DISCRETE AND CONTINUOUS MODELS AND APPLIED COMPUTATIONAL SCIENCE

Volume 28 Number 4 (2020)

Founded in 1993
Founder: PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA

DOI: 10.22363/2658-4670-2020-28-4

Edition registered by the Federal Service for Supervision of Communications, Information Technology and Mass Media
Registration Certificate: ПИ № ФС 77-76317, 19.07.2019

ISSN 2658-7149 (online); 2658-4670 (print)
4 issues per year.
Language: English.

Publisher: Peoples' Friendship University of Russia (RUDN University).
Indexed in Ulrich's Periodicals Directory (http://www.ulrichsweb.com), in https://elibrary.ru, EBSCOhost (https://www.ebsco.com), CyberLeninka (https://cyberleninka.ru).

Aim and Scope

Discrete and Continuous Models and Applied Computational Science arose in 2019 as a continuation of RUDN Journal of Mathematics, Information Sciences and Physics. RUDN Journal of Mathematics, Information Sciences and Physics arose in 2006 as a merger and continuation of the series "Physics", "Mathematics", "Applied Mathematics and Computer Science", "Applied Mathematics and Computer Mathematics".

Discussed issues affecting modern problems of physics, mathematics, queuing theory, the Teletraffic theory, computer science, software and databases development.

It's an international journal regarding both the editorial board and contributing authors as well as research and topics of publications. Its authors are leading researchers possessing PhD and PhDr degrees, and PhD and MA students from Russia and abroad. Articles are indexed in the Russian and foreign databases. Each paper is reviewed by at least two reviewers, the composition of which includes PhDs, are well known in their circles. Author's part of the magazine includes both young scientists, graduate students and talented students, who publish their works, and famous giants of world science.

The Journal is published in accordance with the policies of COPE (Committee on Publication Ethics). The editors are open to thematic issue initiatives with guest editors. Further information regarding notes for contributors, subscription, and back volumes is available at http://journals.rudn.ru/miph.

E-mail: miphj@rudn.ru, dcm@sci.pfu.edu.ru.

EDITORIAL BOARD

Computer Design: *A. V. Korolkova, D. S. Kulyabov*

Address of editorial board:
Ordzhonikidze St., 3, Moscow, Russia, 115419
Tel. +7 (495) 955-07-16, e-mail: publishing@rudn.ru
Editorial office:
Tel. +7 (495) 952-02-50, miphj@rudn.ru, dcm@sci.pfu.edu.ru
site: http://journals.rudn.ru/miph

Paper size 70×100/16. Offset paper. Offset printing. Typeface "Computer Modern".
Conventional printed sheet 8,55. Printing run 500 copies. Open price. The order 1040.
PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA
6 Miklukho-Maklaya St., 117198 Moscow, Russia
Printed at RUDN Publishing House:
3 Ordzhonikidze St., 115419 Moscow, Russia,
Ph. +7 (495) 952-04-41; e-mail: publishing@rudn.ru

Discrete & Continuous Models
& Applied Computational Science
2020, 28 (4) 304
http://journals.rudn.ru/miph

Contents

Discrete & Continuous Models
& Applied Computational Science

2020, 28 (4) 305–318
http://journals.rudn.ru/miph

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2020-28-4-305-318

Queueing systems with different types of renovation mechanism and thresholds as the mathematical models of active queue management mechanism

Hilquias Viana Carvalho Cravid[1], **Ivan S. Zaryadov**[1,2], **Tatiana A. Milovanova**[1]

[1] *Peoples' Friendship University of Russia (RUDN University)*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation
[2] *Institute of Informatics Problems, FRC CSC RAS*
44-2, Vavilova St., Moscow 119333, Russian Federation

(received: October 27, 2020; accepted: November 12, 2020)

This article is devoted to some aspects of using the renovation mechanism (different types of renovation are considered, definitions and brief overview are also given) with one or several thresholds as the mathematical models of active queue management mechanisms.

The attention is paid to the queuing systems in which a threshold mechanism with renovation is implemented. This mechanism allows to adjust the number of packets in the system by dropping (resetting) them from the queue depending on the ratio of a certain control parameter with specified thresholds at the moment of the end of service on the device (server) (in contrast to standard RED-like algorithms, when a possible drop of a packet occurs at the time of arrivals of next packets in the system).

The models with one, two and three thresholds with different types of renovation are under consideration. It is worth noting that the thresholds determine not only from which place in the buffer the packets are dropped, but also to which the reset of packets occurs. For some of the models certain analytical and numerical results are obtained (the references are given), some of them are only under investigation, so only the mathematical model and current results may be considered.

Some results of comparing classic RED algorithm with renovation mechanism are presented.

Key words and phrases: random early detection, active queue management, queuing system, general renovation, threshold mechanism, drop function, congestion control

1. Introduction

In modern communication networks the problem of congestion avoidance does not have a satisfying solution [1] and the development of active queue management (AQM) algorithms appears to be the actual task for researches and practitioners.

A numerous number of AQM schemes have been proposed [2], some of them were investigated and standardised by IETF working group on "Active Queue Management and Packet Scheduling" [3].

For the most AQM algorithms models the performance analysis is performed by simulation (for example, [2]) and that is why the bridges between the available use-case results and the available analytic results are very few (see, for example, [4]–[7]).

In this paper the mathematical models of RED-like (in contrast to standard RED algorithm, when a possible reset occurs at the time of the next packet arrival and the control parameter is an exponentially weighted average queue length [8]–[10], in our models the decision about a possible packet drop is made at the momemnts of service completions) algorithms with renovation and one or several thresholds (which determine not only the place in the buffer from which the packets are dropped, but also the place to which the reset of packets occurs).

The structure of the article is following. The section 2 gives the brief description active queue algorithms, especially of the classic RED algorithm. The section 3 consists of foyr subsections: in 3.1 the definitions of different types of renovation are given as well as a brief overview of queueing systems with renovation; in 3.2 the models with renovation and a single thresholds are described; the subsectionn 3.3 is devoted to models with renovation and two thresholds (for one of these models some analitycal results are presented); the last subsection 3.3 describes renovation models with three thresholds. In section 4 the comparison of results for RED algorithm and some renovation models is presented, based on experimental results from [11]–[13]. These results confirm assumptions that the use of the renovation mechanism in the single server queues under heavy overload conditions and some other constraints allows one to achieve at least the same performance level as by the classical random early detection algorithm. The last section 5 concludes the paper with the short discussion.

2. Active queue management algorithms. The brief description of RED algorithm module

According to RFC 7567 [1] active queue management (AQM) is considered as a best practice of network congestion avoidance (reducing) in Internet routers. The active queue management is the policy of dropping packets inside a buffer associated with a network interface controller (NIC) before that buffer becomes full (or gets close to becoming full) and this policy is based on some rules (algorithms) such as:

— Random Early Detection (random early discard or random early drop) (RED) [8]–[10], [14] — is a queuing discipline for a network scheduler suited for congestion avoidance by pre-emptively dropping packets before

the buffer becomes completely full based on predictive models to decide which packets to drop;
— Explicit Congestion Notification (ECN) [9], [15] and its modifications [16]–[19] — allows end-to-end notification of network congestion without dropping packets (opposed to RED);
— controlled delay (CoDel) [20] and its modifications [21], [22] — a scheduling algorithm for the network scheduler, designed to overcome bufferbloat in networking hardware by setting limits on the delay network packets experience as they pass through buffers in this equipment;
— BLUE [23] and its modifications [24]–[26] — operates by randomly dropping or marking packet with explicit congestion notification mark before the transmit buffer of the network interface controller overflows and it requires little or no tuning to be performed by the network administrator;
— CAKE (Common Applications Kept Enhanced)[27], [28] — is a shaping-capable queue discipline which uses both AQM and FQ, itcombines COBALT, which is an AQM algorithm combining Codel and BLUE

Early AQM disciplines (notably RED and SRED) require careful tuning of their parameters in order to provide good performance, but modern AQM disciplines (Blue, CoDel, CAKE and new modifications of RED — the overview may be seen in [29], [30]) are self-tuning, so they can be run with their default parameters.

The classic RED [8] is an active queue management algorithm with two thresholds (Q_{min} and Q_{max}), which monitors the average queue size and drops packets (or marks packets when it used in conjunction with ECN) based on statistical probabilities $p(\hat{Q})$ [8]:

$$p(\hat{Q}) = \begin{cases} 0, & 0 \leqslant \hat{Q} \leqslant Q_{\min}, \\ \dfrac{\hat{Q} - Q_{\min}}{Q_{\max} - Q_{\min}} p_{\max}, & Q_{\min} < \hat{Q} \leqslant Q_{\max}, \\ 1, & \hat{Q} > Q_{\max}, \end{cases}$$

where $p_{\max}$ is the maximum of dropped probability (probability of marking), and $\hat{Q}$ is the queue length control parameter (low-pass filter) calculated by formula:

$$\hat{Q}_{k+1} = (1 - w_q)\hat{Q}_k + w_q \hat{Q}_k, \quad k = 0, 1, 2, \ldots,$$

here w_q $(0 < w_q < 1)$ is a weight coefficient of the exponentially weighted moving-average.

3. Renovation mechanism with thresholds

3.1. The definition of renovation mechanism

The renovation mechanism was introduced in [31]: at the moment of the end of its service the packet on the server may either just leave the system with some non-zero probability p, or may empty the buffer with the renovation

probability $q = 1 - p$. In [31] the steady-state probability distributions for several types of queueing systems were presented.

Later the model of renovation with feedback (or repeated service) was introduced in 2007 by professor P. P. Bocharov in [32]. The renovation with feedback means that after emptying the buffer with probability q the served packet starts another round of service. The main characteristics in matrix-analytical form were obtained. Another system with renovation and feedback was studied in [33], [34], where the retrial queueing system with renovation and recurrent input flow was investigated.

The generalisation of renovation mechanism was proposed in 2008 by professor A. V. Pechinkin [35]: at the moment of the end of service the packet may drop from the buffer with probability $q(i)$ exactly i ($i \geqslant 1$) other packets and leaves the system or may just just leave the system without any effect on it with the complementary probability $p = 1 - \sum_{i=1}^{r} q(i)$.

In [12], [13], [35]–[39] queueing systems with different types of renovation and service disciplines were presented. It was proved that for $Gi/M/n/\infty$ systems with general renovation (or just renovation) the steady-state probability distribution has the geometric form and sojourn time distribution of a served (lossed) packet has an exponential form.

The study of queueing systems with renovation mechanism and thresholds was started in [12], [13].

3.2. The mathematical model with renovation and one threshold

The simplest mathematical models of renovation mechanism with thresholds are the models with one threshold Q_1 for queue length.

If the current queue length $\hat{q}$ is larger than Q_1 ($\hat{q} > Q_1$) then a packet on the server at the moment of the end of its service may either leave the system with probability p or may drop a packet (packets) from the buffer and leave (renovation mechanism).

The next types of renovation may be considered:

— the "simple" renovation of the first type — the fixed size group of k ($k \geqslant 1$) packets will be dropped with fixed probability q_k;
— the "simple" renovation of the second type — the fixed size group of k ($k \geqslant 1$) packets will be dropped with fixed probability q_k (but Q_1 will remain anyhow)
— the "complete" ("full") renovation (or just renovation) — all packets from the buffer will be dropped with probability $q = 1 - p$, the system will be empty;
— the "incomplete" renovation — all packets (that are after the threshold) willbe dropped with probability $q = 1 - p$, Q_1 packets will remain in the system;
— the general renovation of the first type — exactly k ($k \geqslant 1$) packets from the buffer will be dropped with probability $q(k)$ if in the buffer there are more then k packets or the system will be empty if in the buffer were were less (or equal) then k packets;
— the general renovation of the second type — exactly k ($k \geqslant 1$) packets from the buffer may be dropped with probability $q(k)$ if there were more

then $k+Q_1$ packets in the buffer, or Q_1 packets will remain in the system if in the buffer were less (or equal) then $k+Q_1$ packets.

For the last two cases of general renovation the following condition on probabilies p and $q(k)$ $(k \geqslant 1)$ must be met:

$$p + \sum_{k=1}^{\infty} q(k) = 1,$$

p is the probability, that a served packet will just leave the system without dropping other packets.

The minus of the first, the third and the fifth drop mechanism is that too many packets may be dropped, the minus of the second, the fourth and the sixth drop mechanism is that the buffer may remain overflowed. Models with two thresholds allow to get rid of these minuses.

3.3. The mathematical model with renovation and two thresholds

In this part of the article we will discuss models of RED-like algorithm based on queueing systems with renovation and two thresholds.

The thresholds of the first type: the first threshold Q_1 determines when packets start to be dropped, the second threshold $Q_2(Q_2 > Q_1)$ sets the maximum value of drop probability.

The types of renovation policy. At the moment of the end of a packet service the current queue length $\hat{q}$ is compared with thresholds and if $\hat{q} \leqslant Q_1$ then no one of the packets from the buffer is dropped.

— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k $(k \geqslant 1)$ packets will be dropped with fixed probability q_k. If $\hat{q} \geqslant Q_2 + 1$ then the fixed size group of k packets in the buffer will be dropped with maximal probability $q_{\max}(q_k < q_{\max} \leqslant 1$.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k packets will be dropped with probability $q_k(\hat{q})$, $0 < q_k(\hat{q}) < q_{\max}$. If $\hat{q} \geqslant Q_2 + 1$ then the fixed size group of k packets in the buffer will be dropped with maximal probability $q_{\max}$ (this model was considered in [12] for $k = 1$ (the last packet is dropped) and the steady-state probability distribution of packets in the system(for imbedded Markov chain) as well as some probabilistic (the probability that the arriving packet will be dropped or served) and sojourn time characteristics for served packets were obtained).
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k packets will be dropped with fixed probability q_k. If $\hat{q} \geqslant Q_2 + 1$ then all packets except the first Q_1 will be dropped with probability q.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_1$ then the fixed size group of k packets will be dropped with probability $q_k(\hat{q})$. If $\hat{q} \geqslant Q_2 + 1$ then all packets except the first Q_1 will be dropped with probability q.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the arbitrary size group of k packets will be dropped with given probability $q(k)$ $(0 < q(k) < 1)$. If $\hat{q} \geqslant Q_2 + 1$ then all packets except the first Q_1 will be dropped with probability q.

For all types of renovation policy the following condition on probabilies q_k, $q_k(\hat{q})$, $q(k)$, $q_{\max}$, q $(k \geqslant 1)$ must be met:

$$p_k + q_k = 1, \quad p_k(\hat{q}) + q_k(\hat{q}) = 1, \quad k \geqslant 1, \quad p + \sum_{k=1}^{\infty} q(k) = 1,$$

where p_k, $p_k(\hat{q})$ $(k \geqslant 1)$ and p are the probabilities that the served packet will leave the system without dropping any other packets from the buffer (if $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$), and $p^{(2)}_{\max} + q_{\max} = 1$, $p^{(2)} + q = 1$, $p^{(2)}_{\max}$ and $p^{(2)}$ are the probabilities that the served packet will leave the system without dropping any other packets from the buffer (if $\hat{q} \geqslant Q_2 + 1$).

The thresholds of the second type: the first threshold Q_1 determines when packets start to be dropped, the second threshold Q_2 determines the place to which packets will be dropped from the queue $(Q_1 \geqslant Q_2)$. If the current queue size $\hat{q}$ is less than Q_1, then a served packet just leave the system. But if $\hat{q} \geqslant Q_1 + 1$ then the following types of renovation (dropping) mechanism may be applied:

— the fixed size group of k $(k \geqslant 1)$ packets will be dropped with fixed probability q_k (anyhow Q_2 packets will remain in system);
— with probability q only Q_2 packets will remain in system, all others packets will be dropped;
— the arbitrary size group of k packets will be dropped with given probability $q(k)$ $(0 < q(k) < 1)$ (anyhow Q_2 packets will remain in system).

For all types of renovation policy the following condition on probabilies q_k, $q_k(\hat{q})$ $(k \geqslant 1)$ and q must be met:

$$p_k + q_k = 1, \quad p + \sum_{k=1}^{\infty} q(k) = 1, \quad p^{(2)} + q = 1,$$

where p_k $(k \geqslant 1)$, p and $p^{(2)}$ are the probabilities that the served packet will leave the system without dropping any other packets from the buffer (if $Q_1 + 1 \leqslant \hat{q}$).

The minus of these policies is that the drop probability ceases to depend on the queue length and it is impossible to set the maximum drop probability. Models with three thresholds allow to get rid of this minus.

For the third model in [13] the steady-state probability distribution of packets p_i $(i \geqslant 0)$ in the system(for imbedded Markov chain), some probabilities were obtained and represented by geometric form (when the threshold Q_1 is overcomed):

$$p_i = \sum_{j=i-1}^{Q_1} p_j(-\mu)^{j+1-i}\alpha^{(j+1-i)}(\mu) + \\ + p_{Q_1+1}g^{i-Q_1-2}\left(g - \alpha(\mu) - \int_0^{\infty} A(g,x)e^{-\mu x}dA(x)\right), \quad i = \overline{1, Q_1+1},$$

$$p_i = p_{Q_1+1} g^{i-(Q_1+1)}, \quad i > Q_1 + 1,$$

where g is the unique solution of the equation $g = \alpha(\mu(1 - gQ(g)))$, and belongs to interval $(0; 1)$, $\alpha(s)$ is the Laplase–Stieltjes transformation of interrarrival time distribution function $A(x)$, $Q(g)$ is the probability generating function for probabilities $\pi(l, k)$ that l packets will be served and k packets will be dropped from the buffer,

$$A(g, x) = \sum_{l=1}^{Q_1+1-i} \frac{(\mu x g)^l}{l!} \sum_{j=0}^{Q_1+1-i-l} \pi(l, j) g^j.$$

Also the probability that the arriving packet will be dropped and sojourn time characteristics for dropped packets were obtained in form of integral equations.

3.4. The mathematical model with renovation and tree thresholds

In this part of the article models with renovation and three thresholds (Q_1, Q_2 and Q_3) will be formulated. The first threshold Q_1 determines when packets start to be dropped, the second threshold $Q_2 (Q_2 > Q_1)$ sets the maximum value of drop probability, the third threshold $Q_3 (Q_3 < Q_1)$ determines the place to which packets may be dropped from the queue.

The following types of renovation mechanism may be defined. At the moment of the end of a packet service the current queue length $\hat{q}$ is compared with thresholds and if $\hat{q} \leqslant Q_1$ then no one of the packets from the buffer is dropped.

— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k $(k \geqslant 1)$ packets will be dropped with fixed probability q_k. If $\hat{q} \geqslant Q_2 + 1$ then the fixed size group of k packets in the buffer will be dropped with maximal probability $q_{\max} (q_k < q_{\max} \leqslant 1$, but Q_3 packeta anyhow will remain in the system.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k packets will be dropped with probability $q_k(\hat{q})$, $0 < q_k(\hat{q}) < q_{\max}$. If $\hat{q} \geqslant Q_2 + 1$ then the fixed size group of k packets in the buffer will be dropped with maximal probability $q_{\max}$, but Q_3 packeta anyhow will remain in the system.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k packets will be dropped with fixed probability q_k. If $\hat{q} \geqslant Q_2 + 1$ then all packets except the first Q_1 will be dropped with probability q.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the fixed size group of k packets will be dropped with probability $q_k(\hat{q})$. If $\hat{q} \geqslant Q_2 + 1$ then all packets except the first Q_1 will be dropped with probability q.
— If $Q_1 + 1 \leqslant \hat{q} \leqslant Q_2$ then the arbitrary size group of k packets will be dropped with given probability $q(k)$ $(0 < q(k) < 1)$. If $\hat{q} \geqslant Q_2 + 1$ then all packets except the first Q_1 will be dropped with probability q.

For all types of renovation policy the following condition on probabilies q_1, q_k, $q(\widehat{q})$, $q_k(\widehat{q})$, $q(k)$, $q_{\max}$, q $(k \geqslant 1)$ must be met:

$$p_k + q_k = 1, \quad p_k(\widehat{q}) + q_k(\widehat{q}) = 1, \quad k \geqslant 1, \quad p + \sum_{k=1}^{\infty} q(k) = 1,$$

where p_k, $p_k(\widehat{q})$ $(k \geqslant 1)$ and p are the probabilities that the served packet will leave the system without dropping any other packets from the buffer (if $Q_1 + 1 \leqslant \widehat{q} \leqslant Q_2$), and $p^{(2)}_{\max} + q_{\max} = 1$, $p^{(2)} + q = 1$, $p^{(2)}_{\max}$ and $p^{(2)}$ are the probabilities that the served packet will leave the system without dropping any other packets from the buffer (if $\widehat{q} \geqslant Q_2 + 1$).

4. The comparison of some renovation models with RED algorithm

In this section we will compare values of the probability of packet being dropped from the system p^{loss} for RED and TailDrop algorithms [4], [6], [7] and values of the probability p^{loss}, obtained by formulas derived for queueing system with general renovation [11], queueing system with general renovation and feedback [33], [34] and queueing system with general renovation and two thresholds [12].

As can be see from the table 1, according to the values of the p^{loss}, renovation mechanism can perform as good as RED in the wide range of the offered load ρ.

Table 1

The value of the loss probability for Taildrop, RED, general renovation, general renovation with feedback, general renovation with two thresholds

Loss probability				
Taildrop	RED	renov.	ren-fd.	ren-thr.
$\rho = 0.5$				
0	0.002	0.002	0.002	0.003
$\rho = 1$				
0.051	0.091	0.104	0.11	0.109
$\rho = 2$				
0.500	0.500	0.502	0.54	0.524
$\rho = 3$				
0.667	0.667	0.667	0.71	0.0.679

5. Conclusion

The main task of the authors in this work was to formulate different types of models with various renovation policies and one or more threshold values in a buffer of a system. These thresholds allows to control the drop probability of renovation mechanism. Some of the described models have already been researched or are in the process of research (links to scientific publications are provided)

The presented numerical experiments show that the results remain qualitatively the same for RED-type AQM with other dropping functions. Being defined by N parameters, the renovation mechanism is very flexible and this constitutes its strength and weakness. By varying the values of the renovation probabilities $q(i)$, it is possible to carry out conditional optimisation, but good searching procedures are required here.

Acknowledgments

The publication has been funded by RFBR according to the research project No. 19-07-00739.

References

[1] F. Baker and G. Fairhurst. (Jul. 2015). "IETF Recommendations Regarding Active Queue Management. RFC 7567.," [Online]. Available: https://tools.ietf.org/html/rfc7567.

[2] M. Menth and S. Veith, "Active Queue Management Based on Congestion Policing (CP-AQM)," in. Jan. 2018, pp. 173–187. DOI: 10.1007/978-3-319-74947-1_12.

[3] R. Adams, "Active Queue Management: A Survey," *Communications Surveys & Tutorials, IEEE*, vol. 15, pp. 1425–1476, Jan. 2013. DOI: 10.1109/SURV.2012.082212.00018.

[4] A. Chydzinski and L. Chrost, "Analysis of AQM queues with queue size based packet dropping," *Applied Mathematics and Computer Science*, vol. 21, pp. 567–577, Sep. 2011. DOI: 10.2478/v10006-011-0045-7.

[5] A. Chydzinski and P. Mrozowski, "Queues with dropping functions and general arrival processes," *PloS one*, vol. 11, e0150702, Mar. 2016. DOI: 10.1371/journal.pone.0150702.

[6] M. Konovalov and R. Razumchik, "Numerical analysis of improved access restriction algorithms in a GI/G/1/N system," *Journal of Communications Technology and Electronics*, vol. 63, pp. 616–625, Jun. 2018. DOI: 10.1134/S1064226918060141.

[7] M. Konovalov and R. Razumchik, "Comparison of two active queue management schemes through the M/D/1/N queue," *Informatika i ee Primeneniya (Informatics and Applications)*, vol. 12, no. 4, pp. 9–15, 2018. DOI: 10.14357/19922264180402.

[8] S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, pp. 397–413, Sep. 1993. DOI: 10.1109/90.251892.

[9] K. Ramakrishnan, S. Floyd, and D. Black. (2001). "RFC3168: The Addition of Explicit Congestion Notification (ECN) to IP," [Online]. Available: https://tools.ietf.org/html/rfc3168.

[10] S. Floyd, R. Gummadi, and S. Shenker, "Adaptive RED: An Algorithm for Increasing the Robustness of RED's Active Queue Management," Sep. 2001.

[11] M. Konovalov and R. Razumchik. (2017). "Queueing systems with renovation vs. queues with RED. Supplementary Material." arXiv: 1709.01477, [Online]. Available: https://arxiv.org/abs/1709.01477.

[12] I. Zaryadov and A. Korolkova, "The Application of model with general renovation to the analysis of characteristics of active queue management with random early detection (RED) [Primeneniye modeli s obobshchennym obnovleniyem k analizu kharakteristik sistem aktivnogo upravleniya ocheredyami tipa Random Early Detection (RED)]," *T-Comm*, no. 7, pp. 84–88, 2011, In Russian.

[13] H. Viana Carvalho Cravid *et al.*, "The General Renovation as the Active Queue Management Mechanism. Some Aspects and Results," in *Communications in Computer and Information Science*. 2019, vol. 1141, pp. 488–502. DOI: 10.1007/978-3-030-36625-4_39.

[14] V. Jacobson, K. Nichols, and K. Poduri, "RED in a Different Light," Tech. Rep., 1999.

[15] G. Fairhurst and M. Welzl. (2017). "The Benefits of Using Explicit Congestion Notification (ECN). RFC 8087," [Online]. Available: https://tools.ietf.org/html/rfc3168.

[16] C. So-In, R. Jain, and J. Jiang, "Enhanced Forward Explicit Congestion Notification (E-FECN) scheme for datacenter Ethernet networks," in *2008 International Symposium on Performance Evaluation of Computer and Telecommunication Systems*, 2008, pp. 542–546.

[17] C. Gomez, X. Wang, and A. Shami, "Intelligent active queue management using explicit congestion notification," in *2019 IEEE Global Communications Conference (GLOBECOM)*, Sep. 2019, pp. 1–6. DOI: 10.20944/preprints201909.0077.v1.

[18] S. Shahzad, E.-S. Jung, J. F. Chung M., and R. Kettimuthu, "Enhanced Explicit Congestion Notification (EECN) in TCP with P4 Programming," in *2020 International Conference on Green and Human Information Technology (ICGHIT)*, Feb. 2020. DOI: 10.1109/ICGHIT49656.2020.00015.

[19] S. Wang, J. Zhang, T. Huang, T. Pan, J. Liu, and Y. Liu, "A-ECN Minimizing Queue Length for Datacenter Networks," *IEEE Access*, vol. 8, pp. 49 100–49 111, 2020. DOI: 10.1109/ACCESS.2020.2979216.

[20] K. Nichols and V. Jacobson, "Controlling queue delay," *Communications of The ACM – CACM*, vol. 55, pp. 42–50, May 2012. DOI: 10.1145/2209249.2209264.

[21] T. Hoeiland-Joergensen *et al.* (2018). "The Flow Queue CoDel Packet Scheduler and Active Queue Management Algorithm. RFC 8290," [Online]. Available: https://www.rfc-editor.org/info/rfc8290.

[22] S. Jung, J. Kim, and J.-H. Kim, "Intelligent active queue management for stabilized QoS guarantees in 5G mobile networks," *IEEE Systems Journal*, vol. PP, pp. 1–10, Aug. 2020. DOI: 10.1109/JSYST.2020.3014231.

[23] W.-C. Feng, D. Kandlur, D. Saha, and K. Shin, "BLUE: a new class of active queue management algorithms," University of Michigan, Tech. Rep., Sep. 2000. DOI: 10.1109/TNET.2002.801399.

[24] W.-C. Feng, D. Kandlur, and D. Saha, "The BLUE active queue management algorithms," *Networking, IEEE/ACM Transactions on*, vol. 10, pp. 513–528, Sep. 2002. DOI: 10.1109/TNET.2002.801399.

[25] C. Zhang, J. Yin, and Z. Cai, "RSFB: a Resilient Stochastic Fair Blue algorithm against spoofing DDoS attacks," in *2009 9th International Symposium on Communications and Information Technology*, 2009, pp. 1566–1567. DOI: 10.1109/ISCIT.2009.5341033.

[26] G. Da-gang, "A New Adaptive BLUE Algorithm," in *2010 International Conference on Electrical and Control Engineering*, 2010, pp. 2601–2605. DOI: 10.1109/iCECE.2010.638.

[27] T. Hoiland-Jorgensen, D. Taht, and J. Morton, "Piece of CAKE: A Comprehensive Queue Management Solution for Home Gateways," in *2018 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN)*, Jun. 2018, pp. 37–42. DOI: 10.1109/LANMAN.2018.8475045.

[28] J. Palmei, S. Gupta, P. Imputato, J. Morton, M. Tahiliani, S. Avallone, and D. Taht, "Design and Evaluation of COBALT Queue Discipline," in *2019 IEEE International Symposium on Local and Metropolitan Area Networks (LANMAN)*, Jul. 2019, pp. 1–6. DOI: 10.1109/LANMAN.2019.8847054.

[29] W.-C. Feng, "Improving Internet Congestion Control and Queue Management Algorithms," Ph.D. dissertation, The University of Michigan, 1999.

[30] A. V. Korolkova, D. S. Kulyabov, and A. I. Chernoivanov, "On the classification of RED algorithms," *Bulletin of Peoples' Friendship University of Russia. Series Mathematics. Information Sciences. Physics*, no. 3, pp. 34–46, 2009, In Russian.

[31] A. Kreinin, "Queueing systems with renovation," *Journal of Applied Mathematics and Stochastic Analysis*, vol. 10, pp. 431–443, Jan. 1997. DOI: 10.1155/S1048953397000464.

[32] P. P. Bocharov and I. S. Zaryadov, "Probability Distribution in Queueing Systems with Renovation," *Bulletin of Peoples' Friendship University of Russia. Series Mathematics. Information Sciences. Physics*, pp. 15–25, 2007.

[33] E. Bogdanova *et al.*, "Characteristics of lost and served packets for retrial queueing system with general renovation and recurrent input flow," *Communications in Computer and Information Science*, vol. 919, pp. 327–340, 2018. DOI: 10.1007/978-3-319-99447-5_28.

[34] I. Zaryadov, E. Bogdanova, T. Milovanova, S. Matushenko, and D. Pyatkina, "Stationary Characteristics of the GI/M/1 Queue with General Renovation and Feedback," in *10th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, Nov. 2018, pp. 1–6. DOI: 10.1109/ICUMT.2018.8631244.

[35] I. Zaryadov and A. Pechinkin, "Stationary time characteristics of the $GI/M/n/\infty$ system with some variants of the generalized renovation discipline," *Automation and Remote Control*, vol. 70, pp. 2085–2097, Dec. 2009. DOI: 10.1134/S0005117909120157.

[36] I. Zaryadov, "Queueing systems with general renovation," in *International Conference on Ultra Modern Telecommunications and Workshops*, Oct. 2009, pp. 1–4. DOI: 10.1109/ICUMT.2009.5345382.

[37] I. Zaryadov, R. Razumchik, and T. Milovanova, "Stationary waiting time distribution in G|M|n|r with random renovation policy," in *Communications in Computer and Information Science*, vol. 678, Feb. 2016, pp. 349–360. DOI: 10.1007/978-3-319-51917-3_31.

[38] E. V. Bogdanova, T. A. Milovanova, and I. S. Zaryadov, "The analysis of queuing system with general service distribution and renovation," *RUDN Journal of Mathematics, Information Sciences and Physics*, vol. 25, no. 1, pp. 3–8, 2017. DOI: 10.22363/2312-9735-2017-25-1-3-8.

[39] I. S. Zaryadov, E. V. Bogdanova, and T. A. Milovanova, "Probability-time characteristics of M|G|1|1 queueing system with renovation," in *CEUR Workshop Proceedings*, vol. 1995, 2017, pp. 125–131.

For citation:

H. Viana Carvalho Cravid, I. S. Zaryadov, T. A. Milovanova, Queueing systems with different types of renovation mechanism and thresholds as the mathematical models of active queue management mechanism, Discrete and Continuous Models and Applied Computational Science 28 (4) (2020) 305–318. DOI: 10.22363/2658-4670-2020-28-4-305-318.

Information about the authors:

Viana Carvalho Cravid, Hilquias — post-graduate student of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: hilvianamat1@gmail.com, phone: +7(495)9550927, Scopus Author ID: 57212930802)

Zaryadov, Ivan S. — Candidate of Physical and Mathematical Sciences, assistant professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University); Senior Researcher of Institute of Informatics Problems of Federal Research Center "Computer Science and Control" Russian Academy of Sciences (e-mail: zaryadov-is@rudn.ru, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0002-7909-6396, ResearcherID: B-8154-2018, Scopus Author ID: 35294470000)

Milovanova, Tatiana A. — Candidate of Physical and Mathematical Sciences, lecturer of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: milovanova-ta@rudn.ru, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0002-9388-9499, Scopus Author ID: 26641495400)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2020-28-4-305-318

Системы массового обслуживания с различными видами обновления и порогами как математические модели алгоритмов активного управления очередями

Илкиаш Виана Карвалью Кравид[1], **Иван С. Зарядов**[1, 2], **Татьяна А. Милованова**[1]

[1] *Российский университет дружбы народов*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия
[2] *Институт проблем информатики*
Федеральный исследовательский центр «Информатика и управление» РАН
ул. Вавилова, д. 44, корп. 2, Москва, 119333, Россия

Работа посвящена некоторым аспектам использования механизма обновления (различные варианты обновления рассмотрены, определения и краткий обзор представлены) с одним или несколькими порогами в качестве математических моделей механизмов активного управления очередями.

