DISCRETE AND CONTINUOUS MODELS AND APPLIED COMPUTATIONAL SCIENCE

Volume 31 Number 4 (2023)

Founded in 1993
Founder: PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA

DOI: 10.22363/2658-4670-2023-31-4

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 named after Patrice Lumumba (RUDN University).
Indexed by Scopus (https://www.scopus.com), Ulrich's Periodicals Directory (http://www.ulrichsweb.com), Directory of Open Access Journals (DOAJ) (https://doaj.org), Russian Index of Science Citation (https://elibrary.ru), 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.

© RUDN University, 2023

EDITORIAL BOARD

Editor-in-Chief

Yury P. Rybakov, Doctor of Sciences in Physics and Mathematics, Professor, Honored Scientist of Russia, Professor of the Institute of Physical Research & Technologies, RUDN University, Moscow, Russia

Vice Editors-in-Chief

Leonid A. Sevastianov, Doctor of Sciences in Physics and Mathematics, Professor, Professor of the Department of Computational Mathematics and Artificial Intelligence, RUDN University, Moscow, Russia

Dmitry S. Kulyabov, Doctor of Sciences in Physics and Mathematics, Docent, Professor of the Department of Probability Theory and Cyber Security, RUDN University, Moscow, Russia

Members of the editorial board

Konstantin E. Samouylov, Doctor of Sciences in Technical Sciences, Professor, Head of Department of Probability Theory and Cyber Security, RUDN University, Moscow, Russia
Yulia V. Gaidamaka, Doctor of Sciences in Physics and Mathematics, Professor, Professor of the Department of Probability Theory and Cyber Security, RUDN University, Moscow, Russia
Gleb Beliakov, PhD, Professor of Mathematics at Deakin University, Melbourne, Australia
Michal Hnatič, DrSc, Professor of Pavol Jozef Safarik University in Košice, Košice, Slovakia
Datta Gupta Subhashish, PhD in Physics and Mathematics, Professor of Hyderabad University, Hyderabad, India
Olli Erkki Martikainen, PhD in Engineering, member of the Research Institute of the Finnish Economy, Helsinki, Finland
Mikhail V. Medvedev, Doctor of Sciences in Physics and Mathematics, Professor of the Kansas University, Lawrence, USA
Raphael Orlando Ramírez Inostroza, PhD, Professor of Rovira i Virgili University (Universitat Rovira i Virgili), Tarragona, Spain
Bijan Saha, Doctor of Sciences in Physics and Mathematics, Leading Researcher in Laboratory of Information Technologies of the Joint Institute for Nuclear Research, Dubna, Russia
Ochbadrah Chuluunbaatar, Doctor of Sciences in Physics and Mathematics, Leading Researcher in the Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Mongolia

---

Computer Design: *Anna V. Korolkova*, *Dmitry S. Kulyabov*

English text editors: *Nikolay E. Nikolaev*, *Ivan S. Zaryadov*, *Konstantin P. Lovetskiy*

Address of editorial board:
3 Ordzhonikidze St., 115419 Moscow, Russia
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 70×108/16. Offset paper. Offset printing. Typeface "Computer Modern".
Conventional printed sheet 10,35. Printing run 500 copies. Open price. The order 1602.
PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA NAMED AFTER PATRICE LUMUMBA
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
ISSN 2658-7149 (online), 2658-4670 (print)
2023, 31 (4) 304
http://journals.rudn.ru/miph

Contents

© RUDN University, 2023

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 305–331
http://journals.rudn.ru/miph

Review

UDC 519.71:519.872
DOI: 10.22363/2658-4670-2023-31-4-305-331
EDN: FPQLOH

Chronology of the development of Active Queue Management algorithms of RED family. Part 1: from 1993 up to 2005

Ivan S. Zaryadov[1,2], **Hilquias C.C. Viana**[1], **Anna V. Korolkova**[1], **Tatiana A. Milovanova**[1]

[1] *RUDN University,*
6 Miklukho-Maklaya St, Moscow, 117198, Russian Federation
[2] *Federal Research Center "Computer Science and Control" of the RAS,*
44 Vavilova St., bldg. 2, Moscow, 119333, Russian Federation

(received: September 13, 2023; revised: October 25, 2023; accepted: December 29, 2023)

Abstract. This work is the first part of a large bibliographic review of active queue management algorithms of the Random Early Detection (RED) family, presented in the scientific press from 1993 to 2023. The first part will provide data on algorithms published from 1993 to 2005.

Key words and phrases: active queue management, AQM, random early detection, RED, congestion control

1. Introduction

This work is a brief bibliographic review of algorithms of the Random Early Detection (RED) family, compiled according to the dates of publication of scientific works (articles and conference proceedings) in which the algorithms in question were presented to the public.

The authors do not claim that the prepared review includes all existing algorithms, but is the most complete of those published previously, since it includes bibliographic data on 240 algorithms.

Let's briefly talk about other reviews that were published earlier, in which not only algorithms of the RED family were considered, but also other algorithms for active queue management.

The first of these works can be considered the article [1] published in 1995, although the term "active queue management (AQM) algorithms" was not used in it. Instead the term "congestion control algorithms" was used.

In the survey [1] the packet dropping policies for asynchronous transfer mode (ATM) and IP networks were discussed and compared in terms of fairness.

© Zaryadov I. S., Viana H. C.C., Korolkova A. V., Milovanova T. A., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

Another work [2], dated 1999, was the dissertation for the degree of Doctor of Philosophy, in which existing AQM algorithms (and in particular algorithms of the RED family) were analyzed, and ideas were formulated that allowed the development of new algorithms.

