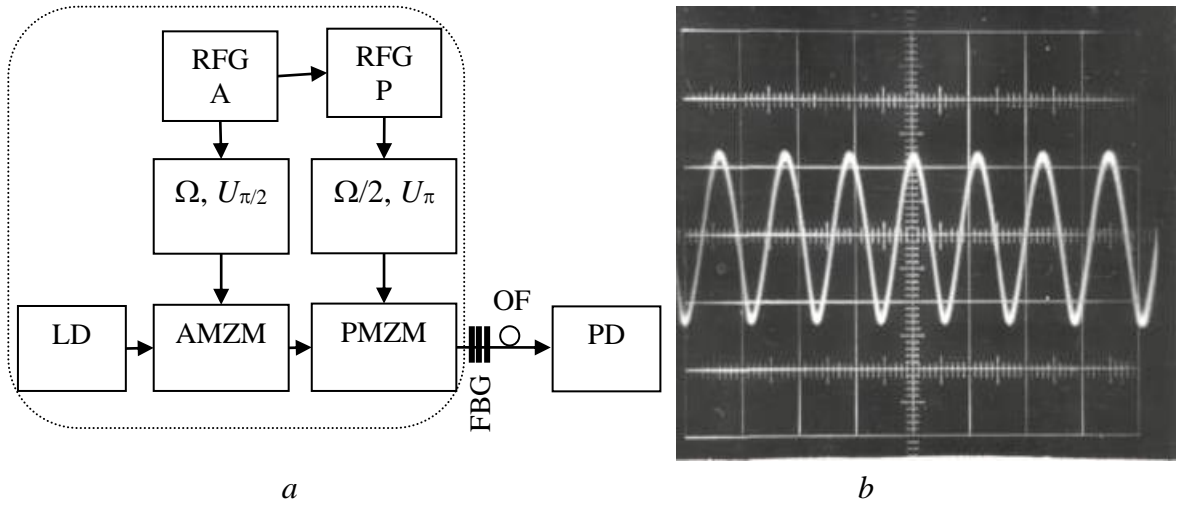


Fig. 4. Generator of dual frequency symmetrical radiation and waveform radiofrequency carrier on photodetector output (its frequency is two time less than put on fig. 3,c)

Components of higher harmonics can be ignored because of the smallness of their amplitude. We achieved a narrowing of the difference frequency in two times in comparison with classical schemes of its doubling, used in the practice of known microwave photonics systems, for example, while using a single amplitude Mach-Zehnder modulator applied to suppress the carrier in a "zero" point of the modulation characteristics, as indicated in (4) and (6) and fig. 1,a, fig. 3,b,c and fig. 4,b.

5. Electro-optic modulators for microwave photonic generators

In microwave photonics applications dual channel electro-optic modulator based on Mach-Zehnder interferometer are mainly used implementing amplitude, phase and amplitude-phase modulation. When electric field is applied to one of two parallel electrodes there is a change in the first optical channel of the refractive index and the phase difference occurs relatively to the optical wave propagating in the second channel, which determines the intensity and phase of signal at the output of the modulator.

Currently electro-optic modulator based on lithium niobate LiNbO_3 , as well as, on gallium arsenide GaAs and indium phosphide InP, optimized for the wavelengths of 1.3–1.6 μm , are actively used. In case of using the InP substrate it is possible to create photonic integrated circuits, including lasers, modulators, multiplexers in wavelength, waveguides and drive electronics. For effective modulation in a compact integrated modulator, it is necessary to

ensure the greatest change of refractive index in its modulatory section, with a slight increase in its absorption coefficient and at the smallest possible voltage applied to the modulator.

The main characteristics of LiNbO₃ modulators are half-wave voltage, bandwidth, optical insertion loss, temperature and working point stability, the optical sensitivity of the refractive index. Their main purpose – operation in the frequency range up to 40-60 GHz. The technology of the specified class modulators is best suited to work with signals of microwave and millimeter ranges. However, there are several disadvantages associated with the presence of phase distortions and non-uniformity of the working characteristics near the upper limit of the frequency range. The main manufacturers of LiNbO₃ modulators are *JDSU*, *Avanex*, *Sumitomo*, *Photline* and *EOSPACE*, etc.

Gallium arsenide GaAs and indium phosphide InP, due to their optical properties and high integration potential into compact multi-function devices, and the lack of symmetry in the crystal structure, play a significant role in expanding frequency range of the processed signals and creation of advanced systems of microwave photonics. Conventionally, their frequency response can be estimated as up to 100 GHz for GaAs and more than 100 GHz for InP. Despite the relatively small value of the electro-optic coefficient ($r_{41}=1,4$ pm/V) and high refractive index (up to 3.4), insuring of 100% by crystal's volume of interaction of optical and electric fields define the high electro-optical efficiency of GaAs and InP semiconductors, comparable to that attainable in LiNbO₃. However, the obtained "tension-length" ($U_{\pi}L_{\pi}$) setting is on average 9 V·cm, while estimated at 5,5 V·cm, which allows you to do electro-optic modulator of smaller length and sizes than modulators based on LiNbO₃. A major provider of electro-optic modulator based on gallium arsenide GaAs and indium phosphide InP is *Bookham*.

A special place in microwave photonics systems takes *Versawave Technologies* polarization modulator [42]. Polarization modulator (PolM) is a special phase modulator that supports both modes TM and TE, providing the phase modulation in each of them, but with the opposite sign. Such modulation is achieved by tuning the polarization vector of the source radiation by the polarization controller so that its position equals to 45° with one of the principal axes of polarization modulator. The normalized optical field at the polarization modulator output can be described as follows:

$$\begin{bmatrix} \frac{E_x}{E_y} \end{bmatrix} \propto \frac{\sqrt{2}}{2} \begin{bmatrix} \exp(j\omega_0 t + j\beta_1 \cos \omega_{RF} t) \\ \exp(j\omega_0 t - j\beta_1 \cos \omega_{RF} t) \end{bmatrix}, \quad (9)$$

where β_1 is the index of phase modulation in polarization modulator, which is defined as $\beta_1 = \pi U_p / U_{\pi}$ with appropriate amplitude values of the modulating voltage U_p and a half-wave voltage U_{π} . The polarization controller at the polarization modulator output allows you to combine both components or to highlight each of them.

Polarization modulator is fabricated on a GaAs technology, has a contrast ratio of 20, the optical insertion loss of 3.5 dB and half-wave voltage up to 3.5 V in the frequency range from 40 to 80 GHz.

When using two cascaded modulators amplitude Mach-Zehnder modulator and phase Mach-Zehnder modulator, realizing Il'in-Morozov's method [10-13], all the modulation types are simply implemented, including amplitude-phase, obtained in LiNbO₃ crystals and presented in [43].

6. Conclusion

The review study presented in this article showed that the development of microwave photonics systems is constrained by the difficulty of forming stable in frequency and synchronized in radiation phase sources of probing or the local oscillations with the possibility of frequency adjustment in the general case by the any law and any speed. This

requirement is determined by the necessity of receiving signals or measurements with the required sensitivity and signal-to-noise ratio in a wide frequency range.

These problems can be solved by building microwave photonic systems based on single frequency cw lasers and the amplitude-phase electro-optical modulators (one can use the polarization modulators). In the basis of the task solution is the use of a unique method [10-13], which had been called Il'in-Morozov's method, and was proposed by the authors in 1983 [43].

It is based on amplitude modulation and phase manipulation of the single-frequency optical carrier for the synthesis of symmetrical dual-frequency and poly-harmonic radiations. The structure of the output poly-harmonic radiation at the Il'in-Morozov's method realization contains two symmetrical components relatively suppressed source, with equal amplitudes and alternative phases and a difference frequency equal to the frequency of the modulating radiofrequency signal or its double value. It is characterized by a high degree of spectral purity when occur the deviations of the transformation parameters from optimal. Method development was carried out based on a detailed study of the modulation characteristics of an amplitude electro-optical modulator according to the electric field intensity, unlike widely used at that time intensity analysis of modulation characteristics. It was shown that when the amplitude electro-optic modulator operated in the «zero» (minimum) point of the modulation characteristics the output radiation is not amplitude-modulated with side frequencies at a double modulation frequency, but it is dual frequency with suppressed carrier and the components separated by doubled modulation frequency.

In this article, we shown the possibility of microwave frequencies signals generation with low phase noise level and generation of poly-harmonic radiation with radio frequency components. Moreover, we studied the possibility of applying the technologies based on lithium niobate, gallium arsenide and indium phosphide for creation of elements of the advanced modulation microwave photonics systems in different radio frequency bands. The applications of methods and means in different microwave photonic systems presented by us in [44-53], as in joint works with our colleagues [54-56].

Acknowledgements

The work was executed at financial support of the Ministry of education and science of the Russian Federation in the framework of the state task on services rendering (performance of works) on the scientific research organizations carried out by Kazan National Research Technical University n.a. A.N. Tupolev – KAI in the departments of Radiophotonics and Microwave Technologies, Radio-Electronic and Quantum Devices and in the R&D Institute of Applied Electrodynamics, Photonics and Living Systems (programs «Radiophotonics» 3.1962.2014/K and «Groza» 3.2426.2014/K).

Abbreviations list

RoF – Radio over Fiber	NF –notch filter
UWBoF – ultra wideband over fiber	DSB-SC – double sideband with suppressed carrier
MF – microwave frequencies	AEOM – amplitude electro-optic modulator
MPG – microwave photonics generators	RZ – return-to-zero
PLL – phase-locked loop	CSRZ – carrier-suppressed return-to-zero
MZM – Mach-Zehnder modulator	AMZM – amplitude Mach-Zehnder modulator
FBG – fiber Bragg gratings	PMZM – phase Mach-Zehnder modulator
OF – optical fiber	RFG – radiofrequency generator
LD – laser diode	
PD – photodetector	

References

1. A. Seeds, "Microwave photonics," *IEEE Trans. Microw. Theory Tech.* 50(3), 877-887 (2002).
2. J. Capmany, D. Novak, "Microwave photonics combines two worlds," *Nat. Photonics* 1(6), 319-330 (2007).
3. J. Yao, "Microwave photonics," *J. Lightwave Technol.* 27(3), 314-335 (2009).
4. T. Berceli, P.R. Herczfeld, "Microwave photonics – a historical perspective," *IEEE Trans. Microw. Theory Tech.* 58 (11), 2992-3000 (2010).
5. J. Capmany, et al., "Microwave photonic signal processing," *J. Lightwave Technol.* 31(4), 571-586 (2013).
6. R.A. Minasian, E.H.W. Chan, X. Yi, "Microwave photonic signal processing," *Opt. Express* 21(19), 22918-22936 (2013).
7. L. Maleki, "The optoelectronic oscillator," *Nat. Photonics* 5(12), 728-730 (2011).
8. C.H. Lee, "Optical generation and control of microwave and millimeter-wave signals," *IEEE MTT-S International Microwave Symposium Digest* 2, 811-814 (1987).
9. P.R. Herczfeld, et al., "Indirect subharmonic optical injection locking of a millimeter-wave IMPATT oscillator," *IEEE Trans. Microw. Theory Tech.* 34(12), 1371-1376 (1986).
10. G.I. Il'In, O.G. Morozov, A.G. Il'In, "Theory of symmetrical two-frequency signals and key aspects of its application," *Proc. of SPIE* 9156, 91560M (2014).
11. O.G. Morozov, D.L. Aybatov, "Spectrum conversion investigation in lithium niobate Mach-Zehnder modulator," *Proc. of SPIE* 7523, 75230D (2010).
12. O.G. Morozov, et al., "Methodology of symmetric double frequency reflectometry for selective fiber optic structures," *Proc. of SPIE* 7026, 70260I (2008).
13. O.G. Morozov, et al., "Metrological aspects of symmetric double frequency and multi frequency reflectometry for fiber Bragg structures," *Proc. of SPIE*, 7026, 70260J (2008).
14. J.J. O'Reilly, et al., "Optical generation of very narrow line width millimeter wave signals," *Electron. Lett.* 28(25), 2309–2311 (1992).
15. J.J. O'Reilly, P.M. Lane, "Fiber-supported optical generation and delivery of 60 GHz signals," *Electron. Lett.* 30(16), 1329-1330 (1994).
16. L. Goldberg, et al., "Microwave signal generation with injection locked laser diodes," *Electron. Lett.* 19(13), 491-493 (1983).
17. L. Goldberg, et al., "35 GHz microwave signal generation with injection locked laser diode," *Electron Lett.* 21(18), 714-715 (1985).
18. R.T. Ramos, A.J. Seeds, "Fast heterodyne optical phase-lock loop using double quantum well laser diodes," *Electron. Lett.* 28(1), 82-83 (1992).
19. H. Rideout, et al., "Discriminator-aided optical phase-lock loop incorporating a frequency down-conversion module," *IEEE Photon. Technol. Lett.* 18(22), 2344-2346 (2006).
20. A.C. Bordonalli, C. Walton, A.J. Seeds, "High-performance phase locking of wide line width semiconductor lasers by combined use of optical injection locking and optical phase-lock loop," *J. Lightw. Technol.* 17(2), 328-342 (1999).
21. X. Chen, Z. Deng, J.P. Yao, "Photonic generation of microwave signal using a dual-wavelength single-longitudinal-mode fiber ring laser," *IEEE Trans. Microw. Theory Tech.* 54(2), 804-809 (2006).
22. X. Chen, Z. Deng, J.P. Yao, "Ultrannarrow dual-transmission-band fiber Bragg grating filter and its application in a dual-wavelength single-longitudinal-mode fiber ring laser," *Opt. Lett.* 30(16), 2068-2070 (2005).
23. G. Qi, et al., "Generation and distribution of a wide-band continuously tunable mm-wave signal with an optical external modulation technique," *IEEE Trans. Microw. Theory Tech.* 53(10), 3090-3097 (2005).

24. G. Qi, et al., "Optical generation and distribution of continuously tunable millimeter-wave signals using an optical phase modulator," *J. Lightw. Technol.* 23(9), 2687-2695 (2005).
25. O.G. Natanson, et al., "Development problems of frequency reflectometry for monitoring systems of optical fiber structures," *Proc. of SPIE* 5854, 215-223 (2005).
26. O. G. Natanson, et al., "Reflectometry in open and fiber mediums: technology transfer," *Proc. of SPIE* 5854, 205-214 (2005).
27. O.G. Morozov, et al., "Structural minimization of fiber optic sensor nets for monitoring of dangerous materials storage," *Proc. of SPIE* 7992, 79920E (2011).
28. O.G. Morozov, et al., "Characterization of stimulated Mandelstam-Brillouin scattering spectrum using a double-frequency probing radiation," *Proc. of SPIE* 8787, 878709 (2013).
29. O.G. Morozov, T.S. Sadeev, "All-optical microwave photonic filter based on two-frequency optical source," *Proc. of SPIE* 7992, 79920C (2011).
30. Diddams, S.A., "The evolving optical frequency comb," *J. Opt. Soc. Amer. B.* 27(11), B51-B62 (2010).
31. F.Z. Zhang, et al., "Flat optical frequency comb generation and its application for optical waveform generation," *Opt. Commun.* 290(1), 37-42 (2013).
32. C. He, et al., "Ultraflat optical frequency comb generated based on cascaded polarization modulators," *Opt. Lett.* 37(18), 3834-3836 (2012).
33. M. Fujiwara, et al., "Optical carrier supply module using flattened optical multicarrier generation based on sinusoidal amplitude and phase hybrid modulation," *J. Lightwave Technol.* 21(11), 2705 (2003).
34. Y.J. Dou, H.M. Zhang, M.Y. Yao, "Generation of flat optical-frequency comb using cascaded intensity and phase modulators," *Opt. Lett.* 36, 2749-2751 (2011).
35. R. Wu, et al., "Generation of very flat optical frequency combs from continuous-wave lasers using cascaded intensity and phase modulators driven by tailored radio frequency waveforms," *Opt. Lett.* 35(19), 3234-3236 (2010).
36. T. Yamamoto, et al., "Multicarrier light source with flattened spectrum using phase modulators and dispersion medium," *J. Lightwave Technol.* 27(19), 4297-4305 (2009).
37. C. Chen, et al., "Generation of a flat optical frequency comb based on a cascaded polarization modulator and phase modulator," *Opt. Lett.* 38(16), 3137-3140 (2013).
38. T.S. Sadeev, O.G. Morozov, "Investigation and analysis of electro-optical devices in implementation of microwave photonic filters," *Proc. of SPIE* 8410, 841007 (2012).
39. E.M. Dianov, et al., "Generation of a train of fundamental solitons at a high repetition rate in optical fibers," *Opt. Lett.* 14(18), 1008-1010 (1989).
40. O.G. Morozov, "RZ, CS-RZ and soliton generation for access networks applications: problems and variants of decisions," *Proc. of SPIE* 8410, 84100P (2012).
41. D.L. Aybatov, O.G. Morozov, T.S. Sadeev, "Dual port MZM based optical comb generator for all-optical microwave photonic devices," *Proc. of SPIE* 7992, 799202 (2011).
42. X. Zou, J. Yao, "Repetition-rate-tunable return-to-zero and carrier-suppressed return-to-zero optical pulse train generation using a polarization modulator," *Opt. Lett.* 34(3), 313-315 (2009).
43. O.G. Morozov, et al., "Synthesis of two-frequency symmetrical radiation and its application in fiber optical structures monitoring," *InTech*, 2012, <http://www.intechopen.com/books/fiber-optic-sensors/synthesis-of-two-frequency-symmetricalradiation-and-its-application-in-fiber-opticalstructures-mon> (10 December 2015).
44. O.G. Morozov, D.L. Aibatov, "Two-frequency scanning of FBG with arbitrary reflection spectrum," *Proc. of SPIE* 6605, 660506 (2007).
45. O.A. Stepustchenko, et al., "Optical refractometric FBG biosensors: problems of development and decision courses," *Proc. of SPIE* 7992, 79920D (2011).

46. O.G. Morozov, et al., "Instantaneous frequency measurement of microwave signals in optical range using "frequency-amplitude" conversion in the π -phase-shifted fiber Bragg grating," Proc. of SPIE 9136, 91361B (2014).
47. O.G. Morozov, et al., "Instantaneous microwave frequency measurement with monitoring of system temperature," Proc. of SPIE 9156, 91560N (2014).
48. O.G. Morozov, et al., "Instantaneous frequency measurement using double-frequency probing," Proc. of SPIE 8787, 878708 (2013).
49. O.G. Morozov, Y.E. Pol'ski, "Perspectives of fiber sensors based on optical reflectometry for nondestructive evaluation," Proc. of SPIE 2944, 178-183 (1996).
50. O.G. Morozov, et al., "Two-frequency analysis of fiber-optic structures," Proc. of SPIE 6277, 62770E (2006).
51. D.L. Aybatov, et al., "Distributed temperature fiber Bragg grating sensor," Proc. of SPIE 7374, 73740B (2009).
52. O.G. Morozov, et al., "Double mode system for FWM reducing," Proc. of SPIE 7026, 70260B (2008).
53. O.G. Morozov, et al., "Poly-harmonic Analysis of Raman and Mandelstam-Brillouin Scatterings and Bragg Reflection Spectra," InTech, 2015, <http://www.intechopen.com/books/advances-in-optical-fiber-technology-fundamental-optical-phenomena-and-applications/poly-harmonic-analysis-of-raman-and-mandelstam-brillouin-scatterings-and-bragg-reflection-spectra> (10 December 2015).
54. V.A. Andreev, et al., "Investigation of cross-phase modulation in multichannel dispersion managed soliton systems," Proc. of SPIE 7374, 737405-8 (2009).
55. A.Kh. Sultanov, et al., "Measuring of chromatic and polarization mode dispersion in optical fiber," Proc. of SPIE 4589, 192 (2011).
56. V.A. Burdin, M.V. Dashkov, "Polarization optical time domain reflectometer with linear extension of pulse width," Proc. of SPIE 7523, 75230K (2010).

Oleg Gennadyevich Morozov, Doctor of Engineering Sciences, Professor, the Head of the Radiophotonics and Microwave Technologies Department, the Head of R&D Institute of Applied Electrodynamics, Photonics and Living Systems at Kazan National Research Technical University named after A.N. Tupolev-KAI (Kazan, Russian Federation). The sphere of scientific interests is information-measuring and telecommunication systems of optical and microwave ranges. The author of more than 240 publications.

E-mail: OGMorozov@kai.ru

German Ivanovich Il'in, Honoured Science Worker of Russian Federation, Honoured Science and Engineering Worker of Republic of Tatarstan, Doctor of Engineering Sciences, Professor, the Head of Radio Electronics and Quantum Devices Department at Kazan National Research Technical University named after A. N. Tupolev-KAI (Kazan, Russian Federation). The sphere of scientific interests is laser and radio technique systems, theory of potential interference and noise immunity. The author of more than 150 publications.

E-mail: Gillin@kai.ru

Gennady Alexandrovich Morozov, Honoured Educational Worker of Russian Federation, Honoured Science and Engineering Worker of Republic of Tatarstan, Doctor of Engineering Sciences, Professor, the Professor of Radio Electronics and Telecommunication Systems Department at Kazan National Research Technical University named after A. N. Tupolev-KAI (Kazan, Russian Federation). The sphere of scientific interests is antennas and microwave technologies, biotechnologies. The author of more than 150 publications.

E-mail: GAMorozov@kai.ru

Автоматизированное проектирование широкополосного диодного удвоителя частоты в диапазоне (26÷40) ГГц

В.В. Березин¹, А.М. Щитов², А.С. Нечаев³

¹ОАО «ФНПЦ «ННИИРТ», г. Н. Новгород
Россия, 603950, Нижний Новгород, ул. Шапошникова, 5

²ОАО «ФНПЦ «ННИПИ «Кварц» им. А.П. Горшкова», г. Н. Новгород

³Нижегородский государственный технический университет им. Р.Е. Алексеева
г. Н. Новгород, Российская Федерация, 603950, ГСП-41, ул. Минина, д. 24

Аннотация. Предложена методика расчета и проектирования волноводного диодного удвоителя частоты в диапазоне (26÷40) ГГц на базе программ автоматизированного проектирования HFSS и Designer фирмы ANSYS, которая сочетает электродинамический расчет волноведущих структур с анализом нелинейных характеристик устройства. Приводится сравнение расчетных параметров удвоителя частоты с экспериментальными данными.

Ключевые слова: удвоитель частоты, автоматизированное проектирование, диодный умножитель.

1. Введение

В [1] предложена методика расчета и проектирования волноводного диодного удвоителя частоты в диапазоне (75÷110) ГГц, сочетающая электродинамический расчет волноведущих структур с анализом нелинейных характеристик устройства. В [2] с применением данной методики рассчитан и оптимизирован удвоитель частоты в диапазоне (110÷170) ГГц. Экспериментальные данные измерения потерь преобразования хорошо согласуются с расчетом.

В настоящей работе предпринята попытка распространить методику для расчета и проектирования волноводных диодных удвоителей крайне высокочастотного диапазона (КВЧ). Ставилась задача подтверждения работоспособности методики, а так же определение потенциальных возможностей совершенствования технических характеристик (в первую очередь эффективности преобразования) ранее экспериментально разработанной модели волноводного удвоителя частоты [3-4].

2. Схемно-конструктивные особенности удвоителя частоты

Удвоитель частоты выполнен по схеме двухполупериодного выпрямителя (рисунок 1). Конструктивно удвоитель частоты выполняется в виде комбинации двух линий передачи: входной – копланарного волновода и выходной – щелевой линии, имеющей плавный или ступенчатый переход на волноводный тракт WR-28 (рисунок 2). Линии передачи изготавливаются на тонкой диэлектрической подложке, размещенной в разрезе по центру широкой стенки волновода; диоды размещены в плоскости соединения линий. Входной сигнал подается через коаксиально-полосковый переход. Со стороны входа "земляные" обкладки копланарного волновода замкнуты перемычкой на расстоянии l от плоскости размещения диодов, образуя отрезок линии передачи четвертьволновой длины на средней частоте выходного сигнала. Короткозамкнутый отрезок, благодаря шунтирующему действию, уменьшает уровень 4-й гармоники в спектре выходного сигнала, так как составляет для нее половину длины волны.

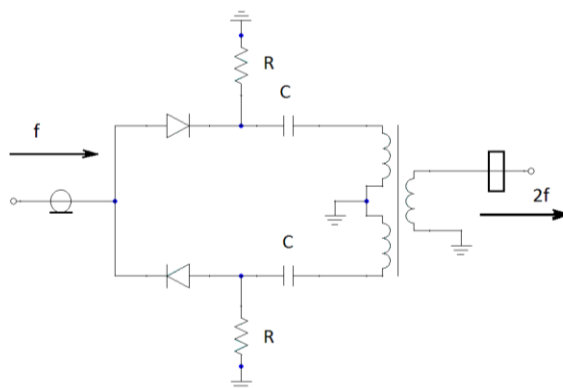


Рис. 1. Эквивалентная схема удвоителя частоты диапазона (26÷40) ГГц

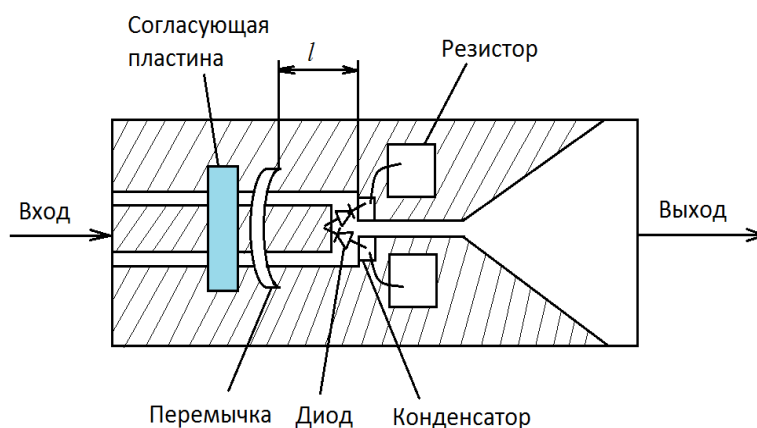


Рис. 2. Эскиз удвоителя частоты для диапазона (26÷40) ГГц

Фильтрация 1-й гармоники осуществляется, помимо баланса схемы, выходным волноводным трактом, так как для нее волновод является запердельным. С целью достижения минимума потерь преобразования в заданном динамическом диапазоне входных сигналов используют цепи внешнего смещения или внутреннего самосмещения. До сих пор в литературе нет детального теоретического анализа и расчета основных характеристик умножителя частоты. На Рисунке 3 показан внешний вид удвоителя.



Рис. 3. Внешний вид удвоителя частоты для диапазона (26÷40) ГГц

Наиболее приоритетными техническими характеристиками широкополосных умножителей частоты являются: эффективность (потери) преобразования, максимально достижимая выходная мощность, неравномерность выходной мощности в диапазоне частот и уровень паразитных составляющих в спектре выходного сигнала. Повышение и оптимизация этих характеристик удвоителя частоты являются основными задачами данной работы.

3. Моделирование удвоителя частоты

При моделировании удвоителя использовалась система автоматизированного проектирования компании Ansys (HFSS и Designer). На первом этапе моделирования удвоителя частоты использовалась среда HFSS.

Процесс проектирования с помощью HFSS включает в себя ряд стандартных шагов:

1. Создание модели анализируемой структуры (рисунок 4), в том числе:
 - создание трехмерной графической модели структуры (создание чертежа);
 - задание параметров материалов, из которых состоит структура.
2. Определение электродинамических параметров структуры, включающее:
 - задание граничных условий на поверхностях, формирующих анализируемый объект;
 - определение и калибровка портов;
 - задание параметров решения.
3. Электродинамический анализ исследуемого объекта, в том числе:
 - анализ объекта в полосе частот;
 - параметрический анализ объекта;
 - параметрическая оптимизация объекта.
4. Получение S -параметров устройства, визуализация распределения электромагнитного поля в продольном сечении тракта.

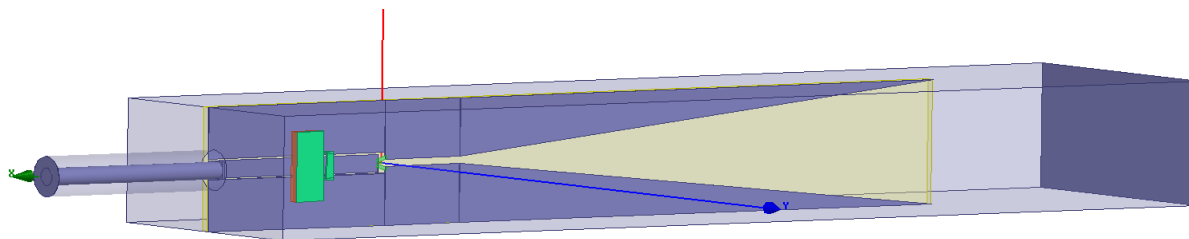


Рис. 4. Модель из программы HFSS

На втором этапе использовалась программа Designer. Процесс проектирования с помощью Designer включает в себя следующее:

1. Создание схемы умножителя с цепями автосмещения (Рисунок 5), в том числе:
 - экспорт решения модели из среды HFSS;
 - задание параметров диодов и элементов цепи.
2. Определение параметров структуры, включающее:
 - определение и калибровка портов;
 - задание параметров решения.
3. Анализ исследуемого объекта, включает:
 - анализ объекта в полосе частот;
 - параметрический анализ объекта;
 - параметрическая оптимизация объекта.
4. Получение уровня мощности гармоник и визуализация их в полосе частот.

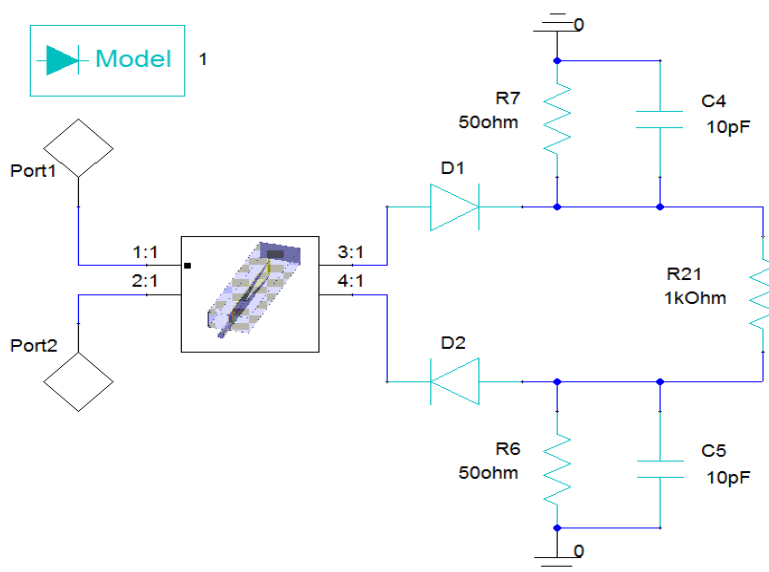


Рис. 5. Модель из программы Designer

4. Результаты анализа

Анализ расчета удвоителя проведен при следующих условиях:

- диапазон частот входного сигнала (13÷20) ГГц;
- уровень мощности входного сигнала 100 мВт (20 дБм);
- диоды предполагались идентичными, соответствующими использованным в экспериментальных образцах A92220-3 с параметрами $C_0=0,2$ пФ, $R_s=3$ Ом (НПО «Салют» г. Н. Новгород), при расчетах использовалась модель D LEVEL=7 из программы Designer;
- номиналы элементов цепей автосмещения (емкость и сопротивление): $C = 10$ пФ, $R = 50$ Ом.

Результатом анализа является рассчитанный уровень мощности гармоник входного сигнала на выходе удвоителя частоты. На рисунке 6 показан уровень мощности второй гармоники при наличии и при отсутствии регулировочных элементов (регулируемые элементы: перемычка или пластина). Размеры элементов и их местоположение выбирались оптимальными, с использованием нескольких итераций для достижения максимального и равномерного уровня мощности в диапазоне частот (26÷40) ГГц.

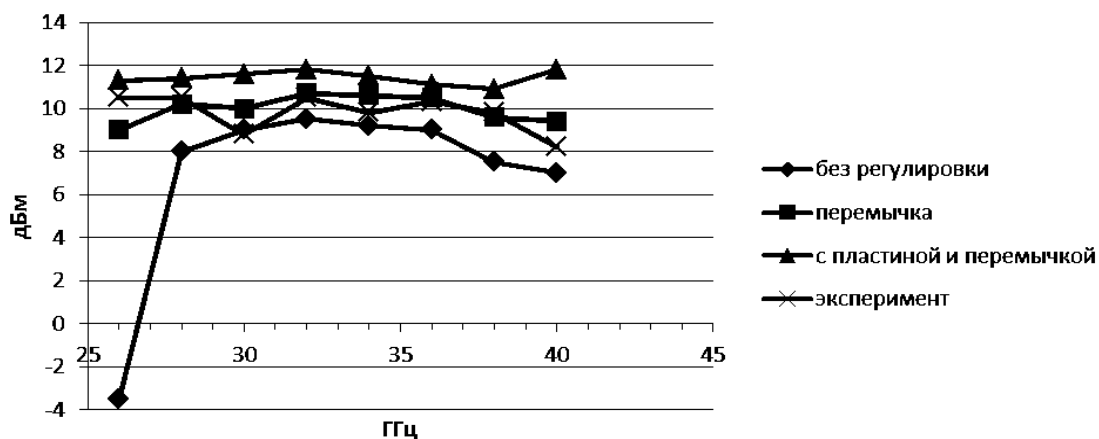


Рис. 6. Уровень мощности 2-й гармоники в зависимости от установленных регулировочных элементов

Из графиков следует, что установка регулировочных элементов при их оптимальном расположении относительно диодов позволяет снизить неравномерность уровня выходной мощности в рассматриваемом диапазоне частот, а также повысить его средний уровень на ~ 3 дБ. Также показана (рис. 6) экспериментальная зависимость уровня мощности, усредненная по десяти образцам. Сравнение показывает, что потенциально выходная мощность экспериментальных образцов может быть повышена как минимум на 2 дБ.

На рисунке 7 показан уровень 4-й гармоники.

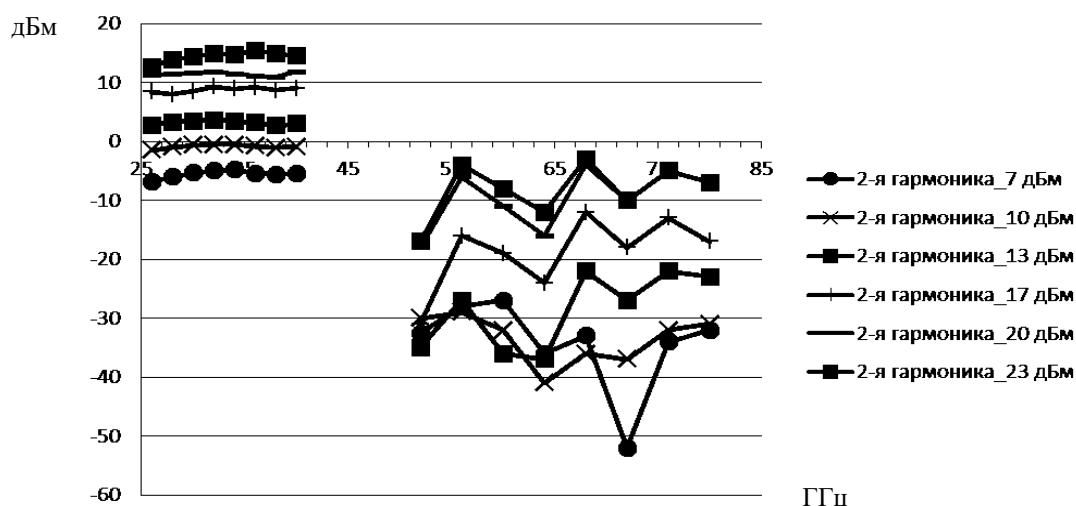


Рис. 7. Уровень 4-й гармоники относительно 2-ой

При $P_{\text{вх.}} = 20$ дБм она подавлена относительно второй на $(17 \div 35)$ дБ, что согласуется с экспериментальными результатами. При уменьшении $P_{\text{вх.}}$ до 10 дБм подавление 4-й гармоники существенно возрастает.

Проанализируем зависимость уровня выходной мощности удвоителя частоты от вариации основных параметров диодов.

На рисунке 8 приведены расчетные характеристики уровня выходной мощности для различных значений C_0 диодов при фиксированном $R_s = 3$ Ом, а на рисунке 9 – для различных R_s при $C_0 = 0,2$ пФ.

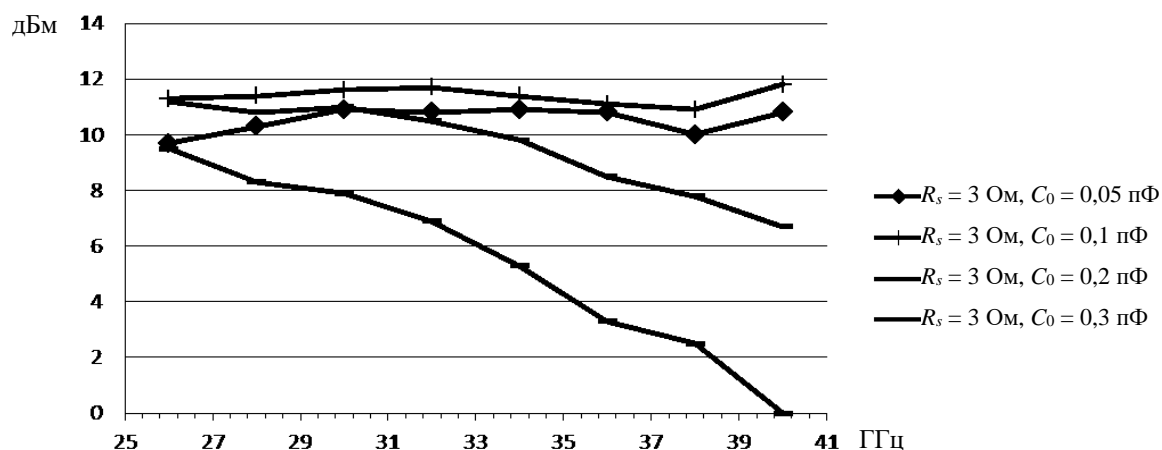


Рис. 8. Уровень выходной мощности при $P_{\text{вх.}}=100$ мВт, $R_s=3$ Ом и различных C_0

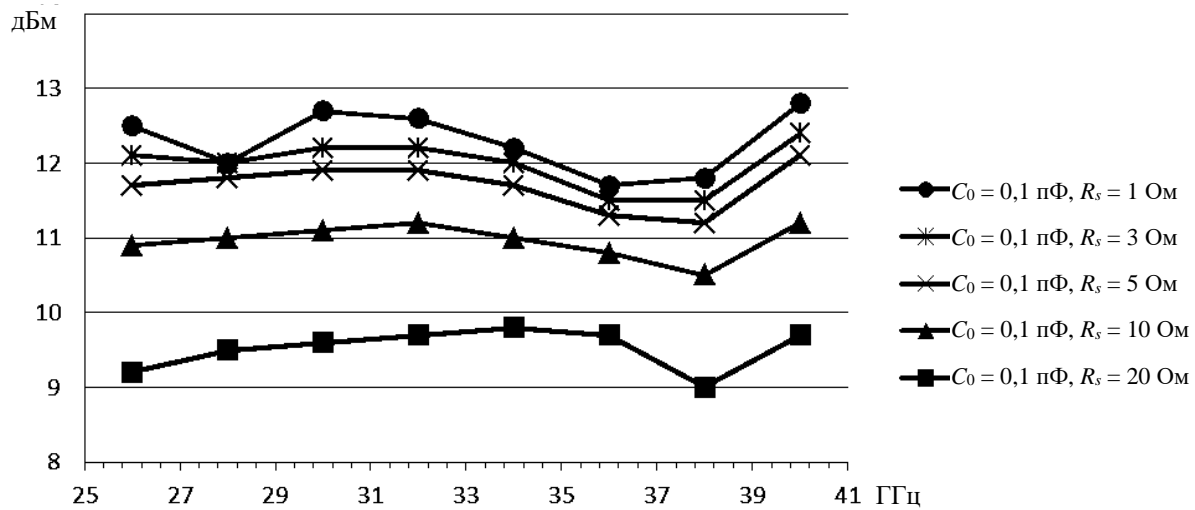


Рис. 9. Уровень выходной мощности при значениях $P_{вх} = 100$ мВт, $C_0 = 0,1$ пФ и различных R_s

Очевидно, что одновременное уменьшение C_0 и R_s для повышения эффективности преобразования умножителя невозможно без снижения допустимой мощности рассеяния диодов. Поэтому оптимальные значения C_0 и R_s для данного диапазона частот удвоителя и максимального уровня входной мощности 100 мВт выбираются из компромисса: $C_0 = 0,1$ пФ и $R_s = 3$ Ом.

При идентичных параметрах диодов в удвоителе частоты нечетные гармоники входного сигнала не генерируются. Однако на практике всегда имеет место разброс параметров диодов, даже подобранных в пары по вольт-амперной характеристике (ВАХ) и вольт-фарадной характеристике (ВФХ). На рисунке 10 даны зависимости уровня мощности третьей гармоники при различной степени неидентичности диодов по C_0 и R_s .

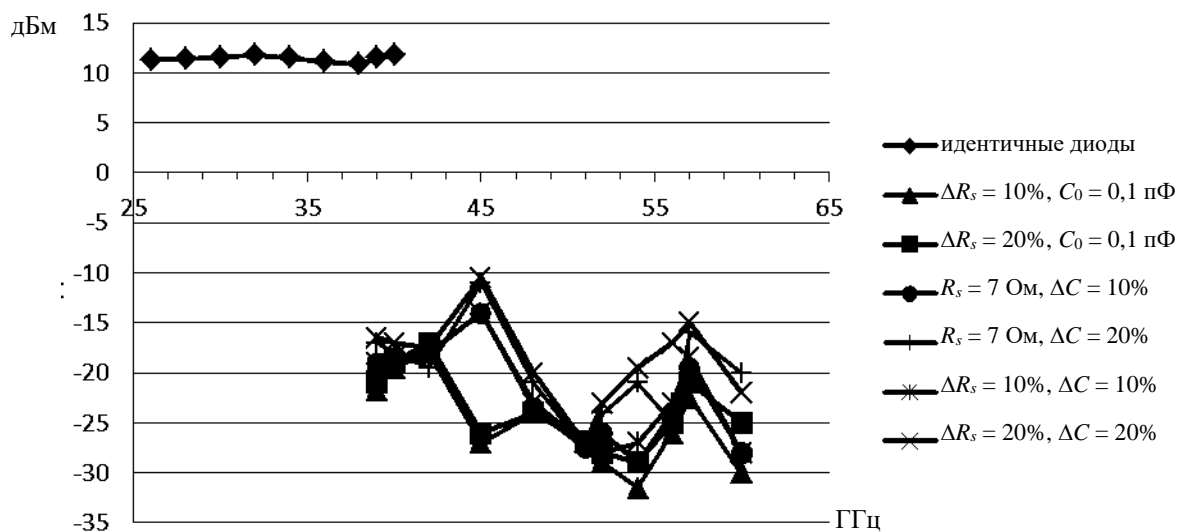


Рис. 10. Уровень третьей гармоники в зависимости от неидентичности параметров диодов по C_0 и R_s

Анализ результатов показывает, что при минимально достижимой на практике неидентичности C_0 и R_s в 10% третья гармоника относительно второй подавлена на 25 дБ. Заметим, что в широкополосном удвоителе частоты 26-40 ГГц третья гармоника не может быть отфильтрована, поскольку находится в рабочем диапазоне частот, поэтому требования к неидентичности параметров диодов весьма высоки.

Эффективность преобразования в значительной степени зависит от характера изменения крутизны вольт-фарадной характеристики (ВФХ) умножительных диодов. Рассмотрены три различных вольт-амперной характеристики: номинальная, соответствующая реальным диодам А92220-3; пологая и крутая (Рисунок 11).

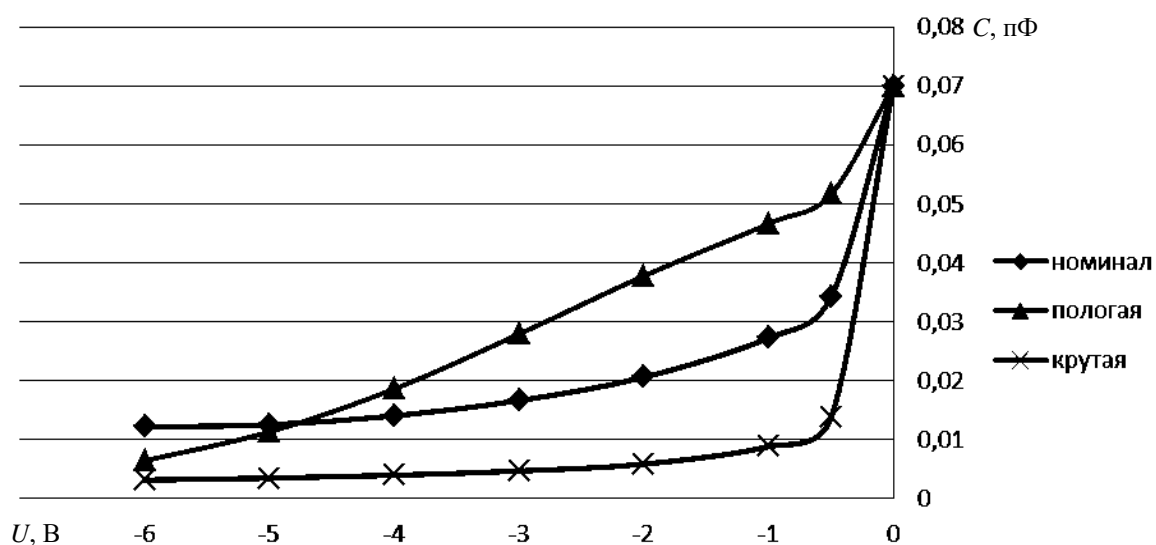


Рис. 11. Семейство вольт-фарадных характеристик диодов

На Рисунке 12 приведены характеристики уровня выходной мощности в зависимости от крутизны ВФХ.

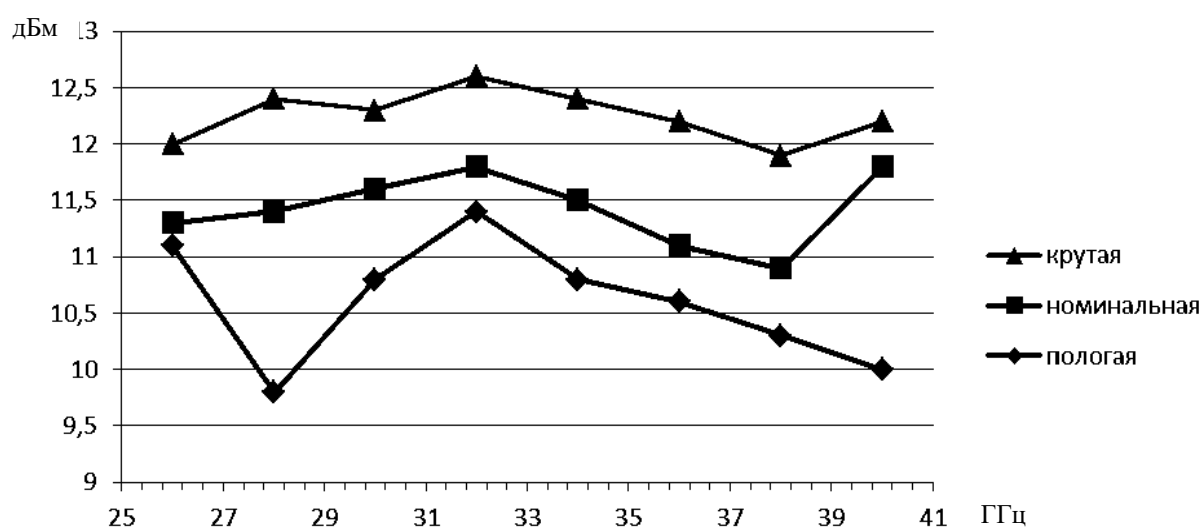


Рис. 12. Уровень выходной мощности в зависимости от крутизны ВФХ

Сравнение результатов показывает, что при более крутой вольт-фарадной характеристике эффективность преобразования повышается примерно на 1,5 дБ.

Важным этапом в проектировании удвоителя частоты является оптимальный выбор номиналов элементов цепей автосмещения диодов R , C . На Рисунке 13 приведены характеристики выходного уровня мощности утроителя при $C = 10$ пФ и различных номиналах R .

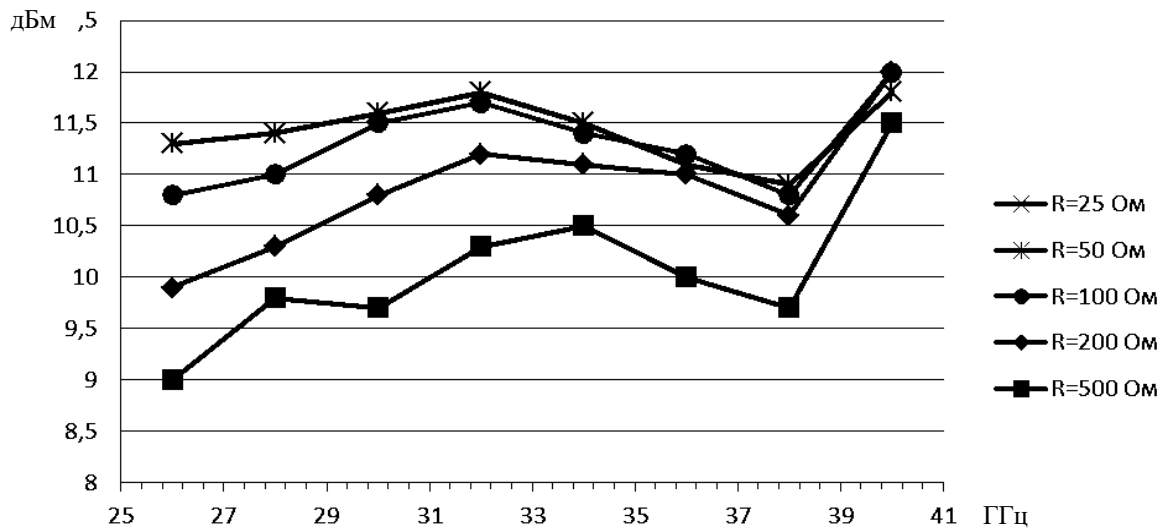


Рис. 13. Графики уровня выходной мощности при фиксированной емкости $C=10$ пФ и различных номиналах сопротивления R

Очевидно, оптимальным значением сопротивления для достижения равномерной $P_{\text{вых}}$ при $P_{\text{вх}}=100$ мВт является 50 Ом.

На рисунке 14 представлены зависимости уровня выходной мощности при $R = 50$ Ом и различных номиналах C .

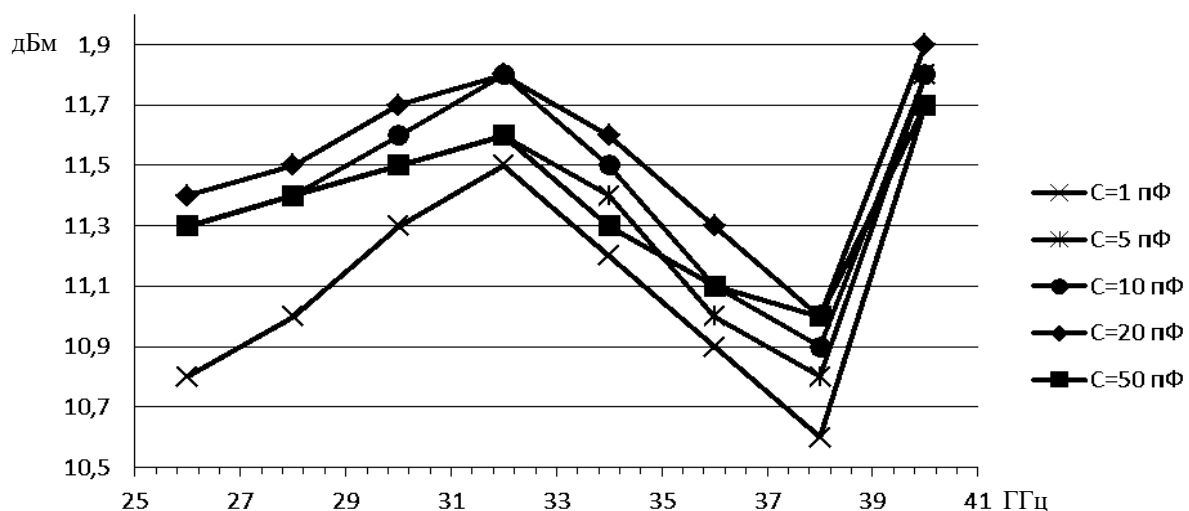


Рис. 14. Графики уровня выходной мощности при фиксированном сопротивлении $R = 50$ Ом и различных номиналах емкости C

При вариации емкости в широких пределах от 5 пФ до 50 пФ результат практически не меняется. Оптимальным значением является $C = 20$ пФ.

Одной из важнейших характеристик умножителей частоты является динамический диапазон входных сигналов. На рисунке 15 даны зависимости уровня мощности на выходе удвоителя частоты от уровня мощности на входе.

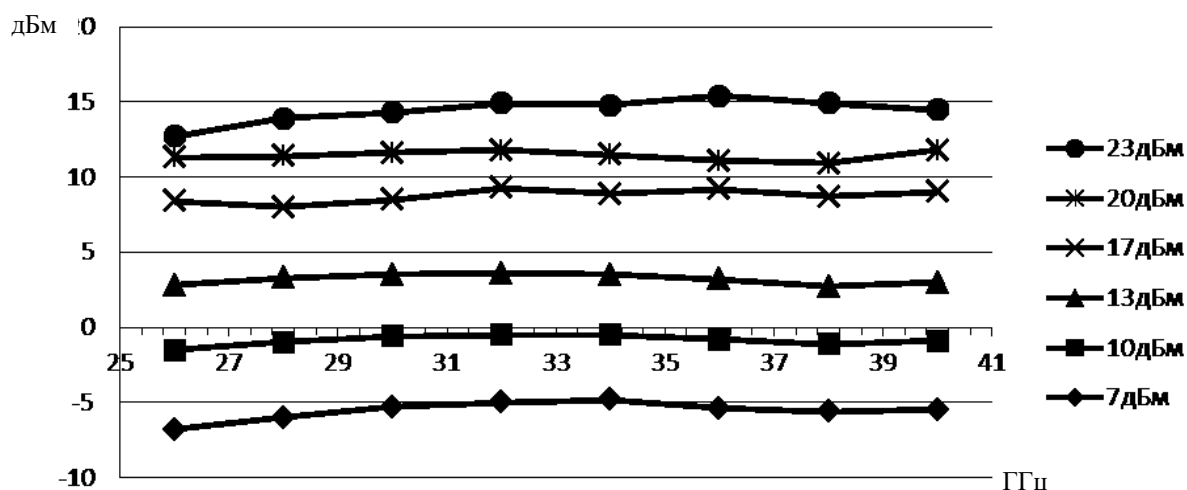


Рис. 15. Мощность выходного сигнала в зависимости от уровня входного сигнала

В диапазоне входных мощностей $P_{вх}=15-23$ дБм зависимость $P_{вых}$ от $P_{вх}$ близка к линейной. С уменьшением $P_{вх}$ линейная зависимость нарушается. Следует заметить, что большой динамический диапазон (не менее 30 дБ) изменения уровня выходной мощности с однозначной зависимостью $P_{вых}$ от $P_{вх}$ позволяет осуществлять регулировку уровня выходной мощности изменением входной. Это обстоятельство существенно упрощает проектирование широкополосных источников сигналов на умножителях частоты с регулируемым и стабилизированным уровнем выходной мощности.

5. Заключение

Хорошее совпадение расчетных и экспериментальных результатов по всем основным характеристикам удвоителя частоты позволяет констатировать, что предложенная в [1] методика может быть распространена для расчета и проектирования волноводных диодных удвоителей КВЧ-диапазона.

Методика позволяет проводить оптимизацию итерационным методом любых схемно-конструктивных элементов умножителя частоты как волноведущих трактов, так и параметров нелинейных элементов (диодов), включая степень и характер нелинейности по вольт-амперной и вольт-фарадной характеристике, и позволяет априорно определять потенциально возможные характеристики нелинейного устройства без проведения дорогостоящих и трудоемких экспериментальных исследований.

Результаты работы доложены на II российско-белорусской научно-технической конференции «Элементная база отечественной радиоэлектроники: импортозамещение и применение» им. О.В. Лосева (Нижний Новгород, 17-19 ноября 2015 г.) и на 52-м Научно-техническом семинаре-миниконференции (Нижний Новгород, 14-15 декабря 2015 г.)

Список аббревиатур

ВАХ – вольт-амперная характеристика

ВФХ – вольт-фарадная характеристика

Список литературы

1. В.В. Березин, А.М. Щитов, С.В. Оболенский. Методика расчета и проектирования волноводного диодного утроителя частоты 75-110 ГГц // XXI Международная научно-техническая конференция «Информационные системы и технологии (ИСТ-2015)». НГТУ им. Р.Е.Алексеева: Материалы конференции. 17 апреля 2015г., г. Нижний Новгород. 2015. С. 62-63.
2. А.М. Щитов, В.В. Березин. Автоматизированный расчет широкополосных волноводных диодных утроителей частоты в диапазонах 75-110 ГГц и 110-170 ГГц // СевГУ: 25-я Международная конференция «СВЧ-техника и телекоммуникационные технологии (Крымико 2015)» 12 сентября 2015г. г. Севастополь. 2015. С. 75-76.
3. А.М. Щитов, А.Е. Серебряков. Широкополосные умножители частоты миллиметрового диапазона (26,5-110 ГГц). – Вестник ВВО АТН РФ, серия Высокие технологии в радиоэлектронике, 1997, №1(3), с. 91-96.
4. А.М. Щитов. Диодные умножители частоты. – Системы и средства связи телевидения и радиовещания, №1, 2, 2002. С. 45-51.

Виталий Витальевич Березин, инженер ОАО «ФНПЦ «ННИИРТ», г. Н. Новгород. Область научных интересов: исследование и разработка СВЧ приборов, теоретические и математические задачи специальных структур для микроволновых измерений, диодные умножители частоты миллиметрового диапазона длин волн.

Аркадий Максимович Щитов, доктор технических наук, ведущий научный сотрудник (ОАО «ФНПЦ «ННИПИ «Кварц» им. А.П. Горшкова», г. Н. Новгород). Область научных интересов: исследование и разработка СВЧ приборов, теоретические и математические задачи специальных структур для микроволновых измерений, диодные умножители частоты миллиметрового диапазона длин волн.

Артем Сергеевич Нечаев, магистрант (Нижегородский государственный технический университет им. Р.Е. Алексеева). Область научных интересов: анализ и разработка интегральных цепей и антенн СВЧ и миллиметрового диапазона.

Computer-aided design of a broadband diode frequency doubler in the range (26÷40) GHz

V.V. Berezin¹, A.M. Schitov², A.S. Nechaev³

¹JSC «FSPC «NSRIR», Nizhny Novgorod

²JSC «FSPC«NSRIMI «Kvarts» n.a. A.P. Gorshkov», Nizhny Novgorod

³Nizhny Novgorod State Technical University n.a. R.E. Alexeev,

24 Minin Str., Nizhny Novgorod, 603950, Russia

Abstract. The calculation methods and design strategy of a wave diode 26-40 GHz frequency doubler based on computer-aided design from ANSYS (HFSS and Designer) is proposed. This strategy combines electrodynamic calculation of waveguide structures and analysis of nonlinear device characteristics. The calculated parameters of a frequency doubler are compared with experimental data.

Key words: frequency doubler, computer-aided design, diode multiplier.

1. Introduction

The calculation and design strategy of a wave diode 75-110 GHz frequency tripler, combining the electrodynamic calculation of waveguide structures and the analysis of nonlinear device characteristics is proposed in [1]. In [2] the strategy is used to design and optimize a frequency tripler ranging from 110 to 170 GHz. Experimental measurement data of conversion loss are in good agreement with the calculation.

In the work the authors make an attempt to apply the calculation and design strategy of a millimeter-wave diode doubler. The goal is to confirm the strategy performance as well as to determine the potential for improving technical characteristics (first of all conversion efficiency) of a previously developed experimental model of a waveguide frequency doubler [3-4].

2. Design features of a frequency doubler

A frequency doubler is designed on scheme (Figure 1) with a full-wave rectifier.