Описаны системы массового обслуживания, в которых реализован механизм обновления с порогами, позволяющий управлять числом заявок в системе путем их сброса из накопителя в зависимости от значения некоторого управляющего параметра и пороговых значений. Сброс заявок из накопителя происходит в момент окончания обслуживания заявки на приборе, что отличает данный механизм сброса от RED-подобных алгоритмов, для которых сброс возможен в момент поступления в систему.

Представлены модели с одним, двумя или тремя порогами. В этих моделях пороговые значения определяют не только место, с которого в накопителе начинается сброс заявок, но и до какой позиции заявки могут быть сброшены. Для некоторых из описываемых моделей уже получены аналитические и численные результаты (ссылки на работы представлены), но большая часть моделей находится в процессе изучения, поэтому представлены только описания и некоторые текущие данные.

Приведены некоторые результаты сравнения классического алгоритма RED с механизмом обновления.

Ключевые слова: система массового обслуживания, активное управление очередью, обновление, обобщенное обновление, функция сброса, пороговый механизм, контроль перегрузок сети

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2020-28-4-319-326

Stochastic analysis of a single server unreliable queue with balking and general retrial time

Mohamed Boualem

Research Unit LaMOS (Modeling and Optimization of Systems)
Faculty of Technology, University of Bejaia, Algeria

(received: June 8, 2020; accepted: November 12, 2020)

In this investigation, we consider an $M/G/1$ queue with general retrial times allowing balking and server subject to breakdowns and repairs. In addition, the customer whose service is interrupted can stay at the server waiting for repair or leave and return while the server is being repaired. The server is not allowed to begin service on other customers until the current customer has completed service, even if current customer is temporarily absent. This model has a potential application in various fields, such as in the cognitive radio network and the manufacturing systems, etc. The methodology is strongly based on the general theory of stochastic orders. Particularly, we derive insensitive bounds for the stationary distribution of the embedded Markov chain of the considered system.

Key words and phrases: Retrial queue, Markov chain, balking, breakdowns and repairs, stochastic orders, bounds, ageing classes

1. Introduction

The study on queueing models have become an indispensable area due to its wide applicability in real life situations. Retrial queues occupy an intermediate situation between an Erlang model with loss and classical model with wait, which constitute their limiting models in the case of low and high retrial rates. Retrial queueing systems are characterized by the requirement that customers finding the service area busy must join the retrial group and retry for service at random intervals. Queues in which customers are allowed to retry have been extensively used to model many problems in telephone switching systems, cognitive radio network, manufacturing systems, telecommunication networks and computer systems for competing to gain service from a central processor unit [1]–[3].

Retrial queueing systems with general service times and non-exponential retrial time distributions have been received little attention because of the complexity of the known results. Indeed, in most cases, we are faced with systems of equations whose resolution is complex, or having solutions not

easily interpretable. For instance, Pollaczek-Khintchine formula requires a numerical inversion of the Laplace transform to compute the distribution of the waiting time. In many cases, even the Laplace transform or probability generating functions are not available in explicit forms. To overcome these difficulties, approximation methods are often used to obtain quantitative and/or qualitative estimates for certain performance measures. For all these reasons, in this study, a particular interest is devoted to the stochastic comparison method based on the general theory of stochastic orders [4], [5].

The stochastic comparison method is a mathematical tool used to study the performance of some systems modeled by continuous or discrete time Markov chains. The general idea of this method is to bound a complex system with a new system that is simpler to solve providing qualitative bounds for these performance measures. These methods represent one of the main research activities in various scientific fields, such as economy, biology, operation research, reliability theory, decision theory, retrial queues and queueing networks [4]–[15].

In the present study, stochastic comparison analysis of an $M/G/1$ queue with server subject to breakdowns, general service times and non-exponential retrial time distributions by considering the both balking and reneging behavior of the customer is presented.

The rest of the paper is organized as follows. In the next Section, we describe the considered queueing system. In Section 3, we introduce an overview on stochastic orders. In Section 4, we present some lemmas that will be used in what follows. Section 5 focusses on monotonicity properties of the transition operator. Insensitive bounds for the stationary number of customers in the system at a departure epochs are discussed in Section 6.

2. Mathematical model description

Primary customers arrive in a Poisson process with rate λ. If an arriving primary customer finds the server idle, it begins service immediately and leaves the system after service completion. If the server is found to be blocked, the arriving primary customer either enters a retrial queue (according to a FCFS discipline) with probability p or leaves the system with probability $1-p$. The service times follow a general distribution $B(x)$ with corresponding Laplace-Stieltjes transform $B^\star(s)$ and the first two moments β_1 and β_2.

The customers in the orbit try to require the service later and the inter retrial times have a general distribution $A(x)$ with corresponding Laplace-Stieltjes transform $A^\star(s)$. The retrial customer cancels its attempt for service if a primary customer arrives first and either returns to its position in the retrial queue with probability q or quits the system with probability $1-q$.

We assume that the server may fail (or break down), but only when a customer is in service. The time until failure is exponentially distributed with mean $1/\mu$, but failure of the server can occur only when a customer is being served. When the server fails, the repair starts immediately and the customer just under service waits for the server until repair completion in order to accomplish its remaining service. The repair time is a random variable with probability distribution function $C(x)$, Laplace-Stieltjes transform $C^\star(s)$ and the first two moments γ_1 and γ_2.

Upon failure of the server, the customer in service either remains in the service position with probability α until the server is up or enters a service retrial orbit with probability $1-\alpha$ and keeps returning at times exponentially distributed with mean $1/\theta$, until the server is repaired. If the customer in service enters the service retrial orbit upon server failure, after repair the server must wait for the customer to return. We refer to this time as the reserved time.

The performance characteristics of our queueing system are available in Wu et al. [16]. Under the stability condition $p\lambda\beta_1(1+\mu(\frac{1-\alpha}{\theta}+\gamma_1)) < 1-q+qA^\star(\lambda)$, the one-step transition probabilities are given as:

$$P_{ij} = \begin{cases} r_j, \quad \text{if} \quad i=0, \quad j \geqslant 0, \\ (1-q+qA^\star(\lambda))r_0, \quad \text{if} \quad i>0, \quad j=i-1, \\ q(1-A^\star(\lambda))r_{j-1} + (1-q+qA^\star(\lambda))r_{j-i+1}, \text{if } i>0,\ j>i-1, \\ 0, \quad \text{otherwise}, \end{cases} \tag{1}$$

where

$$r_m = \int_0^{+\infty} \frac{(p\lambda x)^m}{m!} \exp\{-p\lambda x\}\, dB^\bullet(x), \quad m=0,1,2,...$$

is the distribution of the number of primary customers which arrive during the generalized service time of the $(n+1)th$ customer. We define the generalised service time as the length of time from when a customer begins service until service completion (that includes: some eventual repair times (X_i) and reserved times (Y_i)) with common distribution

$$B^\bullet(x) = \sum_{i=0}^{\infty}\sum_{j=0}^{i}\int_0^x \binom{i}{j}\alpha^j(1-\alpha)^{i-j}\frac{(\mu y)^i}{i!}e^{-\mu y}C_{i,j}^{(2)}(x-y)dB(y).$$

3. Overview on stochastic orders and ageing notions

Stochastic ordering is useful for studying internal changes of performance due to parameter variations, to compare distinct systems, to approximate a system by a simpler one, and to obtain upper and lower bounds for the main performance measures of systems. For a comprehensive discussion on these stochastic orders, see [4], [5].

Let X and Y be two random variables non-negative with distribution functions F and G, respectively.

Laplace order $(\leqslant_L)$: $X \leqslant_L Y \Leftrightarrow \int_0^{+\infty} e^{-sx}dF(x) \geqslant \int_0^{+\infty} e^{-sx}dG(x), \forall s \geqslant 0.$

Stochastic order $(\leqslant_{st})$: $X \leqslant_{st} Y \Leftrightarrow F(x) \geqslant G(x), \forall x \in \mathbb{R}.$

Convex order $(\leqslant_v)$: $X \leqslant_v Y \Leftrightarrow \int_x^{\infty}(1-F(t))dt \leqslant \int_x^{\infty}(1-G(t))dt, \forall x \geqslant 0.$

Proposition 1. Let X be a random variable with distribution function F and finite mean m.

(a) F is New Better than Used in Expectation (NBUE) iff $F \leqslant_v F^*$,
(b) F is New Worse than Used in Expectation (NWUE) iff $F^* \leqslant_v F$,
(c) F is of class $\mathcal{L}$ iff $F \geqslant_L F^*$, where F^* is the exponential distribution function with the same mean as F.

4. Comparison bounds for the probability r_m

This subsection presents several useful lemmas which will be used later in establishing the main results. For this, we consider two $M/G/1$ queue with general retrial times allowing balking and server subject to breakdowns and repairs with parameters $\lambda^{(i)}$, $q^{(i)}$, $\alpha^{(i)}$, $\mu^{(i)}$, $C^{(i)}$, $B^{(i)}$, for $i = 1, 2$ respectively.

Lemma 1. *If* $\lambda^{(1)} \leqslant \lambda^{(2)}$, $q^{(1)} \leqslant q^{(2)}$, $\alpha^{(1)} \geqslant \alpha^{(2)}$, $\mu^{(1)} \leqslant \mu^{(2)}$, $C^{(1)} \leqslant_{st} C^{(2)}$, $B^{(1)} \leqslant_{st} B^{(2)}$, *then* $\{r_m^{(1)}\} \leqslant_{st} \{r_m^{(2)}\}$.

Lemma 2. *If* $\lambda^{(1)} \leqslant \lambda^{(2)}$, $q^{(1)} \leqslant q^{(2)}$, $\alpha^{(1)} \geqslant \alpha^{(2)}$, $\mu^{(1)} \leqslant \mu^{(2)}$, $C^{(1)} \leqslant_v C^{(2)}$, $B^{(1)} \leqslant_v B^{(2)}$, *then* $\{r_m^{(1)}\} \leqslant_v \{r_m^{(2)}\}$.

Proof. To show both lemmas, we have to establish the usual numerical inequalities

$$\overline{r}_m^{(1)} = \sum_{n \geqslant m} r_n^{(1)} \leqslant \overline{r}_m^{(2)} \quad (\text{for} \quad \leqslant_s = \leqslant_{st}),$$

$$\overline{\overline{r}}_m^{(1)} = \sum_{n \geqslant m} \overline{r}_n^{(1)} \leqslant \overline{\overline{r}}_m^{(2)} \quad (\text{for } \leqslant_s = \leqslant_v).$$

By definition, we have

$$\overline{r}_m^{(i)} = \sum_{j \geqslant m} r_j^{(i)} = \int_0^{+\infty} \sum_{j \geqslant m} \frac{(p^{(i)}\lambda^{(i)}x)^j}{j!} e^{-p^{(i)}\lambda^{(i)}x} dB^{\bullet(i)}(x), \ i = 1, 2.$$

$$\overline{\overline{r}}_m^{(i)} = \sum_{j \geqslant m} \overline{r}_j^{(i)} = \int_0^{+\infty} \sum_{j \geqslant m} \sum_{l \geqslant j} \frac{(p^{(i)}\lambda^{(i)}x)^l}{l!} e^{-p^{(i)}\lambda^{(i)}x} dB^{\bullet(i)}(x), \ i = 1, 2.$$

The rest of the proof is known in the more general setting of a random summation. □

5. Monotonicity properties of the transition operator

Let $\mathbf{T}$ be the transition operator which associates to every distribution $\omega = \{\omega_j\}_{j \geqslant 0}$ a distribution $\mathbf{T}\omega = \{\nu_j\}_{j \geqslant 0}$ such that $\nu_j = \sum_{i \geqslant 0} \omega_i P_{ij}$.

Theorem 1. *The transition operator* $\mathbf{T}$ *is monotone with respect to the orders stochastic* ($\leqslant_{st}$) *and convex* ($\leqslant_v$).

Proof. It is well known that an operator is monotone with respect to $\leqslant_{st}$ if and only if $\overline{P}_{ij} - \overline{P}_{i-1j} \geqslant 0$, and is monotone with respect to $\leqslant_v$ if and only if $\overline{\overline{P}}_{i-1j} + \overline{\overline{P}}_{i+1j} - 2\overline{\overline{P}}_{ij} \geqslant 0,\ \forall\ i,\ j.$

In our case, $\overline{P}_{ij} = \sum_{l=j}^{\infty} P_{il}$ and $\overline{\overline{P}}_{ij} = \sum_{l=j}^{\infty} \overline{p}_{il}$. □

Now, we consider two $M/G/1$ queue with general retrial times allowing balking and server subject to breakdowns and repairs with parameters $\lambda^{(i)}$, $q^{(i)}$, $\alpha^{(i)}$, $\mu^{(i)}$, $C^{(i)}$, $A^{(i)}$, $B^{(i)}$, for $i = 1, 2$ respectively. In the following theorem, we give comparability conditions of two transition operators $\mathbf{T}^1$ and $\mathbf{T}^2$.

Theorem 2. *If $\lambda^{(1)} \leqslant \lambda^{(2)}$, $q^{(1)} \leqslant q^{(2)}$, $\alpha^{(1)} \geqslant \alpha^{(2)}$, $\mu^{(1)} \leqslant \mu^{(2)}$, $C^{(1)} \leqslant_s C^{(2)}$, $B^{(1)} \leqslant_s B^{(2)}$ and $A^{(1)} \leqslant_L A^{(2)}$, then $\mathbf{T}^1 \leqslant_s \mathbf{T}^2$, where $\leqslant_s$ is one of the symbols $\leqslant_{st}$ or $\leqslant_v$. I.e. for any distribution ω, we have $\mathbf{T}^1\omega \leqslant_s \mathbf{T}^2\omega$.*

Proof. It is well known that to prove $\mathbf{T}^1 \leqslant_s \mathbf{T}^2$, we have to show the following numerical inequalities for the one-step transition probabilities $P_{ij}^{(1)}$, $P_{ij}^{(2)}$ (see Stoyan [5]): $\overline{P}_{ij}^{(1)} \leqslant \overline{P}_{ij}^{(2)}$, $\forall i,\ j$, (for $\leqslant_s = \leqslant_{st}$), and $\overline{\overline{P}}_{ij}^{(1)} \leqslant \overline{\overline{P}}_{ij}^{(2)}$, $\forall i, j$, (for $\leqslant_s = \leqslant_v$).

After algebraic manipulations, we obtain the result searched. □

6. Bounds for the stationary distribution

Now, we add the the corresponding stationary distributions of the number of customers in the system $\pi_n^{(1)}$ and $\pi_n^{(2)}$.

Theorem 3. *If $\lambda^{(1)} \leqslant \lambda^{(2)}$, $q^{(1)} \leqslant q^{(2)}$, $\alpha^{(1)} \geqslant \alpha^{(2)}$, $\mu^{(1)} \leqslant \mu^{(2)}$, $C^{(1)} \leqslant_s C^{(2)}$, $B^{(1)} \leqslant_s B^{(2)}$ and $A^{(1)} \leqslant_L A^{(2)}$, then $\{\pi_n^{(1)}\} \leqslant_s \{\pi_n^{(2)}\}$, where $\leqslant_s$ is one of the symbols $\leqslant_{st}$ or $\leqslant_v$.*

Proof. Using Theorems 1 and 2 which state that $\mathbf{T}^i$ are monotone with respect to the order $\leqslant_s$ and $\mathbf{T}^1 \leqslant_s \mathbf{T}^2$, we have by induction $\mathbf{T}^{1,n}\omega \leqslant_s \mathbf{T}^{2,n}\omega$ for any distribution ω. Taking the limit, we obtain the stated result. □

Theorem 4. *We consider an $M/G/1$ queue with general retrial times allowing balking and server subject to breakdowns and repairs. If we have, the service time distribution $B(x)$ and the repair time distribution $C(x)$ are NBUE and the retrial time distribution is of class $\mathcal{L}$, then $\{\pi_n\} \leqslant_v \{\pi_n^*\}$, where $\{\pi_n^*\}$ is the stationary distribution in the Markovian retrial queue with the same parameters.*

Proof. Consider an auxiliary Markovian $M/M/1$ retrial queue allowing balking and server subject to breakdowns and repairs, having the same parameters as the model considered in our study , but with exponentially distributed retrial time $A^*(x)$, service time $B^*(x)$ and repair time $C^*(x)$. We have also, if $B(x)$ and $C(x)$ are $NBUE$ and $A(x)$ is $\mathcal{L}$, then $B(x) \leqslant_v B^*(x)$, $C(x) \leqslant_v C^*(x)$ and $A(x) \leqslant_L A^*(x)$. Therefore, by using Theorem 3, we deduce the statement of this theorem. □

7. Conclusions

In this work, we investigated the monotonicity of the transition operator of the embedded Markov chain with respect to stochastic and convex orderings. In addition, we done comparability conditions of the transition operators of the considered systems, associated with two Markov chains. Further, we obtained comparability conditions for which the stationary distribution in the considered model is bounded by the stationary distribution of the Markovian retrial queue, if the service time and the repair time distributions are NBUE and the retrial time distribution is of class $\mathcal{L}$.

Acknowledgments

The authors are pleased to thank the anonymous referees and the editor for their valuable comments and suggestions, which improved the content and the presentation of this paper.

References

[1] J. R. Artalejo and A. Gómez-Corral, *Retrial queueing system: A computational approach*. Berlin: Springer, Berlin, Heidelberg, 2008, 318 pp.

[2] A. A. Nazarov, S. V. Paul, and O. D. Lizyura, "Two-way communication retrial queue with unreliable server and multiple types of outgoing calls," *Discrete and Continuous Models and Applied Computational Science*, vol. 28, no. 1, pp. 49–61, 2020. DOI: 10.22363/2658-4670-2020-28-1-49-61.

[3] D. Zirem, M. Boualem, K. Adel-Aissanou, and D. Aïssani, "Analysis of a single server batch arrival unreliable queue with balking and general retrial time," *Quality Technology & Quantitative Management*, vol. 16, pp. 672–695, 2019. DOI: 10.1080/16843703.2018.1510359.

[4] M. Shaked and J. G. Shanthikumar, *Stochastic Orders*. New York: Springer-Verlag, 2007, 473 pp.

[5] D. Stoyan, *Comparison methods for queues and other stochastic models*. New York: Wiley, 1983, 217 pp.

[6] L. M. Alem, M. Boualem, and D. Aïssani, "Bounds of the stationary distribution in $M/G/1$ retrial queue with two-way communication and n types of outgoing calls," *Yugoslav Journal of Operations Research*, vol. 29, no. 3, pp. 375–391, 2019. DOI: 10.2298/YJOR180715012A.

[7] L. M. Alem, M. Boualem, and D. Aïssani, "Stochastic comparison bounds for an $M_1, M_2/G_1, G_2/1$ retrial queue with two way communication," *Hacettepe Journal of Mathematics and Statistics*, vol. 48, no. 4, pp. 1185–1200, 2019. DOI: 10.1572/HJMS.2018.629.

[8] M. Boualem, "Insensitive bounds for the stationary distribution of a single server retrial queue with server subject to active breakdowns," *Advances in Operations Research*, vol. 2014, no. 1, pp. 1–12, 2014. DOI: 10.1155/2014/985453.

[9] M. Boualem, A. Bareche, and M. Cherfaoui, "Approximate controllability of stochastic bounds of stationary distribution of an $M/G/1$ queue with repeated attempts and two-phase service," *International Journal of Management Science and Engineering Management*, vol. 14, no. 2, pp. 79–85, 2019. DOI: 10.1080/17509653.2018.1488634.

[10] M. Boualem, M. Cherfaoui, and D. Aïssani, "Monotonicity properties for a single server queue with classical retrial policy and service interruptions," *Proceedings of the Jangjeon Mathematical Society*, vol. 19, no. 2, pp. 225–236, 2016.

[11] M. Boualem, M. Cherfaoui, N. Djellab, and D. Aïssani, "A stochastic version analysis of an $M/G/1$ retrial queue with Bernoulli schedule," *Bulletin of the Iranian Mathematical Society*, vol. 43, no. 5, pp. 1377–1397, 2017.

[12] M. Boualem, M. Cherfaoui, N. Djellab, and D. Aïssani, "Inégalités stochastiques pour le modèle d'attente $M/G/1/1$ avec rappels," French, *Afrika Matematika*, vol. 28, pp. 851–868, 2017. DOI: 10.1007/s13370-017-0492-x.

[13] M. Boualem, N. Djellab, and D. Aïssani, "Stochastic inequalities for $M/G/1$ retrial queues with vacations and constant retrial policy," *Mathematical and Computer Modelling*, vol. 50, no. 1–2, pp. 207–212, 2009. DOI: 10.1016/j.mcm.2009.03.009.

[14] M. Boualem, N. Djellab, and D. Aïssani, "Stochastic approximations and monotonicity of a single server feedback retrial queue," *Mathematical Problems in Engineering*, vol. 2012, 12 pages, 2012. DOI: 10.1155/2012/536982.

[15] M. Boualem, N. Djellab, and D. Aïssani, "Stochastic bounds for a single server queue with general retrial times," *Bulletin of the Iranian Mathematical Society*, vol. 40, no. 1, pp. 183–198, 2014.

[16] X. Wu, P. Brill, M. Hlynka, and J. Wang, "An $M/G/1$ retrial queue with balking and retrials during service," *International Journal of Operational Research*, vol. 1, pp. 30–51, 2005. DOI: 10.1504/IJOR.2005.007432.

For citation:

M. Boualem, Stochastic analysis of a single server unreliable queue with balking and general retrial time, Discrete and Continuous Models and Applied Computational Science 28 (4) (2020) 319–326. DOI: 10.22363/2658-4670-2020-28-4-319-326.

Information about the authors:

Mohamed Boualem — Full Professor, Professor of Applied Mathematics at the Department of Technology at the University of Bejaia, Algeria. (e-mail: robertt15dz@yahoo.fr, phone: +213792991121, ORCID: https://orcid.org/https://orcid.org/0000-0001-9414-714X, ResearcherID: A-3176-2019, Scopus Author ID: 26633399700)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2020-28-4-319-326

Стохастический анализ системы типа «клиент–сервер» с ненадёжной очередью с блокировкой и общим временем обновления

Мохамед Буалем

Исследовательское подразделение LaMOS
Факультет технологий, Университет Беджая, Алжир

В статье рассматривается система массового обслуживания типа $M/G/1$ с обобщённым временем обновления, допускающая блокировку, выход из строя и возобновление работы сервера. Кроме того, клиент, обслуживание которого прервано, может оставаться на сервере в ожидании восстановления его работы, а может покинуть систему и вернуться в период восстановления работы сервера. Серверу не разрешается начинать обслуживание других клиентов до тех пор, пока текущий клиент не завершит обслуживание, даже если он временно отсутствует. Эта модель имеет потенциальное применение в различных областях, таких как сеть когнитивного радио, производственные системы и т. д. Методология строго базируется на общей теории стохастических порядков. В частности, получены оценки стационарного распределения вложенной цепи Маркова рассматриваемой системы.

Ключевые слова: очередь с обновлением, цепь Маркова, блокировка, выход из строя и восстановление, стохастический порядок, границы, классы старения

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2020-28-4-327-345

On the realization of explicit Runge–Kutta schemes preserving quadratic invariants of dynamical systems

Yu Ying[1], **Mikhail D. Malykh**[2]

[1] *Kaili University*
3, Kaiyuan Road, Kaili, 556011, China
[2] *Peoples' Friendship University of Russia (RUDN University)*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: October 23, 2020; accepted: November 12, 2020)

We implement several explicit Runge–Kutta schemes that preserve quadratic invariants of autonomous dynamical systems in Sage. In this paper, we want to present our package ex.sage and the results of our numerical experiments.

In the package, the functions rrk_solve, idt_solve and project_1 are constructed for the case when only one given quadratic invariant will be exactly preserved. The function phi_solve_1 allows us to preserve two specified quadratic invariants simultaneously. To solve the equations with respect to parameters determined by the conservation law we use the elimination technique based on Gröbner basis implemented in Sage. An elliptic oscillator is used as a test example of the presented package. This dynamical system has two quadratic invariants. Numerical results of the comparing of standard explicit Runge–Kutta method RK(4,4) with rrk_solve are presented. In addition, for the functions rrk_solve and idt_solve, that preserve only one given invariant, we investigated the change of the second quadratic invariant of the elliptic oscillator. In conclusion, the drawbacks of using these schemes are discussed.

Key words and phrases: Explicit Runge–Kutta method, quadratic invariant, dynamical system, Sage

1. Quadratic invariant and conservative RK scheme

One of most widespread mathematical models is an autonomous system of ordinary differential equations, i.e., the system of the form

$$\begin{cases} \dfrac{dx}{dt} = f(x), \quad t \geqslant 0, \\ x(0) = x^0, \end{cases} \tag{1}$$

where: t is an independent variable, commonly interpreted as time; x is a vector $(x_1, \dots, x_n)$; f is a vector function $(f_1, f_2, \dots, f_n)$, when in applications

its element f_i $(i = 1, 2, \dots, n)$ is often taken as a rational or an algebraic function of the coordinates $x_1, \dots, x_n$ or one can be reduced to this form by some transformation of variables.

Definition 1 (Goriely [1]). If there exists a function I of x, such that, for any solution $x(t)$ of system (1), the condition

$$\nabla I(x) f(x) = \sum_{k=1}^{n} \frac{dx_k}{dt} \frac{\partial I}{\partial x_k} = \sum_{k=1}^{n} f_k(x(t)) \frac{\partial I}{\partial x_k} = 0,$$

holds, then I is called the first integral or invariant of the system (1). If $x(t)$ is any exact solution of system (1) then $I(x(t))$ is independent of t. If I is a polynomial of degree 2 with respect to x then it is called a quadratic invariant.

Any quadratic invariant after a linear transformation can be rewritten in the form

$$I(x(t)) = x(t)^T S x(t) = \langle Sx(t), x(t) \rangle = \text{const}, \tag{2}$$

where $\langle \cdot\ , \cdot \rangle$ denote the Euclidean inner product on $\mathbb{R}^n$ and $S \in \mathbb{R}^{n \times n}$ is a symmetric, constant matrix.

To determine an uniform grid (with a step Δt) of the time interval $[0, T]$ we take

$$t_n = n\Delta t \quad (n = 0, \dots, N).$$

We will interpret $\{x^n\}$ as an approximation to the exact solution $x(t)$ at time $t_0 + n\ \Delta t$, i.e.

$$x(t_0 + n\Delta t) \approx x^n.$$

For the system (1) Runge–Kutta scheme (RK scheme) with s stages can be written as

$$k_i = x^n + \Delta t \sum_{j=1}^{s} a_{ij} f\left(k_j\right), \quad i = 1, 2, \dots, s \tag{3}$$

and

$$x^{n+1} = x^n + \Delta t \sum_{i=1}^{s} b_i f\left(k_i\right). \tag{4}$$

Below the parameters a_{ij} and $b_i (i = 1, 2, ..., s, j = 1, 2, ..., s)$ will be arranged in an array

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \cdots & a_{1s} \\ c_2 & a_{21} & a_{22} & \cdots & a_{2s} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_s & a_{s1} & a_{s2} & \cdots & a_{ss} \\ \hline & b_1 & b_2 & \cdots & b_s \end{array}$$

where

$$c_i = \sum_{j=1}^{s} a_{ij},$$

is known as the Butcher table [2], [3] and will be called the coefficients of the RK scheme. The RK method is explicit if $a_{ij} = 0$ for $i \leq j$, otherwise is implicit.

We want to construct numerical solutions $x^0, x^1, \dots, x^N$ such that the quadratic invariant $I(x)$ is preserved numerically, i.e.

$$\langle Sx^n, x^n \rangle = \langle Sx^0, x^0 \rangle \; i = 1, \dots, \; N. \tag{5}$$

In this case the RK method will be called S-conservative RK scheme.

Ref. [4]–[6] indicated that the RK method preserves the quadratic first integrals of system (1) iff the coefficients of such RK method satisfy

$$b_i a_{ij} + b_j a_{ji} - b_i b_j = 0, \quad i, j = 1, \dots, s. \tag{6}$$

Such Runge–Kutta methods are called *symplectic*.

Obviously, no explicit RK schemes satisfies the symplectic condition [6], [7]. Unfortunately, during using the implicit schemes, we must solve a system of non-linear algebraic equations at each step. This is very complex problem, so implicit schemes require more resources than explicit RK schemes [8]. Furthermore numerical solutions of nonlinear system (for ex., by the Newton method) introduce new errors that sometimes we cannot estimate effectively [9]. Thus, the integrals could not be preserved exactly. For this reason, many authors try to construct numerical methods for solving the system of differential equations (1) with the preservation of algebraic integrals without the need to solve nonlinear algebraic equations.

To overcome these difficulties Buono and Mastroserio [10] suggested a method that uses explicit RK schemes for the construction of new finite-difference schemes which exactly preserve invariants. Below we will call it the Buono method for shorthand. Of course, these new schemes are not standard RK schemes, but they are usually called an *explicit RK scheme preserving invariants* [8]. These schemes preserve only one specified invariant. We implemented several such schemes in Sage and investigated what happens to other invariants. Next, we investigated the method from th article [11] by Calvo et al. which is an extension of the Buono method and can be used as a conservation one or more invariants. Below we will call it the Calvo method for shorthand.

2. Explicit RK scheme of preserving one quadratic invariant

2.1. The Buono method

To make the explicit RK scheme conservative, we follow to Buono and Mastroserio [10]. We scale the weights b_i by a parameter $\gamma_n \in \mathbb{R}$ at the step n, i.e. use

$$x_\gamma^{n+1} = x^n + \Delta t \sum_{i=1}^{s} \gamma_n b_i f(k_i) \tag{7}$$

instead of x^{n+1} obtained by the formula (4) numerical solution after one time step.

Using the shorthand

$$\Delta_n = \sum_{i=1}^{s} b_i f(k_i),$$

the parameter γ_n could be estimated by the conservative condition, i.e.

$$\langle Sx_\gamma^{n+1}, x_\gamma^{n+1}\rangle - \langle Sx^n, x^n\rangle = \gamma_n \Delta t(2\langle Sx^n, \Delta_n\rangle + \gamma_n \Delta t\langle \Delta_n, \Delta_n\rangle). \tag{8}$$

Thus, we preserve the invariant $\langle Sx, x\rangle$, if we take

$$\gamma_n = -\frac{2\langle Sx^n, \Delta_n\rangle}{\Delta t\langle \Delta_n, \Delta_n\rangle}. \tag{9}$$

For Runge–Kutta schemes of order p this expression is close to 1, i.e

$$\gamma_n = 1 + \mathcal{O}(\Delta t^{p-1})$$

as $\Delta t \to 0$, see Buono at al. [10, prop. 4] for $p = 4$ and Zhang at al. [8, lemma 3.3]. Thus, the new numerical solution x_γ^{n+1} can be considered as an approximation either to $x(t_n + \Delta t)$ with the RK weights scaled by γ_n, or to $x(t_n + \gamma_n \Delta t)$ with the time Δt scaled by γ_n. Zhang at al. [8] denote the method defined by (3) and (7) with the interpretation

$$x_\gamma^{n+1} \approx x(t_n + \gamma_n \Delta t)$$

as *the relaxation Runge–Kutta method* (RRK), while the method using

$$x_\gamma^{n+1} \approx x(t_n + \Delta t)$$

this is called *the increment direction technique* (IDT). Note that the value of the scalar parameter γ_n at each step depends on the quadratic invariant that appears in *rrk or idt method.*

Theorem 1 (Zhang et al. [8]). *Let the original RK scheme be defined by* (3) *and* (4) *has order p, then the method defined by* (3) *and* (7) *has:*

- *(RRK method) If the solution x_γ^{n+1} is interpreted as an approximation to $x(t_n + \gamma_n \Delta t)$, the method has order p.*
- *(IDT method) If the solution x_γ^{n+1} is interpreted as an approximation to $x(t_n + \Delta t)$, the method has order $p - 1$.*

In our package `ex.sage` for Sage [12] the function `rrk_solve(P1,F,ics)` returns the numeric points $(0, x^0), (\gamma_0 \Delta t, x^1), \cdots$ with the parameters:

- `P1` is a quadratic invariant;
- `F` is the right sides of system (1)
- `ics` is the initial condition
- the default $\Delta t = 0.1, T = 10$
- 4-stage explicit RK scheme with the Butcher table

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
1	0	0	1	0
	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$

Function `idt_solve()` returns the numerical points $(0, x^0), (\Delta t, x^1), \cdots$ with the parameters:

— `P1` is a quadratic invariant;
— `F` is the right sides of the system (1)
— `ics` is the initial condition
— the default $\Delta t = 0.1, T = 10$
— 4-stage explicit RK scheme with coefficients the same as the previous table.

Users can redefine these variables in both functions, for ex., by adding `dt=0.01` or new explicit RK method.