The following survey [3], published in 2003, was devoted to the aspects of congestion control with the emphasis on the active queue management. The such AQM problems as parameter tuning, insensitivity to input traffic load variation, mismatch between macroscopic and microscopic queue length behaviour and their implications were summarized and discussed. The attention was also paid to topics that are still relevant and open today: fairness, convergence and implementation complexity, interoperability and robustness, stability, assumptions of network dynamics and link characteristics.

In the 2004 survey [4] the AQM algorithms for responsive and unresponsive TCP flows and aggressive UDP flows were discussed and compared based on the fairness criterion. Also the classification, based on this criterion, for AQM schemes was proposed.

In the work [5], published in 2010, one of the authors of this review proposed the classification of RED algorithms according to different criteria (for example, the type of probability drop function, the type of queue function).

In the survey [6] published in 2013, the author tried to plot the development trajectory of active control algorithms from the first Random Early Detection (RED) algorithm in 1993 to the algorithms presented to the general public in 2011. The algorithms were classified according to various criteria and the general attributes of AQM schemes as well as the design approaches (for example, heuristic, control-theoretic and deterministic optimization) were presented.

In 2016 the comprehensive review of fairness-driven queue management algorithms was presented in [7] with a new taxonomy of categorizing fairness-driven queue management algorithms. The design approaches and key attributes were discussed, compared and analyzed.

Among the works of the last three years in which various algorithms are analyzed and compared, it is worth mentioning the following [8–10].

The review is structured as follows. The structure of the work is as follows. Each section is dedicated to one year, and it presents algorithms of the RED family, scientific publications (articles in scientific journals, conference proceedings, technical reports, etc.) on which were presented this year. In Section 12 the authors discussed the results and the future research directions are highlighted.

2. 1993

This year, the work [11] was published in which the classical Random Early detection (RED) algorithm was presented.

The classic RED (random early detection or random early discard or random early drop) is a queueing discipline with two thresholds ($Q_{\min}$ and $Q_{\max}$) and a low-pass filter to calculate the average queue size $\hat{Q}$ [11]:

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

where w_q, $0 < w_q < 1$ is a weight coefficient of the exponentially weighted moving-average and determines the time constant of the low-pass filter. As said in [11] RED monitors the average queue size and drops (or marks when used in conjunction with ECN) packets based on statistical probabilities $p(\hat{Q})$:

$$p(\hat{Q}) = \begin{cases} 0, & 0 \leqslant \hat{Q} < Q_{\min}, \\ \dfrac{\hat{Q} - Q_{\min}}{Q_{\max} - Q_{\min}} p_{\max}, & Q_{\min} \leqslant \hat{Q} < Q_{\max}, \\ 1, & \hat{Q} \geqslant Q_{\max}, \end{cases} \tag{2}$$

$p_{\max}$ — the fixed maximum value of drop (marking) probability if the threshold $Q_{\max}$ is overcome.

Analysis and criticism of proposed AQM algorithm are presented in the works [12–18].

Suggestions for tuning and optimizing the key parameters of the algorithm are proposed in the following works [19–30]. The implementation of RED in the Next Generation Passive Optical Network (NG-PON) was presented in [31].

Further modifications of the RED algorithm consisted, as a rule, either in changing the number and/or value of thresholds, or in changing the type of drop function (a single linear function was replaced by several linear or nonlinear ones, or combinations of linear and nonlinear functions), or in replacing the average queue size $\hat{Q}$ by the current (instant) queue size q, either in the simultaneous use of the average $\hat{Q}$ and current q queue lengths, or in the dynamic change of one or several parameters (threshold values $Q_{\min}$ and $Q_{\max}$, maximum drop probability $p_{\max}$) depending on control parameters (queue size, incoming rate, rate of queue size change), or in the use of methods of fuzzy logic, Q-learning, neural networks to determine the optimal algorithm parameter values.

The changes also affected whether the new algorithm was being developed to manage a single incoming traffic flow or multiple incoming flows with different priorities.

3. 1997

The development of RED for several flows — Fair RED [32] or Flow RED [33] was introduced in 1997. It uses per-active-flow accounting to impose on each flow a loss rate that depends on the flow's buffer use.

The idea of adaptive active queue management algorithms which may reduce loss rates for congested links was formulated in [34] (as Adaptive RED) and further developed in [35]. The main idea was to adapt $p_{\max}$ of RED in order to keep the average queue size between thresholds $Q_{\min}$ and $Q_{\max}$.

4. 1998

The first Recommendations on Queue Management and Congestion Avoidance in the Internet was given in [36].

RIO (RED with In|Out) algorithm was introduced in [37]. The core of the idea was to monitor the traffic of each user as it enters the network and tag packets as either *in* or *out* based on their service allocation profiles, then at each congested router, preferentially drop packets that are marked as being *out*. RIO used the same drop mechanism as in RED but is configured with two sets of parameters ($p_{\max}$, $Q_{\min}$ and $Q_{\max}$), one for *in* packets and one for *out* packets. Among the other works devoted to modeling, performance analysis and optimal tuning of the algorithm parameters, the following scientific publications [38–41] should be mentioned.

The modifications of RIO algorithm — RIO-C (RED with In|Out and Coupled Virtual Queues) and RIO-DC (RED with In|Out and Decoupled Virtual Queues) are described in [42] and the comparison of these algorithms with the algorithm WRED (1999) in terms of drop of packets is given in [43].