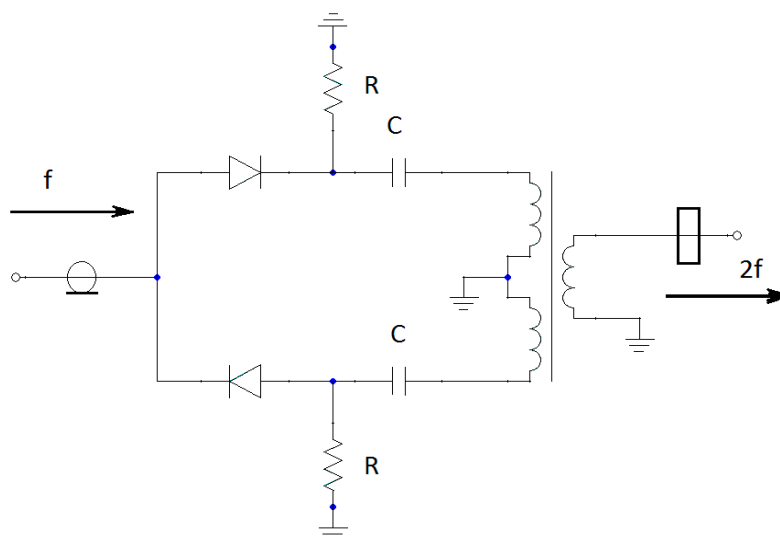


Fig. 1. Analogue circuit of a 26-40 GHz doubler

In terms of design, a frequency doubler is performed as a combination of two transmission lines: the input - a coplanar waveguide and the output – a slotline with a smooth or step junction to the waveguide WR-28 (Figure 2). Lines are fabricated on a thin dielectric substrate placed in a broad-wall section; diodes are placed where lines connect. An input signal is given through a coaxial-to-stripline transition. From the entrance a grounded coplanar waveguide is connected by a jumper link at the distance l from diodes switching. L is quarter-wave length with the average frequency of the output signal. Having a shunting effect l reduces the 4th harmonic level in the output spectrum and equals a half wavelength.

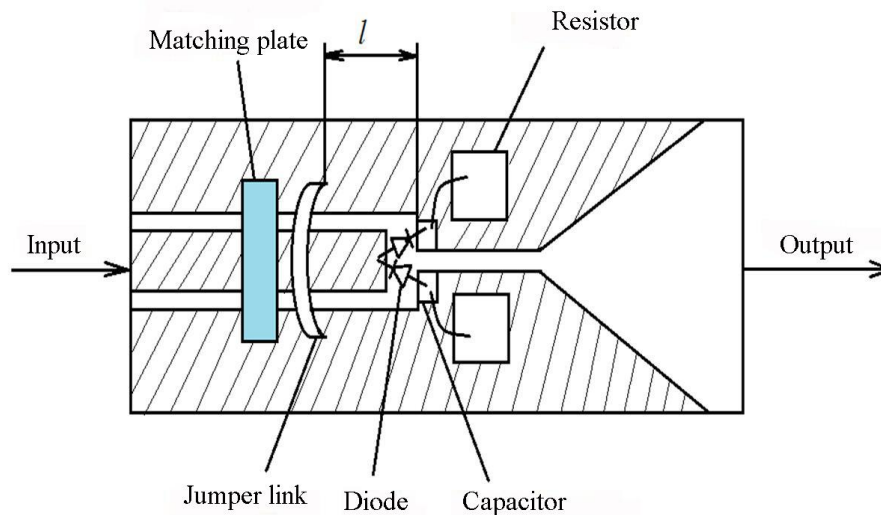


Fig. 2. A 26-40 GHz frequency doubler rough drawing

The waveguide being below-cutoff, the 1st harmonic filtering is carried out by both a balance circuit and an output waveguide. To achieve minimum conversion loss in a given dynamic range of input signals external bias or internal self-bias circuits are used. So far there is no a detailed theoretical analysis and frequency multiplier calculations in the scientific literature. Physical form of a (26÷40) GHz frequency doubler is shown in Figure 3.

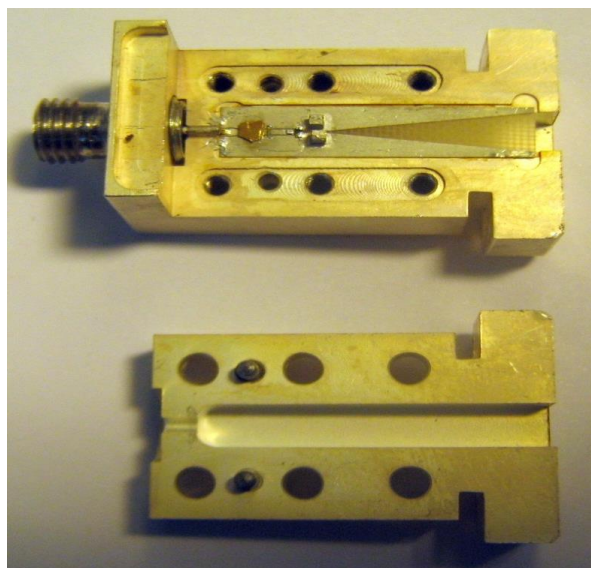


Fig. 3. Physical form of a 26-40 GHz frequency doubler

More important technical characteristics of broadband frequency multipliers are: conversion efficiency (loss), the maximum attainable power output, power output imbalance within the frequency range and the spurious components level in output spectrum. The goal of the article is improvement and optimization of these frequency doubler characteristics.

3. Frequency doubler simulation

Computer-aided design from Ansys (HFSS and Designer) was used for doubler simulation. HFSS was used at the first stage of frequency doubler simulation.

HFSS simulation includes a number of standard steps:

1. Making the model of doubler (Figure 4), including:
 - making a 3D model (drawing);
 - setting material parameters.
2. Identification the electrodynamic doubler parameters, comprising:
 - setting boundary conditions on the surfaces forming the test subject;
 - ports definition and calibration;
 - setting decisions.
3. Electrodynamic analysis of the investigated doubler, including:
 - analysis in a frequency band;
 - parametric analysis of the doubler ;
 - its parametric optimization.
4. Obtaining device S-parameters, electromagnetic field imaging in a longitudinal cross section of a waveguide.

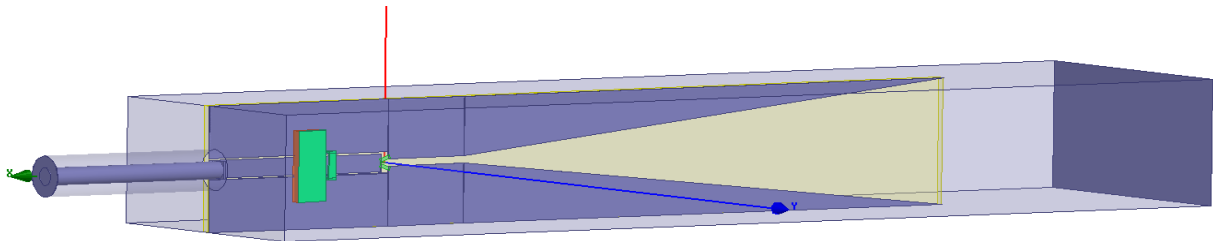


Fig. 4. The model from HFSS

Designer was used at the second stage. Designer simulation includes the following:

1. Making the multiplier model with autobias circuits (Figure 5), including:
 - export the model from HFSS;
 - setting the diodes and circuit elements parameters.
2. Identification the multiplier parameters, comprising:
 - ports definition and calibration;
 - setting decisions.
3. Analysis of the test multiplier includes:
 - analysis in a frequency band;
 - parametric analysis of the multiplier;
 - its parametric optimization.
4. Obtaining the harmonics power level and their imaging in a frequency band.

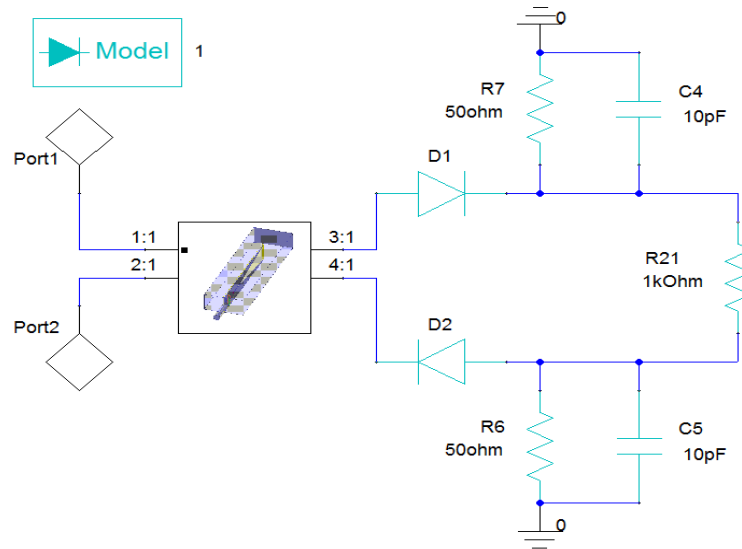


Fig. 5. The model from Designer

4. Experimental data

The analysis of frequency doubler was carried out in the following conditions:

- band input signal frequency range – (13÷20) GHz;
- input signal power - 100 mW (20 dBm);
- identical diodes A92220-3 $C_0=0,2$ pF, $R_s=3$ Ohm (R&D enterprise "Salut" N. Novgorod); D LEVEL =7 was used from Designer;
- values of circuits autobias: $C = 10$ pF, $R = 50$ Ohm.

The result is the calculated harmonics level of the input signal in frequency doubler output. Figure 6 shows the second harmonic level with and without the adjusting elements. The dimensions of the elements and their optimal location were fully satisfactory and chosen with several iterations to achieve maximum and uniform power level in the frequency range of 26-40 GHz.

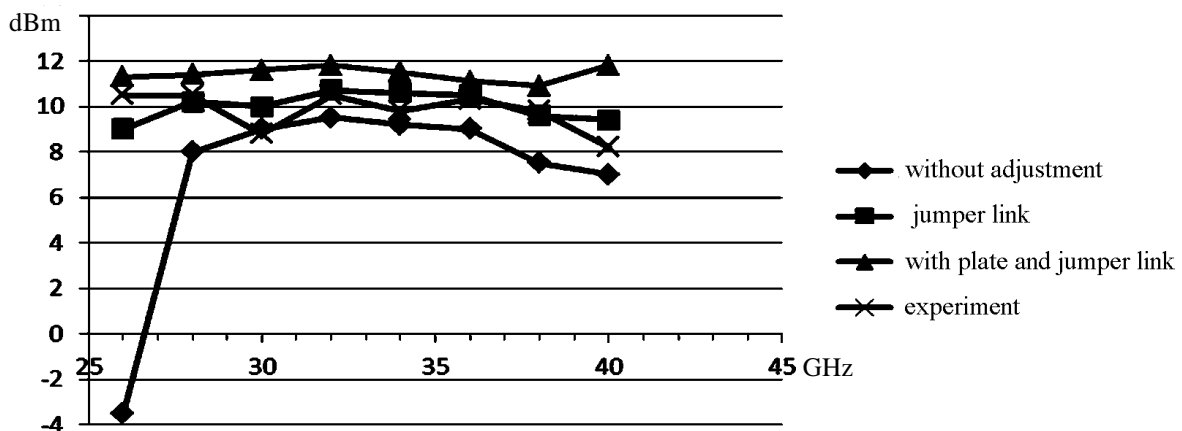


Fig. 6. The 2nd harmonic level depending on fixed adjusting elements

According to the curves, installation of the adjusting elements when they have optimal location relative to diodes allows to reduce an output power level imbalance in a frequency

range as well as to improve its average level by ~ 3 dB. Also figure 6 shows the experimental dependence of power level. This experimental dependence is ten-examples averaged. The comparison shows that the potential output power of experimental examples can be increased by at least 2 dB.

Figure 7 shows the 4th harmonic level.

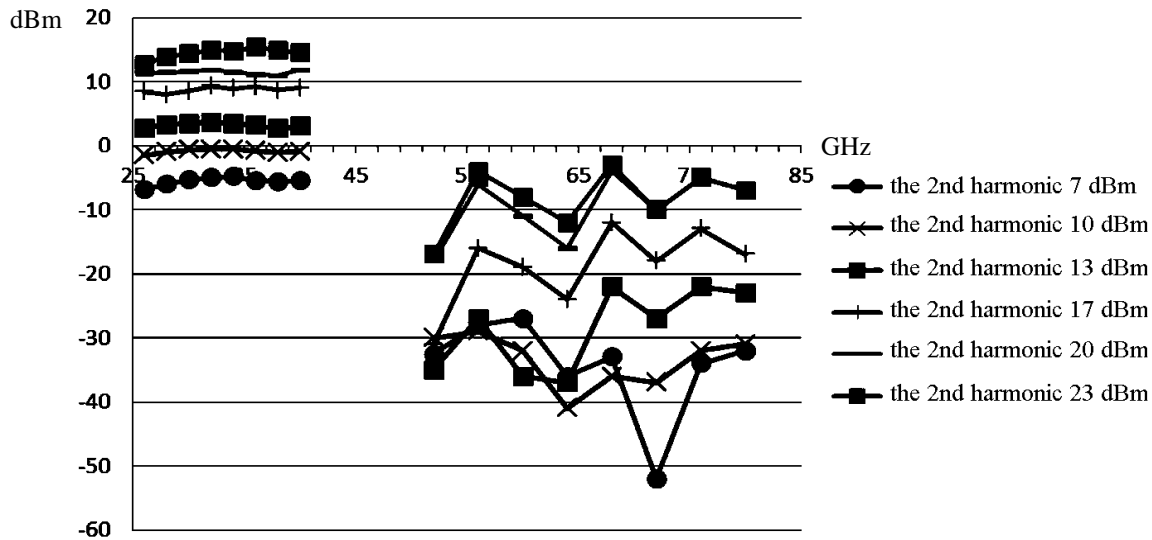


Fig. 7. The level of the 4th harmonic in comparison with the level of the 2nd harmonic

When $P_{in.} = 20$ dBm, the level of the 4th harmonic is 17-35 dB less than the level of the 2nd harmonic. Experimental data show it. If $P_{in.}$ reduces up to 10 dBm, the 4th harmonic cancellation increases significantly.

Let us consider a doubler output power level versus variations of the main diodes parameters. Figure 8 shows the characteristics lines of output power level for different values of C_0 diodes when $R_s = 3$ Ohm, a in Figure 9 – for various R_s when $C_0 = 0,2$ pF.

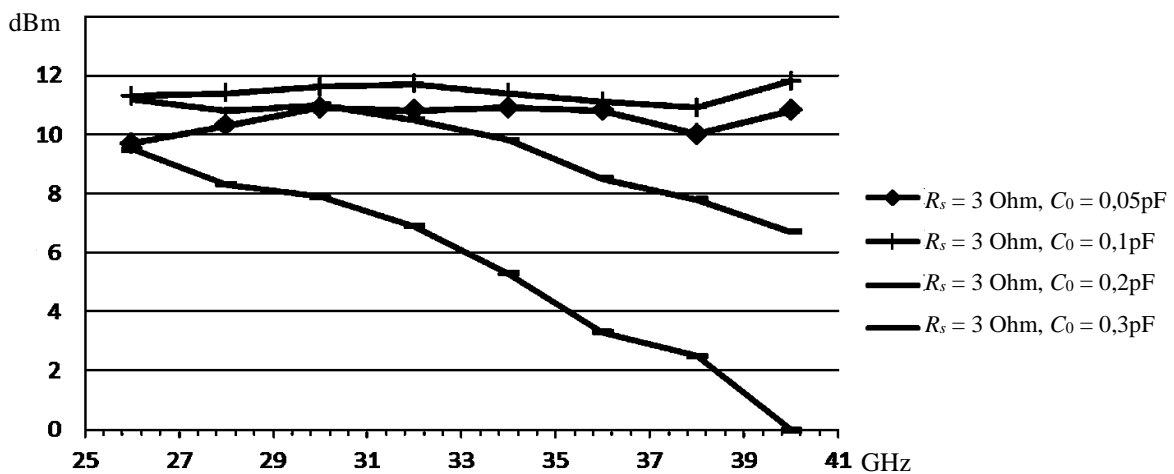


Fig. 8. Output power level when $P_{in}=100$ mW, $R_s=3$ Ohm and various C_0

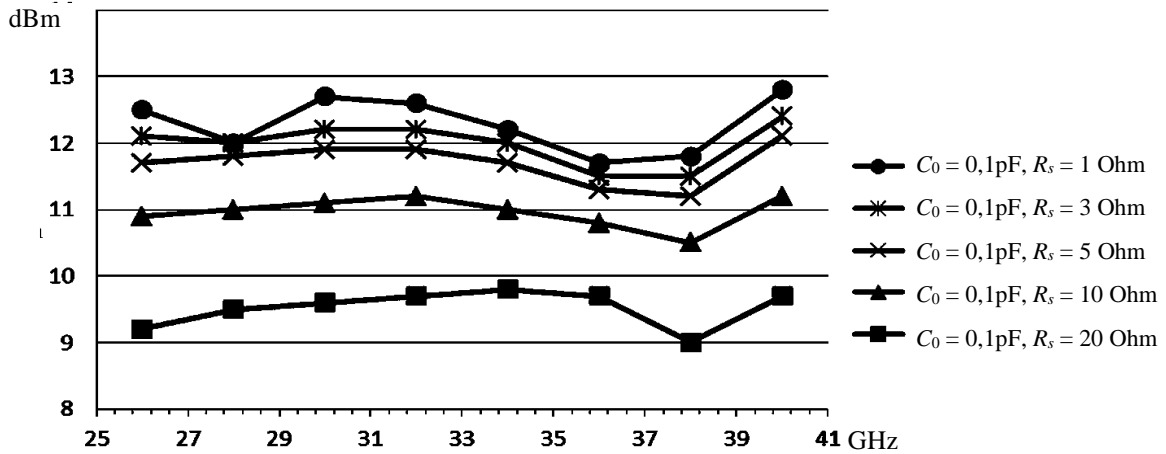


Fig. 9. Output power level when $P_{in}=100 \text{ mW}$, $C_0=0,1 \text{ pF}$ and various R_s

It is obvious the following: to decrease C_0 , R_s and enhance multiplier conversion efficiency at the same time it is necessary to reduce power diode dissipation. Therefore, the optimal values of C_0 and R_s for a given frequency range of a doubler and maximum input power (100 mW) come out of: $C_0 = 0,1 \text{ pF}$ and $R_s = 3 \text{ Ohm}$.

When diodes parameters in a frequency doubler are identical, odd harmonics of an input signal are not generated. However, in practice there is a variability of diodes parameters, even pair matched in terms of volt-ampere characteristic (VAC) and capacity-voltage characteristic (CVC). Figure 10 shows the curves of the power levels of the third harmonic when there are different degrees of diodes mismatch on C_0 and R_s .

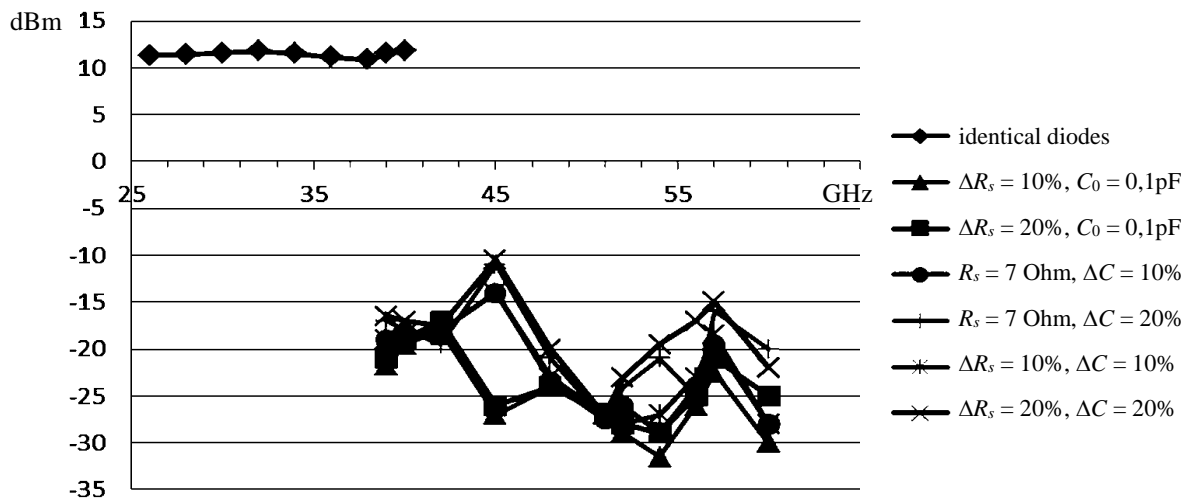


Fig. 10. The level of the 3rd harmonic versus diodes mismatch on C_0 and R_s

When mismatch of C_0 and R_s is 10% in practice, the 3rd harmonic relative to the second one is suppressed by 25 dB based on the results of the analysis performed. It is necessary to point out that the 3rd harmonic in a 26-40 GHz broadband frequency doubler cannot be filtered, since it is located within the operating band, so diodes mismatch requirements is very high.

Conversion efficiency (loss) relies heavily on capacitance-voltage curve of multiplier diodes. Three different volt-ampere characteristic are investigated: nominal conforming to real A92220-3 diodes; flat and abrupt (Figure 11).

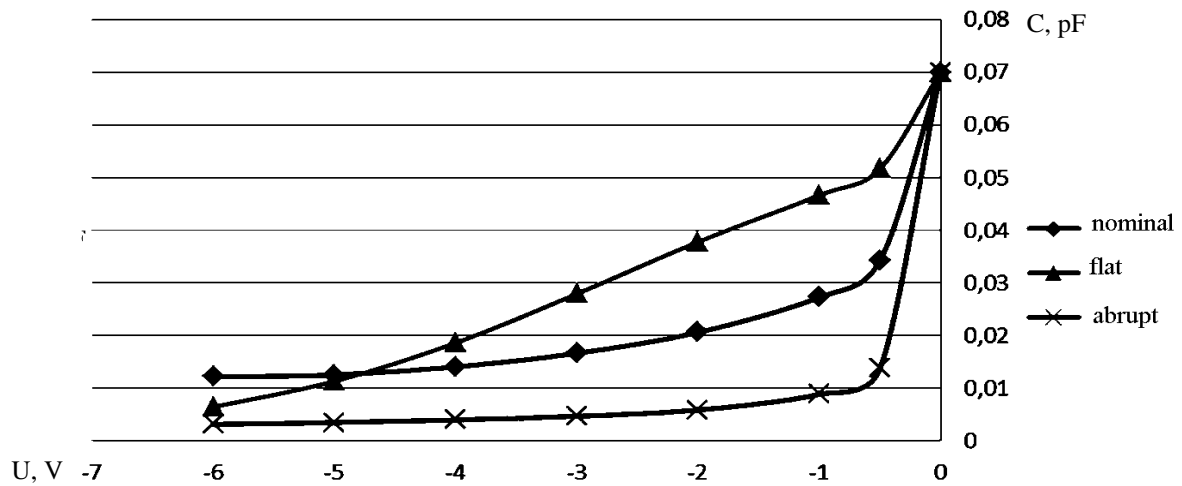


Fig. 11. Capacitance-voltage curve of multiplier diodes

Figure 12 shows the characteristic lines of output power level, which depend on capacitance-voltage curve.

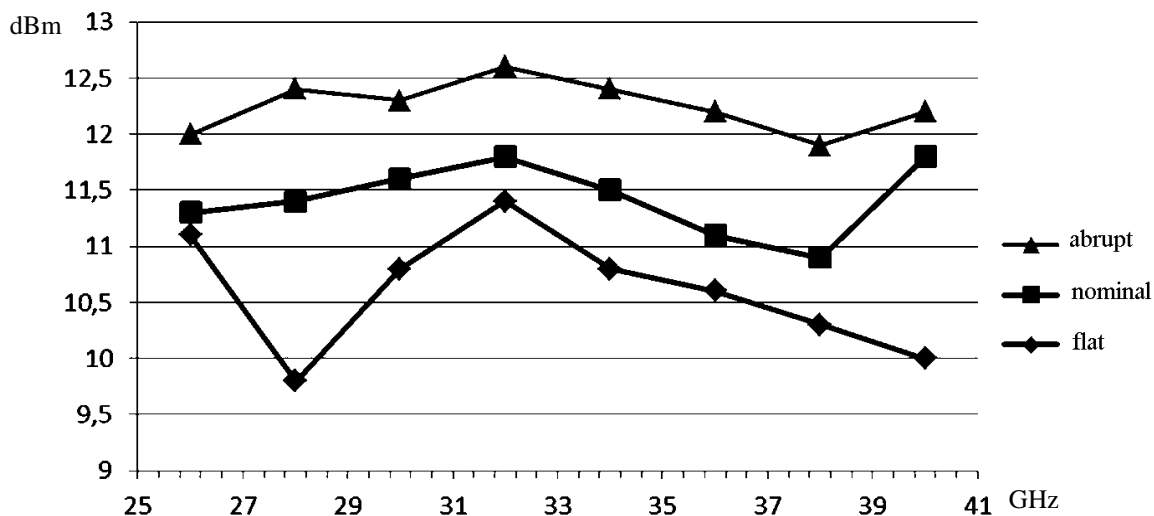


Fig. 12. Output power level depending on capacitance-voltage curve

Comparison of the results shows that more abrupt capacity-voltage characteristic allows to increase conversion efficiency (loss) by about 1.5 dB.

Premier choice of specific autobias circuits of R, C diodes is an important step in frequency doubler design. Figure 13 shows the curves of output power level for a tripler when $C = 10$ pF and different values of R.

Obviously, the optimum resistance to achieve even P_{out} when $P_{in}=100$ mW is 50 Ohm.

Figure 14 shows curves of output power level when $R = 50$ Ohm and different values of C.

If capacity variations are long-range (from 5 pF to 50 pF, the result scarcely change). Optimum is when $C = 20$ pF.

Dynamic range of an input signal is an essential characteristic of frequency multiplier. Figure 15 shows a power level in output of a doubler versus power level in input.

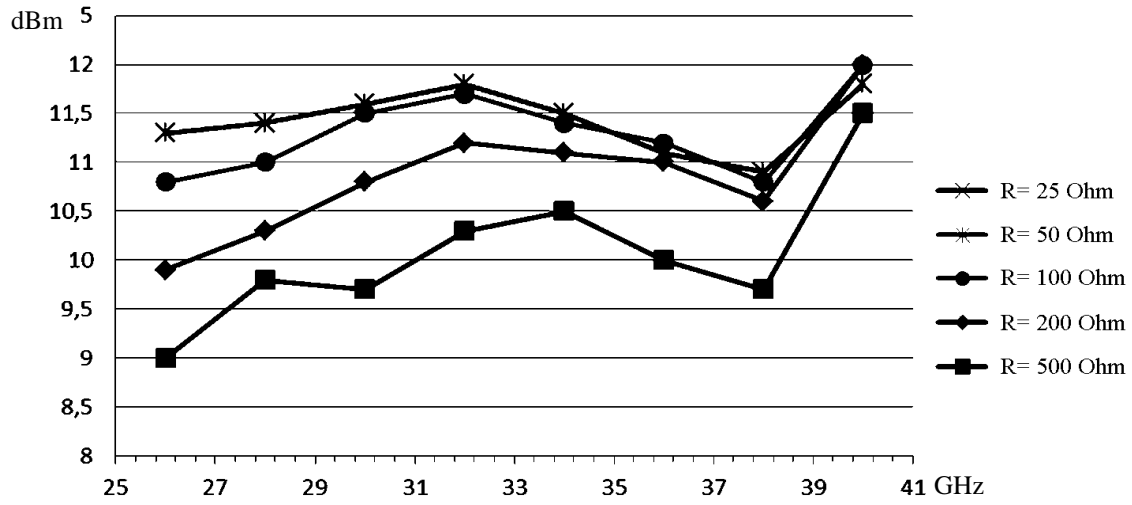
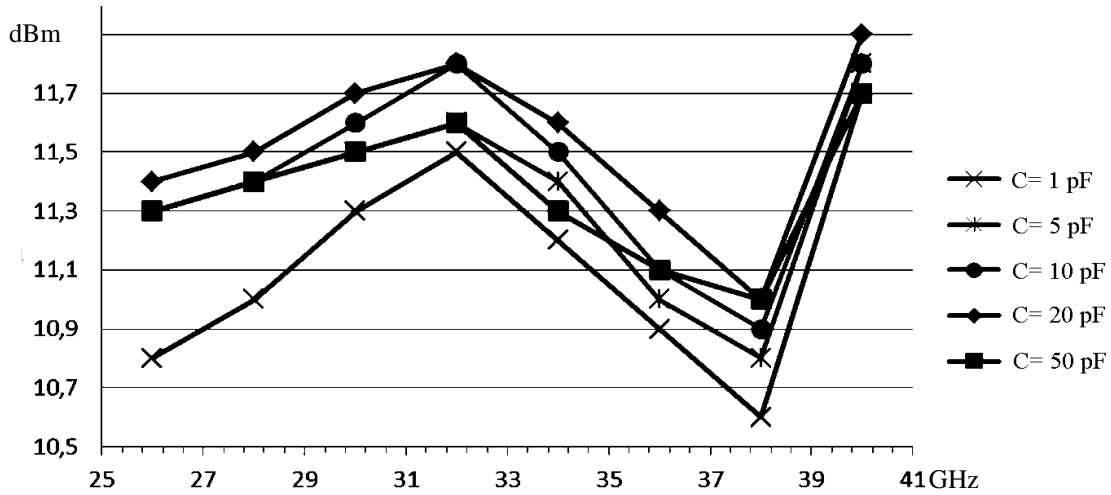
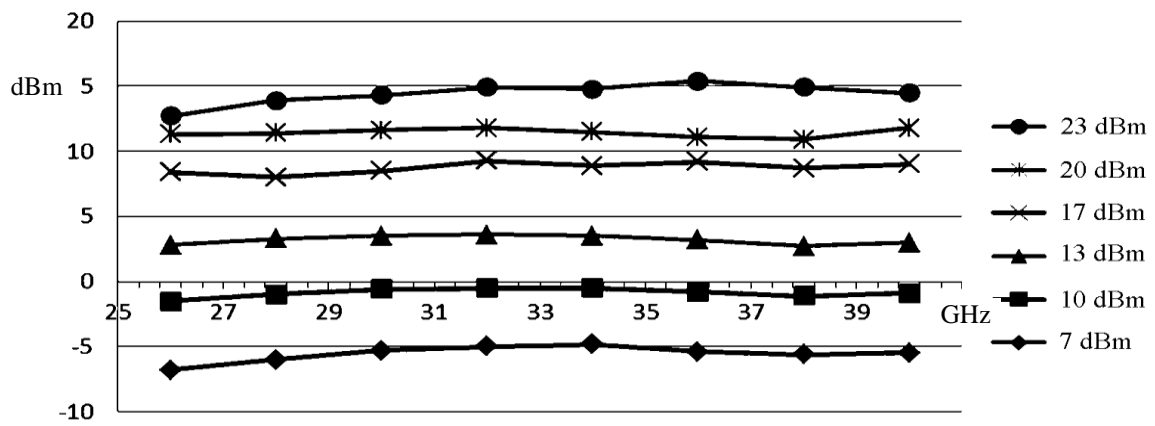
Fig. 13. Curves of output power level when $C=10$ pF and different values of R Fig. 14. Curves of output power level when $R = 50$ Ohm and different values of C 

Fig. 15. Output signal power versus an input signal level

If $P_{in}=15-23$ dBm, P_{out} versus P_{in} is near-linear. P_{in} decreased, linear dependence is violated. It should be noted that a large dynamic output power range (not less than 30 dB) with P_{out} versus P_{in} allows to adjust output power level by changing the input. This fact significantly

simplifies the construction of broadband signal sources in frequency multipliers with an adjustable and stabilized output power level.

5. Conclusion

The good agreement between the calculated and experimental results for all main characteristics of a frequency doubler goes to show that the proposed strategy in [1] can be used to calculate and design millimeter-wave diode doublers.

The iteration method of the strategy allows to realize optimization of any circuit-designed components of multiplier both waveguides and the parameters of nonlinear diode elements including the degree and phenomena of nonlinearity according to volt-ampere characteristic and capacity-voltage characteristic. Also the strategy allows to determine the potential characteristics of nonlinear devices without high-priced and time-consuming experiments of a priori.

The results of the work were presented at the II Russian-Belarusian scientific-technical conference "Hardware components of domestic radioelectronics: imports phase-out and use" n.a. O.V. Losev (Nizhny Novgorod, from 17 to 19 November 2015) and at the 52nd scientific-technical conference (Nizhny Novgorod, Russia, 14-15 December 2015)

Abbreviation list

VAC – volt-ampere characteristic

CVC – capacity-voltage characteristic

References

1. V.V. Berezin, A.M. Schitov, S.V. Obolensky. Metodika rascheta i proektirovaniya volnovodnogo diodnogo utroitelya chastoty 75-110 GGts [The calculation and design strategy of a wave diode 75-110 GHz frequency tripler] // the XXIst International scientific-technical conference «Information systems and technologies (IST-2015)». NGTU n.a. R.E. Alexeev: Proceedings. 17 april 2015, N. Novgorog. 2015. pp. 62-63. (In Russian)
2. A.M. Schitov, V.V. Berezin. Avtomatizirovannyi raschet shirokopolosnykh volnovodnykh diodnykh utroiteley chastoty v diapazonakh 75-110 GGts i 110-170 GGts [Automated analysis of broadband wave diode 75-110 GHz and 110-170 GHz frequency triplers] // SevGu: the 25th International conference «Microwave technology and telecom technology (Ktimiko 2015)» 12 september 2015г., Sevastopol. 2015. p.p. 75-76. (In Russian)
3. A.M. Schitov, A.E. Serebrykov. Shirokopolosnye umnozhiteli chastoty millimetrovogo diapazona (26,5-110 GGts) [Broadband millimeter-wave multiplier (26,5-110 GHz)]. – Reporter VVO ATN RF, серия High technologies in radioelectronics, 1997, №1(3), p.p. 91-96. (In Russian)
4. A.M. Schitov. Diodnye umnozhiteli chastoty [Diode frequency multipliers]. – Systems and means of tele- and radiocommunication, №1, 2, 2002, p.p. 45-51. (In Russian)

Vitaly Vital'evich Berezin, Engineer in JSC «FSPC «NSRIR», Nizhny Novgorod. His research interests include the analysis and design of various microwave devices, theoretical and mathematical problems of special structures for microwave measurement techniques, diode frequency multipliers in millimeter waveband.

Arkady Maksimovich Schitov, Leading Researcher in JSC «FSPC«NSRIMI «Kvarts» n.a. A.P. Gorshkov», Nizhny Novgorod (Doctor of Engineering). His research interests include the analysis and design of various microwave devices, theoretical and mathematical problems of special structures for microwave measurement techniques, diode frequency multipliers in millimeter waveband.

Artem Sergeevich Nechaev, undergraduate of Nizhny Novgorod State Technical University n.a. R.E. Alexeev. His research interests include the analysis and design of microwave and millimeter-wave integrate circuits and antenna.

Моделирование нелинейного распространения мод в многомодовых волоконно-оптических линиях дальней связи

В.А. Бурдин, А.В. Бурдин

Поволжский государственный университет телекоммуникаций и информатики
Россия, Самара, Московское шоссе 77, 443090

Аннотация. Предложена модификация метода расщепления по физическим процессам для численного решения системы связанных нелинейных уравнений Шредингера. Представлен пример результатов моделирования маломодовой передачи в кусочно-регулярной многомодовой волоконно-оптической линии дальней связи с помощью этого метода.

Ключевые слова: нелинейное уравнение Шредингера, связанные нелинейные уравнения Шредингера, эрбиевый усилитель, кусочно-регулярная модель, маломодовая волоконно-оптическая линия связи, теория связи мод, Гауссово приближение, метод расщепления по физическим процессам.

1. Введение

В современных системах связи для высокоскоростных оптических каналов со скоростью передачи 100 Гбит/с проблемы дисперсионных искажений практически полностью решаются методами электронной компенсации в сочетании с применением устойчивых к дисперсии форматов модуляции, оптимизации кодирования, демодуляции и коррекции ошибок. В этом случае основным фактором ошибок в оптическом канале становится нелинейность [1-7]. Устранение нелинейных искажений оптического канала для оптических одномодовых волокон является чрезвычайно сложной задачей [5-11]. Известно, что одним из кардинальных способов снижения влияния нелинейных эффектов является увеличение площади поперечного сечения сердечника оптического волокна. В частности, для этой цели предлагается использовать так называемые маломодовые оптические волокна [8, 12-15]. С этой точки зрения, также представляют интерес многомодовые оптические волокна, функционирующие в маломодовом режиме. В этом случае площадь поперечного сечения больше и, соответственно, – больше степень снижения действия факторов нелинейности [8, 6-18].

Распространение оптических импульсов в волоконных световодах обычно описывается с помощью нелинейного уравнения Шредингера [19, 20]. Для маломодового и многомодовых оптических волокон, этот процесс описывается системой связанных нелинейных уравнений Шредингера (CNLS) [19-25]. Различные формы записи системы различаются в основном способами определения линейной и нелинейной связей мод. В общей постановке задачи система M - связанных нелинейных уравнений Шредингера решается численно – методом расщепления по физическим процессам [19-25]. При использовании традиционного метода расщепления по физическим процессам для решения нелинейного уравнения Шредингера предполагается, что параметры оптического волокна не зависят от параметров передаваемого сигнала [19, 20]. Тем не менее очевидно, что при распространении нескольких мод в результате влияния нелинейности происходит перераспределение мощности по сечению и изменение показателя преломления профиля оптического волокна. Это приводит к изменениям параметров мод и параметров модовой связи в оптическом волокне под действием оптического импульса.

В работе предлагается модификация метода расщепления по физическим процессам для решения системы М - уравнений, описывающих нелинейное распространение нескольких мод в оптических волокнах. В качестве примера представлены результаты моделирования на основе предложенной модификации для волоконно-оптической линии с многомодовым оптическим волокном, функционирующим в маломодовом режиме. Для сравнения также приведены результаты моделирования с применением данной модификации метода для линии связи с одномодовым оптическим волокном.

2. Система связанных нелинейных уравнений Шредингера и метод расщепления по физическим процессам

Процесс маломодового распространения оптических импульсов в линии обычно описывают с помощью системы уравнений типа нелинейных уравнений Шредингера [19-25]. Эта система может быть записана в виде [21-22]:

$$j \frac{\partial \Psi_{l,m}}{\partial \eta} + j \beta_{1l,m} \frac{\partial \Psi_{l,m}}{\partial \tau} + \beta_{2l,m} \frac{\partial^2 \Psi_{l,m}}{\partial \tau^2} + j \bar{\alpha}_{l,m} \Psi_{l,m} + \sum k_{l,m;p,q} |\Psi_{p,q}|^2 \Psi_{l,m} = 0, \quad (1)$$

$$\eta = \frac{z}{L_0}; \quad \tau = \frac{(t - z/v_0)}{T_0};$$

$$\bar{\alpha}_{l,m} = \alpha_{l,m} L_0; \quad \beta_{2l,m} = \frac{L_0}{2T_0} \frac{\partial v_{l,m}^{-1}}{\partial \omega}; \quad \beta_{1l,m} = \frac{L_0}{T_0} \left(\frac{1}{v_{l,m}} - \frac{1}{v_0} \right);$$

$$k_{l,m;p,q} = K_{l,m}(E_0) \cdot B_{l,m;p,q}; \quad K_{l,m} = \frac{w_0^2 n \cdot n_2 L_0 E_0^2}{c^2 \beta_{l,m} B_{l,m;l,m}},$$

здесь $\Psi_{l,m}$ – нормированная комплексная огибающая l, m – моды;

$\alpha_{l,m}$ – коэффициент затухания l, m – моды;

$v_{l,m}$ – групповая скорость распространения l, m – моды;

$\beta_{l,m}$ – постоянная распространения l, m – моды;

n, n_2 – невозмущенное значение показателя преломления и его коэффициент нелинейности;

$\beta_{1l,m}, \beta_{2l,m}$ – нормированные обратные значения групповой скорости и параметра дисперсии l, m – моды;

$B_{l,m;p,q}$ – коэффициент связи l, m and p, q – мод;

l, m – азимутальный и радиальный порядок моды, соответственно.

Согласно методу расщепления по физическим процессам (SSFM) уравнение (1) формально можно представить в виде:

$$\frac{\partial \Psi_{l,m}}{\partial \eta} = (\hat{D} + \hat{N}) \Psi_{l,m}, \quad (2)$$

где \hat{D} – дифференциальный оператор, который учитывает дисперсию и потери в линейной среде, а \hat{N} – нелинейный оператор, который учитывает влияние нелинейностей волокна на распространение импульса. Вообще дисперсия и нелинейность действуют на длине волокна совместно. Однако при использовании метода расщепления по физическим процессам предполагается, что при распространении оптического импульса на небольшом расстоянии h дисперсионные и нелинейные эффекты действуют независимо друг от друга. При этом моделирование распространения от точки z до точки $z+h$ выполняется в два этапа. На первом этапе

учитывают только нелинейность и $\hat{D}=0$. На втором этапе учитывают только дисперсию и $\hat{N}=0$. Оператор \hat{D} может быть выполнен в частотной области. Использование для этого алгоритма быстрого преобразования Фурье (FFT) позволяет получать численные оценки с помощью оператора \hat{D} относительно быстро. Известно две схемы метода расщепления по физическим процессам – несимметричная и симметричная. Обычно симметричная схема метода дает более точные результаты. Применение описанного выше традиционного варианта метода расщепления предполагает, что параметры мод и параметры связей мод не зависят от параметров передаваемого сигнала. Очевидно, что это не так, и при анализе процессов распространения в многомодовых и маломодовых волокнах эта зависимость должна быть учтена. Здесь предложена модификация метода расщепления по физическим процессам, в которой этот недостаток частично устранен.

3. Модифицированный метод расщепления по физическим процессам

Предположим, что волоконно-оптическая линия передачи может быть описана с помощью кусочно-регулярной модели и, таким образом, в рамках отдельного регулярного участка линии передачи параметры оптического волокна и параметры распространяющихся мод являются постоянными. Предположим, что параметры мод можно вычислить с помощью модифицированного приближения Гаусса [26]. Тогда в случае линейного распространения моды для произвольного профиля показателя преломления оптического волокна уравнение для эквивалентного радиуса пятна моды $R_{l,m}$ имеет вид:

$$\left\{ \sum_{i=1}^N k^2 \Delta n_i^2 r_i^2 \left[F_m^{(l)}(x_i) \right]^2 \right\} - C = 0, \quad (3)$$

и постоянная распространения моды описывается соотношением:

$$\beta_{l,m}^2 = k^2 n_0^2 - \frac{m!}{(l+m-1)!} \left\{ \left[\sum_{i=1}^N k^2 \Delta n_i^2 \Phi_m^{(l)}(x_i) \right] + \frac{C}{a^2 \cdot R_{l,m}^2} \right\}, \quad (4)$$

$$C = m \frac{(m+l-1)!}{(m-1)!} + (m+l-1) \frac{(m+l-2)!}{(m-2)!} + 2l^2 \sum_{i=0}^{m-2} \frac{(i+l-1)!}{i!},$$

Здесь

$$\left[F_m^{(l)}(x) \right]^2 = x^l \left[L_{m-1}^{(l)}(x) \right]^2 \exp(-x),$$

$$\Phi_m^{(l)}(x) = \exp(-x) \cdot \sum_{q=0}^{2n} g_q \cdot \left[x^q + \sum_{k=1}^{q+l} \frac{(q+l)!}{(q+l-k)!} x^{q+l-k} \right],$$

$$g_q = \sum_{j=0}^q b_j \cdot b_{q-j}; \quad b_q = \frac{(l+m)!}{q!(m-q)!(l+q)!}$$

$$\Delta n_i^2 = n_i^2 - n_{i+1}^2; \quad x = R^2 / R_{l,m}^2; \quad R = r/a$$

где $L_p^{(l)}(x)$ – полиномы Лагерра;

a – радиус сердцевины;

r – радиальная координата.

Известно, что $n_E(R) = n(R) + n_2 |E(R)|^2$. Мы предполагаем, что $n_2 |E(R)|^2 \ll n(R)$.

Тогда $n_E^2(R) \approx n^2(R) + 2 \cdot n(R) \cdot n_2 |E(R)|^2$. Отсюда для нелинейного распространения получаем систему уравнений для радиусов пятна мод:

$$\left\{ \sum_{i=1}^N k^2 \Delta n_i^2 r_i^2 \left[F_m^{(l)}(x_i) \right]^2 \right\} + \left\{ \sum_{i=1}^N k^2 n_2 E_0^2 \Delta n_i r_i^2 \left[F_m^{(l)}(x_i) \right]^2 \cdot \sum_{p=0}^P \sum_{q=0}^Q B_{l,m;p,q} \left[F_q^{(p)} \left(\frac{R_{l,m}^2}{R_{p,q}^2} x_i \right) \right]^2 \right\} - C = 0 \quad (5)$$

Формула для постоянной распространения принимает следующий вид:

$$\beta_{l,m}^2 = k^2 n_0^2 - \frac{m!}{(l+m-1)!} \left\{ \left[\sum_{i=1}^N k^2 \Delta n_i^2 \Phi_m^{(l)}(x_i) \right] + \left[\sum_{i=1}^N k^2 n_2 E_0^2 \Delta n_i \sum_{p=0}^P \sum_{q=0}^Q B_{l,m;p,q} \cdot G(x_i) \right] + \frac{C}{a^2 \cdot R_{l,m}^2} \right\} \quad (6)$$

$$G(y) = \sum_{k_1=0}^{2m-2} \left\{ \sum_{k_2=0}^{2q-2} g_R \cdot \exp \left[-y \cdot \left(1 + \frac{R_{l,m}^2}{R_{p,q}^2} \right) \right] \cdot \sum_{k_3=0}^{l+p+k_1+k_2} \frac{(l+p+k_1+k_2)!}{(l+p+k_1+k_2-k_3)!} \cdot \frac{y^{(l+p+k_1+k_2-k_3)}}{(1+R_{l,m}^2/R_{p,q}^2)^{k_3+1}} \right\}.$$

$$g_R = \bar{g}(l, m, R_{l,m}, p, q, R_{p,q}).$$

При отсутствии эксцентриситета коэффициенты связи мод одного и того же азимутального порядка рассчитываются по формуле [27]:

$$B_{l,m;l,q} = \frac{\Gamma(m)\Gamma(q)}{\Gamma(l+m)\Gamma(l+q)} \cdot (2\rho_m \rho_q)^{2l+2} \cdot \left[\frac{(\rho_m - \rho_q)^{m+q-2}}{\rho_m^2 + \rho_q^2} \right]^2 \times \left\{ \sum_{k=0}^{\min(m-1, q-1)} (-1)^k \frac{\Gamma(m+q+l+k-1)}{\Gamma(m-k)\Gamma(q-k)k!} \cdot \left(\frac{\rho_m^2 + \rho_q^2}{\rho_m^2 - \rho_q^2} \right)^k \right\}^2, \quad (7)$$

В отличие от обычного метода расщепления по физическим процессам для некоторого i -го участка линии вычисление модифицированным методом осуществляется в четыре шага. Первым шагом является расчет параметров мод при линейном распространении. Определяются коэффициенты связи мод и значения мощности каждой моды на входе участка. На втором шаге выполняется оператор \hat{N} . На третьем этапе, выполняется расчет параметров мод при нелинейном распространении с учетом амплитуды каждой моды. На четвертом шаге выполняется оператор \hat{D} с параметрами мод, рассчитанными для нелинейного режима распространения.

Затем вычисления повторяются для $(i+1)$ -го участка линии. Коэффициенты связи мод на входе $(i+1)$ -го участка, вычисляются для эквивалентного радиуса пятна моды, рассчитанного для нелинейного распространения на i -том участке и эквивалентного радиуса пятна моды на $(i+1)$ -м участке для линейного распространения. В этом случае схема расчета не является симметричной.

4. Моделирование распространения сигнала в маломодовом режиме передачи многомодовой волоконно-оптической линии связи

В данной работе были получены результаты моделирования протяженных линий с рассматриваемыми оптическими волокнами при передаче в одном оптическом канале

со скоростью 100 Гбит/с. Моделирование было выполнено для общей длины линии 2000 км. Предполагалось, что используются оптические эрбиевые усилители (EDFA). Расстояния между усилителями одинаковы. Длина усилительного участка составляла 100 км. Модель оптического эрбиевого усилителя учитывала лишь усиление и шумы усилителя (ASE). Рабочая длина волны оптического канала 1550 нм. Длительность импульса была принята равной 25 пс. Моделировали систему передачи с форматом модуляции DQPSK.

Предполагалось, что фильтр эквалайзера был настроен при инсталляции линии передачи. Соответственно, моделирование для каждого варианта проводили в два этапа. На первом этапе определяли характеристики эквалайзера. На втором этапе оценивали частоту появления ошибочных битов (BER) для эквалайзера с параметрами, найденными на первом этапе.

При моделировании линий с многомодовыми оптическими волокнами, функционирующими в маломодовом режиме, полагали, что ввод и вывод оптического излучения в многомодовое волокно линии осуществляется через стандартные ступенчатые одномодовые оптические волокна SMF28e. Линейные оптические эрбиевые усилители также включены через стандартные ступенчатые одномодовые волокна SMF28e как компонент между линейными волокнами соседних участков. Процесс распространения оптических импульсов в линиях с одномодовыми волокнами в одномодовом режиме был описан аналогично (1) с помощью уравнения Шредингера, но только для основной моды.

В случае применения одномодового волокна моделирование было проведено для хорошо известного оптического волокна SMF28e. Профиль показателя преломления таких оптических волокон приведен на рисунке 1. Технические данные для этих волокон были взяты из техпаспорта [28]. Моделирование линии передачи с многомодовым оптическим волокном проводилось для волокон OM3, профиль показателя преломления которых был получен в результате измерений типичного образца волокна такого рода [29]. Профиль исследуемого оптического волокна приведен на рисунке 2. На рисунке 3 представлен типичный пример полученных результатов моделирования.

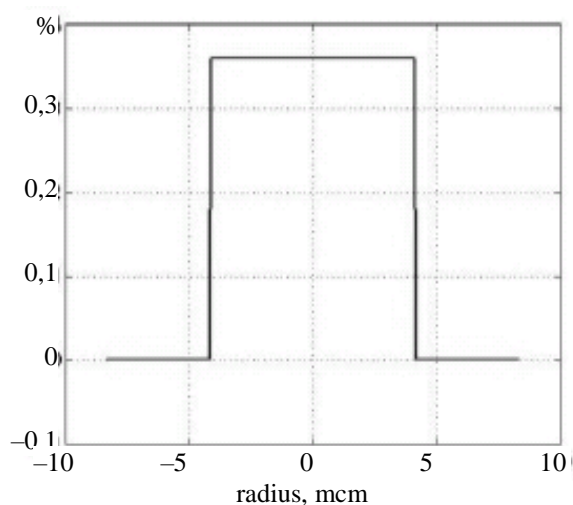


Рис. 1. Профиль показателя преломления одномодового ступенчатого оптического волокна

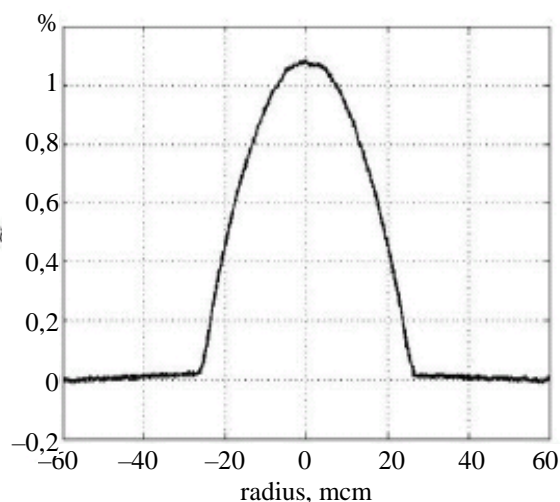


Рис. 2. Профиль показателя преломления многомодового оптического волокна

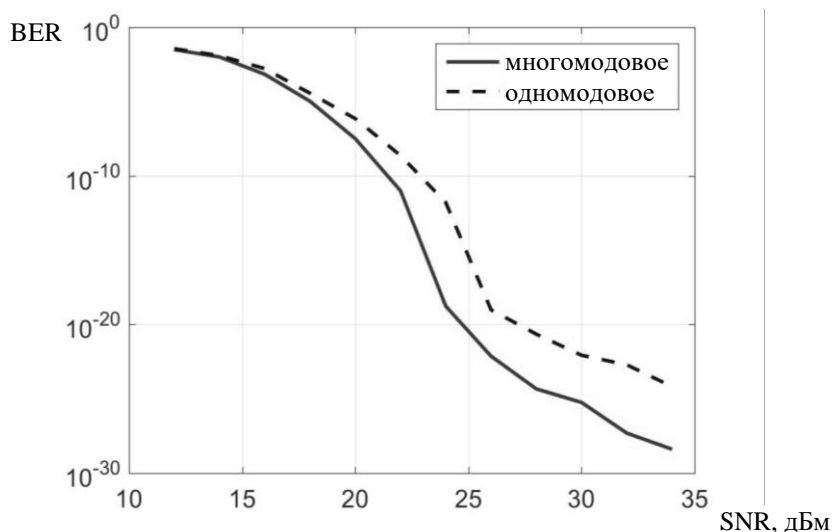


Рис. 3. Пример результатов моделирования

5. Заключение

Описано решение системы нелинейных уравнений Шредингера с помощью модифицированного метода расщепления по физическим процессам для маломодового режима передачи. Приведен пример применения предложенного метода для моделирования передачи по оптической линии связи с одномодовым волокном и многомодовым волокном, функционирующем в маломодовом режиме.

Благодарности

Исследования выполнены при финансовой поддержке РФФИ в рамках научного проекта № 16-37-6001515 мол_а_дк.

Список сокращений

CNLS – система связанных нелинейных уравнений Шредингера (coupled nonlinear Schrödinger equations)	FFT – быстрое преобразования Фурье (fast Fourier transform)
SSFM – метод расщепления по физическим процессам (split-step Fourier method)	EDFA – оптический эрбиевый усилитель (Erbium optical amplifier)
	BER – ошибочные биты (bit error rate)

Литература

1. J.M. Kahn, P.H. Keang. A bottleneck for optical fibres. Nature 411, 2001, pp. 1007-1010. doi: 10.1038/35082671.
2. P.P. Mitra, J.B.Stark. Nonlinear limits to the information capacity of optical fibre communications. Nature 411, 2001, pp.1027-1030. doi: 10.1038/35082518.
3. A.D. Ellis, J. Zhao, D. Cotter. Approaching the Non-Linear Shannon Limit. JLT 11648, R1 2009. doi: 10.1109/JLT.2009.2030693.
4. R.-J. Essiambre, G. Kramer, P.J. Winzer, G.J. Foschini, B. Goebel. Capacity Limits of Optical Fiber Networks. Journal of Lightwave Technology, vol.28, n.4, 2010, pp. 662-700.

5. R.-J. Essiambre, R.W. Tkach. Capacity Trends and Limits of Optical Communication Networks. *Proc. of the IEEE*, vol.100, n.5, 2012, pp. 1035-1055. doi: 10.1109/JLT.2009.2039464.
6. Mecozzi, R.-J. Essiambre. Nonlinear Shannon Limit in Pseudolinear Coherent Systems. *Journal of Lightwave Technology*, vol. 30, n.12, 2012, pp. 2011-2024.
7. M.A. Sorokina, S.K. Turitsyn. Regeneration limit of classical Shannon capacity. *Nature Communications*, n 5, 2014. doi: 10.1038/ncomms4861
8. A.D. Ellis. The nonlinear Shannon limit and the need for new fibres. *Proc. of SPIE* 8434, 84340H, 2012. doi: 10.1117/12.928093
9. K.S. Turitsyn, S.K. Turitsyn. Nonlinear communication channels with capacity above the linear Shannon limit. *Optics Letters*, vol.37, n.17, 2012, pp. 3600-3602. doi: 10.1364/OL.37.003600
10. C. Schubert, J.K. Fischer, C. Schmidt-Langhorst, R. Elschner, L. Molle, M. Nolle, T. Richter. New Trends and Challenges in Optical Digital Transmission Systems. *ECOC Technical Digest*, 2012.
11. M. Shoda, I. Yokota, Y. Sugaya, Sh. Saito. Ultra-High-Speed Core Transport to Handle Rapid Growth in Traffic. *Fujitsu Sci. Tech. J.* vol. 49, n.3, 2013, pp. 342-349.
12. R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, Y. Miyamoto. Coherent Optical Single Carrier Transmission Using Overlap Frequency Domain Equalization for Long-Haul Optical Systems. *Journal of Lightwave Technology*, vol. 27, n.16, 2009, pp. 3721-3728.
13. R. Ryf, N.K. Fontaine, M.A. Mestre, S. Randel, X. Palou, C. Bolle, A.H. Gnauck, S. Chandrasekhar, X. Liu, B. Guan, R.-J. Essiambre, P.J. Winzer, S.G. Leon-Saval, J. Bland-Hawthorn, R. Delbue, P. Pupalais, A. Sureka, Y. Sun, L. Gruner-Nielsen, R.V. Jensen, R. Lingle. Transmission over 130-km Few-Mode Fiber. *Frontiers in Optics 2012, Laser Science XXVIII, OSA Technical Digest (online)*, paper FW6C.4.
14. X. Chen, A. Li, J. Ye, A. Al Amin, W. Shieh. Reception of mode-division multiplexed superchannel via few-mode compatible optical add/drop multiplexer. *Optics Express*, vol.20, n. 13, 2012, pp. 14302-14307. doi: 10.1364/OE.20.014302
15. M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Astruc, L. Provost, F Cerou, G. Charlet. Transmission at 2x100Gb/s, over Two Modes of 40km-long Prototype Few-Mode Fiber, using LCOS-based Mode Multiplexer and Demultiplexer. *Optics Express*, vol. 19, n.17, 2011, pp. 16593-16600. doi: 10.1364/OFC.2011.PDPB9
16. В.А. Андреев, А.В. Бурдин. Многомодовые оптические волокна. Теория и приложения на высокоскоростных сетях связи. Радио и Связь, Москва, 248, 2004.
17. A.V. Bourdine. Mode coupling at the splice of diverse optical fibers. *Proc. SPIE* 8787, *Optical Technologies for Telecommunications 2012*, 878706 (2013). doi: 10.1117/12.2018553
18. В.А. Андреев, В.А. Бурдин, А.В. Бурдин. Маломодовый режим передачи по оптическим волокнам: применение на высокоскоростных ВОЛС. *Электросвязь* 12, 2013, с. 27-30.
19. G.P. Agrawal. *Nonlinear fiber optics*. Academic Press, 2001, pp. 467.
20. G.P. Agrawal. *Applications of nonlinear fiber optics*. Academic Press, 2011, pp. 459.
21. И.Н. Сисакян, А.В. Шварцбург. Нелинейная динамика пикосекундных импульсов в волоконно-оптических световодах. *Квантовая электроника*, т.11, №9, 1984, с.1703-1721.
22. С.М. Широков. Приближенные параметрические модели динамики самовоздействия импульсов в нелинейных оптических средах с модовой дисперсией. *Компьютерная оптика*, том 14-15(2), 1995, с.117-124.

23. F. Ferreira, S. Jansen, P. Monteiro and H. Silva. Nonlinear semi-analytical model for simulation of few-mode fiber transmission. *IEEE Photonics Technology Letters*, vol. 24, n. 4, 2012, pp. 240-242. doi: 10.1109/LPT.2011.2177250.
24. S. Mumtaz, R-J. Essiambre and G. P. Agrawal. Nonlinear propagation in multimode and Multicore fibers: generalization of the Manakov equations. *Journal of Lightwave Technology*, vol. 31, n. 3, 2013, pp.398-406. doi: 10.1109/JLT.2012.2231401.
25. V. A. Burdin and A.V. Bourdine. Modeling and simulation of a few-mode long-haul fiber optic transmission link. *Proc. of SPIE*, vol. 9533, 2015, pp. 953307-1 - 953307-6. doi: 10.1117/12.2181127.
26. A.V. Bourdine and V.A. Burdin. Solution for arbitrary order guided mode of circle optical fiber based on Gaussian approximation. *Proc. of SPIE* 8410, 2012, pp. 841009-1 – 841009-7. doi: 10.1117/12.923235.
27. A.V. Bourdine, Coupled-mode theory application for splice of different optical fibers with arbitrary refractive index profiles. *Proc. of SPIE* 4425, 2001, pp. 224-23. doi: 10.1117/12.429728.
28. А.В. Листвин, В.Н. Листвин, Д.В. Швырков. Оптические волокна для линий связи. ЛЕСАР арт, Москва, 288, 2003.
29. А.В. Бурдин, К.А. Яблочкин. Исследование дефектов профиля показателя преломления многомодовых оптических волокон кабелей связи. *Инфокоммуникационные технологии*. т. 18, №2, 2010, с. 22 – 27.

Владимир Александрович Бурдин, доктор технических наук, профессор (проректор по науке и инновациям, Поволжский государственный университет телекоммуникаций и информатики). Научные интересы: волоконная оптика в телекоммуникациях, новые технологии проектирования, строительства и эксплуатации оптических линий, анализ и синтез волоконных световодов, управляемые дисперсией солитоны в оптических волокнах телекоммуникационных линий, тестирование оптических кабелей, прогноз срока службы оптических кабелей, измерения в волоконной оптике. Он – автор и соавтор более 500 публикаций, включая 134 опубликованных статьи в технических журналах и сборниках, 1 монографию, 3 учебника, 50 патентов; senior member обществ OSA и SPIE.

Антон Владимирович Бурдин, доктор технических наук (профессор кафедры линий связи и измерений в технике связи, Поволжский государственный университет телекоммуникаций и информатики), помощник ректора по инновациям. Научные интересы: многомодовые и маломодовые эффекты в оптических волокнах, высокоскоростная передача сигналов лазеров по нерегулярным линиям с многомодовыми и маломодовыми волокнами, управление дифференциальной модовой задержкой. Он – автор и соавтор более 100 публикаций в реферируемых научных журналах, 25 патентов, 2 монографии; член международных обществ OSA и SPIE; научный руководитель студенческой ячейки SPIE в ПГУТИ.