2.2. Elliptic function test

To test this routine, we investigate a nonlinear oscillator. By the definition of Jacobi functions [13], $p = \operatorname{sn} t,\ q = \operatorname{cn} t,\ r = \operatorname{dn} t$ is a particular solution to a nonlinear autonomous system of differential equations

$$\dot{p} = qr,\ \dot{q} = -pr,\ \dot{r} = -k^2 pq \tag{10}$$

with the initial conditions

$$p = 0,\ q = r = 1 \text{ at } t = 0.$$

This autonomous system has two quadratic integrals of motion

$$p^2 + q^2 = 1 \quad \text{and} \quad k^2 p^2 + r^2 = 1 \tag{11}$$

We can solve the autonomous system (10) by rrk or idt methods that preserve only the first or second integral. For certainly, we take $k = 1/2$ and indicated above initial condition.

```
sage: var('p,q,r')
sage: load('ex.sage')
sage: k=1/2
sage: s=4
sage: F=[r*q,-p*r,-k^2*p*q]
sage: list_of_integral=[p^2+q^2,k^2*p^2+r^2]
sage: ics=[p==0,q==1,r==1]
sage: idt_solve(P1=list_of_integral[0],ics=ics,F=F,dt=0.2,T=1)
sage: rrk_solve(P1=list_of_integral[0],ics=ics,F=F,dt=0.2,T=1)
sage:
 ↪  B1=rrk_solve(P1=list_of_integral[0],ics=ics,F=F,dt=0.1,T=40)
sage:
 ↪  B2=rrk_solve(P1=list_of_integral[1],ics=ics,F=F,dt=0.1,T=40)
```

```
sage: G=line([[t,p] for [t,p,q,r] in
 ↪  B1],color='red',axes_labels=['t','p'])+point([[t,p] for
 ↪  [t,p,q,r] in B2],frame=true)
sage: max(abs(k^2*p^2+r^2-1) for [t,p,q,r] in B2 )
sage: max(abs(p^2+q^2-1) for [t,p,q,r] in B2)
sage: max(abs(k^2*p^2+r^2-1) for [t,p,q,r] in B1 )
sage: max(abs(p^2+q^2-1) for [t,p,q,r] in B1)
sage: G1=line([[t,k^2*p^2+r^2-1] for [t,p,q,r] in
 ↪  B1],axes_labels=['$t$','$k^2p^2+r^2-
 ↪  1$'],tick_formatter=[None,RR(10e-
 ↪  18).n(digits=1)],frame=true)
sage: G2=line([[t,p^2+q^2-1] for [t,p,q,r] in
 ↪  B2],axes_labels=['$t$','$p^2+q^2-
 ↪  1$'],tick_formatter=[None,RR(10e-
 ↪  18).n(digits=1)],frame=true)
```

In figure 1 we can see a graph of the solution found by rrk method with exact conservation of $p^2+q^2=1$. Rrk give a condensation of the greed points in those arches of the graph where the curvature has a maximum.

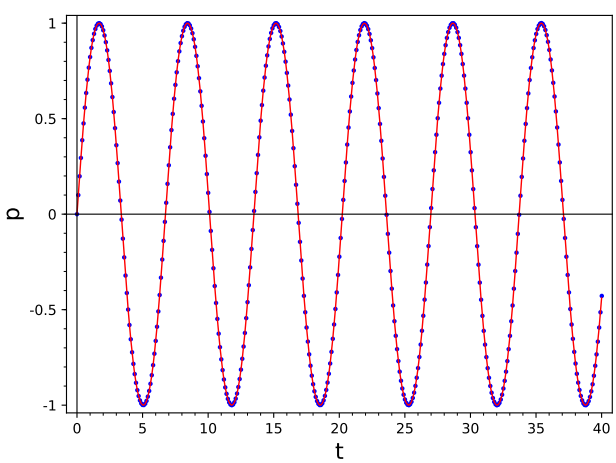


Figure 1. Graph of $p(t)$, rrk, $dt=0.1$

The second integral $k^2p^2+r^2-1$ is not exactly preserved, but its value fluctuates with a small amplitude of 10^-7 (figure 2). We also use rrk with exact conservation of $k^2p^2+r^2=1$. In this case first integral grows by leaps bounds and quickly becomes larger than 10^{-3} (figure 3). Thus, the preservation of one integral does not preserve others.

Let's compare the Buono method with the standard rk4 at the same step size dt (which denote Δt in the rrk method).

```
sage: var('p,q,r,t')
sage: QRK=desolve_system_rk4(F,[p,q,r], ics=[0,0,1,1,], ivar=t,
 ↪  end_points=20)
```

```
sage: G3=line([[t,p^2+q^2] for [t,p,q,r] in
 ↪  B2],color='red')+point([[t,p^2+q^2] for  [t,p,q,r] in
 ↪  QRK],axes_labels=['$t$','$p^2+q^2$'], tick_formatter=[None,
 ↪  RR(10e-18).n(digits=1)],frame=true)
sage: G4=line([[t,k^2*p^2+r^2-1] for [t,p,q,r] in
B1],axes_labels=['$t$','$k^2p^2+r^2-1$'],tick_formatter=[None,
 ↪  RR(10e-18).n(digits=1)], color='red') +
 ↪  point([[t,k^2*p^2+r^2-1] for  [t,p,q,r] in QRK],frame=true)
```

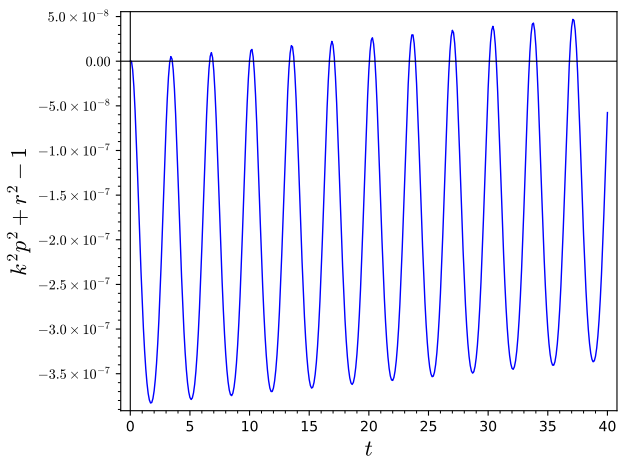


Figure 2. Second invariant for rrk method with exact conservation of first invariant $p^2 + q^2 = 1$

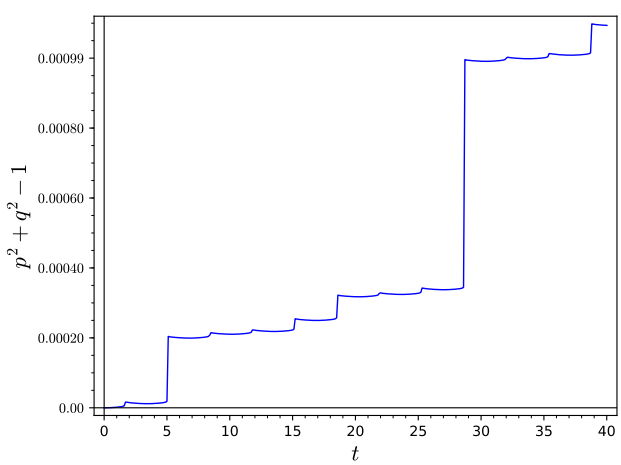


Figure 3. First integral for rrk with exact conservation of $k^2p^2 + r^2 = 1$

The rrk method with the exact conservation of the first invariant $p^2+q^2 = 1$ preserves both integrals better than rk4 (figure 5) but the rrk with the exact conservation of $k^2p^2 + r^2 = 1$ preserves the first integral worse than rk4 (figure 4).

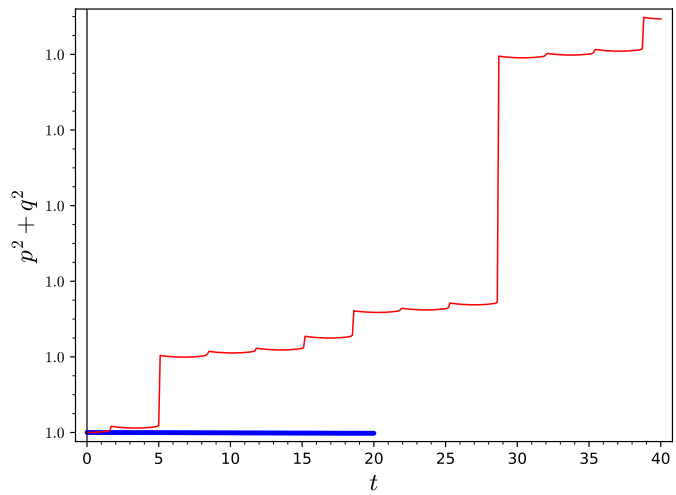


Figure 4. First invariant for rrk method with exact conservation of second invariant $k^2p^2 + r^2 = 1$ (red) and for standard rk4

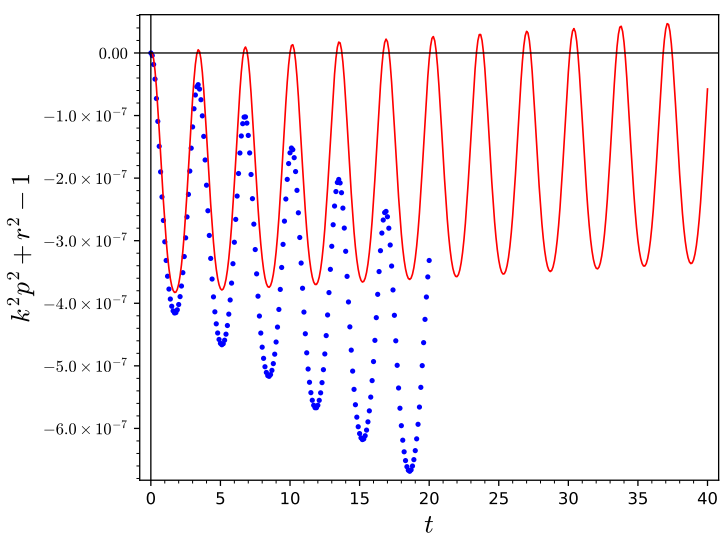


Figure 5. Second invariant for rrk method with exact conservation of first invariant $p^2 + q^2 = 1$ (red) and for standard rk4 (blue)

From the numerical experiments, we came to the conclusion that preserving multiple integrals requires a different approach.

3. The Calvo method for one invariant

3.1. The Calvo method

Let us take two popular explicit RK methods for examples: the RK4 and Euler methods. RK4 method has 4 stages and 4th approximation order. We calculate four axillary quantities at nth step

$$\begin{aligned} k_1&=x^n,\\ k_2&=x^n+\tfrac{1}{2}f(k_1)\Delta t,\\ k_3&=x^n+\tfrac{1}{2}f(k_2)\Delta t,\\ k_4&=x^n+f(k_3)\Delta t \end{aligned}$$

and then the quantity

$$\phi(x^n)=x^n+\Delta t\left(\frac{1}{6}f(k_1)+\frac{1}{3}f(k_2)+\frac{1}{3}f(k_3)+\frac{1}{6}f(k_4)\right) \tag{12}$$

which is used in standard way as x^{n+1}. Similarly, by the Euler method, we calculate in the step n one axillary quantity $k_1=x^n$ and the quantity

$$\phi_1(x^n)=x^n+f(k_1)\Delta t \tag{13}$$

which is used in standard way as x^{n+1}. We can describe this scheme by Butcher table with additional row:

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
1	0	0	1	0
ϕ	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
ϕ_1	1	0	0	0

Calvo et al. [11] extend the Buono method by coupling these two schemes: in the step n we take

$$x_\lambda^{n+1}=\phi(x^n)-\lambda_n(\phi(x^n)-\phi_1(x^n)), \tag{14}$$

where $\lambda_n\in\mathbb{R}$ is the scalar parameter that can be determined by the conservation law, i.e.

$$I(x_\lambda^{n+1})=I(x^n). \tag{15}$$

Using (12) and (13) we have

$$\begin{aligned} x_\lambda^{n+1}=x^n+\left(\frac{1}{6}-\frac{5}{6}\lambda_n\right)f(k_1)\Delta t+\left(\frac{1}{3}+\frac{1}{3}\lambda_n\right)f(k_2)\Delta t+\\ +\left(\frac{1}{3}+\frac{1}{3}\lambda_n\right)f(k_3)\Delta t+\left(\frac{1}{6}+\frac{1}{6}\lambda_n\right)f(k_4)\Delta t. \end{aligned} \tag{16}$$

Thus, (15) gives an algebraic equation to determinating the parameter λ_n at each step. Calvo et al [11] proved that approximation order of this new scheme is equal to 3. They investigated more general case when coupling any two explicit RK scheme.

If I is a quadratic invariant then we have a quadratic equation for λ_n, one of the roots of which goes to 0 at $\Delta t \to 0$ and the other goes to ∞. In numerical experiments we choose the parameter λ_n as real number which is close to the value 0. In CAS sagemath, we use symbolic calculation to solve this equation (15) with respect to λ_n, and use the function `roots` to get this value, since our calculation is performed in the ring $\mathbb{R}$ (Real Field with 53 bits of precision). So, there is a small error, since in $\overline{\mathbb{Q}}$ will get the exact value, but it is very time-consuming.

3.2. Elliptic function test

We implement the described scheme in Sage as the function `project_1()`.

Function `project_1()` returns the numerical points $(0, x^0), (\Delta t, x^1), \cdots$ with the parameters:

— `list_of_integral` is an invariant required to be conserved;
— `F` is the right sides of the system (1)
— `ics` is the initial condition
— dt is the step size,T is the end point of time t.

Let us take the elliptic function for example.

```
sage: load('ex.sage')
sage: var('p,q,r,k')
sage: k=1/2
sage: list_of_integral=[p^2+q^2,k^2*p^2+r^2]
sage: F=[r*q,-p*r,-k^2*p*q]
sage: ics=[p==0,q==1,r==1]
sage: B1=project_1(list_of_integral[0],F,ics,dt=0.1,T=20)
sage: max([abs(p^2+q^2-1) for [t,p,q,r] in B1])
sage: P=line([[t,p] for [t,p,q,r] in
 ↪  B1],axes_labels=['$t$','$p$'],tick_formatter=[None,RR(10e-
 ↪  15).n(digits=1)],frame=true, color='red')+point([[t,p] for
 ↪  [t,p,q,r] in B1])
sage: P1=line([[t,p^2+q^2-1] for [t,p,q,r] in
 ↪  B1],axes_labels=['$t$','$p^2+q^2-
 ↪  1$'],tick_formatter=[None,RR(10e-
 ↪  15).n(digits=1)],frame=true,
 ↪  color='red')
sage: P2=line([[t,k^2*p^2+r^2-1] for [t,p,q,r] in
 ↪  B1],axes_labels=['$t$','$k^2*p^2+r^2-
 ↪  1$'],tick_formatter=[None,RR(10e-
 ↪  15).n(digits=1)],frame=true)
sage: B2=project_1(list_of_integral[1],F,ics,dt=0.1,T=20)
sage: max([abs(p^2+q^2-1) for [t,p,q,r] in B2])
sage: P0=line([[t,p] for [t,p,q,r] in
 ↪  B2],axes_labels=['$t$','$p$'],tick_formatter=[None,RR(10e-
 ↪  15).n(digits=1)],frame=true, color='red')+point([[t,p] for
 ↪  [t,p,q,r] in B2])
```

```
sage: P01=line([[t,p^2+q^2-1] for [t,p,q,r] in
  ↪ B2],axes_labels=['$t$','$p^2+q^2-
  ↪ 1$'],tick_formatter=[None,RR(10e-
  ↪ 15).n(digits=1)],frame=true)
sage: P02=line([[t,k^2*p^2+r^2-1] for [t,p,q,r] in
  ↪ B2],axes_labels=['$t$','$k^2*p^2+r^2-
  ↪ 1$'],tick_formatter=[None,RR(10e-
  ↪ 16).n(digits=1)],frame=true)
```

In figure 6 we can see that the solution founded by the Calvo method with exact conservation of the first invariant $p^2 + q^2 = 1$ and the second invariant $k^2p^2 + r^2 = 1$ gives a condensation of greed points in those arches of the graph where the curvature has a maximum.

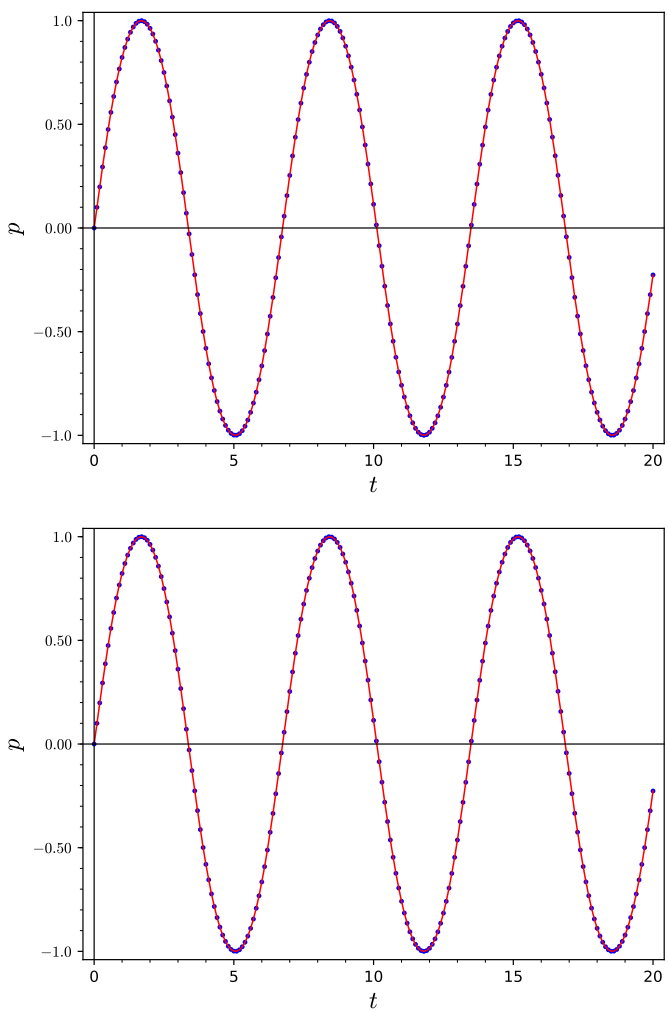


Figure 6. Graph of $p(t)$, for the Calvo method with exact conservation of $p^2 + q^2 = 1$, $dt = 0.1$(up one) and that of $k^2p^2 + r^2 = 1$ (down one)

The second integral $k^2p^2 + r^2 = 1$ in not be exactly preserved, but its value fluctuates with a small amplitude 10^{-7} (figure 7) by the Calvo method with exact conservation of the first invariant $p^2 + q^2 = 1$. The Calvo method with exact conservation of the second invariant $k^2p^2 + r^2 = 1$ shows that the first invariant $p^2 + q^2 = 1$ grows with an error of no more than 10^{-5} (figure 8). Thus, preserving one integral does not preserve the other.

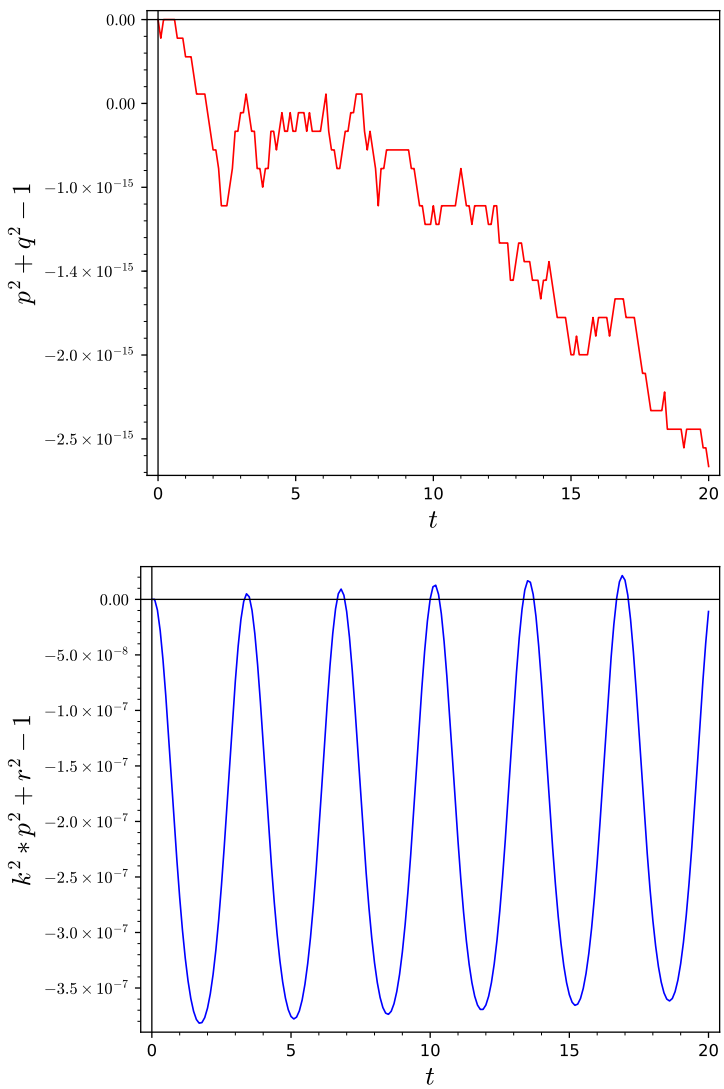


Figure 7. First and second integrals for the Calvo method with exact conservation of $p^2 + q^2 = 1$

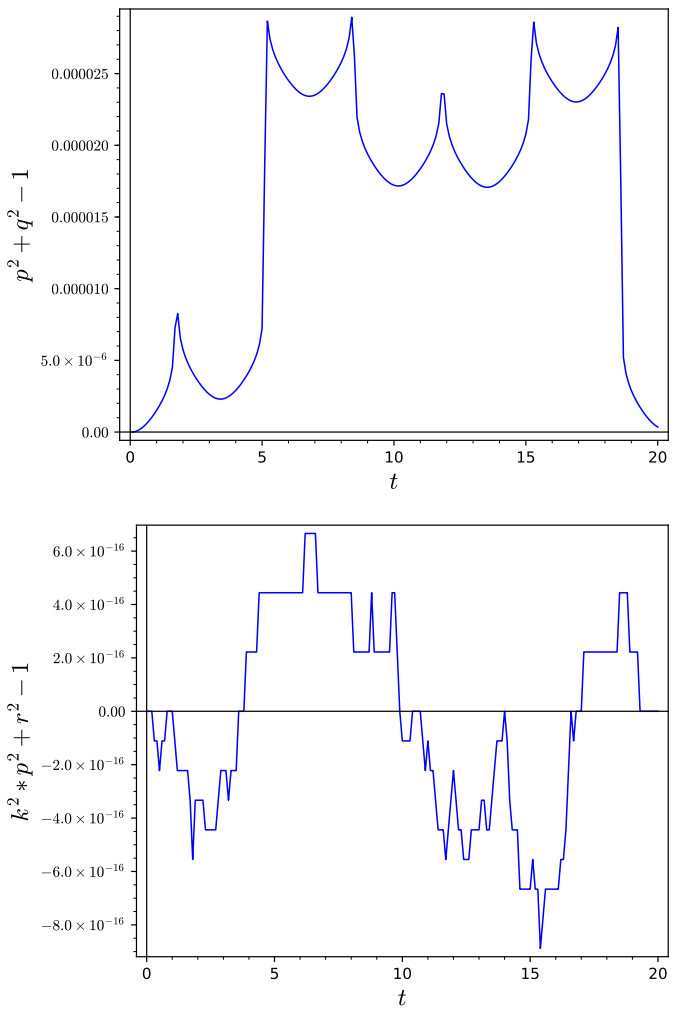


Figure 8. First and second integrals for the Calvo method with exact conservation of $k^2p^2+r^2=1$

4. Scheme for preserving two invariants

Theoretically, the Calvo method allows to construct schemes that preserve several invariants. We take two pairs of RK methods defined by the following two extended Butcher tables:

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
1	0	0	1	0
ϕ	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
ϕ_1	1	0	0	0

and

0	0	0	0	0
$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
$\frac{1}{2}$	0	$\frac{1}{2}$	0	0
0	1	0	1	0
ϕ	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{6}$
ϕ_2	0	1	0	0

The first embedded RK method corresponding to the first table is constructed by combining the standard RK4 method, which has order 4, with the Euler method, which has order 1. The second embedded RK method corresponding to second table is constructed by combining the standard RK4 method, which has order 4, with a second order explicit RK scheme.

We are trying to find an explicit RK scheme of the type:

$$x_{\alpha\beta}^{n+1} = \phi(x^n) - \alpha_n(\phi(x^n) - \phi_1(x^n)) - \beta_n(\phi(x^n) - \phi_2(x^n)), \qquad (17)$$

where $\alpha_n, \beta_n \in \mathbb{R}$ are two scalar parameters, which can be determined by using the conservation laws, i.e.

$$I_i(x_{\alpha\beta}^{n+1}) = I_i(x^n), \quad i = 1, 2, \qquad (18)$$

where I_1 and I_2 are two invariants of system (1).

By definition, we have

$$\phi(x^n) - \phi_1(x^n) = \left(-\frac{5}{6}f(k_1) + \frac{1}{3}f(k_2) + \frac{1}{3}f(k_3) + \frac{1}{6}f(k_4)\right)\Delta t$$

and

$$\phi(x^n) - \phi_2(x^n) = \left(\frac{1}{6}k_1 - \frac{2}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4\right)\Delta t.$$

Thus, (18) gives us a system of two equations with respect to two unknowns α_n and β_n. In numerical experiment, we choose the parameters α_n, β_n as real numbers close to the value 0.

In this way we have a system of algebraic equations for calculating parameters and we must solve this system at each step. Thus, we lose the main advantage of the exact methods. Sage has numerous tools for applying operations on the field of ideals. In our numerical experiments we use the elimination technique based on Gröbner basis [14] to solve (18). Namely, in each step n, after constructing multivariate polynomial ideal in variables α_n, β_n generated by (18), we use Sage built-in function `elimination_ideal` to obtain an univariate equation in variable β_n. The function `roots` over ring $\mathbb{R}$ is used to solve this equation. Substituting one of the value of β_n which is close to 0 to (18), the value of another parameter α_n can be obtained by using the function `roots` again.

Consider elliptic function test.

We construct the function `phi_solve_1()` to implement the routine described above.

We implement the described scheme in Sage as the function `phi_solve_1()`. Function `phi_solve_1()` returns the numerical points $(0, x^0), (\Delta t, x^1), \cdots$ with the parameters:

— `list_of_integral` is two invariants that are required to be conserved;
— `F` is the right sides of the system (1);
— `ics` is the initial condition;
— dt is the step size,T is the end point of time t.

Let us take the elliptic function as example.

```
sage: var('p,q,r,t')
sage: k=1/2
sage: s=4
sage: F=[r*q,-p*r,-k^2*p*q]
sage: list_of_integral=[p^2+q^2,k^2*p^2+r^2]
sage: ics=[p==0,q==1,r==1]
sage: L=phi_solve_1(list_of_integral=list_of_integral, F=F,
 ↪  ics=ics, dt=0.1, T=20)
sage: max([abs(k^2*p^2+r^2-1) for [t,p,q,r] in L])
sage: G4=line([[t,k^2*p^2+r^2-1] for [t,p,q,r] in
 ↪  L],axes_labels=['$t$','$k^2*p^2+r^2-
 ↪  1$'],tick_formatter=[None,RR(10e-
 ↪  20).n(digits=1)],frame=true)
sage: max([abs(p^2+q^2-1) for [t,p,q,r] in L])
sage: G5=line([[t,p^2+q^2-1] for [t,p,q,r] in
 ↪  L],axes_labels=['$t$','$p^2+q^2-
 ↪  1$'],tick_formatter=[None,RR(10e-
 ↪  20).n(digits=1)],frame=true)
```

The Calvo method allows to preserve both invariants and thus significantly surpasses the other methods presented in the previous sections. Figure 9 shows that the error $k^2p^2 + r^2 - 1$ remains constant in size at 10^{-16}, while figure 10 shows the error $p^2 + q^2 - 1$ remains constant in size at 10^{-13}. We can say that these errors are due to the implementation of the calculation in solving equation (18) over the ring $\mathbb{R}$ instead of the algebraic closed field $\overline{\mathbb{Q}}$ in CAS Sage [15]. From this point of view, we can conclude that this method can be considered as a method that exactly preserves exactly both quadratic invariants in the elliptic function test.

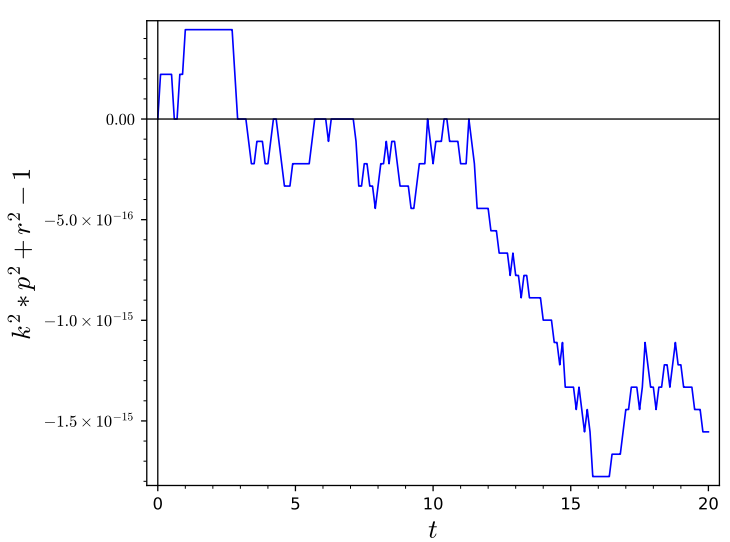


Figure 9. The first integral $k^2p^2 + r^2 - 1$ for $T = 20, dt = 0.1$

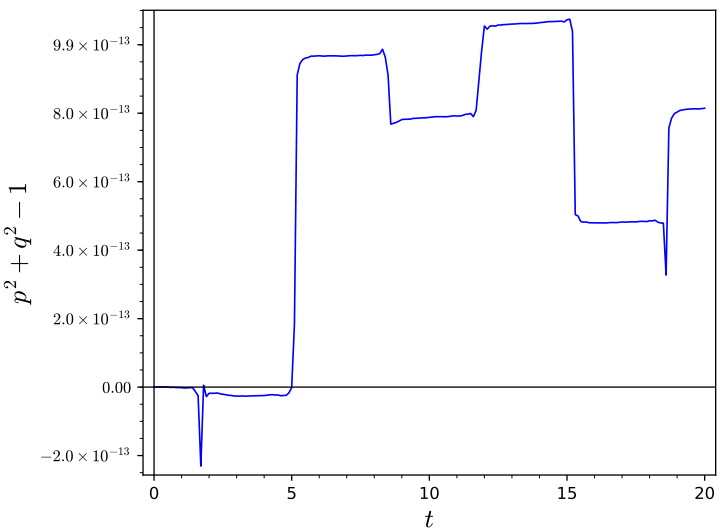


Figure 10. The second integral $p^2 + q^2 - 1$ for $T = 20, dt = 0.1$

5. Conclusion

We have investigated several implementation of explicit RK schemes that preserve invariants. To preserve one invariant the Buono and Calvo methods require solving one algebraic equation with one unknown at each step. In our example, this equation is quadratic, so we can find its numerical solution without any difficulties. From the numerical experiments, we concluded that the exact conservation of one invariant is not an obstacle for changing of the other invariants. Thus, the conservation of multiple integrals requires a different approach.

The Calvo method which preserves several invariants has a drawback: it requires solving a system of algebraic equations with several unknown variables at each step, i.e. has the same drawback as standard implicit RK methods have. Fortunately, in our tests the system we obtained is much simpler than the system described by the midpoint scheme.

Acknowledgments

The authors are thanks Prof. L. A. Sevastianov for his discuss and suggestions with this paper. This paper was supported by the foundation of KaiLi university (Z1602) and by the Russian Foundation for Basic Research (RFBR) according to the research project No 18-07-00567.

References

[1] A. Goriely, *Integrability and nonintegrability of dynamical systems*. Singapore; River Edge, NJ: World Scientific, 2001. DOI: 10.1142/3846.

[2] J. C. Butcher, "On Runge–Kutta processes of high order," *Journal of the Australian Mathematical Society*, vol. 4, no. 2, pp. 179–194, 1964. DOI: 10.1017/S1446788700023387.

[3] E. Hairer, G. Wanner, and S. P. Nørsett, *Solving ordinary differential equations I: Nonstiff Problems*, 3rd ed. New York: Springer, 1993, vol. 1. DOI: 10.1007/978-3-540-78862-1.

[4] K. Burrage and J. C. Butcher, "Stability criteria for implicit Runge–Kutta methods," *SIAM Journal on Numerical Analysis*, vol. 16, no. 1, pp. 46–57, 1979. DOI: 10.1137/0716004.

[5] G. Cooper, "Stability of Runge–Kutta methods for trajectory problems," *IMA journal of numerical analysis*, vol. 7, no. 1, pp. 1–13, 1987. DOI: 10.1093/imanum/7.1.1.