Fair-buffering random early detection (FB-RED) algorithm was presented in [44] to solve the problem of unfairness among links. Although FB-RED results in fairness among links, it however needs to track the information for all the links.

The Fair RED algorithm, which relies on usage of buffer spaces by the different flows (per-active flow accounting) to determine the drop rate of the each flow, was presented in [45]. Although it achieves a fair drop rate for different flows, it needs to track the state of each flow which results in scalability problems similar to those in [44].

5. 1999

Weighted Random Early Detection (WRED) algorithm was introduced by Cisco and the specification can be seen in [46] (or in Cisco IOS Quality of Service Solutions Configuration Guide, Release 12.2 [47]) and was one of the predominant AQM scheme implemented. The algorithm WRED is an extension to RED and was designed to handle traffic of various priorities. In this algorithm, for each type of traffic, its own sets (coinciding or not coinciding) of control parameters are specified (threshold values $Q_{\min}$ and $Q_{\max}$, parameter $p_{\max}$)). The modifications of WRED are Distributed WRED (DWRED), which is the Cisco high-speed version of WRED [47], and Flow-Based WRED (forces WRED to afford greater fairness to all flows on an interface in regard to how packets are dropped) [47]. The other works on WRED are [43, 48–52].

In [53] two versions of RIO algorithm were introduced: (r,RTT)-adaptive RIO algorithm and dynamic RIO (DRIO) algorithm. In [54] DRIO was applied to aggregated traffic instead of the individual flows as in [53].

In [55] to solve the scalability problem of FB-RED [44] and FRED [45] the Stabilized RED (SRED) was introduced. SRED like RED discards packets with a load-dependent probability when a buffer in a router seems congested

and stabilizes the buffer occupation at a level independent of the number of active connections by estimating the number of active connections (flows).

In [56, 57] Balanced RED (BRED) algorithm that drops packet preventively in order to actively penalize the non-adaptive traffic that attempts to "steal" buffer space, and therefore bandwidth from the adaptive traffic flows, was presented.

In [58] Class-Based Threshold RED (CBT-RED) algorithm was presented in order to reduce congestion in routers and to protect TCP from all UDP flows while also ensuring acceptable throughput and latency for well-behaved UDP flows. This algorithm sets the $Q_{\min}$ and $Q_{\max}$ thresholds according to the traffic type and its priority.

A RED discard strategy for ATM networks (ATM-RED) was introduced in [59].

In [60] the Refined RED (Re-RED) algorithm was proposed in order to prevent buffer overflow at a gateway, the RED framework was refined in such a way that the gateway can detect a transient congestion in a timely manner and take actions to quench it when the queue is near full.

In [61] the idea of adaptive active queue management algorithms, started in [34], was continued (Self Configuring RED). The dependence of the effectiveness of RED on the appropriate parameterization of the RED queue was shown. It was proved that there were no single set of RED parameters that work well under different congestion scenarios. As a result the authors proposed and experiment some adaptive RED gateways which self-parameterize themselves based on the traffic mix.

In [62] the modification of RED probability drop function was proposed. The linear drop function has been replaced by a parabolic one, so the new algorithm is called Parabolic RED (PRED). This algorithm was implemented in Cisco routers.

6. 2000

The Gentle RED algorithm (GRED), based on ideas from [59], was proposed by Sally Floyd in [63] and the double maximum threshold parameter was introduced in order to overcome the limitations of the RED algorithm. The comparison of tail drop and active queue management RED and GRED algorithms performance for bulk-data and Web-like Internet traffic was conducted in [64]. In [65] the performance GRED (Gentle RED), DRED (Dynamic-RED) [66] and SRED (Stabilized RED) [55] was analyzed. It was clarified how the performance of AQM mechanisms (such steady state performance measures as the average queue length and the packet loss probability) is affected by a setting of control parameters. In [67] the discrete-time queuing model of GRED algorithm was considered. The performance analysis of GRED (as well as other active queue and passive queue management algorithms) for multi-hop wireless relay networks was presented in [68]. The gentle parameter of GRED was reconsidered in [69].

Weighted RED with Thresholds (WRT) as the development of Weighted RED (WRED) [46] was introduced in [70] and compared with classicsl WRED and RED In and Out (RIO) [37] algorithms.

The Gentle RED algorithm with instantaneous queue size (GRED-I) (instead of exponentially weighted average queue size as in GRED) was proposed in [71].

The modification of RIO [37] algorithm called RI+O was presented in [72] to reduce the effect of an inadequacy in the packet differentiation (between high-profile flows and low-profile flows) RIO algorithm used in the Diff-Serv routers.

The congestion algorithm that uses fuzzy logic based control theory (Fuzzy RED) in order to achieve finer tuning for packet discarding behaviours for individual flows and to provide better quality of service to different kinds of traffic was introduced in [73]. The revised version of [73] is [74]. Fuzzy logic based approach for more predictable congestion control implementation within the DiffServ architecture was presented in [75, 76].

A Rate Based RED Mechanism (Rb-RED) that reduce the number of RED parameters to only one was introduced in [77]. The basic idea of this algorithm is that the packet drop probability was defined as a function of the long-term average arrival rate.

The Double Slope RED (DSRED) as the active queue management scheme for next generation networks (homogeneous TCP/IP networks) was described in [78]. It was proposed to use the two segment drop function and three thresholds instead of the single segment linear drop function with two thresholds as in RED [11] and dynamically change the slope of the packet drop probability curve based on the level of congestion in the buffer. The case of heterogeneous networks and DSRED was considered in [79]. The other works on DSRED are [80] and [81], where the influence of the way packets were chosen to be dropped (end of the tail, head of the tail) on the response time was investigated.