Modelling nonlinear propagation of modes in long-haul multimode fiber optic links

V.A. Burdin , A.V. Bourdine

Povolzhskiy State University of Telecommunications and Informatics,
Samara, Russia, 77 Moscow av., 443090

Abstract. Modified split-step Fourier method for the numerical simulation of the coupled nonlinear Schrodinger equation is introduced. Examples of simulation for piecewise regular model of the few-mode long-haul fiber optic transmission link by this method are presented.

Keywords: nonlinear Schrodinger equation, coupled nonlinear Schrodinger equations, Erbium amplifier, piecewise regular model, few-mode fiber optic link, mode coupling theory, Gaussian approximation, split-step Fourier method.

1. Introduction

In modern communication systems in the optical 100G channel the problem of dispersion distortion compensation is almost completely solved by methods of electronic compensation in combination with modulation formats application of steady against dispersion, optimization of coding, demodulation and detection. At this case the nonlinearity becomes a major factor of errors in the optical channel [1-7]. The reducing of the nonlinear distortions of the optical channel in the single-mode optical fibers is a extremely difficult task [5-11]. It is known that one of cardinal ways of decrease in influence of nonlinear effects is the increase of the cross section area of an optical fiber core. In particular, for this purpose it is proposed to use so-called few-mode optical fibers [8, 12-15]. From this point of view also of interest multimode optical fibers in few-mode regime. In this case the cross-sectional area is more increased and degree of nonlinearity factors decrease is larger [8, 6-18].

Propagation of optical pulses in optical fibers is usually described by nonlinear Schrödinger equation (NLS) [19, 20]. For few-mode and multimode optical fibers, this process is described by a system of coupled nonlinear Schrödinger equations (CNLS) [19-25]. Various forms of the system recording differ mainly in the definition methods of linear and non-linear mode coupling. In general formulation of the problem the system of M - coupled nonlinear Schrödinger equations is solved numerically by the split-step Fourier method [19-25]. When using the classical split-step Fourier method for the solution of the nonlinear Schrödinger equation it is assumed that the parameters of the optical fiber does not depend on the parameters of the transmitted signal [19, 20]. However, it is obvious that when the few-mode propagation as result of the non-linearity action takes place of the power redistribution over the cross section and the changes of the refractive index profile of optical fiber. This leads to changes of the modes parameters and the parameters of mode coupling in an optical fiber under action of the optical pulse.

In this paper, we propose a modification of the split-step method for solving the system of M - equations describing the nonlinear propagation of few modes in optical fibers. As an example, the results of simulation on the basis of the proposed modification for the optical fiber link with multimode optical fiber in a few-mode regime are presented. For comparison, the results of simulation by this modification for the optical link with a single-mode optical fiber are presented too.

2. Coupled nonlinear Schrodinger equations and split-step Fourier method

Process of propagation of optical impulses in the few-mode mode in the line was described by means of the system of the equations of Schrodinger type[19-25]. This system may be written as [21-22]:

$$\begin{aligned}
j \frac{\partial \Psi_{l,m}}{\partial \eta} + j \beta_{1l,m} \frac{\partial \Psi_{l,m}}{\partial \tau} + \beta_{2l,m} \frac{\partial^2 \Psi_{l,m}}{\partial \tau^2} + j \bar{\alpha}_{l,m} \Psi_{l,m} + \sum k_{l,m;p,q} |\Psi_{p,q}|^2 \Psi_{l,m} &= 0, \\
\eta = \frac{z}{L_0}; \quad \tau = \frac{(t - z/v_0)}{T_0}; \\
\bar{\alpha}_{l,m} = \alpha_{l,m} L_0; \quad \beta_{2l,m} = \frac{L_0}{2T_0} \frac{\partial v_{l,m}^{-1}}{\partial \omega}; \quad \beta_{1l,m} = \frac{L_0}{T_0} \left(\frac{1}{v_{l,m}} - \frac{1}{v_0} \right); \\
k_{l,m;p,q} = K_{l,m}(E_0) \cdot B_{l,m;p,q}; \quad K_{l,m} = \frac{\omega_0^2 n \cdot n_2 L_0 E_0^2}{c^2 \beta_{l,m} B_{l,m;l,m}},
\end{aligned} \tag{1}$$

where

$\Psi_{l,m}$ – is the normalized complex envelope of the l, m – mode;

$\alpha_{l,m}$ – is the attenuation coefficient of the l, m – mode;

$v_{l,m}$ – group speed of the l, m – mode;

$\beta_{l,m}$ – propagation constant of the l, m – mode;

n, n_2 – nonperturbed value of refractive index and coefficient in his nonlinear term;

$\beta_{1l,m}, \beta_{2l,m}$ – is the normalized inverse group speed and dispersion parameter of the l, m –mode;

$B_{l,m;p,q}$ – is the coupling mode coefficient for the l, m and p, q –mode;

l, m – are the azimuthal and the radial order of mode, respectively.

According to the split-step Fourier method (SSFM) the equation (1) can be formally represented as a:

$$\frac{\partial \Psi_{l,m}}{\partial \eta} = (\hat{D} + \hat{N}) \Psi_{l,m}, \tag{2}$$

Where \hat{D} is a differential operator that accounts for dispersion and absorption in a linear medium and \hat{N} is a nonlinear operator that governs the effect of fiber nonlinearities on pulse propagation. In general, dispersion and nonlinearity act together along the length of the fiber. The split-step Fourier method obtains an approximate solution by assuming that in propagating the optical field over a small distance h , the dispersive and nonlinear effects can be pretended to act independently. More specifically, propagation from z to $z+h$ is carried out in two steps. In the first step, the nonlinearity acts alone, and $\hat{D} = 0$. In the second step, dispersion acts alone, and $\hat{N} = 0$. The operator D can be evaluated in the Fourier domain. The use of the fast Fourier transform (FFT) algorithm makes numerical evaluation by operator D relatively fast. It is known two schemes of split-step Fourier method. There are known as non-symmetrized and symmetrized or Usually symmetrized split-step Fourier method is more accurate.

Application of the above-described usual split-step Fourier method assumes that the parameters of modes and hence the parameters of mode coupling do not depend on the parameters of the transmitted signal. Obviously, this is not true and at the analysis of processes of propagation in multimode and few-mode fibers this dependence must be considered. Here we introduce the modified split-step Fourier method, in which this disadvantage is partially eliminated.

3. Modified split-step Fourier method

We assume that the fiber-optic line can be described by a piecewise regular model and thus within a separate regular span of a line parameters of the optical fiber and parameters

propagating modes are constant. Let us assume that the parameters of guided modes can be calculated by a modified Gaussian approximation [26]. Then, in the case of linear propagation for an arbitrary refractive index profile of optical fiber the equation for the equivalent radius of mode spot $R_{l,m}$ has the form:

$$\left\{ \sum_{i=1}^N k^2 \Delta n_i^2 r_i^2 \left[F_m^{(l)}(x_i) \right]^2 \right\} - C = 0, \quad (3)$$

and the mode propagation constant is described by the formula:

$$\beta_{l,m}^2 = k^2 n_0^2 - \frac{m!}{(l+m-1)!} \left\{ \left[\sum_{i=1}^N k^2 \Delta n_i^2 \Phi_m^{(l)}(x_i) \right] + \frac{C}{a^2 \cdot R_{l,m}^2} \right\}, \quad (4)$$

$$C = m \frac{(m+l-1)!}{(m-1)!} + (m+l-1) \frac{(m+l-2)!}{(m-2)!} + 2l^2 \sum_{i=0}^{m-2} \frac{(i+l-1)!}{i!},$$

Here

$$\begin{aligned} \left[F_m^{(l)}(x) \right]^2 &= x^l \left[L_{m-1}^{(l)}(x) \right]^2 \exp(-x), \\ \Phi_m^{(l)}(x) &= \exp(-x) \cdot \sum_{q=0}^{2n} g_q \cdot \left[x^q + \sum_{k=1}^{q+l} \frac{(q+l)!}{(q+l-k)!} x^{q+l-k} \right], \\ g_q &= \sum_{j=0}^q b_j \cdot b_{q-j} ; \quad b_q = \frac{(l+m)!}{q!(m-q)!(l+q)!} \\ \Delta n_i^2 &= n_i^2 - n_{i+1}^2 ; \quad x = R^2 / R_{l,m}^2 ; \quad R = r / a \end{aligned}$$

where $L_p^{(l)}(x)$ —Laguerre polynomials;

a — core radius;

r — radial coordinate.

It is known that $n_E(R) = n(R) + n_2 |E(R)|^2$. We believe that $n_2 |E(R)|^2 \ll n(R)$.

Then $n_E^2(R) \approx n^2(R) + 2 \cdot n(R) \cdot n_2 |E(R)|^2$. For nonlinear propagation we obtain a system of equations for the equivalent radius of mode spot for a set of modes:

$$\begin{aligned} & \left\{ \sum_{i=1}^N k^2 \Delta n_i^2 r_i^2 \left[F_m^{(l)}(x_i) \right]^2 \right\} + \\ & + \left\{ \sum_{i=1}^N k^2 n_2 E_0^2 \Delta n_i^2 r_i^2 \left[F_m^{(l)}(x_i) \right]^2 \cdot \sum_{p=0}^P \sum_{q=0}^Q B_{l,m;p,q} \left[F_q^{(p)} \left(\frac{R_{l,m}^2}{R_{p,q}^2} x_i \right) \right]^2 \right\} - C = 0 \end{aligned} \quad (5)$$

The formula for the propagation constant takes the following form:

$$\begin{aligned} \beta_{l,m}^2 &= k^2 n_0^2 - \\ & - \frac{m!}{(l+m-1)!} \left\{ \left[\sum_{i=1}^N k^2 \Delta n_i^2 \Phi_m^{(l)}(x_i) \right] + \left[\sum_{i=1}^N k^2 n_2 E_0^2 \Delta n_i^2 \sum_{p=0}^P \sum_{q=0}^Q B_{l,m;p,q} \cdot G(x_i) \right] + \frac{C}{a^2 \cdot R_{l,m}^2} \right\}, \quad (6) \\ G(y) &= \sum_{k_1=0}^{2m-2} \left\{ \sum_{k_2=0}^{2q-2} g_R \cdot \exp \left[-y \cdot \left(1 + \frac{R_{l,m}^2}{R_{p,q}^2} \right) \right] \cdot \sum_{k_3=0}^{l+p+k_1+k_2} \frac{(l+p+k_1+k_2)!}{(l+p+k_1+k_2-k_3)!} \cdot \frac{y^{(l+p+k_1+k_2-k_3)}}{(1+R_{l,m}^2/R_{p,q}^2)^{k_3+1}} \right\}. \\ g_R &= \bar{g}(l, m, R_{l,m}, p, q, R_{p,q}). \end{aligned}$$

In the absence of eccentricity the coefficients of mode coupling for mode of the same azimuthal order is calculated according to the formula [27]:

$$B_{l,m;l,q} = \frac{\Gamma(m)\Gamma(q)}{\Gamma(l+m)\Gamma(l+q)} \cdot (2\rho_m\rho_q)^{2l+2} \cdot \left[\frac{(\rho_m - \rho_q)^{m+q-2}}{\rho_m^2 + \rho_q^2} \right]^2 \times \left\{ \sum_{k=0}^{\min\left(\frac{m-1}{2}, \frac{q-1}{2}\right)} (-1)^k \frac{\Gamma(m+q+l+k-1)}{\Gamma(m-k)\Gamma(q-k)k!} \cdot \left(\frac{\rho_m^2 + \rho_q^2}{\rho_m^2 - \rho_q^2} \right)^k \right\}^2, \quad (7)$$

In contrast to the conventional split-step Fourier method for the i -th certain span of line the calculation by modified method carried out in four steps. The first step is the calculation of the mode parameters at the linear propagation. Coefficients of mode coupling and power of mode at the span input calculated. The second step is performed operator \hat{N} . In the third step, the calculation modes parameters in the nonlinear propagation with regard to their amplitude. The fourth step is the implementation of operator \hat{D} with parameters of mode calculated for the nonlinear propagation.

Then calculations are repeated for $(i+1)$ -th span. Coefficients of mode coupling at the input $(i+1)$ -th span calculated for equivalent radius of mode spot calculated for the nonlinear propagation at the i -th span and the equivalent radius of mode spot at the $(i+1)$ -th span calculated for the linear propagation. In this case, the scheme of calculation is non symmetrical.

4. Simulation of signal propagation in multimode fiber optic link at few-mode regime

In this work results of modeling of long-haul lines with considered optical fibers by transmission in the optical 100G channel are presented. At modeling the total link length makes 2000 km. Erbium optical amplifiers (EDFA) are used. Distances between amplifiers are identical. Length of the amplifying section (AS) made 100 km. The erbium optical amplifier model considered only a amplification gain and optical amplifier noise (ASE). Researches carried out on the operating wavelength of the optical channel $\lambda = 1550$ nm. Pulse duration was accepted equal 25 ps. Transmission system with a modulation format DQPSK was modeled.

It is assumed that the equalizer filter is tuned at installation of a transmission line. Respectively, modeling for each variation was carried out in two stages. At the first stage characteristics of the equalizer filter is defined. At the second stage bit error rate (BER) is estimated for the equalizer filter with the parameters found at the first stage.

At modeling of lines with the multimode optical fibers operating in the few-mode-regime believed that input and output of optical emission in link fiber is carried out through standard step single-mode fiber SMF28e. Linear optical erbium amplifiers turn on also through standard step single-mode fiber SMF28e as a component between line fibers of adjoined spans. Propagation process for optical impulses of links with single-mode fibers in the single-mode regime was described similarly to (1) by means of the Schrodinger equation, but only for the fundamental mode.

In the case of single-mode fiber simulation was carried out for the well known optical fiber SMF28e. The refractive index profile of such optical fibers is given on figure 1. Technical data for these fibers took according to the data sheet [28]. Modeling of a transmission line with multimode optical fiber was carried out for OM3 fibers which refractive index profile was received as a result of measurements of a typical fiber sample of this kind [29]. The profile of the researched optical fiber is given on figure 2. On figure 3 the typical example of simulations results is presented.

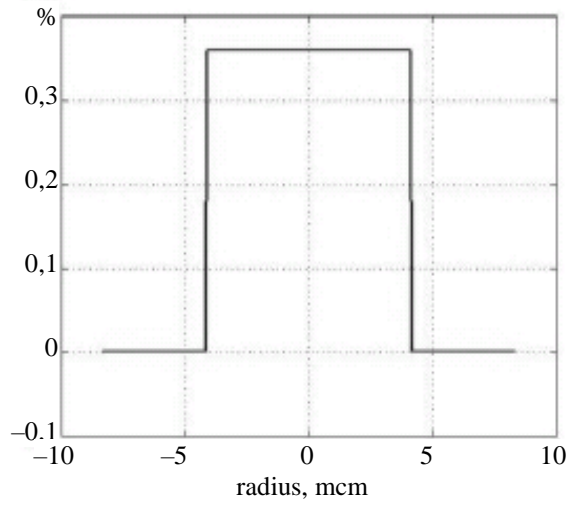


Fig. 1. Refractive index profile of single mode fiber

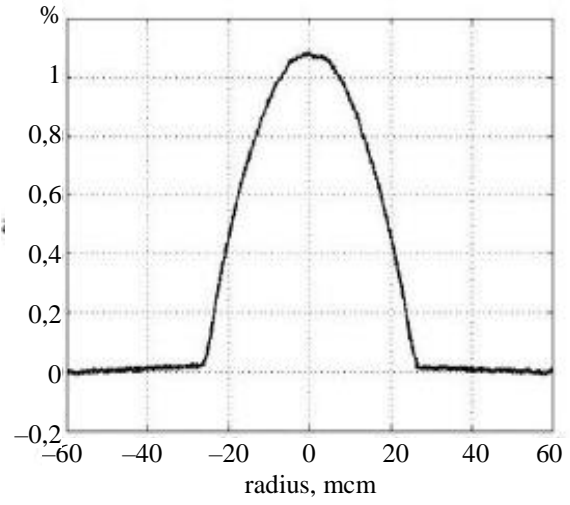


Fig. 2. Refractive index profile of multimode fiber

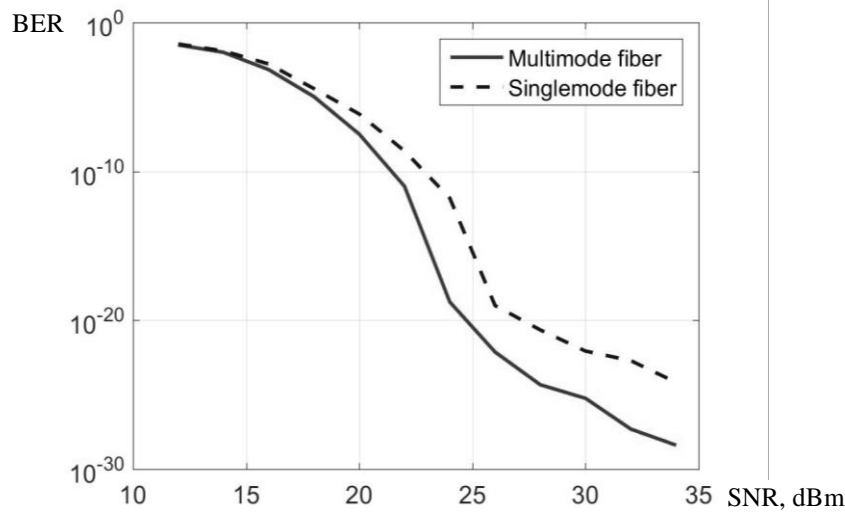


Fig. 3. Results of simulations

5. Conclusions

Solution of CNLS equations set by using the modified for few-mode transmission split-step Fourier method is presented. Examples of proposed method application on optical fiber link with single-mode fiber and multimode fiber under few-mode regime are represented.

6. Acknowledgements

The reported study was funded by RFBR according to the research project No. 16-37-60015 mol_a_dk.

Abbreviation list

NLS – nonlinear Schrödinger equation
CNLS – coupled nonlinear Schrödinger equations
SSFM – split-step Fourier method

FFT – fast Fourier transform
EDFA – Erbium optical amplifier
AS – amplifying section
BER – bit error rate

References

1. J.M. Kahn, P.H. Keang. A bottleneck for optical fibres. *Nature* 411, 2001, pp. 1007-1010. doi: 10.1038/35082671.
2. P.P. Mitra, J.B.Stark. Nonlinear limits to the information capacity of optical fibre communications. *Nature* 411, 2001, pp.1027-1030. doi: 10.1038/35082518.
3. A.D. Ellis, J. Zhao, D. Cotter. Approaching the Non-Linear Shannon Limit. *JLT* 11648, R1 2009. doi: 10.1109/JLT.2009.2030693.
4. R.-J. Essiambre, G. Kramer, P.J. Winzer, G.J. Foschini, B. Goebel. Capacity Limits of Optical Fiber Networks. *Journal of Lightwave Technology*, vol.28, n.4, 2010, pp. 662-700.
5. R.-J. Essiambre, R.W. Tkach. Capacity Trends and Limits of Optical Communication Networks. *Proc. of the IEEE*, vol.100, n.5, 2012, pp. 1035-1055. doi: 10.1109/JLT.2009.2039464.
6. R.-J. Mecozi, Essiambre. Nonlinear Shannon Limit in Pseudolinear Coherent Systems. *Journal of Lightwave Technology*, vol. 30, n.12, 2012, pp. 2011-2024.
7. M.A. Sorokina, S.K. Turitsyn. Regeneration limit of classical Shannon capacity. *Nature Communications*, n 5, 2014. doi: 10.1038/ncomms4861
8. A.D. Ellis. The nonlinear Shannon limit and the need for new fibres. *Proc. of SPIE* 8434, 84340H, 2012. doi: 10.1117/12.928093
9. K.S. Turitsyn, S.K. Turitsyn. Nonlinear communication channels with capacity above the linear Shannon limit. *Optics Letters*, vol.37, n.17, 2012, pp. 3600-3602. doi: 10.1364/OL.37.003600
10. C. Schubert, J.K. Fischer, C. Schmidt-Langhorst, R. Elschner, L. Molle, M. Nolle, T. Richter. New Trends and Challenges in Optical Digital Transmission Systems. *ECOC Technical Digest*, 2012.
11. M. Shoda, I. Yokota, Y. Sugaya, Sh. Saito. Ultra-High-Speed Core Transport to Handle Rapid Growth in Traffic. *Fujitsu Sci. Tech. J.* vol. 49, n.3, 2013, pp. 342-349.
12. R. Kudo, T. Kobayashi, K. Ishihara, Y. Takatori, A. Sano, Y. Miyamoto. Coherent Optical Single Carrier Transmission Using Overlap Frequency Domain Equalization for Long-Haul Optical Systems. *Journal of Lightwave Technology*, vol. 27, n.16, 2009, pp. 3721-3728.
13. R. Ryf, N.K. Fontaine, M.A. Mestre, S. Randel, X. Palou, C. Bolle, A.H. Gnauck, S. Chandrasekhar, X. Liu, B. Guan, R.-J. Essiambre, P.J. Winzer, S.G. Leon-Saval, J. Bland-Hawthorn, R. Delbue, P. Pupalaikis, A. Sureka, Y. Sun, L. Gruner-Nielsen, R.V. Jensen, R. Lingle. Transmission over 130-km Few-Mode Fiber. *Frontiers in Optics 2012, Laser Science XXVIII, OSA Technical Digest (online)*, paper FW6C.4.
14. X. Chen, A. Li, J. Ye, A. Al Amin, W. Shieh. Reception of mode-division multiplexed superchannel via few-mode compatible optical add/drop multiplexer. *Optics Express*, vol.20, n. 13, 2012, pp. 14302-14307. doi: 10.1364/OE.20.014302
15. M. Salsi, C. Koebele, D. Sperti, P. Tran, P. Brindel, H. Mardoyan, S. Bigo, A. Boutin, F. Verluise, P. Sillard, M. Astruc, L. Provost, F Cerou, G. Charlet. Transmission at 2x100Gb/s, over Two Modes of 40km-long Prototype Few-Mode Fiber, using LCOS-based Mode Multiplexer and Demultiplexer. *Optics Express*, vol. 19, n.17, 2011, pp. 16593-16600. doi: 10.1364/OFC.2011.PDPB9
16. V.A. Andreev, A.V. Bourdine. Mnogomodovye opticheskie volokna. Teoriya i prilozheniya na vysokoskorostnykh setyakh svyazi [Multimode optical fibers. The theory and appendices on high-speed communication networks]. *Radio i svyaz*, Moscow, 248, 2004. (In Russian)
17. A.V. Bourdine. Mode coupling at the splice of diverse optical fibers. *Proc. SPIE* 8787, *Optical Technologies for Telecommunications 2012*, 878706 (2013). doi: 10.1117/12.2018553

18. V.A. Andreev, V.A. Burdin, A.V. Bourdine. Few-mode transmission regime in optical fibers: application on high-speed fiber optic lines. *Electrosvyaz* 12, 2013, pp. 27-30. (In Russian)
19. G.P. Agrawal. *Nonlinear fiber optics*. Academic Press, 2001, pp. 467.
20. G.P. Agrawal. *Applications of nonlinear fiber optics*. Academic Press, 2011, pp. 459.
21. I.N. Sisakyan, A.V. Shvartburg. Nelineynaya dinamika pikosekundnykh impul'sov v volokonno-opticheskikh svetovodakh [Nonlinear dynamics of picosecond impulses in fiber-optical lightguides]. *Kvantovaya elektronika*, vol. 11, n.9, 1984, pp. 1703-1721. (In Russian)
22. S.M. Shirokov. Priblizhennye parametricheskie modeli dinamiki samovozdeystviya impul'sov v nelineynykh opticheskikh sredakh s modovoy dispersiei. [Approximate parametrical models of dynamics of self-influence of impulses in nonlinear optical environments with mode dispersion]. *Komputernaya optika*, 14-15(2), 1995, pp.117-124. (In Russian)
23. F. Ferreira, S. Jansen, P. Monteiro and H. Silva. Nonlinear semi-analytical model for simulation of few-mode fiber transmission. *IEEE Photonics Technology Letters*, vol. 24, n. 4, 2012, pp. 240-242. doi: 10.1109/LPT.2011.2177250.
24. S. Mumtaz, R-J. Essiambre and G. P. Agrawal. Nonlinear propagation in multimode and Multicore fibers: generalization of the Manakov equations. *Journal of Lightwave Technology*, vol. 31, n. 3, 2013, pp.398-406. doi: 10.1109/JLT.2012.2231401.
25. V. A. Burdin and A.V. Bourdine. Modeling and simulation of a few-mode long-haul fiber optic transmission link. *Proc. of SPIE*, vol. 9533, 2015, pp. 953307-1 - 953307-6. doi: 10.1117/12.2181127.
26. A.V. Bourdine and V.A. Burdin. Solution for arbitrary order guided mode of circle optical fiber based on Gaussian approximation. *Proc. of SPIE* 8410, 2012, pp. 841009-1 – 841009-7. doi: 10.1117/12.923235.
27. A.V. Bourdine, Coupled-mode theory application for splice of different optical fibers with arbitrary refractive index profiles. *Proc. of SPIE* 4425, 2001, pp. 224-23. doi: 10.1117/12.429728.
28. A.V. Listvin, V.N. Listvin, D.V. Shvyrykov, *Opticheskie volokna dlya liniy svyazi*. [Optical fibers for communication lines]. LESARart, Moscow, 2003, 288. (In Russian)
29. A.V. Bourdine, K.A. Yablochkin. Issledovanie defektov profilya pokazatelya prelomleniya mnogomodovykh opticheskikh volokon kabeley svyazi [Research of defects of a refractive index profile of multimode optical fibers of communication cables]. *Infocommunication technologies*, vol.8, n.2, 2010, pp. 22-27. (In Russian)

Vladimir Alexandrovich Burdin, Doctor of Technical Sciences, Professor (Vice-Rector of Scientific Activity of Povolzhskiy State University of Telecommunications and Informatics). The research interests: fiber optics in telecommunications, advanced technologies of fiber optic link design, installation and technical maintenance, optical waveguide analysis and design, dispersion managed solitons in fiber optic of telecommunication links, fiber optic cable test, fiber optic cable lifetime, measurement in fiber optics. He is author or co-author of more than 500 publications including 134 printed papers in technical journals and proceedings, 1 monograph, 3 college textbooks, 50 patents; senior member of OSA and SPIE.

Anton Vladimirovich Bourdine, Doctor of Technical Sciences, Professor (Povolzhskiy State University of Telecommunications and Informatics, Department of Communication Lines, Assistant to the Rector for Innovation). The research interests: multi- and few-mode effects in optical fibers, laser-based high bit rate data transmission over irregular fiber optic links with multi- and few-mode fibers, differential mode delay management. He is author or co-author of more than 100 publications in referred scientific journals, 25 patents, 2 monographs; regular OSA and SPIE member; scientific advisor of PSUTI SPIE Student Chapter.

Об условиях трансверсальности в решении вариационных задач

М.В. Левский

НИИ Космических Систем ГКНПЦ им. М.В. Хруничева
г. Королев Московской области, ул. Тихонравова, 27

Аннотация: Тематика работы связана с исследованиями вариационных задач и, в частности, с методами математической теории оптимальных процессов. С использованием принципа максимума Л.С. Понтрягина детально изучены условия трансверсальности, их роль, место и значение в общей процедуре решения задач поиска оптимальных функций. Представленная работа позволяет устранить имеющийся методический пробел, связанный с неполным использованием условий трансверсальности. На конкретных примерах вариационных задач иллюстрируется, что условия трансверсальности (как одно из необходимых условий оптимальности в форме принципа максимума) являются эффективным математическим инструментом (а в ряде случаев – единственным) при определении характерных свойств, закономерностей и ключевых характеристик (параметров, констант, интегралов движения) оптимального решения систем обыкновенных дифференциальных уравнений (динамических систем).

Ключевые слова: критерий оптимальности, принцип максимума, условия трансверсальности, функция Гамильтона, условия оптимальности, краевая задача

1. Актуальность проблемы

В настоящее время при решении многих оптимизационных задач вариационными методами (например, с помощью принципа максимума Л.С. Понтрягина) значение условий трансверсальности явно недооценивается. Как правило, условия трансверсальности не используются должным образом и не участвуют в определении количественных характеристик оптимального процесса (параметров, констант, первых интегралов и пр.), и они не привлекаются для расчета оптимальных функций. При этом о полноте такого полученного решения (найденного в результате применения принципа максимума без условий трансверсальности) судить, видимо, проблематично. Цель данной работы – обеспечить понимание того, что условия трансверсальности – это не простая формальность, участвующая в процедуре нахождения оптимальных временных функций (присутствующая в системе необходимых условий оптимальности, записанных в соответствии с принципом максимума). Это – эффективное средство (в ряде случаев – единственное) для определения характеристик и свойств оптимального процесса, для вычисления его ключевых параметров.

Многие авторы и исследователи (как правило, в прикладных областях) ограничиваются лишь тем, что проверяют выполнение условий трансверсальности (как необходимых условий оптимальности) только в конце решения задач построения оптимальных процессов, т.е. – после того, как само решение получено, формализовано конкретными математическими соотношениями. Проверку делают путем подстановки найденного оптимального решения исходной системы уравнений и сопряженных функций в уравнения, соответствующие условиям трансверсальности. Между тем условия трансверсальности являются такими же равноправными необходимыми элементами оптимальности, как и общепризнанные: условие максимума функции Гамильтона и сопряженная система дифференциальных уравнений. Условия трансверсальности необходимо включать в систему необходимых условий оптимальности с самого начала

(наравне с сопряженной системой уравнений и условием максимума функции Гамильтона), сразу после введения сопряженных переменных, записи функции Гамильтона и сопряженной системы уравнений. Такой порядок действий и дальнейшего применения процедуры принципа максимума является наиболее правильным и целесообразным. Как показывает практическое применение формализма принципа максимума, во многих случаях включение условий трансверсальности в общую систему необходимых условий оптимальности еще до формализации закономерностей оптимального движения динамической системы, описываемой исходной системой дифференциальных уравнений (задаваемой формулировкой оптимизационной задачи), является единственно верным. Это позволяет найти решение сложнейших задач. Решение краевой задачи принципа максимума также будет учитывать условия трансверсальности, что обеспечит полноту и достоверность результатов решения исходной задачи оптимизации, гарантирует законченность полученного решения.

2. Постановка задачи

Пусть динамический процесс задан дифференциальным уравнением вида

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y}, t) \quad (1)$$

где \mathbf{x} – вектор состояния (его компоненты x_i – фазовые переменные, отражающие текущее состояние динамической системы, $i = \overline{1, n}$); \mathbf{y} – вектор управления, у которого компоненты y_j – искомые функции времени, которые необходимо выбрать оптимальными ($j = \overline{1, l}$). Физическое содержание конкретной оптимизационной задачи определяется видом уравнений $\dot{x}_i = f_i(\mathbf{x}, \mathbf{y}, t)$. Пусть динамический процесс описывается следующей системой уравнений

$$\begin{aligned} \dot{x}_1 &= y_3 x_4 - y_2 x_7, \quad \dot{x}_2 = y_3 x_5 - y_2 x_8, \quad \dot{x}_3 = y_3 x_6 - y_2 x_9 \\ \dot{x}_4 &= y_1 x_7 - y_3 x_1, \quad \dot{x}_5 = y_1 x_8 - y_3 x_2, \quad \dot{x}_6 = y_1 x_9 - y_3 x_3 \\ \dot{x}_7 &= y_2 x_1 - y_1 x_4, \quad \dot{x}_8 = y_2 x_2 - y_1 x_5, \quad \dot{x}_9 = y_2 x_3 - y_1 x_6 \end{aligned} \quad (2)$$

Система (2) определяет изменение функций $x_i(t)$ при известных функциях $y_j(t)$ ($i = \overline{1, 9}$, $j = \overline{1, 3}$). Функция $\mathbf{x}(t)$ определяет решение дифференциального уравнения (1). Для конкретизации на решения уравнения (1) необходимо наложить некоторые ограничения, в частности, задать условия в начальный $t = 0$ и конечный $t = T$ моменты времени. Рассмотрим широко распространенный класс задач, когда краевые условия для траектории движения $\mathbf{x}(t)$ формализуются равенствами

$$\mathbf{x}(0) = \mathbf{x}_0 \quad (3)$$

$$\mathbf{x}(T) = \mathbf{x}_T \quad (4)$$

где \mathbf{x}_0 , \mathbf{x}_T – фиксированные векторы, задающие состояние динамической системы в начальный и конечный моменты времени.

Заметим, что в большинстве случаев уравнение (1) имеет не одно, а бесконечно много решений, удовлетворяющих заданным условиям (3), (4). Из всего множества решений системы уравнений (1), (3), (4) интересны только те, которые являются оптимальными (наилучшими) в том или ином смысле. Обычно критерий оптимальности выражается функционалом качества

$$G = \int_0^T g(\mathbf{x}, \mathbf{y}, t) dt \quad (5)$$

где $g(\mathbf{x}, \mathbf{y}, t)$ – неотрицательная функция; T – время окончания оптимизации изучаемого процесса.

В самом общем случае задача оптимизации формулируется следующим образом: требуется перевести динамическую систему, описываемую уравнением (1), из состояния (3) в состояние (4) так, чтобы интеграл (5) был минимальным.

3. Процедура решения задачи

Предположим функции y_j в уравнениях (2) кусочно-непрерывными функциями времени. Используем принцип максимума Л.С. Понтрягина [1]. В соответствии с этим введем сопряженные переменные q_i , соответствующие компонентам x_i вектора состояния \mathbf{x} , и запишем функцию Гамильтона.

$$H = \varphi_0 g + q_1 (y_3 x_4 - y_2 x_7) + q_2 (y_3 x_5 - y_2 x_8) + q_3 (y_3 x_6 - y_2 x_9) + \\ + q_4 (y_1 x_7 - y_3 x_1) + q_5 (y_1 x_8 - y_3 x_2) + q_6 (y_1 x_9 - y_3 x_3) + \\ + q_7 (y_2 x_1 - y_1 x_4) + q_8 (y_2 x_2 - y_1 x_5) + q_9 (y_2 x_3 - y_1 x_6)$$

где $\varphi_0 = \text{const} < 0$ ($\varphi_0 \neq 0$).

Уравнения для сопряженных переменных q_i , согласно принципу максимума [1], имеют вид:

$$\dot{q}_i = -\partial H / \partial x_i \quad (i = \overline{1,9} \quad j = \overline{1,3}) \quad (6)$$

Сопряженные переменные q_i должны удовлетворять условиям трансверсальности

$$q_i(0) = \sum_{h=1}^m w_h \frac{\partial F_h(\mathbf{x}(0))}{\partial x_i} \Big|_{t=0}; \quad q_i(T) = \sum_{k=1}^p z_k \frac{\partial \Phi_k(\mathbf{x}(T))}{\partial x_i} \Big|_{t=T} \quad (7)$$

где $F_h(\mathbf{x}(0))$ – функции, задающие краевое условие на левом конце траектории $\mathbf{x}(t)$ в форме $F_h(\mathbf{x}(0)) = 0$, причем $h = \overline{1,m}$ ($m \leq n$); w_h – числа не все равные нулю; $\Phi_k(\mathbf{x}(T))$ – функции, задающие краевое условие для правого конца траектории $\mathbf{x}(t)$ в форме $\Phi_k(\mathbf{x}(T)) = 0$, причем $k = \overline{1,p}$ ($p \leq n$); z_k – числа не все равные нулю; для задачи в постановке (2)–(5) число фазовых переменных $n = 9$.

Если время T окончания оптимального процесса не фиксировано, то оптимальные сопряженные функции $q_i(t)$ должны удовлетворять еще одному необходимому условию оптимальности, выраженному условием трансверсальности

$$H(\mathbf{x}(T), \mathbf{y}(T), \mathbf{q}(T), T) = - \sum_{k=1}^p v_k \frac{\partial \Phi_k(\mathbf{x}(T), T)}{\partial T} \quad (8)$$

где $\mathbf{q} = \{q_1, q_2, \dots, q_9\}$ – вектор сопряженных переменных; $\Phi_k(\mathbf{x}(T), T)$ – функции, задающие краевое условие на правом конце траектории $\mathbf{x}(t)$ в форме $\Phi_k(\mathbf{x}(T), T) = 0$, причем $k = \overline{1,p}$ ($p \leq n$); v_k – числа не все равные нулю (напомним, что для задачи (2)–(5) $n = 9$, а $l = 3$).

Из структуры функции Гамильтона, уравнений (6) и условий трансверсальности (7), (8) следует, что оптимальные функции q_i пропорциональны отрицательной константе φ_0 . Поэтому справедливо принять $\varphi_0 = -1$.

Так как рассматриваемая задача (2)–(5) – это задача с закрепленными левым и правым концами траектории $\mathbf{x}(t)$, то условие трансверсальности (7) для сопряженных переменных q_i на момент времени $t = 0$ имеет вид [2]: $q_i(0) = w_i$ ($i = \overline{1,9}$), где w_1, w_2, \dots

w_9 – некоторые числа, одновременно не равные нулю. Для сопряженных переменных q_i на момент времени $t = T$ условие трансверсальности (7) имеет аналогичный вид [2]: $q_i(T) = z_i$ ($i = 1, 9$), где хотя бы одно из чисел z_1, z_2, \dots, z_9 отлично от нуля; а условие трансверсальности (8) принимает форму

$$H(T) = 0 \quad (9)$$

Первые два условия трансверсальности эквивалентны неравенствам

$$\sum_{i=1}^9 q_i^2(0) > 0 \quad \text{и} \quad \sum_{i=1}^9 q_i^2(T) > 0 \quad (10)$$

и становятся очевидными, если краевые условия (3), (4) переписать в форме равенств

$$x_i(0) - x_{i0} = 0 \quad \text{и} \quad x_i(T) - x_{iT} = 0$$

где x_{i0} – элементы вектора \mathbf{x}_0 ; x_{iT} – элементы вектора \mathbf{x}_T ; отсюда $F_k = x_i(0) - x_{i0}$; $\Phi_k = x_i(T) - x_{iT}$; $\partial F_k / \partial x_i = \partial \Phi_k / \partial x_i = 1$ для $k = i$ и $\partial F_k / \partial x_i = \partial \Phi_k / \partial x_i = 0$, если $k \neq i$.

Последнее условие трансверсальности (требование (9)) получается, если граничное условие (4) для правого конца траектории представить в виде $\Phi_k = 0$, где $\Phi_k = x_i(T) - x_{iT}$; откуда становится очевидным утверждение, что $\partial \Phi_k / \partial T \equiv 0$ для всех k , а значит, условия (8) и (9) являются эквивалентными.

Задача нахождения оптимальной функции $\mathbf{y}(t)$ и оптимального движения $\mathbf{x}(t)$ системы (2) сводится к решению уравнений (2), (6) (с одновременной максимизацией в каждый момент времени t функции Гамильтона), удовлетворяющего условиям трансверсальности (10) (если время T фиксировано) или (9), (10) (если время T не фиксировано).

Краевая задача принципа максимума заключается в определении таких начальных значений $q_i(0)$ для сопряженных функций, при которых решение $\mathbf{x}(t)$, $\mathbf{q}(t)$ системы уравнений (2), (6) с начальными условиями (3), удовлетворяющее условию максимума функции Гамильтона в каждый текущий момент времени t и условиям трансверсальности (10) (или (9), (10) для задач с нефиксированным временем T) в начальный и конечный моменты времени, удовлетворяло бы граничному условию (4).

Предположим, что функционал (5) не содержит переменных состояния x_i . В этом случае сопряженная система дифференциальных уравнений для q_i запишется в следующей форме:

$$\begin{aligned} \dot{q}_1 &= y_3 q_4 - y_2 q_7, \quad \dot{q}_2 = y_3 q_5 - y_2 q_8, \quad \dot{q}_3 = y_3 q_6 - y_2 q_9 \\ \dot{q}_4 &= y_1 q_7 - y_3 q_1, \quad \dot{q}_5 = y_1 q_8 - y_3 q_2, \quad \dot{q}_6 = y_1 q_9 - y_3 q_3 \\ \dot{q}_7 &= y_2 q_1 - y_1 q_4, \quad \dot{q}_8 = y_2 q_2 - y_1 q_5, \quad \dot{q}_9 = y_2 q_3 - y_1 q_6 \end{aligned} \quad (11)$$

Нетрудно показать, что в этом частном варианте постановки задачи (2)–(5) (когда функционал (5) не зависит явным образом от фазовых переменных x_i) условия трансверсальности $\mathbf{q}(0) \neq 0$ и $\mathbf{q}(T) \neq 0$ выполняются автоматически, если считать функцию $\mathbf{q}(t)$ не тождественно равной нулю (\mathbf{q} – вектор сопряженных переменных q_i). Чтобы убедиться в этом, достаточно проанализировать сумму

$$\sum_{i=1}^9 q_i^2(t) \quad (12)$$

Взяв производную по времени от суммы (12) и заменив в получившемся выражении величины \dot{q}_i правыми частями уравнений (11), получим тождество

$$\sum_{i=1}^9 \dot{q}_i q_i \equiv 0$$

означающее, что сумма (12) есть величина постоянная. Поэтому, если допустить, что хотя бы одно из условий (10) не выполняется, то вектор \mathbf{q} сопряженных переменных окажется тождественно равным нулю ($\mathbf{q} \equiv 0$), а такие сопряженные функции $q_i(t)$, очевидно, не могут быть оптимальными. Следовательно, любое решение $\mathbf{q} \neq 0$, удовлетворяющее системе (11), неизбежно будет удовлетворять условиям трансверсальности (7).

Системы уравнений (2) и (11) однотипны относительно коэффициентов y_1, y_2, y_3 . Решения уравнений (2) для переменных x_i и уравнений (11) для переменных q_i отличаются начальными условиями.

В функции H сгруппируем множители для y_1, y_2, y_3 , содержащие фазовые переменные x_i , и обозначим их r_1, r_2, r_3 . Выделив в явном виде множители r_j при y_j , представим функцию Гамильтона H в следующем виде:

$$H = -g + y_1 r_1 + y_2 r_2 + y_3 r_3$$

где $r_1 = q_4 x_7 + q_5 x_8 + q_6 x_9 - q_7 x_4 - q_8 x_5 - q_9 x_6$; $r_2 = q_7 x_1 + q_8 x_2 + q_9 x_3 - q_1 x_7 - q_2 x_8 - q_3 x_9$; $r_3 = q_1 x_4 + q_2 x_5 + q_3 x_6 - q_4 x_1 - q_5 x_2 - q_6 x_3$.

Для переменных r_i справедливы дифференциальные уравнения:

$$\dot{r}_1 = y_3 r_2 - y_2 r_3, \quad \dot{r}_2 = y_1 r_3 - y_3 r_1, \quad \dot{r}_3 = y_2 r_1 - y_1 r_2 \quad (13)$$

Дифференцируя по времени равенства для r_j ($j = \overline{1,3}$) и подставляя в них выражения для \dot{x}_i и \dot{q}_i , вычисленные по уравнениям (2) и (11) соответственно, получим необходимые дифференциальные уравнения для функций $r_j(t)$.

$$\begin{aligned} \dot{r}_1 = & q_4 \dot{x}_7 + \dot{q}_4 x_7 + q_5 \dot{x}_8 + \dot{q}_5 x_8 + q_6 \dot{x}_9 + \dot{q}_6 x_9 - q_7 \dot{x}_4 - \dot{q}_7 x_4 - q_8 \dot{x}_5 - \dot{q}_8 x_5 - \\ & - q_9 \dot{x}_6 - \dot{q}_9 x_6 = y_2 (q_4 x_1 + q_5 x_2 + q_6 x_3 - q_1 x_7 - q_2 x_8 - q_3 x_9) + \\ & + y_3 (q_7 x_1 + q_8 x_2 + q_9 x_3 - q_1 x_7 - q_2 x_8 - q_3 x_9) = y_3 r_2 - y_2 r_3 \end{aligned}$$

(для переменных r_2, r_3 получаем аналогичные выражения). Изменение вектора $\mathbf{r} = \{r_1, r_2, r_3\}$ определяется решением уравнения $\dot{\mathbf{r}} = \mathbf{r} \times \mathbf{y}$, где $\mathbf{y} = \{y_1, y_2, y_3\}$.

Уравнения (13) для компонент r_j вектора \mathbf{r} заменяют сопряженную систему уравнений (11). Как нетрудно видеть из (13), для вектора \mathbf{r} справедливо свойство $|\mathbf{r}| = \text{const}$. Для существования нетривиального решения необходимо, чтобы при любом $t \in [0, T]$ не выполнялись равенства $x_i = \chi q_i$, потому что при их выполнении $r_1 = r_2 = r_3 \equiv 0$ и дальнейшее решение задачи теряет смысл. Поэтому, при поиске оптимального решения векторы \mathbf{x} и \mathbf{q} считаются такими, что $\mathbf{x} \neq \chi \mathbf{q}$, где χ – скалярная величина (в противном случае $\mathbf{r} \equiv 0$), и рассматривается только вариант $\mathbf{r} \neq 0$ (т.е. когда функции q_i не пропорциональны переменным x_i). Понятно, что если $\mathbf{r} \neq 0$, то $\mathbf{q}(t) \neq 0$, и условия трансверсальности (7) удовлетворяются.

Задача нахождения оптимальной функции $y(t)$ и оптимального движения $\mathbf{x}(t)$ системы (2) свелась к решению уравнений (2), (13) с одновременной максимизацией в каждый момент времени t функции Гамильтона и удовлетворением условия трансверсальности (9), если время T не фиксировано. Краевая задача принципа максимума заключается в определении такого значения вектора $\mathbf{r}(0) \neq 0$, при котором решение $\mathbf{x}(t)$ системы уравнений (2), (13) с начальными условиями (3), удовлетворяющее условию максимума функции Гамильтона в каждый момент времени t и условию трансверсальности (9) (для задач с нефиксированным временем T), удовлетворяет граничному условию (4).

4. Примеры задач, когда условия трансверсальности определяют ключевые характеристики оптимального процесса

Для примера приведем решение задачи (2)–(5), в которой минимизируемый функционал (5) имеет вид:

$$G = \int_0^T \sqrt{k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2} dt \quad (14)$$

где $k_1 > 0, k_2 > 0, k_3 > 0$ – постоянные положительные коэффициенты.

Особенностью минимизируемого функционала (14) является то, что его значение не зависит от длительности T оптимального процесса; значение G определяется исключительно краевыми условиями (векторами \mathbf{x}_0 и \mathbf{x}_T) и коэффициентами k_1, k_2, k_3 (данное утверждение будет доказано ниже).

Примем, что y_j – кусочно-непрерывные функции времени. Запишем функцию Гамильтона для задачи (2)–(4), (14)

$$H = -\sqrt{k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2} + r_1 y_1 + r_2 y_2 + r_3 y_3$$

Для построения искомой функции $\mathbf{y}(t)$ и нахождения оптимальной функции $\mathbf{r}(t)$ сделаем замену переменных $u_j = y_j \sqrt{k_j}$; $\mu_j = r_j / \sqrt{k_j}$ и введем векторы $\mathbf{u} = \{u_1, u_2, u_3\}$ и $\boldsymbol{\mu} = \{\mu_1, \mu_2, \mu_3\}$. Функцию Гамильтона перепишем в виде:

$$H = \mathbf{u} \cdot \boldsymbol{\mu} - |\mathbf{u}| = |\mathbf{u}|(|\boldsymbol{\mu}| \cos \gamma - 1)$$

где γ – угол между векторами \mathbf{u} и $\boldsymbol{\mu}$. Функция Гамильтона H максимальна, если $\gamma = 0$ (когда векторы \mathbf{u} и $\boldsymbol{\mu}$ имеют одинаковое направление). Оптимальные функции y_j и r_j удовлетворяют равенствам

$$y_j = a r_j / k_j \quad (15)$$

где $a \geq 0$ – скалярная величина, кроме $a \equiv 0$.

Для всех t из промежутка $t \in [0, T]$ модуль оптимального вектора $\boldsymbol{\mu}$ остается постоянным $|\boldsymbol{\mu}| = \text{const}$. Оптимальные функции r_j удовлетворяют условию:

$$r_1^2 / k_1 + r_2^2 / k_2 + r_3^2 / k_3 = \text{const} \quad (16)$$

Убедимся в этом, продифференцировав по времени левую часть равенства (16) и заменив по формулам (13) производные компонент r_j вектора \mathbf{r} , а затем в полученном выражении компоненты y_i вектора \mathbf{y} заменяются с использованием соотношений (15), связывающим r_j и y_j . В результате получаем

$$\begin{aligned} & r_1 \dot{r}_1 / k_1 + r_2 \dot{r}_2 / k_2 + r_3 \dot{r}_3 / k_3 = \\ & = y_1 r_2 r_3 / k_3 - y_1 r_2 r_3 / k_2 + y_1 r_2 r_3 / k_2 - y_2 r_1 r_3 / k_3 + y_2 r_1 r_3 / k_3 - y_1 r_2 r_3 / k_3 \equiv 0 \end{aligned}$$

Получили ключевое свойство оптимального по критерию (14) процесса.

В задаче (2)–(4), (14) величина T не фиксирована. Поэтому оптимальные векторы \mathbf{u} и $\boldsymbol{\mu}$ должны удовлетворять условию трансверсальности (9). Чтобы было $H(T) = 0$, надо выполнить хотя бы одно из условий $|\boldsymbol{\mu}(T)| = 1$ или $|\mathbf{u}(T)| = 0$. Функция H является линейной по аргументу $|\mathbf{u}|$. Если $|\boldsymbol{\mu}(T)| > 1$, то при оптимальном решении $|\mathbf{u}(T)| \rightarrow \infty$, как того требует условие максимума функции Гамильтона. Но тогда $H \rightarrow \infty$, т.е. $H(T) \neq 0$, и условие трансверсальности (9) будет нарушено. Если $|\boldsymbol{\mu}(T)| < 1$, то в силу (16) условие максимума функции H в каждый текущий момент времени t требует, чтобы $|\mathbf{y}(t)| \equiv 0$; но в этом случае будет нарушено краевое условие (4). Единственным допустимым для оптимальных функций r_j вариантом является $|\boldsymbol{\mu}| = 1$. Учитывая свойство (16), приходим к выводу, что оптимальное значение $\mathbf{r}(0)$ обязано удовлетворять условию

$$r_1^2(0)/k_1 + r_2^2(0)/k_2 + r_3^2(0)/k_3 = 1 \quad (17)$$

Из свойств оптимальных функций r_j следует, что при оптимальном решении $y(t)$ максимум функции Гамильтона H не зависит от модуля $|y|$ (а зависит только от направления искомого вектора y), и равенство $H=0$ имеет место на всем отрезке времени $[0, T]$. Величина $|y|$ не влияет на выполнение необходимых условий оптимальности, а значит, значение функционала качества (14) не зависит от значения функции $a(t)$ в конкретный момент времени t .

Можем утверждать, что любое решение системы (2), (13), (15), удовлетворяющее краевым условиям (3), (4) и равенству (17), является оптимальным в смысле минимума интеграла (14). Условие (17) определяет значение ключевой характеристики $r_0 = |r| = |r(0)|$ для оптимальных функций $r_j(t)$. Линия, соединяющая точки x_0 и x_T , остается в конфигурационном гиперпространстве R^9 одной и той же независимо от вида масштабирующей функции $a(t)$, определяющей темп приближения к конечной точке x_T . Длина оптимальной кривой из точки x_0 в точку x_T , вычисленная в соответствии с (14), определяется в процессе решения краевой задачи принципа максимума одновременно с оптимальным значением $r(0)$. Она зависит только от x_0, x_T и коэффициентов k_1, k_2, k_3 .

Задача минимизации интеграла (14) свелась к решению системы уравнений (2), (13), (15) с одновременным выполнением краевых условий (3), (4) и требования (17). Сформулированная задача (2)–(4), (14) решается с точностью до неотрицательной функции $a(t)$.

Произвол в выборе функции $a(t)$ позволяет сделать так, чтобы построенная оптимальная функция $y(t)$ могла удовлетворять каким-то дополнительным условиям (не оговоренным в рамках задачи (2)–(4), (14)), что может оказаться весьма полезным при решении прикладных задач, связанных с оптимизацией физических процессов. Кроме требования $a(t) \geq 0$ функция $a(t)$, присутствующая в уравнениях (15), будет оптимальной тогда и только тогда, когда она удовлетворяет равенству

$$\int_0^T a(t) dt = G_{\text{opt}} \quad (18)$$

где G_{opt} – оптимальное значение интеграла (14), рассчитанное в результате решения краевой задачи принципа максимума (напомним, что $a(t)$ не может быть тождественно равной нулю).

Свойство инвариантности интеграла (14) по отношению к характеру поведения скалярной функции $a(t)$ – несомненное преимущество оптимизации по критерию (14). Кроме того, произвольность функции $a(t)$ и независимость значения (14) от величины $a(t)$ упрощает решение краевой задачи принципа максимума (нахождение оптимального значения $r(0)$) и определение характеристики (14) для оптимального процесса. Пользуясь указанным свойством независимости оптимальных значений $r(0)$ и (14) от функции $a(t)$, при решении краевой задачи принципа максимума мы можем принять $a(t) = \text{const} = a_0$. Тогда $G = a_0 T$, где величина T задает конечный момент времени для решения $x(t)$ системы (2). При этом функции y_1, y_2, y_3 вычисляются по выражениям (15) с учетом (13), (17) и условия $a = a_0$. Если $a_0 = 1$, то $G = T$. Отсюда следует, что мы можем интерпретировать характеристику (14) оптимального процесса (т.е. G) так: для оптимального закона изменения функции $y(t)$ обобщенные затраты G_{opt} численно равны величине интервала времени, необходимого для достижения равенства $x(t) = x_T$ для решения системы уравнений (2), (13), (17), (15), где $a = 1$ (начальное

состояние $\mathbf{x}(0)$ зафиксировано условием (3)). При произвольном поведении функции $a(t)$ условие (18) определяет продолжительность оптимального процесса T .

Данный пример наглядно иллюстрирует, как условия трансверсальности (и только они) дают возможность выявить ключевые свойства оптимального процесса. Для случая минимизации интеграла (14) такими отличительными свойствами оптимального решения являются равенство (17) и независимость значения (14) от неотрицательной, тождественно не равной нулю, функции $a(t)$, и, как следствие, от модуля $|y|$ в конкретный момент времени t . На примере задачи (2)–(4), (14) видно, что без условия трансверсальности (9) однозначно определить оптимальное значение вектора $\mathbf{r}(0)$ и решить краевую задачу принципа максимума было бы невозможно.

Еще одним примером является задача оптимального перевода динамической системы (2) из состояния (3) в состояние (4) с минимальным значением функционала вида

$$S = \int_0^T (k_0 + k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2) dt \quad (19)$$

где $k_0 > 0, k_1 > 0, k_2 > 0, k_3 > 0$ – постоянные положительные коэффициенты.

Принятый нами функционал (19) соответствует квадратичному критерию качества с добавлением затраченного времени, умноженного на постоянный положительный коэффициент ($k_0 \neq 0$). Фактор времени, присутствующий в критерии оптимальности (19), ограничивает длительность T оптимального процесса некоторым конечным значением T_{opt} .

Как и в предыдущей задаче, оптимальное решение $y(t)$ ищем в классе кусочно-непрерывных функций. Для задачи (2)–(4), (19) функция Гамильтона имеет вид:

$$G = r_1 y_1 + r_2 y_2 + r_3 y_3 - k_0 - k_1 y_1^2 - k_2 y_2^2 - k_3 y_3^2.$$

Поскольку G – квадратичная функция искомого вектора y , то её максимальное значение достигается в точке локального экстремума. Из необходимых условий экстремума $\partial G / \partial y_j = 0$ находим, что функция G максимальна, если $r_j - 2k_j y_j = 0$. Оптимальные функции y_j связаны с r_j соотношениями

$$y_j = r_j / 2k_j, \quad (20)$$

Сравнив условия оптимальности (20) с равенствами (15), заключаем, что для оптимального по критерию (19) решения справедливо свойство (16). Поэтому функция Гамильтона G является постоянной. Из условия трансверсальности (9), которое имеет вид $G(T) = 0$, находим необходимое условие оптимальности для функций $r_j(t)$.

$$(r_1^2 / k_1 + r_2^2 / k_2 + r_3^2 / k_3) / 4 - k_0 = 0$$

Следовательно, оптимальное значение $\mathbf{r}(0)$ обязано удовлетворять соотношению

$$r_1^2(0) / k_1 + r_2^2(0) / k_2 + r_3^2(0) / k_3 = 4k_0$$

Оптимальные функции $y_j(t)$ удовлетворяют соотношению

$$k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2 = k_0 \quad (21)$$

Из полученного свойства оптимальных функций $y_j(t)$ находим, что значение S функционала (19) связано с длительностью оптимального процесса соотношением $S = 2k_0 T$. Учитывая, что равенства (20) – частный случай условий (15), определим величину оптимальной длительности T_{opt} через известную характеристику G_{opt} , рассчитанную при решении задачи (2)–(4), (14): $T_{\text{opt}} = G_{\text{opt}} / \sqrt{k_0}$ (функции $y_j(t)$, оптимальные в смысле минимума (19), являются оптимальными по критерию (14)).

Соответственно значение показателя (19) при оптимальном поведении функций y_1, y_2, y_3 равно $S_{\text{opt}} = 2\sqrt{k_0} G_{\text{opt}}$. Значения коэффициентов k_1, k_2, k_3 влияют только на значение G_{opt} и определяют его совместно с заданными начальной x_0 и конечной x_T точками траектории $x(t)$.

Таким образом, рассмотренные примеры демонстрируют, что условия трансверсальности – исключительно эффективный инструмент в определении оптимальных функций $y(t), r(t)$. В ряде случаев они являются незаменимым (а порой и единственным) средством, чтобы выявить ключевые свойства оптимального процесса.

Заметим, что если под фазовыми переменными x_i понимать направляющие косинусы между связанной системой координат твердого тела и инерциальной системой координат, то вектор y будет соответствовать угловой скорости вращения связанной системы координат относительно инерциальной системы координат, а управляющие функции y_j интерпретируются как проекции вектора абсолютной угловой скорости твердого тела на оси связанной системы координат. В этом случае рассмотренные выше задачи оптимизации приобретают вполне конкретный физический смысл. Если при этом коэффициенты k_1, k_2, k_3 в задаче (2)–(4), (19) пропорциональны значениям J_1^2, J_2^2, J_3^2 , где J_1, J_2, J_3 – моменты инерции твердого тела, то условие трансверсальности $\Gamma(T) = 0$ определяет модуль кинетического момента во время оптимального по критерию (19) движения твердого тела из положения (3) в положение (4) (уравнение (21) наглядно подтверждает сказанное). Из уравнения (21) следует закономерность: для оптимального движения системы (2) с минимальным значением (19) длительность T перехода из состояния (3) в состояние (4) и значение интеграла (19) связаны прямо-пропорциональной зависимостью $S_{\text{opt}} = 2 k_0 T_{\text{opt}}$.

5. Заключение

Рассматриваемые в статье вопросы относятся к математической теории оптимальных процессов и ее методам. В частности, обсуждаются крайне актуальные аспекты методологии принципа максимума и особенности процедуры его практического применения. Проведенное исследование устраняет имеющийся методический пробел, связанный с неполным использованием условий трансверсальности.

Отмеченный методологический недостаток, присутствующий в рассуждениях при решении вариационных задач, наблюдается у многих авторов. В подавляющем большинстве случаев считается, что условия трансверсальности усложняют решение задачи поиска оптимальных функций, а не упрощают его. Такое мнение сложилось, видимо, из-за общеизвестной проблемы размерности, предполагающей, что чем больше условий (уравнений, ограничений), тем трудней решить исходную задачу (с математической или вычислительной точек зрения). Выше было показано, что это далеко не так и является общепринятым заблуждением, ставшим уже чуть ли не традиционным. Поскольку с увеличением числа уравнений, формализующих условия оптимальности, решение задачи оптимизации затрудняется, то размерность краевой задачи принципа максимума стараются сделать минимально возможной, из-за чего условия трансверсальности включают в систему необходимых условий оптимальности только в крайних случаях. А обычно ограничиваются лишь проверкой готового решения, найденного без учета условий трансверсальности, на предмет удовлетворения условиям трансверсальности (если того требует конкретная постановка оптимизационной задачи). По существу, такой подход аналогичен декомпозиции, которую успешно применяют в теории систем алгебраических уравнений, в теории больших систем (как

одного из разделов целого направления, именуемого «ИССЛЕДОВАНИЕ ОПЕРАЦИЙ») и других дисциплинах. Однако повторим, что гораздо целесообразнее и грамотнее (с чисто математической точки зрения) изначально сформировать полную систему уравнений, записанных в соответствии с принципом максимума, включая условия трансверсальности (как необходимые условия оптимальности наравне с условием максимума функции Гамильтона и дифференциальными уравнениями для сопряженных переменных), а затем уже внутри полученной системы уравнений производить декомпозицию, в том числе, выделив в отдельную независимую подсистему условия трансверсальности (если это становится возможным и допустимым). Вопрос о выгоде разделения условий оптимальности на отдельные группы решается в каждом конкретном случае самостоятельно, исходя из структуры исходной полной системы уравнений (включая условия трансверсальности), определяющей оптимальные функции. Справедливости ради заметим, что большое число прикладных задач оптимизации допускает оговоренную декомпозицию и выделение условий трансверсальности в отдельную подсистему. Тем не менее важно подчеркнуть, что известны нередкие примеры практических задач (в том числе, некоторые задачи оптимального управления), для которых условия трансверсальности являются незаменимым математическим средством в определении не только количественных характеристик оптимальных функций (отдельных параметров, констант, коэффициентов и т.д.), но и ключевых свойств и уравнений оптимального процесса.