[6] E. Hairer, G. Wanner, and C. Lubich, *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations*. Berlin Heidelberg New York: Springer, 2000.

[7] J.-M. Sanz-Serna and M.-P. Calvo, *Numerical hamiltonian problems*. Courier Dover Publications, 2018.

[8] H. Zhang, X. Qian, J. Yan, and S. Song, "Highly efficient invariant-conserving explicit Runge–Kutta schemes for nonlinear Hamiltonian differential equations," *Journal of Computational Physics*, p. 109 598, 2020.

[9] W. H. Press, B. P. Flannery, S. A. Teukolsky, W. T. Vetterling, *et al.*, *Numerical recipes*. Cambridge University Press Cambridge, 1989.

[10] N. Del Buono and C. Mastroserio, "Explicit methods based on a class of four stage fourth order Runge–Kutta methods for preserving quadratic laws," *Journal of Computational and Applied Mathematics*, vol. 140, pp. 231–243, 2002. DOI: 10.1016/S0377-0427(01)00398-3.

[11] M. Calvo, D. Hernández-Abreu, J. I. Montijano, and L. Rández, "On the preservation of invariants by explicit Runge–Kutta methods," *SIAM Journal on Scientific Computing*, vol. 28, no. 3, pp. 868–885, 2006. DOI: 10.1137/04061979X.

[12] Y. Ying. (2020). "Package ex.sage," [Online]. Available: https://malykhmd.neocities.org.

[13] Y. S. Sikorsky, *Elements of the theory of elliptic functions with applications to mechanics [Elementy teoriy ellipticheskikh funktsiy s prilozheniyami v mekhanike]*. Moscow–Leningrad: ONTI, 1936, In Russian.

[14] D. Cox, J. Little, D. O'Shea, and M. Sweedler, "Ideals, varieties, and algorithms," *American Mathematical Monthly*, vol. 101, no. 6, pp. 582–586, 1994.

[15] The Sage Developers, *Sagemath, the Sage Mathematics Software System (Version 7.4)*, https://www.sagemath.org, 2016.

For citation:

Y. Ying, M. D. Malykh, On the realization of explicit Runge–Kutta schemes preserving quadratic invariants of dynamical systems, Discrete and Continuous

Models and Applied Computational Science 28 (4) (2020) 327–345. DOI: 10.22363/2658-4670-2020-28-4-327-345.

Information about the authors:

Ying, Yu — Assistant Professor of Department of Algebra and Geometry, Kaili University (e-mail: yingy6165@gmail.com, phone: +7(495)9550927)

Malykh, Mikhail D. — doctor of Physical and Mathematical Sciences, Assistant Professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University); (e-mail: malykh_md@pfur.ru, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0001-6541-6603, ResearcherID: P-8123-2016, Scopus Author ID: 6602318510)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2020-28-4-327-345

О реализации явных схем Рунге–Кутты с сохранением квадратичных инвариантов динамических систем

Юй Ин[1], М. Д. Малых[2]

[1] *Университет Кайли*
Кайли, 556011, Китай
[2] *Российский университет дружбы народов*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

Авторами реализовано несколько явных схем Рунге–Кутты, которые сохраняют квадратичные инварианты автономных динамических систем в Sage. В статье представлен пакет ex.sage и результаты численных экспериментов.

В пакете функции rrk_solve, idt_solve и project_1 построены для случая, когда только один заданный квадратичный инвариант будет сохранён точно. Функция phi_solve_1 позволяет сохранить одновременно два указанных квадратичных инварианта. Для решения уравнений относительно параметров, определяемых законом сохранения, использована методика исключения на основе базисов Грёбнера, реализованная в Sage. В качестве тестового примера представленного пакета используется эллиптический осциллятор. Эта динамическая система имеет два квадратичных инварианта. Представлены численные результаты сравнения стандартного явного метода Рунге–Кутты RK(4,4) с rrk_solve. Кроме того, для функций rrk_solve и idt_solve, сохраняющих только один инвариант, исследовано изменение второго квадратичного инварианта эллиптического осциллятора. В заключение рассматриваются недостатки использования этих схем.

Ключевые слова: явный метод Рунге–Кутты, квадратичный инвариант, динамическая система, Sage

Research article

UDC 519.7:338.67
DOI: 10.22363/2658-4670-2020-28-4-346-360

On methods of quantitative analysis of the company's financial indicators under conditions of high risk of investments

Eugeny Yu. Shchetinin

Financial University under the Government of Russian Federation
49, Leningradsky Prospect, Moscow 125993, Russian Federation

(received: October 30, 2020; accepted: November 12, 2020)

The paper investigates the methods of quantitative analysis of hidden statistical relationships of the financial indicators of companies under conditions of high investment risk. A new semi-parametric method for estimating tail dependence indicators using BB1 and BB7 dependence structures is proposed. For a dataset containing the cost indicators of leading Russian companies, computer experiments were carried out, as a result of which it was shown that the proposed method has a higher stability and accuracy in comparison with other considered methods. Practical application of the proposed risk management method would allow financial companies to assess investment risks adequately in the face of extreme events.

Key words and phrases: financial indicators, dependency structures of extremal type, tail ratio, copula

1. Introduction

The global changes currently taking place in the world financial markets caused by the global pandemic of the coronavirus COVID-19, along with a catastrophic decline in oil prices, will obviously entail the need for serious changes in the business structure of both individual companies and entire industries, regional economies and countries. In the Russian conditions of high volatility of financial markets, a successful solution to this problem is impossible without the application of an analysis of the company's resistance to the effects of the external environment, the implementation of constant monitoring of the behavior of a number of indicators of the enterprise profitability. In particular, trigger analysis [1] is one of such effective methods of analysis. The use of trigger analysis to study the sensitivity of a business structure to the impact of disturbing factors is important, since it allows identifying in advance its most weakly protected, most risky business lines, the socalled trigger points. Examples of these are individual lines of business of a company, credit and debt obligations, etc., precisely those areas of activity that potentially pose a threat of an emergency in the company. Therefore, in order to weaken

their influence on the structure as a whole, and on its individual lines, we proposed instead of the trigger point in its vicinity on the scatter diagram to form an aggregate set with the statistical properties of the meta-elliptic type dependence structure [2]. This allows weakening to certain limits the impact of the statistical dependence of the rest of the business lines on the problematic one and redistributing the aggregate damage to the rest of the business structure lines. In addition, the transition from an extreme dependence to a meta-elliptic one simplifies the calculation of structural risk indicators by the usual summation of the marginal risks for this structure. Examples of such an approach to managing extreme risks are debt-restructuring, transfer of fixed assets to subsidiaries and offshore companies, insurance, hedging (issuance of credit derivatives), limiting and securitization operations in the banking sector, etc.

This paper considers effective methods for analyzing the financial performance of companies in the face of increased volatility in stock markets associated with the global COVID-19 pandemic, as well as declining oil prices. Using the example of analyzing the value of shares of leading Russian companies, the advantages of their application are shown in comparison with the classical multivariate analysis using a Gaussian distribution.

2. Methods for the quantitative analysis of structure indicators of statistical dependences

Let (X, Y) be a two-dimensional random variable characterized by a joint distribution function F and partial distribution functions F_1 and F_2, respectively. Then the coefficient of the lower tail dependence and the coefficient of the upper tail dependence are respectively the limits

$$\lambda_L = \lim_{v \to 0+0} P\big(X \leqslant F_1^{-1}(v) | Y \leqslant F_2^{-1}(v)\big)\,, \tag{1}$$

$$\lambda_U = \lim_{v \to 1-0} P\big(X > F_1^{-1}(v) | Y > F_2^{-1}(v)\big)\,. \tag{2}$$

From equations (1)–(2) it obviously follows that the coefficients λ_L and λ_U can take the values within the limits from 0 to 1. The case $\lambda_L > 0$ ($\lambda_U > 0$) is referred to as the presence of a tail dependence or the appearance of a contagion [2] between the random variables X and Y. The situation $\lambda_L = 1$ ($\lambda_U = 1$) corresponds to full contagion [3].

Using the concepts of copula theory [3], we can write expressions for the coefficients of the tail dependence in the following form:

$$\lambda_L = \lim_{v \to 0+0} \frac{C(v, v)}{v}, \tag{3}$$

$$\lambda_U = \lim_{v \to 1-0} \frac{1 - 2v + C(v, v)}{1 - v}, \tag{4}$$

where C is the copula of the joint distribution of random variables. It follows from this representation that the tail dependence coefficients are a property of the dependence structure and do not depend on the partial distributions.

In the general case of d-dimensional distribution of random variables $\left(X^{(1)}, ..., X^{(d)}\right)$, to describe the nature of the extreme dependence using the coefficients of the tail dependence, it is necessary to estimate the $\dfrac{d(d-1)}{2}$ pairs of values of the coefficients λ_{ij}, $1 \leqslant i, j \leqslant d$. When using nonparametric methods, each pair of tail coefficients can be estimated separately, while when using para-metric and semi-parametric approaches, it is necessary to take into account the structure of the relationship between all d components of the multivariate distribution.

3. Nonparametric estimation methods

Let $\mathbf{X_n}, \mathbf{Y_n}$ be samples of independent identically distributed random variables. We define the empirical copula function as

$$C_n(u, v) = \frac{1}{n} \sum_{j=1}^{n} 1_{\{\mathrm{Rank}(X_j) \leqslant n \cdot u, \mathrm{Rank}(Y_j) \leqslant n \cdot v\}}$$

Introducing the notation $U_j = \mathrm{Rank}(X_j)$, $V_j = \mathrm{Rank}(Y_j)$, by definition of tail dependence coefficients we get

$$\hat{\lambda}_{L,n}(k) = \frac{n}{k} \cdot C_n \left(\frac{k}{n}, \frac{k}{n}\right) = \frac{1}{k} \cdot \sum_{j=1}^{n} 1_{\{U_j \leqslant k, V_j \leqslant k\}}, \tag{5}$$

$$\hat{\lambda}_{L,n}(k) = \frac{n}{k} \cdot C_n \left(\left(1 - \frac{k}{n}, 1\right] \times \left(1 - \frac{k}{n}, 1\right]\right) = \\ = \frac{1}{k} \sum_{j=1}^{n} 1_{\{U_j > n-k, V_j > n-k\}} = 2 - \frac{1}{k} \sum_{j=1}^{n} 1_{\{U_j > n-k \text{ or } V_j > n-k\}}, \tag{6}$$

where $C_n\left((a, b] \times (c, d]\right)$ is the empirical probability measure of the copula function on a rectangle $(a, b] \times (c, d]$, $k = k(n) \to \infty$, $\dfrac{k}{n} \to 0$ at $n \to \infty$. The consistency and normality of the estimates were proved in [4]. Logarithmic parametric estimates for λ_L and λ_U can be obtained using equations (3), (4). Thus, the expression for the coefficient of the lower tail dependence can be represented as

$$\lambda_L = \lim_{v \to 0+0} \frac{C(v, v)}{v} = 2 - \lim_{v \to 0+0} \frac{-2v + C(v, v)}{-v} = 2 - \lim_{v \to 0+0} \frac{\ln(1 - 2v + C(v, v))}{\ln(1 - v)}.$$

According to this formula, the value of λ_L can be estimated as

$$\hat{\lambda}_{L,n}^{\mathrm{LOG}}(k) = 2 - \frac{\ln\left(C_n\left(\left(\frac{k}{n}, 1\right] \times \left(\frac{k}{n}, 1\right]\right)\right)}{\ln\left(1 - \frac{k}{n}\right)} = 2 - \frac{\ln\left(\frac{1}{n} \sum\limits_{j=1}^{n} 1_{\{U_j > k, V_j > k\}}\right)}{\ln\left(1 - \frac{k}{n}\right)}, \tag{7}$$

where index $k = k(n) \to \infty$, $\dfrac{k}{n} \to 0$ when $n \to \infty$. Similarly, for the coefficient of the upper tail dependence, the representation

$$\lambda_U = \lim_{v \to 1-0} \frac{1 - 2v + C(v,v)}{1-v} = 2 - \lim_{v \to 1-0} \frac{C(v,v) - 1}{v - 1} = 2 - \lim_{v \to 1-0} \frac{\ln(C(v,v))}{\ln(v)}$$

is valid, as well as the estimate

$$\hat{\lambda}^{\mathrm{LOG}}_{U,n}(k) = 2 - \frac{\ln\left(C_n\left(1 - \dfrac{k}{n}, 1 - \dfrac{k}{n}\right)\right)}{\ln\left(1 - \dfrac{k}{n}\right)} = 2 - \frac{\ln\left(\dfrac{1}{n}\sum\limits_{j=1}^{n} 1\{U_j \leqslant n - k \mathrm{or} V_j \leqslant n - k\}\right)}{\ln\left(1 - \dfrac{k}{n}\right)}, \quad (8)$$

$k = k(n) \to \infty$, $\dfrac{k}{n} \to 0$ at $n \to \infty$. Estimates (7), (8) have the property that they are sharp for all k in the limiting cases of statistical independence and comonotonicity. Indeed,

$$\lambda^{\mathrm{LOG}}_{\mathrm{indep},L} = 2 - \frac{\ln(1 - 2v + \Pi(v,v))}{\ln(1-v)} = 2 - \frac{\ln(1 - 2v + v^2)}{\ln(1-v)} = 0,$$

$$\lambda^{\mathrm{LOG}}_{\mathrm{com},L} = 2 - \frac{\ln(1 - 2v + M(v,v))}{\ln(1-v)} = 2 - \frac{\ln(1 - 2v + v)}{\ln(1-v)} = 1,$$

$$\lambda^{\mathrm{LOG}}_{\mathrm{indep},U} = 2 - \frac{\ln(\Pi(v,v))}{\ln(v)} = 2 - \frac{\ln(v^2)}{\ln(v)} = 0,$$

$$\lambda^{\mathrm{LOG}}_{\mathrm{com},U} = 2 - \frac{\ln(M(v,v))}{\ln(v)} = 2 - \frac{\ln(v)}{\ln(v)} = 1,$$

where $\Pi(u,v) = uv$ is the function of the copula of independent random variables, $M(u,v) = \min(u,v)$ is the copula of comonotonic (completely dependent) random variables. The consistency and asymptotic normality of the estimates was proved in [4]. Figure 1 plots the values of the tail ratios estimates depending on the choice of the threshold k for the joint distribution of fifteen-minute logarithmic increments in the value of Rosneft and Lenta equities in the period from December 15, 2019 to September 30, 2020. In the range of stability $30 \leqslant k \leqslant 70$ of estimates, the value of the tail ratios is significantly greater than zero, which indicates a strong dependence of the investigated financial indicators in the area of extreme values.

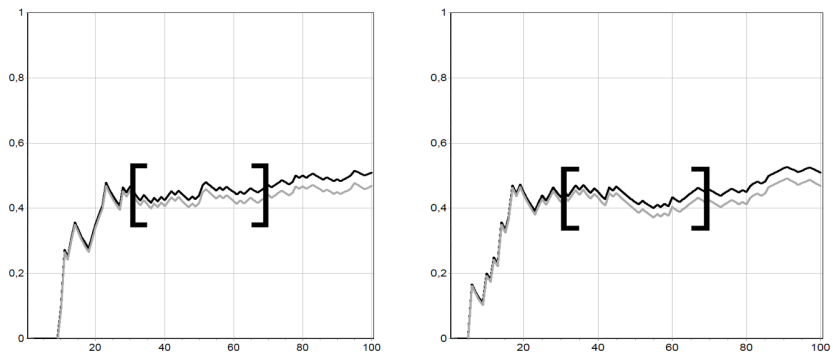


Figure 1. Estimation of indicators of the lower (left) and upper (right) tail coefficients of the joint distribution of the logarithmic increments of the value of Rosneft and Lenta shares depending on the parameter k. ━ standard estimate (6), (7), ━ logarithmic estimate (8), (9)

4. Parametric estimation methods

The parametric approach to estimating the values of tail coefficients is based on the use of a mathematical model of the initial data structure in the form of a parametric copula C_θ. In this case, estimates of the lower and upper coefficients of the tail dependence can be found as a function of the model parameters C_θ: $\hat{\lambda} = \lambda(\hat{\theta})$. To obtain estimates of the parameters of the dependence structure function, it is necessary to pass from the set of initial samples $\mathbf{X}_\mathbf{n}^{(\mathbf{i})} = \left(X_1^{(i)}, ..., X_n^{(i)}\right)$, $1 \leqslant i \leqslant d$ to samples of uniformly distributed on $[0, 1]$ random variables $\mathbf{U}_\mathbf{n}^{(\mathbf{i})} = \left(U_1^{(i)}, ..., U_n^{(i)}\right)$, $1 \leqslant i \leqslant d$. It is necessary to characterize the partial distributions F_i, $1 \leqslant i \leqslant d$. If any of the partial distribution functions is characterized by a set of parameters ϑ_i, $1 \leqslant i \leqslant d$, the approach is referred to as fully parametric. Otherwise, if the partial distributions are replaced with empirical ones (i.e., the ranging operation is applied) the approach is referred to as semiparametric.

Parametric estimation can be done in one or two stages. In the first case, the parameters of the partial distributions ϑ_i, $1 \leqslant i \leqslant d$, as well as the parameters θ of the dependence structure, are estimated together. Typically, the maximum likelihood method is used. In the second case, the assessment takes place in two stages. Due to the very useful property of the copula function, according to which the copula function does not depend on the partial distributions, it is possible to separate the operations of estimating the parameters of the partial distributions ϑ_i, $1 \leqslant i \leqslant d$ and the parameters of the dependency structure θ. At the first stage, the parameters ϑ_i of the partial distribution functions F_i are estimated for $1 \leqslant i \leqslant d$. Then, based on the estimates found, a set of samples is formed

$$\left(\hat{\mathbf{U}}_\mathbf{n}^{(\mathbf{1})}, ..., \hat{\mathbf{U}}_\mathbf{n}^{(\mathbf{d})}\right) = \left(F_{1,\hat{\vartheta}_1}\left(\hat{\mathbf{X}}_\mathbf{n}^{(\mathbf{1})}\right), ..., F_{d,\hat{\vartheta}_d}\left(\hat{\mathbf{X}}_\mathbf{n}^{(\mathbf{d})}\right)\right),$$

using which the parameters θ of the dependency structure are estimated. This method is also known as the pseudo-maximum likelihood method.

The semi-parametric approach implies that partial distributions F_i, $1 \leqslant i \leqslant d$ are replaced by empirical distribution functions. Then, as in the previous method, the parameters θ of the dependency structure are estimated using the maximum likelihood method. As shown in [5], the estimate of the model parameters obtained in this way, as well as the estimate of the tail coefficients, is asymptotically stable and normal. Numerical experiments [4] show that in their properties semi-parametric estimates are almost identical to fully parametric estimates. It should be noted that the use of parametric models of partial distributions F_i, $1 \leqslant i \leqslant d$ can lead to significant errors at the stage of estimating the parameters of the structure of dependence, and, as a consequence, inaccurate and inadequate estimates of the values of the coefficients λ. The semi-parametric approach is more stable in this sense, since it does not have the described disadvantage.

It is convenient to use functions from the number of two-parameter Archimedean copulas [3] as models of the structures of statistical dependence capable of simulating the tail dependence in the two-dimensional case. Below are the expressions for the copula function and the Archimedean generator of models BB1 and BB7

$$C_{BB1}(u,v) = \left(1 + \left((u^{-\theta}-1)^{\delta} + (v^{-\theta}-1)^{\delta}\right)^{\frac{1}{\delta}}\right)^{-\frac{1}{\theta}}, \tag{9}$$

$$\delta \geqslant 1, \theta > 0, \phi_{BB1}(w) = (w^{-\theta}-1)^{\delta};$$

$$C_{BB7}(u,v) = \\ = 1 - \left(1 - \left(\left(1-(1-u)^{\theta}\right)^{-\delta} + \left(1-(1-u)^{\theta}\right)^{-\delta} - 1\right)^{-\frac{1}{\delta}}\right)^{\frac{1}{\theta}}, \tag{10}$$

$$\delta > 0, \quad \theta \geqslant 1, \quad \phi_{BB7}(w) = (1-(1-w)^{\theta})^{-\delta}.$$

These models are convenient in that they allow one to obtain explicit expressions for the tail dependence coefficients. Thus, for model (9)

$$\lambda_U = 2 - 2^{\frac{1}{\delta}}, \quad \lambda_L = 2^{-\frac{1}{\delta\theta}}, \quad \lambda_L = 2^{-\frac{1}{\delta}}, \quad \lambda_U = 2 - 2^{\frac{1}{\theta}}.$$

As can be seen from the last formula, the coefficient of the upper tail dependence λ_U of the model (10) depends only on the model parameter θ and does not depend on the model parameter δ, whereas the coefficient of the lower tail dependence λ_L depends only on the parameter δ and does not depend on θ. This allows parameterization of the model using the coefficients λ_L and λ_U:

$$\delta = -\frac{1}{\log_2 \lambda_L}, \quad \theta = \frac{1}{\log_2(2-\lambda_U)},$$

as well as construction of the modified copula models of Clayton, BB7 and some others [1]. Figure 2 plots the density of the copula BB7 with the corresponding parameters of the tail relationship between the Rosneft and AO Lukoil equities.

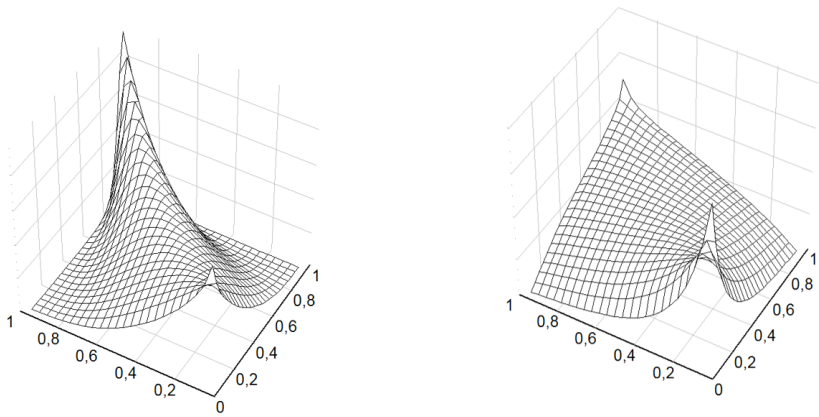


Figure 2. Density of copula BB7, parameterized by the coefficients of the tail dependence. Parameter values: left $\lambda_L = 0.2$, $\lambda_U = 0.8$, right $\lambda_L = 0.5$, $\lambda_U = 0.1$

5. Parametric approach for elliptic distributions

The case when the investigated distribution $\left(X^{(1)}, ..., X^{(d)}\right)$ is elliptic deserves a separate consideration. Since distributions of elliptic type are symmetric, it is obvious that for $\forall i, j\ \lambda_{ij,L} = \lambda_{ij,U}$ (below the notation λ_{ij} is used). In [5], it was shown that the coefficient λ_{ij} of the tail dependence between the components $X^{(i)}$ and $X^{(j)}$ depends only on the index v of regular variation of the elliptic generator, if it is regularly changing, and the value of the parameter ρ_{ij}. In this case, the tail dependence coefficient is expressed through these parameters as follows [6]:

$$\lambda_{ij} = 2 \cdot \bar{t}_{v+1}\left(\sqrt{v+1}\sqrt{\frac{1-\rho_{ij}}{1+\rho_{ij}}}\right),$$

where $\bar{t}_{v+1}$ is the survival function of the Student's distribution with $v+1$ degrees of freedom.

When studying the behavior of the logarithmic increments in the value of shares of the largest issuers of the Russian stock market, we found that the joint distribution of these indicators could be described by the multidimensional Student's distribution, which is known to belong to the class of elliptic distributions [7]. We selected shares of companies Rosneft, Lukoil, Lenta, Mosenergo, Rostelecom; the data were obtained on the website of the Finam company [8]. The results of estimating the parameters of the joint distribution are shown in table 1. The values correspond to the period from 12/15/2019 to 09/30/2020 (ARCH (2) time series model). Table 2 shows the tail dependence matrix calculated from these values, containing pairwise coefficients λ_{ij}.

The values of λ_{ij} in table 2 are significantly greater than zero, which indicates the presence of an extreme type dependence structure. The investigated

statistical dependence cannot be correctly described by the multivariate normal distribution. In practice, this can lead to a significant underestimation of the risks of extremely high losses when investing in this block of shares.

Table 1

Parameters of joint distribution of logarithmic increments in the value of shares of Russian companies

Extreme index $v = 4,1$					
	Correlation matrix				
	Lukoil	Mosenergo	Rostelecom	Lenta	Rosneft
Lukoil	1	0,59	0,70	0,72	0,70
Mosenergo	0,59	1	0,57	0,54	0,48
Rostelecom	0,70	0,57	1	0,67	0,60
Lenta	0,72	0,54	0,67	1	0,62
Rosneft	0,70	0,48	0,60	0,62	1

Table 2

Tail dependence matrix of the logarithmic increments distribution of the value of shares of Russian companies

	Lukoil	Mosenergo	Rostelecom	Lenta	Rosneft
Lukoil	1	0,30	0,39	0,40	0,39
Mosenergo	0,30	1	0,29	0,27	0,24
Rostelecom	0,39	0,29	1	0,36	0,31
Lenta	0,40	0,27	0,36	1	0,32
Rosneft	0,39	0,24	0,31	0,32	1

6. Methods for estimating tail dependence indicators using the theory of extreme values

Let us consider the case when the considered distribution function of random variables F lies in the attraction domain of the distribution of extreme values G [3]: $F \in DA(G)$, i.e.,

$$\lim_{n\to\infty} P\left(\frac{\max_{1\leqslant j\leqslant n} X_n - b_n}{a_n} \leqslant x, \frac{\max_{1\leqslant j\leqslant n} Y_n - d_n}{c_n} \leqslant y\right) = \\ = \lim_{n\to\infty} F^n(a_n x + b_n, c_n y + d_n) = G(x, y), \quad (11)$$

where $a_n, c_n > 0$, $b_n, d_n \in \mathbb{R}$. Let us introduce the concept of the limiting indicator of the tail dependence: $\lambda_U^{EV} = \lim_{v\to 1-0} P(X > G_1^{-1}(v) | Y > G_2^{-1}(v))$, where G_1, G_2 – partial distribution functions of the extreme joint distribution function G. Then the following theorem is true:

Theorem 1. *For a distribution function lying in the attraction domain of the distribution function of extreme values G, the tail dependence index λ_U coincides with the extreme tail dependence index λ_U^{EV}.*

Proof. We use the transformation proposed in [9], passing to the distribution

$$F_*(x,y) = F\left(\left(\frac{1}{1-F_1}\right)^{-1}(x), \left(\frac{1}{1-F_2}\right)^{-1}(y)\right),$$

that lies in the attraction domain of the distribution

$$G_*(x,y) = G\left(\left(\frac{1}{-\ln G_1}\right)^{-1}(x), \left(\frac{1}{-\ln G_2}\right)^{-1}(y)\right). \tag{12}$$

The partial distributions G_{*1},G_{*2} have the form of Frechet distribution functions with the parameter $\gamma = 1$ (standard Frechet distribution):

$$G_{*1} = G_{*2} = \Phi_1(x) \begin{cases} e^{-\frac{1}{x}}, & x > 0, \\ 0, & x \leqslant 0. \end{cases}$$

Limit relation (**??**) for functions F_* and G_* takes the form

$$\lim_{n\to\infty} F_*^n(nx, ny) = G_*(x,y),$$

or, similarly

$$\lim_{n\to\infty} n\,(1 - F_*(nx, ny)) = -\ln G_*(x,y). \tag{13}$$

Omitting the intermediate derivations, we obtain from equation (13)

$$-\ln G_*(t,t) = \lim_{n\to\infty} n\left(1 - F\left(\left(\frac{1}{1-F_1}\right)^{-1}(nt), \left(\frac{1}{1-F_2}\right)^{-1}(nt)\right)\right) =$$
$$= \lim_{n\to\infty} n\left(\frac{1}{nt} + \frac{1}{nt} - \lambda_U \frac{1}{nt}\right) = \frac{2-\lambda_U}{t}.$$

By definition (12), the distribution of extreme values G_* is obtained from the original distribution G using a monotonic transformation. This means that the copula functions of these distributions coincide. Using this fact and the result obtained above for $\ln G_*(t,t)$, we have:

$$\lambda_U^{EV} = \lim_{v\to 1-0} \frac{1 - 2v + C_{EV}(v,v)}{1-v} = \lim_{v\to 1-0} \frac{1 - 2v + G_*\left(-\frac{1}{\ln v}, -\frac{1}{\ln v}\right)}{1-v} =$$

$$= \lim_{v \to 1-0} \frac{1 - 2v + e^{\ln v(2-\lambda_U)}}{1-v} = \lim_{v \to 1-0} \frac{1 - 2v + v^{2-\lambda_U}}{1-v} = \lambda_U.$$

A similar result is, of course, also valid for the lower tail coefficient and the limiting distribution of block minima. This property allows, when estimating the tail coefficient, to go from the entire available sample to its extreme values (block extrema). At the same time, a distinction is made between semi- and fully parametric approaches to estimating the parameters of the resulting distribution. In the case of using a semi-parametric approach, the parameters of the limiting structure of dependence [3] are estimated from

$$\left(\hat{U}_{\max i}, \hat{V}_{\max i}\right) = \left(\mathrm{Rank}\left(\hat{X}_{\max i}\right), \mathrm{Rank}\left(\hat{Y}_{\max i}\right)\right), \quad 1 \leqslant i \leqslant \left\lfloor \frac{n}{k} \right\rfloor,$$

where $\hat{X}_{\max i} = \bigvee\limits_{j=1}^{k} X_{(i-1)k+j}$, $\hat{Y}_{\max i} = \bigvee\limits_{j=1}^{k} Y_{(i-1)k+j}$, $1 \leqslant k < n$, $1 \leqslant i \leqslant \left\lfloor \frac{n}{k} \right\rfloor$. As a model of the structure of the dependence of the limiting distribution of extrema, we have proposed various functions of copula of the extreme type, in particular, the Gumbel model (logistic model):

$$C_{EV}(u,v) = \exp\left(-\left((-\ln u)^R + (-\ln v)^R\right)^{\frac{1}{R}}\right), \quad R \geqslant 1. \tag{14}$$

The estimate of the coefficient of the upper tail dependence for model (14) has the form $\lambda_U^{EV}(k) = 2 - 2^{\frac{1}{R_U(k)}}$. The formulas for estimating the coefficient of the lower tail dependence are completely similar. They were obtained by passing to the distribution of the maxima of the quantities $(-X, -Y)$:

$$\hat{X}_{\min i} = \max_{1 \leqslant j \leqslant n}\left(-X_{(i-1)k+j}\right), \quad \hat{Y}_{\min i} = \max_{1 \leqslant j \leqslant n}\left(-Y_{(i-1)k+j}\right),$$

$$1 \leqslant k < n, \quad 1 \leqslant i \leqslant \left\lfloor \frac{n}{k} \right\rfloor,$$

$$\left(\hat{U}_{\min i}, \hat{V}_{\min i}\right) = \left(\mathrm{Rank}\left(\hat{X}_{\min i}\right), \mathrm{Rank}\left(\hat{y}_{\min i}\right)\right), \quad 1 \leqslant i \leqslant \left\lfloor \frac{n}{k} \right\rfloor,$$

$$\lambda_L(k) = 2 - 2^{\frac{1}{R_L(k)}}.$$

An alternative approach to characterizing the distribution function of extreme values is the peaks over threshold (POT) method. A detailed description of this approach can be found in [10]. We only note that the idea of the method is to characterize the limiting distribution of excesses that have exceeded a given threshold value. In [4], it was proposed to evaluate the value of the upper tail coefficient by considering the values of the points of the investigated structure of the dependence (U, V) lying in the region $[t, 1]^2$, where t is the threshold parameter tending to 1:

$$\left(\hat{\mathbf{U}}_{\mathbf{U}}^{*}, \hat{\mathbf{V}}_{\mathbf{U}}^{*}\right) = \left\{(U_j, V_j) : (U_j, V_j) \in [t,1] \times [t,1]\right\}, \quad 0 \leqslant t < 1.$$

To estimate the lower tail coefficient, values from the region $[0,t]^2$, $t \to 0$ should be considered:

$$\left(\hat{\mathbf{U}}^*_{\mathbf{L}}, \hat{\mathbf{V}}^*_{\mathbf{L}}\right) = \left\{(U_j, V_j) : (U_j, V_j) \in [0,t] \times [0,t]\right\}, \quad 0 < t \leqslant 1.$$

In [9], a relation was proved that makes it possible to estimate the parameter λ_L in the case when the quantities (X, Y) under consideration are related by the Archimedean structure of the dependence. The authors showed that if a copula C has an Archimedean generator with an index of regular variation $\alpha > 0$, then the lower threshold copula

$$C_{L,t}(u,v) = P(U \leqslant u, V \leqslant v | U \leqslant t, V \leqslant t), \tag{15}$$

converges in the limit at $t \to 0$ to the Clayton copula C_α with parameter $\delta = \alpha$:

$$\lim_{t \to 0} C_{L,t}(u,v) = C_{Cl}(u,v), \quad \forall u, v \in [0,1],$$

where $C_{Cl}(u,v) = \left(u^{-\delta} + v^{-\delta} - 1\right)^{-\frac{1}{\delta}}$ is an Archimedean copula with generator $\phi(w) = w^{-\delta} - 1$, $\delta > 0$. This dependence structure model has a lower tail dependence with coefficient $\lambda_L = 2^{-\frac{1}{\delta}}$.