Random Early Adaptive Detection RED/ECN (READ) algorithm as an adaptive queue management scheme for maintaining of high throughput and low round-trip delays under dynamic traffic loads was introduced in [82].

7. 2001

The Adaptive RED (ARED) algorithm, based on ideas from [34] and [61], was proposed by Sally Floyd, Ramakrishna Gummadi and Scott Shenker in [35]. Some algorithmic modifications, while leaving intact the basic Feng idea of $p_{\max}$ adaption in order to keep the average queue size between thresholds $Q_{\min}$ and $Q_{\max}$, were made. The proposed version of RED algorithm, according to the authors, removed the sensitivity to parameters that affect RED's performance and could reliably a specified target average queue length in a wide variety of traffic scenarios. The main differences from the proposed algorithm were the following:

1) $p_{\max}$ was adapted to keep the average queue size within a target range half way between $Q_{\min}$ and $Q_{\max}$;
2) $p_{\max}$ was adapted slowly, over time scales greater than a typical round-trip time, and in small steps;
3) $p_{\max}$ was constrained to be within the range $[0.01; 0.5]$;

4) for $p_{\max}$ an additive-increase multiplicative-decrease (AIMD) policy was used instead of multiplicative-increase and multiplicative-decrease (MIMD) policy.

The tuning of ARED parameters was considered in [83–85]. The ARED comparison with other algorithms was presented in [86–88].

Dynamic RED (DRED) algorithm, which randomly discards packets with a load-dependent probability when a buffer in a router gets congested so a router queue occupancy is stabilised at a level independent of the number of active TCP connections, was introduced in [66]. The comparison of the DRED algorithm with GRED [63] and [55] was carried out in the work [89]. The comparison of DRED algorithm with other algorithms that do not belong to the RED family is presented in [90]. The analytical discrete-time queuing models of DRED algorithm were developed in [91, 92].

The Modified RED (MRED) algorithm computing the packet drop probability based on the heuristic method was presented in [93]. MRED controls queue by using packet loss information and link utilization history information with small queue size. The simulation results presented by authors proved MRED ability to improve fairness, throughput and delay.

The Least Recently Used Cache RED (LRU-RED) algorithm introduced in [94] empowers the routers to contain high bandwidth flows at the time of congestion. Also this algorithm lowers the drop probabilities of short-lived flows and also of responsive high bandwidth flows.

In [95] the new version of RED algorithm — RED-PD (Random Early Detection-Preferential Dropping) was proposed. This algorithm controls the throughput of the high-bandwidth flows by using the packet drop history at the router in order to detect these flows in times of congestion and preferentially drop packets from them.

8. 2002

Rate-based RIO (Rb-RIO) algorithm [96] is an extension of Rate-based RED [77] for traffic with different priority classes (MPEG video stream), so the main idea of RIO [37] was used. The proposed algorithm was compared with Drop-Tail and RED in terms of transport layer throughput, system fairness, application layer throughput and video stream quality.

Extended drop slope random early detection (ExRED) [97] was proposed in order to overcome such RED mechanism problems as low throughput achievement and high number of consecutive drop. The main idea was to modify the drop probability function as a second order polynomial function of the average queue size $\hat{Q}$ in order to keep packet drop rate increasing smoothly but continue with a higher rate when the queue size is more closed to the limit of buffer size ($\hat{Q} \geqslant Q_{\max}$).

In order to improve fairness in high-speed networks the EASY RED algorithm was developed in [98]. It was proposed to use the instantaneous queue size and the single threshold $Q_{\min}$ instead of average queue size and two thresholds ($Q_{\min}$ and $Q_{\max}$) as in RED [11], also the drop probability was defined as a constant when the instantaneous queue length is greater or equal to $Q_{\min}$.

Multi-class RED (MRED) algorithm for several classes of traffic (each traffic class may comprise a number of flows) was introduced in [99]. For each class of traffic, its own set of RED parameters (maximum drop probability and the minimum and maximum thresholds) were specified, which can either coincide or differ.

In [100] for Active RED (ARED, AcRED) algorithm it was proposed to use the heuristic method instead of static one for parameters setting.

In order to improve overall QoS support at the router by satisfying the average performance requirements of incoming packets in terms of throughput and delay in [101] the extension of Adaptive RED (ARED) [35] called RED-Worcester was introduced. The RED-Worcester algorithm based on queuing delays provides a moving target queue size instead of fixed target queue size in ARED, so, when incoming traffic is mostly throughput-sensitive, RED-Worcester tries to maintain a higher average queue to improve the overall throughput, or, when incoming traffic is mostly delay-sensitive, RED-Worcester tries to lower the average queue size to reduce the average queuing delays.

9. 2003

In [102] the new version of Adaptive RED [34, 35], also called Adaptive RED (A-RED), was introduced. It was proposed to adaptive vary not only the maximum packet drop probability $p_{\max}$ (as in [34, 35]), but also the weight coefficient of the exponentially weighted moving-average w_q. As a result the improvement in terms of packet loss rates and queue stability without adversely affecting the link utilization was achieved.

In [103] the flow-based congestion control scheme, called RED with dual-fairness metrics (DRED), was proposed in order to dissolve the unfairness per flow and so provide a feasible QoS. DRED explicitly considers both the instantaneous (the amount of network resources that each flow occupies at the considered time) and the historical (the amount of network resources that each flow has consumed up until the considered point of time) use of network resources for the purpose of dissolving unfairness per flow within the same class, and thereby improving the throughput of each flow.