В работе убедительно продемонстрировано, что для максимально корректного и полного решения задач оптимального управления условия трансверсальности, являясь необходимыми условиями оптимальности наравне с условием максимума функции Гамильтона и сопряженной системой дифференциальных уравнений, должны учитываться с самого начала. Они должны быть включены в систему уравнений, формализующих принцип максимума, сразу после введения сопряженных переменных, записи функции Гамильтона и сопряженной системы уравнений, непосредственно определяя свойства оптимальных сопряженных функций и управляющих переменных. Рассмотренные в статье конкретные примеры указывают на исключительно важную роль (нередко ключевую роль) условий трансверсальности в общем алгоритме процедуры применения принципа максимума.

Список литературы:

1. Л.С. Понтрягин, В.Г. Болтянский, Р.В. Гамкрелидзе, Е.Ф. Мищенко. Математическая теория оптимальных процессов. М.: Наука, 1983.
2. Н.Н. Моисеев. Численные методы в теории оптимальных систем. М.: Наука, 1971.

Левский Михаил Валерьевич, кандидат технических наук, старший научный сотрудник; ведущий научный сотрудник Государственного Космического научно-производственного центра им. М.В. Хруничева (Научно-исследовательский институт Космических систем, Московская область). Автор более 150 научных трудов (научные статьи) и автор 25 изобретений и патентов, лауреат премии С.П. Королева, международный Педагог Года 2007; представлен в изданиях Who's Who in Science and Engineering, Who's Who in the World, 2006–2008, 2014–2016 гг. Специалист в области теории оптимального управления и приборов и систем автоматического управления. Сфера научных интересов: механика космического полета, теория систем ориентации, навигации и стабилизации летательных аппаратов, движение вокруг центра масс космических аппаратов, разгонных блоков и орбитальных станций.

About transversality conditions for solving variation problems

M.V. Levskii

Research Institute of Space systems, Khrunichev State research and production Space Center
Tihonravov street, 27, Korolev, Moscow region, 141092, Russia 27

Abstract. The article concerns variational problems and methods of the mathematical theory of optimal processes. The role, place and significance of transversality conditions in the general procedure of finding the optimal functions are studied in details taking Pontryagin's maximum principle as an example. This work eliminates the existing methodical defect associated with incomplete use of the transversality conditions. Using specific examples, the transversality conditions (as one of the necessary conditions of optimality in the form of the maximum principle) have been shown to be an extremely efficient mathematical tool (and even the only one in some cases) for determining the characteristic properties, laws and key characteristics (parameters, constants, integrals of motion) of optimal solution to the systems of ordinary differential equations (of dynamic systems).

Key words: optimality criterion, maximum principle, transversality conditions, Hamiltonian, conditions of optimality, boundary value problem

1. Relevance of the problem

When solving many optimization problems by the variational methods (e.g., using Pontryagin's maximum principle), significance of transversality conditions has been clearly underestimated so far. As a rule, the transversality conditions are not used properly and do not participate in determination of quantitative characteristics of optimal process (parameters, constants, first integrals, etc.), they are not involved in computation of optimal functions. Therefore it is difficult to estimate completeness of such solution derived using the procedure of the maximum principle without transversality conditions. The aim of this work is to provide understanding that transversality conditions are not a mere formality participating in procedure of finding the optimal time functions (which are present in the system of necessary optimality conditions formulated according to the maximum principle), but they are an efficient tool (but in some cases the only and irreplaceable one) for determination of characteristics and properties of optimal process, for calculation of its key parameters (constants).

When constructing an optimal process, many authors and researchers (mainly in applied areas) only check transversality conditions (as the necessary conditions for optimality) at the end of solution procedure after the solution has been already obtained and formalized with certain mathematical relations. Such a test is made by substitution of the obtained optimal solution to the initial system of equations and the conjugate functions into equations corresponding to transversality conditions. Meanwhile, the transversality conditions along with the condition of Hamiltonian maximum and the conjugate system of differential equations are necessary conditions of optimality, and they must be included in the system of necessary conditions of optimality since the very beginning (together with the conjugate system of equations and the condition of Hamiltonian maximum), after defining the conjugate variables, the Hamiltonian and the conjugate system of equations. Such an algorithm and further application of the procedure of maximum principle is the most correct and expedient. As the formalism of maximum principle application demonstrates, in many cases the inclusion of transversality conditions into the general system of necessary conditions of optimality before formalization of laws of dynamic system optimal motion described by the

initial system of differential equations (given by statement of optimization problem) is the only correct one and it allows finding the solution to the most complicated problems. In this case, solution to the boundary value problem of the maximum principle will take into account the transversality conditions, it will provide reliability of the stated optimization problem solution results and guarantee the completeness of the obtained solution.

2. Problem statement

Usually, the dynamical process is determined by a differential equation

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y}, t) \quad (1)$$

where \mathbf{x} is the state vector (its components x_i are the phase variables, which reflect the current state of the dynamic system, $i = \overline{1, n}$); \mathbf{y} is the control vector with components y_j , which are the sought functions of time (we should choose their optimum, $j = \overline{1, l}$). Physical sense of the specific optimization problem is determined by the form of equations $\dot{x}_i = f_i(\mathbf{x}, \mathbf{y}, t)$. Let the dynamical process be described by the following system of equations

$$\begin{aligned} \dot{x}_1 &= y_3 x_4 - y_2 x_7, \quad \dot{x}_2 = y_3 x_5 - y_2 x_8, \quad \dot{x}_3 = y_3 x_6 - y_2 x_9 \\ \dot{x}_4 &= y_1 x_7 - y_3 x_1, \quad \dot{x}_5 = y_1 x_8 - y_3 x_2, \quad \dot{x}_6 = y_1 x_9 - y_3 x_3 \\ \dot{x}_7 &= y_2 x_1 - y_1 x_4, \quad \dot{x}_8 = y_2 x_2 - y_1 x_5, \quad \dot{x}_9 = y_2 x_3 - y_1 x_6 \end{aligned} \quad (2)$$

The system (2) determines variation of the time functions $x_i(t)$ under known variation of the functions $y_j(t)$ ($i = \overline{1, 9}, j = \overline{1, 3}$). The motion $\mathbf{x}(t)$ is the solution to the differential equation (1). To make the problem more specific, it is necessary to make some restrictions concerning the solution to the equation (1), in particular, to set boundary conditions at initial time $t = 0$ and final time $t = T$. Let us consider a widespread class of problems when the boundary conditions for the trajectory of motion $\mathbf{x}(t)$ are formalized by the equalities

$$\mathbf{x}(0) = \mathbf{x}_0 \quad (3)$$

$$\mathbf{x}(T) = \mathbf{x}_T \quad (4)$$

where $\mathbf{x}_0, \mathbf{x}_T$ are the fixed vectors, which specify the state of a dynamic system at initial and final time.

Note that in most cases the equation (1) has infinitely many solutions satisfying conditions (3) and (4). From all the solutions to the system of equations (1), (3), (4), we are interested only in the solutions that are optimal (or best) in some sense. Usually, the criterion of optimality is expressed by the functional of quality

$$G = \int_0^T g(\mathbf{x}, \mathbf{y}, t) dt \quad (5)$$

where $g(\mathbf{x}, \mathbf{y}, t)$ is the non-negative function; T is the termination time of the optimal process. In the most general case, the optimization problem is formulated as follows: it is necessary to transfer the dynamic system described by the equation (1) from the state (3) into the state (4) so as to minimize the functional (5).

3. Procedure of solving

If we suppose that functions y_j in the right-hand parts of equations (2) are piecewise continuous functions of time, then we can use Pontryagin's maximum principle [1].

According to this method, we define the conjugate variables q_i corresponding to the components x_i of the state vector \mathbf{x} and write a Hamilton function (Hamiltonian).

$$H = \varphi_0 g + q_1 (y_3 x_4 - y_2 x_7) + q_2 (y_3 x_5 - y_2 x_8) + q_3 (y_3 x_6 - y_2 x_9) + \\ + q_4 (y_1 x_7 - y_3 x_1) + q_5 (y_1 x_8 - y_3 x_2) + q_6 (y_1 x_9 - y_3 x_3) + \\ + q_7 (y_2 x_1 - y_1 x_4) + q_8 (y_2 x_2 - y_1 x_5) + q_9 (y_2 x_3 - y_1 x_6)$$

where $\varphi_0 = \text{const} < 0$ ($\varphi_0 \neq 0$).

According to the maximum principle [1], the equations for conjugate functions q_i have the form

$$\dot{q}_i = -\partial H / \partial x_i \quad (i = \overline{1,9} \quad j = \overline{1,3}) \quad (6)$$

The conjugate functions q_i must satisfy the transversality conditions

$$q_i(0) = \sum_{h=1}^m w_h \left. \frac{\partial F_h(\mathbf{x}(0))}{\partial x_i} \right|_{t=0}; \quad q_i(T) = \sum_{k=1}^p z_k \left. \frac{\partial \Phi_k(\mathbf{x}(T))}{\partial x_i} \right|_{t=T} \quad (7)$$

where $F_h(\mathbf{x}(0))$ are the functions specifying the boundary condition at the left endpoint of the trajectory $\mathbf{x}(t)$ in the form $F_h(\mathbf{x}(0)) = 0$, and $h = \overline{1,m}$ ($m \leq n$); w_h are certain numbers that do not vanish simultaneously; $\Phi_k(\mathbf{x}(T))$ are the functions specifying the boundary condition for the right endpoint of the trajectory $\mathbf{x}(t)$ in the form $\Phi_k(\mathbf{x}(T)) = 0$, and $k = \overline{1,p}$ ($p \leq n$); z_k are certain numbers that do not vanish simultaneously; for the problem stated as (2)–(5), the number of phase variables is $n = 9$.

If the ending time T of the optimal process is not fixed, then the optimal conjugate functions $q_i(t)$ must satisfy yet another necessary condition expressed as the transversality condition

$$H(\mathbf{x}(T), \mathbf{y}(T), \mathbf{q}(T), T) = - \sum_{k=1}^p v_k \frac{\partial \Phi_k(\mathbf{x}(T), T)}{\partial T} \quad (8)$$

where $\mathbf{q} = \{ q_1, q_2, \dots, q_9 \}$ is the vector of conjugate variables; $\Phi_k(\mathbf{x}(T), T)$ are the functions specifying the boundary condition for the right endpoint of the trajectory $\mathbf{x}(t)$ in the form $\Phi_k(\mathbf{x}(T), T) = 0$, and $k = \overline{1,p}$ ($p \leq n$); v_k are certain numbers that do not vanish simultaneously ($n = 9$, and $l = 3$ for the problem (2)–(5)).

From the structure of the Hamiltonian, equations (6) and transversality conditions (7), (8) we conclude that optimal functions q_i are proportional to the negative constant φ_0 . Hence, we can assume $\varphi_0 = -1$.

Since the considered problem (2)–(5) is a problem with fixed left and right endpoints of the trajectory $\mathbf{x}(t)$, the transversality conditions (7) for the conjugate variables q_i at $t = 0$ have the form (see [2]): $q_i(0) = w_i$ ($i = \overline{1,9}$), where w_1, w_2, \dots, w_9 are certain numbers that do not vanish simultaneously. For the conjugate variables q_i at the time $t = T$, the transversality conditions (7) have similar form (see [2]): $q_i(T) = z_i$ ($i = \overline{1,9}$), where at least one number z_1, z_2, \dots, z_9 is non-zero; and the transversality condition (8) has the form

$$H(T) = 0 \quad (9)$$

First two conditions are equivalent to the inequalities

$$\sum_{i=1}^9 q_i^2(0) > 0 \quad \text{и} \quad \sum_{i=1}^9 q_i^2(T) > 0 \quad (10)$$

and they are obvious if we rewrite boundary conditions (3), (4) as equalities

$$x_i(0) - x_{i0} = 0, \text{ and } x_i(T) - x_{iT} = 0$$

where x_{i0} are elements of the vector \mathbf{x}_0 ; x_{iT} are elements of the vector \mathbf{x}_T ; whence $F_k = x_i(0) - x_{i0}$; $\Phi_k = x_i(T) - x_{iT}$; $\partial F_k / \partial x_i = \partial \Phi_k / \partial x_i = 1$ for $k = i$, and $\partial F_k / \partial x_i = \partial \Phi_k / \partial x_i = 0$ if $k \neq i$.

The last transversality condition (requirement (9)) is obtained if we represent the boundary condition (4) for the trajectory right endpoint in the form $\Phi_k = 0$, where $\Phi_k = x_i(T) - x_{iT}$; whence we have $\partial \Phi_k / \partial T \equiv 0$ for all k , it is obvious, therefore conditions (8) and (9) are equivalent.

The problem of finding the optimal function $\mathbf{y}(t)$ and optimal motion $\mathbf{x}(t)$ of the system (2) is reduced to solving equations (2), (6) with the simultaneous maximization of the Hamiltonian H at each moment of time t and with satisfaction of transversality conditions (10) (if the time T is fixed) or (9), (10) (if the time T is not fixed).

The boundary value problem of the maximum principle is to find such initial values $q_i(0)$ for conjugate functions at which the boundary condition (4) is satisfied for the solution to the system of differential equations (2), (6) $\mathbf{x}(t)$, $\mathbf{q}(t)$ (with initial condition (3)) that satisfies the condition of the Hamiltonian maximum at each moment of time t and transversality conditions (10) at initial and final moments of time (or (9), (10) when the time T is not fixed). Let us suppose that the functional (5) does not include state variables x_i . In this case, the conjugate system of differential equations for q_i is written as

$$\begin{aligned} \dot{q}_1 &= y_3 q_4 - y_2 q_7, \quad \dot{q}_2 = y_3 q_5 - y_2 q_8, \quad \dot{q}_3 = y_3 q_6 - y_2 q_9, \\ \dot{q}_4 &= y_1 q_7 - y_3 q_1, \quad \dot{q}_5 = y_1 q_8 - y_3 q_2, \quad \dot{q}_6 = y_1 q_9 - y_3 q_3, \\ \dot{q}_7 &= y_2 q_1 - y_1 q_4, \quad \dot{q}_8 = y_2 q_2 - y_1 q_5, \quad \dot{q}_9 = y_2 q_3 - y_1 q_6 \end{aligned} \quad (11)$$

It is easy to show that in this particular version of problem statement (2)–(5) (when functional (5) does not explicitly depend on phase variables x_i), the transversality conditions $\mathbf{q}(0) \neq 0$ и $\mathbf{q}(T) \neq 0$ are satisfied automatically if the function $\mathbf{q}(t)$ is not identically equal to zero (\mathbf{q} is the vector of conjugate variables q_i). To check this, it is sufficient to analyse the sum

$$\sum_{i=1}^9 q_i^2(t) \quad (12)$$

Taking the time derivative of the sum (12) and replacing quantities \dot{q}_i by right parts of equations (11) in the obtained expression, we obtain the equality

$$\sum_{i=1}^9 \dot{q}_i q_i \equiv 0$$

Hence, we conclude that the sum (12) is constant. Therefore, if it is assumed that at least one condition (10) is not satisfied, then the vector \mathbf{q} of conjugate variables will be identically equal to zero $\mathbf{q} \equiv 0$; but such conjugate functions $q_i(t)$ can not be optimum, it is obvious. Thus, any solution $\mathbf{q} \neq 0$ satisfying the system (11) will satisfy transversality conditions (7).

The systems of equations (2) and (11) are identical in terms of coefficients y_1 , y_2 , and y_3 . The solutions to equations (2) for variables x_i and to equations (11) for variables q_i differ in initial conditions.

We group multipliers for y_1 , y_2 , y_3 in the function H so that they contain state variables x_i , and we denote such multipliers as r_1 , r_2 , and r_3 . Taking multipliers r_j for y_j in an explicit form, we rewrite the Hamiltonian H as follows:

$$H = -g + y_1 r_1 + y_2 r_2 + y_3 r_3$$

where $r_1 = q_4 x_7 + q_5 x_8 + q_6 x_9 - q_7 x_4 - q_8 x_5 - q_9 x_6$;

$$r_2 = q_7 x_1 + q_8 x_2 + q_9 x_3 - q_1 x_7 - q_2 x_8 - q_3 x_9;$$

$$r_3 = q_1 x_4 + q_2 x_5 + q_3 x_6 - q_4 x_1 - q_5 x_2 - q_6 x_3.$$

For variables r_i , we have the following differential equations:

$$\dot{r}_1 = y_3 r_2 - y_2 r_3, \quad \dot{r}_2 = y_1 r_3 - y_3 r_1, \quad \dot{r}_3 = y_2 r_1 - y_1 r_2 \quad (13)$$

Taking the time derivative of equalities for r_j ($j = \overline{1,3}$) and substituting expressions for \dot{x}_i and \dot{q}_i calculated by (2) and (11), respectively, we obtain necessary differential equations for time functions $r_j(t)$.

$$\begin{aligned} \dot{r}_1 = & q_4 \dot{x}_7 + \dot{q}_4 x_7 + q_5 \dot{x}_8 + \dot{q}_5 x_8 + q_6 \dot{x}_9 + \dot{q}_6 x_9 - q_7 \dot{x}_4 - \dot{q}_7 x_4 - q_8 \dot{x}_5 - \dot{q}_8 x_5 - \\ & - q_9 \dot{x}_{23} - \dot{q}_{33} x_6 = y_2 (q_4 x_1 + q_5 x_2 + q_6 x_3 - x_4 q_1 - x_5 q_2 - x_6 q_3) + \\ & + y_3 (q_7 x_1 + q_8 x_2 + q_9 x_3 - x_7 q_1 - x_8 q_2 - x_9 q_3) = y_3 r_2 - y_2 r_3 \end{aligned}$$

(we have similar expressions for variables r_2 and r_3). Variation of the vector $\mathbf{r} = \{r_1, r_2, r_3\}$ is determined by the solution to the equation $\dot{\mathbf{r}} = \mathbf{r} \times \mathbf{y}$, where $\mathbf{y} = \{y_1, y_2, y_3\}$.

The equations (13) for components r_j of the vector \mathbf{r} replace the conjugate system of equations (11). We see from (13) that the vector \mathbf{r} has the property $|\mathbf{r}| = \text{const}$. For existence of non-trivial solution, it is necessary that equalities $x_i = \chi q_i$ do not hold at any time $t \in [0, T]$ since in this case $r_1 = r_2 = r_3 \equiv 0$, and solving the problem becomes meaningless. Therefore, searching for the optimal solution, we assume vectors \mathbf{x} and \mathbf{q} such that $\mathbf{x} \neq \chi \mathbf{q}$, where χ is a scalar (otherwise $\mathbf{r} \equiv 0$), and only the case $\mathbf{r} \neq 0$ is considered (i.e., when functions q_i are not proportional to variables x_i). It is obvious, that $\mathbf{q}(t) \neq 0$ if $\mathbf{r} \neq 0$, and transversality conditions (7) are met.

The problem of finding the optimal function $\mathbf{y}(t)$ and the optimal motion $\mathbf{x}(t)$ of the system (2) is reduced to solving equations (2), (13) with the simultaneous maximization of the Hamiltonian H at each moment of time t and with satisfaction of the transversality condition (9) if the time T is not fixed. The boundary value problem of the maximum principle is to find such a vector $\mathbf{r}(0) \neq 0$ at which the solution to the system of equations (2), (13), $\mathbf{x}(t)$ (with initial condition (3)) that meets the condition of the Hamiltonian maximum at each moment of time t and the transversality condition (9) (for the problems with not fixed time T) satisfies the boundary condition (4).

4. Examples of problems when transversality conditions determine key characteristics of an optimal process

As an example, we give solution to the problem (2)–(5), in which the minimized functional (5) has the form:

$$G = \int_0^T \sqrt{k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2} dt \quad (14)$$

where $k_1 > 0$, $k_2 > 0$, and $k_3 > 0$ are constant positive coefficients.

A feature of the minimized functional (14) is that its value is independent of optimal process duration T ; the value G is determined only by boundary conditions (by vectors \mathbf{x}_0 and \mathbf{x}_T) and by coefficients k_1 , k_2 , and k_3 (this conclusion will be proved further).

We assume that y_j are piecewise continuous functions of time. For the problem (2)–(4), (14), the Hamiltonian takes the form

$$H = -\sqrt{k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2} + r_1 y_1 + r_2 y_2 + r_3 y_3$$

In order to construct a sought-for function $\mathbf{y}(t)$ and find an optimal function $\mathbf{r}(t)$, we replace variables $u_j = y_j \sqrt{k_j}$; $\mu_j = r_j / \sqrt{k_j}$, and introduce vectors $\mathbf{u} = \{u_1, u_2, u_3\}$ and $\boldsymbol{\mu} = \{\mu_1, \mu_2, \mu_3\}$. We can rewrite the Hamiltonian in the following form

$$H = \mathbf{u} \cdot \boldsymbol{\mu} - |\mathbf{u}| = |\mathbf{u}|(|\boldsymbol{\mu}| \cos \gamma - 1)$$

where γ is the angle between vectors \mathbf{u} and $\boldsymbol{\mu}$. The Hamiltonian H takes the maximum value at $\gamma = 0$ (i.e. when vectors \mathbf{u} and $\boldsymbol{\mu}$ have the same direction). Optimal functions y_j and r_j satisfy the equalities

$$y_j = a r_j / k_j \quad (15)$$

where $a \geq 0$ is a scalar, but it should not be $a \equiv 0$.

At any moment of time $t \in [0, T]$, the absolute value of the optimal vector $\boldsymbol{\mu}$ is constant $|\boldsymbol{\mu}| = \text{const}$. The optimal functions r_j satisfy the condition

$$r_1^2 / k_1 + r_2^2 / k_2 + r_3^2 / k_3 = \text{const} \quad (16)$$

To prove the validity of this property, we take the time derivative of the left-hand side of the equality (16) and substitute formulas (13) for the derivatives of components r_j of the vector \mathbf{r} ; next, in the resulting expression we replace components y_i of the vector \mathbf{y} with their expressions given by (15), which give relation between r_j and y_j . This yields

$$r_1 \dot{r}_1 / k_1 + r_2 \dot{r}_2 / k_2 + r_3 \dot{r}_3 / k_3 =$$

$$= y_1 r_2 r_3 / k_3 - y_1 r_2 r_3 / k_2 + y_1 r_2 r_3 / k_2 - y_2 r_1 r_3 / k_3 + y_2 r_1 r_3 / k_3 - y_1 r_2 r_3 / k_3 \equiv 0$$

As a result, we derive a key property of an optimal process in terms of the criterion (14).

In the problem (2)–(4), (14), the process end time T is not fixed. Therefore, optimal vectors \mathbf{u} and $\boldsymbol{\mu}$ must satisfy the transversality condition (9). To meet the equality $H(T) = 0$, it is necessary to satisfy at least one of conditions $|\boldsymbol{\mu}(T)| = 1$ or $|\mathbf{u}(T)| = 0$. The Hamiltonian H is a linear function of argument $|\mathbf{u}|$. If $|\boldsymbol{\mu}(T)| > 1$, then for the optimal solution we have $|\mathbf{u}(T)| \rightarrow \infty$ (since the Hamiltonian must be maximum), but under such vector \mathbf{u} we have $H \rightarrow \infty$, i.e. $H(T) \neq 0$, and the transversality condition (9) is violated. If $|\boldsymbol{\mu}(T)| < 1$, then, due to the property (16), the condition of the Hamiltonian maximum H at each moment of time t requires the equality $|\mathbf{y}(t)| \equiv 0$, which implies that the boundary condition (4) can never be satisfied. For optimal functions r_j , the only feasible option is then $|\boldsymbol{\mu}| = 1$. Taking into account (16), we conclude that the optimal value $\mathbf{r}(0)$ must satisfy the condition

$$r_1^2(0)/k_1 + r_2^2(0)/k_2 + r_3^2(0)/k_3 = 1 \quad (17)$$

The properties of optimal functions r_j imply that, in the case of the optimal solution $\mathbf{y}(t)$, the maximum of the Hamiltonian H is independent of the absolute value $|\mathbf{y}|$ (it depends only on the direction of the sought-for vector \mathbf{y}), and the equality $H = 0$ holds for the entire interval of time $[0, T]$. The magnitude $|\mathbf{y}|$ does not affect the validity of necessary optimality conditions; therefore, the value of functional (14) is independent of the value of the function $a(t)$ (at the given time t).

We claim that any solution to the system (2), (13), (15) satisfying boundary conditions (3), (4) and the equality (17) is optimal in the sense of minimum of the integral (14). The condition (17) determines value of key characteristic $r_0 = |\mathbf{r}| = |\mathbf{r}(0)|$ for optimal functions $r_j(t)$. The line, which connects points \mathbf{x}_0 and \mathbf{x}_T , remains invariable in configuration hyperspace R^9 irrespective of the form of the scaling function, $a(t)$, which determines the speed of approach

to the final point \mathbf{x}_T . The length of an optimum curve from the point \mathbf{x}_0 to the endpoint \mathbf{x}_T calculated according to (14), is determined in the course of solving the boundary value problem of the maximum principle simultaneously with the optimum value $\mathbf{r}(0)$ and depends only on the boundary values $\mathbf{x}_0, \mathbf{x}_T$, and coefficients k_1, k_2 , and k_3 .

The problem of minimization of the integral (14) is reduced to solving the system of equations (2), (13), (15) with the simultaneous satisfaction of boundary conditions (3), (4) and the requirement (17). The formulated problem (2)–(4), (14) is solved up to the nonnegative function $a(t)$.

The arbitrariness in the choice of the function $a(t)$ makes it possible to ensure that the designed function $\mathbf{y}(t)$ satisfies additional conditions (which were not mentioned in the statement of the problem (2)–(4), (14)), it can be rather useful for the solution to applied problems associated with optimization of physical processes. Except for the constraint $a(t) \geq 0$, the function $a(t)$, which appears in equations (15), is optimal if and only if it satisfies the equality

$$\int_0^T a(t) dt = G_{\text{opt}} \quad (18)$$

where G_{opt} is the optimal value of the integral (14), which is calculated in the course of solving the boundary value problem of the maximum principle (remember that $a(t)$ cannot be identically equal to zero).

The property that the integral (14) is invariant relative to the behavior of the scalar function $a(t)$, is a doubtless benefit of optimization by the criterion (14). Furthermore, arbitrariness of the function $a(t)$ and independence of the value (14) of quantity $a(t)$ simplifies solution to the boundary value problem of the maximum principle (finding the optimal value $\mathbf{r}(0)$) and determination of the characteristic (14) for the optimal process. Using this specified property of independence of optimal values $\mathbf{r}(0)$ and (14) of the function $a(t)$, when solving the boundary value problem of the maximum principle, we can assume $a(t) = \text{const} = a_0$. Then $G = a_0 T$, where T is the time of satisfying the condition (4) for the solution $\mathbf{x}(t)$ to the system (2), at which functions y_1, y_2 , and y_3 are calculated by (15) taking into account (13), (17) and condition $a = a_0$. If $a_0 = 1$, then $G = T$. Therefore, we can interpret the characteristic (14) of optimal process as follows: for the optimal law of the function $\mathbf{y}(t)$ variation, generalized expenditures G_{opt} are numerically equal to the time of satisfying the equality $\mathbf{x}(t) = \mathbf{x}_T$ for solution to the system of equations (2), (13), (17), (15) in which $a = 1$ (the initial state $\mathbf{x}(0)$ is fixed by the condition (3)). For the arbitrary behavior of the function $a(t)$, condition (18) determines the duration of optimal process T .

The example presented above illustrates the fact that transversality conditions (and only them) make it possible to specify key properties of optimal process. For the case when the integral (14) is minimized, such unique properties of optimal solution are the equality (17) and independence of value (14) of the nonnegative function $a(t)$, which is not identically equal to zero, and as a consequence, of the absolute value $|\mathbf{y}|$ at specific moment of time t . Using the problem (2)–(4), (14) as an example, we demonstrate the following conclusion: it would be impossible to unequivocally specify the optimal value of the vector $\mathbf{r}(0)$ and to solve the boundary value problem of the maximum principle without the transversality condition (9).

Another example is the problem of optimal transfer of the dynamic system (2) from the state (3) to the state (4) with minimal value of the functional

$$S = \int_0^T (k_0 + k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2) dt \quad (19)$$

where $k_0 > 0$, $k_1 > 0$, $k_2 > 0$, and $k_3 > 0$ are constant positive coefficients.

The functional (19), which we have assumed, corresponds to the quadratic criterion of quality with addition of transfer time multiplied by the constant positive coefficient ($k_0 \neq 0$). Time factor, which is present in the criterion of optimality (19), limits duration T of optimal process to a certain value T_{opt} .

Like in the previous problem, we search for the optimal solution $\mathbf{y}(t)$ in the class of piecewise continuous functions. For the problem (2) - (4), (19), the Hamiltonian is

$$\Gamma = r_1 y_1 + r_2 y_2 + r_3 y_3 - k_0 - k_1 y_1^2 - k_2 y_2^2 - k_3 y_3^2.$$

Since Γ is a quadratic function of the vector \mathbf{y} , its maximum value is attained at the local extremum point. Using necessary conditions of extremum $\partial \Gamma / \partial y_j = 0$, we find that the function Γ has a maximum if $r_j - 2k_j y_j = 0$. Hence, optimal functions y_j correlate with r_j by expressions

$$y_j = r_j / 2k_j, \quad (20)$$

Having compared conditions of optimality (20) with equalities (15), we conclude that the property (16) is true for optimal solution in terms of the criterion (19). Therefore the Hamiltonian Γ is constant. From the transversality condition (9), which has the form $\Gamma(T) = 0$, we find the necessary condition of optimality for functions $r_j(t)$.

$$(r_1^2 / k_1 + r_2^2 / k_2 + r_3^2 / k_3) / 4 - k_0 = 0$$

Hence, the optimal value $\mathbf{r}(0)$ must satisfy the equation

$$r_1^2(0) / k_1 + r_2^2(0) / k_2 + r_3^2(0) / k_3 = 4k_0$$

The optimal functions $y_j(t)$ satisfy the relation

$$k_1 y_1^2 + k_2 y_2^2 + k_3 y_3^2 = k_0 \quad (21)$$

From the obtained property of optimal functions $y_j(t)$, we find that the value S of the functional (19) correlates with duration of optimal process by the formula $S = 2k_0 T$. Taking into account that equalities (20) are a particular case of conditions (15), we determine the optimal time T_{opt} using known characteristic G_{opt} , which is calculated in the course of solving the problem (2)–(4), (14): $T_{\text{opt}} = G_{\text{opt}} / \sqrt{k_0}$ (functions $y_j(t)$, which are optimal in the sense of minimum (19), are optimum for the criterion (14)). Accordingly, value of the index (19) under optimal variation of functions y_1 , y_2 , and y_3 is equal $S_{\text{opt}} = 2\sqrt{k_0} G_{\text{opt}}$. The values of coefficients k_1 , k_2 , and k_3 influence only the value G_{opt} and determine it together with the given initial \mathbf{x}_0 and final \mathbf{x}_T endpoints of the trajectory $\mathbf{x}(t)$.

Thus, the considered examples demonstrate that transversality conditions are very efficient for determining optimal functions $\mathbf{y}(t)$, $\mathbf{r}(t)$, and in some cases, they are irreplaceable (sometimes, it is the only one tool) to find key properties of an optimal process.

Note that, if phase variables x_i are considered as directing cosines between the coordinate system of a solid body and the inertial coordinate system, then the vector \mathbf{y} will correspond to the angular velocity of the body's coordinate system rotating relative to the inertial coordinate system, and control functions y_j are considered as projections of the vector of absolute angular velocity on the axes of the body coordinate system. In this case, the optimization problem considered above acquires specific physical sense. Moreover, if coefficients k_1 , k_2 , and k_3 in the problem (2)–(4), (19) are proportional to values J_1^2 , J_2^2 , and J_3^2 , where J_1 , J_2 , and J_3 are moments of inertia of the solid body, then the transversality condition $\Gamma(T) = 0$ determines the absolute value of angular momentum during optimal rotation of the solid body from the

position (3) into the position (4), in terms of criterion (19) (the equation (21) directly confirms the derived conclusion). From the equation (21), we have the following regularity: for optimal motion of the system (2) with minimal value of (19), duration of transfer from the state (3) to the state (4) and value of the integral (19) are related by direct-proportional dependence $S_{\text{opt}} = 2 k_0 T_{\text{opt}}$.

5. Conclusions

Aspects considered in the article concern the mathematical theory of optimal processes and its methods. In particular, extremely topical aspects of methodology of the maximum principle and issues of its practical application are discussed. The presented work eliminates the existing methodical defect associated with incomplete use of transversality conditions.

The above methodological defect, which is present in solution to variational problems, is observed in many papers. In overwhelming majority of cases, transversality conditions are assumed to complicate the search for optimum functions instead of simplification. Such an opinion is probably due to a well-known problem of dimension, which assumes that the more conditions there are (equations, constraints), the more difficult it is to solve the initial problem (in the mathematical or computational sense). Above we have shown that this is far from it, and it is a widespread misperception, which has already become almost traditional.

Since the increase of number of equations formalizing the conditions of optimality leads to complication of the optimization problem solution, researchers try to make the dimension of the boundary value problem of the maximum principle as low as possible, hence transversality conditions are included into the system of necessary conditions of optimality only in extreme cases. As a rule, the final solution found without transversality conditions is only checked for meeting transversality conditions (if that is required by specific statement of optimization problem). Essentially, such an approach is similar to decomposition which is successfully applied in the theory of systems of algebraic equations, in the theory of complicated systems (as a section of OPERATIONAL RESEARCH) and other disciplines. However, it is much more expedient and competent (from mathematical point of view) to initially build a full system of equations formulated according to the maximum principle, including transversality conditions (as necessary conditions of optimality as well as the condition of the Hamiltonian maximum and differential equations for conjugate variables), and then to make decomposition within the obtained system of equations, for example, to separate independent subsystem of transversality conditions (if it becomes possible and admissible). In each particular case, the question about benefit of separating the transversality conditions into separate groups is solved individually depending on the structure of initial full system of equations (including the transversality conditions), which determines optimal functions. To be fair, we note that many applied optimization problems permit the mentioned decomposition and localization of transversality conditions in a separate subsystem. Nevertheless, it is important to stress that there are some applied practical problems (including some problems of optimal control) for which the transversality conditions are irreplaceable mathematical means of determination of not only numerical characteristics of optimal functions (certain parameters, constants, coefficients, and others) but also of key properties and equations of optimal process.

In the present article, we have convincingly proved the following conclusion: for the most correct and complete solution to optimal control problems, the transversality conditions, being the necessary conditions of optimality similarly to the condition of the Hamiltonian maximum and the conjugate system of differential equations, should be taken into account since the very beginning and included into the system of equations that formalize the maximum principle

directly after defining the conjugate variables, the Hamiltonian and the conjugate system of equations, when determining the properties of optimal conjugate functions and control variables. Specific examples considered in the article confirm the crucial role of transversality conditions in general algorithm of the maximum principle application.

References

1. L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze and E.F. Mishchenko. *Matematicheskaya teoriya optimal'nykh protsessov* [Mathematical theory of optimal processes], Moscow: Nauka, 1983. (In Russian)
2. N.N. Moiseev. *Chislennyye metody v teorii optimal'nykh sistem* [Numerical methods in theory of optimal systems]. Moscow: Nauka, 1971. (In Russian)

Mikhail Valer'evich Levskii, PhD (engineering science), senior researcher; leading researcher at Khrunichev State Research-and-Production Space Center (Research Institute of Space Systems, Moscow region). Published more than 150 scientific papers (scientific articles), author of 25 inventions (Russian patents), laureate of Korolev Prize, International Educator of the Year 2007; member of the Marquis Who's Who family (9th and 10th Anniversary Editions of **Who's Who in Science and Engineering**, 25th Silver Anniversary Edition of **Who's Who in the World**, and **Who's Who in the World** 2014, 2015, 2016). He specializes in theory of optimal control and automatic control systems and devices. Area of scientific interests: mechanics of spaceflight, theory of attitude control systems, navigation and stabilization of flying vehicles, motion around a center of mass of spacecrafts, booster units and orbital stations.

Сравнение различных двигателей и методы повышения термического КПД

В.И. Панченко, Ли Цзывань

Казанский национальный исследовательский технический университет
им. А.Н.Туполева (КНИТУ-КАИ),
Россия, 420111, Казань, К. Маркса, 10

Аннотация. Рассмотрены термический коэффициент полезного действия и циклы четырёх разных типов тепловых двигателей. Сделан вывод, что даже у самого современного теплового двигателя термический коэффициент полезного действия не высок, и есть много возможностей для его увеличения. Из рассмотрения термодинамических циклов следует, что для увеличения термического коэффициента полезного действия можно либо повысить температуру (и соответственно, давление) высокотемпературного источника тепла (тепловой энергии), либо снизить температуру (давление) низкотемпературного источника тепла. У современных тепловых двигателей обычно применяется первый способ. Но этот метод требует применения новых дорогих материалов, новых технологий охлаждения. В статье рассмотрено другое направление увеличения термического коэффициента полезного действия двигателя: снизить температуру (давление) низкотемпературного источника тепла. В статье рассмотрены конкретные способы увеличения термического коэффициента полезного действия воздушно-реактивных двигателей, например, с помощью эжектора.

Ключевые слова: тепловой двигатель, воздушно-реактивный двигатель, термический коэффициент полезного действия (КПД), ресурс, снижение температуры (давления), эжектор

1. Введение

Тепловые двигатели незаменимы для современного производства и инженерно-технических применений. Они широко используются в авиационном, морском, наземном транспорте и других случаях. Тепловые двигатели имеют различное устройство, у каждого типа двигателя – свои преимущества и недостатки. Предмет анализа – рабочие циклы, термические коэффициенты полезного действия (КПД), преимущества, недостатки и сферы применения четырёх различных типов тепловых двигателей. Обсуждаются способы увеличения термических коэффициентов полезного действия. Отметим, что с точки зрения моделирования тепловые двигатели – это сложные системы междисциплинарного характера. Здесь необходимы высокий профессионализм и понимание, основанные на знаниях из различных областей науки.

2. Анализ некоторых типов тепловых двигателей

2.1. Воздушно-реактивный двигатель

Схема и цикл воздушно-реактивного двигателя (ВРД) в диаграммах $p-v$ и $T-s$ представлен на рисунке 1.

На схеме показаны основные узлы воздушно-реактивного двигателя: воздухозаборник – 1, компрессор – 2, камера сгорания – 3, турбина – 4, реактивное сопло – 5. Сечения на входе и выходе каждого узла обозначены следующим образом:

Н – сечение невозмущенного потока перед двигателем;

В – на выходе из воздухозаборника (на входе в компрессор);

К – на выходе из компрессора (на входе в камеру сгорания);

Г – на выходе из камеры сгорания (на входе в турбину);

Т – на выходе из турбины (на входе в сопло);

С – на выходе из сопла.

Параметры состояния рабочего тела в этих сечениях в термодинамических диаграммах $p-v$ и $T-s$ соответствуют точкам на схеме, обозначенным этими же буквами (p – давление, v – удельный объем, T – абсолютная температура, s – удельная энтропия рабочего тела).

Здесь термический коэффициент полезного действия воздушно-реактивного двигателя, как и других тепловых двигателей, это – отношение удельного количества теплоты, превращенного в удельную механическую энергию, к удельному количеству теплоты, подведенному к рабочему телу [4]:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{T_c - T_H}{T_\Gamma - T_K} = 1 - 1 / \pi_\Sigma^{\frac{k-1}{k}},$$

где: q_1 и q_2 – удельная теплота, подведенная к рабочему телу (воздуху) и отведенная от него; T_c , T_H , T_Γ , T_K – температуры рабочего тела в сечениях С, Г, Н, К; $\pi_\Sigma = p_K^*/p_H$ – суммарная степень повышения давления в цикле; p_K^* – полное давление в сечении К; p_H – статическое давление в сечении Н; k – отношение теплоемкостей при постоянном давлении и постоянном объеме.

Из приведенной формулы видно, что чем больше степень повышения давления (π_Σ), тем больше и термический коэффициент полезного действия. Степень повышения давления современного воздушно-реактивного двигателя достигает 80, идеальный термический коэффициент полезного действия равен 0.714 [5]. Цикл воздушно-реактивного двигателя в диаграммах $p-v$ и $T-s$ показан на рисунке 1. Но, до сих пор, реальный термический коэффициент полезного действия меньше 0.5.

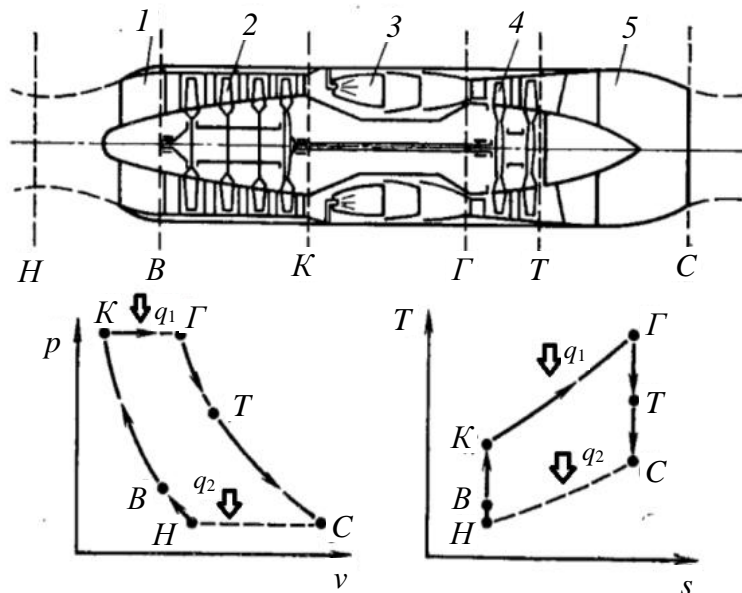


Рис. 1. Цикл ВРД в диаграммах $p-v$ и $T-s$

2.2. Двигатель внутреннего сгорания

2.2.1. Бензиновый двигатель

Цикл бензинового двигателя в диаграммах $p-v$ и $T-s$ показан на рис. 2. Точки 1,2,3,4 на этих диаграммах соответствуют параметрам состояния рабочего тела: начальному

состоянию (точка 1), в конце процесса сжатия (точка 2), в конце процесса подвода теплоты q_1 (точка 3), в конце процесса расширения (точка 4). Концу теплоподвода q_2 и возвращению рабочего тела в первоначальное состояние соответствует точка 1.

Термический коэффициент полезного действия для бензинового двигателя [4]:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{T_4 - T_1}{T_3 - T_2} = 1 - 1 / \left(\frac{v_1}{v_2} \right)^{k-1} = 1 - 1 / \varepsilon^{k-1},$$

где ε – степень сжатия, $\varepsilon = v_1/v_2$, (v_1, v_2 – удельные объемы рабочего тела в состояниях 1 и 2) т.е. чем выше степень сжатия, тем больше термический коэффициент полезного действия. Степень сжатия современного бензинового двигателя достигает 18, термический коэффициент полезного действия равен ~ 0.685 . Но из-за потерь фактический термический коэффициент полезного действия бензинового двигателя около 0.4. Цикл бензинового двигателя в диаграммах $p-v$ и $T-s$ показан на рисунке 2.

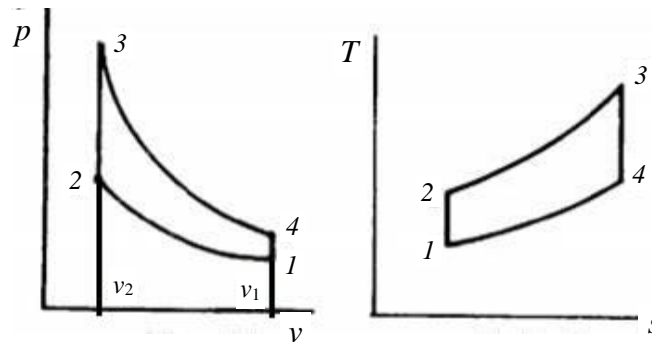


Рис. 2. Цикл бензинового двигателя в диаграммах $p-v$ и $T-s$

Преимуществами бензинового двигателя являются: большой ресурс, небольшой объем, вес и малый шум. Они широко применяются как стационарные, автотракторные, судовые, тепловозные, авиационные и др.

2.2.2. Дизельный двигатель

Термический коэффициент полезного действия [4]:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{c_v(T_4 - T_1)}{c_p(T_3 - T_2)} = 1 - \frac{\rho^k - 1}{k(\rho - 1)\varepsilon^{k-1}},$$

где $\rho = v_3/v_2$, $\varepsilon = v_1/v_2$, (ρ – степень предварительного расширения рабочего тела).

Степень повышения давления современного дизельного двигателя достиг 25, а фактический термический коэффициент полезного действия – 0.55. Так, дизельный двигатель B&W компании S80ME-C7 тратит только 155 г на кВт*ч [6].

Цикл дизельного двигателя в диаграммах $p-v$ и $T-s$ показан на рисунке 3.

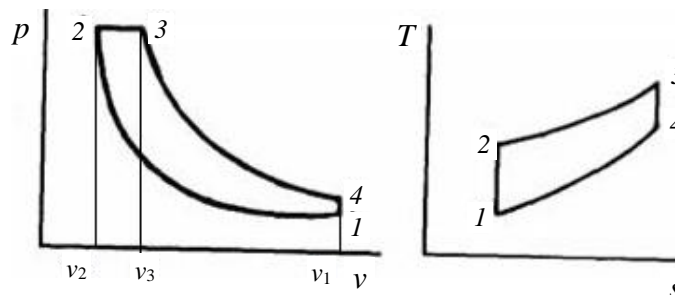


Рис. 3. Цикл дизельного двигателя

Преимуществами дизельного двигателя являются: высокий КПД, высокий крутящий момент. Дизельные двигатели обычно применяются для привода стационарных силовых установок, на рельсовых (тепловозы, дизелевозы, дизель-поезда, автодрезины) и безрельсовых (автомобили, автобусы, грузовики) транспортных средствах, самоходных машинах и механизмах (тракторы, комбайны, асфальтовые катки, скреперы и т. д.), а также в судостроении в качестве главных и вспомогательных двигателей.

2.3. Двигатель Стирлинга

Двигатель Стирлинга (рисунок 4) использует цикл Стирлинга (рисунок 5), который по термодинамической эффективности не уступает циклу Карно и даже обладает преимуществом. Дело в том, что цикл Карно состоит из мало отличающихся между собой изотерм и адиабат. Практическое воплощение этого цикла малоперспективно. Цикл Стирлинга позволил получить работающий на практике двигатель с приемлемыми размерами. Термический коэффициент полезного действия двигателя Стирлинга:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{T_2}{T_1}$$

Фактический термический коэффициент полезного действия современного двигателя Стирлинга достигает 0.47 [7]

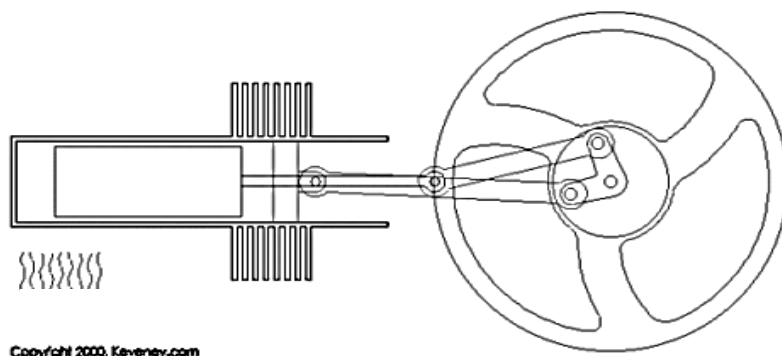


Рис. 4. Двигатель Стирлинга

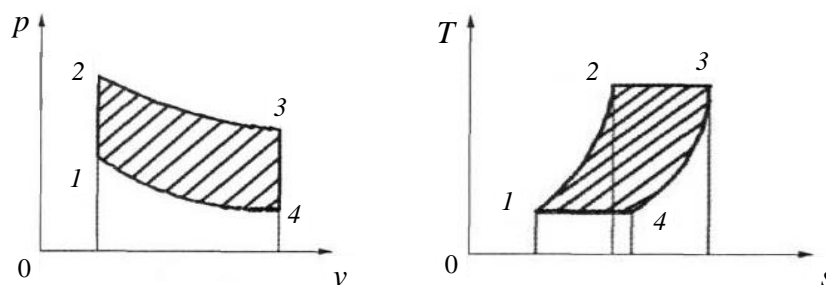


Рис.5. Цикл Стирлинга

Преимуществами двигателя Стирлинга являются: «всеядность» двигателя (любое топливо), простота конструкции, Увеличенный ресурс, экономичность, экологичность. Двигатели Стирлинга обычно применяются на субмаринах, солнечных электростанциях и космических кораблях.

2.4. Паровая машина Ползунова

Паровая машина Ползунова (рисунок 6) является первым в мире двухцилиндровым паровым двигателем. Принцип работы: поршни в цилиндрах паровой машины Ползунова работали в противофазе. Когда в один из цилиндров наполненных паром, впрыскивали воду, пар конденсировался в нем, и в цилиндре создавалось разрежение. Под действием атмосферного давления поршень начинал опускаться вниз. В этот момент, в другой цилиндр, находившийся в нижнем положении, начинал поступать пар, и он начинал двигаться вверх.

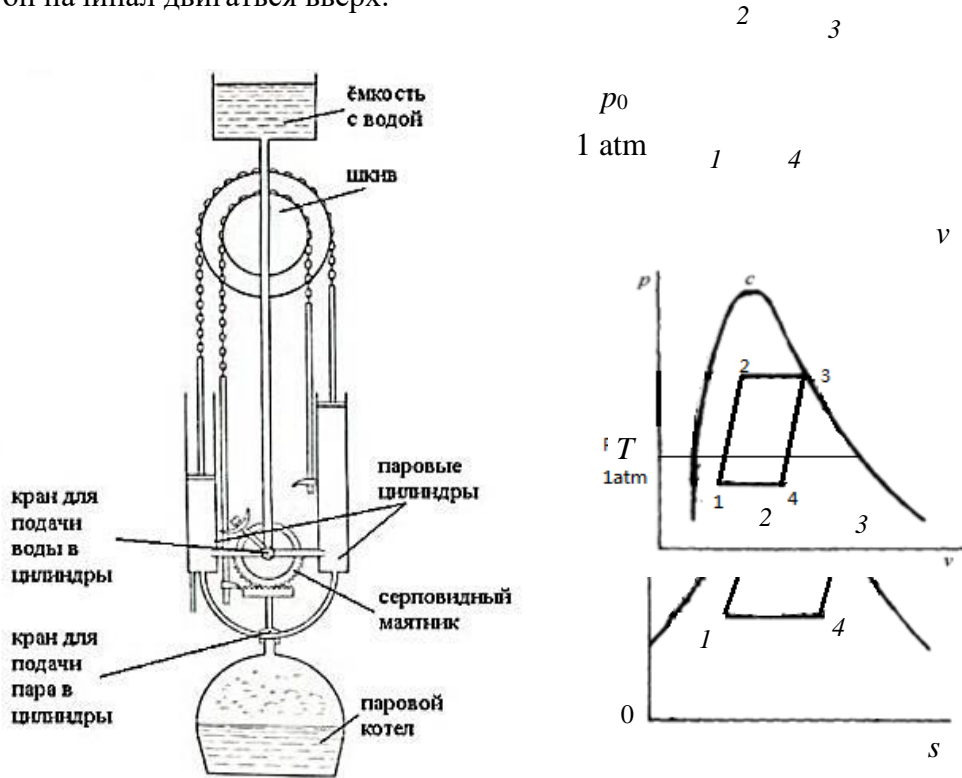


Рис. 6. Паровая машина Ползунова

Термический коэффициент полезного действия паровой машины Ползунова невелик из-за малой разницы температур и давлений между высокотемпературным источником тепла и низкотемпературным. В настоящее время такой двигатель не применяется.

3. Способы повышения КПД

3.1. Повышение температуры (давления) высокотемпературного источника тепла

При одной и той же температуре (и, соответственно, при одном и том же давлении) низкотемпературного источника тепла повышение температуры (давления) высокотемпературного источника тепла позволяет увеличить площадь рабочего цикла и увеличить полезную работу (рисунок 7). Этот способ широко применяется в разных двигателях. Но требуется применение новых дорогих материалов. С другой стороны, уменьшается ресурс двигателя.

На рис. 7 точки 1, 2, 3, 4 соответствуют сечениям Н, К, Г, С на рис. 1. При повышении температуры (давления) точки 2, 3 на рис. 7 перемещаются в положение 2', 3'.

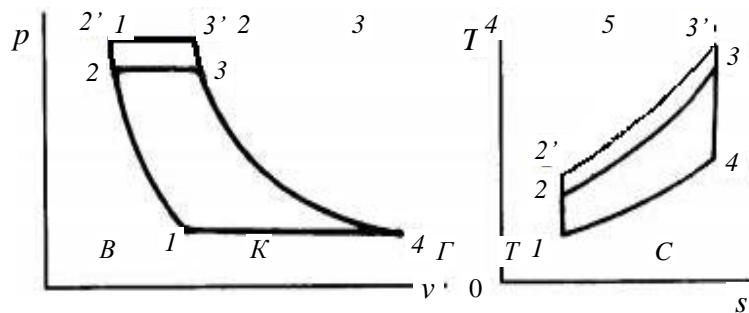


Рис. 7. Сравнение разных условий высокотемпературного источника тепла

3.2. Понижение температуры (давления) низкотемпературного источника тепла

При одной и той же температуре (и, соответственно, при одном и том же давлении) высокотемпературного источника тепла, понижение температуры (давления) низкотемпературного источника тепла тоже позволяет увеличить площадь цикла, т.е. увеличить работу цикла (рисунок 8). Но такой способ пока мало применяется на практике.

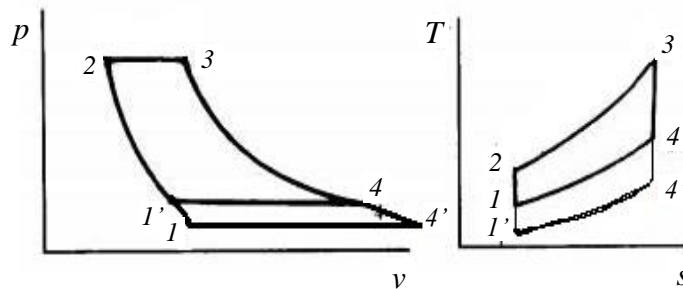


Рис. 8. Сравнение разных условий низкотемпературного источника тепла

4. Способ повышения КПД для ВРД

Для воздушно-реактивного двигателя обычно повышают термический КПД через увеличение температуры перед турбиной. В современных двигателях температура перед турбиной воздушно-реактивного двигателя достигает 2000 К. Это очень «жесткое» условие для рабочих лопаток турбины. Этот способ требует применения новых материалов.

Другим способом является понижение температуры (давления) за соплом. Т.е. создают зону низкого давления и температуры за соплом. Впрыск воды на выходе сопла или постановка эжектора на выходе сопла (рисунок 9) позволит увеличить степень расширения газа на турбине и увеличить скорость газа на срезе сопла. Преимуществами такого способа являются: увеличение тяги ВРД при M меньше 0,4 (M – число Маха полета), увеличение КПД двигателя, большой ресурс и малая стоимость.

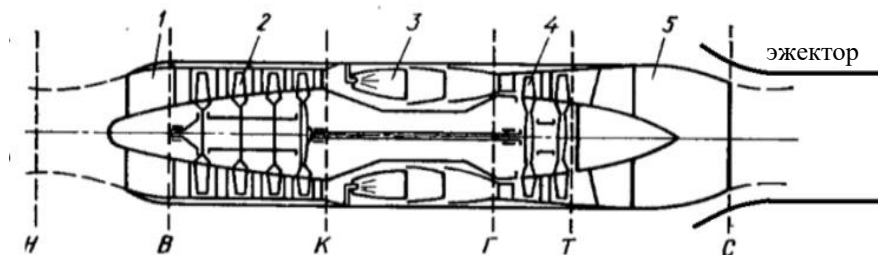


Рис. 9. Создание зоны низкого давления с помощью эжектора

Все обозначения на рис. 9 такие же, как и на схеме воздушно-реактивного двигателя на рис. 1.

5. Заключение

В работе рассмотрены направления увеличения термического коэффициента полезного действия тепловых двигателей разных типов:

- за счет повышения температуры высокотемпературного источника тепла;
- за счет понижения температуры низкотемпературного источника тепла.

Преимуществами второго способа увеличения термического коэффициента полезного действия воздушно-реактивного двигателя являются:

- увеличение тяги воздушно-реактивного двигателя при числах Маха полета меньших 0,4;
- увеличение ресурса двигателя;
- уменьшение стоимости двигателя.

Аббревиатуры

ВРД – воздушно-реактивный двигатель;
КПД – коэффициент полезного действия.

Узлы ВРД

- 1 – воздухозаборник;
- 2 – компрессор;
- 3 – камера сгорания;
- 4 – турбина;
- 5 – реактивное сопло.

Сечения потока

- Н – невозмущенный поток перед двигателем;
- В – выход из воздухозаборника (вход в компрессор);
- К – выход из компрессора (вход в камеру сгорания);
- Г – выход из камеры сгорания (вход в турбину);
- Т – выход из турбины (вход в сопло);
- С – выход из сопла.

Литература

1. В.М. Акимов, В.И. Бакулев, Р.И. Курзинер и др. Теория и расчет воздушно-реактивных двигателей . - М.: Машиностроение, 1987. - 568 с.
2. Шен Вэй-Дао, Тонг Чуен-Гэн. Техническая термодинамика / Шен Вэйдао, Тонг Чуенгэн. - М.: Высшее образование, 2006.
3. Г.Н.Абрамович. Прикладная газовая динамика. - М.: Наука, 1991. - 600 с.
4. В.В. Нащокин, Техническая термодинамика и теплопередача. - М.: Высшая школа, 1975. - 496с.
5. Wikipedia-overall pressure ratio, 2011 http://en.wikipedia.org/wiki/Overall_pressure_ratio
6. Wikipedia-diesel engine, 2008 http://en.wikipedia.org/wiki/Diesel_engine
7. Ли Хай-Вэй. Развитие и применение двигателей Стирлинга / Ли Хай-Вэй, Ши Лин-Со, Ли Яй-Чи: Технология Энергии, 2012.

Владимир Иванович Панченко, кандидат технических наук, профессор (кафедра реактивных двигателей и энергических установок (РДЭУ), Казанский национальный исследовательский технический университет им. А.Н. Туполева-КАИ). Специалист в области механики жидкости и газа.

Ли Цзывань, магистрант (кафедра РДЭУ, КНИТУ-КАИ), научная тема - исследование газодинамической стабилизации пламени, научный руководитель - доцент КНИТУ-КАИ В.Л. Варсегов

Comparison of different engines and the methods of increasing the thermal efficiency

V.I. Panchenko, Li Ziwan

Kazan National Research Technical University named after A.N. Tupolev
K. Marx, 10, Kazan, 420111, Russia

Abstract. It is considered the thermal efficiency of four typical heat engines and their working cycles. We came to conclusion that even the most modern heat engine doesn't have enough thermal efficiency whereas there are a lot of opportunities to increase it. Examining thermodynamic cycles we found that if we want to improve thermal efficiency we can increase temperature (and correspondingly pressure) of high temperature heat source, or reduce temperature (pressure) of low temperature heat source. The first method is usually used in modern heat engines. But this method requires new expensive materials and cooling technologies. The findings of our research illustrate another way how to improve thermal efficiency - decrease in temperature (pressure) of low temperature heat source. As a result, we offer detailed technologies of thermal efficiency improvement in jet engines, using an ejector to create a low pressure zone behind the nozzle.

Keywords: Heat engine, jet engine, the thermal efficiency, costs, reducing the temperature (pressure), ejector.

1.Introduction

Thermal engines are indispensable engineering construction for modern production and applications. They are widely used in the aviation, maritime, land transport and other areas. Heat engines have different forms, each of ones has its advantages and disadvantages. The article analyzes the working cycles, thermal efficiency, advantages, disadvantages and application areas of four different types of heat engines. Different ways to increase the thermal efficiency of heat engines are considered.