In the case when the type of the investigated structure of dependence (U, V) is not Archimedean, one should choose another model of the limiting copula capable of simulating the tail dependence. For example, it is convenient to use the Gumbel model in the form (14) to model the upper threshold copula

$$C_{U,t}(u,v) = P(U > u, V > v | U > t, V > t), \quad 0 \leqslant t < 1, \tag{16}$$

and the inverse Gumbel copula $\overline{C}_{EV}(u,v) = C_{EV}(1-u, 1-v) + u + v - 1$ for modeling the lower threshold copula (15).

The disadvantage of the threshold method is that when the threshold parameter t is close to 0 (when estimating the coefficient of the lower tail dependence) and 1 (when estimating the coefficient of the upper tail dependence), too few points fall into the region where the threshold is exceeded, which makes it impossible to estimate tail parameters with sufficient accuracy. If the threshold is too low, the extreme copula (14) ceases to be an adequate model of the truncated structure of the dependence, which leads to significant systematic errors. The so-called bias-variance problem arises. To solve it, it was proposed to use a combined threshold approach [3]. The essence of the method is that the threshold model of the structure of dependence could adequately describe both the behavior of the entire sample and its points lying in the region of extreme values for different values of the parameters. Then, even at low values of the threshold t, one can count on the stability of the estimates obtained. We propose to use models (9), (10) as such flexible structures of dependence. These models are very versatile and make it possible to simulate both the integral structure of dependence and its behavior in extreme areas.

To test the described models of tail dependence, we used samples $\mathbf{X_n}$, $\mathbf{Y_n}$ from a two-dimensional distribution with the structure of a mixed-type dependence

$$C_{mix}(u,v) = \frac{1}{2}C_{Cl_1}(u,v) + \frac{1}{2}\overline{C}_{Cl_2}(u,v), \tag{17}$$

where C_{Cl_1} is the Clayton copula with parameter $\delta_1 = 3.11$, $\overline{C}_{Cl_2}$ is the inverse Clayton copula with the parameter $\delta_2 = 1.36$. The coefficients of the lower and upper tail dependences of copula (17) are, respectively, equal to

$$\lambda_L^* = \frac{1}{2}\lambda_{L,1} + \frac{1}{2}\lambda_{L,2} = \frac{1}{2}\cdot 2^{-\frac{1}{\delta_1}} + \frac{1}{2}\cdot 0 = 0.40,$$

$$\lambda_U^* = \frac{1}{2}\lambda_{U,1} + \frac{1}{2}\lambda_{U,2} = \frac{1}{2}\cdot 0 + \frac{1}{2}\cdot 2^{-\frac{1}{\delta_2}} = 0.30.$$

Figures 3, 4 show the plots of the estimates obtained by the block extremum method and the threshold method. As can be seen, the combined threshold method provides greater stability of estimates depending on the parameter t.

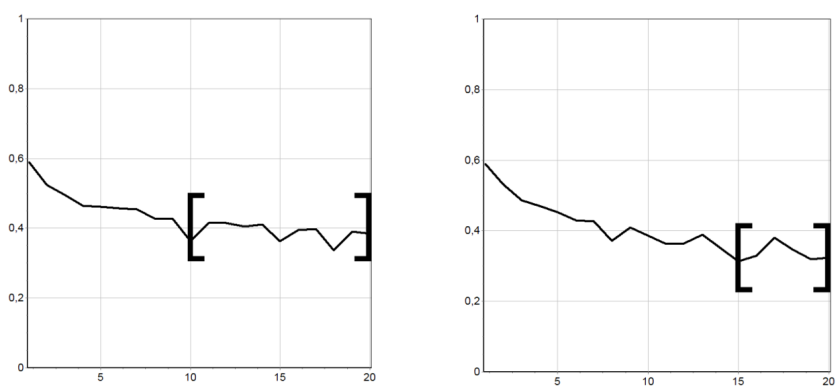


Figure 3. Estimation of indicators of tail distribution coefficients of block minima (left) and block maxima (right) depending on the parameter k

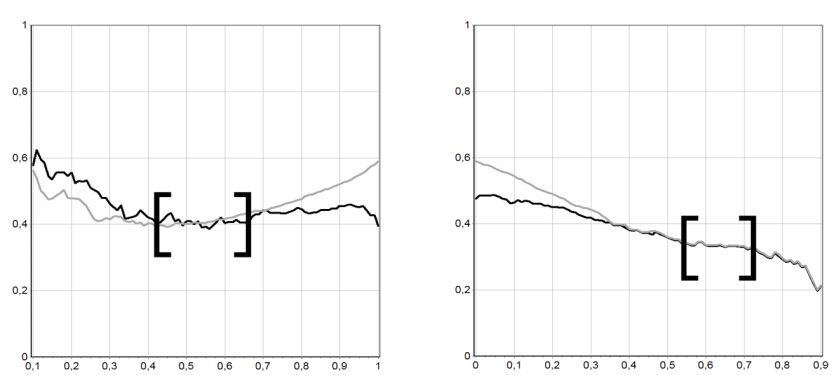


Figure 4. Estimation of indicators of the lower (left) and upper (right) tail distribution coefficients depending on the threshold parameter t. ━ model BB1, ━ model Gumbel

7. Conclusion

This paper proposes and investigates methods for analyzing the financial performance of companies in the context of increased volatility in stock markets associated with the global COVID-19 pandemic, as well as a decline in oil prices during the period 12/15/2019 – 09/30/2020. Computer experiments are carried out to assess the riskiness of investments in leading Russian companies and to analyze the value of shares of leading Russian companies, and the advantages of their application are shown in comparison with the classical multivariate analysis using a Gaussian distribution. Analysis of the properties of estimates of the tail dependence coefficients showed their high sensitivity to extreme changes in the value of companies' shares, which makes it possible to use them as indicators of the occurrence of extreme events in the stock markets and to make timely decisions on the management of their investment projects.

References

[1] E. Y. Shchetinin, "On new approaches to company management in emergency situations [Novyh podhodah k upravleniyu kompaniej v chrezvychajnyh situaciyah]," *Finansy i kredit*, vol. 30, no. 198, pp. 71–75, 2005, In Russian.

[2] E. Y. Shchetinin, K. M. Nazarenko, and A. V. Paramonov, "Instrumental methods of stochastic analysis of extreme events [Instrumental'nye metody stohasticheskogo analiza ekstremal'nyh sobytij]," *Vestnik NNGU, Matematicheskoye modelirovaniye i optimal'noye upravleniye*, vol. 2, no. 29, pp. 56–63, 2004, In Russian.

[3] V. A. Akimov, A. A. Bykov, and E. Y. Shchetinin, *Introduction to statistics of extreme values, EMERCOM of Russia [Vvedenie v statistiku ekstremal'nyh znachenij i ee prilozheniya]*. Moscow: FGU VNII GOChS (FTs), 2009, 524 pp., In Russian.

[4] R. Schmidt and U. Stadtmüller, "Non-parametric estimation of tail dependence," *Scandinavian Journal of Statistics*, vol. 33, no. 2, pp. 307–335, 2006.

[5] R. Schmidt, *Tail dependence. In Statistical tools in finance and insurance*, W. Hardle, P. Cizek, and R. Weron, Eds. Springer Verlag, 2003.

[6] G. Frahm, M. Junker, and A. Szimayer, "Elliptical copulas: Applicability and limitations," *Statistics & Probability Letters*, vol. 63, no. 3, pp. 275–286, 2003. DOI: 10.1016/S0167-7152(03)00092-0.

[7] S. Resnick, *Extreme values, regular variation and point processes*. Berlin: Springer, 1987.

[8] E. Y. Shchetinin, "Vine copulas structures modeling on Russian stock market," *Discrete and Continuous Models and Applied Computational Science*, vol. 27, no. 4, pp. 343–354, 2019. DOI: 10.22363/2658-4670-2019-27-4-343-354.

[9] E. Y. Shchetinin, "Modeling of D-branching structures in the Russian stock market [Modelirovanie struktur D-vetvlenij na rossijskom fondovom rynke]," *Vestnik komp'yuternykh i informatsionnykh tekhnologiy*, vol. 8, no. 182, pp. 38–45, 2019, In Russian.

[10] C. Genest, K. Ghoudi, and L. Rivest, "A semiparametric estimation procedure of dependence parameters in multivariate families of distributions," *Biometrika*, vol. 82, no. 3, pp. 543–552, 1995. DOI: 10.1093/biomet/82.3.543.

For citation:

E. Yu. Shchetinin, On methods of quantitative analysis of the company's financial indicators under conditions of high risk of investments, Discrete and Continuous Models and Applied Computational Science 28 (4) (2020) 346–360. DOI: 10.22363/2658-4670-2020-28-4-346-360.

Information about the authors:

Shchetinin, Eugeny Yu. — Doctor of Physical and Mathematical Sciences, lecturer of Department of Data Analysis, Decision Making and Financial Technologies (e-mail: riviera-molto@mail.ru, phone: +7(917)5390698, ORCID: https://orcid.org/0000-0003-3651-7629, ResearcherID: O-8287-2017, Scopus Author ID: 16408533100)

УДК 519.7:338.67
DOI: 10.22363/2658-4670-2020-28-4-346-360

О методах количественного анализа финансовых показателей компании в условиях высокой рискованности инвестиций

Е. Ю. Щетинин

Финансовый университет при Правительстве Российской Федерации
Ленинградский проспект, д. 49, Москва, 125993, Россия

В работе исследованы методы количественного анализа скрытых статистических связей финансовых показателей компаний в условиях высокой рискованности инвестирования. Предложен новый полупараметрический метод оценивания показателей хвостовой зависимости с использованием моделей структур зависимости BB1 и BB7. Для набора данных, содержащих стоимостные показатели ведущих российских компаний, проведены компьютерные эксперименты, в результате которых показано, что предложенный метод обладает более высокой устойчивостью и точностью по сравнению с другими рассмотренными методами. Практическое применение представленного метода управления рисками позволило бы финансовым компаниям адекватно оценивать инвестиционные риски в условиях наступления экстремальных событий.

Ключевые слова: финансовые показатели, глубокие статистические связи, структуры зависимости, хвостовой коэффициент, копула

Research article

UDC 519.61:519.62:621.372.82
DOI: 10.22363/2658-4670-2020-28-4-361-377

Single-mode propagation of adiabatic guided modes in smoothly irregular integral optical waveguides

Anton L. Sevastianov

Peoples' Friendship University of Russia (RUDN University)
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: June 10, 2020; accepted: November 12, 2020)

This paper investigates the waveguide propagation of polarized electromagnetic radiation in a thin-film integral optical waveguide. To describe this propagation, the adiabatic approximation of solutions of Maxwell's equations is used. The construction of a reduced model for adiabatic waveguide modes that retains all the properties of the corresponding approximate solutions of the Maxwell system of equations was carried out by the author in a previous publication in DCM & ACS, 2020, No 3. In this work, for a special case when the geometry of the waveguide and the electromagnetic field are invariant in the transverse direction. In this case, there are separate nontrivial TE- and TM-polarized solutions of this reduced model. The paper describes the parametrically dependent on longitudinal coordinates solutions of problems for eigenvalues and eigenfunctions – adiabatic waveguide TE and TM polarizations. In this work, we present a statement of the problem of finding solutions to the model of adiabatic waveguide modes that describe the stationary propagation of electromagnetic radiation. The paper presents solutions for the single-mode propagation of TE and TM polarized adiabatic waveguide waves.

Key words and phrases: waveguide propagation of polarized light, integral optical waveguide, adiabatic approximation, eigenvalues and eigenfunctions, Kantorovich method, single-mode regime

1. Introduction

In works [1]–[5] a cycle of studies of the propagation of polarized light in integrated-optical smoothly irregular thin-film waveguides was carried out within the framework of the model of adiabatic waveguide waves. They showed the advantages of the model and its advantages over other models in the description of open dielectric waveguides [6]–[8]. At the same time, until recently, the question of substantiating this model remained open. In work [9] the substantiation of the model was carried out, which is a reduction of a more complex in use general model based on Maxwell's equations. In the present work, within the framework of the model of adiabatic waveguide waves, the problem of stationary propagation of polarized light in a smoothly irregular

integral-optical waveguide is posed, an auxiliary problem for eigenvalues and eigenfunctions (adiabatic waveguide modes) is formulated and solved. The solution of the stationary problem by the generalized Kantorovich method is proposed, its solution is obtained in the single-mode propagation mode.

2. Basic concepts and notation

Waveguide propagation of monochromatic polarized electromagnetic radiation in integrated optical waveguides is described by Maxwell's equations. The electromagnetic field is described using complex amplitudes. The material environment is considered, consisting of dielectric subdomains that fill the entire three-dimensional space. The latter means that the dielectric constants of the subdomains are different and real, and the magnetic permeability is everywhere equal to the magnetic permeability of the vacuum. It follows from the foregoing that in the absence of external currents and charges, the induced currents and charges are equal to zero.

In the absence of external charges and currents, the scalar Maxwell equations follow from the vector ones, and the boundary conditions for the normal components follow from the boundary conditions for the tangential components. The constitutive equations of connection in the case under consideration are assumed to be linear. Thus, the electromagnetic field in a space filled with dielectrics in the Gaussian system of units is described by the equations [10]:

$$\mathrm{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \quad \mathrm{rot}\mathbf{H} = -\frac{1}{c}\frac{\partial \mathbf{D}}{\partial t}, \quad \mathbf{D} = \varepsilon\mathbf{E}, \quad \mathbf{B} = \mu\mathbf{H}, \tag{1}$$

where $\mathbf{E}, \mathbf{H}$ are the vectors of electric and magnetic field strength; $\mathbf{D}$ is the electric displacement vector, $\mathbf{B}$ is the magnetic flux density vector; c is the velocity of electromagnetic wave propagation in vacuum.

In this case, the boundary conditions

$$\mathbf{H}_\tau|_1 = \mathbf{H}_\tau|_2, \quad \mathbf{E}_\tau|_1 = \mathbf{E}_\tau|_2. \tag{2}$$

and the asymptotic boundary conditions at infinity

$$\|\mathbf{E}\| \xrightarrow[|x|\to\infty]{} 0, \quad \|\mathbf{H}\| \xrightarrow[|x|\to\infty]{} 0, \tag{3}$$

are assumed to hold for guided modes, which ensures the uniqueness of the solution to problem (1)–(2).

In equations (1): ε is the permittivity of the medium; μ is the permeability of the medium, $\mathbf{E}, \mathbf{H}$ are the electric and magnetic field strength vectors. We denote by $n = \sqrt{\mu\varepsilon}$ the refractive index of the medium (here and below — of a layer of the multilayer dielectric structure under consideration).

All subdomains are infinite and bonded by planes parallel to the yOz-plane and surfaces, asymptotically parallel to the yOz-plane, so that below we have $\varepsilon = \varepsilon(x)$, $\mu = 1$.

3. Model of adiabatic guided modes

In Ref. [9] the adiabatic approximation of the guided solution of Maxwell's equations is found in the form:

$$\left\{\begin{matrix}\vec{E}(x,y,z,t)\\ \vec{H}(x,y,z,t)\end{matrix}\right\} = \left\{\begin{matrix}\vec{E}_0(x;y,z)\\ \vec{H}_0(x;y,z)\end{matrix}\right\} \exp\left\{i\omega t - ik_0\varphi(y,z)\right\}, \tag{4}$$

where

$$\begin{aligned}
\varepsilon\frac{\partial E_0^y}{\partial x} &= -ik_0\left(\frac{\partial\varphi}{\partial y}\right)\left(\frac{\partial\varphi}{\partial z}\right)H_0^y - ik_0\left(\varepsilon\mu - \left(\frac{\partial\varphi}{\partial y}\right)^2\right)H_0^z,\\
\varepsilon\frac{\partial E_0^z}{\partial x} &= ik_0\left(\varepsilon\mu - \left(\frac{\partial\varphi}{\partial z}\right)^2\right)H_0^y + ik_0\left(\frac{\partial\varphi}{\partial z}\right)\left(\frac{\partial\varphi}{\partial y}\right)H_0^z,\\
\mu\frac{\partial H_0^y}{\partial x} &= ik_0\left(\frac{\partial\varphi}{\partial y}\right)\left(\frac{\partial\varphi}{\partial z}\right)E_0^y + ik_0\left(\varepsilon\mu - \left(\frac{\partial\varphi}{\partial y}\right)^2\right)E_0^z,\\
\mu\frac{\partial H_0^z}{\partial x} &= -ik_0\left(\varepsilon\mu - \left(\frac{\partial\varphi}{\partial z}\right)^2\right)E_0^y - ik_0\left(\frac{\partial\varphi}{\partial z}\right)\left(\frac{\partial\varphi}{\partial y}\right)E_0^z,\\
E_0^x &= -\frac{\partial\varphi}{\partial y}\frac{1}{\varepsilon}H_0^z + \frac{\partial\varphi}{\partial z}\frac{1}{\varepsilon}H_0^y, \quad H_0^x = \frac{\partial\varphi}{\partial y}\frac{1}{\mu}E_0^z - \frac{\partial\varphi}{\partial z}\frac{1}{\mu}E_0^y.
\end{aligned} \tag{5}$$

After additional differentiations from four first-order ODEs, four second-order ODEs are obtained, two of which take the form:

$$\begin{aligned}
\frac{\partial^2 E_0^y}{\partial x^2} + k_0^2\left(\varepsilon\mu - \left(\frac{\partial\varphi}{\partial y}\right)^2 - \left(\frac{\partial\varphi}{\partial z}\right)^2\right)E_0^y = 0,\\
\frac{\partial^2 H_0^y}{\partial x^2} + k_0^2\left(\varepsilon\mu - \left(\frac{\partial\varphi}{\partial y}\right)^2 - \left(\frac{\partial\varphi}{\partial z}\right)^2\right)H_0^y = 0.
\end{aligned} \tag{6}$$

in the case when the layers (three or four) of a multilayer waveguide are homogeneous.

The rest four components are calculated from the system of linear algebraic equations (SLAE)

$$\begin{aligned}
ik_0\varepsilon E_0^x + ik_0\frac{\partial\varphi}{\partial y}H_0^z = ik_0\frac{\partial\varphi}{\partial z}H_0^y, \quad ik_0\varepsilon E_0^z - ik_0\frac{\partial\varphi}{\partial y}H_0^x = \frac{\partial H_0^y}{\partial x},\\
ik_0\frac{\partial\varphi}{\partial y}E_0^z - ik_0\mu H_0^x = ik_0\frac{\partial\varphi}{\partial z}E_0^y, \quad -ik_0\mu H_0^z - ik_0\frac{\partial\varphi}{\partial y}E_0^z = \frac{\partial E_0^y}{\partial x}.
\end{aligned} \tag{7}$$

4. Waveguides regular in y and electromagnetic fields

First, let us consider the case when neither the integrated optical waveguide geometry, nor the solutions to Maxwell's equations for the adiabatic guided mode (AGM) depend on one of the horizontal coordinates, i.e., the case $\partial/\partial y \equiv 0$.

For fields harmonic in time in the Cartesian system of coordinates, the system of Maxwell's equations has the form:

$$\begin{cases} \dfrac{\partial H_z}{\partial y} - \dfrac{\partial H_y}{\partial z} = ik_0\varepsilon E_x, & \dfrac{\partial E_z}{\partial y} - \dfrac{\partial E_y}{\partial z} = -ik_0\mu H_x, \\ \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} = ik_0\varepsilon E_y, & \dfrac{\partial E_x}{\partial z} - \dfrac{\partial E_z}{\partial x} = -ik_0\mu H_y, \\ \dfrac{\partial H_y}{\partial x} - \dfrac{\partial H_x}{\partial y} = ik_0\varepsilon E_z, & \dfrac{\partial E_y}{\partial x} - \dfrac{\partial E_x}{\partial y} = -ik_0\mu H_z, \end{cases} \tag{8}$$

In the case $\partial/\partial y \equiv 0$, system (8) takes the form

$$\begin{cases} -\dfrac{\partial H_y}{\partial z} = ik_0\varepsilon E_x, & -\dfrac{\partial E_y}{\partial z} = -ik_0\mu H_x \\ \dfrac{\partial E_x}{\partial z} - \dfrac{\partial E_z}{\partial x} = -ik_0\mu H_y, & \dfrac{\partial H_x}{\partial z} - \dfrac{\partial H_z}{\partial x} = ik_0\varepsilon E_y \\ \dfrac{\partial H_y}{\partial x} = ik_0\varepsilon E_z, & \dfrac{\partial E_y}{\partial x} = -ik_0\mu H_z \end{cases} \tag{9}$$

A substitution of the first and the third equations (9) into the second ones leads to the equivalent systems

$$\begin{cases} \dfrac{\partial^2 E_y}{\partial z^2} + \dfrac{\partial^2 E_y}{\partial x^2} + k_0^2\varepsilon\mu E_y = 0, \\ -\dfrac{\partial E_y}{\partial z} = -ik_0\mu H_x, \\ \dfrac{\partial E_y}{\partial x} = -ik_0\mu H_z \end{cases} \tag{10}$$

for TE polarization and

$$\begin{cases} \dfrac{\partial^2 H_y}{\partial z^2} + \dfrac{\partial^2 H_y}{\partial x^2} + k_0^2\varepsilon\mu H_y = 0, \\ -\dfrac{\partial H_y}{\partial z} = ik_0\varepsilon E_x \\ \dfrac{\partial H_y}{\partial x} = ik_0\varepsilon E_z \end{cases} \tag{11}$$

for TM polarization.

In this case, the system of equations (6), (7) is split into two independent subsystems:
for TE polarization:

$$\begin{cases} \dfrac{d^2E_0^y}{dx^2} + k_0^2\left(\varepsilon\mu - \left(\dfrac{d\varphi}{dz}\right)^2\right)E_0^y = 0, \\ \mu H_0^x = -\dfrac{d\varphi}{dz}E_0^y, \\ -ik_0\mu H_0^z = \dfrac{dE_0^y}{dx} \end{cases} \tag{12}$$

and for TM polarization:

$$\begin{cases} \dfrac{d^2H_0^y}{dx^2} + k_0^2\left(\varepsilon\mu - \left(\dfrac{d\varphi}{dz}\right)^2\right)H_0^y = 0, \\ ik_0\varepsilon E_0^z = \dfrac{dH_0^y}{dx}, \\ \varepsilon E_0^x = \dfrac{d\varphi}{dz}H_0^y. \end{cases} \tag{13}$$

For thin film multilayer waveguide consisting of optically homogeneous layers, the system of equations (12) and (13) should be completed with the conditions of the electromagnetic field matching at the interfaces between the media that follow from (2):

$$\vec{n} \times \vec{E}^- + \vec{n} \times \vec{E}^+ = 0, \tag{14}$$

$$\vec{n} \times \vec{H}^- + \vec{n} \times \vec{H}^+ = 0 \tag{15}$$

and the asymptotic conditions that follow from (3):

$$E_y^0, E_z^0, H_y^0, H_z^0 \xrightarrow[x\to\pm\infty]{} 0. \tag{16}$$

5. Setting of the physical problem

The solution of the first equation of system (10) is found using the generalized Kantorovich method [11], [12], which is analogous to the method of separation of variables proposed in [13]. We find the solutions to an auxiliary problem analogous to a problem of finding eigenvalues and eigenfunctions that depend on a parameter.

Auxiliary eigenvalue problem

For each fixed z we consider the problem

$$\left(\frac{d^2}{dx^2} + k_0^2 n^2(x,z)\right)E_y^j(x;z) = k_0^2\beta_j^2(z)E_y^j(x;z), \quad -\infty < x < \infty, \tag{17}$$

where $\beta(z) = \dfrac{d\varphi}{dz}(z)$, with the boundary (asymptotic) conditions

$$E_y^j(x;z) \xrightarrow[x\to\pm\infty]{} 0. \tag{18}$$

Let us normalize the eigenfunctions with a condition

$$\langle E_y^j, E_y^k \rangle = \int_{-\infty}^{\infty} E_y^j(x;z)\bar{E}_y^k(x;z)dx = \delta_{jk}. \tag{19}$$

It is known that for any fixed $z \in \mathbb{R}$ problem (17)–(19) is a problem of finding normal guided TE modes of a regular planar reference waveguide [14], [15]. At any real-valued ε, μ and any finite thickness of the reference waveguide it allows a finite number N^{TE} of forward and N^{TE} of backward TE modes [16]–[18].

We restrict ourselves to considering such smoothly irregular waveguides in which the number of guided modes is constant throughout the change and the degree of irregularity is so small that the transformation of the energy of guided modes is limited by the adiabatic approximation. In this case, we seek a solution to problem in the form of an expansion:

$$E_y^0(x,z) = \sum_{-N^{\mathrm{TE}}}^{N^{\mathrm{TE}}} C_j^{\mathrm{TE}}(z)E_y^j(x;z), \tag{20}$$

satisfying the condition

$$\frac{\partial E_y^0}{\partial z}(x,z) = -i\sum_{-N^{\mathrm{TE}}}^{N^{\mathrm{TE}}} k_0\beta_j^{\mathrm{TE}}(z)C_j^{\mathrm{TE}}(z)E_y^j(x;z). \tag{21}$$

Substituting expansion (20),(21) into Eqs. (10) with relations (19) taken into account, after cumbersome but not complicated transformations, we arrive at a system of ODEs for the expansion coefficient functions $C_j^{\mathrm{TE}}(z)$.

6. Solution of a single-mode problem for zero contribution to the AGM

As an example, let us carry out the above calculations for the particular case (20), (21), which describes the single-mode propagation of a TE-polarized AGM, namely:

$$\tilde{E}_y^j(x,z) = C_j^{\mathrm{TE}}(z)E_y^j(x;z), \tag{22}$$

satisfying condition

$$\frac{\partial \tilde{E}_y^j}{\partial z}(x,z) = -ik_0\beta_j^{\mathrm{TE}}(z)C_j^{\mathrm{TE}}(z)E_y^j(x;z). \tag{23}$$

We differentiate (23) with respect to z and substitute the result into the first equation of the system (10) taking into account (17) (following [13]). As a result, we arrive at a system consisting of a single ODE

$$\frac{dC_j^{\mathrm{TE}}}{dz}(z)\beta_j^{\mathrm{TE}}(z) + \frac{d\beta_j^{\mathrm{TE}}}{dz}(z)C_j^{\mathrm{TE}}(z) = 0. \tag{24}$$

We substitute the solution of Eq. (24) into relation (22) and obtain the ultimate explicit form of the component $\tilde{E}_y^j$ of the electromagnetic field of the single-mode TE polarized AGM:

$$\tilde{E}_y^j(x,z) = \frac{\beta_j^{\mathrm{TE}}(0)}{C_j^{\mathrm{TE}}(0)}\frac{E_y^j(x;z)}{\beta_j^{\mathrm{TE}}(z)}. \tag{25}$$

By means of the second equation of the system (10) we obtain

$$\tilde{H}_z^j(x,z) = -\frac{1}{ik_0\mu}\frac{d\tilde{E}_y^j}{dx}(x,z) \tag{26}$$

and using the third equation of the system (10) we obtain

$$\tilde{H}_x^j(x,z) = \frac{\beta_j^{\mathrm{TE}}(z)}{\mu}\tilde{E}_y^j(x,z). \tag{27}$$

In analogy with the above calculations, we solve the system (11) using the auxiliary problem

$$\left(\frac{d^2}{dx^2} + k_0^2 n^2(x,z)\right) H_y^j(x;z) = k_0^2\beta_j^2(z)H_y^j(x;z), \quad -\infty < x < \infty, \tag{28}$$

with boundary (asymptotic) conditions

$$H_y^j(x;z) \xrightarrow[x\to\pm\infty]{} 0. \tag{29}$$

The eigenfunctions are normalized by the condition

$$\langle H_y^j, H_y^k\rangle = \int_{-\infty}^{\infty} H_y^j(x;z)\bar{H}_y^k(x;z)dx = \delta_{jk}. \tag{30}$$

The single-mode solution for the TM-polarized AGM has the form

$$\tilde{H}_y^j(x,z) = \frac{\beta_j^{\mathrm{TM}}(0)}{C_j^{\mathrm{TM}}(0)}\frac{H_y^j(x;z)}{\beta_j^{\mathrm{TM}}(z)}, \tag{31}$$

$$\tilde{E}_z^j(x,z) = \frac{1}{ik_0\varepsilon}\frac{d\tilde{H}_y^j}{dx}(x,z) \tag{32}$$

$$\tilde{E}_x^j(x,z) = -\frac{\beta_j^{\mathrm{TM}}(z)}{\varepsilon}\tilde{H}_y^j(x,z). \tag{33}$$

7. Search for adiabatic guided mode phase using the Cauchy method

In the model of adiabatic guided modes, in addition to the coefficient functions $C_j^{\mathrm{TE}}(z)$ and $C_j^{\mathrm{TM}}(z)$, the explicit dependence on the horizontal coordinate is present in the exponential factor $\exp\{i\omega t - ik_0\varphi(y,z)\}$ in expression (4) [9]. The mode evolution in the horizontal direction, besides the dependences (24) for the coefficient functions in the case $\partial/\partial y \equiv 0$, is formed by the dependence of $d\varphi/dz$ on the obtained eigenvalues of problems (17). It is convenient to formulate the description of phase evolution law $\varphi(z)$ in terms of the algebraic model of adiabatic guided modes [9] for thin-film waveguides. A constructive representation of this dependence for thin-film waveguides consisting of homogeneous layers is obtained in Refs. [19]–[21]. Namely, the general solutions in the homogeneous layers have the form

$$\begin{pmatrix} E_y^s \\ H_z^s \\ H_y^s \\ E_z^s \end{pmatrix} = \begin{pmatrix} -i\mu_s \exp\{k_0\eta_s^j(z)x\}A_s^1 \\ \eta_s^j(z)\exp\{k_0\eta_s^j(z)x\}A_s^1 \\ i\varepsilon_s \exp\{k_0\eta_s^j(z)x\}A_s^2 \\ \eta_s^j(z)\exp\{k_0\eta_s^j(z)x\}A_s^2 \end{pmatrix}, \tag{34}$$

$$\begin{pmatrix} E_y^f \\ H_z^f \\ H_y^f \\ E_z^f \end{pmatrix} = \begin{pmatrix} i\mu_f \left(\exp\{-k_0\eta_f^j(z)x\}A_f^3 - \exp\{-k_0\eta_f^j(z)x\}A_f^1\right) \\ \eta_f^j(z)\left(\exp\{-k_0\eta_f^j(z)x\}A_f^3 + \exp\{-k_0\eta_f^j(z)x\}A_f^1\right) \\ i\varepsilon_f \left(\exp\{-k_0\eta_f^j(z)x\}A_f^2 - \exp\{-k_0\eta_f^j(z)x\}A_f^4\right) \\ \eta_f^j(z)\left(\exp\{-k_0\eta_f^j(z)x\}A_f^4 + \exp\{-k_0\eta_f^j(z)x\}A_f^2\right) \end{pmatrix}, \tag{35}$$

$$\begin{pmatrix} E_y^c \\ H_z^c \\ H_y^c \\ E_z^c \end{pmatrix} = \begin{pmatrix} i\mu_c \exp\{k_0\eta_c^j(z)(-x+h)\}A_c^3 \\ \eta_c^j(z)\exp\{k_0\eta_c^j(z)(-x+h)\}A_c^3 \\ -i\varepsilon_c \exp\{k_0\eta_c^j(z)(-x+h)\}A_c^4 \\ \eta_c^j(z)\exp\{k_0\eta_c^j(z)(-x+h)\}A_c^4 \end{pmatrix}, \tag{36}$$

where $\eta_\alpha^j(z) = \sqrt{\beta_j^2(z) - n_\alpha^2}$, $\alpha = s, f, c$.