In [104] the new version of Adaptive RED [35] scheme — Proportional derivative RED controller (PD-RED), based on the proportional derivative (PD) control principle, was introduced.

In [105] the new adaptive fuzzy-based control RED algorithm (AFRED) was designed. This algorithm computes the packet drop probability according to pre-configured fuzzy logic by using the instantaneous queue size as input variable. The ability to dynamically readjust the fuzzy rule in order to make AFRED itself extensively stable for many dynamic environments was also introduced.

The new version of RIO (RED with In|Out) [37] with the ability of self-configuring out-of-profile thresholds, called Adaptive-RIO (A-RIO), was developed in [106]. The main objective of Adaptive-RIO was to increase best-effort throughput by utilizing the available buffer spaces of the core routers in DiffServ networks.

The combination of the Adaptive RED (A-RED) algorithm [35] and the RIO-C algorithm [42], suitable for building an Assured Forwarding (AF) per-hop behaviour (PHB), was proposed in [107] for solving the following tasks: the simplification of the configuration of DiffServ-enabled routers by alleviating the parameter settings problem; the automatic translation a delay parameter into a set of router parameters, the stabilization of the queue occupation around a target value under heavy network load, irrespective of traffic profile. The further study of the proposed algorithm was carried out in [108].

10. 2004

Another version of the Adaptive RED algorithm [35], working with joint co-operation between sources and network routers and therefore called Dynamically Adaptive RED (DARED), was presented in the work [109]. The aim of this algorithm was to guarantee the distribution of the available network resources to flows of different classes with different declared QoS requirements.

The Modified Random Early Detection (MRED) algorithm, designed to provide better control over the burstiness level, was presented in [110]. The proposed modification was that the drop function (probability) takes into account not only the average queue size $\hat{Q}$, but also the instantaneous queue size Q (the incoming packet is dropped if $\hat{Q} > Q_{\max}$ and $Q > Q_{\max}$).

The algorithm, called Priority Random Early Detection (PRED), was described in [111] for different priority types of traffic, and for each type of traffic, its own set of control parameters is set, which can dynamically change values depending on network load.

The new AQM scheme, Short-lived Flow Friendly RED (SHRED), targeted at providing better network performance for short-lived Web traffic, was presented and analyzed in [112]. Using an edge hint to indicate the congestion window size in each packet sent by the flow source or by an edge router, SHRED preferentially drops packets from short-lived Web flows (flows with small TCP windows) with a lower probability than packets from long-lived flows (flows with large TCP windows). Thus SHRED protects short-lived flows from low transmission rates, and provides fairer bandwidth allocation among flows.

The development of Flow RED algorithm [33] — RED with Dynamic Thresholds (RED-DT), was proposed in [113]. This algorithm dynamically adapts queue parameters to achieve a more fair distribution of the link capacity. In order to identify unresponsive and greedy flows, RED-DT maintains per-flow state for active flows. Flow is considered to be active if it has at least one packet in the queue. For each active flow, there is an entry in a flow table that contains the instantaneous queue size Q_i, average queue size $\hat{Q}_i$ and maximum drop probability $p^i_{\max}$. Similar to RED, RED-DT maintains minimum and maximum thresholds $Q_{\min}$ and $Q_{\max}$, but these thresholds are dynamically changed upon each packet arrival.

In [114] the modification of the classic Random Early Detection (RED) algorithm [11] with hyperbolic drop function instead of linear one (RED) was introduced.

In [115, 116] two versions of Adaptive RIO algorithm [107] were presented: ARIO-D (Adaptive RIO for Delay) and ARIO-L (Adaptive RIO for Loss). The first algorithm (ARIO-D) takes specific average queue length range as control target, keeps $Q_{\max}$ as constant and adaptively adjusts $p_{\max}$ for *out*-packets to meet the expected steady state. The second algorithm (ARIO-L) takes specific loss ratio range as control target, for *out*-packets keeps $p_{\max}$ as constant and adaptively adjusts $Q_{\max}$ to meet the expected steady state.

The new RED scheme, called Loss Ratio Based RED (LRED), which measures not only the queue size, but also the latest packet loss ratio, and uses both these parameters in order to dynamically adjust packet drop probability, was presented in [117]. The further development of this algorithm is presented in [118].

The Class-Guided RED (CGRED) algorithm was introduced in [119]. In CGRED the packet drop operation is developed into a class-differential one by introducing a couple of class-specific adjusting decisions or guidelines: the adjustment direction (whether to increase or decrease or just maintain the RED drop probability) and the guiding probability (the intensity of the given adjustment). Thus the RED packet drop probability calculated is considered as a reference value and the guiding probability is considered as a limiting bound for choosing the actual dropping probability.

The Proxy-RED algorithm [120] was proposed as a solution for reducing the AQM overhead from the access point and as a development of ARED [34, 35]. The average queue size in Proxy-RED is calculated periodically but not at the moment of each packet arrival as in RED [11] or ARED [34, 35]. Also the modifications for drop probability function were made. The development of Proxy RED was presented in [121].

11. 2005

In [122] the Exponential-RED (E-RED) AQM (Active Queue Management) algorithm, as a decentralized network congestion control algorithm with dynamic adaptations at both user ends and link ends, was presented. In this algorithm the packet dropping probability was set as the exponential function of the virtual queue length and the capacity of the virtual queue was slightly smaller than the link capacity. E-RED was the first AQM schema for which the ability to stabilize TCP-Reno for a general topology network with heterogeneous delays has been proven. For a TCP-Reno network with Exponential-RED control, a discrete-time dynamical feedback system model with delay was studied in [123].