2. Analysis of different heat engines

2.1.The jet engine

The state parameters in these sections in terms of thermodynamic diagrams, p-v and T-s (here: p – pressure, v – specific volume, T is the absolute temperature, s is specific entropy working body) correspond to the points on the diagram marked with the same letters. Scheme and cycle jet engine diagrams p-v and T-s are shown in figure 1. The diagram shows the main nodes of the air jet engine: air-intake – 1, compressor 2, combustion chamber – 3, turbine 4, the jet nozzle – 5. Cross-section at the inlet and exit of each node are marked as follows: N- cross section of the unperturbed flow before the engine; In-at the exit of the inlet (at the inlet to the compressor); K – on the outlet of the compressor (inlet to the combustion chamber); Mr. at the exit of the combustion chamber (turbine inlet); T – at the exit of the turbine (inlet nozzle); - At the exit of the nozzle. The state parameters in these sections in terms of thermodynamic diagrams, p-v and T-s (here: p – pressure, v – specific volume, T is the absolute temperature, s is specific entropy working body) correspond to the points on the diagram marked with the same letters.

As all know, the thermal efficiency of jet engine [4]:

$$\eta_t = 1 - \frac{q_c}{q_1} = 1 - \frac{T_c - T_H}{T_r - T_K} = 1 - 1 / \pi_\Sigma^{\frac{k-1}{k}}.$$

That means, the higher the overall pressure ratio, the higher the thermal efficiency. The overall pressure ratio of modern jet engine can reach 80, the ideal thermal efficiency reach 0,714 [5]. The working cycle of jet engine in the diagrams p-v and T-s are shown in Figure 1. But, in fact, the real thermal efficiency of jet engine is less than 0.5.

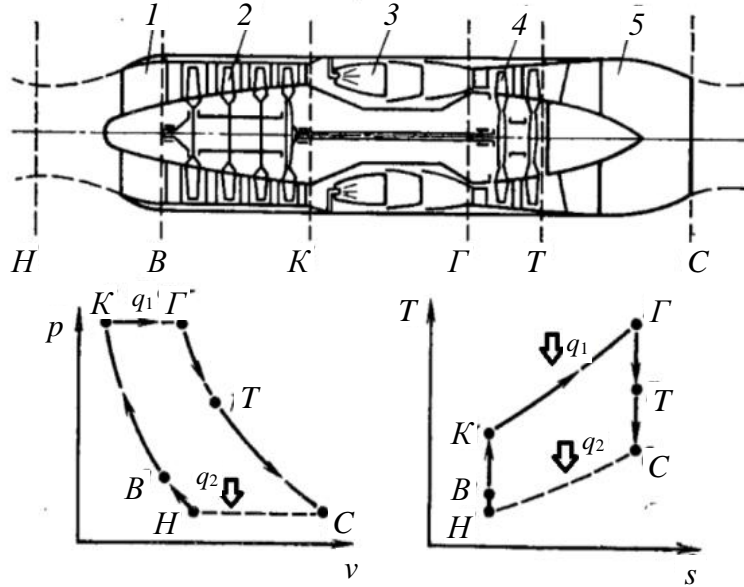


Fig. 1. The cycle of the WFD in diagrams p-v and T-s

2.2. Internal combustion engine

2.2.1. Gasoline engine. The thermal efficiency of gasoline engine:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{T_4 - T_1}{T_3 - T_2} = 1 - 1 / \left(\frac{v_1}{v_2} \right)^{k-1} = 1 - 1 / \varepsilon^{k-1}$$

[4] ε – the compression ratio, $\varepsilon = v_1 / v_2$.

That means, the greater the compression ratio, the greater the thermal efficiency. The compression ratio of the modern gasoline engine reach 18, and the thermal efficiency can reach the value 0.685 [5]. But, because of various losses, the actual thermal efficiency of gasoline engines is about 0.4

The working cycle of gasoline engine in the diagrams p-v and T-s are shown in Figure 2.

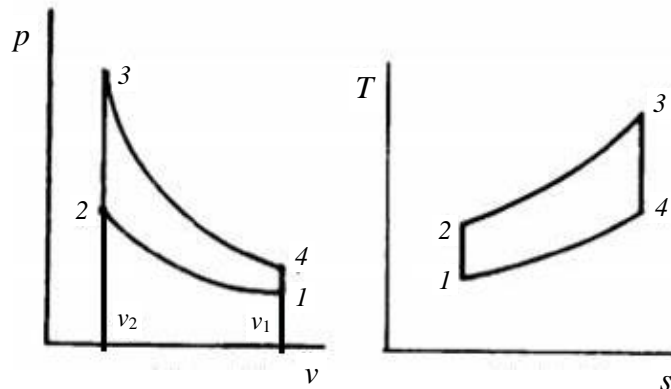


Fig. 2. Cycle gasoline engine in diagrams p-v and T-s

The advantages of gasoline engines: long service life, small volume and weight, and low noise. They are widely used as stationary, automotive, marine, aviation and other transports.

2.2.2. Diesel engine

The thermal efficiency of gasoline engine[4]:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{c_v(T_4 - T_1)}{c_p(T_3 - T_2)} = 1 - \frac{\rho^k - 1}{k(\rho - 1)\epsilon^{k-1}} \cdot \rho = v_3 / v_2, \epsilon = v_1 / v_2.$$

The compression ratio of the modern diesel engine can reach 25, and the actual thermal efficiency can reach the value 0.55. The diesel engine S80ME-C7 of company B & W, spends only – 155 g/ kWh, efficiency reaching – 54.4% [6]

The working cycle of diesel engine in the diagrams p-v and T-s are shown in Figure 3.

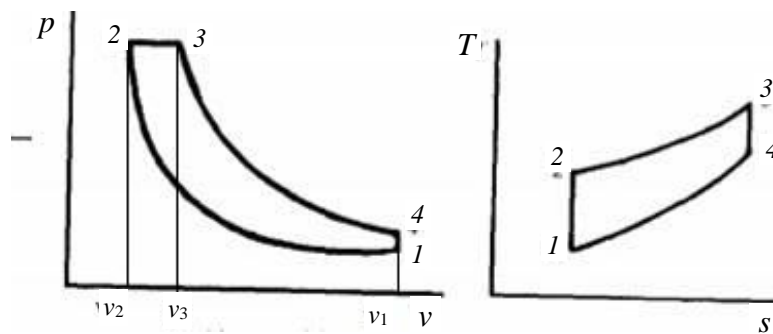


Fig. 2. Cycle diesel engine

The advantages of the diesel engine are: high efficiency, high torque. Diesel engines are usually used to drive stationary power plants, the rail traffic (locomotives, diesel trains, rail cars) and trackless transportation (cars, buses, trucks), self-propelled machines and mechanisms (tractors, combines, rollers, scrapers, etc).

2.3. Stirling engine

Stirling engine (Figure 4) uses a Stirling cycle (Figure 5), which is not inferior to of the Carnot cycle by the thermodynamic efficiency, and even has an advantage. In fact, the Carnot cycle consists of a slightly different between the isotherms and adiabatic. The practical realization of this cycle is unpromising. Stirling engine can work in practice to a reasonable size. The thermal efficiency of Stirling engine:

$$\eta_t = 1 - \frac{q_2}{q_1} = 1 - \frac{T_2}{T_1}.$$

The actual thermal efficiency of Stirling engine can reach 0.47 [7].

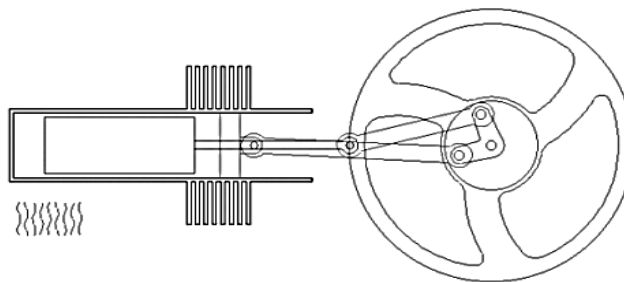


Fig. 4. Stirling Engine

The advantages of the Stirling engine are: omnivorous engine (any fuel), simple structure, long service life, economy, ecology. Stirling engines are commonly used on submarines, solar power stations and spacecraft.

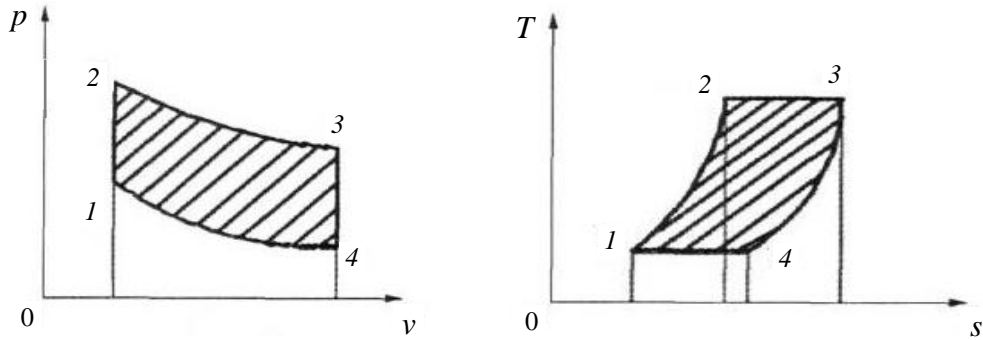


Fig. 5. Stirling Cycle

2.4. Polzunov's steam engine

Polzunov steam engine (Figure 6) is the first two-cylinder Steam engine in the world. Working principle: piston in the cylinder of Polzunov steam engine worked opposite phase. When one cylinder filled with steam, inject water to the cylinder and then vapor condenses there and a vacuum is created in the cylinder. Under atmospheric pressure, the piston starts to descend. At this point, the other cylinder is in the lower position, steam begins to flow in, and piston begins to move up.

The thermal efficiency of Polzunov steam engine is not high, because of the small difference of the temperature (pressure) between high-temperature heat source and the low-temperature heat source. Currently, this engine is not yet used.

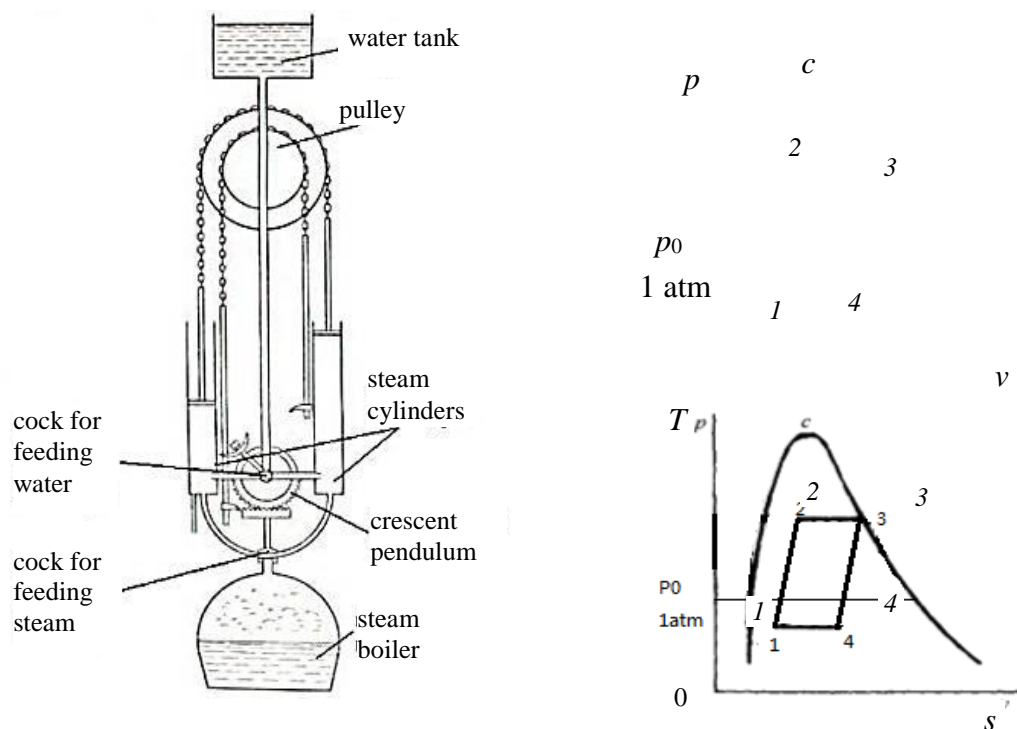


Fig. 6. Polzunov's steam engine

3. Methods to improve efficiency

3.1. Improving the temperature and pressure of the high-temperature heat source

Keeping the same temperature (pressure) of low-temperature heat source, increasing the temperature (pressure) of high-temperature heat source can increase the area of the working cycle and increase the useful work (Figure 7). This method is widely used in different engines. But it requires the use of expensive new materials. On the other hand, it may reduce the service life of engines.

In Fig.7 points 1,2,3,4 correspond to the points H,K, Γ ,C in Fig. 1. When the temperature (pressure) points 2, 3 in Fig. 7 move to position 2', 3'.

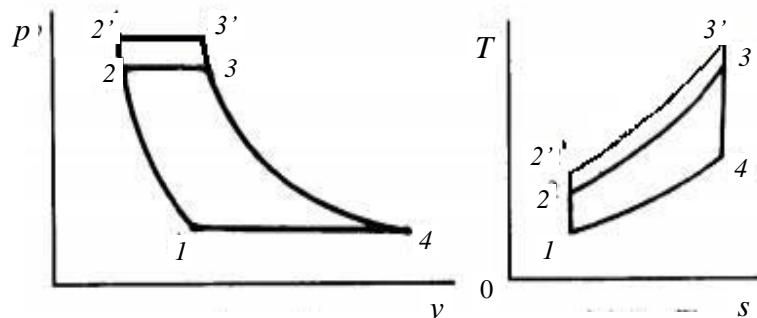


Fig. 7. Comparison of different condition in high temperature heat source

3.2. Decrease the temperature and pressure of the low-temperature heat source

Keeping the same temperature (pressure) of high- temperature heat source, decreasing the temperature (pressure) of low-temperature heat source can also increase the area of working cycle and increase the useful work (Figure 8). But, nowadays this method is not widely used in practice.

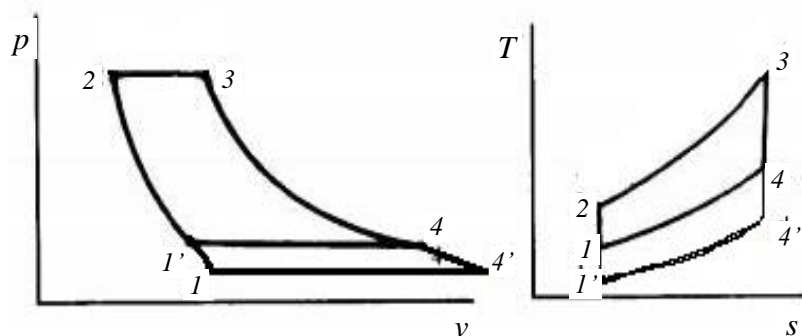


Fig. 8. Comparison of different condition in low temperature heat source

4. Methods for improving the efficiency of jet engine

For a jet engine, generally increase the thermal efficiency by increasing the temperature in the inlet of turbine. In modern engines, the temperature at the inlet of turbine reaches 2000K. This is a very strict condition for turbine blades. And requires the use of new materials

Another direction improving the efficiency of jet engine is to reduce the temperature (pressure) behind the nozzle. That means create a zone of lower pressure and temperature behind the nozzle. For example, injecting water on the outlet of nozzle or put a ejector at the outlet of nozzle, allows increase the expansion ratio of the gas in the turbine and increase gas

velocity at the outlet of nozzle. The advantages of this method are: increased the thrust force (at less than 0.4 Ma), increase the efficiency of jet engine, long service life and low cost.

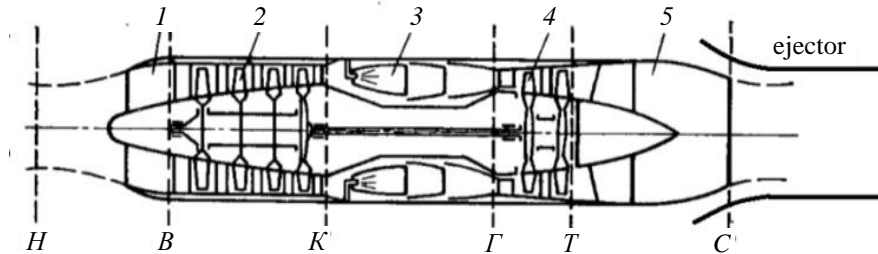


Fig. 9. Creating a zone of low pressure by using the ejector

All designations in Fig. 9 are the same as in the scheme of the WFD in Fig. 1.

Nodes of the WFD: 1 – air-intake, 2 – compressor, 3 – combustion chamber, 4 – turbine, 5 – the jet nozzle.

Conclusion

In the work we offer detailed technologies of thermal efficiency improvement in jet engines, using an ejector to create a low pressure zone behind the nozzle, using two ways:

- to increase temperature (and correspondingly pressure) of high temperature heat source,
- to reduce temperature (pressure) of low temperature heat source.

References

1. V.M.Akimov, Teoriya i raschet vozdušno-reaktivnyh dvigatelej / V.M.Akimov, V.I.Bakulev, R.I.Kurziner and dr. M.: Mashinostroenie, 1987. – 568s. (In Russian)
2. Sheng Weidao, Tong Jungeng. Tekhnicheskaya termodinamika. M.: Vysshaya shkola, 2006 (In Russian)
3. G.N.Abramovich. Prikladnaya gazovaya dinamika. – M.: Nauka, 1991-600s. (In Russian)
4. V.V.Naschokin, Tekhnicheskaya termodinamika i teploperedacha. – M.: c.1975. – 496s. (In Russian)
5. Wikipedia-overall pressure ratio http://en.wikipedia.org/wiki/Overall_pressure_ratio
6. Wikipedia-diesel engine http://en.wikipedia.org/wiki/Diesel_engine
7. Li Haiwei, Shi Linsuo, Li Yaqi. Razvitie i primeneniye dvigatelej Stirlinga / Li Haj-Vehj, SHi Lin-So, Li YAj-CHi: Tekhnologiya Energii, 2012 (In Chinese)

Vladimir Ivanovich Panchenko, Prof. (Department of jet engines and energetic plants, Kazan National Research Technical University named after AN Tupolev (KNRTU-KAI), Russia). Specialist in the field of fluid dynamics.

Li Ziwan, master (Department of jet engines and energetic plants, Kazan National Research Technical University named after A.N. Tupolev (KNRTU-KAI), Russia). Research topic: the research of gas dynamics flame stabilization, scientific supervisor: assistant professor Varsegov V.L. (KNRTU-KAI)

О конвертировании авиационных двигателей, отработавших ресурс в лётной эксплуатации

П.Г. Великанов, Чжан Бин

Казанский национальный исследовательский технический университет
им. А.Н.Туполева (КНИТУ-КАИ),
Россия, 420111, Казань, К. Маркса 10

Аннотация: Выполнены сравнительный анализ прочности и циклической долговечности дисков турбин авиационных газотурбинных двигателей и наземных газотурбинных установок для газоперекачивающих агрегатов. Для расчета на прочность применены метод интегральных уравнений и метод конечных элементов (МКЭ) с использованием конечно-элементного комплекса «ANSYS». Расчет циклической долговечности выполнен с использованием эмпирической зависимости Мэнсона. Для сравнения повреждаемости дисков в лётной эксплуатации и на наземной установке использован принцип линейного суммирования повреждений. Установлено, что после конвертирования авиационных двигателей, отработавших ресурс в лётной эксплуатации, применение дисков турбин целесообразно продолжить на наземной газотурбинной установке.

Ключевые слова: диск, турбина, прочность, долговечность, конечный элемент, уравнение Мэнсона, суммирование повреждений

Введение

В настоящее время авиационные газотурбинные двигатели (ГТД), отработавшие свой летный ресурс, находят применение для привода газоперекачивающих агрегатов, электрогенераторов, газоструйных установок, устройств для очистки карьеров, снегоочистителей и т.д.

Массовое применение авиадвигателей, отработавших летный ресурс и сохранивших способность к дальнейшему использованию, позволяет в масштабах содружества независимых государств решить поставленную задачу, поскольку в условиях общего спада производства сохранение овеществленного в двигателях труда и экономия дорогостоящих материалов, используемых при их создании, позволяет не только затормозить дальнейший экономический спад, но и достичь роста экономики [1].

Основная часть

Объектами исследования в данной работе являются двигатель НК-8-2У, применяемый на самолете ТУ-154Б, и наземная газотурбинная установка (ГТУ) НК-16СТ, разработанная на базе двигателя НК-8-2У, использующаяся как привод для газоперекачивающего агрегата. В качестве предмета исследования принят диск турбины ротора низкого давления.

1. Обоснование выбора диска турбины для проведения расчетных оценок

Детали турбин авиационных ГТД испытывают в процессе эксплуатации воздействие различных видов нагружения.

Например, детали статора турбин (сопловые лопатки, корпуса, диафрагмы и т.п.) испытывают воздействие газовых нагрузок и температуры. В отличие от них роторные детали (рабочие лопатки, диски, валы) испытывают, кроме вышеперечисленных нагрузок, дополнительные нагрузки от воздействия центробежных сил.

Поэтому нагруженность роторных деталей обычно выше, чем у статорных.

С точки зрения тяжести последствий разрушения деталей турбины наиболее тяжелые последствия вызывают разрушения дисков, т.к. при разрушении дисков их фрагменты обладают значительной кинетической энергией и, как правило, не удерживаются корпусами турбины.

В связи с этим прочностная надежность дисков турбин во многом определяет и надежность работы двигателя, а иногда и лимитирует его ресурс.

Конструкция двигателя НК-16СТ разработана на основе конструкции авиационного двигателя НК-8-2У, имеющего ресурс 18000 часов в летной эксплуатации.

Оба двигателя двухкаскадные, т.е. у них есть турбины высокого и низкого давлений. Турбина двигателя НК-8-2У имеет три диска: один диск относится к турбине высокого давления и два диска – к турбине низкого давления.

В турбине двигателя НК-16СТ диск третьей ступени ротора низкого давления не имеет лопаточного венца и его нагруженность значительно ниже, чем у диска третьей ступени турбины двигателя НК-8-2У.

Поэтому для сравнительной оценки был выбран диск второй ступени ротора низкого давления. Этот диск, в сравнении с диском ротора высокого давления, имеет более высокую рабочую температуру и более высокий температурный перепад между ободом и ступицей.

2. Анализ условий нагружения дисков турбин в режиме летной эксплуатации и на наземной установке

В процессе эксплуатации диски подвергаются циклическому воздействию температурных и механических нагрузок. При этом в наиболее нагруженных сечениях дисков действующие напряжения могут превышать предел пропорциональности материала. В этом случае происходит пластическое деформирование материала с накоплением пластических деформаций от цикла к циклу. Механизм повреждения материала в этих условиях называется малоцикловой усталостью. Число циклов до разрушения, характерное для малоцикловой усталости, составляет, как правило, $10^3 \dots 10^5$ циклов.

Обобщенным полетным циклом нагружения авиационного двигателя называется цикл нагружения от запуска до останова (рис.1).

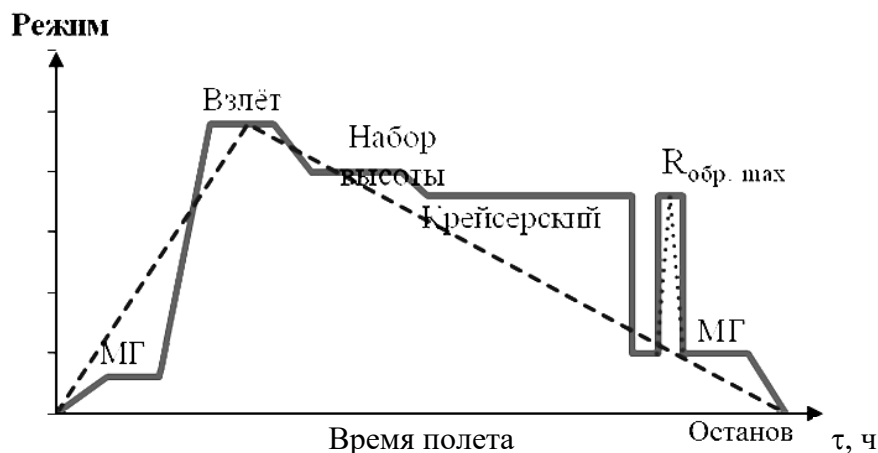


Рис. 1. Цикл нагружения авиационного двигателя от запуска до останова (МГ – малый газ на режиме взлета, взлетный (максимальный), набор высоты (максимальный продолжительный, номинальный), крейсерский, $R_{обр. max}$ – реверс обратной тяги, МГ – малый газ в режиме полета)

Основной повреждающий подцикл двигателя показан на рисунке 1 пунктирной линией, его обычно обозначают «0-Взлет-0». Расчет циклической долговечности дисков обычно выполняют именно для основного подцикла; остальные подциклы, например «Малый газ- $R_{обр. \max}$ -Малый газ», показанный на рисунке 1 точечной линией, вносят меньшее повреждение в диск и ими можно пренебречь. Подобные подциклы могут иметь место при заходе самолета на посадку. Однако следует отметить, что иногда, если самолет уходит на второй круг и совершает несколько подциклов при заходе на посадку, то и такие подциклы суммарно могут внести существенное повреждение, и при уточненных расчетах эти подциклы также принимают во внимание [2].

В отличие от авиационного двигателя цикл нагружения энергетических установок более простой. На рисунке 2 представлен цикл нагружения газотурбинного привода газоперекачивающего агрегата или электрогенератора.

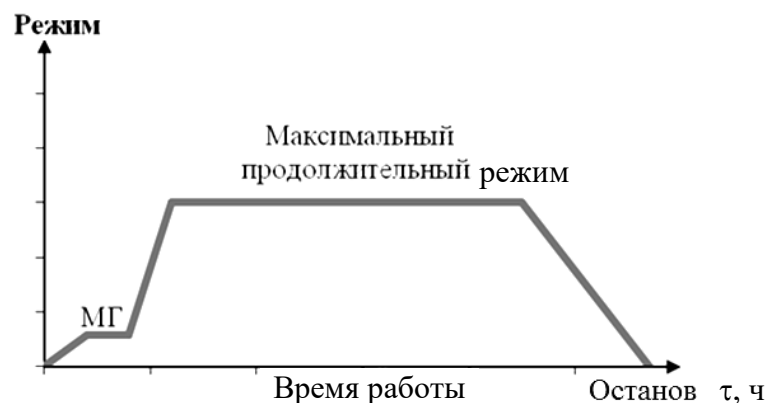


Рис. 2. Цикл нагружения газотурбинного привода газоперекачивающего агрегата

Основные отличия циклов нагружения авиационного двигателя и газотурбинного привода состоят в следующем:

1. Угловая скорость вращения роторов газотурбинных двигателей при работе двигателя на взлетном режиме выше, чем при работе газотурбинного привода на максимальном продолжительном режиме.
2. Время работы авиационного двигателя на взлетном режиме составляет от 1,5 до 5 мин, а время работы газотурбинного привода на максимальном продолжительном режиме может составить от 1000 до 1500 ч как время работы между профилактическими осмотрами.
3. Средняя продолжительность полетного цикла авиационного двигателя обычно 2-4 ч, а средняя продолжительность цикла нагружения газотурбинного привода, как было сказано ранее, 1000-1500 ч.

Таким образом, если авиационный двигатель имеет назначенный ресурс 20000 ч и в зависимости от средней продолжительности полета (цикла нагружения) назначенный ресурс в циклах от 10000 до 5000 циклов, газотурбинный привод, даже при характерных для него ресурсах 50000 ... 100000 ч, будет иметь число циклов нагружения не более 100.

Из этого следует, что для работы наземных газотурбинных установок (ГТУ) циклическое нагружение не является определяющим.

Другое дело, если речь идет об авиационном двигателе. Опыт эксплуатации авиационных ГТД показал, что для дисков этих двигателей достаточно часто имеет место именно малоцикловая усталость, лимитирующая ресурс диска, а иногда и двигателя в целом.

3. Оценка расчета дисков на прочность и циклическую долговечность

3.1. Расчет дисков на прочность

Напряжения в дисках возникают от наличия сил инерции самого диска и присоединенных к нему лопаток, от неравномерного нагрева диска по радиусу, от натяга, создаваемого прессовой посадкой диска на вал, если диск соединен с валом таким способом.

Значительное усложнение в расчет дисков вносит наличие высокой температуры и ее изменение вдоль радиуса дисков. В связи с этим появилась необходимость развития таких методов расчета, в которых учитывается по длине радиуса изменение: температуры, упругих постоянных и прочностных характеристик материала. Во многих случаях возникла необходимость определения напряжений с учетом пластических деформаций материала.

В данной работе применен достаточно эффективный инженерный метод расчета дисков, основанный на численном решении интегральных уравнений путем использования последовательных приближений разности радиальных и окружных (тангенциальных) напряжений.

Предметом исследования является диск второй ступени турбины низкого давления двигателя НК-8-2У и наземной ГТУ НК-16СТ, изготовленный из жаропрочного сплава ЭИ698-ВД. Процедура расчета диска реализована в программе «Расчет дисков с отверстием методом интегральных уравнений» [3]. Результаты показаны в рисунке. 3.

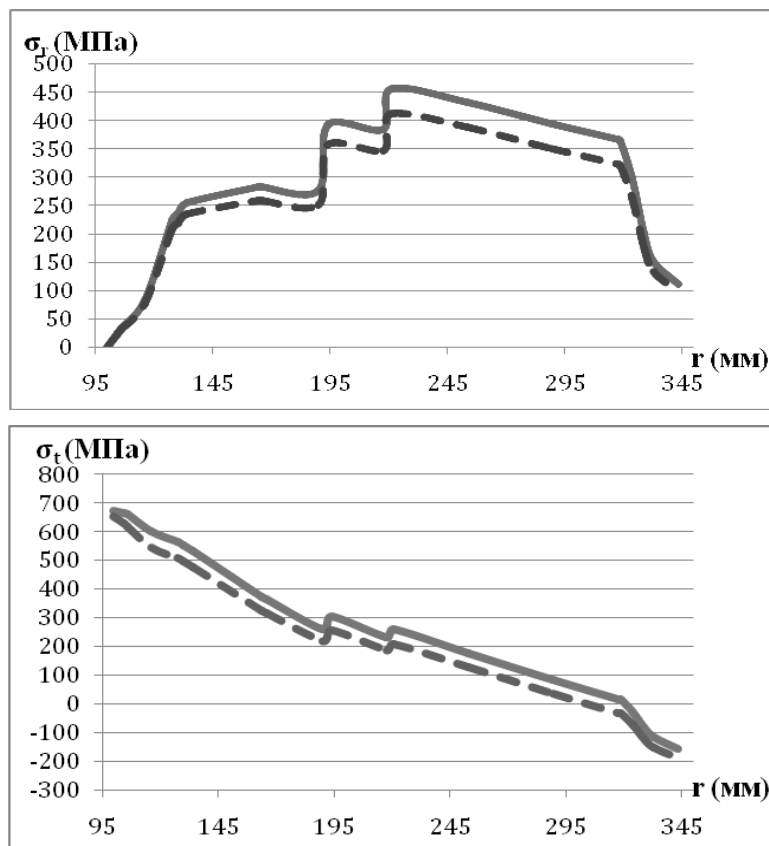


Рис. 3. Изменение радиальных напряжений σ_r и тангенциальных напряжений σ_t вдоль радиуса диска (— для двигателя НК-8-2У, - - - для наземной ГТУ НК-16СТ)

Объёмное напряженное состояние дисков, имеющих развитую ступицу, выполнено методом конечных элементов с использованием комплекса “ANSYS”. Результаты расчета приведены в таблице 1 (МЦУ – малоцикловая усталость, МКЭ – метод конечных элементов) [4].

Таблица 1

Результаты расчета

Тип двигателя	R (мм)	МЦУ (МПа) Номин. напряжение	МКЭ (МПа) Макс. напряжение
НК-8-2У	155	440	730
НК-16СТ	155	376	624

3.2. Определение циклической долговечности диска турбины

Увеличение ресурса авиационных ГТД, работающих в условиях циклического нагружения, опасность последствий разрушения деталей ротора, вращающихся с высокими скоростями, особенно дисков, требуют совершенствования расчетных методов прогнозирования долговечности.

В результате действия больших нестационарных нагрузок в дисках в процессе эксплуатации появляются необратимые деформации пластичности и ползучести и происходит накопление повреждений, которые могут привести к разрушению конструкции. Разрушение при этом происходит при количестве циклов нагружения в интервале $N=10^3 \dots 10^5$ циклов.

Разрушение деталей из-за МЦУ (малоцикловой усталости) начинается чаще всего в местах концентрации напряжений.

Для определения циклической долговечности необходимо провести исследование напряженно-деформированного состояния в зонах выступов замковых соединений, отверстий под болты и под охлаждающий воздух, лабиринтных выступов и в местах с резкими изменениями конфигурации деталей.

Долговечность N_p до образования трещины МЦУ можно найти как по экспериментальным кривым малоцикловой усталости, так и используя эмпирическую зависимость, предложенную Мэнсоном [5]:

$$\Delta \varepsilon = 3.5 \frac{\sigma_b - \sigma_m}{E} N_p^{-0.12} + \left(\ln \frac{1}{1 - \psi} \right)^{0.6} N_p^{-0.6}$$

где $\Delta \varepsilon$ – изменение полных деформаций в цикле; σ_b – предел прочности материала;

$\sigma_m = \frac{\sigma_{\max} + \sigma_{\min}}{2}$ – среднее напряжение в цикле «0-Взлет-0»; где σ_{\min} соответствует

режиму – «0», σ_{\max} соответствует режиму – «Взлет». E – модуль упругости Юнга материала; ψ – коэффициент относительного сужения материала; N_p – циклическая долговечность до момента образования трещины МЦУ.

Для использования зависимости необходимо определить характеристики цикла упругопластического деформирования в расчетной точке детали.

В таблице 2 приведены результаты расчета циклической долговечности дисков двигателя НК-8-2У и НК-16СТ по программе «Определение циклической долговечности» [3] (здесь T – температура (°C), $1/N_p$ – повреждаемость (1/цикл), остальные обозначения приведены выше в тексте).

Таблица 2

Результаты расчета

Номер подцикла	«Вес» подцикла	Упругие параметры подцикла				Пластические параметры подцикла				Прогнозируемые параметры	
		σ_{\max}	T	σ_{\min}	T	σ_{\max}	σ_{\min}	σ_m	$\Delta\varepsilon$	N_p	Повреждаемость $1/N_p$
		кг/мм ²	°C	кг/мм ²	°C	кг/мм ²	кг/мм ²	кг/мм ²	отн. ед.	цикл	1/цикл
1	1.00	98.00	298.00	0.00	20.00	71.90	-27.5	22.20	0.00439	0.137E+06	7.30E-06
2	1.00	82.82	298.00	0.00	20.00	69.70	-14.3	27.70	0.00371	0.409E+06	2.45E-06

Циклическая долговечность диска наземной установки НК-16СТ в 3 раза больше, чем диска двигателя НК-8-2У.

4. Принцип линейного суммирования повреждений

В реальной эксплуатации авиационные ГТД работают при сложном изменении режимов нагружения в процессе эксплуатационного цикла (см. рис. 4). Для деталей двигателя каждый режим работы характеризуется своим уровнем температур, статических и вибрационных нагрузок. Каждый режим вносит свою долю в истощение ресурса, поэтому оценка ресурса должна вестись с учетом наработки на всех режимах работы двигателя.

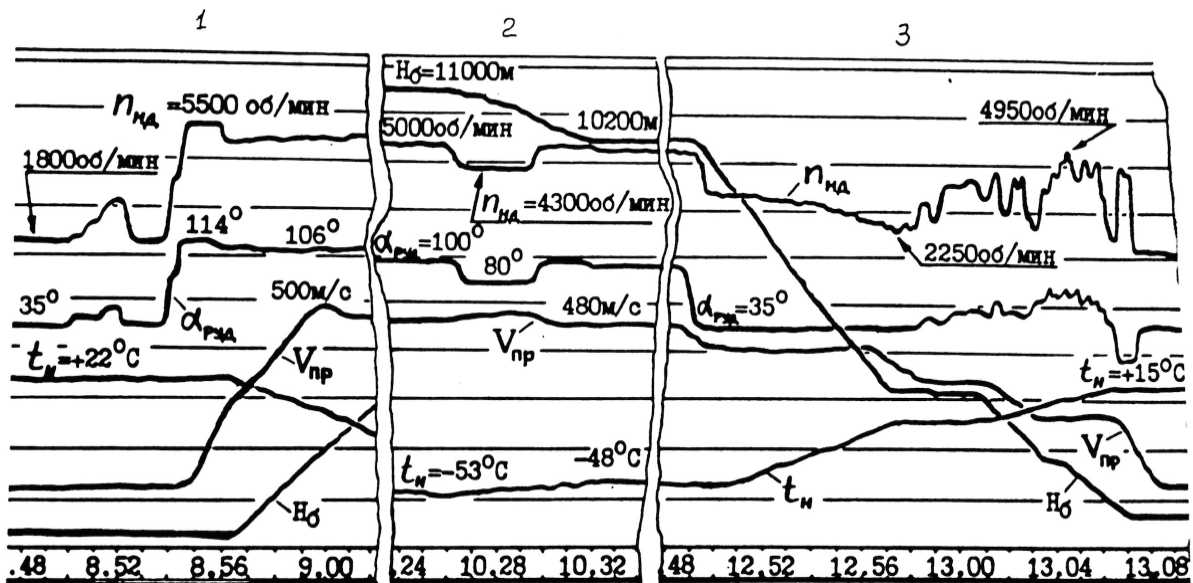


Рис.4. Типичное изменение режимов работы двигателя НК-86 в процессе полета: 1 – руление на исполнительный старт; взлет; набор высоты, 2 – крейсерский по маршруту, 3 – снижение, заход на посадку, посадка, включение реверса [6]

Методы, позволяющие вести оценку ресурса деталей, работающих на различных режимах нагружения, имеют в своей основе различные гипотезы, применяемые для суммирования повреждений. Под повреждением здесь подразумевается доля истощения ресурса, вносимая отдельным режимом нагружения.

Впервые метод расчета долговечности при изменяющихся режимах нагружения (программном нагружении) был разработан Пальмгреном применительно к подшипникам качения в 1924 году, затем независимо от него и друг от друга аналогичный метод был предложен применительно к высокочастотной усталости в работах А. Майнера в 1945 году и С.В. Серенсена в 1944 году. В настоящее время метод, разработанный Пальмгреном, С.В. Серенсеном и А. Майнером, носит название гипотезы линейного суммирования повреждений [6].

В инженерной практике наибольшее распространение имеет линейная гипотеза суммирования повреждений, в соответствии с которой условие разрушения при программном статическом нагружении представляется в виде:

$$P_{\tau} = \sum_{i=1}^n \frac{\tau_i}{\tau_{pi}} = 1$$

где P_{τ} – суммарное повреждение от длительного статического нагружения; i – номер режима нагружения; n – количество режимов нагружения; τ_i – суммарное время работы на i режиме; τ_{pi} – время до разрушения при непрерывной работе на i -ом режиме.

При программном усталостном нагружении:

$$P_N = \sum_{i=1}^n \frac{N_i}{N_{pi}} = 1,$$

где P_N – суммарное повреждение от малоциклового нагружения; N_i – число циклов нагружения на i -ом режиме; N_{pi} – число циклов нагружения до разрушения при непрерывной работе на i -ом режиме.

Величины τ_{pi} и N_{pi} определяются по кривым длительной или усталостной прочности при соответствующих условиях нагружения [7].

Линейная гипотеза «суммирования» повреждений не учитывает взаимного влияния переходов с одного уровня нагрузок на другой и чередования нагрузок на долговечность, и поэтому она может вносить определенную погрешность в расчеты. Для учета взаимного влияния различных уровней нагрузок в рамках линейной гипотезы «суммирования» повреждений предлагаются различные методы корректировки величины предельного накопленного повреждения при разрушении, например, в случае усталостного нагружения. Однако, учитывая большой разброс характеристик длительной прочности, малоциклового долговечности и сопротивления усталости, точность расчетов с использованием линейной гипотезы суммирования повреждений вполне достаточна для инженерных расчетов, тем более, что работоспособность конструкций оценивается с учетом нормируемых коэффициентов запасов.

При статическом нагружении «суммарная» повреждаемость диска двигателя НК-8-2У равна 0,025, суммарная повреждаемость диска наземной ГТУ НК-16СТ равна 0,5.

При усталостном нагружении «суммарная» повреждаемость диска двигателя НК-8-2У равна 0,065, суммарная повреждаемость диска наземной ГТУ НК-16СТ равна 0,000244.

Таким образом, суммарная повреждаемость диска при статическом нагружении после летной эксплуатации и при работе на наземной ГТУ будет не больше $P_{\tau} = 0.525$, а при циклическом нагружении $P_N = 0.006524$, что значительно меньше единицы.

Заключение

Установлено, что после конвертирования авиационных двигателей, отработавших ресурс в лётной эксплуатации, применение дисков турбин целесообразно продолжить на наземной ГТУ.

Список сокращений

ANSYS – универсальная программная система конечно-элементного анализа; ГТД – газотурбинный двигатель; ГТУ – газотурбинная установка; МКЭ – метод конечных элементов; МЦУ – малоцикловая усталость.

Список литературы

1. Е.А. Гриценко. Конвертирование авиационных ГТД в газотурбинные установки наземного применения / Е.А. Гриценко, Б.Л. Данильченко, С.В. Лукачев и др. - Самара: СН РАН, 2004. - 266 с.
2. И.В. Демьянушко. Прочность и долговечность дисков авиационных двигателей и энергетических установок: учебное пособие для студентов вузов / И.В. Демьянушко, Н.П. Великанова. – М., Казань: МАДИ, КНИТУ-КАИ, 2008. – 144 с.
3. Н.П. Великанова. Расчёт на прочность основных деталей газотурбинных двигателей: учебное пособие для студентов вузов / Н.П. Великанова, Ф.К. Закиев, П.Г. Великанов. - Казань: КНИТУ-КАИ, 2011. – 72 с.
4. Н.П. Великанова. Прогнозирование долговечности дисков турбин авиационных ГТД на основе вероятностного подхода: дис. ... канд. техн. наук. Москва, 250 с. 1994.
5. С.С. Мэнсон. Температурные напряжения и малоцикловая усталость / С.С. Мэнсон. - М.: Машиностроение, 1975. – 344 с.
6. Н.Д. Кузнецов. Установление и увеличение ресурсов авиационных ГТД: Учебное пособие / Н.Д. Кузнецов, Е.А. Гриценко, А.А. Корноухов и др. – Самара: СГАУ, 1998. – 207 с.
7. Е.Н. Каблов. Жаропрочность никелевых сплавов / Е.Н. Каблов, Е.Р. Голубовский. – М.: Машиностроение, 1998. – 464 с.

Петр Геннадиевич Великанов, кандидат физ.-мат. наук, доцент (КНИТУ-КАИ, кафедра реактивных двигателей и энергетических установок, РДЭУ). Область научных интересов: исследование напряженно-деформированного состояния деталей методами конечных и граничных элементов, расчет на прочность основных деталей газотурбинных двигателей.

Чжан Бин, магистрант (КНИТУ-КАИ, кафедра реактивных двигателей и энергетических установок, РДЭУ). Область научных интересов: расчет на прочность основных деталей газотурбинных двигателей.

About conversion of aircraft engines after life cycle in the flight operation

P.G. Velikanov, Zhang Bing

Kazan National Research Technical University named after A.N. Tupolev
K. Marx, 10, Kazan, 420111, Russia

Abstract. In the given material the strength and cycle durability of gas turbine disk, comparing of the difference between aircraft and ground-based gas turbines is analyzed. The analysis is performed by integral equation method and the finite element method in the ANSYS software. The calculation of cycle durability is performed by the empirical Manson equation. To compare the damage of disk in flight operations and ground-based gas turbines, the principle linear summation of damage was used. It is founded that after the conversion of aircraft engines, exhaust the resource in flight operations, the use of turbine disks is worth continuing on the ground-based gas turbines.

Key words: disk, turbine, strength, durability, finite element, Manson equation, summation of damage, ground-based gas turbines

Introduction

Currently, aircraft gas turbine engines, exhaust their flight hours, are used to drive the gas compressor units, generators, gas-jet systems, devices for cleaning pits, snow blowers, etc. The massive use of aircraft engines, exhaust flight hours and retain the ability to continue to use, allows the scale of the Commonwealth of Independent States to solve the problem, since a general decline in production preservation embodied in engines work and saving of expensive materials used in their creation, can not only slow down further the economic downturn, but also to achieve economic growth [1].

The objects of study in this paper are the engine NK-8-2U, operated on Tu-154B, and ground gas turbine NK-16ST, developed on the basis of the engine NK-8-2U used as a drive for gas compressor units. As a subject of study adopted a turbine disk of low pressure rotor.

1. Justification of the choice of the turbine disk for estimates

Details turbines GTE experience in the operation of various types of impact loading.

For example, a turbine stator parts (nozzle vanes, casings, diaphragm, etc.) are exposed to gas temperatures and loads. In contrast, the rotor parts (rotor blades, disks, shafts) test, except the above loads, the load on the still and the centrifugal forces.

Therefore, the loading of the rotor parts are generally higher than that of the stator.

From the point of view of the severity of the consequences of the destruction of parts of the turbine, the most serious consequences caused by the breakdown of disks, because disk fracture fragments have considerable kinetic energy, and typically are not confined turbine housing.

In this regard, the strength reliability of turbine disks largely determines the reliability of the engine, and sometimes limits its lifetime.

The design of the engine NK-16ST is based on the design of the aircraft engine NK-8-2U, available resources 18000 hours of flight operation.

Both engines are two-stage, they have the turbine high and low pressure. Turbine engine NK-8-2U has three disks - one disk refers to the high pressure turbine and two drives to the low pressure turbine.

In the turbine engine NK-16ST disk stage 3th low-pressure rotor does not have blade row and its loading is significantly lower than that of the disk 3th stage turbine engine NK-8-2U. Therefore, for the comparative evaluation was chosen second stage rotor disk low pressure. This disk, in comparison with the high-pressure rotor disk has a higher operating temperature and higher temperature difference between the rim and the hub.

2. The analysis of the loading drives turbines conditions in flight operations and ground installation

During operation, disks cycle the thermal and mechanical loads. In the most loaded sections of disk operating stress may exceed the proportional limit of the material. In this case, plastic deformation occurs with the accumulation of plastic material deformations from cycle to cycle. The mechanism of damage to the material under these conditions is called a low-cycle fatigue. The number of cycles to failure, typical for low-cycle fatigue, usually 103... 105 cycles.

Generalized full cycle of loading an aircraft engine is called the loading cycle from startup to shutdown (Figure 1).

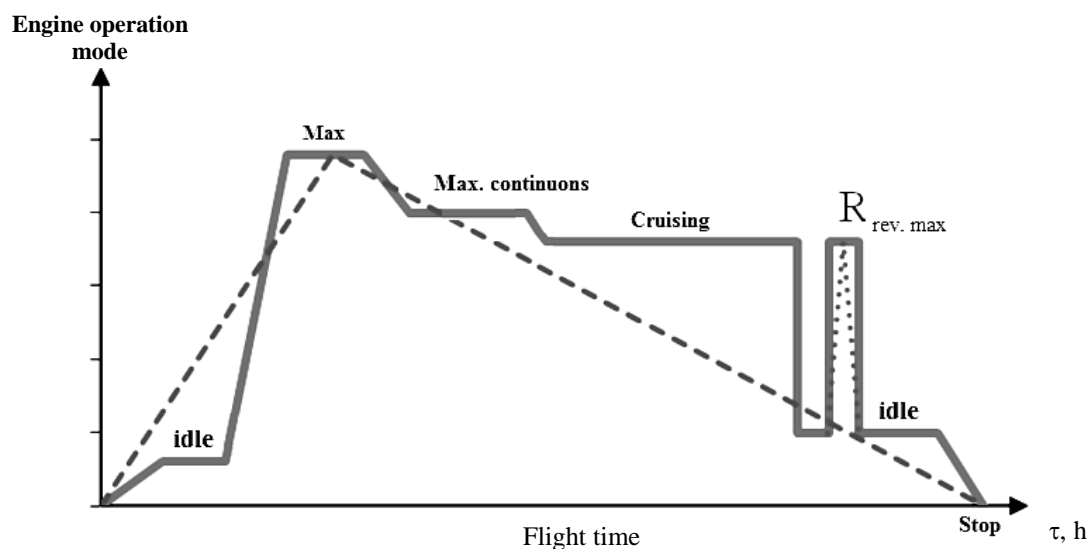


Figure 1. Loading cycles of aircraft engine from startup to shutdown (runway idle, max, max continuous, cruising, $R_{rev.max}$ – reverse thrust, flight small gas)

The main damaging the engine subframe is shown in Figure 1 by the dashed line, it is usually denoted by "0-Rise-0". Calculation of the cyclic durability of disks is usually performed for the main subcycle; the remaining sub-cycles, such as "Idle-Rrev.max-Idle", as shown in Figure 1 by the dotted line, make less damage to the disk and can be ignored. Such sub-cycles can occur when entering the aircraft to land. But it is notable that sometimes if the plane goes to the second round and makes several sub-cycles during landing, and such total sub-cycles can make significant damage, and refined these calculations sub-cycles also take into account [2]. Unlike aircraft engine loading cycle power plants easier. Figure 2 shows the loading cycle gas turbine drive gas compressor unit or electric.

The main differences of loading cycles of aircraft engines and gas turbine drive are as follows:

1. Rotor speed turbine engines with the engine at takeoff is higher than when using a gas turbine drive at maximum continuous operation.

2. Hours aircraft engine at takeoff is from 1.5 to 5 minutes, and while the gas turbine drive at maximum continuous operation could range from 1000 to 1500 hours as the time between routine inspections.

3. The average duration of the flight of the aircraft engine cycle is usually 2-4 hours, and the average duration of the loading cycle gas turbine drive, as mentioned earlier, 1000-1500 h.

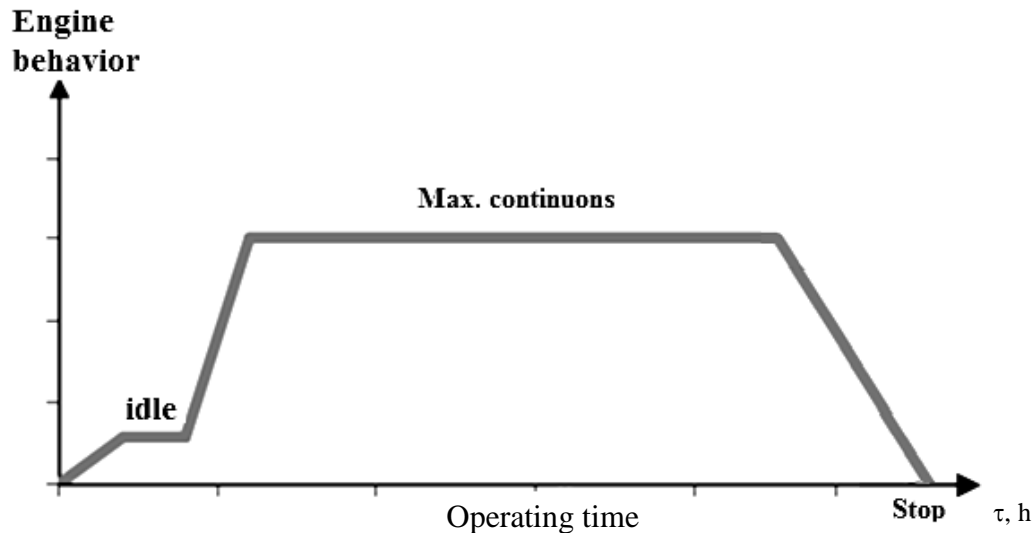


Figure 2. Loading cycle of gas turbine drive of gas pumping unit

Thus, if the aircraft engine has an assigned resource, and 20,000 hours depending on the average length of the flight (load cycle) assigned resource in cycles from 5000 to 10000 cycles turbine drive characteristic even when a resource 50000 ... 100000 h will have a number of cycles load no more than 100.

From this it follows that for land gas turbines cyclic loading is not determinative.

Another thing, when it comes to aircraft engines. Operating experience GTE showed that the disk these engines often took place exactly low cycle fatigue, limiting disk resource, and sometimes the engine as a whole.

3. The disks analysis evaluation on strength and cycle life

3.1. Calculation of disks on strength

Stresses arise in disks of the inertia forces of the disk and blades connected thereto, the uneven heating of the disk along a radius of the interference produced by press fitting a disk on the shaft, when connected to the disk shaft in such a way.

A significant complication in the calculation of disk makes the presence of high temperature and its variation along the radius of the disk. In this regard, there was a need for the development of such methods of calculation, which takes into account the change in the radial temperature, the elastic constants and the strength characteristics of the material. In many cases, it became necessary to determine the stresses based on the plastic deformation of the material.

In this paper we used a very efficient engineering method for calculating the drive based on the numerical solution of integral equations by using successive approximations of the difference of radial and circumferential (tangential) stress.

The subject of the study is to disk the second stage low-pressure turbine engine NK-8-2U and ground gas turbine NK-16ST, made of heat-resistant alloy EI698-VD. The procedure for calculating the disk is implemented in the "Calculation of disks with a hole method of integral equations" [3]. The results are shown in Figure. 3.

Three-dimensional stress state disks having developed the hub, the finite element method using a complex "ANSYS". The results of the calculation are given in Table 1 [4].

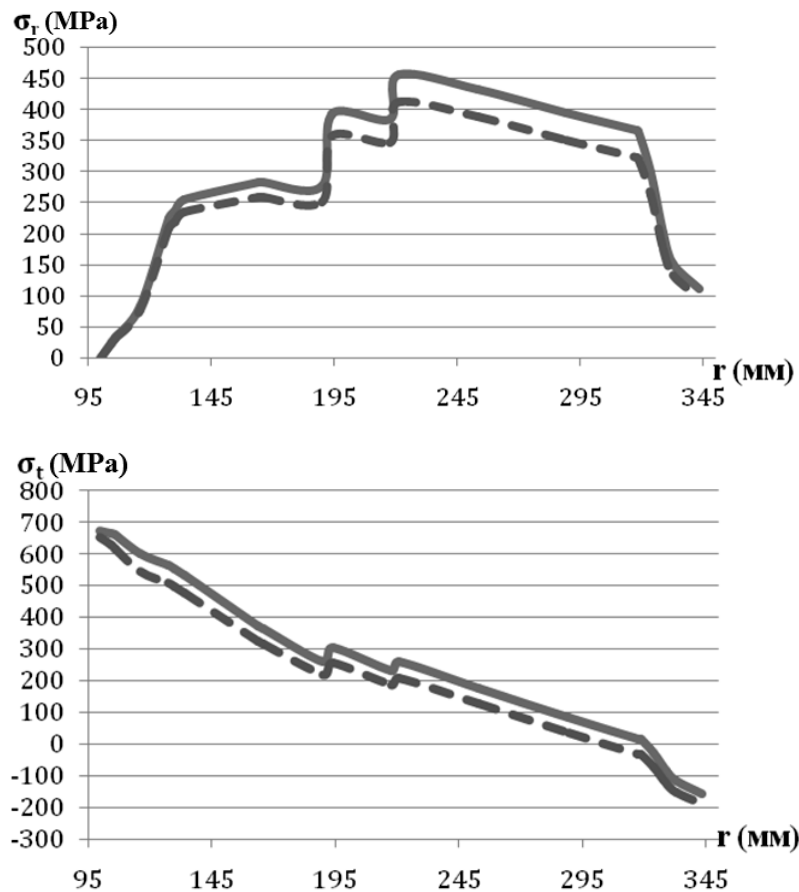


Fig. 3. Change the radial stresses σ_r and hoop stress σ_t along the radius of the disk (—engine NK-8-2U- - - ground gas turbine NK-16 ST)

Table 1

The results of calculation

Type of engine	R (mm)	LCF (MPa) Nominal stress	FEM (MPa) Maximal stress
NK-8-2U	155	440	730
NK-16ST	155	376	624

3.2. Determination of cyclic durability of the turbine disk

Increase resource GTE working under cyclic loading, the risk of consequences of destruction of parts of the rotor rotates at a high peripheral speeds, especially drives, need to be improved computational methods predict longevity.

As a result of the high transient loads during operation drives appear irreversible deformation and ductility and creep damage accumulation occurs, which can lead to structural failure.

Destruction at the same time when the number of cycles in the range $N=10^3 \dots 10^5$ cycles. Such destruction is called low-cycle fatigue (LCF).

The destruction of parts due to LCF often begins in stress concentration field.

To determine the cyclic durability necessary to undertake a study of the stress-strain state in the areas of projections interlocks, bolt holes and a cooling air labyrinth projections and in places with sharp changes in the configuration details.

Durability N_p to LCF crack formation can be found both on the experimental curves of low cycle fatigue, and using the empirical relationship proposed by Manson[5]:

$$\Delta \varepsilon = 3.5 \frac{\sigma_b - \sigma_m}{E} N_p^{-0.12} + \left(\ln \frac{1}{1 - \psi} \right)^{0.6} N_p^{-0.6}$$

where $\Delta \varepsilon$ – full scope of strain in the cycle; σ_b – tensile strength of the material;

$\sigma_m = \frac{\sigma_{\max} + \sigma_{\min}}{2}$ – the average stress in the cycle "0-Rise-0"; where σ_{\min} corresponds to – "0"; σ_{\max} – corresponding to "Rise". E – modulus of elasticity; ψ – relative narrowing of material; N_p – cycle durability.

For use according to the characteristics necessary to determine the cycle elastic-plastic deformation at the design point details.

Table 2 shows the results of the calculation of the cyclic durability of the engine disks NK8-2U and NK-16ST program "Defining cyclic life" [3].

Table 2

The results of calculation

Number subcycle	Weight subcycle	Elastic parameters subcycle				Plastic parameters subcycle				Projected parameters	
		σ_{\max}	T	σ_{\min}	T	σ_{\max}	σ_{\min}	σ_m	$\Delta \varepsilon$	N_p	Defectiveness $1 / N_p$
		kg/mm ²	°C	kg/mm ²	°C	kg/mm ²	kg/mm ²	kg/mm ²	rel. units.	cycle	1 / cycle
1	1.00	98.00	298.00	0.00	20.00	71.90	-27.5	22.20	0.00439	0.137E+06	7.30E-06
2	1.00	82.82	298.00	0.00	20.00	69.70	-14.3	27.70	0.00371	0.409E+06	2.45E-06

Cyclic durability ground installation disk NK-16 ST 3 times larger than the drive motor NK-8-2U.

4. The principle of linear summation of damage

In real operation the GTE work under complex loading conditions change during the life cycle. See Figure 4. Engine parts for each mode of operation is characterized by its level of temperature, static and vibration loads. Each mode brings its share of the exhaustion of the resource, so the resource assessment should be carried out taking into account the developments in all modes of operation of the engine.

Methods to conduct an assessment of the resource components operating at different loading conditions, based on various hypotheses damage summation. Under the damage here means the proportion of resource exhaustion, introduced a separate mode of loading.

For the first time the method of calculation of durability under varying loading conditions (loading software) was developed Palmgren applied to rolling bearings in 1924, and then independently of him and each other similar method was proposed in relation to the high-frequency fatigue in the works A. Mayner in 1945 and S.V. Serensen in 1944. At present, the

method developed Palmgren, S.V. Serensen and A. Mayner is called the hypothesis of linear damage summation [6].

In engineering practice, the most widespread hypothesis of linear damage summation, according to which the condition of failure in program static loading can be written as:

$$\Pi_{\tau} = \sum_{i=1}^n \frac{\tau_i}{\tau_{pi}} = 1$$

where Π_{τ} – the total damage from prolonged static loading; i — number loading mode; n — the number of load cases; τ_i — total operating time on the i mode; τ_{pi} — time to failure in continuous operation on the i -th mode.

A soft fatigue loading:

$$\Pi_N = \sum_{i=1}^n \frac{N_i}{N_{pi}} = 1,$$

where Π_N – the total damage from low-cycle loading; N_i – number of cycles of loading at i -th mode; N_{pi} – the number of load cycles to failure in continuous operation in the j -th mode.

The magnitude τ_{pi} и N_{pi} of the curve or prolonged fatigue strength under appropriate conditions of loading [7].

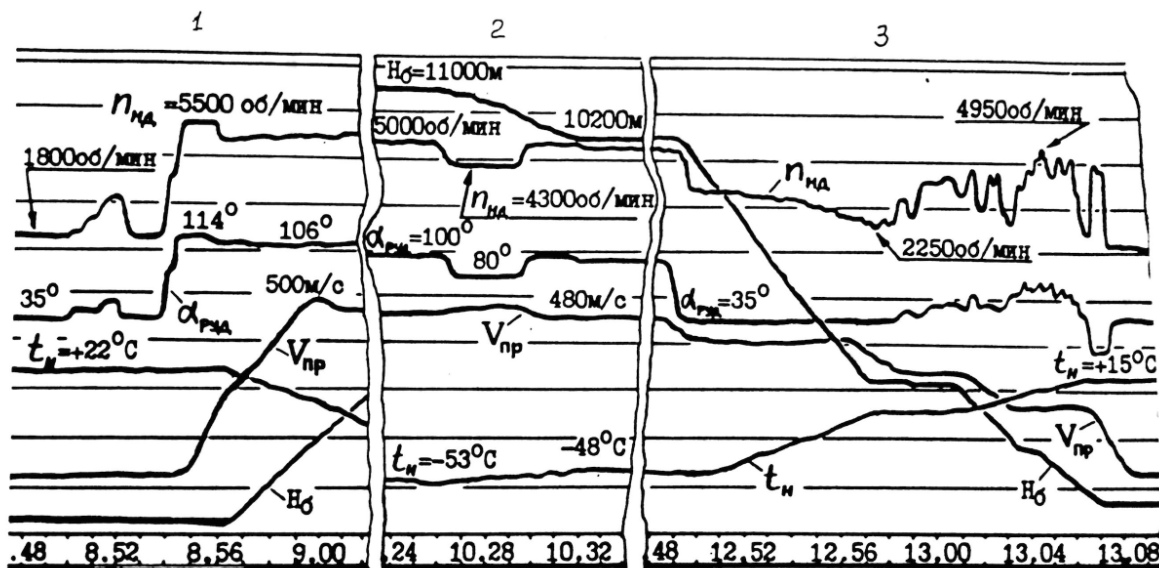


Figure 4. Typical change of modes of engine NK-86 in flight: 1 – lining up; takeoff; climb, 2 – cruising, 3 – descent, landing approach, landing, selection of reverse thrust [6]

The hypothesis of damage linear summation does not account for the mutual influence of the transition from one level to another and loads of alternating loads on durability, and therefore it can make the certain error in calculations. To take into account the mutual influence of different levels of stress within the linear hypothesis of damage summation various methods of adjusting the limit of accumulated damage at failure, for example, in the case of fatigue loading. However with given the large scatter characteristics of long-term strength, durability and low cycle fatigue resistance, this accuracy of calculations using linear damage summation hypothesis is sufficient for engineering calculations, moreover. The efficiency of structures is assessed in the light of standardized coefficients of stocks.