The system of boundary Maxwell's equations (2) with the explicit form of solutions (34),(36) taken into account leads [22]–[24] to the systems of homogeneous linear algebraic equations

$$\hat{M}^{\mathrm{TE}}\left(A_s^1\ A_f^1\ A_f^3\ A_f^3\right)^T = (0\ 0\ 0\ 0)^T \tag{37}$$

for TE modes, where

$$\hat{M}^{\mathrm{TE}} = \begin{pmatrix} i\mu_c & i\mu_f \exp\{k_0\eta_f^j(z)h\} & -i\mu_f \exp\{-k_0\eta_f^j(z)h\} & 0 \\ 0 & i\mu_f & -i\mu_f & -i\mu_s \\ 0 & -\eta_f^j(z) & -\eta_f^j(z) & \eta_s^j(z) \\ \eta_c^j(z) & -\eta_f^j(z)\exp\{k_0\eta_f^j(z)h\} & -\eta_f^j(z)\exp\{-k_0\eta_f^j(z)h\} & 0 \end{pmatrix} \tag{38}$$

and

$$\hat{M}^{\mathrm{TM}} \left(A_s^2\ A_f^2\ A_f^4\ A_c^4\right)^T = (0\ 0\ 0\ 0)^T \tag{39}$$

for TM modes, where

$$\hat{M}^{\mathrm{TM}} = \begin{pmatrix} 0 & -i\varepsilon_f & i\varepsilon_f & i\varepsilon_s \\ i\varepsilon_s & -i\varepsilon_f \exp\{k_0\eta_f^j(z)h\} & i\varepsilon_f \exp\{-k_0\eta_f^j(z)h\} & 0 \\ \eta_c^j(z) & -\eta_f^j(z)\exp\{k_0\eta_f^j(z)h\} & -\eta_f^j(z)\exp\{-k_0\eta_f^j(z)h\} & 0 \\ 0 & -\eta_f^j(z) & -\eta_f^j(z) & \eta_s^j(z) \end{pmatrix}. \tag{40}$$

The systems of homogeneous linear algebraic equations (37) and (39) have nontrivial solutions under the conditions

$$\det \hat{M}^{\mathrm{TE}} \left(d\varphi^{\mathrm{TE}}/dz\right) = 0, \tag{41}$$

$$\det \hat{M}^{\mathrm{TM}} \left(d\varphi^{\mathrm{TM}}/dz\right) = 0 \tag{42}$$

In terms of ordinary differential equations for $d\varphi^{\mathrm{TE}}/dz$ and $d\varphi^{\mathrm{TM}}/dz$, equations (41) and (42) are written in the form

$$F\left(\frac{d\varphi}{dz}(z);\ h(z), \frac{dh}{dz}(z);\ \varepsilon_\alpha, \mu_\alpha\right) = 0, \tag{43}$$

where F is a transcendental function of $\dfrac{d\varphi}{dz}$, including expressions $\eta_\alpha^j(z) = \sqrt{\beta_j^2(z) - n_\alpha^2}$, $\alpha = s, f, c$ and $\exp\{\pm k_0\eta_\alpha^j(z)h(z)\}$, $\alpha = s, f, c$ i.e., radicals of $\left(\dfrac{d\varphi}{dz}\right)^2$ and exponential functions of these radicals.

Equations (41) and (42) of the form (43) are solved symbolic-numerically using the Cauchy method (see [25], [26]).

8. Basic equations of the adiabatic guided mode model

The problem of propagation of polarized electromagnetic radiation in regular waveguides was successfully solved both in closed [27]–[30] and in open waveguides [31]. In both cases the base method is the method of separation of variables, which reduces the initial problem to an auxiliary

spectral problem with a discrete spectrum in the case of closed waveguides and a mixed spectrum in the case of open ones.

Problems with dielectric and magnetic filling of waveguides, intricate in the cross section but regular in the longitudinal direction (along the radiation propagation axis) were reduced to considerably more complicated spectral problems, which have been subsequently also solved (see, e.g., [32]–[39]). These physical models gave rise to new mathematical problems and description of new electromagnetic phenomena, which are beyond the present discussion.

In contrast to this line of study, the mathematical modeling of waveguides irregular in the longitudinal direction began to develop from the middle of the 20th century. Here we should mention the pioneering works [40]–[42] in the English-language literature. However, the dominant contribution to the development of mathematically substantiated methods was made by Russian-language authors. Their works can be conventionally divided into congeneric, in fact, to the ideology of Kantorovich method [43] generalizing the Fourier method of separation of variables (e.g., [44]–[54]) and those using the ideology of asymptotic approach (see, e.g., [55]). It is worth noting that the studies in the first line have acquired rigorous mathematical completeness in the works by Sveshnikov and his disciples (see, e.g., [56]–[66]).

The present paper combines the Sveshnikov ideas in both approaches and bases on technique of exploiting the Kantorovich method, implicitly used by Fedoryuk.

Acknowledgments

The publication has been prepared with the support of the Russian Foundation for Basic Research (RFBR) according to the research project No 19-01-00645. The author is grateful to his colleague Dmitry Divakov for providing the results of computer calculations.

References

[1] L. A. Sevastianov and A. A. Egorov, "Theoretical analysis of the waveguide propagation of electromagnetic waves in dielectric smoothly-irregular integrated structures," *Optics and Spectroscopy*, vol. 105, no. 4, pp. 576–584, 2008. DOI: 10.1134/S0030400X08100123.

[2] A. A. Egorov and L. A. Sevastianov, "Structure of modes of a smoothly irregular integrated optical four-layer three-dimensional waveguide," *Quantum Electronics*, vol. 39, no. 6, pp. 566–574, 2009. DOI: 10.1070/QE2009v039n06ABEH013966.

[3] A. A. Egorov, K. P. Lovetskiy, A. L. Sevastianov, and L. A. Sevastianov, "Simulation of guided modes (eigenmodes) and synthesis of a thin-film generalised waveguide Luneburg lens in the zero-order vector approximation," *Quantum Electronics*, vol. 40, no. 9, pp. 830–836, 2010. DOI: 10.1070/QE2010V040N09ABEH014332.

[4] A. A. Egorov *et al.*, "Adiabatic modes of smoothly irregular optical waveguide: zero approximation of vector theory [Adiabaticheskie mody plavno-neregulyarnogo opticheskogo volnovoda: nulevoe priblizhenie vektornoj teorii]," Russian, *Matem. modelirovaniye*, vol. 22, no. 8, pp. 42–54, 2010, [in Russian].

[5] A. A. Egorov, A. L. Sevast'yanov, and L. A. Sevast'yanov, "Stable computer modeling of thin-film generalized waveguide Luneburg lens," *Quantum Electronics*, vol. 44, no. 2, pp. 167–173, 2014. DOI: 10.1070/QE2014v044n02ABEH015303.

[6] A. A. Egorov, K. P. Lovetsky, A. L. Sevastyanov, and L. A. Sevastyanov, "Luneberg Thin-Film Waveguide Lens: From Problem Statement to Solution. Theory and mathematical modeling of adiabatic modes [Tonkoplenochnaya volnovodnaya linza Lyuneberga: ot postanovki problemy do ee resheniya. Teoriya i matematicheskoe modelirovanie adiabaticheskih mod]," Russian, in *Trudy RNTORES im. A.S. Popova. Vyp. 5. The 5th International Conference "Acousto-Optical and Radar Measurement and Information Processing Methods" (ARMIMP-2012), Moscow-Suzdal*, [in Russian], 2012, pp. 186–190.

[7] A. A. Egorov, A. L. Sevastyanov, E. A. Ayryan, and L. A. Sevastyanov, "Stable computer modeling of thin-film generalized waveguide Luneburg lens [Ustojchivoe komp'yuternoe modelirovanie tonkoplenochnoj obobshchennoj volnovodnoj linzy Lyuneberga]," Russian, *Matem. modelirovaniye*, vol. 26, no. 11, pp. 37–44, 2014, [in Russian].

[8] E. Ayryan, G. Dashitsyrenov, E. Laneev, K. Lovetskiy, L. Sevastianov, and A. Sevastianov, "Mathematical synthesis of the thickness profile of the waveguide Lüneburg lens using the adiabatic waveguide modes method," in *Saratov Fall Meeting 2016: Laser Physics and Photonics XVII; and Computational Biophysics and Analysis of Biomedical Data III*, V. L. Derbov, D. E. Postnov, V. L. Derbov, and D. E. Postnov, Eds., International Society for Optics and Photonics, vol. 10337, SPIE, 2017, pp. 134–145. DOI: 10.1117/12.2267920.

[9] A. L. Sevastyanov, "Asymptotic method for constructing a model of adiabatic guided modes of smoothly irregular integrated optical waveguides," *Discrete and Continuous Models and Applied Computational Science*, vol. 28, no. 3, pp. 252–273, 2020. DOI: 10.22363/2658-4670-2020-28-3-252-273.

[10] A. S. Il'inskii, V. V. Kravtsov, and A. G. Sveshnikov, *Mathematical Models of Electrodynamics [Matematicheskie modeli elektrodinamiki]*, Russian. Moscow: Vyssh. Shkola, 1991, [in Russian].

[11] L. A. Sevastyanov, A. A. Egorov, and A. L. Sevastyanov, "Method of adiabatic modes in studying problems of smoothly irregular open waveguide structures," *Physics of Atomic Nuclei*, vol. 776, no. 2, pp. 224–239, 2013. DOI: 10.1134/S1063778813010134.

[12] L. V. Kantorovich and V. I. Krylov, *Approximate Methods of Higher Analysis*. New York: Wiley, 1964.

[13] M. V. Fedoryuk, "A justification of the method of transverse sections for an acoustic wave guide with nonhomogeneous content," *Mathematical Physics*, vol. 13, no. 1, pp. 162–173, 1973. DOI: 10.1016/0041-5553(74)90012-3.

[14] B. Z. Katsenelenbaum, *Theory of Irregular Waveguides with Slowly Varying Parameters [Teoriya neregulyarnyh volnovodov s medlenno menyayushchimisya parametrami]*, Russian. Moscow: Akad. Nauk SSSR, 1961, [in Russian].

[15] V. V. Shevchenko, *Continuous Transitions in Open Waveguides [Plavnye perekhody v otkrytyh volnovodah]*, Russian. Moscow: Nauka, 1969, [in Russian].

[16] M. J. Adams, *An Introduction to Optical Waveguides*. New York: Wiley, 1981.

[17] T. Tamir, "Guided-wave optoelectronics," in *Integrated Optics*, T. Tamir, Ed., Berlin: Springer-Verlag, 1990.

[18] A. W. Snyder and J. D. Love, *Optical Waveguide Theory*. New York: Chapman and Hall, 1983.

[19] D. V. Divakov and A. L. Sevastianov, "The Implementation of the Symbolic-Numerical Method for Finding the Adiabatic Waveguide Modes of Integrated Optical Waveguides in CAS Maple," in *Computer Algebra in Scientific Computing*, M. England, W. Koepf, T. M. Sadykov, W. M. Seiler, and E. V. Vorozhtsov, Eds., Cham: Springer International Publishing, 2019, pp. 107–121. DOI: 10.1007/978-3-030-26831-2_8.

[20] D. V. Divakov, A. A. Tiutiunnik, and A. L. Sevastianov, "Symbolic-Numeric Study of Geometric Properties of Adiabatic Waveguide Modes," in *Computer Algebra in Scientific Computing*, F. Boulier, M. England, T. M. Sadykov, and E. V. Vorozhtsov, Eds., Cham: Springer International Publishing, 2020, pp. 228–244. DOI: 10.1007/978-3-030-60026-6_13.

[21] D. V. Divakov and A. A. Tiutiunnik, "Symbolic study of eigenvectors for constructing a general solution to a system of ODEs with a symbolic matrix of coefficients," *Programmirovaniye*, no. 2, pp. 3–16, 2020.

[22] L. A. Sevastyanov, A. L. Sevastyanov, and A. A. Tyutyunnik, "Analytical Calculations in Maple to Implement the Method of Adiabatic Modes for Modelling Smoothly Irregular Integrated Optical Waveguide Structures," in *Computer Algebra in Scientific Computing*, V. P. Gerdt, W. Koepf, W. M. Seiler, and E. V. Vorozhtsov, Eds., Cham: Springer International Publishing, 2014, pp. 419–431. DOI: 10.1007/978-3-319-10515-4_30.

[23] A. A. Egorov, K. P. Lovetskii, A. L. Sevastianov, and L. A. Sevastianov, *Integrated optics: theory and computer modeling [Integral'naya optika: teoriya i komp'yuternoe modelirovanie]*, Russian. Moscow: PFUR Publishing house, 2015, [in Russian].

[24] A. A. Egorov, K. P. Lovetskii, A. L. Sevastyanov, and L. A. Sevastyanov, "Model of a smoothly irregular multilayer integrated-optical waveguide in the zero vector approximation: theory and numerical analysis [Model' mnogoslojnogo plavno-neregulyarnogo integral'no-opticheskogo volnovoda v nulevom vektornom priblizhenii: teoriya i chislennyj analiz]," Russian, *Zhurnal radioelektroniki*, no. 3, 2019, [in Russian]. DOI: 10.30898/1684-1719.2019.3.11.

[25] M. D. Malykh, "On integration of the first order differential equations in a finite terms," *Journal of Physics: Conference Series*, vol. 788, p. 012 026, Jan. 2017. DOI: 10.1088/1742-6596/788/1/012026.

[26] A. D. Polyanin and V. E. Nazaikinskii, *Handbook of linear partial differentialequations for engineers and scientists, 2nd Edition*. BocaRaton, London: CRC Press, 2016.

[27] A. A. Samarskiy and A. N. Tikhonov, "Excitation of radio waveguides. I [O vozbuzhdenii radiovolnovodov. I]," Russian, *Zhurnal tekhnicheskoy fiziki*, vol. 17, no. 11, pp. 1283–1296, 1947, [in Russian].

[28] A. A. Samarskiy and A. N. Tikhonov, "Excitation of radio waveguides. II [O vozbuzhdenii radiovolnovodov. II]," Russian, *Zhurnal tekhnicheskoy fiziki*, vol. 17, no. 12, pp. 1431–1440, 1947, [in Russian].

[29] A. A. Samarskiy and A. N. Tikhonov, "Excitation of radio waveguides. III [O vozbuzhdenii radiovolnovodov. III]," Russian, *Zhurnal tekhnicheskoy fiziki*, vol. 18, no. 7, pp. 971–983, 1948, [in Russian].

[30] A. A. Samarskii and A. N. Tikhonov, "Representation of the field in a waveguide as the sum of the TE and TM fields [O predstavlenii polya v volnovode v vide summy polej TE i TM]," Russian, *Zhurnal tekhnicheskoy fiziki*, vol. 18, no. 7, pp. 959–970, 1948, [in Russian].

[31] V. V. Shevchenko, "Spectral decomposition in eigen- and associated functions of a nonselfadjoint problem of Sturm–Liouville type on the entire axis [O spektral'nom razlozhenii po sobstvennym i prisoedinennym funkciyam odnoj nesamosopryazhennoj zadachi tipa Shturma–Liuvillya na vsej osi]," Russian, *Differ. Uravn.*, vol. 15, no. 11, pp. 2004–2020, 1979, [in Russian].

[32] E. M. Karchevskii, "Determination of the propagation constants of dielectric-waveguide eigenmodes by methods of potential theory," *Computational Mathematics and Mathematical Physics*, vol. 38, no. 1, pp. 132–136, 1998.

[33] R. Z. Dautov and E. M. Karchevskii, "On a spectral problem of the theory of dielectric waveguides," *Computational Mathematics and Mathematical Physics*, vol. 39, no. 8, pp. 1293–1299, 1999.

[34] E. M. Karchevskii, "Analysis of the eigenmode spectra of dielectric waveguides," *Computational Mathematics and Mathematical Physics*, vol. 39, no. 9, pp. 1493–1498, 1999.

[35] E. M. Karchevskii, "Investigation of a numerical method for solving a spectral problem in the theory of dielectric waveguides," *Russian Mathematics (Izvestiya VUZ. Matematika)*, vol. 43, no. 1, pp. 8–15, 1999.

[36] R. Z. Dautov and E. M. Karchevskii, "Existence and properties of solutions to the spectral problem of the dielectric waveguide theory," *Computational Mathematics and Mathematical Physics*, vol. 40, no. 8, pp. 1200–1213, 2000.

[37] R. Z. Dautov and E. M. Karchevskii, "Solution of the vector problem of the natural waves of cylindrical dielectric waveguides based on a nonlocal boundary condition," *Computational Mathematics and Mathematical Physics*, vol. 42, no. 7, pp. 1012–1027, 2002.

[38] E. M. Karchevskii and S. I. Solov'ev, "Existence of eigenvalues of a spectral problem in the theory of dielectric waveguides," *Russian Mathematics (Izvestiya VUZ. Matematika)*, vol. 47, no. 3, pp. 75–77, 2003.

[39] E. M. Karchevskii, A. I. Nosich, and G. W. Hanson, "Mathematical analysis of the generalized natural modes of an inhomogeneous optical fiber," *SIAM Journal on Applied Mathematics*, vol. 65, no. 6, pp. 2033–2048, 2005. DOI: 10.1137/040604376.

[40] A. F. Stevenson, "General Theory of Electromagnetic Horns," *Journal of Applied Physics*, vol. 22, no. 12, p. 1447, 1951. DOI: 10.1063/1.1699891.

[41] S. A. Schelkunoff, "Conversion of Maxwell's equations into generalized Telegraphist's equations," *The Bell System Technical Journal*, vol. 34, no. 5, pp. 995–1043, 1955. DOI: 10.1002/j.1538-7305.1955.tb03787.x.

[42] A. Yariv, "Coupled-mode theory for guided-wave optics," *IEEE Journal of Quantum Electronics*, vol. 9, pp. 919–933, 1973. DOI: 10.1109/JQE.1973.1077767.

[43] L. V. Kantorovich, "A direct method for the approximate solution of the problem of the minimum of a double integral [Odin pryamoj metod priblizhennogo resheniya zadachi o minimume dvojnogo integrala]," Russian, *Izvestiya Akademii nauk SSSR. VII seriya. Otdeleniye matematicheskikh i yestestvennykh nauk*, no. 5, pp. 647–652, 1933, [in Russian].

[44] B. Z. Katsenelenbaum, "Irregular waveguides with slowly varying parameters [Neregulyarnye volnovody s medlenno menyayushchimisya parametrami]," Russian, *Doklady Akademii Nauk SSSR*, vol. 102, no. 4, p. 711, 1955, [in Russian].

[45] B. Z. Katsenelenbaum, "On the general theory of irregular waveguides [K obshchej teorii neregulyarnyh volnovodov]," Russian, *Doklady Akademii Nauk SSSR*, vol. 116, no. 2, pp. 203–206, 1957, [in Russian].

[46] A. G. Sveshnikov, "An approximate method for calculating a weakly irregular waveguide [Priblizhennyj metod rascheta slabo neregulyarnogo volnovoda]," Russian, *Doklady Akademii Nauk SSSR*, vol. 110, no. 2, pp. 197–199, 1956, [in Russian].

[47] G. Y. Lyubarsky and A. Y. Povzner, "On the theory of wave propagation in irregular waveguides [K teorii rasprostraneniya voln v neregulyarnyh volnovodah]," Russian, *Zhurnal tekhnicheskoy fiziki*, no. 29, pp. 170–179, 1959, [in Russian].

[48] N. E. Maltsev, "Some modifications of the method of cross sections [Nekotorye modifikacii metoda poperechnyh sechenij]," Russian, *Akusticheskii zhurnal*, no. 16, pp. 102–109, 1970, [in Russian].

[49] B. Z. Katsenelenbaum, "Curved waveguides of constant cross-section [Izognutye volnovody postoyannogo secheniya]," Russian, *Radiotekhnika i elektronika*, no. 2, pp. 171–185, 1956, [in Russian].

[50] B. Z. Katsenelenbaum, "Symmetric dielectric transition in a circular waveguide for the H_{01} wave [Simmetrichnyj dielektricheskij perekhod v volnovode kruglogo secheniya dlya volny H_{01}]," Russian, *Radiotekhnika i elektronika*, no. 3, p. 339, 1956, [in Russian].

[51] B. Z. Katsenelenbaum, "Long symmetric waveguide transition for the H_{01} wave [Dlinnyj simmetrichnyj volnovodnyj perekhod dlya volny H_{01}]," Russian, *Radiotekhnika i elektronika*, no. 5, pp. 531–546, 1957, [in Russian].

[52] A. G. Sveshnikov, "On the propagation of radio waves in weakly curved waveguides [O rasprostranenii radiovoln v slaboizognutyh volnovodah]," Russian, *Radiotekhnika i elektronika*, vol. 1, no. 9, p. 1222, 1956, [in Russian].

[53] A. G. Sveshnikov, "Waves in curved pipes [Volny v izognutyh trubah]," Russian, *Radiotekhnika i elektronika*, vol. 3, no. 5, p. 641, 1958, [in Russian].

[54] A. G. Sveshnikov, "Irregular waveguides [Neregulyarnye volnovody]," Russian, *Izv. Vuzov. Radiofizika*, vol. 2, no. 5, p. 720, 1959, [in Russian].

[55] A. G. Sveshnikov, "An approximate method for calculating a weakly irregular waveguide [Priblizhennyj metod rascheta slabo neregulyarnogo volnovoda]," Russian, *Doklady Akademii Nauk SSSR*, vol. 80, no. 3, pp. 345–347, 1956, [in Russian].

[56] A. G. Sveshnikov, "On the proof of a method of calculation for irregular waveguides," *USSR Computational Mathematics and Mathematical Physics*, vol. 3, no. 1, pp. 219–232, 1963.

[57] A. G. Sveshnikov, "A substantiation of a method for computing the propagation of electromagnetic oscillations in irregular waveguides," *USSR Computational Mathematics and Mathematical Physics*, vol. 3, no. 2, pp. 413–429, 1963. DOI: 10.1016/0041-5553(63)90027-2.

[58] A. G. Sveshnikov, "On the bending of waveguides," *USSR Computational Mathematics and Mathematical Physics*, vol. 1, no. 3, pp. 882–888, 1962.

[59] A. G. Sveshnikov and A. S. Il'inskii, "Calculation of waveguide transition of composite form," *USSR Computational Mathematics and Mathematical Physics*, vol. 3, no. 3, pp. 635–649, 1963.

[60] A. S. Il'inskii and A. G. Sveshnikov, "Methods for investigating irregular waveguides," *USSR Computational Mathematics and Mathematical Physics*, vol. 8, no. 2, pp. 167–180, 1968.

[61] A. G. Sveshnikov, "The incomplete Galerkin method [Nepolnyj metod Galerkina]," Russian, *Dokl. Akad. Nauk SSSR*, vol. 236, no. 5, pp. 1076–1079, 1977, [in Russian].

[62] A. N. Bogolyubov and A. G. Sveshnikov, "Application of an iteration method to the investigation of plane waveguides with inhomogeneous filling," *USSR Computational Mathematics and Mathematical Physics*, vol. 14, no. 4, pp. 125–133, 1974.

[63] A. N. Bogolyubov and A. G. Sveshnikov, "Justification of a finite-difference method for analyzing optical waveguides," *USSR Computational Mathematics and Mathematical Physics*, vol. 19, no. 6, pp. 139–150, 1979.

[64] A. N. Bogolyubov, A. L. Delitsyn, and A. G. Sveshnikov, "On the problem of excitation of a waveguide filled with an inhomogeneous medium," *Computations mathematics and mathematical physics*, vol. 39, no. 11, pp. 1794–1813, 1999.

[65] A. N. Bogolyubov and M. D. Malykh, "Remark on the Radiation Conditions for an Irregular Waveguide," *Computations mathematics and mathematical physics*, vol. 43, no. 4, pp. 560–563, 2003.

[66] A. N. Bogolyubov and M. D. Malykh, "Theory of Perturbations of Spectral Characteristics of Waveguide Systems," *Computations mathematics and mathematical physics*, vol. 43, no. 7, pp. 1049–1061, 2003.

For citation:

A. L. Sevastianov, Single-mode propagation of adiabatic guided modes in smoothly irregular integral optical waveguides, Discrete and Continuous Models and Applied Computational Science 28 (4) (2020) 361–377. DOI: 10.22363/2658-4670-2020-28-4-361-377.

Information about the authors:

Sevastianov, Anton L. — Candidate of Physical and Mathematical Sciences, assistant professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: `sevastianov-al@rudn.ru`, ORCID: https://orcid.org/0000−0002−0280−485X, Scopus Author ID: 50462435500)

УДК 519.61:519.62:621.372.82
DOI: 10.22363/2658-4670-2020-28-4-361-377

Одномодовый режим распространения адиабатических волноводных мод плавно-нерегулярных интегрально-оптических волноводов

А. Л. Севастьянов

Российский университет дружбы народов
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия

В работе представлено исследование волноводного распространения поляризованного электромагнитного излучения в тонкоплёночном интегрально-оптическом волноводе. Для описания этого распространения используется адиабатическое приближение решений уравнений Максвелла. Построение редуцированной модели для адиабатических волноводных мод, сохраняющей все свойства соответствующих приближённых решений системы уравнений Максвелла, было проведено автором в предыдущей публикации в DCM&ACS, 2020, № 3. В настоящей работе исследование проведено для частного случая, когда геометрия волновода и электромагнитное поле инвариантны в поперечном направлении. В этих условиях существуют раздельные нетривиальные TE- и TM-поляризованные решения указанной редуцированной модели. В работе описываются параметрически зависящие от продольных координат решения задач на собственные значения и собственные функции — адиабатические волноводные TE- и TM-поляризации. В работе приводится постановка задачи отыскания решений модели адиабатических волноводных мод, описывающих стационарное распространение электромагнитного излучения. Представлены решения для одномодового распространения TE- и TM-поляризованных адиабатических волноводных волн.

Ключевые слова: волноводное распространение поляризованного света, интегральный оптический волновод, адиабатическое приближение, собственные значения и собственные функции, метод Канторовича, одномодовый режим

UDC 535:535.3:681.7
DOI: 10.22363/2658-4670-2020-28-4-378-397

Solving the inverse problem for determining the optical characteristics of materials

Konstantin P. Lovetski[1], Andrey A. Zhukov[2], Michael V. Paukshto[3], Leonid A. Sevastianov[1], Anastasiia A. Tiutiunnik[1]

[1] *Peoples' Friendship University of Russia (RUDN University)*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation
[2] *ITL Consulting*
16, Olkhovskaya St., bldg. 5, Moscow, 105066, Russian Federation
[3] *Fibralign Corporation*
32930, Alvarado-Niles Rd., Suite 350, Union City, CA 94587, USA

(received: November 10, 2020; accepted: November 12, 2020)

The paper describes a methodology for determining the optical and physical properties of anisotropic thin film materials. This approach allows in the future designing multilayer thin-film coatings with specified properties. An inverse problem of determining the permittivity tensor and the thickness of a thin film deposited on a glass substrate is formulated. Preliminary information on the belonging of a thin-film coating to a certain class can significantly reduce the computing time and increase the accuracy of determining the permittivity tensor over the entire investigated range of wavelengths and film thickness at the point of reflection and transmission measurement

Depending on the goals, it is possible to formulate and, therefore, solve various inverse problems:

– determination of the permittivity tensor and specification of the thickness of a thick (up to 1 cm) substrate, often isotropic;

– determination of the permittivity tensor of a thin isotropic or anisotropic film deposited on a substrate with known optical properties.

The complexity of solving each of the problems is very different and each problem requires its own specific set of measured input data. The ultimate results of solving the inverse problem are verified by comparing the calculated transmission and reflection with those measured for arbitrary angles of incidence and reflection.

Key words and phrases: transmittance, reflectance, refractive indices determination, thin films, multilayers, optical coatings, optical properties

1. Introduction

The efficiency of production of existing devices for solid-state micro- and nanoelectronics and successful creation of new ones largely depend on the level of development of the technology for manufacturing layers of various

materials with a thickness from several nanometers to tens of micrometers [1]. The design and manufacture of multilayer structures with desired properties from dielectric and/or metal films requires an accurate knowledge of the optical parameters of each layer [2], [3].

Methods for evaluating the electrophysical parameters of dielectric and semiconductor thin-film materials [4] based on regularized methods [5] for solving inverse problems allow accurate determination of the electrophysical parameters of thin-film semiconductor materials [6]–[8]. It becomes possible to create multilayer structures with predetermined properties [9].

The advantage of non-contact methods, which include spectrophotometric and polarimetric methods, is the possibility to carry out measurements without destroying the material and without changing its properties. When using these methods, the interaction of electromagnetic waves in the optical range with the sample material is considered and the intensities of the transmitted and reflected waves are measured. The obtained intensities can be then used to calculate both optical and geometric parameters of the samples [10]–[12].

The advantage of spectrophotometric measurements is the possibility to determine several parameters using one measuring device and one sample [13]. To determine the thickness, permittivity, and electrical conductivity of nanometer films in layered structures, one can use the results of measurements of the reflection and transmission spectra of the optical radiation interacting with them, provided that the mathematical model of their interaction is known [14], [15].

Finding the electrophysical parameters of layered structures from the reflection and transmission spectra of electromagnetic waves is associated with the need to solve inverse ill-posed problems of electrodynamics.

The developed program "Multilayer" serves both for modeling the transmission of light through multilayer thin-film layered media [16]–[18] and for determining the dielectric (permittivity tensor of anisotropic films) and geometric (film thickness) parameters of various thin-film coatings. The program was created based on many years of experience of collaboration with organizations engaged in the design of thin-film coatings [11] used in the production of liquid crystal displays.

2. Formulation of extended inverse problem

Methods described in [19] for description of the transmission of an electromagnetic wave through an optical system are also used in solving the inverse problem for determining the optical characteristics of materials. Let us consider formulation of the inverse problem to determine optical parameters of thin film coating.

Within the framework of the inverse problem, it is required to determine, using data on the transmission $T_m(\lambda_i)$ and reflection $R_m(\lambda_i)$ for various angles of incidence and various polarizations ($m \in [1, \dots, n]$, here n is the number of different spectrophotometric measurements for each wavelength $\lambda_i, i \in [1, \dots, p]$ from desired range $[\lambda_{\text{beg}}, \lambda_{\text{end}}]$), the elements of the permittivity tensor $\varepsilon(\lambda)$ of a homogeneous material in a preset wavelength range $[\lambda_{\text{beg}}, \lambda_{\text{end}}]$.

Mathematically, this problem reduces to minimization of the variance functional with respect to the unknown parameter $\varepsilon(\lambda_i)$.

$$F(\varepsilon(\lambda)) = \\ = \int\limits_{\lambda_j=\lambda_{\text{beg}}}^{\lambda_{\text{end}}} \left(\sum_{i=1}^{n} \left(T_i(\lambda) - T_i(\varepsilon_j(\lambda), \lambda)\right)^2 + \sum_{i=1}^{n} \left(R_i(\lambda) - R_i(\varepsilon_j(\lambda), \lambda)\right)^2 \right) d\lambda. \quad (1)$$

As was noted above, an additional unknown parameter can be the layer thickness d, which has to be especially thoroughly determined for thin films. In such cases, the target functional is written as $F(\varepsilon, d)$.

An important feature of the problem under consideration is that the optical parameters have to be found in a continuous wavelength interval $[\lambda_{\text{beg}}, \lambda_{\text{end}}]$, rather than at separate points in this interval. Typically, the interval of interest covers the visible spectral range from 400 to 800 nm. The search for parameters defined over a given spectral range implies the requirement that the spectral dependences of these parameters would be smooth functions, which influences the structure of solutions.

Solving the task of minimization [20] of the functional (1) frequently involves procedures requiring considerable computational time. The time consumption can be significantly reduced by taking into account certain special features of the system (isotropic vs. anisotropic materials, thick vs. thin layers), that is, by selecting a proper model of light propagation [21]–[23]. Problems related to the classification of materials are considered in the next section.

Note. From mathematical standpoint, the inverse problem of restoring the parameters of a differential operator belongs to the class of so-called ill-posed problems [5], which implies that small variations in the initial data may lead to large changes in the coefficients that have to be calculated. Such behavior of the solution is called unstable. The problem of reconstruction of the elements of the permittivity tensor $\varepsilon(\lambda)$ is the typical example of an ill-posed problem. The problem of correctness poses additional requirements both to the accuracy of initial data and to the stability of solution algorithms. In order to effectively solve an ill-posed problem, it is also highly desirable to take into account all the information known a priori about the system. Finally, it should be noted that poor (uncertain) initial data on the transmission $T_m(\lambda_i)$ and reflection $R_m(\lambda_i)$ of light in the system (for example, in the case of a significant level of light scattering) may not allow eliminating all the difficulties related to incorrect formulation of the problem.

3. Classification of media

In the adopted approach to formulation of the inverse problem, the properties of materials determined by the permittivity tensor are considered as most important [7], [24], [25]. For the Berreman matrix method [26], special features of a material (isotropic vs. anisotropic) are not of principal significance in solving the direct problem. However, the knowledge about such features may significantly simplify solution of the inverse problem by reducing the number of unknown parameters and by making possible the use of

simplified computational schemes [27]. In connection with this, we present a classification of the materials of interest [25].

The classification is closely related to the relative values and orientations of the principal axes of the refractive index ellipsoid (figure 1).

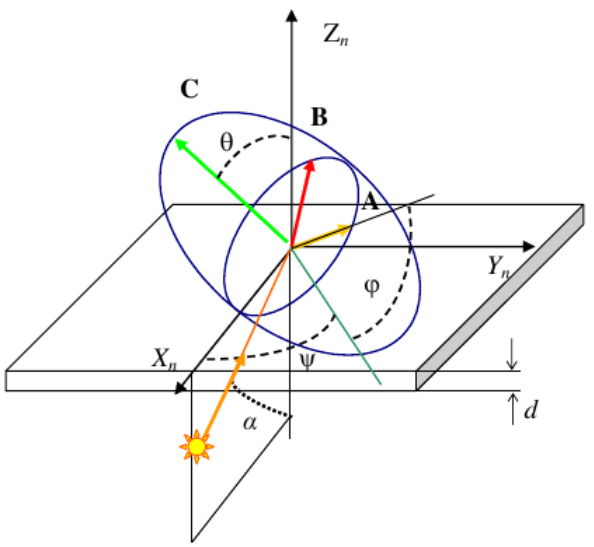


Figure 1. A general orientation of a principal axes of the permittivity ellipsoid. θ nutation angle, ψ precession angle, ϕ spin angle

Classification with respect to the principal values of the permittivity tensor:

— Biaxial anisotropic material. In this case, all three principal axes of the refractive index ellipsoid are different, for example, $\varepsilon_A < \varepsilon_B < \varepsilon_C$. Such materials have two optical axes.