In order to maximize the throughput and to minimize the packet drop and delay the new algorithm — Adaptive RED with Dynamic Threshold Adjustment (ARDTA) was introduced in [124]. For this algorithm the thresholds were dynamically modified by using an exact expression of average queue size for a given burst size and number of nodes. The minimum threshold $Q_{\min}$ was set by an expression for a given burst size, the maximum thresholds $Q_{\max}$ was changed dynamically based on traffic conditions and buffer size with also taking into account the burst size. The assumption was made that the maximum threshold $Q_{\max}$ would be reached when the instantaneous queue size was equal to the maximum buffer size. The other work on this algorithm is [125].

Piecewise Linear RED (PL-RED) [126] is the modification of Gentle RED (GRED) [63]. In this algorithm the drop probability function was a piecewise linear function with $N = 5$ segments.

The another variant of Gentle RED (GRED) [63] — Adaptive Exponential RED (AERED), was also presented in [126], where the drop probability function was an exponential function with parameter β depending on average queue size and thresholds and p (the amount of concavity of an exponential function).

The dynamic modification of Weighted RED (WRED) [46, 47] was proposed in [127] and was called as Dynamic Weighted RED (DWRED). For this algorithm the TCP Window-Aware Marker (TWAM) was introduced for distribution of the resources available for the total traffic of an AF FEC (Assured Forwarding Forward Equivalence class) among the individual AF FEC flows in a fair manner. Based on TWAM the thresholds and $p_{\max}$ are dynamically configured. So the proposed WRED configuration mechanism responds to fluctuations in available resources, allowing the use of excessive resources whenever they are available, in a way that achieves a bounded average queuing delay for packets.

The Revised version of Adaptive RED [35] — RARED, was described in [128]. To alleviate the effect of w_q the RARED takes the input rate besides queue occupancy into account to detect significant changes in the network's condition.

The multi-class signaling overload control algorithm (Signaling RED — SiRED) for telecommunication switches as a modified version of WRED [46, 47] was proposed in [129]. SiRED measures the system load using queue lengths.

The modification of RED [11], based on game theory, for Internet switching with selfish users was presented in [130]. The new algorithm was called as Preemptive RED (PRED). The main feature of PRED is extra drop mechanism that drops an additional packet of the same user from the buffer when its packet is dropped by RED in order to penalize users that do not respond to congestion signals.

Subsidized RED (SubRED) algorithm, designed for short-lived (fragile) flows (e.g. most HTTP flows) in order to keep the link utilization high while reducing the average flow response time, was proposed in [131]. Subsidized RED (SubRED) identifies short-lived flows that have recently lost packets and/or are in their slow start phases, and protects them from being further punished unnecessarily for a short duration of time (flow subsidy).

The burst-sensitive RED algorithm for GPRS links in a heterogeneous mobile environment was presented in [132] and was called Burst-sensitive RED (BSRED). The new parameter to the classic RED [11] was added — Burst Threshold B_t, which represents the threshold of the number of consecutive packets en-queued (B_p) without a packet de-queued. Thus the new rule was added to the RED algorithm: on each packet's arrival, if $B_p \geqslant B_t$, a packet will be dropped, irrespective of the current queue length. In addition, B_p is reset to zero only when there is a packet de-queued. That is, incoming packets are continually dropped if there is no packet de-queued once $B_p \geqslant B_t$.

The nonlinear version of ARED [35] with nonlinear power packet dropping probability function was proposed in [133] and called as POWer Adaptive

Random Early Detection (POWARED). In order to enable POWARED to cater for dynamics of bursty network traffic and intelligently differentiate between levels of congestion occurred, the decrement or increment adjustment was based on the ratio of deviation between current average queue size and target queue size (steady-state queue size).

The virtual queue management approach, named Virtual Queue RED (VQ-RED, VQRED) to address the fairness problems (downlink/uplink fairness and fairness among flows in the same direction) was introduced in [134]. VQ-RED treats all the competing flows (uplink flows and downlink flows) fairly through managing their corresponding virtual queues. It punishes the arbitrary flows and gives more benefits to the weak flows. In this way it guarantees the fairness among the flows.

12. Conclusions

The presented bibliographical chronological review of active control algorithms of the RED family is the most complete both in terms of the number of algorithms reviewed (more than two hundred) and in terms of the number of scientific publications analyzed and presented. This review will be useful to researchers in the field of the congestion control.

Active queue management algorithms of the RED family are not something new for the authors of this work, as evidenced by the publications presented below [5, 135–142].

In the future, the authors plan not only to classify the considered algorithms based on the classification criteria presented in [5–7], but also to review and classify other active queue management algorithms.

References

[1] C.-Q. Yang and A. V. S. Reddy, "A taxonomy for congestion control algorithms in packet switching networks," *IEEE Network*, vol. 9, no. 4, pp. 34–45, 1995. DOI: 10.1109/65.397042.

[2] W.-C. Feng, "Improving Internet Congestion Control and Queue Management Algorithms. Doctor of Philosophy dissertation," The University of Michigan, Tech. Rep., 1999.

[3] S. Ryu, C. Rump, and C. Qiao, "Advances in internet congestion control," *IEEE Communications Surveys & Tutorials*, vol. 5, no. 1, pp. 28–39, 2003. DOI: 10.1109/COMST.2003.5342228.