Under static loading total defect drive motors NK-8-2U equal to 0.025, the total disk defect ground GTU NK-16ST is 0.5.

When fatigue loading total defect drive motors NK-8-2U equal to 0.065, the total disk defect ground GTU NK-16ST is 0.000244.

Thus the total disk defect (under static loading and after the flight manual when working on the ground GTU) will be no more than $\Pi_{\tau} = 0.525$. And under cyclic loading we have $\Pi_N = 0.006524$, which is much less than unity.

Conclusions

It is found that after the conversion of aircraft engines, exhaust the resource in flight operations, the use of turbine disks is possible to continue on the ground-based gas turbines.

Abbreviations

ANSYS – universal software system for finite element analysis; GTE – gas turbine engine; GTU – gas turbine unit; FEM – finite element method; LCF – low cycle fatigue.

References

1. E.A. Gritsenko, B.L. Danilchenko, S.V. Lukachev, ets. Konvertirovanie aviacionnyh GTD v gazoturbinnye ustanovky nazemnogo primeneniya. (in Russian) [Conversion of aircraft GTE into gas turbine units for terrestrial applications] Samara: *SN RAN*, 2004. p. 266.
2. I.V. Demijanushko, N.P. Velikanova. Prochnost i dolgovechnost diskov aviacionnyh dvigateley i energeticheskikh ustanovok: uchebnoe posobie dlja studentov vuzov. (in Russian) [The strength and durability of disks of aircraft engines and gas turbine engines] Moscow-Kazan: *MARU-KNITU-KAI*, 2008. p. 144.
3. N.P. Velikanova, F.K. Zakiev, P.G. Velikanov. Raschet na prochnost osnovnyh detaley gazoturbinnnyh dvigateley: uchebnoe posobie dlja studentov vuzov. (in Russian) [Calculate the strength of main parts of gas turbine engines] Kazan: *KNITU-KAI*, 2011. p. 72.
4. N.P. Velikanova. Prognozirovanie dolgovechnosti diskov turbin aviacionnyh GTU na osnove veroyatnostnogo podhoda: dissertaciya kandidata technicheskikh nauk. (in Russian) [Forecasting of durability of gas turbine discs based on probabilistic approach: dis. ... candidate. tech. sciences]. Moscow, 1994. p. 250.
5. S.S. Manson Temperaturnye napriyazheniya i malociklovaya ustalost. (in Russian) [Temperature stress and low-cycle fatigue] Moscow: *Mashinostroenie*, 1975. p. 344
6. N.D. Kuznetsov, E.A. Gritsenko, A.A. Kornoukhov ets. Ustanovlenie i uvelichenie resursov aviacionnyh GTD: uchebnoe posobie. (in Russian) [Establishment and resources of aircraft gas-turbine engines] Samara: *SGAU*, 1998. p. 207.
7. E.N. Kablov, E.R. Golubovsky. Zharoprochnost nikelovyh splavov. (in Russian) [The heat-resistant Nickel alloys] Moscow: Mechanical Engineering, 1998. p. 464.

Petr Gennad'evich Velikanov, Candidate of Physics and Mathematical Sciences, associate professor (KNRTU-KAI, Department of JEPS). The research interests: determination of stress-strain state of the parts finite and boundary element methods, strength calculation of main parts of gas turbine engines..

Zhang Bing, master, (KNRTU-KAI, Department of JEPS). The research interests: strength calculation of main parts of gas turbine engines.

К проблеме жаропрочности рабочих лопаток турбины авиационных ГТД (статистический анализ)

Н.П. Великанова, Фань Вэйвэй

Казанский национальный исследовательский технический университет
им. А.Н.Туполева (КНИТУ-КАИ),
Россия, 420111, Казань, К. Маркса, 10

Аннотация. В работе рассматривается влияние длительной эксплуатационной наработки газотурбинных двигателей (ГТД) для гражданской авиации на долговечность (жаропрочность) материала рабочих лопаток турбин (материал – литейный жаропрочный сплав ЖС6У-ВИ) на никелевой основе равноосной кристаллизации. Исследован конкретный интервал наработки, выполнен статистический анализ результатов испытаний эмпирических распределений для величины жаропрочности материала; установлен закон эмпирических распределений в каждом интервале наработки; построены графики эмпирических распределений с оценкой влияние наработки на величину долговечности.

Ключевые слова: прочность, долговечность, лопатка, турбина, жаропрочный сплав, статистический метод

Введение

Предметом исследования являются рабочие лопатки (РЛ) первой ступени турбины высокого давления двигателя НК-86, изготовленные из литейного жаропрочного сплава ЖС6У-ВИ с равноосной структурой материала. Этот сплав в настоящее время широко используются для изготовления рабочих лопаток турбин различных двигателей для самолетов гражданской авиации.

Анализ исходного уровня долговечности в часах и анализ уровней после эксплуатационной наработки материала рабочих лопаток включает в себя установление закона распределения характеристик долговечности и проверки принадлежности выборок к одной общей генеральной совокупности. Цель анализа – сравнение исходного уровня долговечности сплава ЖС6У-ВИ и уровня после наработки в эксплуатации.

Основная часть

Исходный уровень долговечности оценивался по результатам испытаний образцов, вырезанных из холодной зоны лопаток, в которой структура металла не претерпевает изменений. Уровень долговечности сплавов ЖС6У-ВИ после наработки в эксплуатации оценивался по результатам испытаний образцов, вырезанных из горячей зоны пера лопатки, где особенно сказывается влияние наработки.

Испытания образцов проводились в соответствии с ГОСТ 10145-81 «Металлы. Метод испытания на длительную прочность». В качестве базового уровня использовалось значение температуры $T = 1248$ К, близкой к расчетной, которая имеет место в наиболее опасном сечении. В качестве уровня действующих напряжений использовался уровень в 23 кг/мм^2 для сплава ЖС6У-ВИ. Испытание проводилось до разрушения образца.

В результате испытаний получились наборы экспериментальных данных по времени до разрушения τ_p , которые были подвергнуты статистической обработке согласно рекомендациям [1]. Статистический анализ включал в себя установление закона распределения и проверку принадлежности выборок к одной генеральной совокупности.

Результаты испытаний долговечности материала рабочих лопаток (ЖС6У-ВИ) были подвергнуты статистической обработке с целью определения влияния наработки на закон и числовые характеристики распределения долговечности. Для проверки вида закона распределения долговечности после наработки использовался графический метод и критерий Шапиро–Уилка [2], предназначенный для проверки гипотезы о нормальном или логарифмически нормальном распределении при ограниченном объеме выборки ($n < 50$); он является более мощным, чем другие критерии [1].

Для вычисления статистики критерия $W_{\text{раас}} = b^2 / S^2$ определяют критерий проверки однородности дисперсий S^2

$$S^2 = \sum_{i=1}^n (x_i - \bar{x}_i)^2, \quad \text{где } \bar{x}_i \text{ — выборочное среднее, и}$$

$$b = \sum_{i=1}^k a_{n-i+1} (x_{n-i+1} - x_i).$$

Гипотезу проверяют на основании неравенства $W_{\text{раас}} \geq W_\alpha$, где W_α — критическое значение критерия Шапиро–Уилка. Если выполняется неравенство, то нулевую гипотезу не бракуют.

Построенные графики достаточно хорошо аппроксимируются прямыми линиями, что свидетельствует о том, что закон распределения долговечности в часах τ_p может являться логарифмически нормальным, τ_p — характеристика долговечности (жаропрочности) материала образцов из рабочих лопаток турбин.

Проверка соответствия опытных данных логарифмически нормальному закону распределения по критерию Шапиро–Уилка показала, что закон распределения рассматриваемых характеристик после наработки является логарифмически нормальным при уровне значимости $\alpha = 0,05$. Согласно [1], уровень значимости α — это вероятность совершить ошибку первого рода, которая заключается в отбрасывании нулевой гипотезы в то время, когда на самом деле она верна.

В качестве иллюстрации результатов проведенного статистического анализа влияния наработки на долговечность сплава ЖС6У-ВИ на рисунках 1...4 показаны графики эмпирических распределений величины $\lg \tau_p$, в таблице 1 даны результаты проверки нормальности распределений величины $\lg \tau_p$ по критерию Шапиро–Уилка.

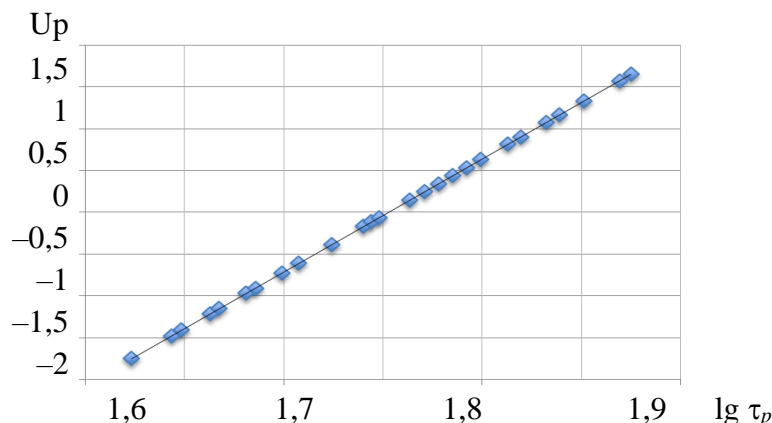


Рис. 1. График эмпирического распределения десятичного логарифма исходной долговечности τ_p сплава ЖС6У-ВИ; здесь U_p — функция распределения, безразмерная величина

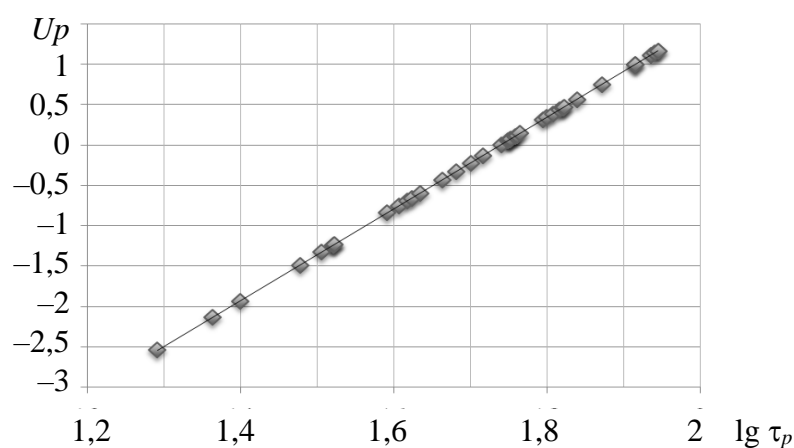


Рис. 2. График эмпирического распределения десятичного логарифма долговечности τ_p сплава ЖС6У-ВИ после эксплуатационной наработки в пределах от 2400 до 5000 часов

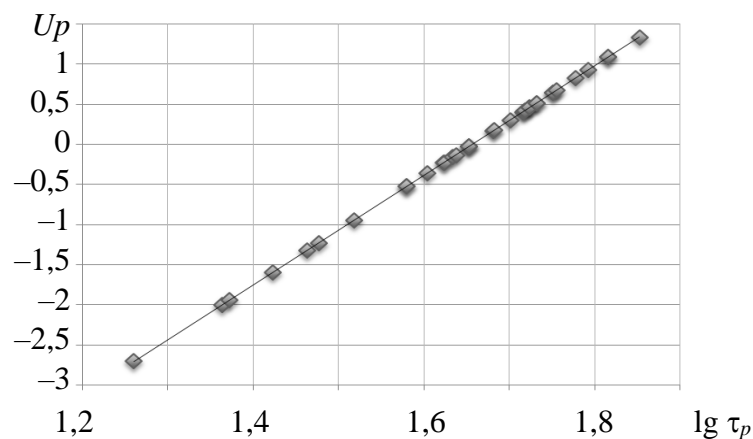


Рис. 3. График эмпирического распределения десятичного логарифма долговечности τ_p сплава ЖС6У-ВИ после эксплуатационной наработки в пределах от 5000 до 8500 часов

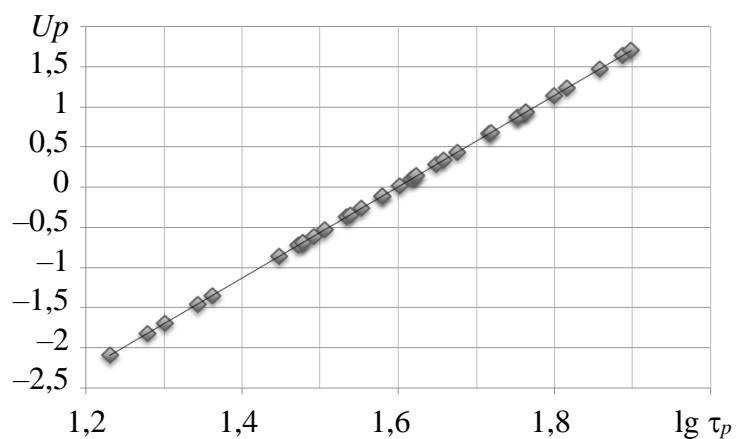


Рис. 4. График эмпирического распределения десятичного логарифма долговечности τ_p сплава ЖС6У-ВИ после эксплуатационной наработки в пределах от 8500 до 10500 часов

Таблица 1

**Проверка нормальности распределения логарифма долговечности $\lg \tau_p$
в исходном состоянии и после эксплуатационной наработки сплава ЖС6У-ВИ**

Наработка	Количество образцов	к	b^2	S^2	$W_{рас}$	W_α	H_0
0	26	13	0.131	0.136	0.963	0.920	+
2400-5000	48	24	1.439	1.460	0.985	0.947	+
5000-8500	37	18	0.728	0.774	0.940	0.936	+
8500-10500	34	17	0.991	1.019	0.972	0.933	+

В таблице 1 использованы следующие обозначения: к – параметр, определяемый по количеству образцов; H_0 – гипотеза о нормальности распределения.

Таким образом, наработка двигателей не оказывает влияния на закон распределения долговечности, который является логарифмически нормальным.

Также необходимо проверить, принадлежат ли выборки к одной общей совокупности. Для проверки равенства дисперсий ряда совокупностей при неодинаковом числе образцов в отдельных выборках использовался параметрический критерий Бартлета [1]. Проверка проводилась при уровне значимости $\alpha = 0,05$. Результаты проверки однородности дисперсий логарифмов долговечности при различных уровнях наработки приведены в таблице 2.

Таблица 2

**Проверка однородности дисперсий логарифмов долговечности сплава ЖС6У-ВИ
по критерию Бартлета**

Параметр	Наработка	n	$\chi^2_{расч}$	$\chi^2_{0,05}$	H_0
$\lg \tau_{p\tau_i}$	0	26	5,87	5,99	+
	2400-5000	48			
	5000-8500	37			
	8500-10500	34			

В таблице 2 использованы следующие обозначения: $\chi^2_{расч}$ – расчетная статистика по критерию Бартлета; $\chi^2_{0,05}$ – расчетная статистика по критерию Бартлета при уровне значимости $\alpha=0,05$; H_0 – гипотеза об однородности дисперсий.

Как видно из приведенных результатов проверки, в данном случае выполняется условие:

$$\chi^2_{расч} < \chi^2_\alpha.$$

Значит, гипотеза о том, что все m генеральных совокупностей долговечностей после наработки для сплава ЖС6У-ВИ имеют равные дисперсии, не отвергается. Проверка по критерию Бартлета показала, что для материала (сплава ЖС6У-ВИ) рабочих лопаток первой ступени двигателя НК-86 имеет место однородность дисперсий логарифмов $\lg \tau_{p\tau}$.

На основе результатов проверки однородности дисперсий рассматриваемой величины проведен анализ однородности средних значений этой величины путем использования процедуры однофакторного дисперсионного анализа.

В таблице 3 представлены результаты однофакторного дисперсионного анализа. Проверка нулевой гипотезы о равенстве средних значений долговечности выполнена с помощью F-критерия [3].

Из таблицы следует, что для всех исследуемых величин гипотеза об однородности средних значений логарифмов долговечностей для выборок во всех случаях отвергается. Значит, эксплуатационная наработка рабочих лопаток двигателя НК-86 из сплава ЖС6У-ВИ оказывает значимое влияние на среднее значение долговечности исследуемого сплава.

Таблица 3

Проверка однородности средних значений логарифмов долговечности сплава ЖС6У-ВИ

Параметр	Число степеней свободы		S_1^2	S_2^2	$F_{расч} = S_1^2 / S_2^2$	$F_{0,95}$	H_0	$F_{0,99}$	H_0
	K_1	K_2							
$\lg \tau_p$	3	144	0,2685	0,0138	19,46	2,68	-	3,94	-

В таблице 3 приняты следующие обозначения: S_1, S_2 – параметры F-критерия; $F_{расч}$ – расчетная статистика F-критерия; $F_{0,95}$ – расчетная статистика F-критерия при доверительной вероятности $P_d=0,95$; H_0 – гипотеза об однородности средних значений; $F_{0,99}$ – расчетная статистика F-критерия при доверительной вероятности $P_d=0,99$.

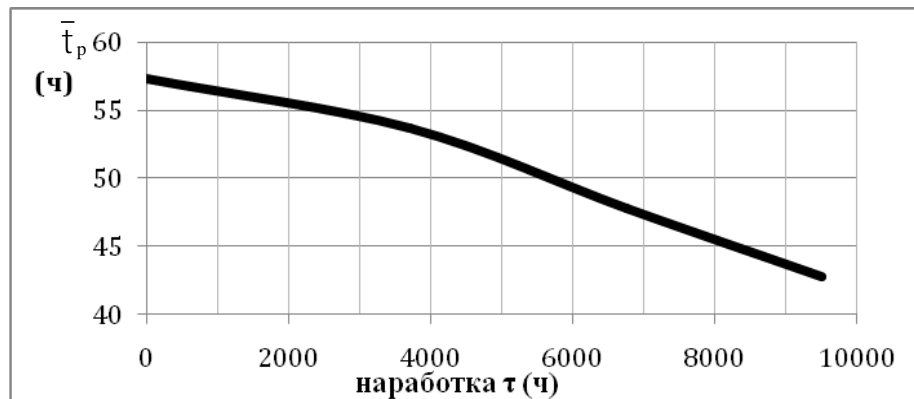


Рис. 5. Графики изменения средних значений долговечности (жаропрочности) \bar{t}_p сплава ЖС6У-ВИ при различных уровнях наработки

На рис. 5 приведен график изменения средних значений долговечности сплава ЖС6У-ВИ в зависимости от наработки.

Из приведенных на рисунке 5 данных следует, что в интервале эксплуатационной наработки до 10500 часов жаропрочность материала \bar{t}_p рабочих лопаток турбин сплава ЖС6У-ВИ снижается на 25% [4].

Выводы

Нарботка в эксплуатации не оказывает влияния на закон распределения долговечности.

При эксплуатационной наработке характеристики рассеяния (в частности, среднее квадратическое отклонение величины долговечности) не изменяются в зависимости от наработки.

Эксплуатационная наработка оказывает влияние на среднее значение долговечности. При увеличении наработки до 10500 часов долговечность сплава снижается на 25%.

Список сокращений

ГТД – газотурбинный двигатель

РЛ – рабочие лопатки

Список литературы

1. М.Н. Степнов. Статистические методы обработки результатов механических испытаний: Справочник / М.Н. Степнов. — М.: Машиностроение, 1985. - 232 с.
2. Г. Хан. Статистические модели в инженерных задачах / Г. Хан, С. Шапиро. Пер. с англ. под ред. Налимова В.В. - М.: Мир, 1969. - 395 с.
3. П. Мюллер. Таблицы по математической статистике: пер. с нем. / П. Мюллер, П. Нойман, Р. Шторм. – М.: Финансы и статистика, 1982. - 278 с.
4. Н.П. Великанова. Анализ влияния эксплуатационной наработки на долговечность жаропрочного сплава рабочих лопаток турбин авиационных ГТД / Н.П. Великанова, А.С. Киселев // Вестник КГТУ им. А.Н. Туполева. – Казань: КГТУ-КАИ, 2011.- №.1. – С. 23-26.

Нина Петровна Великанова, доктор технических наук, профессор, (КНИТУ-КАИ, кафедра реактивных двигателей и энергетических установок, РДЭУ). Область научных интересов: влияние эксплуатационных факторов на долговечность деталей турбин, расчет на прочность основных деталей газотурбинных двигателей; расчет конструкции газотурбинных двигателей.

Фань Вэйвэй, магистрант (КНИТУ-КАИ, кафедра реактивных двигателей и энергетических установок, РДЭУ). Область научных интересов: расчет на прочность основных деталей газотурбинных двигателей.

To heat-temperature strength problem of aviation gas turbine engine blades (statistical analysis)

N.P. Velikanova, Fan Weiwei

Kazan National Research Technical University named after A.N. Tupolev – KAI
10 K.Marx st., Kazan, 420111, Russia

Annotation: The paper examines the influence of long-term operation of civil aviation gas turbine engine (GTE) on the durability (high-temperature strength) of turbine blades material – high-temperature nickel-based ZhS6U-VI casting alloy with equiaxed crystallization. Specific operating time interval has been studied, statistical analysis of empirical distributions of high-temperature strength of metals has been carried out, regularities in empirical distributions have been established for each operating time interval, empirical distributions have been plotted that estimate the effect of operation time on durability.

Key words: strength, durability, blade, turbine, high-temperature alloy, statistical method

Introduction

The subjects of the study are the rotor blades of the first-stage high-pressure turbine engine NK-86, made of ZhS6U-VI superalloy cast with equiaxed structure of the material. This alloy is now widely used to manufacture turbine rotor blades for aircraft engines of civil aviation.

Analysis of the initial level of durability in hours and analysis of the levels after the operation includes the determination of the distribution of the characteristics of durability and test samples belonging to the same general population. The purpose of the analysis is comparison of the initial level of durability of ZhS6U-VI alloy with durability after operation.

The initial level of durability was evaluated based on the results of tests of samples cut from the cold zone of the blades, in which the metal structure remains unchanged. Level of durability of ZhS6U-VI alloys after operation was estimated based on the results of tests of samples cut from the hot zone of the blade, where the effect of operation is the most pronounced.

Samples were tested in accordance with GOST 10145-81 "Metals. Test method for long-term strength". The base level of temperature was $T = 1248 \text{ K}$, which is close to the reference temperature which occurs in the most dangerous section. The level of operating stress was 23 kg/mm^2 for the ZhS6U-VI alloy. The test was performed until the sample destruction.

The tests yielded experimental data on the time to failure τ_p that were subjected to statistical analysis according to the recommendations [1]. In the framework of the statistical analysis, distribution regularities were established and samples were tested for membership in the same general population.

The test results of durability of rotor blades material (ZhS6U-VI) were subjected to statistical analysis to determine the impact of operation on regularities and numerical characteristics of durability distribution. To find the shape of durability distribution after operation, a graphic method and Shapiro-Wilk test criterion [2] intended for the test of the hypothesis of normal or log-normal distribution with a limited sample size ($n < 50$) were applied; the latter is better than other criteria [1].

In order to calculate the statistics of criterion $W_{paa} = b^2 / S^2$, the following criterion is determined

$$S^2 = \sum_{i=1}^n (x_i - \bar{x}_i)^2$$

where \bar{x}_i is the average value, and

$$b = \sum_{i=1}^k a_{n-i+1} (x_{n-i+1} - x_i).$$

Hypothesis is tested on the basis of inequality $W_{paac} \geq W_\alpha$, where W_α is critical Shapiro-Wilk test. If the inequality is satisfied, the null hypothesis is not rejected.

The graphs are well approximated by straight lines, which indicates that the distribution of durability in hours τ_p can be log-normal.

Verification of test data compliance with log-normal distribution by the Shapiro-Wilk test criterion showed that the distribution of the considered characteristics after operation is lognormal with a significance level of $\alpha = 0,05$. According to [1], the significance level α is the probability of making a mistake of the first kind, which consists in rejecting the null hypothesis when in fact it is true.

To illustrate the results of the statistical analysis of operation impact on the ZhS6U-VI alloy durability, Fig. 1 ... 4 show the plots of empirical distributions of $\lg \tau_p$; Table 1 gives the results of testing for normality of $\lg \tau_p$ distribution by Shapiro-Wilk test criterion.

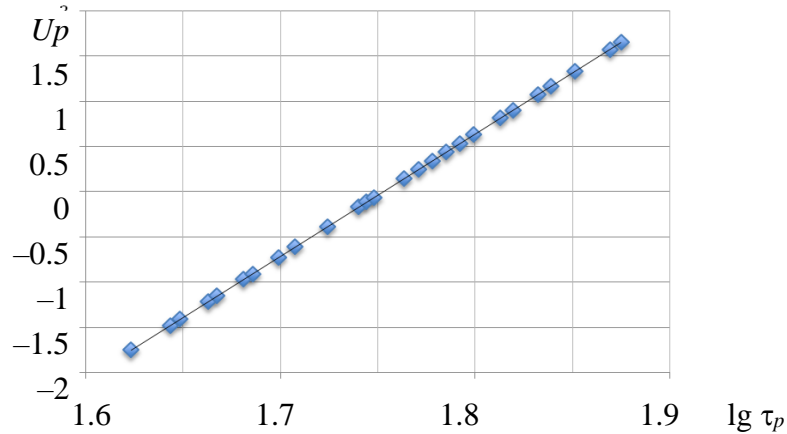


Figure 1. Empirical distribution of logarithm of the original durability of τ_p of ZhS6U-VI alloy; here Up is the distribution function, dimensionless one

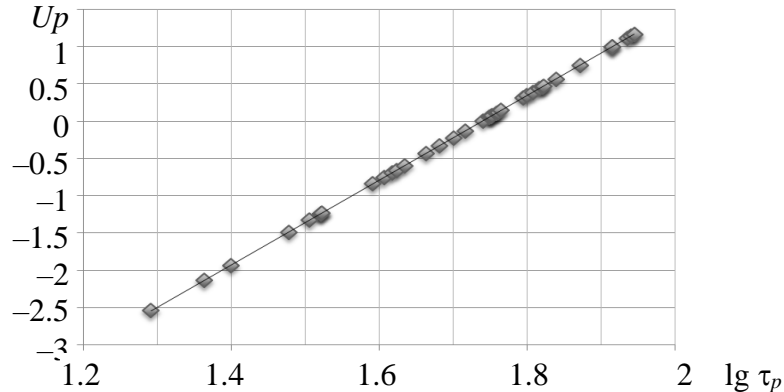


Figure 2. Empirical distribution of logarithm of durability τ_p of ZhS6U-VI alloy after operation for 2400 to 5000 hours

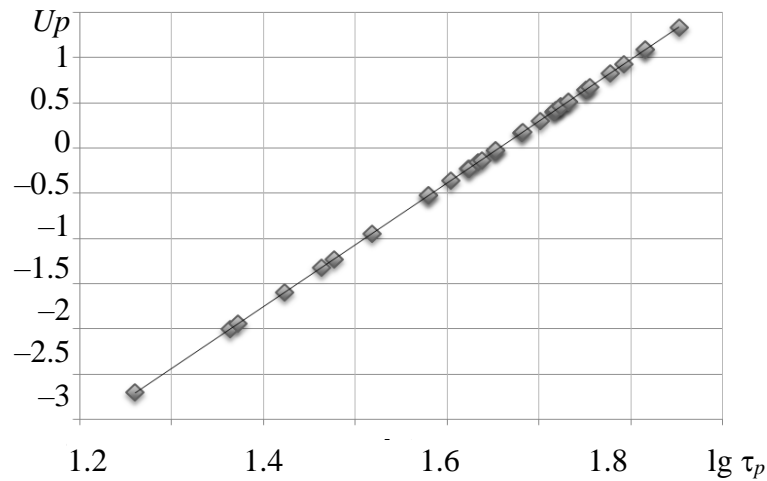


Figure 3. Empirical distribution of logarithm of durability τ_p of ZhS6U-VI alloy after operation for 5000 to 8500 hours

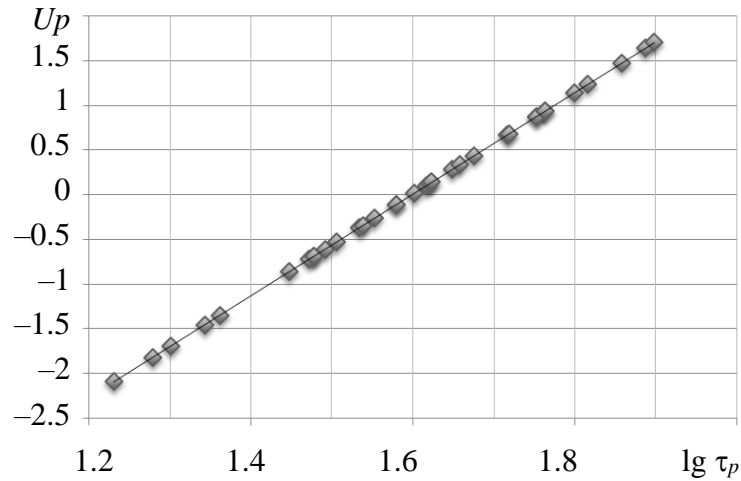


Figure 4. Empirical distribution of logarithm of durability τ_p of ZhS6U-VI alloy after operation for 8500 to 10500 hours

Table 1

Checking the logarithm of durability for normality in the initial condition and after operation for ZhS6U-VI alloy

Working hours	Sample number	k	b^2	S^2	W_{pac}	W_α	H_0
0	26	13	0.131	0.136	0.963	0.920	+
2400-5000	48	24	1.439	1.460	0.985	0.947	+
5000-8500	37	18	0.728	0.774	0.940	0.936	+
8500-10500	34	17	0.991	1.019	0.972	0.933	+

In the Table 1: k is parameter defined on quantity of samples; H_0 is the hypothesis of dispersions normality.

Thus, the engine operating time does not affect the durability of the distribution law which is log-normal.

It is also necessary to check whether the samples belong to the same general population. To test the equality of variances of a number of sets with different numbers of samples, Bartlett's test was performed for individual samples [1].

The testing was conducted at a significance level $\alpha = 0.05$. Testing of homogeneity of variance logarithms of durability under different levels of operation is summarized in Table 2.

Table 2

Bartlett's test of logarithms of ZhS6U-VI alloy durability for homogeneity of variance

Parameter	Working hours	n	χ^2_{calc}	$\chi^2_{0,05}$	H_0
$\lg \tau_{p\tau_i}$	0	26	5.87	5.99	+
	2400-5000	48			
	5000-8500	37			
	8500-10500	34			

Here χ^2_{calc} is calculated statistics on Bartlett's test; $\chi^2_{0,05}$ is calculated statistics on Bartlett's test for significance level $\alpha = 0.05$; H_0 is the hypothesis of dispersions uniformity.

As seen from the above test results, in this case the following condition is satisfied:

$$\chi^2_{calc} < \chi^2_{\alpha}.$$

Hence, the hypothesis that all m general populations of durability after operation of ZhS6U-VI alloy have equal variance is not rejected. Bartlett's test showed that for ZhS6U-VI alloy of rotor blades of the first stage of NK-86 engine there is a uniform dispersion of logarithms $\lg \tau_{p\tau}$.

Based on the results of homogeneity testing of variances of the considered value, the homogeneity of average values of this quantity was analysed by one-way ANOVA procedure. Table 3 presents the results of one-way ANOVA analysis. The null hypothesis of equality of the mean values of durability is tested using the F-test [3].

Table 3

Checking the homogeneity of the average logarithms of ZhS6U-VI alloy durability

Parameter	Number of degrees of freedom		S_1^2	S_2^2	$F_{calc} = S_1^2 / S_2^2$	$F_{0,95}$	H_0	$F_{0,99}$	H_0
	K_1	K_2							
$\lg \tau_p$	3	144	0.2685	0.0138	19.46	2.68	-	3.94	-

In table 3: S_1, S_2 – F-test parameters; F_{calc} is calculated statistics of F-test; $F_{0,95}$ is calculated statistics of F-test for confidence level $P_d = 0.95$; H_0 is the hypothesis of average values uniformity; $F_{0,99}$ is calculated statistics of F-test for confidence level $P_d = 0.99$.

The table shows that for all the studied variables the hypothesis of homogeneity of the average logarithms of samples durability in all cases is rejected. Hence, the operation time of NM-86 engine blades made of ZhS6U-VI alloy has significant impact on the average value of the durability of the considered alloy.

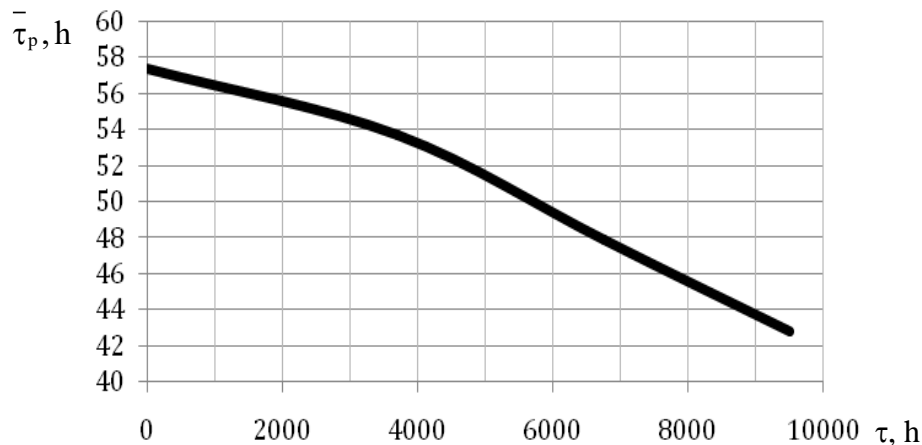


Figure 5. Average durability of ZhS6U-VI alloy for different operation periods

Fig. 5 shows that within the operation time interval up to 10500 hours the high-temperature strength $\bar{\tau}_p$ of ZhS6U-VI alloy of working turbine blades is reduced by 25% [4].

Conclusions

The operation time has not effect on the distribution law of durability.

Operation does not affect the characteristics of scattering – the root-mean-square durability does not vary with the operation time.

Operation affects the average durability. When the operation time increases up to 10500 hours, the alloy durability is reduced by 25%.

References

1. M.N. Stepnov. *Statisticheskie metody obrabotki rezultatov mekhanicheskikh ispytaniy: spravochnik.* (in Russian) [Statistical methods of processing of mechanical test results: Handbook] Moscow: *Mashinostroenie*, 1985. 232 p.
2. G. Hahn, S. Shapiro. *Statisticheskie modeli v inzhenernykh zadachah.* (in Russian) [Statistical models in engineering problems] Moscow: *Mir*, 1969. 395 p.
3. P. Muller, P. Neumann, R. Storm. *Tablicy po matematicheskoy statistike.* (in Russian) [Tables of mathematical statistics] Moscow: Finance and statistics, 1982. 278 p.
4. N.P. Velikanova, A.S. Kiselev. Analiz vliyaniya ekspluatatsionnoy narabotki na dolgovechnost zharoprochnogo splava rabochih lopatok turbin aviacionnykh GTD. (in Russian) [Analysis of the impact of operation on the durability of high-temperature alloy blades of gas turbines // Bulletin of KSTU named after A.N. Tupolev] Kazan: *KSTU-KAI*, 2011. No.1. pp. 23-26.

N.P. Velikanova, Doctor of Technical Sciences, professor (KNRTU-KAI, Department of Jet Engines and Power Systems). The research interests: influence of operational factors on the durability of turbine elements.

Fan Weiwei, graduate student (KNRTU-KAI, Department of Jet Engines and Power Systems). Research interests: calculation of strength of main parts of gas turbine engines.

To problem of science organization (some meditations)

A.A. Bolonkin

V.P. Consulting and Research Co.
1310 Avenue R, #6-F, Brooklyn, NY 11229, USA

A.Einstein. Why socialism (1949):

"...In compliance with legally rules, established as laws, the economic anarchy of the capitalist society leads to appearance of oligarchy, the power of which cannot be efficiently restrained by a democratically organized society. Education in such a society trains not to invent but to worship success in acquisition..."

"...to change the situation by possible only by building the socialist economy and corresponding system of the education..."

Abstract. There is wide discussion in the Russian media about reform and efficiency improvement of Russian science and funds invested in it. According to the Russian habit, many people look to America and offer a copy its mechanisms of research selection and financing of new ideas. But is it all right there in these matters? Failures of the shuttle "Discovery" is only the visible part of the iceberg. Former Soviet scientist who has worked in NASA and in the US public research institutions for a long time, shows as example one of the most important divisions of NASA — Institute of Advanced Concepts (NIAC) – the consequences of the inefficient organization and management and supervising failures.

Key words: reform, Russian science, efficiency, advanced researches, NASA, Space Agency

1. Introduction

NASA Institute of Advance Concepts (NIAC) (critical analysis of NIAC activity and its Director)

*"Secrecy of simple information
by NIAC is concealment of a criminal activity"*

NIAC (NASA Institute for Advance Concepts) is funded and controlled by the U.S. Government Organization (NASA), which is formed also as the department of USRA (Union Space Research Association). The stated official NIAC purpose is:

"Provide an independent, OPEN forum for the external analysis and definition of space and aeronautics advanced concepts.

Complement the advanced concepts activities conducted within the NASA Enterprise."

"Goal: Develop advanced concepts which will result in changes to the nation's future aerospace policies and plans". According with Chapter and Status, NASA is to be a non-military open government organization having civil purposes and falling under the "Freedom of Information Act" (FOIA)(5 U.S.C. § 552). This act obliges to give any information about operations and activities, "... ANY executive department, military department, Government Corporation, Government controlled corporation, or other establishment in the executive branch ...".

What are the facts? Which I had asked numerous times of the NIAC Director, Mr R. Cassanova:

- Why has NIAC not published on a WEB site, the abstracts of all submitted proposals? The NIAC, NASA, other government organizations require special abstracts contained ONLY open information?
- Why doesn't NIAC give names and addresses of organizations, which took part in submissions?
- When, where, were review panels, peer reviewers?
- When, where, and how, were scientists selected for review?
- How many scientists took part in the review process?
- What were the results of voting for every proposal?
- How many reviewers see the proposal?
- Why the debriefings were sent much later, after the awarding process?
- Why debriefing does not contain the scope of the review criteria, and the total scores?

The debriefing does not contain the information: reviewer recommends, or not recommends, this proposal for awarding and development?

Mr. Cassanova did not reply to any of these simple questions. Why??

The explanations are very simple: NO review panels, No peer reviewers, NO scientists who took part in the review process, No voting, No scientists who see the proposals, NO anything! Everything is just fabricated fiction. There is only just Mr. Cassanova in NIAC who changes all reviewers, all scientists (in any scientific fields!), all panels, and all debates. Who distributes tens (hundreds?) of millions government (taxpayers) money to friends and insiders (To grease their palms?). What sort of proposals are awarded by Mr. Cassanova will be considered later?

2. Factious Front (Puppet Organization) NIAC Science Council

Mr. Cassanova has a NIAC Science Council consisting of 9 persons from every considered organization. No information at all (from Mr. Cassanova) who these scientists are? Two of them are named "consultant" without name of affiliated organizations. Where are these consultants?, what scientific field? (unknown?). I requested from Mr. Cassanova, to give the E-mail addresses for this "Science Council". He answered, that this information is secret?? After the large time demands in 2002 Mr. Cassanova names Mr. Whitehead as Council Chairman (who is named only as consultant in NIAC Council list?). I asked Mr. Whitehead – what is the name of the person who appointed him as Council Chairman, and when it was? he doesn't know?! I sent a complaint to the Council and Mr. Whitehead promised to consider it on the next NIAC Council meeting. After many reminders some months later he answered the Council declined your complaint. I asked, when, where was this Council, who was presented at this meeting, what is voting? He doesn't know!

Now Mr. Cassanova has published the new list of the NIAC Science Council (9 persons), which also presented two consultants without organization. I called 7 of the indicated organizations. Two of them answered, none of these names were in their employee directories. Two E-mails did not work. I however, could reach three members of the "Council". All of which did not know who the Chairman of their "Council" was, and did not take part in any panel reviews, and have never seen my proposals.

I asked Mr. Cassanova: who is the Chairman of NIAC Science Council for 2004, the specialty of the Council members, their affiliated organizations, some brief information about their scientific activities, their E-mail addresses, etc. No reply, because the NIAC Science Council is only a scam, a front, a false decoration.

I am very surprised, as to why these scientists take part in a fake and factious NIAC Science Council, and allow Mr. Cassanova to make all the decisions for them, and to use their names for the cover of his criminal activities.

3. False debriefing

In conscientious competition, a funded organization sent proposals to independent reviewers (minimum 2), who wrote dated recommendations, and may be granted this proposal, or not, and why. This special science Committee studies these reviews, studies proposals, compares them, votes on them, and writes protocol why the Committee selected these approved proposals for granting.

I asked Mr. Cassanova many times to send me dated reviews of my proposals, resulting votes, and the decision of the Committee regarding my proposals, date of debriefing, etc.

No reply! Why? Because nothing in NIAC is for real, and is just a front. No reviews, no Committee, no peer reviewers, no discussing, no voting, no protocol. Just Mr. Cassanova knows the winners BEFORE announcing of Call Proposals. Therefore, all the other proposals are called for as decoration, for factious purposes, as a total deception. The debriefing is written AFTER (sometimes a month later) announcing of winner and only for the troublesome applicants. Mr. Cassanova writes all debriefing in all the scientific fields (!!). These “debriefings” show poor knowledge of Mr. Cassanova in many space problems (launch, engine, powers system, space flights, etc). He doesn’t understand or recognize new revolutionary concepts, their prospects or possibilities. He requests unrealizable, idiotic demands, claims in debriefing (for example, to present solutions of all problems in proposal (???), or a space catapult of 1070 km length is the same as aircraft carrier catapults).

Every serious organization gives brief information about its chief and members of Science Council, Committee, Trust, Founders, etc in WEB. I required Mr. Cassanova what is his education, specialty, and list of main scientific works. No reply. Apparently, he is specialist in all sciences.

4. Activity of NIAC and its Director

Every serious organization is lead by a competent Chair, who chooses (picks out) and solves the main problems first of the most importance. The main problems of space exploration and routine access to space, is decreasing the cost of launch. The launch of a 1 Lb payload presently costs about \$10,000, and a 30,000/lb payload at the present time, is a very high and restrictive price. We cannot have routine orbital & space flight and exploration of Space, if we do not decrease the cost of a launch in the magnitude of a hundred times. This means that the main development finances, resources, and money (70-80% of research efforts) must be funded and directed into new & advanced launch technology, engine technology, and power systems.

During the last 6 years of the existence of NIAC, it gave 113 grants: 92 grants of 40-80 thousands dollars and 21 grants of 400-800 thousand dollars. That was a total spent of about 20 millions dollars plus 30% for Mr. Cassanova operations!

What do we see in NIAC? Three grants are given in computer (for example, 3D Viewing of Images on Basis of 2D Images), 4 grants in biology (spent more \$1,5 millions!) (for example, bio-suit), 10 grants in Astronomy?), 5 grants in life Sciences (for example, spent about 1 millions dollars only in chameleon suit), 10 grants in Earth Scientist, 12 grants in Robotics, 8 grants in aeronautics (for example, Planetary Exploration Using Biomimetics).

Most grants declared as advanced propulsion do not have any relation to space propulsion (for example, Cyclical Visits to Mars via Astronaut Hotels, spent about 1 million dollars).

In launch & propulsion system’s Mr. Cassanova grants only old ideas having tens and hundreds of research articles. For example, Space Elevator (Tsiolkovsky offered in 1894, Artsatyunov published a detail article 50 years ago, Cassanova granted about one million in 2002 (?)), tether system (15 years old, it has about hundred articles, it granted 1,5 millions

dollars), magnetic sail (10 year old), etc. The most NIAC reports are compilation of published articles, contain only results are gotten by others authors.

Take, for example, NIAC Grant Awards list in 2001:

1. *Adaptive Observation Strategies for Advanced Weather Prediction.*
2. *Controlling the Global Weather.*
3. *Global Observation...* There are special weather Institute (NOAA) for this topics. NSF also accepts this topic, but does not accept the revolutionary launch and propulsion systems. Why do NIAC give the grants (awards) them?
4. *3D Viewing of Images on the Basis of 2D Images.* There are a lot of organizations which award in computer science (include special Department USRA). Why Mr. Cassanova awarded them?

The same question applies to the following topics:

5. *A Flexible Architecture for Plant...*
 6. *Astronaut Bio-Suit System ...*
 7. *A chameleon Suit ...*
 8. *Directed Application of Nanobiotechnology...*
 9. *Achieving Comprehensive Mission Robustness.*
 10. *Architecture of Intelligent Earth Observation Satellite for Common Users in 2010-2050.*
 11. *Ultrahigh Resolution X-ray Astronomy....*
- etc. That is only 11 from 18.**

For all these topics there are a lot of specialized organizations, which accept, develop, give grants, and awards to these types of proposals. However, there is only ONE organization (NIAC from NASA), which “**MUST**” support new revolutionary launch and propulsion systems. Why does NIAC (Its Director Mr. Cassanova?) award limited and precious resources to other non-related topics? Does he help just his friends and NASA insiders?

Maybe there are just NO good proposals, which suggest new launch and propulsion systems? I know more 18 proposals which offer new revolutionary launch, propulsion ideas. These concepts decrease the launch cost in thousands of times, and offer no expensive installations. They are cheaper then rocket systems by tens of times.

For example: building space elevator from ground without rockets and space flights; transport system Earth-Moon without rockets; the “Space Launcher” which allows to launch in space thousands tons of payload annually in price 1-2 dollars/lb; method which allows to change speed and trajectory of space ships and probes without expending fuel by employing the use of any space bodies (small asteroids, meteorites, comets) (increase ship speeds up in 1-2 km/sec, get any direction and energy from small space bodies). That increases the probe capability in ten million times! All can be done with current technology.

These real revolutionary concepts need support, development, initial small funding. Mr. Cassanova does not understand them and makes the idiotic demands: “first, solve all problems of these concepts?”.

For example: I received debriefings, which show the full incompetence of Mr. Cassanova in physics, in propulsion systems, and in space. If it is necessary, I can present the details of the answers of nonsense, which he wrote, and I am ready for the open discussion about new ideas, which can make a new revolution in an access to Orbit, Planets, and Space.

Overview: The NIAC spent more 30 millions dollars in 7 years, but they did not suggest new concepts or ideas?! Now the NIAC is just a private manger for “friends” and has spent 90% of government (taxpayers) money, not very effectively, and specifically in fraudulent and criminal ways.

5. Connivance from NASA

I saw the NASA WEB site. There are many NASA self-clarifications. However, “NO” simple information about E-mail of leaders and the chairman of Departments, Advisory Committees, Science Councils.

I asked NASA who lead NIAC from NASA? What is the E-mail of Deputy Administrator, Mr. F. Gregory; Chief of Technology - Mr. M. Hirschbein; Chief of Advisory Council? No reply for my questions!

The press wrote about the criminal activity of Mr. Cassanova and many scientists complained in NASA about NIAC activity. However, in 2004 the NASA makes a new contract with NIAC?! I asked for a copy of NASA-NIAC Agreement about funding NIAC by NASA. NASA must give FREE this information according The Freedom of Information Act if it is less 100 pages and not request more 2 hours for copying. NASA cannot find this agreement in a half of year (?!). After many reminders, NASA sent me 10 pages Agreement + 60 pages of garbage, and then begins to require \$100 dollars!

As I understand it, there is “NO” control over the tens of millions of dollars, which NASA gives Mr. Cassanova for distributions as grants. Mr. Cassanova can spend them for any of his friends or just plain un-useful people and unproductive efforts.

I worked at NASA, and I know a lot of NASA specialists who have very low qualifications. That is the results of NASA illegal employee policy. Scientists can get positions in NASA only on the quiet, under the table, or with class protection. I have a doctor of space science and have lots of real experience, and a lot of scientific articles, inventions, and applied for a vacant position for a research project manager; but NASA took a pilot without any scientific experience, scientific works, or inventions?! When this scientist wanted the position of science engineer, NASA wrote more support to a young bachelor who doesn’t have any experience, scientific works, or inventions.

This is explanation why NASA has a lot of failures (for example, in Mars missions, when NASA “scientist” put in computer pounds and foots instead of international standard kilograms and meters and probe of cost \$250 millions flights in incorrect direction). You can tell: NASA has also successes; but “Don’t Forget”, **NASA has gigantic financing, more then all others countries together. If you will use conventional criteria: use/money you would see that NASA has the worst efficiency compared to all the other space agencies in World.**

For example, in the Big Space Race the Soviet Space Agency had 3 to 5 times less funding then NASA, but they were the leader for many years, while America reached for the Moon. After this the Soviet Space Agency had 5 to 7 times less financing then NASA. But the USSR launched up to 1989 (collapse of USSR) 2 to 3 times more Earth satellites then all countries together, including the USA.

Mr. Sean O’Keefe forced his way into the position of NASA Administrator, even through disparate critics of NASA activity attempted to block. However, he only makes small decreases in the NASA defects.

6. Connivance of Union Space Research Association (USRA) and its President

NIAC is also formally the Department of USRA. May be Mr. Black the President of USRA doesn’t know about the criminal situation in NIAC? There were some complaints to Mr. Black, but no replies. The E-mail addresses of USRA Executive director Mr. David Cummings and Chief Engineer Mr. Lewis Peach indicated in USRA WEB side are false. Mr. Black did not answer any simple questions, such as: who leads Mr. Cassanova?; who selects (or appoints) chair and members NIAC Science Council and NIAC Grant Committee?; what

are E-addresses of Chair and members of Science Council, Committee?; when and where was (are, will be) NIAC peer reviewers?; etc.

7. Conclusion

Some final mediations. I offered USRA (and NASA) a detailed plan on how to improve the work of NIAC, make more open and useful their work, and to change the dismal situation when one man (Mr.Cassanova) distributes tens millions of taxpayer money with no safe guards or oversight. There was No Reply?!

In this situation, it is the best decision, to stop the wasteful and ineffective financing of NIAC and pass their functions to another organization, **for example**, the growing and historically relevant and important **International Space Agency (I.S.A.) Organization**, (<http://www.international-space-agency.org>)-or-(<http://www.isa-hq.net>) which would be better suited, and able, to stimulate, enable, and promote advanced space launch, propulsion, power, orbital, and planetary grant disbursement, research, development, and implementation. This, based on an ever-increasing need for global cooperation, collaboration, effort, and common viewpoint. The International Space Agency (I.S.A.) Directives, Charter, Purpose, Goals, and Certificate of Incorporation reflects this reality far better than the USRA or NIAC directives or charters. The many millions in Government (Tax) & Private Money and Resources would be better used, and better served, under the management and oversight of the International Space Agency (I.S.A.) Organization.

The President and Congress of the United States of America, needs to, and must, thoroughly investigate the NIAC situation and must punish, and remove, NASA and USRA leaders who allow, and create the abuse and corruption from, and by, NIAC. I am ready to present to a Special Investigation Commission the documents which confirm the statements presented and outlined in this article.

For epilog. **Award on the quiet.** **Soviet rocket scientist Konstantin Eduardovich Tsiolkovsky is considered the father of human space flight. He was also the one to conceive the idea of Space elevator one hundred years ago (1894). Half a century afterward, another Russian scientist Yuri Artsutanov provided all required technical aspects for the idea to be put into work (1960). Today, American scientist Bradley Edwards is awarded a prize (more than half a million dollars) for the invention.**

Nowadays, a rather simple cabin intended to transport people to space seems to cause major controversy. Apparently, few are aware that Konstantin Tsiolkovsky was the one to come up with the idea back in 1894. In his famous work "Dreams of the Earth and the Sky" he mentioned the idea of the cosmic lift for the first time. A bit more than sixty years have passed and another Russian scientist Yuri Artsutanov (at the time, he was a post-graduate student of Leningradsky Technical Institute) in a Sunday's supplemental issue to the "Komsomolskaya Pravda" newspaper on 31 July 1960, developed and provided all the necessary technical calculations for the idea to be put to work.

Space elevator in itself is a space station situated on or above the geostationary orbit and attached to the Earth by means of a strong cable. The system is devised in such a way that its center of gravity is located above the geostationary orbit. Since the Earth rotates, centrifugal force tightens the cable, similar to a stone tightening a sling, even if it is a child holding the sling. In other words, it is of utmost importance to create such strong cable in the first place. The stuff that will be carried by it appears secondary. Artsutanov suggested the idea of a power train attached to the cable to transport space tourists to and fro the space station. In fact, the idea is so simple, it seems almost unreal. Endless disputes have been evolving around it for the longest time. However, not so long ago, scientific world was quite shocked to learn

that an American by the name of Bradley Edwards had been awarded a prize for the idea of space elevator.

How could this happen? Did Mr. Edwards make some significant changes/additions to the idea that had been devised by the Russian scientists? As it turns out, director of one of NASA's research centers: NASA Institute of Advance Concepts (NIAC) someone by the name of *R. Cassanova has been the one to decide to award Mr. Edwards the prize.*

Mr. Cassanova is a well-known man within the scientific world. He is mostly known for his multimillion dollar machinations involving taxpayer's money, which had been initially planned to cover space research. Press has already reported Mr. Cassanova's "financial" approach to scientific discoveries in article entitled "NASA Institute of Advanced Concepts (NIAC) as a front for friends." (<http://NASA-NIAC.narod.ru>).

This time, the award appeared to be substantial - \$645 000 USD! In addition, Mr. Cassanova has been awarded 30% "for his efforts."

But perhaps, Mr. Cassanova and his accomplice conducted such major scientific research work, which had outshined hundreds of magnificent works of scientists in the field of aerospace, and what if this award is a mere excuse to thank them for their efforts in other fields of research? No. This could easily be spotted on the Internet site of the American Institute off Aeronautics and Astronautics, the main space organization entitled to register all scientific research works and conferences (<http://www.aiaa.org> , search "Cassanova"). According to it, Cassanova "Ph.D" has written only one (!!!) insignificant work in his life. That article has been by Cassanova and three other co-authors in 1976!

Are you still doubtful? Well, let me direct you to the NIAC site (http://www.niac.usra.edu/files/studies/final_report/521Edwards.pdf),

where you can personally see for yourself what sort of work has been awarded such high prize. And don't be freaked out by the idea of encountering some mind-boggling scientific calculations. There aren't any! The 42-page report, half of which is mere illustrations, represents a mere explanation of the idea of the space elevator intended for elementary school pupils. It lacks the actual results. It is very incongruously in regard to Konstantin E. Tsiolkovsky...

This, however, is not the first award of Cassanova's. Perhaps, not the last either. In the meantime, hundreds of scientists are left without funding to continue their truly meaningful and significant research. This should be quite interesting food for thought for American taxpayers.

Summary. As final result of this analysis on problem of effective Science organization it is *necessary* to take the thesis of Albert Einstein, Nobel Prize laureate, that is presented in the beginning of this paper. Albert Einstein gives the useful recommendation to activity for Mankind evolution in whole.

Alexander Bolonkin, Dipl.Eng., Ph.D.; Senior Research Associate of the USA National Research Council at Eglin AFB, Former Senior Researcher of NASA; he graduated Kazan Aviation Institute (1958); he worked as engineering leader at Antonov Aircraft Construction Bureau (Kiev), chairman of reliability department at Glushko Rocket Engine Construction Bureau (Moscow), as a professor (teacher) in different Russian and American Universities.

Dr. Bolonkin is member of Director Board of International Space Agency; the area of scientific interests is the problems of aviation systems; nine of his reports on actual problems have been approved by the World Space Congress (2002, USA).

Как организовать науку (некоторые размышления)

А.А. Болонкин

Международное космическое агентство
1310 Avenue R, #F-6, Brooklyn, NY 11229, USA

А.Эйнштейн. «Почему социализм» (1949):
«...Экономическая анархия капиталистического общества приводит с помощью правил, устанавливаемых на уровне закона, к появлению олигархии, чью власть демократически организованное общество не может эффективно ограничивать. А образование в таком обществе учит не изобретать, а поклоняться успеху приобретательства...»
«... изменить ситуацию можно только путем создания социалистической экономики и соответствующей системы образования...»

Аннотация. В Российской прессе широко обсуждается вопрос о реформе и повышении эффективности российской науки и вкладываемых в нее средств. По российской привычке многие смотрят на Америку и предлагают копировать ее механизмы отбора исследований и финансирования новых идей. Но так ли там все хорошо в этих вопросах? Неудачи с шаттлом «Дискавери» – это только видимая часть айсберга. Бывший советский ученый, проработавший в НАСА и государственных научных учреждениях США длительное время, показывает на примере одного из важнейших подразделений НАСА – Института Перспективных Исследований (NIAC) – к чему приводят неэффективная организация и просчеты в менеджменте и в контроле.

Ключевые слова: реформа, российская наука, эффективность, перспективные исследования, НАСА, Космическое Агентство

1. Введение

Институт Перспективных Исследований НАСА – NIAC (критический анализ деятельности NIAC и его директора)

*«Скрытие простой несекретной информации
открытым учреждением – верный признак
криминальной деятельности.»*

Институт Перспективных Концепций НАСА (NASA Institute for Advanced Concepts, NIAC, или НИАС) – это «виртуальная» организация, финансируемая и контролируемая правительственной организацией США НАСА (NASA, Национальная Авиационно-Космическая Администрация). Одновременно NIAC является отделом Союза Университетов для Космических Исследований (Union Space Research Association, USRA, или УСРА). Цели NIAC, провозглашенные в ее уставе, следующие:

«Проводить независимый ОТКРЫТЫЙ форум для всеобщего анализа и отбора космических и авиационных перспективных исследований. Дополнять передовыми концепциями деятельность НАСА.»

«Цель: Развивать передовые идеи, которые будут корректировать национальную космическую политику и планы освоения космоса». Оказывать помощь индивидуальным ученым и малым организациям.

Согласно Уставу и Закону НАСА является гражданской открытой государственной организацией, имеющей только гражданские цели и подпадающей под действие «Акта о Свободе Информации» (FOIA)(5 U.S.C. № 552). Этот акт обязывает руководителей, сотрудников, исполнительные отделы давать любую информацию о деятельности организации. В законе сказано: «Любой исполнительный отдел, военный отдел, государственная корпорация, финансируемая правительством организация обязаны давать требуемые сведения».

Каковы действительные факты? Проанализируем ситуацию на конкретных примерах. В свое время директор NIAC м-р **Казанова** (Cassanova) отказался отвечать на простые вопросы о деятельности своей организации (состоявшей из него и его помощника), о принципах и методах отбора предложений, и, более того, он не отвечает даже авторам проектов по их предложениям. Например, мы многократно запрашивали у него ответы на такие вопросы:

1. Почему NIAC не публикует на своем сайте краткое содержание (аннотации) представленных предложений (проектов)? Даже Министерство Обороны требует от подателей предложений представлять в абстрактах только открытую информацию. А NIAC (как и НАСА) по Уставу обязана проводить ОТКРЫТЫЙ форум поданных предложений.
2. Почему NIAC отказывается давать имена и E-адреса (электронные адреса) организаций и лиц, которые подают свои предложения и проекты?
3. Когда и где проводится открытое обсуждение представленных предложений?
4. Кто входит в Отборочный Комитет?
5. Когда, где, как и какие ученые участвуют в отборе предложений и проектов?
6. Какое количество специалистов рецензирует каждое предложение?
7. Почему NIAC отказывается представить эти рецензии авторам предложений?
8. Сколько ученых входит в Отборочный Комитет?
9. Каковы результаты голосования по каждому предложению?
10. Почему краткий перечень положительных и отрицательных сторон предложения, дебрифинг (если NIAC и представляет его в редких случаях особо настойчивым участникам) высылается через длительное время после объявления победителей?
11. Почему дебрифинг не содержит количества очков (или голосов Отборочного Комитета), которое получило данное предложение?
12. Почему дебрифинг не содержит главной информации: рецензент рекомендует или не рекомендует финансировать данное предложение?

Г-н Казанова не ответил ни на один из этих вопросов... Почему?