— Uniaxial anisotropic material. Two of the three optical axes are equal to each other, for example, $\varepsilon_A = \varepsilon_B \neq \varepsilon_C$. Such materials have a single optical axis coinciding with C axis.

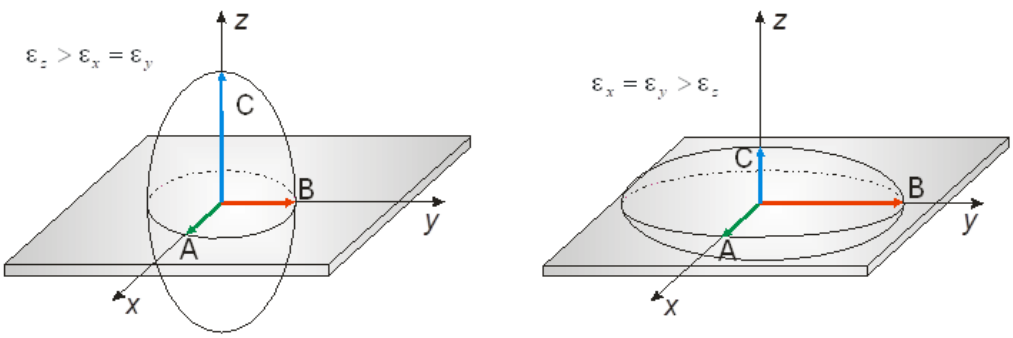
— Isotropic material. In this case, all three principal axes of the refractive index ellipsoid are equal to each other: $\varepsilon_A = \varepsilon_B = \varepsilon_C$.

In uniaxial materials, directions corresponding to equal principal axes of the refractive index ellipsoid (or equal values of the refractive indices) are called ordinary ($\varepsilon_O = \varepsilon_A = \varepsilon_B$ or n_o), while the remaining direction is called extraordinary ($\varepsilon_e = \varepsilon_C$ or n_e). It should be recalled that $\varepsilon_\sigma = n_\sigma^2$, where $\varepsilon_\sigma, n_\sigma$ are complex quantities and n_σ is the refractive index. Such materials are frequently characterized by the value of birefringence (or double refraction), which is defined as the difference $\Delta n = n_e - n_o$.

For $n_e > n_o$ we deal with a positive birefringence, while $n_e < n_o$ corresponds to a negative birefringence. Now we will consider the commonly accepted classification of uniaxial anisotropic materials with respect to the ratio of ε_O and ε_e and the orientation of a special axis. Here, the main parameter determining the type of a material is the angle of nutation θ.

Classification of uniaxial materials with respect to orientation of the principal axes of the refractive index ellipsoid:

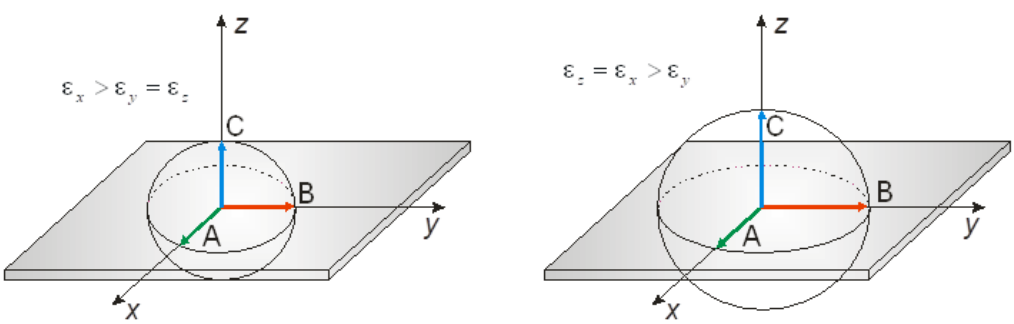
— o-Plate:
This case corresponds to the general arrangement of a uniaxial refractive index ellipsoid with an angle of nutation within $0 < \theta < \pi/2$, figure 1.
In cases, when the principal axes of the ellipsoid coincide with the axes of the laboratory frame, the notation $\varepsilon_A, \varepsilon_B, \varepsilon_C$ will be changed to $\varepsilon_x, \varepsilon_y, \varepsilon_z$.
— C-Plate:



(a) C-plate positive (b) C-plate negative

Figure 2. [$\theta = 0$; $\psi = 0$; ψ (the angle of rotation measured from OX axis) has arbitrary value

— A-Plate:



(a) Positive A-plate (b) Positive A-plate

Figure 3. A-plate (positive): $\theta = \pi/2$; $\psi = 0$; $\varphi = 0$, A-plate (negative): $\theta = \pi/2$; $\psi = \pi$; $\varphi = 0$

The above classification of uniaxial materials, depending on the orientation of the principal axes of the refractive index ellipsoid of the samples, allows the calculations to be limited to a small number of options for choosing calculation schemes, provided that there is a preliminary information about the belonging of the material under study to one of the classes.

4. Classification of inverse problems

In solving the inverse problem, the material should be classified primarily with respect to two factors: anisotropy and thickness. It should be also taken into account that thin materials (with a thickness of several microns and below) cannot be considered without a substrate bearing this layer. In such cases, calculations of the properties of a thin layer will involve the parameters (refractive index and thickness) of the substrate.

In accordance with a scheme presented in figure 4, the inverse problems are classified as follows.

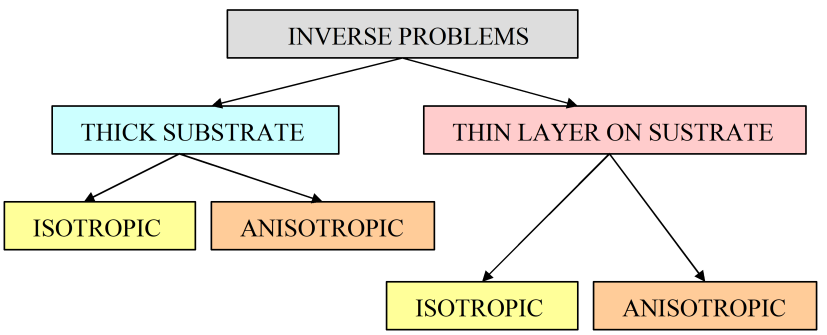


Figure 4. Classification of inverse problems

Problems of determination of refractive indices

Calculation of optical parameters for a thick layer – usually substrate (one-layer model).

— Calculation of the optical parameters (refractive indices) of an isotropic substrate.
— Calculation of the effective optical parameters of an anisotropic substrate for a given direction in the plane.
— Calculation of the optical parameters (permittivity tensor) of an anisotropic substrate.

Calculation of optical parameters for a thin film on thick substrate (two-layer model).

— Calculation of the optical parameters of a thin isotropic film on isotropic substrate with known parameters.
— Calculation of the effective optical parameters of thin anisotropic film on isotropic substrate with known parameters, for a given direction in the plane.
— Calculation of the optical parameters (permittivity tensor) of a thin anisotropic film on isotropic substrate with known parameters.

The problem of calculation of the effective refractive index of an anisotropic material in a given direction on the plane is solved using the same method as that used for determining the parameters of an isotropic material. However, calculations of the effective refractive index in a certain direction, while being of independent interest, may also perform an auxiliary role. On the one

hand, the effective refractive indices calculated in two mutually perpendicular directions (par/per) quite well characterize the given material for the normal incidence of light.

On the other hand, calculation of the principal axes of the refractive index ellipsoid for an anisotropic material in the case of small Euler angles (e.g., for a material with well oriented structure, where deviations of the principal axes from axes of the laboratory coordinate system are small) can start with calculation of the components lying in the sample plane. These par/per components can be used in subsequent calculations as very good initial approximation for the complete calculation of the permittivity tensor.

It should be borne in mind that the inverse problems of various types require using different sets of initial data, which will be considered below.

Note. In present version of the program, calculations of the parameters of a thin layer on substrate can be performed only for an isotropic substrate.

5. Classification of solution structures

Various methods of calculation of the propagation of light [22], [28] described previously (matrix and classical methods) are based on the relations valid for a given wavelength. In this context, calculation of the spectrum of the refractive index can be performed either point wise (in a given wavelength interval) or using a certain preset parametric dependence (figure 5).

Both point wise and parametric methods of calculation have their own advantages and drawbacks.

Pointwise method

Advantages. Provides effective solution of the inverse problem in the case of calculations of the optical parameters of a single thick layer. The solution is sufficiently stable. The level of oscillations in the obtained solution (noise) corresponds to the noise level in the initial data. Requires relatively small computational time.

Drawbacks. Gives highly unstable solutions in the case of calculations of the optical parameters of a thin layer in a two-layer structure (strongly ill-posed problem). Satisfactory solution can be obtained using various additional (physically justified) restrictions and regularization parameters.

Parametric structure of solutions

Advantages. Provides smooth solutions consistent with the physical meaning of the problem. The method is sufficiently universal, especially when the Kramers-Kronig relations are used [29].

Drawbacks. Computational time increases as compared to that required for the point wise calculations. Sometimes it is difficult to select a proper parametric structure of the solution for materials transparent in the entire range $[\lambda_{\text{beg}}, \lambda_{\text{end}}]$.

It should be noted that the requirement of smoothness is especially important for the real part of the complex permittivity, that is, for the refractive index (n). The spectral dependence of the absorption coefficient (k) is usually smooth even for a pointwise solution.

The main, effective method of solution of the inverse problem consists in obtaining a parametric solution using the Kramers–Kronig relations [29]. In this method, an important step is related to simulation of the absorption peaks.

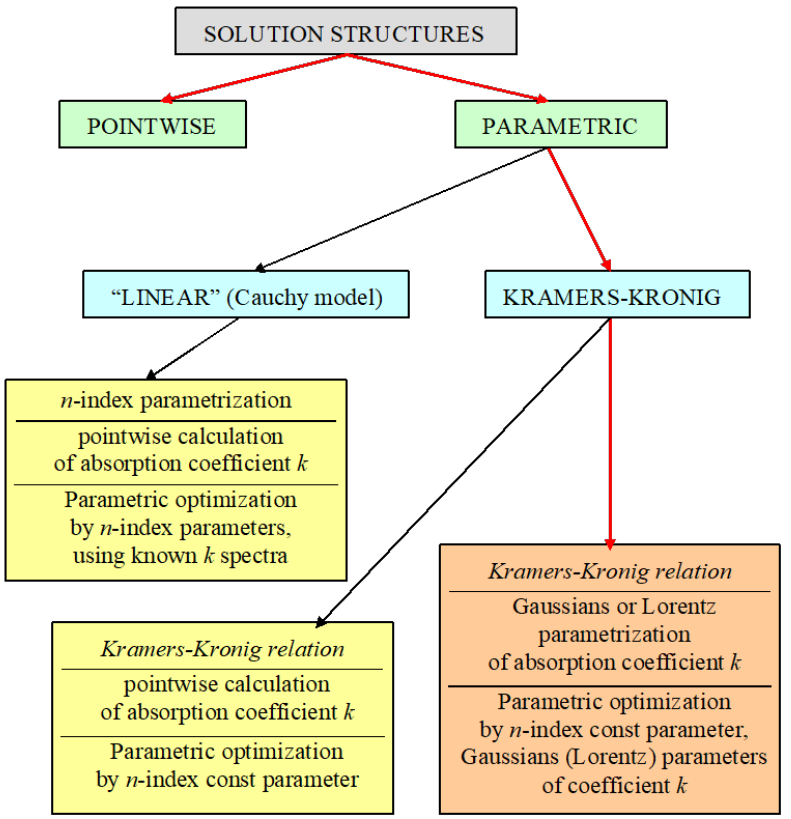


Figure 5. Inverse problems and alternative solution structures. Red field color indicates the most widely used structures

The software "Multilayer" allows the absorption peaks to be approximately described using both Gauss and Lorentz curves. The experience gained in practical calculations shows that the Gauss approximation has to be preferred. Questions pertaining to the Kramers–Kronig relations are considered in more detail in the next Section.

6. Kramers–Kronig relations for permittivity function

Assuming that the permittivity function ($\varepsilon = \varepsilon' + i\,\varepsilon''$) is analytic, we can write the well-known Kramers-Kronig relations between the real (ε') and imaginary (ε'') parts [7]:

$$\begin{aligned} \varepsilon'(\omega) - 1 &= \tfrac{1}{\pi} P \int_{-\infty}^{+\infty} \tfrac{\varepsilon''(x)}{x-\omega} dx, \\ \varepsilon''(\omega) &= -\tfrac{1}{\pi} P \int_{-\infty}^{+\infty} \tfrac{\varepsilon'(x)-1}{x-\omega} dx, \end{aligned} \tag{2}$$

where ω is the circular frequency and P is a symbol indicating the principal value of the integral.

For our purposes, the most important relation is (2) which, by virtue of the odd character of $\varepsilon''(\omega)$ [7], is equivalent to the following expression:

$$\varepsilon'(\omega) - 1 = \frac{2}{\pi} P \int_0^{+\infty} \frac{x\varepsilon''(x)}{x^2 - \omega^2} dx. \tag{3}$$

Experimental measurements can provide sufficiently reliable data on the absorption coefficient $k(\omega)$ (within a certain limited spectral interval $[\omega_{beg}, \omega_{end}]$), which is related to the complex permittivity and the refractive index $n(\omega)$ by the following expressions:

$$\varepsilon = (n(\omega) + ik(\omega))^2 \equiv n(\omega)^2 - k(\omega)^2 + 2in(\omega)k(\omega),$$

or

$$\varepsilon'(\omega) = n(\omega)^2 - k(\omega)^2, \quad \varepsilon''(\omega) = 2n(\omega)k(\omega), \tag{4}$$

where the absorption coefficient k is a dimensionless quantity.

Using equations (4) and (3), one can readily obtain an expression relating the refractive index at a given frequency to the spectrum of the absorption coefficient:

$$n(\omega)^2 - k(\omega)^2 - 1 = \frac{2c}{\pi} P \int_0^{+\infty} \frac{xn(x)k(x)}{x^2 - \omega^2} dx. \tag{5}$$

Thus, the main problem in determining the spectral dependence of the refractive index proceeding from the measured absorption spectrum reduces to solving an integral equation (3) with respect to $n(\omega)$. In practice, the absorption spectrum is measured in a certain limited spectral interval $[\omega_1, \omega_2]$. In such cases, equation (5) can be rewritten as

$$n(\omega)^2 - k(\omega)^2 - 1 = \frac{2c}{\pi} P \int_{\omega\_beg}^{\omega\_end} \frac{xn(x)k(x)}{x^2 - \omega^2} dx + C, \tag{6}$$

where an error arising on the passage to a finite integration interval is represented by a constant C, which, generally speaking, also has to be calculated from experimental data.

Let us consider the results of numerical solution of equation (6) for several model systems, where the absorption coefficients k are defined by a certain set of Gauss functions (Gaussians).

In figure 6 curves drawn in the same type of lines refer to the absorption coefficients (thin lines) and the corresponding refractive index profiles (thick lines) representing solutions of the nonlinear integral equation based on the Kramers–Kronig relations. An interesting example is offered by the pair of curves drawn by dashed lines, which just corresponds to the case under consideration with several absorption bands. The absorption is stronger in

the initial part of the spectral interval and then decreases. The restored refractive index also exhibits several maxima.

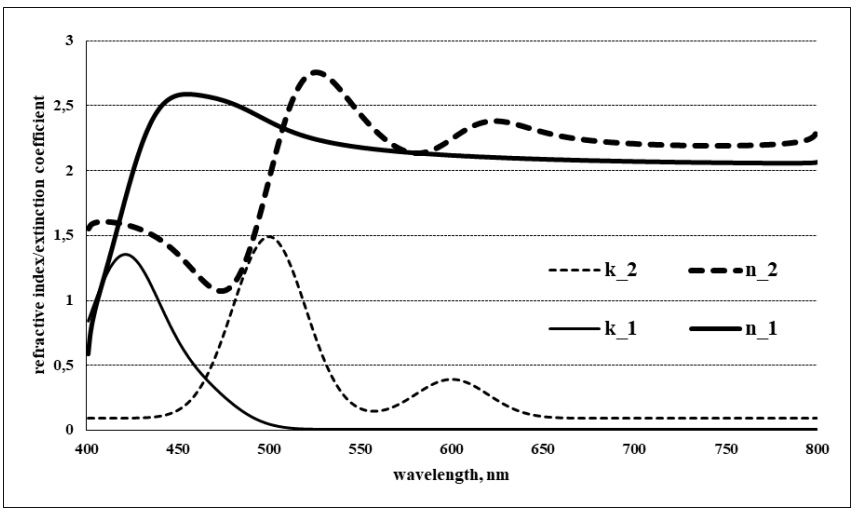


Figure 6. Kramers–Kronig relations. Two solutions (n_1, k_1) and (n_2, k_2) of integral equation (6)

The main approach to solution of the problem of determination of the refractive indices and absorption coefficients using the Kramers-Kronig relations consists in modeling the absorption coefficient by a sum of base functions describing the absorption bands of a given material. Since the spectral interval of measurements $[\lambda_{\text{beg}}, \lambda_{\text{end}}]$ contains by no means all absorption bands of the given material, it is necessary to provide for the possibility of extending the base functions outside this interval to a wider wavelength range $[\lambda^k_{beg},\ \lambda^k_{end}]$, where $\lambda^k_{beg} < \lambda_{\text{beg}} < \lambda_{\text{end}} < \lambda^k_{end}$. It is assumed that the main absorption bands of the materials under consideration fall within the interval $[\lambda_{\text{beg}}, \lambda_{\text{end}}]$ or go slightly outside.

The base functions are usually expressed on the frequency scale and then converted to the wavelength scale. For simulations using the developed software, the base functions can be selected in the following forms.

— *Gauss functions:* $G(\omega, x) = A \cdot \exp\left(-\frac{(\omega-\omega_0)^2}{\Delta}\right)$.

— *Lorentz functions:* $L(\omega, x) = \frac{A}{1+(\frac{\omega-\omega_0}{\Delta})^2}$.

As was noted above, Gaussians are more convenient in use, but Lorentz functions require a shorter time for calculations within each cycle of the iterative process. Use of the Lorentz curves is related to the following disadvantage:

— The shapes of the spectral curves of n and k are less perfect than those obtained by using Gauss functions.

— A significant error arises at the right-hand end of the spectral interval in the course of calculation of the degenerate integral for the refractive index (n). This error can be eliminated by artificially expanding the integration domain (approximately by 40-50 nm) outside the right-hand boundary of the given spectral interval.

7. Calculation of optical parameters for a thick layer. One-layer model

Let us consider some particular examples of calculations of the optical parameters of thick films based on the results of solutions of the corresponding inverse problems.

We'll take as a sample ***Quartz glass*** (KU-7 grade) [30].

The transmission and reflection spectra of a sample were measured on a spectrophotometer. The transmission intensity was measured at a zero incidence angle, while the reflection was measured at an angle of 7° (measured from the normal to the sample surface). The refractive index calculated for the spectral range from 450 to 750 nm is presented in figure 7. The maximum deviation from the standard is 7.0×10^{-4} at a wavelength of 633 nm (for the curve calculated using the Kramers-Kronig algorithm).

The pointwise algorithm provides a strongly oscillating solution, which is related to insufficient accuracy of the spectral measurements (noisy spectrum) used as the initial data.

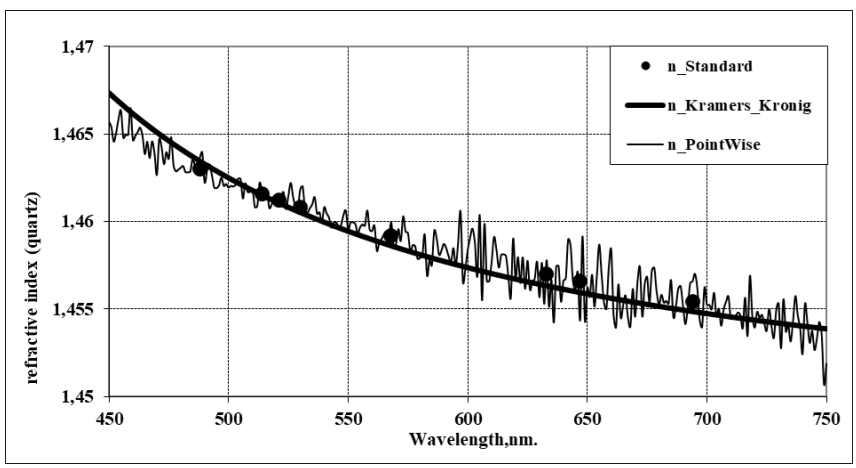


Figure 7. Refractive index of quartz: (Thin) pointwise algorithm; (Thick) Kramers-Kronig algorithm; (Points) standard (tabulated) values

Display glass

The transmission and reflection spectra of a sample were measured on a spectrophotometer using the same angles of incidence as those in the preceding example – see figures 8 and 9.

The refractive index calculated for the spectral range from 400 to 800 nm using the pointwise algorithm and the Kramers-Kronig algorithm (based on the Gaussian approximation of the absorption band) is presented in figure 10. The spectrum of the absorption coefficient is presented in figure 11.

The pointwise algorithm provides a strongly oscillating solution, which is related to insufficient accuracy of the spectral measurements (noisy spectrum) used as the initial data. It should be also noted that the accuracy of measurements in the beginning and at the end of the spectral interval is reduced because the employed polarizers operate more reliably in the middle of this interval than at the ends.

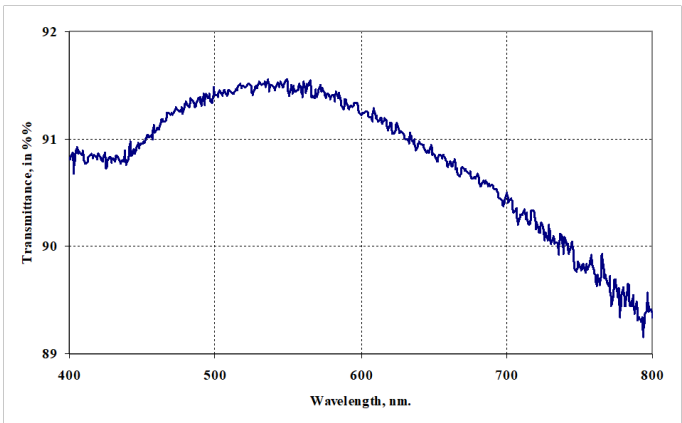


Figure 8. Measured transmission spectra of the glass sample

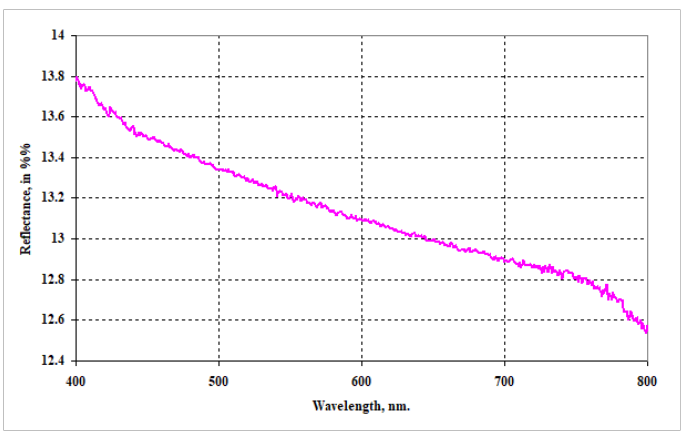


Figure 9. Measured reflection spectra of the glass sample

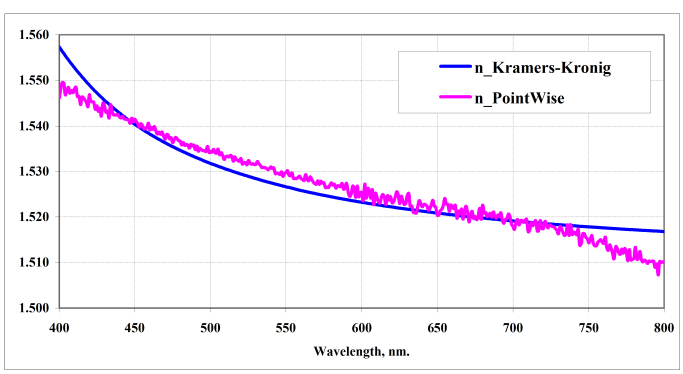


Figure 10. Refractive index of display glass:
(magenta) pointwise algorithm; (blue) Kramers-Kronig algorithm

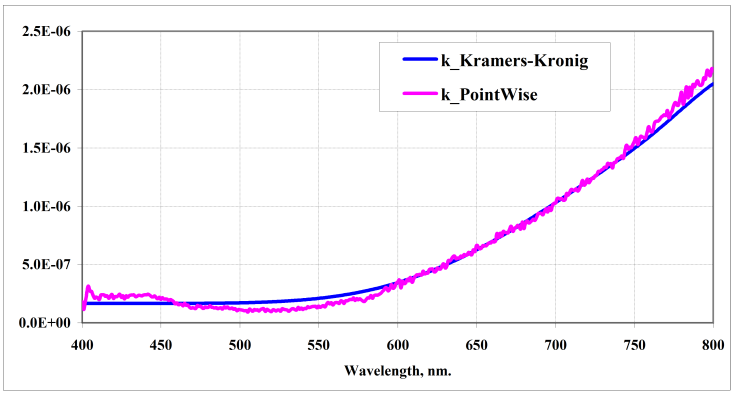


Figure 11. Absorption coefficients of display glass:
(magenta) pointwise algorithm; (blue) Kramers-Kronig algorithm

8. Calculation of optical parameters for thick and thin layers. Two-layers model

Calculations of the refractive indices of thin-film materials is complicated by the fact that their transmission and absorption characteristics cannot be measured directly and independently, since such materials are usually deposited onto substrates (with a thickness on the order of 1 mm) in the course of fabrication. For this reason, the problem is usually solved in at least two steps. First, it is necessary to determine the parameters of a substrate. Then, the calculated values of the refractive indices and absorption coefficients are used in solving the inverse problem for determining the characteristics of a thin coating.

Let us consider examples of the calculation of refractive indices and absorption coefficients for a tested real thin-film material.

In order to simplify and accelerate the computation procedures, the characteristics are first calculated for a slow direction of a sample. This approach provides sufficiently fast and more accurate evaluation of the film thickness as compared to the standard mechanical measurements with the help of profilometer.

Then, the refractive indices and absorption coefficients are calculated in the perpendicular – fast – direction, with the initial thickness approximation obtained in the first step. After that we can take those solutions as initial ones for solving the problem of index estimation in the third direction – along the vertical axis OZ in local coordinate system. The results of calculation are presented on figures 12 for real parts (refraction indices) and 13 for imaginary parts (absorption coefficients). Figures show the calculated values of principal axes of the refractive index ellipsoid as functions of the wavelength.

Optimization with respect to all components of the refractive index ellipsoid, the Euler angles, and the film thickness at each point of measurements gives the results of fitting for measured transmitted and reflected light depicted in figures 14–15.

Figures 14 and 15 illustrate the degree of approximation for all eight measurements used in the solution of the inverse problem. The measured and calculated intensities of light transmitted via a thin film deposited onto

a display glass are compared in figure 14. The intensities of light reflected in various directions for different angles of incidence are presented in figure 15.

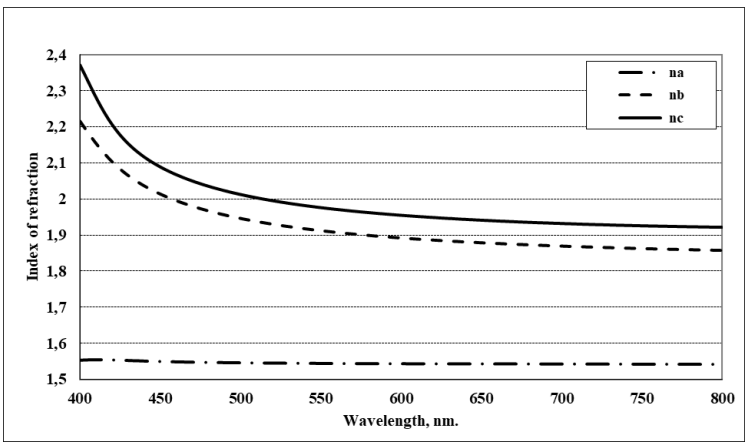


Figure 12. Refractive indices along three axes in local coordinate system

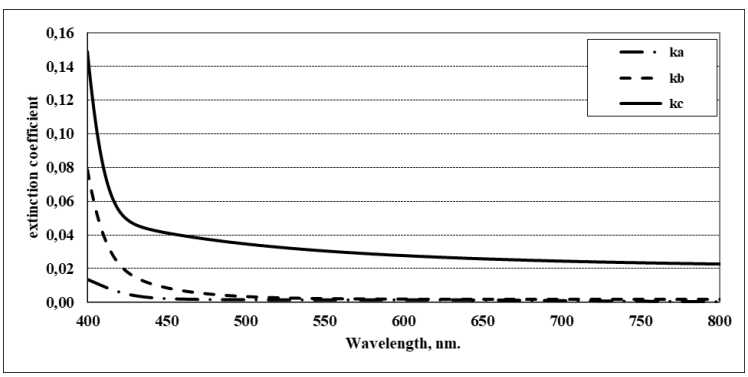


Figure 13. Extinction coefficients along three axes in local coordinate system

It should be noted that the results of fitting with respect to transmission are better than those for reflection, which is related to the fact that the measurements of reflection are more influenced by the scattering of light. Nevertheless, the positions of extrema in the calculated and measured reflection spectra exhibit good coincidence.

9. Conclusion

The paper describes a practically tested methodology for the sequential determination of the optical and physical properties (permittivity and thickness) of anisotropic thin-film materials. This approach ensures the determination of the required parameters with high accuracy, which makes it possible in the future to design multilayer thin-film coatings with specified properties. Such coatings can have the properties of absolute mirrors or absolutely black bodies in a given wavelength range, optical filters with desired characteristics, etc.

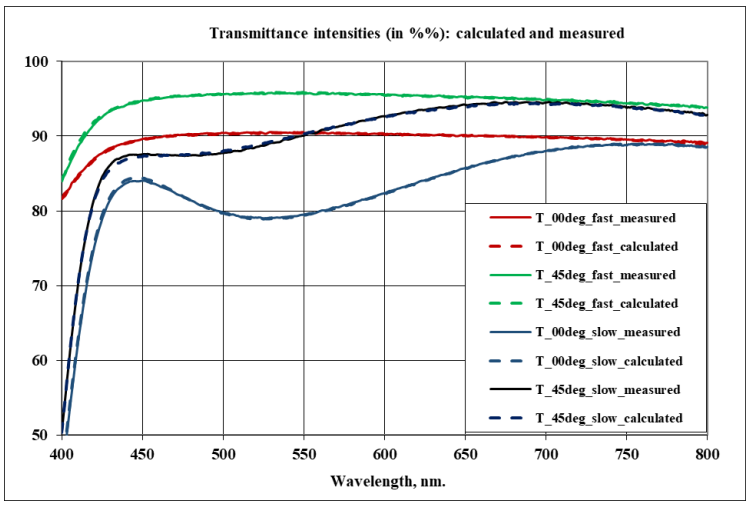


Figure 14. Result of fitting procedure. Calculated (dashed) and measured (solid lines) transmittances for different angles of incident s- and p-polarized light

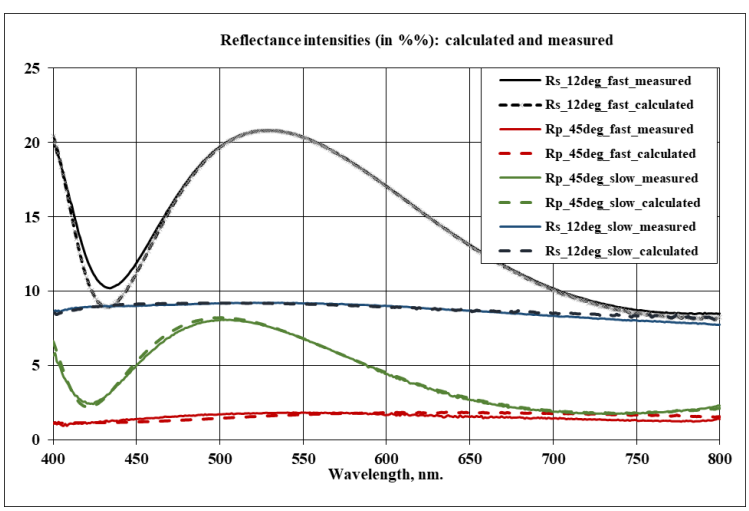


Figure 15. Result of fitting procedure. Calculated (dashed) and measured (solid lines) reflectances for different angles of incident s- and p-polarized light

Depending on the types of the investigated coatings, the user is offered several options (models) for carrying out (executing) calculations. The simplest approach is to solve the inverse problem at each wavelength using the transmission and reflection measured at this wavelength only. In this case, the problem has an (infinite) set of feasible solutions. The choice of the one that is suitable in terms of physical meaning is very difficult for algorithmisation and laborious.