[4] G. Chatranon, M. A. Labrador, and S. Banerjee, "A survey of TCP-friendly router-based AQM schemes," *Computer Communications*, vol. 27, no. 15, pp. 1424–1440, 2004. DOI: 10.1016/j.comcom.2004.05.001.

[5] A. V. Korolkova, D. S. Kulyabov, and A. I. Tchernoivanov, "On the classification of RED algorithms," Russian, *RUDN Journal of Mathematics, Information Sciences and Physics*, vol. 3, pp. 34–46, 2009.

[6] R. Adams, "Active queue management: a survey," *Communications Surveys & Tutorials, IEEE*, vol. 15, pp. 1425–1476, 2013. DOI: 10.1109/SURV.2012.082212.00018.

[7] G. Abbas, Z. Halim, and Z. H. Abbas, "Fairness-driven queue management: a survey and taxonomy," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 324–367, 2016. DOI: 10.1109/COMST.2015.2463121.

[8] D. D. Zala and A. K. Vyas, "Comparative analysis of RED queue variants for data traffic reduction over wireless network," in *Recent Advances in Communication Infrastructure*, A. Mehta, A. Rawat, and P. Chauhan, Eds., ser. Lecture Notes in Electrical Engineering, vol. 618, Singapore: Springer Singapore, 2020, pp. 31–39. DOI: 10.1007/978-981-15-0974-2_3.

[9] S. Sunassee, A. Mungur, S. Armoogum, and S. Pudaruth, "A comprehensive review on congestion control techniques in networking," in *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India: IEEE, 2021, pp. 305–312. DOI: 10.1109/ICCMC51019.2021.9418329.

[10] A. A. Mahawish and H. Hassan, "Survey on: A variety of AQM algorithm schemas and intelligent techniques developed for congestion control," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 3, pp. 1419–1431, Sep. 2021. DOI: 10.11591/ijeecs.v23.i3.pp1419-1431.

[11] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," *IEEE/ACM Transactions on Networking*, vol. 1, no. 4, pp. 397–413, 1993. DOI: 10.1109/90.251892.

[12] M. May, J. Bolot, C. Diot, and B. Lyles, "Reasons not to deploy RED," in *1999 Seventh International Workshop on Quality of Service. IWQoS'99. (Cat. No.98EX354)*, 1999, pp. 260–262. DOI: 10.1109/IWQOS.1999.766502.

[13] T. Bonald, M. May, and J.-C. Bolot, "Analytic evaluation of RED performance," in *Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies (Cat. No.00CH37064)*, vol. 3, IEEE, 2000, pp. 1415–1424. DOI: 10.1109/INFCOM.2000.832539.

[14] S. De Cnodder, O. Elloumi, and K. Pauwels, "RED behavior with different packet sizes," in *Proceedings ISCC 2000. Fifth IEEE Symposium on Computers and Communications*, IEEE, 2000, pp. 793–799. DOI: 10.1109/ISCC.2000.860741.

[15] M. Christiansen, K. Jeffay, D. Ott, and F. D. Smith, "Tuning RED for Web traffic," *IEEE/ACM Transactions on Networking*, vol. 9, no. 3, pp. 249–264, 2001. DOI: 10.1109/90.929849.

[16] C. Hollot, V. Misra, D. Towsley, and W.-B. Gong, "A control theoretic analysis of RED," in *Proceedings IEEE INFOCOM 2001. Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society (Cat. No.01CH37213)*, vol. 3, IEEE, 2001, pp. 1510–1519. DOI: 10.1109/INFCOM.2001.916647.

[17] C. Brandauer, G. Iannaccone, C. Diot, T. Ziegler, S. Fdida, and M. May, "Comparison of tail drop and active queue management performance for bulk-data and Web-like Internet traffic," in *Proceedings. Sixth IEEE Symposium on Computers and Communications*, IEEE, 2001, pp. 122–129. DOI: 10.1109/ISCC.2001.935364.

[18] C. Joo and S. Bahk, "Scalability problems of RED," *Electronics Letters*, vol. 38, no. 21, pp. 1297–1298, 2002. DOI: 10.1049/el:20020744.

[19] R. Vaidya and S. Bhatnagar, "Robust optimization of Random Early Detection," *Telecommunication Systems*, vol. 33, pp. 291–316, Dec. 2006. DOI: 10.1007/s11235-006-9020-2.

[20] L. Tan, W. Zhang, G. Peng, and G. Chen, "Stability of TCP/RED systems in AQM routers," *IEEE Transactions on Automatic Control*, vol. 51, no. 8, pp. 1393–1398, 2006. DOI: 10.1109/TAC.2006.876802.

[21] B. Zheng and M. Atiquzzaman, "A framework to determine bounds of maximum loss rate parameter of RED queue for next generation routers," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 429–445, 2008. DOI: https://doi.org/10.1016/j.jnca.2008.02.003.

[22] B. Zheng and M. Atiquzzaman, "A framework to determine the optimal weight parameter of RED in next-generation Internet routers," *International Journal of Communication Systems*, vol. 21, no. 9, pp. 987–1008, 2008. DOI: https://doi.org/10.1002/dac.932.

[23] G. Min and X. Jin, "Performance Modelling of Random Early Detection Based Congestion Control for Multi-Class Self-Similar Network Traffic," in *2008 IEEE International Conference on Communications*, IEEE, 2008, pp. 5564–5568. DOI: 10.1109/ICC.2008.1043.

[24] X. Chen, S.-C. Wong, and C. K. Tse, "Adding Randomness to Modeling Internet TCP-RED Systems with Interactive Gateways," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 57, no. 4, pp. 300–304, 2010. DOI: 10.1109/TCSII.2010.2043388.