Сейчас понятно, что объяснение очень простое: НЕТ никакого Отборочного Комитета, НЕТ никаких рецензий, НЕТ никакого обсуждения, НЕТ никаких ученых, которые бы принимали участие в обсуждении и отборе действительно ценных идей, предложений, инноваций, НЕТ ученых, которые бы вообще видели эти предложения! НЕТ ничего! Все фикция! Есть только ОДИН г-н Казанова *во всем NIAC*, который заменяет всех рецензентов, всех ученых (во всех областях), проводит он «в единственном числе» заседания, обсуждения, дебаты, голосование. И только он распределяет десятки (сотни?) миллионов государственных денег (денег налогоплательщиков!) среди своих друзей и начальства (это – коррупция, «смазка»). Известно только, что 30% от сумм, выдаваемых ему НАСА, он отчисляет в свою пользу.

За какие предложения г-н Казанова выдает награды в сотни тысяч и миллионы долларов мы рассмотрим позднее.

2. Марионеточный наблюдательный Ученый Совет

М-р Казанова формально имеет в НИАС наблюдательный Ученый Совет из 9 человек из разных организаций, людей проживающих в разных городах. Но Казанова не дает ответа на вопросы: кто эти люди? Двое на его сайте названы консультантами без указания, где они работают, кого и в какой области консультируют. Специальность ни одного члена «Ученого Совета» не указана. Я попросил м-ра Казанова дать электронный адрес этого «Ученого Совета» (и его членов). Он ответил, что это – большой секрет. Например, в 2002 году после многочисленных требований м-р Казанова сообщил, что председателем Совета является г-н Вайтехит (Whitehead) и дал его электронный адрес. Но на сайте НИАС, «сочиненном» г-ном Казанова, м-р Вайтехит числится только консультантом (неизвестно, где и у кого). Я запросил у г-на Вайтехит имя человека, который его назначил председателем Совета, где и когда это было? Он ответил, что *не знает*. Я послал претензию, что г-н Казанова отказывается отвечать на простые вопросы о моих предложениях. Г-н Вайтехит пообещал рассмотреть мое заявление на ближайшем заседании Совета. После многочисленных напоминаний, спустя несколько месяцев он сообщил, что «Совет» отклонил мое заявление. На мой вопрос: когда, где был этот «Совет», кто на нем присутствовал, каков результат голосования? М-р Вайтехит ответить не смог.

В последующее время г-н Казанова объявил на сайте НИАС новый состав «Совета» (9 человек), где по-прежнему присутствуют два консультанта, консультирующих неизвестно кого, неизвестно где, неизвестно по каким вопросам. У 7 членов названы организации, где они якобы работают. Находим сайты этих организаций, их справочники, отделы кадров. Выясняется, что двое никогда не работали в указанных организациях, у двоих – старые, неработающие электронные адреса (т.е. они там давно не работают), трое ответили, что они не знают, кто председатель их «Совета», они не принимали участия в обсуждениях и никогда не видели названных им предложений (в том числе, и по улучшению работы и открытости НИАС и самого Совета).

В свое время я запрашивал г-на Казанова: кто председатель этого нового «Совета», каковы специальности членов, где они работают, краткую информацию об их научной активности, электронные адреса. Нет ответа. Почему? Очевидно, что весь этот «Совет» – просто декорация для прикрытия *криминальной коррупционной* деятельности г-на Казанова. Как этот «Совет» может «наблюдать» за деятельностью г-на Казанова (а тем более, – контролировать его), если он составлен из людей, проживающих в других городах, существует на «подачки» Казанова, собирается раз в год, когда деньги уже распределены среди нужных людей.

Я очень удивлен, как эти «ученые» участвуют в фиктивном «Ученом Совете», позволяют г-ну Казанова делать из них марионеток, решать все за них и использовать их имена для прикрытия его противозаконной деятельности.

3. Фальшивый дебрифинг

Любая уважающая себя организация, а тем более организация, распределяющая государственные (налогоплательщиков!) деньги на основании открытого конкурса, посылает предложения независимым рецензентам (минимум двум), которые определяют, насколько данное предложение ценное и почему, стоит ли его

финансировать, ставят дату рецензирования. Специальный Комитет из ученых изучает поданные предложения, рецензии, анализирует, сравнивает, голосует и составляет протокол – какие предложения и почему он рекомендует для финансирования.

Я запрашивал у г-на Казанова: сколько человек рецензировало предложения, результат голосования, решение Комитета, дату дебрифинга и т.п.

Нет ответов (как и другим людям, и даже авторам проектов...). Почему? Да потому, что повторяем: нет никаких независимых рецензентов, рецензий, никакого отборочного Комитета, нет сравнения полезности предложений, обсуждения, голосования, протокола. Есть только «кормушка» для друзей и начальства. Г-н Казанова знает победителей ДО ТОГО как будет объявлен конкурс. Этот «конкурс» и все документы – это для «декорации», для создания *видимости* деятельности. «Дебрифинг» (рецензии) пишется ПОСЛЕ (иногда через месяц и более) того, как объявлены «победители» и только ПОСЛЕ многократных требований обманутых участников. При этом все рецензии (дебрифинг) пишет один человек (г-н Казанова), причем – во всех научных областях (!!). Эти рецензии показывают полную некомпетентность «доктора» Казанова в большинстве космических проблем (космический запуск, полет, двигательные системы, силовые системы и т.п.). Он не знает, не понимает и не в состоянии определить, какие концепции являются новыми, революционными, их осуществимость, возможности и перспективы. Он предъявляет бессмысленные, нереальные требования в своих «дебрифингах»: например, чтобы решение ВСЕХ проблем было уже представлено в предложении или, например, утверждает, что космическая катапульта длиной 1070 км и обычная корабельная самолетная катапульта, имеющая максимум 150 метров длины, – это одно и то же (???) ...

Любая серьезная и уважающая себя организация дает краткую информацию на своей интернет-странице о своих руководителях, членах Ученого Совета, отборочном Комитете, Опекунском Совете. Я запросил г-на Казанова о его образовании, специальности, список его научных трудов. Нет ответа. На сайте AIAA (Американский Институт Астронавтики и Аэронавтики), который издает в США все авиационно-космические журналы и проводит все научно-технические конференции по данной тематике (<http://www.aiaa.org> , search “Cassanova”), представлено, что м-р Казанова был соавтором (одним из 4-х авторов!) только ОДНОЙ незначительной статьи, написанной очень давно (1976 г.)

4. Деятельность НИАС и директора института

Каждая серьезная организация, если во главе ее стоит компетентный руководитель, выбирает и решает самые важные первоочередные проблемы. **Самая главная задача, которая стоит сейчас перед космонавтикой и тормозит все ее развитие – это снижение стоимости космических запусков.** Стоимость запуска 1 фунта (454 гр) полезной нагрузки колеблется в настоящее время от \$10000 до 30000 долларов США. Это слишком большая цена. Мы никогда не сможем начать программы по широкому освоению космического пространства, по массовым туристическим полетам в космос, пока не снизим стоимость запуска хотя бы в сотни раз. **Это значит, что основные средства и усилия (70-80%) должны вкладываться в новые перспективные технологии запуска, в двигатели и силовые системы.**

Результаты: в течение 6 лет существования НИАС НАСА профинансировала 113 проектов НИАС: 92 проекта по 40-80 тысяч долларов и 21 проект по 400-800 тысяч долларов. Общие расходы – более 30 миллионов долларов плюс 30% мистеру Казанова.

Какие же проекты финансировал НИАС? Три проекта относятся к компьютерным наукам (для примера – 3-х мерный образ на основе 2-х мерного изображения); 4 гранта по биологии (затрачено более 1,5 миллиона долларов!), для примера – мифический биокостюм, – на человека «*напылять*» *слой пластмассы*. Ни одного образца костюма не сделано, только «пустые слова»; 10 грантов на астрономию (?); 5 грантов на науку о жизни (например, 1 миллион долларов на костюм изменяемого цвета); 10 грантов на науку о Земле (?); 12 грантов на роботов; 8 грантов на аэронавтику, например – освоение планет с использованием биометрики.

Многие гранты, декларируемые м-ром Казанова как гранты на проекты по двигательным системам, на самом не имеют отношения к ним (например, проект по тематике «Периодические визиты на Марс через космический отель»..., впустую растрочен 1 миллион долларов!).

В системах запуска г-н Казанова финансирует только старые идеи, опубликованные в десятках и сотнях научных статей (чтобы было откуда «списывать» его друзьям?). Хотя по положению НИАС должен финансировать ТОЛЬКО новые, перспективные идеи и концепции. Например, за давно и хорошо известную идею космического лифта (К.Э.Циолковский предложил эту идею еще в 1894 г., а советский инженер Ю.Н.Арцутанов в 1960 году написал статью «В Космос — на электровозе», где он обсудил концепцию космического лифта как экономически выгодный, безопасный и удобный способ доступа к орбите для облегчения освоения космоса.) г-ну Казанова предоставили финансирование - один миллион долларов за бумажный компиляционный отчет (фактически это – школьный реферат на популярном уровне...).

За бесперспективную тросовую систему (длинная проволока вращается около Земли) (возраст идеи – более 15 лет, написано сотни статей) г-н Казанова «предоставил» друзьям 1,5 миллиона долларов. Он дал гранты на магнитный парус (возраст 10 лет), на электростатический парус (работа содержит только принципиальные ошибки). Он выделил сотни тысяч долларов за проект двигателя на анти-материи (12 страниц пустой болтовни, содержащей грубые принципиальные научные ошибки). Почти все отчеты НИАС представляют компиляцию чужих статей, содержат только результаты других авторов, большинство из которых не входят в «группу» Казанова; при этом он отказывает им в финансировании их действительно *перспективных* исследований.

Рассмотрим для примера типовые гранты НИАС за один год.

1. Порядок наблюдений для предсказания погоды.
2. Управление всемирной погодой.
3. Наблюдение за погодой.

Заметим, имеется специальный общеамериканский государственный институт погоды (NOAA) для исследований по этим темам. Национальный научный фонд (NSF) также принимает предложения на исследования по данной тематике, но не принимает предложения на новые космические пусковые и двигательные системы (отсылая авторов в НИАС). Так почему НИАС выдает огромные деньги на исследования, вообще не имеющие отношения к его тематике? Да потому, что надо придать законный вид коррупционной деятельности.

4. 3-х мерный образ на 2-х мерном экране.

Имеется огромное число организаций финансирующих компьютерные науки (включая специальный отдел USRA). Так почему г-н Казанова финансирует эту тему? Почему

вообще деньги, предназначенные для исследования космоса выделены «умнику», который прислал предложение о развитии компьютерных наук? Ответ приведен выше.

Такие же вопросы возникают при анализе других тем:

5. Гибкая архитектура растений...
6. Космический Био-костюм,
7. Цветовой костюм.
8. Применение нанобиотехнологии.
9. Общая выживаемость команды.
10. Архитектура наблюдательных спутников Земли в 2010-2050 году.
11. Разрешимость рентгеновских лучей в астрономии.

Здесь только 11 из 18 профинансированных Казанова проектов в течении одного года.

Повторяем –

для всех этих тем имеется много специализированных организаций, которые принимают, развивают, дают гранты и финансируют подобные проекты. Однако есть только ОДНА организация (НИАС как подразделение НАСА), которая ДОЛЖНА поддерживать **НОВЫЕ РЕВОЛЮЦИОННЫЕ ИССЛЕДОВАНИЯ ПО ТЕМАТИКЕ ПУСКОВЫХ И ДВИГАТЕЛЬНЫХ СИСТЕМ**. Так почему НИАС (точнее, г-н Казанова) тратит ограниченные и столь необходимые для освоения космоса ресурсы на темы, не имеющие отношения или имеющие весьма отдаленное «непрямое» отношение к космосу? Ответ элементарный – эти деньги (деньги налогоплательщиков) он «передает» себе, друзьям и своим покровителям.

Но может быть нет хороших предложений, новых перспективных идей по космическому запуску, двигательным системам и снижению стоимости запуска? М-р Казанова скрывает все другие предложения, тематику предложений, их аннотации, описания, имена компаний, их адреса. Ибо г-н Казанова сразу же был бы изобличен научной общественностью. Например, **мне известно более 18 прекрасных предложений новых революционных идей по запуску и двигательным системам космических кораблей. Они снижают стоимость запуска в тысячи раз. Пусковые установки для них в десятки раз дешевле, чем затраты, сделанные на развитие пусковых установок для ракет.**

Приведем некоторые примеры таких проектов:

Предложения о строительстве Космического элеватора без использования ракет и космических полетов; Безракетная транспортная система Земля-Луна; Космический пускатель, позволяющий запускать в космос тысячи тонн груза в год по цене 1-2 доллара за фунт (фунт = 454 гр.); Метод, позволяющий использовать любые космические тела (астероиды, метеориты, а их - сотни миллионов в космосе) для изменения скорости и траектории космических аппаратов и кораблей без затраты топлива (изменять скорость на 1-2 км, получать любое новое направление и энергию от астероида, метеорита или любого космического тела); ... Это увеличивает возможности космического аппарата в десятки миллионов раз. Все может быть сделано на основе существующих технологий. Все эти реальные революционные концепции нуждаются в развитии, в начальной поддержке при небольшом финансировании. Г-н Казанова не понимает их и предъявляет бессмысленные требования: «Сначала решите все проблемы, которые могут быть» ...

Более того, авторы проектов переслали мне дебрифинги, показывающий полную некомпетентность др. Казанова в физике, двигательных системах, космонавтике. Если

потребуется, мы готовы представить его ответы и готовы к открытой дискуссии о новых инновационных идеях, которые могут сделать революцию в космических запусках, в освоении космоса и планет солнечной системы. Повторим: м-р Казанова не может определять ценность научных идей, если за всю свою жизнь он написал лишь ОДНУ незначительную научную работу очень давно (в 1976 г!), в соавторстве с тремя авторами.

Отметим, что по Уставу м-р Казанова должен помогать авторам проектов и специалистам, работающим в малых научных организациях. Но 90% денег он отдает людям, работающим в университетах и огромных компаниях, таких как Боинг и Сикорский, которые и без НИАС имеют прямые контракты с НАСА.

Итог анализа: НИАС потратил более 30 миллионов долларов за 7 лет, но не предложил ни одной новой концепции или идеи! НИАС превратился в кормушку для друзей, начальства и тратит 90% государственных (налогоплательщиков!) денег неэффективно, не брезгуя *криминальными методами (это – коррупционная деятельность)*.

Для примера посмотрим типовой Итоговый Отчет научного проекта («исследования») м-ра Эдвардса (Edwards) по давней (1894 г.) идее **К.Э.Циолковского «Космический Элеватор»**, **детально разработанной советским ученым Ю.Н.Арцутановым в 1957г** http://www.niac.usra.edu/files/studies/final_report/521Edwards.pdf

На 42 стр. «брошюрки для школьников начальных классов» нет ни одной формулы, ни одного нового результата или хотя бы простенькой идеи. И, разумеется, никаких ссылок на других авторов (хотя научных работ на эту тему за последние 10 лет – сотни). М-р Казанова предоставляет на этот «ценнейший труд» почти миллион долларов (645 тысяч долларов м-ру Эдвардсу плюс 30% Казанове, самому себе «за труд по отбору»). Получается по \$20 тысяч за каждую страничку пустого текста – «болтовни», изданного как популярная брошюрка для широкой публики; более того, они получили еще и гонорар... Добавим, здесь – 42 страницы, а есть и «научные» отчеты в 3 – 12 страничек, за которые Казанова предоставил финансирование в сотни тысяч долларов.

5. Попустительство НАСА

Имеется сайт НАСА. Здесь много саморекламы. Однако нет элементарной общепринятой информации об электронных адресах руководителей НАСА (или их секретарей), начальников отделов, наблюдательного Комитета, Научного Совета.

На неоднократные запросы в НАСА – кто у них отвечает за работу НИАС; какие электронные адреса помощника Администратора НАСА, начальника отдела технологий, председателя наблюдательного Совета, ... Ответов нет.

Открытая печать уже неоднократно писала о криминальной деятельности г-на Казанова, многие ученые жаловались в НАСА на деятельность НИАС. Однако, например, в 2004 году НАСА вновь заключает контракт с НИАС и дает г-ну Казанова десятки миллионов долларов!! В свое время была запрошена у НАСА копия соглашения НАСА-НИАС относительно финансирования НИАС руководством НАСА. НАСА обязано свободно давать эту информацию всем и бесплатно согласно Акту о свободе информации, если информация менее 100 страниц и ее копирование требует менее 2-х часов. НАСА не могло найти его полгода. После многих напоминаний они прислали 10 страниц Соглашения + 60 страниц макулатуры и потребовали в оплату 100 долларов...

Я понял, если у НАСА нет «желания» проверять, куда идут десятки миллионов долларов, которые они должны выделять на прогресс науки, то г-н Казанова может «безбоязненно» тратить их на себя, своих друзей, начальство и полезных ему людей, на ничемные исследования.

Я работал в НАСА и знаю, что большинство специалистов НАСА имеют низкую квалификацию. Это – результат кадровой политики руководства НАСА. Даже широко известные ученые могут получить работу в НАСА только по протекции, «по благу». Я знаю доктора наук, который имеет большой опыт, большое число научных работ, изобретений, который подал заявление на вакантную должность менеджера проекта. Но НАСА взяла «по благу» на эту должность обычного пилота, который не имеет ни научных знаний, ни научного опыта, никаких научных трудов и изобретений. Когда же ученый подал на должность ученого-инженера, они взяли на это место молодого выпускника техникума, бакалавра (B.D.), не имеющего ни опыта, ни научных работ, ни изобретений.

Я не хочу сказать, что все ученые, принимавшие участие в исследованиях НИАС, – халтурщики. **Но сама система финансирования, реализуемая мистером Казанова, побуждает их к безответственности...**

USRA (NIAC) – это организация университетов и она, естественно, старается давать гранты университетам. Но отметим, в университетах профессора, преподаватели сидят на зарплате. Они загружены лекциями, занятиями со студентами, проверкой контрольных и студенческих работ, экзаменами и зачетами. Научная работа им дается в дополнительную нагрузку. Естественно нет ни времени, ни сил, ни возможности ее выполнять, тем более, что *настоящее исследование требует гигантского времени и напряжения*. Если в первую стадию по проекту (\$100,000) они еще вкладывают какие-то усилия (в надежде получить финансирование на вторую стадию \$1 миллион долларов), то во второй стадии вообще не имеют стимулов и выдают откровенную «мишуру». Да и большинство университетов не имеют и не в состоянии иметь дорогостоящее оборудование, необходимое для большинства современных исследований.

Добавим по нашему опыту: ранее, в СССР, в приказном порядке пытались заставить вузовских ученых бесплатно вести научную работу. Каждый ученый ВУЗа обязан был делать, так называемую, госбюджетную тему. И вот мне в руки попал полстраничный отчет трех кандидатов наук кафедры Вычислительной техники Восточно-Сибирского Технологического Института, где они сообщали, что компьютер – вещь полезная...

Если грант дается университету, то полезной отдачи от ученых можно ожидать только в одном случае, – если их полностью освободить от педагогической деятельности. Научный Совет Америки (NRC) так и делает – переводит университетских докторов наук на 1 – 2 года в государственные научные, хорошо оборудованные лаборатории. Расход средств в 10 раз меньше, а отдача в сотни раз выше.

Видимо, поэтому НАСА имеет много неудач и катастроф, многие из которых скрываются. Например, случай, когда «ученый» НАСА ввел в компьютер для коррекции траектории фунты и футы вместо стандартных международных единиц (килограммов и метров), и 250 миллионный корабль вместо Марса полетел неизвестно куда. Или – ужасная трагедия с «Шаттлом». Куски теплоизоляции неоднократно отлетали от ракеты и в предыдущие запуски. Но *на это не обращали внимания*, пока они не ударили в крыло, что и привело к гибели 7-и человек. Разумеется. Вы можете сказать: «НАСА имеет и успехи». Но надо быть объективным: не забывайте, что **НАСА имеет гигантское финансирование вот уже более полувека. Это финансирование**

(15-20 миллиардов долларов ежегодно) в несколько раз больше, чем космическое финансирование всех остальных стран мира вместе взятых (например, НАСА получает в 30 раз больше денег, чем Космическое Агентство другой космической держава мира – России). Если Вы сделаете оценку обычным критерием: эффективность (польза) деленная на затраченные средства (деньги), то увидите, что коэффициент эффективности НАСА - наихудший в мире.

Пример: во время Большой Космической Гонки бывший Советский Союз тратил на космос в 3 – 5 раз меньше денег, чем НАСА, однако оставался лидером многие годы, пока американцы не высадились на Луне. После этого Советская космическая индустрия также имела в 5-7 раз меньшее финансирование, чем НАСА, но вплоть до развала СССР (1989 г) Советский Союз запускал в 2-3 раза больше спутников, чем все остальные страны вместе взятые, включая США.

Как тут не вспомнить слова из известной песни Высоцкого: «Я бы и в Москве уран нашел при такой повышенной зарплате».

Г-н Син О'Киф (Sean O'Keefe) назначен был в Администраторы НАСА (в свое время он подвергал жесткой критике деятельность НАСА). Однако он так и не сумел улучшить работу НАСА.

6. Попустительство Ассоциации Университетов для Космических Исследований (USRA) и ее Президента

НИАС формально является отделом УСРА (USRA), Президент УСРА – г-н Блэк (Black). Может быть президент УСРА г-н Блэк не знает о криминальной (коррупционной) деятельности г-на Казанова в НИАС? Ему многократно жаловались на м-ра Казанова, но ответа нет. Электронные адреса исполнительного директора Каммингса (D. Cummings) и шеф-инженера Пича (L. Peach), указанные на сайте УСРА, оказались недействующими. М-р Блэк не отвечает на простые вопросы:

Как, кто руководит г-ном Казанова? Кто назначает председателя и членов Научного Совета и отборочного Комитета? Где и когда был, будет обзор (и отбор) в НИАС предложений и рецензий на них? И т.п.

Реальность проста: Г-н Казанова содержит на эти миллионы (деньги налогоплательщиков) как директора, так и верхушку УСРА. *Так что, полная коррупция?*

7. Заключение

В свое время неоднократно нами были представлены для НАСА (и УСРА) детальные планы, как улучшить работу НИАС, сделать ее более открытой, полезной (для общества, разумеется!). Изменить ситуацию, когда один человек (г-н Казанова) **бесконтрольно** распределяет десятки миллионов долларов налогоплательщиков (ради своих фаворитов). Но все руководство мгновенно становится «глухонемым», в ответ не поступают даже простые сообщения о получении этих наших писем и предложений.

В такой ситуации у НАСА может быть только одно решение: немедленно остановить бесполезное и неэффективное финансирование НИАС и передать эти деньги независимой организации для открытого конкурса и финансирования действительно перспективных и революционных инноваций. Можно передать функции выбора предложений для исследований Американской Академии Наук, Американскому Национальному Научному Фонду (как это сделано для всех остальных наук!) или, например, Международному Космическому Агентству (<http://www.international-space->

agency.org), которое продвигает реальные перспективные исследования пусковых, двигательных, силовых, орбитальных систем, не делая из них никаких секретов. Миллионы денег налогоплательщиков будут контролироваться научным обществом и будут лучше использованы.

Видимо, следует изменить и саму форму поддержки ученых, работающих в области космоса. Деньги следует давать *не авансом* под обещание сделать ту или иную работу, а как премию за проделанные лучшие работы, исследования, действительно перспективные идеи, отобранные на открытом конкурсе Научным Советом из авторитетных ученых в данной области. В самом деле, никому еще не выдали Нобелевскую премию авансом под обещание сделать революционное открытие. А м-р Казанова единолично тратит *баснословные суммы (больше Нобелевских премий)* авансом своим приятелям за «обещания и последующую пустую болтовню» и «перепев» старых или ошибочных или абсолютно бесперспективных идей вроде «халата хамелеона» или «обрызгивания космонавта» вместо выдачи ему космического скафандра.

Президент и Конгресс США должны, не откладывая, расследовать ситуацию в НАСА-НИАС и привлечь г-на Казанова и руководство УСРА **к ответственности за злоупотребления, дискриминацию, коррупцию**. Мы готовы представить специальной Комиссии материалы, подтверждающие все сказанное в данной статье.

P.S. В интернете, прессе часто приходится читать восторги по поводу НИАС-овских бумажных «исследований» (биокостюм, ракета на анти-материи, магнитный парус и прочая бессмыслица), на которые мистер Казанова тратит миллионы долларов налогоплательщиков. Я призываю как специалистов, так и любителей внимательно проанализировать эти «инновации» и сообщать результаты на специальный сайт <http://nasa-niac.narod.ru> . E-mail: nasa-niac@yandex.ru

Не пугайтесь научных выкладок. В большинстве отчетов их нет. Повторяем: как написал один школьник о финальном отчете НИАС по Космическому Элеватору – это *школьный реферат по интернетовским материалам*, каждая страница которого оплачена 20 тысячами долларов налогоплательщиков.

Для эпилога – Награждение по-американски

Более ста лет назад К.Э. Циолковский предложил идею Космического Лифта. Полстолетия назад другой русский ученый Ю.Н. Арцутанов дал техническую проработку этой идеи. А награду за все это как за инновационный проект (около миллиона долларов) получил в наше время американец Брэдли Эдвардс (Bradley Edwards).

В 1894г Константин Эдуардович Циолковский в своей знаменитой работе «Грезы о Земле и Небе» предложил космический лифт, а русский ученый Ю.Н. Арцутанов в 1960 г в приложении к газете «Комсомольская Правда» от 31 июля 1960 г. (изданной в 2 миллионах экземпляров), представил свою техническую проработку этой идеи.

Космический лифт представляет собой космическую станцию, расположенную на или выше геостационарной орбиты и соединенную с Землей прочным тросом. Система спроектирована так, что ее центр тяжести находится выше геосинхронной орбиты, а поскольку Земля вращается, то центробежная сила натягивает трос, как камень натягивает прашу, когда ребенок ее вращает рукой. Арцутанов, в частности, предложил электропоезд, который будет ездить по тросу к станции и возить туристов.

И вдруг ученый мир с удивлением узнал, что *награду за эту «инновационную идею» получает американец м-р Брэдли Эдвардс (Bradley Edwards).*

Естественно, ученые заинтересовались как это произошло и что выдающегося внес **лично** г-н Брэдли в данную проблему. Прежде всего, выясняется, что решение о награде единолично и *скрытно* от общественности вынес директор Института Перспективных Исследований (NIAC) м-р **Казанова (R. Cassanova)**, известный в научных кругах своими коррупционными действиями с миллионами денег налогоплательщиков, выделяемых государством на космические исследования (статья в международной газете «Правда» http://english.pravda.ru/mailbox/22/98/396/14417_NASA.html и перевод этой статьи на русский в <http://NASA-NIAC.narod.ru>).

Награда оказалась не маленькой – около миллиона долларов: 645 тысяч долларов м-ру Эдвардсу плюс 30% м-ру Казанова («за труд по отбору проектов»).

Но, может быть, м-р Казанова и его партнер провели выдающееся исследование, которое превзошло сотни других блестящих научных работ ученых в аэрокосмической области? Мы уже отмечали здесь в статье, «доктор наук» Казанова за свою научную жизнь написал только ОДНУ незначительную статью в соавторстве еще с тремя учеными. Странно, что в НАСА не нашли более квалифицированного «ученого», который сейчас *усердно раздает миллионы долларов своим протеже.*

Но может быть отчет, который г-н Казанова наградил, содержит гениальные разработки и результаты? Можно посмотреть его на сайте NIAC:

http://www.niac.usra.edu/files/studies/final_report/521Edwards.pdf

Никаких научных выкладок, уравнений там нет. Это – 42 страницы, наполовину занятые цветными рисунками, где содержится только популярное изложение известной идеи космического элеватора (как для учеников начальной школы). Более того, ни один из ученых, работавших над этой проблемой (даже американских!) не упомянут. Все изобрели Эдвардс и Казанова. Это – крайне неприлично по отношению к К.Э.Циолковскому и к сотням ученых (с учетом того, что м-р Казанова подобным образом награждает и других своих протеже по 2-4 раза за одно и то же), которые оказались *без всякой поддержки* своих действительно полезных исследований и инноваций.

Резюме. Как окончательный результата всего анализа, проведенного в этой статье, по проблеме организации науки следует принять тезис Нобелевского Лауреата Альберта Эйнштейна, который представлен в начале этой статьи, как полезную рекомендацию к действию, необходимую для эволюции Человеческого Общества.

Александр Александрович Болонкин, инженер-механик, д.т.н.; с.н.с. Национального Научного Совета США в Научной Лаборатории Базы Военно-Воздушных Сил США (Eglin, USA); бывший старший научный сотрудник НАСА; окончил Казанский Авиационный Институт (1958); работал ведущим инженером в авиационном ОКБ О.Антонова (Киев), начальником отдела надежности в ОКБ Глушко (Химки, МО); преподавал в МАИ, МАТИ, МВТУ им. Баумана, в американских университетах.

А.Болонкин – член Совета Директоров Международного Космического Агентства; область его научных интересов – проблемы авиационных и аэрокосмических систем, ракетные двигатели для стратегических ракет; девять его докладов по безракетным методам полета в космос были в свое время одобрены Всемирным Конгрессом по аэрокосмическим проблемам (2002, США).

О 70-летию исследований по механике в Казанском Научном Центре РАН

М.А. Ильгамов

Институт механики УНЦ РАН
Россия, 420111, Казань, ул. Лобачевского, д. 2/31

Аннотация. Дан краткий обзор о Казанском филиале Академии наук СССР – Казанском научном центре РАН. Перечислены основные научные направления по механике, разрабатывавшиеся в Казанском физико-техническом институте и Институте механики и машиностроения КазНЦ РАН, упомянуты некоторые результаты.

Ключевые слова: Казанский филиал Академии наук СССР, Казанский научный центр РАН, Институт механики и машиностроения КазНЦ РАН, исследования по механике.

Прошло более 70 лет со времени организации Казанского филиала Академии наук СССР (КФАН СССР). В апреле 1945 года Советом Народных Комиссаров СССР было принято постановление об открытии КФАН СССР. В том же году Президиум АН СССР утвердил его структуру. Были организованы институты: Физико-технический, Химический, Геологический, Биологии, Языка, литературы и истории, Отдел водохозяйственных проблем и энергетики. Председателем Президиума КФАН СССР стал акад. А.Е. Арбузов, его заместителем – проф. Л.М. Миропольский, ученым секретарем – проф. Г.Х. Камай.

Филиал начал работу с 1 января 1946 года. В становлении и развитии КФАН СССР большую помощь оказывали партийные и советские органы республики, ведущие вузы Казани, особенно Казанский университет [1].

В деятельности КФАН СССР – КазНЦ РАН (КазНЦ РАН – Казанский Научный Центр РАН) выделяют *четыре периода*. *Первые 17 лет* (1946-62 годы) были отмечены плодотворной слаженной работой сравнительно небольшого научного коллектива под руководством выдающегося советского ученого А.Е. Арбузова. Библиотека, мастерские, бухгалтерия и все службы были общими для всего филиала.

Второй период – это 10 лет самостоятельной работы институтов, вызванной реформированием АН СССР. КФАН СССР был упразднен.

Третий период начинается с 1973 года, когда деятельность Казанского филиала была возобновлена. В состав его вошли институты: Физико-технический, Органической и физической химии, Биологии, Языка, литературы и истории. Позднее был организован Казанский филиал Института проблем информатики АН СССР. В них сохранились самостоятельные материально-финансовые и хозяйственные отделы.

Четвертый период начинается с 1991 года и продолжается по настоящее время. В 1991 году были созданы Отдел Энергетики и Институт механики и машиностроения (ИММ) Казанского научного центра АН СССР, а Институт языка, литературы и истории, а также Филиал Института проблем информатики АН СССР были переведены в Академию наук Республики Татарстан.

Приведенная периодизация является, естественно, весьма условной. Тем не менее в дальнейшем будем придерживаться ее, причем не будем касаться четвертого периода и последних 25 лет развития механики.

Целью настоящей работы является выделение основных научных направлений по механике, разрабатываемых в Физико-техническом институте и Институте механики и машиностроения и некоторых результатов.

За 70 лет развивались многие направления механики, получены значительные результаты. Однако не все они упоминаются в данной работе, что не означает их меньшей ценности для своих направлений исследований. К сожалению, мало сведений приводится по подземной гидромеханике и моделированию процессов нефтедобычи. Эти данные будут приведены в последующем научном обзоре.

КФТИ (Казанский физико-технический институт) КФАН СССР «начинался с двух комнат» в геометрическом корпусе Казанского университета с несколькими столами и стульями. Позже для филиала было предоставлено здание по ул. Лобачевского, 2/31. Но и оно долго ремонтировалось, возводился пристрой. В период прохождения документов по организации КФАН СССР директором КФТИ был член-корреспондент АН СССР Н.Г. Чеботарев, но с самого начала его функционирования директором являлся профессор Х.М. Муштари.

Сектор механики под руководством Х.М. Муштари работал по развитию нелинейной теории тонких оболочек, вариационных и численных методов анализа деформации, устойчивости. Первыми сотрудниками были К.З. Галимов, С.Г. Винокуров, Р.Г. Суркин, Г.Г. Тумашев.

Сектором математики вначале руководил Н.Г. Чеботарев, а с 1947 г. – Г.С. Салехов. Первыми сотрудниками были А.П. Норден, А.Н. Хованский, Б.А. Сейфуллина, В.Л. Данилов, В.Д. Чугунов, С.Ф. Коротков, Г.П. Цыбульский.

С 1951 года сектор математики начинает исследования по теории фильтрации природных жидкостей в пористых средах. Своими исследованиями в области применения математических методов к вопросам рациональной разработки нефтяных месторождений получила известность школа профессора Г.С. Салехова. Он предложил методы решения задачи управления контуром нефтеносности. В работах сотрудников были даны эффективные методы её решения с учетом основных физико-геологических, технологических и экономических факторов. Развитием работ Г.С. Салехова является постановка и разработка методов решения различных задач оптимизации в нефтедобыче.

С 1947 года начинаются первые публикации работ по теории оболочек, выполненных в секторе механики, в таких авторитетных журналах как «Прикладная механика и математика», «Доклады Академии наук СССР» и других научных изданиях. Эти исследования относятся главным образом к общим вопросам теории оболочек и задачам их устойчивости [2, 3]. Появляются совместные работы Х.М. Муштари с С.Г. Винокуровым, Р.Г. Суркиным. Например, в журнале «Прикладная математика и механика» за 1950 год опубликована одна из важнейших работ: нелинейная теория устойчивости сферической оболочки. Спустя двадцать лет после появления этой работы академик В.В. Новожилов писал [4]: «Данную задачу атаквали многие авторы, но решение, по-видимому, ближе всего подтвержденное опытами, принадлежит Х.М. Муштари. При оценке устойчивости сферических корпусов глубоководных аппаратов американцы пользуются именно формулой Муштари, соглашаясь тем самым, что она должна быть признана наиболее достоверной».

В этой работе Х.М. Муштари дает критический анализ исследований Т. Кармана и Цзяня, указывает на две ошибки, допущенные ими. Указано также на ошибку в работе К. Фридрихса. В другой работе Х.М. Муштари дал критический анализ работы В.И. Феодосьева (1954). Об этом другой корифей теории оболочек Э.И. Григolloк отмечает [3]: «Как выдающийся механик он (Х.М. Муштари) не мог не заметить, что контур

раздела зон деформирования также определяется параметрами процесса деформирования, поэтому варьировать нужно не только прогиб оболочки в центре, но и радиус контура сопряжения зон».

Отметим, что указанные две работы Х.М. Муштари привели в дальнейшем к сотрудничеству с великим американско-китайским ученым Цянем и многолетнему расхождению с выводами выдающегося механика В.И. Феодосьева.

К этому периоду относятся также исследования в области напряжений в пограничной зоне, в области устойчивости оболочек с начальными отклонениями от заданной формы, в области обратных задач теории оболочек, плит и оболочек средней толщины и т.д.

К.З. Галимов получил уравнения равновесия, относящиеся к деформированному и недеформированному состояниям оболочки, доказал вариационные теоремы смешанного типа.

Крупными вехами являются большие циклы исследований в секторе механики в связи с созданием аэрокосмической техники в фирме, руководимой С.П. Королевым. С этими исследованиями были связаны: закрытое издание Трудов Физико-технического института КФАН СССР в 1954 году; издание монографии [5]; проведение Всесоюзной конференции по теории пластин и оболочек (1960 г.), защиты диссертаций.

К концу этого периода относится установка вычислительной машины, с чем было связано начало систематических и крупных работ по исследованию напряженно-деформированного состояния пластин различных форм М.С. Корнишиным и Ф.С. Исанбаевой. До этого М.С. Корнишин провел большое исследование нелинейных задач изгиба и устойчивости пластин и панелей методом Бубнова-Галеркина. В то время нелинейные задачи теории оболочек решались главным образом этим методом.

Проблемы применения метода Бубнова-Галеркина к нелинейному поведению оболочек были предметом анализа И.В. Свирского. Оригинальные результаты он получил также в применении метода последовательных приближений. Различные задачи о напряженно-деформированном состоянии и устойчивости оболочек и пластин рассмотрели Н.С. Ганиев и А.В. Саченков. И.Г. Терегулов изучил поведение оболочек с учетом ползучести материала. Напряженно-деформированное состояние оболочек вращения было предметом исследований М.С. Ганеевой. Исследования Н.К. Галимова связаны с теорией трехслойных пластин. Б.М. Зуев и его сотрудники занимались экспериментальными исследованиями методом тензометрирования и оптическим методом. Позже эта группа была переведена в Институт органической и физической химии, где занималась испытаниями материалов.

Во втором периоде (1963-1972) продолжали развиваться прежние направления. Особо глубокое развитие получили численное моделирование, теория трехслойных пластин, а также начались исследования по теории взаимодействия пластин и оболочек с жидкостью, газом и твердым деформируемым телом. К этому периоду относится начало, например, исследований периодических ударных волн в трубе, их взаимодействие с тонкостенными цилиндрической и конической оболочками. Один из самых ярких результатов в этом направлении – экспериментальное обнаружение явления катастрофического разрушения (в течение порядка 1 с) стальной оболочки, внутри которой возбуждаются воздушные волны (Г.В. Садыков). Теорию этого явления не удалось создать на базе аналитических методов. Но когда развились численные методы, к этой проблеме, к сожалению, не вернулись. Конечно, дело не только в этом; само явление – очень сложное.

Было образовано третье подразделение по механике – Лаборатория теории оболочек, взаимодействующих со средой (1971 г.). Если обе первые лаборатории были созданы

привлеченными для этого выдающимися учеными Х.М. Муштари и Н.Г. Чеботаревым, сделавшими к 1946 году большой вклад в своих областях науки, то третье подразделение возникло через 25 лет в недрах существующей лаборатории механики. С сожалением можно отметить, что не было создано еще одно подразделение – по подземной механике в то благоприятное время (подразделения создавались в КФТИ только по физическим направлениям).

В связи с переездом в новое здание появилась возможность усилить вычислительные средства, создать экспериментальную базу.

Так же как в предыдущие периоды очень активно работал семинар, где каждый четверг докладывались и законченные работы, и ход выполнения исследований в порядке их обсуждения, проблемные вопросы. Пожалуй, ни одна рукопись статьи не отправлялась в журнал без обсуждения на семинаре.

Начиная с 1969 года издавались «Труды семинара по теории оболочек», редакторами которых были М.С. Корнишин и М.А. Ильгамов. Они выходили почти ежегодно. Всего было опубликовано 27 выпусков (последний – в 1992 году). Этим самым ускорялась публикация результатов исследований. В условиях малочисленности центральных журналов это было возможностью публиковать материал. Недостатком была малоизвестность этих публикаций в стране и за рубежом.

Третий период (1973 – 1990 гг.) характеризуется сильным развитием численного моделирования процессов деформирования тонкостенных конструкций и разработки нефтяных месторождений. Это и классические задачи, которые были в поле зрения сотрудников в прежние годы, и новые задачи.

К последним относятся, например, статические и динамические задачи совместной деформации оболочки и упругого сплошного заполнителя (модель твердотопливной ракеты), взаимодействия оболочек с жидкостью и газом (модель камер сгорания реакторов, трубопроводов), а также взаимодействие мягких оболочек с потоком газа (парашютных систем).

Р.Г. Суркиным с группой была построена установка для ударного нагружения сферического сегмента. Результаты по устойчивости сегмента вошли в энциклопедическую монографию [6]. Приятно отметить, что начатое в то время изучение тонких пластин (предохранительных мембран) получило в институте дальнейшее развитие (Н.М. Якупов, Н.К. Галимов).

Установка для создания периодических ударных волн была создана Р.Г. Зариповым, а для изучения гидродинамической штамповки тонкостенных изделий – А.А. Павловым. Развивалась соответствующая теория. В этом принимали участие М.М. Шакирьянов, Н.Н. Шихранов. В частности, обнаруженное явление пузырьковой кавитации жидкости при штамповке легло в основу исследований Ш.У. Галиева в Киеве.

М.Б. Гафуровым был построен испытательный бассейн, Г.В. Садыковым, Р.И. Давыдовым с сотрудниками – аэродинамическая труба. На них были получены новые результаты, по которым развивалась теория. Задачами обтекания деформируемых элементов конструкций занимались Б.В. Гулин, Ж.М. Сахабутдинов, А.Н. Гильманов, А.А. Аганин, Р.Р. Шагидуллин, М.З. Сабитов, А.А. Тукмаков. М.В. Талдыкиным были созданы модели волновых движителей и насосов, теорией которых занимались В.Л. Федяев, М.М. Сулейманова. Эти работы проводились в контакте с ленинградскими морскими организациями и получили высокую оценку со стороны академика М.А. Лаврентьева. Задачами биомеханики занимался Р.Н. Мифтахов.

Во всех этих исследованиях мы двигались слишком быстро, и у нас не хватало времени на хорошее оформление результатов и на публикацию их в центральных и

международных журналах. Поэтому многие прекрасные результаты остались малоизвестными.

Надо отметить, что руководство Казанского филиала АН СССР, особенно КФТИ, не сумело встроиться в проблемы КАМАЗа, крупнейшего объекта машиностроения семидесятых годов, в его технические проблемы и на этой основе получить хорошее развитие. С большим опозданием мы создали в КФТИ лабораторию по проблемам КАМАЗа за счет своих штатных единиц. Но время было упущено, не получилось должного развития за счет предприятия. С учетом этой упущенной возможности мы предприняли активные действия в связи с началом строительства Елабужского тракторного завода в начале восьмидесятых годов. Удалось добиться договоренности о передаче фонда заработной платы и штатной численности из Министерства тракторного и сельскохозяйственного машиностроения СССР в АН СССР с целью организации в КФТИ двух лабораторий (заведующие – Ж.М. Сахабутдинов, В.Л. Федяев). Еще до этого был создан Отдел прикладной математики КФТИ и Казанского педагогического института, где сотрудники работали на хоздоговорные средства, которых у нас тогда было много (заведующий отделом – М.А. Ильгамов). Такая организационная форма была удачной, так как в вузах хозрасчетные деньги частично направлялись на заработную плату, а в институтах АН СССР можно было тратить такие деньги только на оборудование и командировки.

Позже была создана лаборатория по проблемам предприятия, разрабатывающего твердотопливные ракетные двигатели (заведующий – А.Н. Гильманов). Численное моделирование неустойчивого режима работы изделия нашло непосредственное применение на предприятии.

В рамках программы АН СССР и Министерства образования СССР по компьютеризации школьного образования некоторое время функционировало подразделение (заведующий – В.Б. Кузнецов), в штате которого находились некоторые наши сотрудники.

Все эти мероприятия по расширению тематики в актуальных для своего времени направлениях происходили на фоне ухудшения положения в стране в восьмидесятые годы, начавшейся с его второй половины деиндустриализации.

В период разгула псевдодемократии подразделения механики во многом потеряли в институте свои позиции, особенно из-за позиции нового руководства КФТИ. В частности, несколько штатных единиц было передано в физические лаборатории, мы потеряли производственные площади.

Выше сказано о первой Всесоюзной конференции по теории оболочек в 1960 году, приуроченной к 60-летию Х.М. Муштари. Такие конференции проходили в Казани в 1970, 1980, 1995, 2000 годах. Последняя была посвящена 100-летию Х.М. Муштари, 90-летию профессора К.З. Галимова, 80-летию профессора М.С. Корнишина. Можно отметить конференцию по краевым задачам механики сплошных сред, посвященную 100-летию Г.Г. Тумашева и 110-летию Х.М. Муштари (2010 г.). Наши механики были организаторами и других конференций и летних школ. Последние были организованы в лагере «Волга» в 1975 и 1986 годах по проблемам взаимодействия оболочек с жидкостью, газом и твердым сплошным телом. На высоком уровне и с большим количеством участников прошла конференция, посвященная 95-летию М.А. Лаврентьева (1995 г.). Сказанным не исчерпывается перечень конференций, проведенных нами за 70 лет. Но, конечно, совершенно выдающееся событие – это проведение Съезда по теоретической и прикладной механике, впервые в Казани в 2015 году. Особо нужно отметить большую работу сотрудников ИММ КазНЦ РАН во главе с Д.А. Губайдуллиным.

В приложении приведен общий список монографий, изданных механиками КФТИ и ИММ.

Список аббревиатур

КФАН СССР – Казанский филиал Академии наук СССР

ИММ – Институт механики и машиностроения

КФТИ – Казанский физико-технический институт

КазНЦ РАН – Казанский научный центр Российской Академии наук

Литература

1. Казанский филиал Академии наук СССР. Буклет. Казань. Изд-во Тат.ОК КПСС. 1987. – 20 с. (составитель М.А. Ильгамов).
2. *Х.М. Муштари*. Работы казанских ученых по нелинейной теории оболочек // Изв. КФАН СССР. Сер. физ.-мат. и техн. наук. 1958. № 12. С. 5-15.
3. *М.А. Ильгамов*. Профессор *Х.М. Муштари*. М.: Наука. Физматлит. 2001. 192 с.
4. *В.В. Новожилов*. Краткий очерк развития теории оболочек в СССР // исследования по теории пластин и оболочек. Вып. VI-VII. Казань: изд-во КГУ. 1970. С. 3-22.
5. *Х.М. Муштари, К.З. Галимов*. Нелинейная теория упругих оболочек. Казань: Таткнигоиздат. 1957. 431 с.
6. *А.С. Вольмир*. Устойчивость деформируемых систем. М.: Наука, 1967. 984 с.

Приложение

Монографии сотрудников-механиков КФТИ КФАН СССР – ИММ КазНЦ РАН

1. Г.С.Салехов. Вычисление рядов. М.: Гостехиздат, 1955. 144 с.
2. *Х.М.Муштари, К.З.Галимов*. Нелинейная теория упругих оболочек. Казань: Таткнигоиздат, 1957. 432 с. (англ. пер.: 1961).
3. Математический анализ. Функции, пределы, ряды, цепные дроби / В.Л.Данилов, А.Н.Иванова, Е.К.Исакова, Л.А.Люстерник, Г.С.Салехов, А.Н.Хованский, Л.Я.Цлаф, А.Р. Янпольский. М.: Физматгиз, 1961. 439 с.
4. М.С.Корнишин. Нелинейные задачи теории пластин и пологих оболочек и методы их решения. М.: Наука, 1964. 192 с.
5. М.С.Корнишин, Ф.С.Исанбаева. Гибкие пластины и панели. М.: Наука, 1968. 260 с.
6. И.В.Свирский. Методы Бубнова-Галеркина и последовательных приближений и их применение к расчету оболочек. М.: Наука, 1968. 200 с.
7. И.Г.Терегулов. Изгиб и устойчивость тонких пластин и оболочек при ползучести. М.: Наука, 1969. 206 с.
8. М.А.Ильгамов. Колебания упругих оболочек, содержащих жидкость и газ. М.: Наука, 1969. 182 с.
9. В.В.Скворцов. Математический эксперимент в теории разработки нефтяных месторождений. М.: Наука, 1970. 224 с.
10. Г.С.Салехов, Л.М.Муратов : Поспеев В.Е. Вычисление рядов и несобственных интегралов. Казань: Изд-во Казан.ун-та, 1973. 169 с.
11. В.В.Скворцов. Методы экспертных оценок и их приложение в задачах теории фильтрации. Казань: Тат. книж. изд-во, 1976. 216 с.
12. М.А.Ильгамов, В.А.Иванов, Б.В.Гулин. Прочность, устойчивость и динамика оболочек с упругим наполнителем. М.: Наука, 1977. 332 с. (англ. пер.: 1981).
13. У.Н.Закиров. Механика релятивистских космических полетов. М.: Наука, 1984. 152 с.
14. М.А.Ильгамов, В.А.Иванов, Б.В.Гулин. Расчет оболочек с упругим наполнителем. М.: Наука, 1987. 264 с.
15. E.H.Dowell, M.A.Ilgamov. Studies in Nonlinear Aeroelasticity. New York - London - Tokyo, Springer-Verlag, 1988. 456 p.
16. М.С.Корнишин, В.Н.Паймушин, В.Ф.Снигирев. Вычислительная геометрия в задачах механики оболочек. М.: Наука, 1989. 208 с.

17. А.И.Голованов, М.С.Корнишин. Введение в метод конечных элементов статики тонких оболочек. Казань: 1990. 271 с.
18. Х.М.Муштари. Нелинейная теория оболочек. М.: Наука, 1990. 223 с.
19. В.В.Ридель, Б.В.Гулин. Динамика мягких оболочек. М.: Наука, 1990. 205 с.
20. М.А.Ильгамов. Введение в нелинейную гидроупругость. М.: Физматлит, 1991. 200 с. (ит. пер.: 1993).
21. М.С.Ганеева. Прочность и устойчивость оболочек вращения. М.: Наука, 1992. 161 с.
22. Н.М.Якупов. Прикладные задачи механики тонкостенных конструкций. Казань: ИММ КазНЦ РАН, 1994. 124 с.
23. Н.М.Якупов, М.Н.Серазутдинов. Расчет упругих тонкостенных конструкций сложной геометрии. Казань: ИММ КазНЦ РАН, 1993. 208 с.
24. М.А.Ильгамов. Статические задачи гидроупругости. Казань: ИММ КазНЦ РАН, 1994. 208 с.
25. Ж.М.Сахабутдинов. Анализ дискретных моделей движения точки. Казань: ИММ КазНЦ РАН, 1995. 196 с.
26. Обзоры исследований по механике сплошной среды. К 50-летию КНЦ РАН. Казань: ИММ КазНЦ РАН, 1995. 216 с.
27. А.Б.Мазо. Математическое моделирование процессов горячей обработки металлов. Казань: Казанский фонд "Математика". 1996. 209 с.
28. Н.М.Якупов, И.Н.Гатауллин, Р.Н.Хисматуллин. Обследование, анализ и прогнозирование долговечности строительных конструкций и рекомендации по их восстановлению. Методическое руководство. Казань: ИММ КНЦ РАН, 1996. 208 с.
29. Безопасность производственных объектов. Под ред. Н.М.Якупова. Казань: ИММ КНЦ РАН, МЧС РТ, 1996.
30. Ф.А.Сальянов. Основы физики низкотемпературной плазмы, плазменных аппаратов и технологий. М.: Наука, 1997. 240 с.
31. Vector Lyapunov Functions in Stability Theory/R.Z.Abdullin, L.J.Anapolski, R.I.Kozlov, A.I.Malikov, V.M.Matrosov, A.A.Voronov, A.S.Zemljakov. Advanced series in mathematical science and engineering. World Federation Publisher Company. 1996. 394 p.
32. Н.М.Якупов, Н.И.Хисматуллин. Механика Нижнекамскнефтехима. Казань: ИММ КНЦ РАН, 1996. 224 с.
33. Д.А.Губайдуллин. Динамика двухфазных парогазокапельных сред. Казань: Изд-во Казан. математического общества, 1998. 154 с.
34. М.А.Ильгамов. Static Problems of Hydroelasticity. Moscow: Nauka. Fizmatlit, 1998. 208 p.
35. Р.Г.Зарипов. Информация различия и переходы беспорядок - порядок. Казань: Изд-во КГТУ им. А.Н.Туполева, 1999. 155 с.
36. У.Н.Закиров. Есть в космосе и наши следы... Казань: Татар. кн. изд-во, 2000. 62 с.
37. У.Н.Закиров. Релятивистская динамика сосредоточенной переменной массы покоя. Казань: Изд-во Казан. гос. техн. ун-та, 2000. 184 с.
38. А.Н.Гильманов. Методы адаптивных сеток в задачах газовой динамики. М.: Наука-Физматлит, 2000. 248 с.
39. М.А.Ильгамов. Профессор Х.М.Муштари. М.: Наука. Физматлит, 2001. 192 с.
40. Н.М.Якупов, Ш.К.Галимов, Н.И.Хисматуллин. От каменных глыб к тонкостенным конструкциям. Казань: SOS, 2001. 96 с.
41. С.Н.Сидоренко, Н.М.Якупов. Коррозия - союзник аварий и катастроф. М.: Изд-во РУДН, 2002. 93 с.
42. М.А.Ильгамов, А.Н.Гильманов. Неотражающие условия на границах расчетной области. М.: Наука. Физматлит, 2003. 240 с.
43. Р.Г.Зарипов. Самоорганизация и необратимость в неэкстенсивных системах. Казань: Изд-во "Фэн", 2002. 251 с.
44. У.Н.Закиров. Физическая механика межзвездного полёта. Казань: Изд-во "Фэн", 2003. 84 с.
45. В.М.Ларионов, Р.Г.Зарипов. Автоколебания газа в установках с горением. Казань: Изд-во Казан. гос. техн. ун-та, 2003. 227 с.
46. П.П.Осипов. Задачи переноса при свободномолекулярном течении газа и лучистом теплообмене. Казань: Изд-во Казан. гос. техн. ун-та, 2004. 82 с.
47. Р.Г.Зарипов. Новые меры и методы в теории информации. Казань: Изд-во Казан. гос. техн. ун-та, 2005. 364 с.
48. Проблемы инновационной экономики и инфокоммуникационных технологий / И.В.Безденежных, М.Р.Габайдуллин, ..., А.И.Маликов и др. // Под ред. Т.К.Сиразетдинова. М. – Казань: Академия наук риска. 2005. 412 с.

49. Х.Н.Низамов, С.Н.Сидоренко, Н.М.Якупов. Прогнозирование и предупреждение коррозионного разрушения конструкций. М.: Изд-во РУДН, 2006, 355 с.
50. М.Х.Хайруллин, Р.С.Хисамов, М.Н.Шамсиев, Р.Г.Фархуллин. Интерпретация результатов гидродинамических исследований скважин методами регуляризации. М.-Ижевск: НИЦ «Регулярная и хаотическая динамика»; Институт компьютерных исследований, 2006. 172 с.
51. Н.М.Якупов. Лаборатория нелинейной механики оболочек: история и разработки последних лет, Казань: ИММ КазНЦ РАН, 2006. 98 с.
52. В.И.Лукашенко, А.А.Абдюшев и др. Экспертиза, расчет, анализ пространственных конструкций. Казань: КГАСУ, 2006, 321 с.
53. Р.Ш.Гимадиев. Динамика мягких оболочек парашютного типа. Казань: Казан. гос. энерг. ун-т, 2006. 208 с.
54. И.Х.Мифтахутдинов. Оболочки минимальной поверхности в природе и в архитектуре. Казань: ЗАО "Новое Знание", 2007. 144 с.
55. У.Н.Закиров. Космическая хроника 60–70 годов. Автобиографические заметки. Казань: КГТУ, 2008. 99 с.
56. А.Е.Смородов, Р.Н.Галиахметов, М.А.Ильгамов. Физика и химия кавитации. М.: Наука, 2008. 228 с.
57. И.Х.Мифтахутдинов. Наглядная геометрия оболочек минимальной поверхности. Казань: ЗАО "Новое знание", 2009. 40 с.
58. М.А.Ильгамов. Портреты современников. М.: Физматлит, 2009. 276 с.
59. И.Х.Мифтахутдинов. Методы расчета оболочек минимальной поверхности. Формообразование+НДС. Казань: ЗАО "Новое знание", 2009. 68 с.
60. Теория и практика моделирования разработки нефтяных месторождений в различных геолого-физических условиях / Р.С.Хисамов, Р.Р.Ибатуллин, А.И.Никифоров, А.Ф.Иванов, Р.Х.Низаев. Казань: Изд-во «Фэн» Академии наук РТ, 2009. 239 с.
61. Н.М.Якупов. Механика: проблема - идея - практика. Казань: Казан. гос. ун-т, 2010. 161 с.
62. Р.Г.Зарипов. Принципы неэкстенсивной статистической механики и геометрии мер беспорядка и порядка. Казань: Изд-во Казан. гос. техн. ун-та, 2010. 404 с.
63. И.Х.Мифтахутдинов. Аэродинамика оболочек в архитектуре и строительстве. Казань: ЗАО «Новое знание», 2010. 60 с.
64. И.Н.Хакимзянов, Р.С.Хисамов, Р.Р.Ибатуллин, Р.Т.Фазлыев, А.И.Никифоров. Наука и практика применения разветвленных и многозабойных скважин при разработке нефтяных месторождений. Казань: Изд-во "Фэн" АН РТ, 2011. 320 с.
65. А.И.Никифоров, Р.Х.Низаев, Р.С.Хисамов. Моделирование потокоотклоняющих технологий в нефтедобыче. Казань: Изд-во "Фэн" АН РТ, 2011. 224 с.
66. М.Х.Хайруллин, Р.С.Хисамов, М.Н.Шамсиев, Е.Р.Бадертдинова. Гидродинамические методы исследования вертикальных скважин с трещиной гидроразрыва пласта. М.-Ижевск: Институт компьютерных исследований, 2012. 84 с.
67. М.А.Ильгамов. Резонанс. Москва-Уфа: Гилем, 2012. 210 с. (второе издание: 2013).

Марат Аксанович Ильгамов, член-корреспондент РАН. Специалист в области механики гидроупругих систем: динамики и устойчивости оболочек, взаимодействующих со сплошными и сыпучими средами. Им развита теория сильного взаимодействия тонкостенных конструкций с жидкостью и газом с учетом больших перемещений контактной поверхности. Организатор и первый директор ИММ КазНЦ РАН. В настоящее время заведующий лабораторией Механика твердого тела Института механики УНЦ РАН, гл.н.с. ИММ КазНЦ РАН.
e-mail: info@imm.knc.ru

On 70-th Anniversary of investigations on mechanics in Kazan Science Center of RAS

M.A. Ilgamov

IME KazSC RAS,
Russia, 420111, Kazan, Lobachevskii str. 2/31

Abstract. The article presents a brief information about the Kazan Branch of the USSR Academy of Sciences / Kazan Science Center of the Russian Academy of Sciences. Main directions of research on mechanics developed in the Kazan Physical and Technical Institute and the Institute of Mechanics and Engineering of RAS KazSC are enumerated, some results are mentioned.

Keywords: Kazan Branch of the USSR Academy of Sciences (KB), Kazan Science Center of the Russian Academy of Sciences (RAS KazSC), Institute of Mechanics and Engineering (IME) of RAS KazSC, investigations on mechanics.

It has been over 70 years since the organization of the Kazan branch of the USSR Academy of Sciences. In April 1945, the Council of People's Commissars of the USSR adopted a resolution on the opening of Kazan Branch of the USSR Academy of Sciences (KB). That same year, the Presidium of the USSR approved the structure.

The following institutions were organized: Physics and Technology, Chemistry, Geology, Biology, Languages, Literature and History, Department of water problems and energy.

Academician A.E. Arbuzov became the chairman of Kazan Branch of the USSR Academy of Sciences, professor L.M. Miropolskiy became his deputy and professor G.Kh. Kamai was appointed to be a scientific secretary.

The branch started functioning at January 1, 1946. Party and Soviet authorities of Tatarstan republic, leading Kazan universities, especially Kazan State University, contributed greatly to it's establishing and development [1].

The activity of Kazan Branch of the USSR Academy of Sciences - RAS KazSC covers four periods. The first 17 years (1946-62) were marked by fruitful coordinated work of a relatively small research team under the guidance of an outstanding Soviet scientist A.E. Arbuzov. The library, workshops, accounting and all services were common throughout the branch.

The second period covers 10 years of independent work of all institutions, which was caused by the USSR Academy of Sciences reforming. Kazan Branch was abolished.

The third period starts in 1973, when the Kazan Branch was reopened. The structure included the following institutes: Physics and Technology, Organic and Physical Chemistry, Biology, Language, Literature and History.

Kazan Branch of the USSR AS Institute of Informatics Problems was established later. These institutes also had independent economic and financial functions.

The fourth period starts in 1991 and continues till the present. The Department of Energy and Institute of Mechanics and Engineering of Kazan Scientific Center of the USSR Academy of Sciences were established in 1991.

The Institute of Language, the Institute of Literature and History, as well as the Kazan Branch of the USSR AS Institute of Informatics Problems became parts of the Tatarstan Republic Academy of Sciences.

This periodization is, of course, very conditional-relative. Nevertheless, we will stick to it, and without addressing the fourth period, i.e. the last quarter of the mechanics development century.

The aim of this work is to describe the basic research directions in mechanics, developed in the Physical-Technical Institute and the Institute of Mechanics and Engineering, and to mention some results.

For 70 years many areas of mechanics were developed, significant results were obtained. However, not all of them are mentioned in this work, but it doesn't because of its not importance. Unfortunately, little information is reported about subsurface hydromechanics and oil production processes modeling. It is a big gap in this review.