The second approach is to use a priori information on the continuity of the components of the permittivity tensor depending on the wavelength. From a practical point of view, the most successful variant seems to be the approximation of the imaginary part of the tensor by a set of Gaussians (the

parameters of which are to be determined). The real part of the permittivity tensor is calculated in accordance with the Kramers-Kronig relation.

The complexity of solving each of the problems is very different and each requires its own specific set of measured input data. Therefore, the solution to each problem is implemented as a separate option within the software package, although the solution methodology remains the same.Examples of determining the permittivity of two different standard samples, namely, the KU-7 silica glass and the display glass, from the measured transmission and reflection in the visible range are considered. The obtained values of the refractive index coincide with the values declared by the manufacturer with an accuracy of 4 decimal places.

The efficiency of the method and algorithm for the sequential determination of the permittivity tensor of a thin film deposited on a glass substrate is demonstrated. First, the permittivity of the isotropic substrate is determined and its thickness is refined. At the second stage, after the deposition of a thin anisotropic film on the substrate, the parameters of which must be determined, a series of measurements of transmission and reflection from the two-layer surface is carried out at different angles and in mutually perpendicular directions. The ultimate results of solving the inverse problem are verified by comparing the calculated transmission and reflection with those measured for arbitrary angles of incidence and reflection.

Acknowledgments

The publication has been prepared with the support of the Russian Foundation for Basic Research (RFBR) according to the research project No 18-07-00567.

References

[1] D. A. Yakovlev, V. G. Chigrinov, and H. S. Kwok, *Modeling and Optimization of LCD Optical Performance*. New York: Wiley, 2015.

[2] J. A. Dobrowolski, F. C. Ho, and A. Waldorf, "Determination of optical constants of thin film coating materials based on inverse synthesis," *Applied Optics*, vol. 22, no. 20, pp. 3191–3200, 1983. DOI: 10.1364/AO.22.003191.

[3] X. Cheng, B. Fan, J. A. Dobrowolski, L. Wang, and Z. Wang, "Gradient-index optical filter synthesis with controllable and predictable refractive index profiles," *Optics Express*, vol. 16, no. 4, pp. 2315–3221, 2008. DOI: 10.1364/OE.16.002315.

[4] A. Tejada *et al.*, "Determination of the fundamental absorption and optical bandgap of dielectric thin films from single optical transmittance measurements," *Applied Optics*, vol. 58, no. 35, pp. 9585–9594, 2019. DOI: 10.1364/AO.58.009585.

[5] J. B. Bell, A. N. Tikhonov, and V. Y. Arsenin, "Solutions of Ill-Posed Problems," *Mathematics of Computation*, vol. 32, no. 144, pp. 1320–1322, 1978.

[6] S. Nevas, F. Manoocheri, E. Ikonen, A. V. Tikhonravov, M. A. Kokarev, and M. K. Trubetskov, "Optical metrology of thin films using high-accuracy spectrophotometric measurements with oblique angles of incidence," in *Advances in Optical Thin Films*, International Society for Optics and Photonics, vol. 5250, SPIE, 2004, pp. 234–242. DOI: 10.1117/12.512700.

[7] L. D. Landau and E. M. Lifshitz, *Electromagnetic Waves in Anisotropic Media*. Oxford: Pergamon Press, 1984.

[8] A. V. Tikhonravov *et al.*, "Effect of systematic errors in spectral photometric data on the accuracy of determination of optical parameters of dielectric thin films," *Applied Optics*, vol. 41, no. 13, pp. 2555–2560, 2002. DOI: 10.1364/AO.41.002555.

[9] M. Nur-E-Alam, M. M. Rahman, M. K. Basher, M. Vasiliev, and K. Alameh, "Optical and chromaticity properties of metal-dielectric composite-based multilayer thin-film structures prepared by rf magnetron sputtering," *Coatings*, vol. 10, no. 3, p. 251, 2020. DOI: 10.3390/coatings10030251.

[10] S. A. Furman and A. V. Tikhonravov, *Basics of optics of multilayer systems*. Singapore: World Scientific Publishing, 1992.

[11] M. Paukshto, K. Lovetskiy, and A. Zhukov, "P-59: Dielectric Constants of Display Optical Components," *SID Symposium Digest of Technical Papers*, vol. 38, no. 1, pp. 410–413, 2007. DOI: 10.1889/1.2785320.

[12] M. Paukshto, K. Lovetsky, A. Zhukov, V. Smirnov, D. Kibalov, and G. King, "P-168: Simulation of Sub-100nm Gratings Incorporated in LCD Stack," *SID Symposium Digest of Technical Papers*, vol. 37, no. 1, pp. 848–850, 2006. DOI: 10.1889/1.2433649.

[13] A. M. Alsaad *et al.*, "Measurement and ab initio Investigation of Structural, Electronic, Optical, and Mechanical Properties of Sputtered Aluminum Nitride Thin Films," *Frontiers in Physics*, vol. 8, p. 115, 2020. DOI: 10.3389/fphy.2020.00115.

[14] R. M. A. Azzam and N. M. Bashara, *Ellipsometry and polarized light*. Amsterdam: North-Holland Pub. Co., 1977.

[15] Q. M. Al-Bataineh, A. M. Alsaad, A. A. Ahmad, and A. Telfah, "A novel optical model of the experimental transmission spectra of nanocomposite PVC-PS hybrid thin films doped with silica nanoparticles," *Heliyon*, vol. 6, no. 6, p. 04 177, 2020. DOI: 10.1016/j.heliyon.2020.e04177.

[16] A. V. Tikhonravov and M. K. Trubetskov. (2020). "OptiChar Software," [Online]. Available: http://www.optilayer.com.

[17] M. Born and E. Wolf, *Principles of Optics*. London: Pergamon Press, 1980.

[18] T. L. Watkins and J. Fendley, "Refractive index," *Physics Education*, vol. 18, no. 2, p. 56, 1983. DOI: 10.1088/0031-9120/18/2/102.

[19] K. P. Lovetskiy, N. E. Nikolaev, and A. L. Sevastianov, "Optical Characterization of a Thin-Film Material Based on Light Intensity Measurements," *RUDN Journal of Mathematics, Information Sciences and Physics*, vol. 26, no. 3, pp. 252–260, 2018. DOI: 10.22363/2312-9735-2018-26-3-252-260.

[20] J. A. Nelder and R. Mead, "A Simplex Method for Function Minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965. DOI: 10.1093/comjnl/7.4.308.

[21] S.-Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on polar decomposition," *Journal of the Optical Society of America A*, vol. 13, no. 5, pp. 1106–1113, 1996. DOI: 10.1364/JOSAA.13.001106.

[22] P. Yeh, "Extended Jones Matrix Method," *Journal of the Optical Society of America*, vol. 72, no. 4, pp. 507–513, 1982. DOI: 10.1364/JOSA.72.000507.

[23] D. A. Yakovlev and V. G. Chigrinov, "A robust polarization-spectral method for determination of twisted liquid crystal layer parameters," *Journal of Applied Physics*, vol. 102, no. 2, p. 023 510, 2007. DOI: 10.1063/1.2756377.

[24] A. Yariv and P. Yeh, *Optical Waves in Crystals*. New York: John Wiley and Sons, Inc, 2003.

[25] F. I. Fedorov, *Optics of Anisotropic Media [Optika anizotropnykh sred]*. Minsk: Academy of Sciences of Belarus, 1958, in Russian.

[26] D. W. Berreman, "Optics in Stratified and Anisotropic Media: 4×4 – Matrix Formulation," *The Journal of the Optical Society of America*, vol. 62, no. 4, pp. 502–510, 1972.

[27] T. F. Isaev, I. V. Kochikov, D. V. Lukyanenko, A. V. Tikhonravov, and A. G. Yagola, "Comparison of Algorithms for Determining the Thickness of Optical Coatings Online," *Computational Mathematics And Mathematical Physics*, vol. 59, no. 3, pp. 465–474, 2019. DOI: 10.1134/S0965542519030102.

[28] J. L. M. Van Mechelen, A. B. Kuzmenko, and H. Merbold, "Stratified dispersive model for material characterization using terahertz time-domain spectroscopy," *Optics Letters*, vol. 39, no. 13, pp. 3853–3856, 2014. DOI: 10.1364/OL.39.003853.

[29] A. B. Kuzmenko, "Kramers-Kronig constrained variational analysis of optical spectra," *Review of Scientific Instruments*, vol. 76, no. 8, pp. 1–9, 2005. DOI: 10.1063/1.1979470.

[30] G. Ghosh, "Refractive Index of Quartz for Thin Film Thickness Measurement," *Optics Communications*, vol. 163, pp. 95–102, 1999.

For citation:

K. P. Lovetski, A. A. Zhukov, M. V. Paukshto, L. A. Sevastianov, A. A. Tiutiunnik, Solving the inverse problem for determining the optical characteristics of materials, Discrete and Continuous Models and Applied Computational Science 28 (4) (2020) 378–397. DOI: 10.22363/2658-4670-2020-28-4-378-397.

Information about the authors:

Lovetskiy, Konstantin P. — Candidate of Physical and Mathematical Sciences, assistant professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: lovetskiy-kp@rudn.ru, phone: +7(495)9522572, ORCID: https://orcid.org/0000-0002-3645-1060, ResearcherID: A-5725-2017, Scopus Author ID: 18634692900)

Zhukov, Andrey A. — PhD, lead analyst of "ITL Consulting" company (e-mail: a.zhukov@itlc.ru, phone: +7(495)6629522, ORCID: https://orcid.org/0000-0002-4006-2922)

Paukshto, Michael V. — DSc., Physics & Mechanical Engineering, co-founder and CTO of Fibralign Corporation (e-mail: mpaukshto@fibralignbio.com, phone: 650 492 1440, ORCID: https://orcid.org/0000-0002-8090-5402, Scopus Author ID: 6701433110)

Sevastianov, Leonid A. — Doctor of Physical and Mathematical Sciences, professor of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: sevastianov-la@rudn.ru, phone: +7(495)9522572, ORCID: https://orcid.org/0000-0002-1856-4643, ResearcherID: B-8497-2016, Scopus Author ID: 8783969400)

Tiutiunnik, Anastasiia A. — Candidate of Physical and Mathematical Sciences, lecturer of Department of Applied Probability and Informatics of Peoples' Friendship University of Russia (RUDN University) (e-mail: tyutyunnik-aa@rudn.ru, phone: +7(495)9550783, ORCID: https://orcid.org/0000-0002-4643-327X, ResearcherID: B-8441-2019, Scopus Author ID: 57188573409)

УДК 535:535.3:681.7
DOI: 10.22363/2658-4670-2020-28-4-378-397

Решение обратной задачи определения оптических характеристик материалов

К. П. Ловецкий[1], **А. А. Жуков**[2], **М. В. Паукшто**[3], **Л. А. Севастьянов**[1], **А. А. Тютюнник**[1]

[1] *Российский университет дружбы народов*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия
[2] *ITL Consulting*
Ольховская, д. 16, корп. 5, Москва, 105066, Россия
[3] *Fibralign Corporation*
Альварадо-Найлс Роуд, д. 32930, офис 350, Юнион-Сити, CA 94587, США

В работе изложена методология определения оптических и физических свойств анизотропных тонкоплёночных материалов. Такой подход позволяет в дальнейшем проектировать многослойные тонкоплёночные покрытия с заданными свойствами. Сформулирована обратная задача определения тензора диэлектрической проницаемости и толщины тонкой плёнки, нанесённой на стеклянную подложку, с известными оптическими свойствами и толщиной. Предварительная информация о принадлежности тонкоплёночного покрытия к определённому классу позволяет значительно сократить время расчёта и увеличить точность определения тензора диэлектрической проницаемости на всём исследуемом интервале длин волн и толщины плёнки в точке измерения отражения и пропускания.

В зависимости от поставленных целей возможна постановка и, следовательно, решение различных обратных задач:

– определение тензора диэлектрической проницаемости и уточнение толщины толстой (до 1 см) подложки, часто изотропной;

– определение тензора диэлектрической проницаемости тонкой изотропной или анизотропной плёнки, нанесённой на подложку, с известными оптическими свойствами.

Сложность решения каждой из задач весьма различна и каждая требует своего определённого набора измеренных входных данных. Окончательные результаты решения обратной задачи верифицируются с помощью сравнения вычисленных коэффициентов пропускания и отражения с измеренными для произвольных углов падения и отражения.

Ключевые слова: определение коэффициентов пропускания, отражения, показателей преломления, тонкие плёнки, многослойные материалы, оптические покрытия, оптические свойства

Research article

UDC 519.6
DOI: 10.22363/2658-4670-2020-28-4-398-405

Numerical modeling of laser ablation of materials

Ilkizar V. Amirkhanov, Nil R. Sarker, Ibrohim Sarkhadov

Laboratory of Information Technology
Joint Institute for Nuclear Research
6, Joliot-Curie St., Dubna, Moscow region, 141980, Russian Federation

(received: November 10, 2020; accepted: November 12, 2020)

In this paper, we report a numerical simulation of laser ablation of a material by ultrashort laser pulses. The thermal mechanism of laser ablation is described in terms of a one-dimensional nonstationary heat conduction equation in a coordinate system associated with a moving evaporation front. The laser action is taken into account through the functions of the source in the thermal conductivity equation that determine the coordinate and time dependence of the laser source. For a given dose of irradiation of the sample, the profiles of the sample temperature at different times, the dynamics of the displacement of the sample boundary due to evaporation, the velocity of this boundary, and the temperature of the sample at the moving boundary are obtained. The dependence of the maximum temperature on the sample surface and the thickness of the ablation layer on the radiation dose of the incident laser pulse is obtained.

Numerical calculations were performed using the finite difference method. The obtained results agree with the results of other works obtained by their authors.

Key words and phrases: Numerical simulation, ablation, pulsed lasers, heat conduction equation

1. Introduction

In recent years, pulsed laser ablation [1]–[3] (any process of laser-stimulated removal of matter, including the emission of electrons) of various materials has attracted more and more interest from the point of view of fundamental study of processes in matter under extreme conditions of ultrafast energy supply. This implies constructing a new physical theory describing strongly nonlinear effects.

For a detailed analysis of the processes in the experiment, it is required to measure various characteristics of the ablation processes with pico- and femtosecond time resolution, which in itself is a rather difficult task. Therefore, the problem of mathematical modeling of physical phenomena in this area becomes extremely urgent.

To describe the dynamics of fast processes in a substance, the method of molecular dynamics (MD) can be used [4]. MD is quite effective for

microscopic analysis of the mechanisms of melting and evaporation under overheating conditions both in the bulk of the target [5] and for a system with a free surface [6]. The emergence and propagation of pressure waves generated by laser radiation [7], [8], as well as the dynamics of laser ablation [9], is well modeled using MD.

In this paper, we consider continuous methods (various modifications of the heat equation) for modeling the effect of laser radiation on matter.

The evaporation process is mathematically described within the framework of the boundary value problem of thermal conductivity for a condensed medium in a coordinate system associated with a moving solid-vapor interface or a melt-vapor interface on which evaporation occurs. If we do not take into account the lateral removal of the laser radiation energy due to thermal conductivity, which is valid under the strict condition $r_0 \gg \sqrt{a_T \tau}$, where τ is the duration of the laser beam exposure to the material, a_T is the thermal conductivity, r_0 is the radius of the overheating spot, then the problem of the motion of the evaporation boundary can be considered within the framework of the one-dimensional model [3]. In Ref. [10], the primary results of numerical simulation of ablation of materials were published. In this paper, the required work is presented in a more extended form.

2. Setting of the problem

Numerical modeling of laser ablation of materials was carried out based on the heat conduction equation written in a moving coordinate system associated with the evaporation front, with initial and boundary conditions [2]:

$$\rho(T)c(T)\left[\frac{\partial T}{\partial t} - v(T_s)\frac{\partial T}{\partial z}\right] = \frac{\partial}{\partial z}\left(\lambda(T)\frac{\partial T}{\partial z}\right) + A(z,t), \quad 0 < z < z_{\max}, \quad (1)$$

$$T(z,0) = T_0; \quad 0 \leqslant z \leqslant z_{\max}, \quad (2)$$

$$\lambda(T)\left.\frac{\partial T(z,t)}{\partial z}\right|_{z=0} = v(T_s)L_{ev}\rho;$$
$$T(z_{\max},t) = T_0; \quad h = \int_0^t v(t)dt, \quad T_s = T(0,t), \quad (3)$$

where $c(T)$, $\lambda(T)$, $\rho(T)$ are the specific heat, thermal conductivity and density of the material at the temperature $T(z,t)$, $h(t)$, respectively is the depth of the crater on the surface of the sample at time t, z_m is the maximum distance, $v(T_s)$ is the velocity of the boundary displacement due to evaporation, L_{ev} is the specific heat of sublimation. The source function $A(z,t)$ has the form [2]

$$A(z,t) = f_1(z)f_2(t),$$
$$f_1(z) = A_s\alpha e^{-\alpha z}e^{-\alpha_g h}, \quad A_s = 1 - R(T_s), \quad f_2(t) = I_0 f(t). \quad (4)$$

Here I_0 is the laser intensity, $R(T_s)$ is the reflection coefficient of the laser from the sample surface, α, α_g are the absorption coefficients of the laser

pulse in the sample material and in the vapor, respectively. The irradiation dose Φ, the intensity of the source I_0 and the temporal form of the source $f(t)$ are related by the relation:

$$\Phi = I_0 \int_0^\infty f(t)dt.$$

Here the source function has a factorized form, as in the work [11], when the material is affected by a pulsed beam of charged particles rather than by a laser pulse. In general, the heat capacity, thermal conductivity, and density of the material depend on temperature. In a particular case, the dependence of some parameters of the sample material can be neglected. In this work, the temperature dependence of the density of the sample material and the laser reflection coefficient is neglected.

3. Discussion of numerical results

In Ref. [2], problem (1)–(4) was solved by the method of moments for a polyimide material. In our work, this problem was solved using the finite difference explicit scheme [12]. The temporal shape of the source $f(t)$, the temperature dependence of the boundary motion velocity due to evaporation $v_s(T)$, the specific heat $c(T)$ and the thermal conductivity $\lambda(T)$ are taken for the polyimide material similar as in Ref. [2]:

$$f(t) = \frac{t}{t_1}\exp\left\{-\frac{t}{t_1}\right\};\ t_1 = 6.13\,ns,\ c(T) = 2550{-}1590{\cdot}\exp\left\{\frac{300-T}{460}\right\}\ \frac{J}{kgK},$$

$$\lambda(T) = 0.155{\cdot}\left(\frac{T}{300}\right)^{0.28}\ \frac{W}{mK},\ v = v_0 e^{-T_a/T_s},\ v_0 = 3{\cdot}10^4 m/s;\ T_a = 15700\,K.$$

Figure 1 shows plots of these dependencies.

Figure 2 shows the temperature profiles of a polyimide sample at different times: $t_j = j \cdot 5\ ns;\ j = 1.10$, the dynamics of the sample boundary motion due to evaporation, the velocity of this boundary motion, and the sample temperature at the moving boundary $x = h(t)$, when exposed to energy fluence $\Phi = 10^3\ J/m^2$ with parameters $A_s = 0.93$, $\alpha = 4.25 \cdot 10^7\ \mathrm{m}^{-1}$ ($\alpha_g = 0.45\alpha$), $L_{ev} = 5 \cdot 10^5 J/kg$, $\rho = 1420\ kg/m^3$.

Figure 3 shows the dependencies of the maximal temperature at the sample surface $T_{max}(h(t), t)$ and the crater depth $h(t)$ on the radiation dose Φ for four sets of values of parameters A_s and α:

1) $A_s = 0.93$, $\alpha = 4.25 \cdot 10^7\ \mathrm{m}^{-1}$;
2) $A_s = 0.88$, $\alpha = 3.1 \cdot 10^7\ \mathrm{m}^{-1}$;
3) $A_s = 0.89$, $\alpha = 10^7\ \mathrm{m}^{-1}$;
4) $A_s = 0.9$, $\alpha = 0.32 \cdot 10^7\ \mathrm{m}^{-1}$.

The source intensity I_0 in this case varies from $3 \cdot 10^6\ \mathrm{W/cm}^2$ to $3 \cdot 10^7\ \mathrm{W/cm}^2$. The obtained results agree with those of Ref. [2].

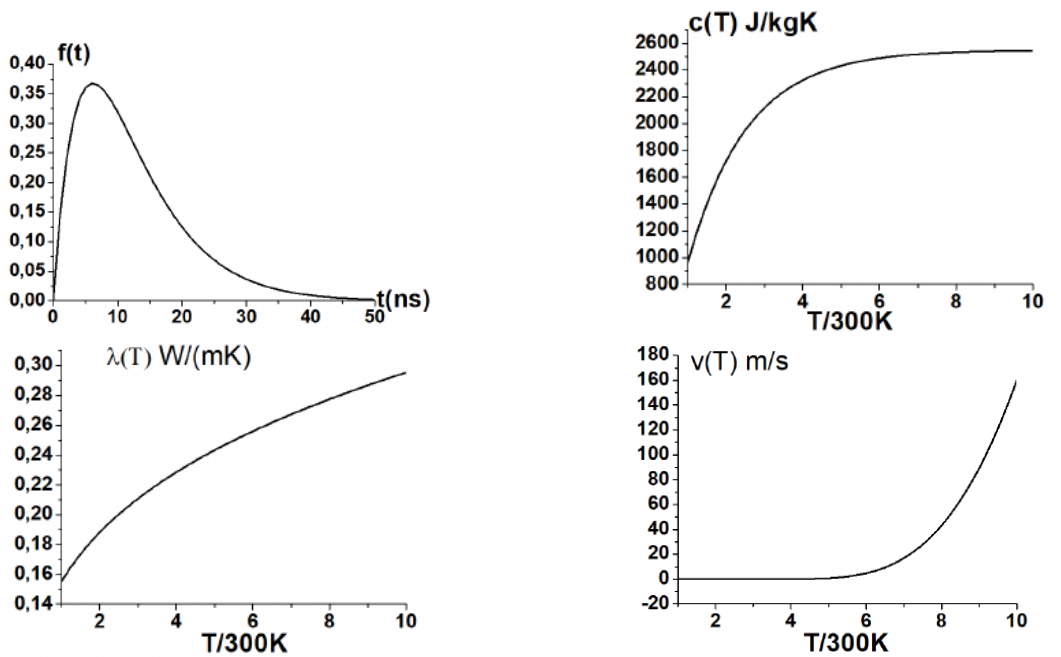


Figure 1. Source temporal shape $f(t)$, temperature dependence of specific heat $c(T)$, thermal conductivity $\lambda(T)$ and boundary motion velocity $v(T)$ due to evaporation

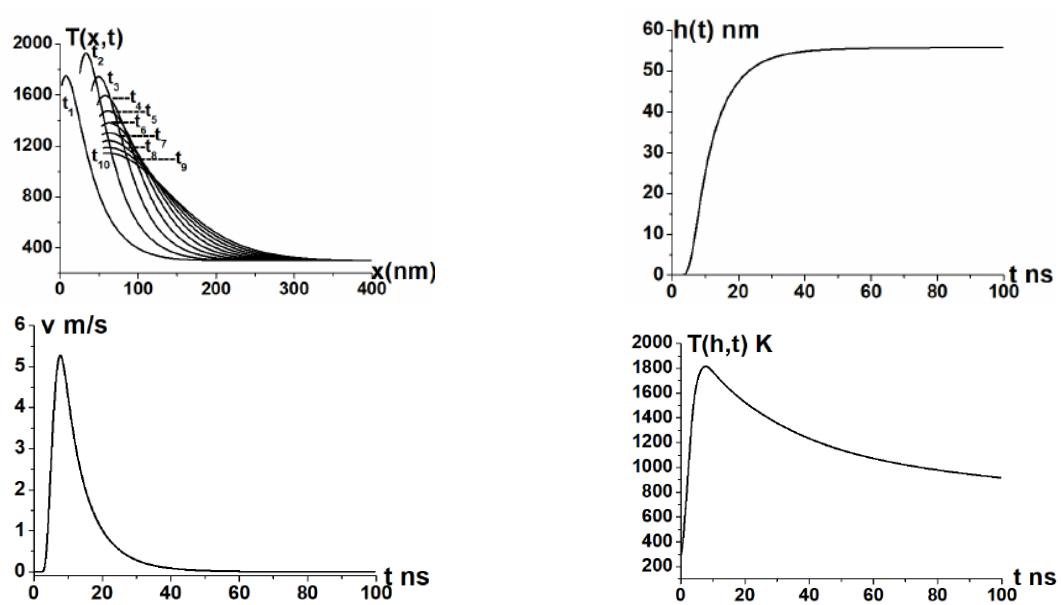


Figure 2. Temperature profiles of polyimide sample at different times: $t_j = j \cdot 5\ ns;\ j = 1, 10$, dynamics of sample boundary motion due to evaporation, velocity of this boundary and the sample temperature at the moving boundary $x = h(t)$ under the exposure to energy fluence $\Phi = 10^3\ J/m^2$

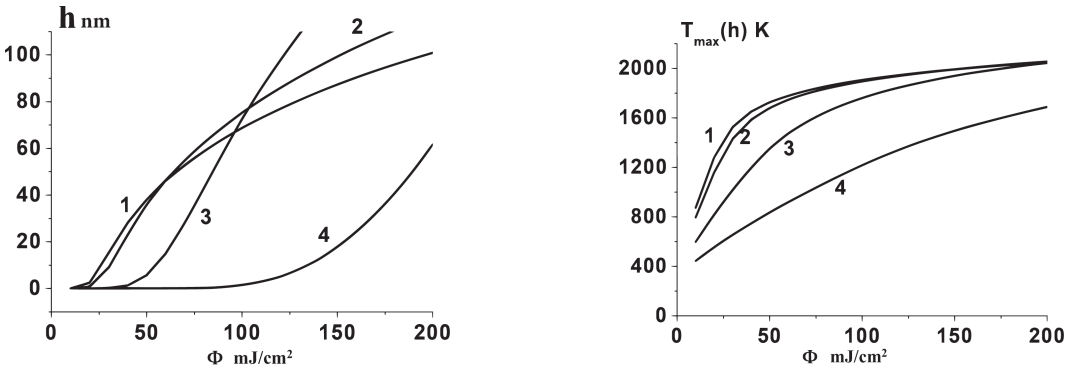


Figure 3. Dependencies of the maximal temperature at the sample surface $T_{max}(h(t), t)$ and the crater depth $h(t)$ on the irradiation doze Φ for four sets of values for A_s, α

4. Conclusion

For a given dose of the sample irradiation, the profiles of the sample temperature at different times, the dynamics of the displacement of the sample boundary due to evaporation, the velocity of this boundary and the temperature of the sample at the moving boundary were obtained. The dependencies of the temperature maximum on the sample surface and the thickness of the ablation layer on the radiation dose of the incident laser pulse are determined.

Numerical calculations were performed using the finite difference method. The obtained results agree with the results of works of other authors. When using shorter laser pulses in the ablation kinetics, arised features that can no longer be described within the framework of the conventional thermal model. In this case, studies are carried out within the framework of other models (two-temperature model, hydrodynamic model, etc.), which is the subject of further research.

Acknowledgments

The work was carried out by financial support of Russian Foundation for Basic Research No. 19-01-00645a and No. 20-51-44001 mong-a.

References

[1] L. A. Zakharov and N. M. Bulgakov, "Numerical simulation of laser ablation of metals and polymers when exposed to pulses of infrared radiation: the effect of the initial temperature of the sample [Chislennoe modelirovanie lazernoj ablyacii metallov i polimerov pri vozdejstvii impul'sami infrakrasnogo izlucheniya: vliyanie nachal'noj temperatury obrazca]," *Vestnik NGU. Seriya: Fizika*, vol. 5, no. 1, pp. 39–47, 2010, in Russian.

[2] S. I. Anisimov and B. S. Lukyanchuk, "Selected problems of laser ablation theory," *Phys. Usp.*, no. 45, pp. 293–324, 2002. DOI: 10.1070/PU2002v045n03ABEH000966.

[3] V. P. Veiko, M. N. Libenson, G. G. Chervyakov, and E. B. Yakovlev, *Interaction of Laser Radiation with Matter. Power Optics [Vzaimodeistvie lazernogo izlucheniya s veshchestvom. Silovaya optika]*. Moscow: Fizmatlit, 2008, in Russian.

[4] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids*. Walton Street, Oxford OX2 6DP: Clarendon Press, 1991.

[5] Z. H. Jin, P. Gumbsch, K. Lu, and E. Ma, "Melting mechanisms at the limit of superheating," *Physical Review Letters*, vol. 87, no. 5, p. 055 703, Jul. 2001. DOI: 10.1103/PhysRevLett.87.055703.

[6] F. F. Abraham and J. Q. Broughton, "Pulsed melting of silicon (111) and (100) surfaces simulated by molecular dynamics," *Physical Review Letters*, vol. 56, no. 7, pp. 734–737, 1986. DOI: 10.1103/PhysRevLett.56.734.

[7] L. V. Zhigilei and B. J. Garrison, "Pressure waves in microscopic simulations of laser ablation," *Materials Research Society Symposium – Proceedings*, vol. 538, pp. 491–496, 1999.

[8] J. I. Etcheverry and M. Mesaros, "Molecular dynamics simulation of the production of acoustic waves by pulsed laser irradiation," *Physical Review B*, vol. 60, no. 13, pp. 9430–9434, 1999. DOI: 10.1103/PhysRevB.60.9430.

[9] L. V. Zhigilei and B. J. Garrison, "Microscopic mechanisms of laser ablation of organic solids in the thermal and stress confinement irradiation regimes," *Journal of Applied Physics*, vol. 88, no. 3, pp. 1281–1298, 2000. DOI: 10.1063/1.373816.

[10] I. V. Amirkhanov, N. R. Sarker, and I. Sarkhadov, "Numerical simulation of laser ablation of materials," in *Proceedings of the 10th International Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems" (ITTMM-2020), Moscow*, in Russian, 2020, pp. 237–239.

[11] I. V. Amirkhanov, E. V. Zemlyanaya, I. V. Puzynin, T. P. Puzynina, and I. Sarhadov, *On the influence of the source shape in the model of phase transitions in metals irradiated with pulsed ion beams [O vliyanii formy istochnika v modeli fazovyh perekhodov v metallah, obluchaemyh impul'snymi puchkami ionov]*. Dubna: JINR Communication P11-2002-78, 2002, p. 18, in Russian.

[12] A. A. Samarskiy, *The theory of difference schemes [Teoriya raznostnyh skhem]*, Russian. Moscow: Nauka, 1983, pp. 258–276, in Russian.

For citation:

I. V. Amirkhanov, N. R. Sarker, I. Sarkhadov, Numerical modeling of laser ablation of materials, Discrete and Continuous Models and Applied Computational Science 28 (4) (2020) 398–405. DOI: 10.22363/2658-4670-2020-28-4-398-405.

Information about the authors:

Amirkhanov, Ilkizar V. — Candidate of Physical and Mathematical Sciences, Head of Sector "Scientific Division of Computational Physics". Laboratory of Information Technologies of the Joint Institute for Nuclear Research

(e-mail: camir@jinr.ru, ORCID: https://orcid.org/0000-0003-2621-144X, Scopus Author ID: 6507929197)

Sarker, Nil R. — Candidate of Physical and Mathematical Sciences, Senior Researcher "Scientific Division of Computational Physics". Laboratory of Information Technologies of the Joint Institute for Nuclear Research (e-mail: sarker@jinr.ru, ORCID: https://orcid.org/0000-0003-0690-2534, Scopus Author ID: 14829498100)

Sarkhadov, Ibrohim — Candidate of Physical and Mathematical Sciences, Senior Researcher "Scientific Division of Computational Physics". Laboratory of Information Technologies of the Joint Institute for Nuclear Research (e-mail: ibrohim@jinr.ru, ORCID: https://orcid.org/0000-0001-5534-3332, Scopus Author ID: 14829718300)

УДК 519.6
DOI: 10.22363/2658-4670-2020-28-4-398-405

Численное моделирование лазерной абляции материалов

И. В. Амирханов, Н. Р. Саркер, И. Сархадов

Лаборатория информационных технологий
Объединённый институт ядерных исследований
ул. Жолио-Кюри, д. 6, Дубна, Московская область, 141980, Россия

В работе проведено численное моделирование лазерной абляции материала под действием ультракоротких лазерных импульсов. Тепловой механизм лазерной абляции описывается в рамках одномерного нестационарного уравнения теплопроводности в системе координат, связанной с движущимся фронтом испарения. Действие лазера учитывается через функции источника в уравнении теплопроводности, задавая координатную и временную зависимости источника лазера. Для заданной дозы облучения образца получены профили температуры образца при разных временах, динамике перемещения границы образца из-за испарения, скорости перемещения этой границы и температуры образца на движущейся границе. Получены зависимость максимума температуры на поверхности образца и толщины слоя абляции от дозы излучения падающего лазерного импульса.

Численные расчеты проведены с применением метода конечных разностей. Полученные результаты согласуются с результатами работ других исследователей.

Ключевые слова: численное моделирование, абляция, импульсные лазеры, уравнение теплопроводности