The history of the Kazan Physical-Technical Institute (KPhTI) of the Kazan Branch began in two rooms at the Kazan University geometric building with just a few tables and chairs. Later a branch was granted a building in 2/31 Lobachevsky street. But it has long been repaired, annex has been built. During the period of the approval of Kazan Branch of the USSR Academy of Sciences Corresponding Member of the USSR Academy N.G. Chebotarev was listed as a director of KPhTI, but since the beginning of functioning professor Kh.M. Mushtari was the director.

Under the leadership of Kh.M. Mushtari Department of Mechanics began working on the development of the nonlinear theory of thin shells, variational and numerical methods of deformation analysis, stability. K.Z. Galimov, S.G. Vinokourov, R.G. Surkin, G.G. Tumashev became the first members. N.G. Chebotarev was the first head of the Mathematics Department, since 1947 the Department was headed by G.S. Salekhov. A.P. Norden, A.N. Khovanskii, B.A. Seifullin, V.L. Danilov, V.D. Chugunov, S.F. Korotkov, G.P. Tsybulskii were the first members.

The Department of mathematics conducts the research on the seepage theory of natural fluids in porous media since 1951. Professor G.S. Salekhov's school gained prominence by the research in the application of mathematical methods to the issues of oil field development. He proposed methods for solving the oil-bearing contour control problem. Effective ways of solution based on the physical-geological, technological and economic factors, were proposed in the members' works. The evolution of G.S. Salekhov's works includes formulation and development of methods for solving various optimization problems of the oil industry.

In 1947 the first publications on the shells theory appear made by the mechanic department members in such prestigious journals as the PMM, DAN USSR and other. These studies relate mainly to general issues of the shells theory and stability problems [2, 3]. Kh.M. Mushtari works in collaboration with S.G. Vinokourov, R.G. Surkin. For example, one of the most important works "Nonlinear theory of the spherical shell stability" was published in PMM in 1950. Academician V.V. Novozhilov wrote twenty years later [4]: "Many scientists have been working on this problem, but the solution seems to be the closest to a proven by track record belongs to Kh.M. Mushtari. When evaluating the submersibles spherical shells stability American scientists use exactly Mushtari formula, thereby agreeing that it should be declared as the most reliable. "

The paper presents a critical analysis of Karman and Tsien research, two mistakes are pointed out. Also Kh.M. Mushtari pointed out a mistake in the Friedrichs work. In another paper Kh.M. Mushtari gave a critical analysis of the V.I. Feodos'ev work (1954). The established authority of the shells theory, E.I. Grigolyuk notes this fact [3]: "As an outstanding mechanical engineer he (Kh.M. Mushtari) could not notice that the deformation zones section contour is also determined by the deformation process, so it is necessary to vary not only sagging in the shell center, but zones conjugation contour radius".

Note that these two Kh.M. Mushtari works resulted in further cooperation with the great American-Chinese scientist Qian and many years of conflict with the outstanding mechanician V.I. Feodos'ev. This period also includes research in the field of the border zone

stresses, stability of shells with initial deviations from the desired shape, inverse problems of the shells theory, medium thickness plates and shells, etc.

K.Z. Galimov got equilibrium equations related to the deformed and non-deformed shell states, proved variational theorem of mixed type.

Major milestones were the performance of the large amount of the research by department of mechanics in connection with the creation of aerospace technology, to the company, headed by SP Korolev, and publication of "Proceedings of the Physico-Technical Institute of the USSR KFAN" in 1954, now closed; monograph publication[5]; holding All-Union Conference on the Theory of plates and shells (1960), thesis defenses.

The end of this period includes the installation of a computer and the beginning of a systematic and large-scale studies on the stress-strain state of various forms plates by M.S. Kornishin and F.S. Isanbaeva. Prior to this, M.S. Kornishin conducted a large study of nonlinear problems of plates and panels bending and stability using the Bubnov-Galerkin method. The nonlinear problems of shell theory were solved mainly by this method.

I.V. Svirsky analyzed problems of the Bubnov-Galerkin method application to the nonlinear behavior of the membranes. He obtained unique results in applying the successive approximations method. N.S. Ganiev and A.V. Sachenkov considered the various tasks of the stress-strain state and stability of shells and plates. I.G. Teregulov studied the behavior of shells, taking into account the creep of the material. M.S. Ganeeva studied stress-strain state of rotational shells. N.K. Galimov researched sandwich plates theory. B.M. Zuev with team engaged in experimental study using strain measurement and optical method. Later, the group was transferred to the Institute of Organic and Physical Chemistry, where they conducted the materials testing.

The same directions continued to develop in the second period (1963-1972). Numerical simulation, sandwich plates theory got a great development, also research on the theory of the plates and shells interaction with liquid, gas and solid deformable body began. This period includes, for example, the research of the repetitive waves in a pipe, its interaction with the thin-walled cylindrical and conical shell. One of the most significant results – experimental detection of the "pest" degradation (for about 1 s) of the steel shell, inside which the air waves are excited. The theory of this phenomenon could not be created on the basis of the analytical methods. But after the development of numerical methods this problem did not attract enough attention. Of course, it is not the only reason, the phenomenon itself is very complex.

The third mechanical department was established - Laboratory of the medium-interacted shells theory (1971). Unlike the first laboratories, which were established by the outstanding scientists – Kh.M. Mushtari and N.G. Chebotarev, who have made a major contribution to their fields of science by 1946, the third department arisen after 25 years in the bowels of the existing mechanics laboratory. With regret we note that it was not created yet another department – department of the underground mechanics at a good time (departments were created in KPhTI only on physical directions).

The move to the new building made it possible to enforce the computational tools and to create experimental basis.

The workshop continued very active work, and every Thursday members reported on finished works and discussed research progress and problem questions. Perhaps none of the manuscript was published without previous workshop discussion.

"Proceedings of the shells theory workshop" were published since 1969, the editors were M.S. Kornishin and M.A. Ilgamov. The volumes came almost every year. A total of 27 issues have been published (last - in 1992). This speeded up the research results publication. In the context of the small number of central journals it was an opportunity. But these proceedings have not been widely known in the country and abroad.

The third period (1973 - 1990) is characterized by a highly developed numerical modeling of deformation of thin-wall structures and the oil fields development. This includes the classic problems, which had been already known in few last years, and new challenges.

These challenges include, for example, the static and dynamic problems of shell and elastic solid fill joint deformation (solid-fuel rocket model), the shells interaction with liquid and gas (combustion chamber reactors and pipelines models), as well as the soft shells interaction with a gas flow (parachute systems).

R.G. Surkin with the team has built plant for the spherical segments shock loading. Results of segment sustainability have entered the encyclopedic monograph [6]. It should be mentioned that studying of the thin plates (fuse membranes) was further developed (Yakupov N.M., Galimov N.K.).

R.G. Zaripov built periodic shock-waves maker, and A.A. Pavlov built a plant for studying the thin-walled products hydrodynamic stamping. Appropriate theory was developed with the participation of M.M. Shakiryanov, N.N. Shihranov took part. In particular, the observed phenomenon of cavitation bubble liquid at forging formed the basis of Sh.U. Galiev's research in Kiev.

M.B. Gafurov built the test pool, G.V. Sadykov, R.I. Davydov with members - a wind tunnel. Using these means new results were obtained, which directed theory development. B.V. Gulin, Zh.M. Sahabutdinov, A.N. Gilmanov, A.A. Aganin, R.R. Shagidullin, M.Z. Sabitov, A.A. Tukmakov were engaged in deformable elements flow problem. M.V. Taldykin has created models of wave propellers and pumps; the theory was developed by V.L. Fedyaev, M.M. Suleymanova. These studies were carried out in contact with the Leningrad sea institutions and have been praised by Academician M.A. Lavrentiev. R.N. Miftakhov was involved in biomechanics problems.

In all of these studies we were moving too fast, and we did not have enough time to unveil the results clearly and to publish it in the national and international journals. Therefore, a lot of excellent results were neglected.

It should be noted that the leadership of the Kazan Branch, especially KPhTI, failed to fit into the problem of KAMAZ, the largest engineering facility of the 70-s, and to get a good development on this basis. We have created the KphTI laboratory of KAMAZ problems with considerable delay. But time was lost, there was not such a good development. Taking into account the missed opportunities we worked actively in connection with the construction of Yelabuga Tractor Plant in the early eighties. It was possible to achieve an agreement on the transfer of payroll and staff number of the Ministry of Trac and Agricultural Engineering to the USSR Academy of Sciences in order to organize two KPhTI laboratories (heads – Zh.M. Sahabutdinov, V.L. Fedyaev). Applied Mathematics Department of KPhTI and Kazan Pedagogical Institute was created, and the employees worked for economic agreement means that we had a lot of (Head – M.A. Ilgamov). This organizational form was advantageous, because self-supporting part of the money were sent to the payroll in universities, and the USSR Academy of Sciences institutes could spend such amounts of money per equipment and business trips.

Later the laboratory of solid rocket motors development factory (Head - AN Gilmanov) was created. Numerical simulation of unstable operation mode found immediate application at the factory.

As part of the USSR Academy of Sciences and the Ministry of Education of the USSR program of the school computerization the department was created (Head - Kuznetsov V.B.).

All of these activities on expanding the topics in the relevant areas have occurred against the background of a national situation deterioration in the country in the eighties, which began with its deindustrialization.

During the pseudodemocracy period mechanical departments lost its positions, particularly because of the new KFTI's head position. In particular, several staff units were transferred to the physical laboratory, and we have lost production facilities.

We mentioned above the first All-Union Conference on the Theory of Shells in 1960, dedicated to the 60th anniversary of Kh.M. Mushtari's birth.

Such conferences were held in Kazan in 1970, 1980, 1995, 2000. The last conference was dedicated to the 100th anniversary of Kh.M. Mushtari's birth, the 90th anniversary of professor K.Z. Galimov's birth, the 80-th anniversary of professor M.S. Kornishin's birth. We can note the Conference on Boundary Problems of continuum mechanics, dedicated to the 100th anniversary of G.G. Tumashev's birth and the 110- th anniversary of Kh.M. Mushtari's birth (2010). Our mechanicians were the organizers of the other, perhaps, smaller conferences and summer schools. Recent conferences were organized at the camp "Volga" in 1975 and 1986 and dedicated to the problems of the membranes interaction with a liquid, gas and solid compact body. Conference dedicated to the 95th anniversary of M.A. Lavrent'ev (1995) was held at a high level and with a large number of participants. But the most remarkable event is the first congress on theoretical and applied mechanics in Kazan, 2015. We should note the great contribution of IME employees, headed by D.A. Gubaidulin,.

List of monographs, published by KPhTI and IMM mechanicians is given in Appendix.

Abbreviations list

KazSC RAS – Kazan Science Center of the Russian Academy of Sciences

IME – Institute of Mechanics and Engineering

KB USSR AS – Kazan Branch of the USSR Academy of Sciences

KPhTI – Kazan Physical-Technical Institute

References

1. Kazanskij filial Akademii nauk SSSR. Buklet. Kazan'. Izd-vo Tat.OK KPSS. 1987. – 20 p. (Redactor M.A. Ilgamov).
2. H.M. Mushtari. *Raboty kazanskih uchenyh po nelinejnoj teorii obolochek* [Kazan scholar proceedings on nonlinear theory of shells] // Izv. KFAN SSSR. Ser. fiz.-mat. i tehn. nauk. 1958. № 12. p. 5-15.
3. M.A. Ilgamov. Professor H.M. Mushtari. M.: Nauka. Fizmatlit. 2001. 192 p.
4. V.V. Novozhilov. *Kratkij ocherk razvitiya teorii obolochek v SSSR // issledovaniya po teorii plastin i obolochek* [A brief sketch of the development of the shells theory in the USSR // research on the plates and shells theory]. Vol. VI-VII. Kazan': izd-vo KGU. 1970. p. 3-22.
5. H.M. Mushtari, K.Z. Galimov. *Nelinejnaja teorija uprugih obolochek* [Nonlinear theory of the elastic shells]. Kazan': Tatknigoizdat. 1957. 431 p.
6. A.S. Vol'mir. *Ustojchivost' deformiruemyh system* [The stability of deformable systems]. M.: Nauka, 1967. 984 p.

Appendix

List of monographs, published by mechanicians of Kazan Physical-Technical Institute of the Kazan Branch of the USSR Academy of Sciences and Institute of Mechanics and Engineering of RAS KazSC

1. G.S.Salekhov. Vychislenie ryadov. Moskva: Gostekhizdat, 1955. 144 p. (In Russian)
2. H.M.Mushtari, K.Z.Galimov. Nelineinaya teoriya uprugikh obolochek. Kazan': Tatknigoizdat, 1957. 432 p. (eng.: 1961)
3. Matematicheskiy analiz. Funktsii, predely, ryady, tsepnye drobi / V.L.Danilov, A.N.Ivanova, E.K.Isakova, L.A.Lyusternik, G.S.Salekhov, A.N.Khovanskiy, L.Ya.Tslaf, A.R.Yanpolskiy. M.: Fizmatgiz, 1961. 439 p. (In Russian)

4. M.S.Kornishin. Nelineinye zadachi teorii plastin i pologikh obolochek i metody ikh resheniya. M.: Nauka, 1964. 192 p. (In Russian)
5. M.S.Kornishin, F.S.Isanbaeva. Gibkie plastiny i paneli. M.: Nauka, 1968. 260 p. (In Russian)
6. I.V.Svirskiy. Metody Bubnova-Galerkina i posledovatelnykh priblizheniy I ikh primenenie k raschetu obolochek. M.:Nauka, 1968. 200 p. (In Russian)
7. I.G.Teregulov. Izgib i ustoichivost' tonkikh plastin i obolochek pri polzuchesti. M.:Nauka, 1969. 206 p. (In Russian)
8. M.A.Ilgamov. Kolebaniya uprugikh obolochek, soderzhachikh zhidkost' i gaz. M.:Nauka, 1969. 182 p. (In Russian)
9. V.V.Skvortsov. Matematicheskiy eksperiment v teorii razrabotki neftnykh mestorozhdenii. M.:Nauka, 1970. 224 p. (In Russian)
10. G.S.Salekhov, L.M.Muratov, V.E.Pospeev. Vychislenie ryadov i nesobstvennykh integralov. Kazan: Izd-vo Kazan. un-ta, 1973. 169 p. (In Russian)
11. V.V.Skvortsov. Metody ekspertnykh otsenok i ih prilozhenie v zadachah teorii filtratsii. Kazan: Tat.knizh.izd-vo, 1976. 216 p. (In Russian)
12. M.A.Ilgamov, V.A.Ivanov, B.V.Gulin. Prochnost', ustoichivost' i dinamika obolochek s uprugim zapolnitelem. M.:Nauka, 1977. 332 p. (eng.: 1981).
13. U.N.Zakirov. Mekhanika relyativistskikh kosmicheskikh poletov.. M.:Nauka, 1984. 152 p. (In Russian)
14. M.A.Ilgamov, V.A.Ivanov, B.V.Gulin. Raschet obolochek s uprugim zapolnitelem. M.:Nauka, 1987. 264 p. (In Russian)
15. E.H.Dowell, M.A.Ilgamov. Studies in Nonlinear Aeroelasticity. New York - London - Tokyo, Springer-Verlag, 1988. 456 p. (In Russian)
16. M.S.Kornishin, V.N.Paimushin, V.F.Snigerev. Vychislitel'naya geometriya v zadachah mekhaniki obolochek, M.:Nauka, 1989. 208 p. (In Russian)
17. A.I.Golovanov, M.S.Kornishin. Vvedenie v metod konechnykh elementov statiki tonkikh obolochek. Kazan: 1990. 271 p. (In Russian)
18. H.M.Mushtari. Nelineinaya teoriya obolochek. M.:Nauka, 1990. 223 p. (In Russian)
19. V.V.Ridel', B.V.Gulin. Dinamika myagkikh obolochek. M.:Nauka, 1990. 205 p. (In Russian)
20. M.A. Ilgamov. Vvedenie v nelineinuyu gidrouprugost'. M.:Fizmatlit, 1991. 200 p. (In Russian; chin.: 1993).
21. M.S.Ganeeva. Prochnost' i ustoichivost' obolochek vrashcheniya. M.:Nauka, 1992. 161 p. (In Russian)
22. N.M.Yakupov. Prikladnie zadachi mekhaniki tonkostennykh konstruktssii. Kazan: IMM KazNC RAN, 1994. 124 p. (In Russian)
23. N.M.Yakupov, M.N.Serazutdinov. Raschet uprugikh tonkostennykh konstruktssii slozhnoi geometrii. Kazan: IMM KazNC RAN, 1993. 208 p. (In Russian)
24. M.A.Ilgamov. Staticheskie zadachi gidrouprugosti. Kazan IMM KazNC RAN, 1994. 208 p. (In Russian)
25. Zh.M.Sakhabutdinov. Analiz diskretnykh modelei dvizheniya tochki. Kazan IMM KazNC RAN, 1995. 196 p. (In Russian)
26. Obzory issledovaniy po mekhanike sploshnoi sredy. K 50-letiyu KNC RAN. Kazan IMM KazNC RAN, 1995. 216 p. (In Russian)
27. A.B.Mazo. Matematicheskoe modelirovanie protsessov goryachei obrabotki metallov. Kazan: Kazanskii fond "Matematika". 1996. 209 p. (In Russian)
28. N.M.Yakupov, I.N.Gataullin, R.N.Khismatullin. Obsledovanie, analiz i prognozirovaniye dolgovechnosti stroitelnykh konstruktssii i rekomendatsii po ikh vosstanovleniyu. Metodicheskoe rukovodstvo, Kazan: IMM KNC RAN, 1996. 208 p. (In Russian)
29. Bezopasnost' proizvodstvennykh ob'ektov. Pod red. N.M.Yakupova. Kazan: IMM KNC RAN, MChS RT, 1996. (In Russian)
30. F.A.Saljanov. Osnovy fiziki nizkoterperaturnoi plazmy, plazmennyykh apparatov i tehnologii. M.:Nauka, 1997. 240 p. (In Russian)
31. Vector Lyapunov Functions in Stability Theory / R.Z.Abdullin, L.J.Anapolski, R.I.Kozlov, A.I.Malikov, V.M.Matrosov, A.A.Voronov, A.S.Zemljakov. Advanced series in mathematical science and engineering. World Federation Publisher Company. 1996. 394 p.
32. N.M.Yakupov, N.I.Khismatullin. Mekhaniki Nizhmekamskneftehima. Kazan: IMM KNC RAN, 1996. 224 p. (In Russian)
33. D.A.Gubaidullin. Dinamika dvuhfaznykh parogazokapelnnykh sred. Kazan: Izd-vo KAzan.matematicheskogo obschestva, 1998. 154 p. (In Russian)
34. M.A.Ilgamov. Static Problems of Hydroelasticity. Moscow: Nauka. Fizmatlit, 1998. 208 p.
35. R.G.Zaripov. Informatsiya pazlichiya i perekhody besporiyadok – poryadok. Kazan: Izd-vo KGTU im.A.N.Tupoleva, 1999. 155 p. (In Russian)
36. U.N.Zakirov. Est' v kosmose i nashi sledy... Kazan: Tatar.kn.izd-vo, 2000. 62 p. (In Russian)

37. U.N.Zakirov. Relyativistskaya dinamika sosredotochennoi peremennoi massy pokoya. Kazan: Izd-vo Kazan. gos. tehn. un-ta, 2000. 184 p. (In Russian)
38. A.N.Gilmanov. Metody adaptivnykh setok v zadachakh gazovoi dinamiki. M.: Nauka-Fizmatlit, 2000. 248 p. (In Russian)
39. M.A.Ilgamov. Professor Kh.M. Mushtari. M.: Nauka-Fizmatlit, 2001. 192 p. (In Russian)
40. N.M.Yakupov, I.N.Gataullin, R.N.Khismatullin. Ot kamennykh glyb k tonkostennym konstrktsiyam. Kazan: SOS, 2001. 96 p. (In Russian)
41. S.N.Sidorenko, N.M.Yakupov. Korroziya – soyuznik avarii i katastrof. M.: Izd-vo RUDN, 2002. 93 p. (In Russian)
42. M.A.Ilgamov, A.N.Gilmanov. Neortazhayuschie usloviya na granitsakh raschetnoi oblasti. M.: Nauka-Fizmatlit, 2003. 240 p. (In Russian)
43. R.G.Zaripov. Samoorganizatsiya i neobratimost' v neekstensivnykh sistemakh. Kazan: Izd-vo "Fen", 2002. 251 p. (In Russian)
44. U.N.Zakirov. Fizicheskaya mekhanika mezhzvezdnogo poleta. Kazan: Izd-vo "Fen", 2003. 84 p. (In Russian)
45. V.M.Larionov, R.G.Zaripov. Avtokolebaniya gaza v ustanovkakh s goreniem. Kazan: Izd-vo Kazan. gos. tehn. un-ta, 2003. 227 p. (In Russian)
46. P.P.Osipov. Zadachi perenosy pri svobodnomolekulyarnom techenii gaza i luchistom teploobmene. Kazan: Izd-vo Kazan. gos. tehn. un-ta, 2004. 82 p. (In Russian)
47. R.G.Zaripov. Novye mery i metody v teorii informatsii. Kazan: Izd-vo Kazan. gos. tehn. un-ta, 2005. 364 p. (In Russian)
48. Problemy innovatsionnoi ekonomiki i infokommunikatsionnykh tekhnologii / I.V.Bezdenezhnyh, M.R.Gabaidullin, ..., A.I.Malikov. I dr. // pod red. T.K. Sirazetdinova. M. – Kazan: Akademiya nauk riska, 2005. 412 p. (In Russian)
49. Kh.N.Nizamov, S.N.Sidorenko, N.M.Yakupov. Prognozirovaniye i preduprezhdeniye korrozionnogo razrusheniya konstruktssii. M.: Izd-vo RUDN, 2006. 355 p. (In Russian)
50. M.Kh.Khairullin, R.S.Khisamov, M.N.Shamsiev, R.G.Farkhullin. Interpretatsiya rezultatov gidrodinamicheskikh issledovaniy skvazhin metodami regilyarizatsii. M.-Izhevsk: NIC "Regulyarnaya i khaoticheskaya dinamika"; Institut komputernykh issledovaniy, 2006. 172 p. (In Russian)
51. N.M.Yakupov. Laboratoriya nelineinoi mekhaniki obolochek: istoriya i razrabotki poslednikh let, Kazan: IMM KazNC RAN, 2006. 98 p. (In Russian)
52. V.I.Lukashenko, A.A.Abdyushev i dr. Ekspertiza, raschet, analiz prostranstvennykh konstruktssii. Kazan: KGASU, 2006. 231 p.
53. R.Sh.Gimadiev. Dinamika myagkikh obolochek parashutnogo tipa. Kazan: Kazan.gos.energ.un-t, 2006. 208 p. (In Russian)
54. I.Kh.Miftakhutdinov. Obolochki minimalnoi poverkhnosti v prirode i v arkhitekture. Kazan: Zao "Novoe Znanie", 2007. 144 p. (In Russian)
55. U.N.Zakirov. Kosmicheskaya khronika 60-70 godov. Avtobiograficheskie zametki. Kazan: KGTU, 2008. 99 p. (In Russian)
56. A.E.Smorodov, R.N.Galiakhmetov, M.A.Ilgamov. Fizika i khimiya kavitatsii. M.: Nauka, 2008. 228 p. (In Russian)
57. I.Kh.Miftakhutdinov. Naglyadnaya geometriya obolochek minimalnoi poverkhnosti. Kazan: ZAO "Novoe Znanie", 2009. 40 p. (In Russian)
58. M.A.Ilgamov. Portrety sovremennikov. M.:Fizmatlit, 2009. 276 p. (In Russian)
59. I.Kh.Miftakhutdinov. Metody rascheta obolochek minimalnoi poverkhnosti. Formoobrazovanie +NDS. Kazan: ZAO "Novoe Znanie", 2009. 68 p. (In Russian)
60. Teoriya i praktika modelirovaniya razrabotki nefityanykh mestorozhdenii v razlichnykh geologo-fizicheskikh usloviyakh / R.S.Khisamov, R.R.Ibatullin, A.I.Nikiforov, A.F.Ivanov, R.Kh.Nizaev. Kazan: Izd-vo "Fen" Akademii nauk RT, 2009. 239 p. (In Russian)
61. N.M.Yakupov. Mekhanika: problema – ideya – praktika. Kazan: Kazan.gos. un-t, 2010. 161 p. (In Russian)
62. R.G.Zaripov. Printsipy neekstensivnoi statisticheskoi mekhaniki i geometrii mer besporyadka i poryadka. Kazan: Izd-vo Kazan. gos. tekhn. un-ta, 2010. 404 p. (In Russian)
63. I.Kh.Miftakhutdinov. Aerodinamika obolochek v arkhitekture i stroitelstve. Kazan: ZAO "Novoe Znanie", 2010. 60 p. (In Russian)
64. I.N.Khakimzyanov, R.S.Khisamov, R.R.Ibatullin, R.T.Fazlyev, A.I.Nikiforov. Nauka i praktika primeneniya razvetvlennykh i mnogozaboynykh skvazhin pri razrabotke nefityanykh mestorozhdenii. Kazan: Izd-vo "Fen" AN RT, 2011. 320 p. (In Russian)
65. A.I.Nikiforov, R.Kh.Nizaev, R.S.Khisamov. Modelirovaniye potokooklonyayuschikh tekhnologii v neftedobyche. Kazan: Izd-vo "Fen" AN RT, 2011. 224 p. (In Russian)

66. M.Kh.Khairullin, R.S.Khisamov, M.N.Shamsiev, E.R.Badertdinova. Gidrodinamicheskie metody issledovaniya vertikalnykh skvazhin s treschinoi gidrorazryva plasta. M.-Izevsk: Institut komputernykh issledovaniy, 2012. 84 p. (In Russian)

67. M.A.Ilgamov. Rezonans. Moskva-Ufa: Gilem, 2012, 210 p. (In Russian, second edition: 2013).

Marat Aksanovich Ilgamov, Corresponding Member of the Russian Academy of Sciences. He is a specialist in the area of mechanics of hydroelastic systems such as dynamics and stability of shells interacting with continuous and loose media. He developed a theory of strong interaction of thin-walled structures with fluids taking into account large displacement of the contact surface. He is organizer and first Director of IME KazSC RAS. At present he is head of "Solid Mechanics" Laboratory, Institute of Mechanics, Ufa Branch RAS, chief research scientist of IME KazSC RAS,
e-mail: info@imm.knc.ru

Advances in operational research for development

(IFORS- EWG ORD Workshop and 28th EURO Conference)

L. Lotero¹, S.S. Ali², E. del Rosario³, G.-W. Weber⁴

¹Facultad de Ingeniería Industrial, Universidad Pontificia Bolivariana
050031, Medellin, Colombia

²New Delhi Institute of Management,
Tughlakabad, New Delhi, India

³14A CyberOneBldg. Eastwood, C5, 1110 Quezon City, Metro Manila, Philippines

⁴Institute of Applied Mathematics, Middle East Technical University
(ODTÜ)06800 Ankara, Turkey

Abstract. The year 2016 marked the first decade of the EURO Working Group on OR for Development. This Workshop was held jointly with IFORS right before the EURO 2016 Conference in Poznan. Techniques presented at the Workshop included causal mapping and Priority Pointing Procedure which were used to gather a common understanding of the problem on how to make Operations Research more effective in addressing issues of development. Various papers were presented using a wide array of OR techniques in the area of health, education, agriculture, poverty, traffic, humanitarian logistics and finance.

Key words: operational research, developing countries, Working Group on OR.

1. Introduction

The 10th Anniversary Workshop of the EURO WG on OR for Development (EWG ORD) was appropriately themed *Highlighting A Decade of Work on OR for Development* (<http://ifors.org/ewg-ord/>). It was jointly sponsored by the EWG ORD (<https://www.euro-online.org/web/ewg/29/or-for-development-ewg-ord>) and IFORS (International Federation of Operational Research Societies) (<http://ifors.org/web/>). The workshop was a satellite event of the 28th EURO conference on Operational Research, (<http://www.euro2016.poznan.pl/>) held in Poznan.

Since it was first put up in 2007, the workshop series aimed to encourage practitioners, academicians and the OR (Operational Research) community worldwide in enriching their knowledge through learning and interaction related to issues that are faced in development, in the areas of education, health, infrastructure, environmental sustainability, energy, population growth, poverty and climate change, among others [1-7].

During the Workshop, each 30-minute presentation of papers was followed by constructive feedback from two previously assigned reactors. Apart from honing their presentation skills, participants benefited from the opportunity to get perspectives of graduate students, academicians, early stage practitioners and more experienced mentors for suggestions on further improvement of their work. Effectual conversations and productive idea sharing were extended after Workshop hours, during the tours with the sights and sounds of Poznan in the background.

2. Diverse application areas at the EWG-ORD Workshop

Showing her vast experience in the field of OR, EWG ORD Chair *Elise del Rosario*, shared her consulting experience through three projects in the area of - globalization of manufacturing facilities, national power dispatching, and city port operations- during the Opening Session. These cases, which had different implementation outcomes, gave a clear view of problems and

successes encountered by an OR analyst in real life situations. From these, lessons from the successful and unsuccessful implementation of recommendations were drawn[8].

This was followed by a Workshop by *Cathal Brugha*, Emeritus and Adjunct Professor at the University College Dublin Center for Business Analytics who engaged participants in an interactive exercise to answer the question: *What is needed to Help OR be more effective in addressing development problems?*[9-10]. He showed how the Priority Pointing Procedure can lead to a strategy.

Ashley I. Carreras, Principal Lecturer in Economics and Decision Analysis at the De Montfort University, presented a Tutorial/Workshop on *A Practicing Academic Perspective on Problem Structuring Methods*. Participants, divided into two groups, learned to do causal mapping, leading to a shared understanding of the issue that was: *What are the key issues that need to be addressed by OR researchers in order to ensure that OR research continues to effectively influence policy makers?*[11]. Carreras then summarized through causal mapping, the discussions of the two groups. All participants had the chance to learn by doing the causal mapping procedure.

Papers presented were carefully selected from various abstract submissions. The sectors addressed were: **Health:** *Designing Blood Supply Chain: A Location-Allocation Model with Collection and Production Considerations* (Osorio) and *Urban Mobility and Disease Spread Dynamics in Medellin, Colombia* (Lotero); **Education:** *Evaluation of Student Affairs Services Program* (Neroza); **Agriculture:** *Increasing the Resiliency of Vietnam's Rice Supply Chain* (Khanh); **Governance:** *Efficiency of Local Government Units to the Attainment of MDG* (Baldemor); *Agent-Based Approach to Modeling Poverty in the Philippines* (Marquez); and *OR Investigation into Traffic Congestion in Cities* (Olabode); **Finance:** *Evolution of Trading, OR and Neuroeconomics in Global Finance Instruction* (Dash and Kajiji); and *Loan and Deposit Activities Modeling for Rationalising Commercial Bank Strategy* (Ilchenko and Nazarenko); **Others:** *Developing Agility and Resilience Framework for Humanitarian Supply Chain under the Moderating Effect of Culture* (Ali) and *Unleashing Third Sector Potential: Using Community Operational Research to Develop Guidelines for Volunteering Best Practices in Aotearoa, New Zealand* (Moore).

Gerhard-Wilhelm Weber from METU and EURO used systems and simulation approach to understand factors that influence the quality of primary education in developing countries, particularly, of migrant pupils in big cities [12-13]. With his coworkers from India, Turkey and Indonesia, he evaluated factors such as local and national potential stability, family narration from rural to urban localities, socio-economic status, and their impact on the quality of education.

Overall, participants presented and discussed various interesting models, most of which are focused on solving major problems faced by developing countries in the areas of health, education, poverty, traffic, disasters, agriculture, banking and finance.

3. The 28th Euro Conference

The Workshop flowed into the 28th European Conference on Operational Research, held from July 3 to July 6, 2016. The Conference attracted more than 1800 participants from 170 countries. In keeping with tradition, the EURO Awards were given during the Opening and Closing Sessions.

The event covered 25 areas and 1600 presentations from Data Analytics to Behavioral OR chaired by very competent and intelligent stream and session organizers from different parts of the world. Participants presented papers and attended streams/sessions of their choice to gain more knowledge and skills on how OR could make an impact in their work and in their lives. The talks shared by three plenary and eleven keynote speakers gave the impression that OR has all the answers.

The equally enriching social program included a welcome reception, gala dinner and farewell party, where old friends and new acquaintances met, exchanged ideas and shared past OR experiences. Apart from the tours of museums, the Old Market Square and other places, a trip to the largest fulfillment center of Amazon in Europe was an opportunity to see first hand the logistics involved in getting millions of goods to countries worldwide. This visit turned out to be both a social event and an informative one.

During EURO 2016, Esteemed Scientist Prof. Dr. Robert Aumann of Hebrew University of Jerusalem, Laureate of Nobel Memorial Prize in Economic Sciences in 2005 presented a lecture on *Why Optimize? An Evolutionary Perspective* suggesting that the poorly understood phenomenon of consciousness has evolved for precisely that reason - to enable efficient optimization of life processes.

Conclusion

The IFORS-EWG ORD workshop provided an opportunity for participants to interact on a deeper level on their respective research areas. Each participant gained insights both from those who were knowledgeable in their specific areas of study and from those whose backgrounds served to enrich perspectives at problem solving approaches and implementation considerations. The program was rich with learning opportunities on methods that are very valuable for environments and situations common in developing countries. The EURO Conference, on the other hand, afforded a multi-dimensional environment that provided academic, social and cultural enrichment for all participants. The combination of the small and large conference participation was a unique education towards becoming a well-rounded OR professional.

References

1. D. DeTombe and G.-W. Weber. An Emerging Field in Operational Research: Methodology of Societal Complexity - with a Note on Operational Research and Development, *International Journal Problems of Nonlinear Analysis in Engineering Systems*, vol. 21, n. 1(43), 2015, pp. 36-48.
2. F.A. Crespo, C.S. Pedamallu, L. Özdamar and G.-W. Weber. Contribution to the Collaborative Work in Virtual Organization - a Case Study, *Organizacija- Journal of Management, Information Systems and Human Resources*, vol. 45, n. 5, 2012, pp. 228-235.
3. H. Smith and G.-W. Weber, OR for Developing Countries: Times of Economic, OR News, n. 46, 2012, pp. 60-61.
4. L. Lotero, E. del Rosario and G.-W. Weber, EURO 2015: Showcase of the Continuing Relevance of OR, *International Journal Problems of Nonlinear Analysis in Engineering Systems*, vol. 21, n. 2(44), 2015, pp. 182-185.
5. M. Kljajic, G.-W. Weber and H. Akar, Recent Advances in Systems, Decision Making, Education and other Complex Problems, *Organizacija- Journal of Management, Information Systems and Human Resources*, Vol. 43, n. 3, 2010.
6. O. Nazarenko, K. Pereverza, A. Pasichny, D. Fishman and G.-W. Weber, Initiatives for OR Education, *International Journal Problems of Nonlinear Analysis in Engineering Systems*, vol. 20, n. 1(41), 2014, pp. 130-135.
7. U. Leopold-Wildburger, G.-W. Weber and M. Zachariasen, OR for Better Management of Sustainable Development, *Central European Journal of Operations Research*, vol. 193, 2007, pp. 647-648.
8. E. del Rosario, G. Rand, IFORS: Bringing the World of OR Together for 50 Years, *Wiley Encyclopedia of Operations Research and Management Science*, 2010.

9. C. Brugha. An introduction to the priority-pointing procedure. *Journal of Multicriteria Decision Analysis*, vol. 9 n. 5, 2000, pp 227.
10. R. Du, S. Ai, and C. Brugha. Priority-Pointing Procedure and Its Application to an Intercultural Trust Project. In *Cutting-Edge Research Topics on Multiple Criteria Decision Making*, Springer Berlin Heidelberg, 2009, pp. 296-303.
11. L. Carreras and P. Kaur. Teaching problem structuring methods: Improving understanding through meaningful learning. *INFORMS Transactions on Education*, vol. 12 n. 1, 2011, pp. 20-30.
12. C.S. Pedomallu, L. Ozdamar, E. Kropat and G.-W. Weber. A system dynamics model for intentional transmission of HIV/AIDS using cross impact analysis, *Central European Journal of Operations Research*, vol. 20, n. 2, 2012, pp. 319-336.
13. C.S. Pedomallu, L. Ozdamar, H. Akar, G.-W. Weber and A. Ozsoy, Investigating academic performance of migrant students: a system dynamics perspective with an application to Turkey, *International Journal of Production Economics*, vol. 139, n. 2, 2012, pp. 422-430.

Laura Lotero is a Professor at the Industrial Engineering Faculty of Universidad Pontificia Bolivariana, Colombia. She is a coordinator of the optimization area of the Industrial Engineering Faculty of the aforementioned university. Her interests and work is oriented to urban mobility modeling and complex systems understanding in developing contexts. She is an active supporter of working groups and conferences organized by EURO and IFORS all over the world. laura.loterov@upb.edu.co

Sadia Samar Ali is Professor, Chair for Operations and Supply Chain at New Delhi Institute of Management, New Delhi. For 12 years, she worked to bridge the gap between academia and industry, providing opportunities for learning through operations and supply chain related real life situations. Associated with IFORS and EURO, she is founding Member of Group on Methodology of Societal Complexity (MSC) for Indian Chapter. She does consulting work for organizations in the area of purchase-incidence models; Bayesian analysis; multi-criteria decision-making; reliability functions; supply chain management; logistics management; research evaluation in management. sadiasamarali@gmail.com

Elise del Rosario does consulting work for the public and private sectors in the Philippines and is currently Chief Financial Officer of the family foundation – One Small Step Forward Foundation, Inc. She worked for 20 years with San Miguel Corporation and retired as its Vice President for Operations Research. She was a lecturer at the University of the Philippines Graduate School of Engineering as well as Accenture's Operations Management for BPO Operations MBA Program. She was founding President of the Operations Research Society of the Philippines and President of the International Federation of Operational Research Societies (2007 –2009). elise.del.rosario@stepforward.ph

Gerhard-Wilhelm Weber is a Professor at IAM of METU. With research areas covering financial mathematics, economics, optimization and control, life sciences, data mining, education and development, he also finds time to help people network by organizing numerous international conferences. He received his Diploma and Doctorate in mathematics, and economics and business administration, at RWTH Aachen, and his Habilitation at TU Darmstadt. At IAM, he is currently in the programs of Financial Mathematics and Scientific Computing and a member of five other graduate schools, institutes and departments of METU. He is EURO Conference Advisor and Chair of the OR for Developing Countries Online Resources of IFORS. gweber@metu.edu.tr

Успехи в области исследования операций (Семинар IFORS- EWG ORD и 28 EURO конференции)

Л. Лотеро¹, С. С. Али², Э. дель Росарио³, Г.-В. Вебер⁴

¹Facultad de Ingeniería Industrial, Universidad Pontificia Bolivariana
050031, Medellín, Colombia (Колумбия)

²New Delhi Institute of Management,
Tughlakabad, New Delhi, India (Индия)

³14A CyberOneBldg. Eastwood, C5, 1110 Quezon City, Metro Manila,
Philippines (Филиппины)

⁴Institute of Applied Mathematics, Middle East Technical University
(ODTÜ)06800 Ankara, Turkey (Турция)

Аннотация. 2016 год завершает первое десятилетие деятельности рабочей группы по развитию EWG ORD (EURO Working Group on OR for Development). Совместный семинар IFORS-EWG ORD Workshop проводился вместе с IFORS (International Federation of Operational Research Societies) непосредственно перед конференцией EURO 2016 (European Conference on Operation Research); (EURO – The Association of European Operation Research Societies) в г.Познань. На Семинаре были представлены методы картографирования и расстановки приоритетов, которые применялись для получения информации о понимании участниками задачи о том, как сделать исследование операций более эффективным в решении вопросов развития. Было представлено большое количество статей, использующих методы операционного исчисления в областях здравоохранения, образования, сельского хозяйства, нищеты, дорожного движения, гуманитарной логистики и финансов.

Ключевые слова: исследование операций, развивающиеся страны. рабочая группа по развитию.

1. Введение

Десятый, юбилейный, Семинар Рабочей группы EURO по исследованию операций для развития (EWG ORD) проходил под названием «Десять лет работы в области исследования операций для развития» (<http://ifors.org/ewg-ord/>). Спонсорами Семинара стали EWG ORD (<https://www.euro-online.org/web/ewg/29/or-for-development-ewg-ord>) и IFORS (<http://ifors.org/web/>). Семинар был мероприятием, сопутствующим 28 конференции EURO, проводившейся в г.Познань.

Со времени своего основания в 2007 г. данный семинар был площадкой для общения, углубления знаний и обмена опытом специалистов в прикладной области, ученых, всего мирового сообщества по исследованию операций. На семинарах обсуждаются вопросы развития в области образования, здравоохранения, инфраструктуры, устойчивости окружающей среды, энергетики, роста населения, нищеты, изменения климата и т.д. [1–7].

За каждой 30-минутной презентацией на Семинаре было организовано выступление заранее отобранных специалистов, которые высказывали свое мнение о доложенной работе. Помимо повышения навыков выступления, участники имели возможность получить отзывы ученых, практиков, студентов и более опытных специалистов. После

Семинара участники продолжили беседы и обмен идеями уже непосредственно во время конференции.

2. Разнообразный спектр прикладных областей Семинара EWG-ORD

На открытии Семинара своим обширным опытом в области исследования операций поделилась председатель EWG ORD Элис дель Росарио. Она рассказывала о трех проектах: глобализация средств производства, национальное управление энергосистемой, операции в городском порту. Эти проекты дали четкое представление о проблемах и успехах, которые могут ждать специалиста в области исследования операций в реальной жизни. Из них были извлечены уроки как успешного, так и неуспешного применения рекомендаций [8].

За этим выступлением следовало заседание под руководством Катал Бруга, заслуженного профессора и внештатного преподавателя University College Dublin Center for Business Analytics. Участники семинара пытались найти ответ на вопрос *«как помочь исследованию операций быть более эффективным в решении проблем развития?»* [9, 10], а ведущий семинара продемонстрировал, как методика расстановки приоритетов может привести к разработке стратегии.

Эшли И. Каррерас, ведущий лектор в области экономики и теории принятия решений в университете De Montfort University, провел тьюторское занятие на тему «Научно-практическая точка зрения на методы структурирования задачи». Участники разделились на две группы и с помощью методов отображения выясняли, каковы ключевые моменты, которые необходимо решить, чтобы исследование операций продолжало эффективно влиять на политических деятелей [11]. Затем Э.И. Каррерас обобщил итоги работы обеих групп.

Представленные на Семинаре работы тщательно отбирались из множества заявок. Рассматривались разные области и вопросы.

Здравоохранение: Проектирование системы сбора крови; модель размещения-распределения с учетом сбора и подготовки (Осорио) и Городская мобильность и динамика распространения болезней в г.Медельин, Колумбия (Лотеро). *Образование:* Оценка программы помощи студентам (Нероза). *Сельское хозяйство:* Повышение отказоустойчивости цепочки поставок риса во Вьетнаме (Хан). *Управление:* Эффективность органов местного самоуправления в достижении Целей в рамках развития MDG (Millennium Development Goals) (Балдемор); Моделирование бедности на Филиппинах с использованием метода Agent – Based Approach (Маркез); Исследование операций в применении к заторам на дорогах крупных городов (Олабоде). *Финансы:* Эволюция торговли, исследование операций и нейроэкономика в глобальной финансовой программе (Дэш, Каджиджи); Моделирование деятельности банков в области вкладов и кредитования для рационализации стратегии коммерческих банков (Ильченко, Назаренко). *Другое:* Разработка оперативной и отказоустойчивой базы для организации гуманитарных поставок с учетом смягчающего влияния культуры (Али); Использование потенциала третьего сектора экономики: применение исследования операций применительно к развитию общества с целью разработки рекомендаций для волонтерской практики в Аотеароа, Новая Зеландия (Мур).

Герхард-Вильгельм Вебер из METU и EURO применил системный подход и моделирование для выяснения, какие факторы влияют на качество начального образования в развивающихся странах, в особенности в отношении детей переселенцев в крупных городах [12-13]. Совместно с коллегами из Индии, Турции и Индонезии он оценил вли-

яние на качество образования таких факторов, как локальная и национальная потенциальная стабильность, переезд семей из сельской в городскую местность, социально-экономический статус.

В целом, участники представили и обсудили разнообразные интересные модели, нацеленные на решение основных проблем, с которыми сталкиваются развивающиеся страны в области здравоохранения, образования, нищеты, дорожного движения, природных бедствий, сельского хозяйства, банков, финансов.

3. 28-я конференция EURO

Работа Семинара плавно перетекла в работу 28-й Европейской конференции по исследованию операций, которая проходила 3-6 июля 2016 г. Конференция собрала более 1800 участников из 170 стран. По традиции награды EURO вручались на открытии и закрытии конференции.

На конференции было представлено 1600 презентаций, затрагивающих 25 различных областей от анализа баз данных до исследования проблем мотивации поведения методами опрационального исчисления. Заседания секций были организованы под руководством квалифицированных специалистов из разных частей света. Участники представили доклады и посетили симпозиумы по своему выбору для обмена опытом и навыками в области методов исследования операций и их значения в работе и жизни. Три пленарных и одиннадцать приглашенных докладов оставили впечатление, что исследование операций располагает всеми необходимыми ответами на актуальные вопросы.

Культурная программа конференции была также чрезвычайно насыщенной и включала приветственный прием, торжественный ужин и прощальный ужин, где встречались старые друзья и совершались новые знакомства, происходил обмен идеями и опытом в области исследования операций. Помимо экскурсии в музеи, на старую рыночную площадь и другие достопримечательности, была организована поездка в крупнейший центр исполнения заказов Интернет-магазина Amazon в Европе. Участники имели возможность своими глазами увидеть логистику поставок миллионов товаров во все страны мира. Эта экскурсия оказалась не только культурным, но и чрезвычайно познавательным мероприятием.

Во время конференции EURO 2016 известный ученый, лауреат премии памяти Альфреда Нобеля в области экономики (2005 г.), профессор, доктор Роберт Ауманн из Еврейского университета (г. Иерусалим) прочитал лекцию на тему «Зачем оптимизировать? Эволюционная точка зрения», в которой предположил, что плохо понимаемый феномен сознания эволюционировал по этой причине – необходима эффективная оптимизация жизненных процессов.

Выводы

Семинар IFORS-EWG ORD стал площадкой взаимодействия участников, специалистов в области исследования операций, с углублением уровня понимания. Каждый участник мог узнать мнение о своей работе специалистов из соответствующей области исследования операций, а также специалистов, способных расширить представления о методах решения задач и вопросах внедрения. Программа Семинара способствовала обучению участников чрезвычайно полезным в развивающихся странах методам исследования. Конференция EURO, с другой стороны, предлагала площадку для полезных научных, социальных и

культурных дискуссий. Участие в небольшом и крупном мероприятиях одновременно давало уникальную возможность совершенствования в области исследования операций.

Список литературы

1. D. DeTombe and G.-W. Weber. An Emerging Field in Operational Research: Methodology of Societal Complexity - with a Note on Operational Research and Development, *International Journal Problems of Nonlinear Analysis in Engineering Systems*, vol. 21, n. 1(43), 2015, pp. 36-48.
2. F.A. Crespo, C.S. Pedomallu, L. Özdamar and G.-W. Weber. Contribution to the Collaborative Work in Virtual Organization - a Case Study, *Organizacija- Journal of Management, Information Systems and Human Resources*, vol. 45, n. 5, 2012, pp. 228-235.
3. H. Smith and G.-W. Weber, OR for Developing Countries: Times of Economic, OR News, n. 46, 2012, pp. 60-61.
4. L. Lotero, E. del Rosario and G.-W. Weber, EURO 2015: Showcase of the Continuing Relevance of OR, *International Journal Problems of Nonlinear Analysis in Engineering Systems*, vol. 21, n. 2(44), 2015, pp. 182-185.
5. M. Kljajic, G.-W. Weber and H. Akar, Recent Advances in Systems, Decision Making, Education and other Complex Problems, *Organizacija- Journal of Management, Information Systems and Human Resources*, Vol. 43, n. 3, 2010.
6. O. Nazarenko, K. Pereverza, A. Pasichny, D. Fishman and G.-W. Weber, Initiatives for OR Education, *International Journal Problems of Nonlinear Analysis in Engineering Systems*, vol. 20, n. 1(41), 2014, pp. 130-135.
7. U. Leopold-Wildburger, G.-W. Weber and M. Zachariasen, OR for Better Management of Sustainable Development, *Central European Journal of Operations Research*, vol. 193, 2007, pp. 647-648.
8. E. del Rosario, G. Rand, IFORS: Bringing the World of OR Together for 50 Years, *Wiley Encyclopedia of Operations Research and Management Science*, 2010.
9. C. Brugha. An introduction to the priority-pointing procedure. *Journal of Multicriteria Decision Analysis*, vol. 9 n. 5, 2000, pp 227.
10. R. Du, S. Ai, and C. Brugha. Priority-Pointing Procedure and Its Application to an Intercultural Trust Project. In *Cutting-Edge Research Topics on Multiple Criteria Decision Making*, Springer Berlin Heidelberg, 2009, pp. 296-303.
11. L. Carreras and P. Kaur. Teaching problem structuring methods: Improving understanding through meaningful learning. *INFORMS Transactions on Education*, vol. 12 n. 1, 2011, pp. 20-30.
12. C.S. Pedomallu, L. Ozdamar, E. Kropat and G.-W. Weber. A system dynamics model for intentional transmission of HIV/AIDS using cross impact analysis, *Central European Journal of Operations Research*, vol. 20, n. 2, 2012, pp. 319-336.
13. C.S. Pedomallu, L. Ozdamar, H. Akar, G.-W. Weber and A. Ozsoy, Investigating academic performance of migrant students: a system dynamics perspective with an application to Turkey, *International Journal of Production Economics*, vol. 139, n. 2, 2012, pp. 422-430.

Лаура Лотеро, профессор факультета промышленной технологии университета Universidad Pontificia Bolivariana, Колумбия. Координатор деятельности по оптимизации факультета. Область интересов: моделирование городской мобильности, понимание

сложных систем в контексте развития. Активно поддерживает функционирование рабочих групп и конференций, организованных EURO и IFORS по всему миру.

Садия Самар Али, профессор института менеджмента, г.Нью-Дели, Индия. В течение 12 лет работала над сближением науки и промышленности, создавая возможности для обучения персонала на реальных примерах из области управления операциями и цепочками поставок. Является членом рабочей группы индийского филиала по методологии социальной комплексности. Консультирует организации в области моделей закупок – реализации, байесовского анализа, многокритериального процесса принятия решений, функций надежности, управления цепочкой поставок, управления логистикой, оценки исследования в маркетинге.

Элиса дель Росарио, консультант государственного и частного сектора на Филиппинах, финансовый директор семейного фонда One Small Step Forward Foundation, Inc. В течение 20 лет работала в San Miguel Corporation; будучи в должности вице-президента этой корпорации в области исследования операций, перешла на работу в университет. Преподает в Филиппинском университете, Graduate School of Engineering, а также в Accenture's Operations в рамках программы MBA по аутсорсингу бизнес-процессов. Президент-основатель сообщества по исследованию операций на Филиппинах, Президент Международной федерации сообществ по исследованию операций (2007-2009).

Герхард-Вильгельм Вебер, профессор Института прикладной математики Ближневосточного технического университета (METU). Область интересов: финансовая математика, экономика, оптимизация и управление, медико-биологические науки, глубинный анализ данных, образование и развитие. Организатор многочисленных международных конференций. Имеет диплом и докторскую степень по математике, экономике и деловому администрированию Рейнско-Вестфальского технического университета Ахена, прошел переподготовку в Дармштадтском техническом университете. В Институте прикладной математики на данный момент участвует в программах по финансовой математике и научным вычислениям; член пяти других школ магистратуры и докторантуры, институтов и кафедр METU. Советник конференций EURO, возглавляет онлайн-ресурс OR for Developing Countries (исследование операций для развивающихся стран) IFORS.

International scientific Journal "Problems of nonlinear Analysis in Engineering Systems" is the periodic Journal, that is founded in 1994 by Russian Scientists, representatives of Kazan Chetayev School of Mechanics and Stability, jointly with foreign Colleagues, is published under aegis of International Federation of nonlinear Analysts and Academy of nonlinear Sciences together with Kazan National Research Technical University of A.N.Tupolev name (Kazan Aviation Institute). It is interdisciplinary scientific Edition, presenting the works on nonlinear problems in all areas of fundamental and applied Sciences, including both natural and humanities disciplines: mathematics, mechanics, physics, chemistry; engineering, biological, medical, social, political sciences; ecology, cosmology, economics; nanoscience and nanotechnology; stability and sustainable development, problems of risk and information protection, operations research, ...

Scientists of different fields are invited for cooperation.

<http://pnaes.kpfu.ru/science/journals/ansj/>

http://www.kcn.ru/tat_en/science/ans/journals/ansj.html

Authors should send their manuscripts (3 clean copies, ~15pp.), prepared to the publication, and a disk (MS Word for Windows, IBM PC). It is also possible to duplicate the submitted paper via e-mail.

The text should be printed on A4 size paper within the margins of 160x235 (mm) (including the title, the author name and affiliation, and the contact address), Times New Roman font, 12pt, single space. Upper margin is 35mm, left margin is 25mm, right margin is 25mm. Illustrations are supposed to be in editable formats of .jpg, .gif, .bmp and placed within the same margins. Pagination should be made on the reverse side by pencil. Short information about the author (3-4 lines) should be given at the end of the paper (covering the area of scientific interests and spheres of application).

The paper should be accompanied with an abstract (2p) prepared according to the same guidelines and printed on separate pages. Authors of papers in French, or German, or Russian should also submit in English version of their article and abstracts (2p) in Russian and English.

Authors reserve their right to copy their publication. The Journal can be sent to the author on request for separate payment or by subscription.

Our contacts (on publications, advertisement or business propositions) -

(7) (843) 236-16-48 Lyudmila K. Kuzmina

(7) (843) 238-44-20 Vladimir I. Panchenko

Address:

L.K.Kuzmina, Kazan National Research Technical University of A.N.Tupolev name
(KNRTU-KAI)
Adamuck, 4-6, Kazan-15, 420015, RUSSIA
Lyudmila.Kuzmina@kpfu.ru

*The Journal has been cataloged:
in Congress Library; the Library of Congress Catalog Number (LCCN) is **98-646147**
in British Library; the British Library Catalog Number (LCCN) is **0133.473700***

Published papers are reviewed in abstract Journal and abstract database of RAS All-Russian Institute of Scientific-Engineering Information
Information about Edition is entered in reference system on periodic Editions "Ulrich's Periodicals Directory" <http://www.ulrichsweb.com>

Online version of Scientific Edition is implemented in cooperation with Kazan Federal University and is available at KFU-server

Edition is carried out with support of ABAK Operating Printing Center

Original-model is prepared for printing
by Humanity Projects and Investigations Center
together with

Foundation of culture development support
under TATARSTAN Republic President
Publishing House **ABAK** (licence No.0195; 03.08.2000)
Kazan, RUSSIA

Международный научный журнал “Проблемы нелинейного анализа в инженерных системах”, основанный в 1994 году по инициативе российских ученых, представителей Казанской Четаевской Школы механики и устойчивости, совместно с зарубежными коллегами, издается под эгидой Международной Федерации Нелинейных Аналитиков, Академии Нелинейных Наук, с участием Казанского национального исследовательского технического университета им.А.Н.Туполева (Казанского Авиационного Института). “ПНАИС” - междисциплинарное научное Издание, представляющее исследования по нелинейным проблемам в целом, во всем многообразии фундаментальных и прикладных наук естественнонаучного и гуманитарного цикла, среди которых: математика, механика, физика, химия; инженерные, биологические, медицинские, социальные, политические науки; экология, космология, экономика; нанонаука и нанотехнология; безопасность и поддерживающее развитие, проблемы риска и защиты информации, исследование операций, ... Приглашаются к сотрудничеству ученые различных областей.

<http://pnaes.kpfu.ru/science/journals/ansj/>

http://www.kcn.ru/tat_en/science/ans/journals/ansj.html

Авторы присылают свои работы, подготовленные к изданию, в трех экземплярах, объемом до 15 страниц (необходима четкая, качественная печать), и дискету (Word for Windows, IBM PC). Возможна дублирующая пересылка по электронной почте.

Формат набора на листе А4: поле текста - 160х235 (мм), включая в начале статьи: заголовок, инициалы и фамилию автора, название института (организации), адрес для контакта. Верхнее поле - 35 мм, левое - 25 мм, правое - 25мм. Основной текст Times New Roman, 12pt, через 1 интервал, рисунки в компьютерном исполнении (файлы .bmp, .jpg, .gif). Страницы нумеруются карандашом на обратной стороне листа, иллюстрации размещаются в тех же размерах. В конце статьи необходимо дать сведения об авторе (3-4 строки), включая область научных интересов, сферы приложения.

Необходима аннотация статьи на языке оригинала (2 стр.), напечатанная по тем же правилам, а также- аннотация (2 стр.) и вариант статьи на английском языке.

За автором сохраняется право копирования своей публикации. Журнал может быть выслан по заказу за отдельную плату или по подписке.

Наши координаты для контактов (по вопросам публикаций, рекламы и деловых предложений)

(7) (843) 236-16-48 Людмила Константиновна Кузьмина

(7) (843) 238-44-20 Владимир Иванович Панченко

Адрес:

Л.К.Кузьмина, Казанский национальный исследовательский технический университет
им.А.Н.Туполева (КНИТУ-КАИ)
Адамюк, 4-6, Казань-15, 420015, РОССИЯ
Lyudmila.Kuzmina@kpfu.ru

Журнал внесен:

*в Каталог Библиотеки Конгресса США; номер в Каталоге (LCCN) - 98-646147
в Каталог Британской Библиотеки; номер в Каталоге (LCCN) - 0133.473700*

Опубликованные статьи реферируются в Реферативном журнале и Базе данных ВИНТИ РАН
Сведения о журнале занесены в справочную систему по периодическим изданиям “Ulrich's Periodicals Directory” <http://www.ulrichsweb.com>
Электронный вариант Научного Издания выполняется в кооперации с Казанским Федеральным Университетом и размещен на сервере КФУ

Издание осуществляется при поддержке Центра Оперативной Печати АБАК

Оригинал-макет подготовлен для печати
Центром Гуманитарных Проектов и Исследований
совместно с
Фондом поддержки развития культуры
при Президенте Республики ТАТАРСТАН
Издательство «АБАК» (лицензия № 0195; 03.08.2000)
Казань, РОССИЯ

СОДЕРЖАНИЕ

А.Феттвейс

Отсутствие потерь в нелинейных цепях Кирхгофа и в теории относительности

О.Г.Морозов, Г.И.Ильин, Г.А.Морозов

Амплитудно-фазовая модуляция лазерного излучения в задачах генерации радиочастотных сигналов

В.В.Березин, А.М.Щитов, А.С.Нечаев

Автоматизированное проектирование широкополосного диодного удвоителя частоты в диапазоне (26-40) ГГц

В.А.Бурдин, А.В.Бурдин

Моделирование нелинейного распространения мод в многомодовых волоконно-оптических линиях дальней связи

М.В.Левский

Об условиях трансверсальности в решении вариационных задач

В.И.Панченко, Ли Цзывань

Сравнение различных двигателей и методы повышения термического КПД

П.Г.Великанов, Чжан Бин

О конвертировании авиационных двигателей, отработавших ресурс в лётной эксплуатации

Н.П.Великанова, Фань Вэйвэй

К проблеме жаропрочности рабочих лопаток турбины авиационных ГТД

А.А.Болонкин

Как организовать науку

НАУЧНО-ИНФОРМАЦИОННЫЙ РАЗДЕЛ

М.А.Ильгамов

О 70-летию исследований по механике в Казанском Научном Центре РАН

Л.Лотеро, С.С.Али, Э. дель Росарио, Г.-В.Вебер

Успехи в области исследования операций

CONTENTS

1 A.Fettweis

Losslessness in nonlinear Kirchhoff circuits and in relativity theory

44 O.G.Morozov, G.I.Ilin, G.A.Morozov

Amplitude-phase modulation of laser radiation in generation problems of radio-frequency signals

70 V.V.Berezin, A.M.Schitov, A.S.Nechaev

Computer-aided design of a broadband diode frequency doubler in the range (26-40) GHz

89 V.A.Burdin, A.V.Bourdine

Modelling nonlinear propagation of modes in long-haul multimode fiber optic links

104 M.V.Levskii

About transversality conditions for solving variation problems

124 V.I.Panchenko, Li Ziwan

Comparison of different engines and the methods of increasing the thermal efficiency

137 P.G.Velikanov, Zhang Bing

About conversion of aircraft engines after life cycle in the flight operation

152 N.P.Velikanova, Fan Weiwei

To heat-temperature strength problem of aviation gas turbine engine blades

163 A.A.Bolonkin

To problem of science organization

SCIENTIFIC-INFORMATION SECTION

181 M.A.Ilgamov

On 70-th Anniversary of investigations on mechanics in Kazan Science Center of RAS

197 L.Lotero, S.S.Ali, E. del Rosario, G.-W.Weber

Advances in operational research for development