[25] S. Woo and K. Kim, "Tight Upper Bound for Stability of TCP/RED Systems in AQM Routers," *IEEE Communications Letters*, vol. 14, no. 7, pp. 682–684, 2010. DOI: 10.1109/LCOMM.2010.07.100375.

[26] J. Shahram and R. Z. Seyed, "An Active Queue Management for High Bandwidth-Delay Product Networks," *International Journal of Computer Theory and Engineering*, vol. 5, no. 5, pp. 763–767, 2013. DOI: 10.7763/IJCTE.2013.V5.792.

[27] Hendrawan and P. Hernandia, "Random Early Detection utilizing genetics algorithm," in *2014 8th International Conference on Telecommunication Systems Services and Applications (TSSA)*, Kuta, Bali, Indonesia: IEEE, 2014, pp. 1–7. DOI: 10.1109/TSSA.2014.7065952.

[28] A. Waheed, N. Habib Khan, M. Zareei, S. Ul Isla, L. Jan, A. Iqbal Umar, and M. M. Ehab, "Traffic queuing management in the Internet of Things: an optimized RED algorithm based approach," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 359–372, 2021. DOI: 10.32604/cmc.2020.012196.

[29] N. G. Goudru, "Tuning Pmax in RED Gateways for QoS Enhancement in Wireless Packet Switching Networks," in *Mathematical Modeling, Computational Intelligence Techniques and Renewable Energy*, M. Sahni, J. M. Merigó, R. Sahni, and R. Verma, Eds., Singapore: Springer Singapore, 2022, pp. 321–334. DOI: 10.1007/978-981-16-5952-2_28.

[30] A. Basheer, H. J. Hassan, and G. Muttasher, "Intelligent Parameter Tuning Using Deep Q-Network for RED Algorithm in Adaptive Queue Management Systems," in *Micro-Electronics and Telecommunication Engineering*, D. K. Sharma, S.-L. Peng, R. Sharma, and D. A. Zaitsev, Eds., Singapore: Springer Nature Singapore, 2022, pp. 439–446. DOI: 10.1007/978-981-16-8721-1_42.

[31] X. Xu, B. Liu, L. Zhang, Y. Mao, X. Wu, J. Ren, S. Han, L. Jiang, and X. Xin, "Self-adaptive bandwidth scheduling based on improved Random Early Detection for NG-PON," in *2019 18th International Conference on Optical Communications and Networks (ICOCN)*, Huangshan, China: IEEE, 2019, pp. 1–3. DOI: 10.1109/ICOCN.2019.8934251.

[32] D. Lin and R. Morris, "Dynamics of Random Early Detection," in *Proceedings of the ACM SIGCOMM'97 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communication*, ser. SIGCOMM'97, Cannes, France: Association for Computing Machinery, 1997, pp. 127–137. DOI: 10.1145/263105.263154.

[33] D. Lin and R. Morris, "Dynamics of Random Early Detection," *SIGCOMM Comput. Commun. Rev.*, vol. 27, no. 4, pp. 127–137, 1997. DOI: 10.1145/263109.263154.

[34] W.-C. Feng, D. D. Kandlur, D. Saha, and K. G. Shin, "Techniques for Eliminating Packet Loss in Congested TCP/IP Networks," The University of Michigan, Tech. Rep., 1997.

[35] S. Floyd, R. Gummadi, and S. Shenker, "Adaptive RED: An Algorithm for Increasing the Robustness of RED's Active Queue Management," AT&T Center for Internet Research at ICSI, Tech. Rep., 2001.

[36] L. Zhang, C. Partridge, S. Shenker, *et al.*, "Recommendations on Queue Management and Congestion Avoidance in the Internet," Internet Engineering Task Force, Tech. Rep. 2309, 1998, 17 pp. DOI: 10.17487/RFC2309.

[37] D. Clark and W. Fang, "Explicit allocation of best-effort packet delivery service," *IEEE/ACM Transactions on Networking*, vol. 6, no. 4, pp. 362–373, 1998. DOI: 10.1109/90.720870.

[38] Y. Hori, T. Ikenaga, and Y. Oie, "Queue management of RIO to achieve high throughput and low delay," in *2001 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (IEEE Cat. No.01CH37233)*, vol. 2, IEEE, 2001, 619–622 vol.2. DOI: 10.1109/PACRIM.2001.953709.

[39] N. Malouch and Z. Liu, "Performance analysis of TCP with RIO routers," in *Global Telecommunications Conference, 2002. GLOBECOM'02. IEEE*, vol. 2, IEEE, 2002, 1623–1627 vol.2. DOI: 10.1109/GLOCOM.2002.1188472.

[40] M. Barbera, A. Lombardo, G. Schembra, and C. A. Trecarichi, "A fluid-flow model of RIO routers loaded by Markov modulated fluid processes," in *2004 Workshop on High Performance Switching and Routing, 2004. HPSR*, 2004, pp. 67–71. DOI: 10.1109/HPSR.2004.1303429.

[41] Y.-I. Joo, K. Hur, J.-K. Kim, D.-S. Eom, and Y. Lee, "RIO configuration optimization for assured service in diffserv networks," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 4, pp. 1968–1972, 2009. DOI: 10.1109/TCE.2009.5373757.

[42] N. Seddigh, B. Nandy, P. S. Pieda, J. H. Salim, and A. Chapman, "Experimental study of assured services in a diffserv IP QoS network," in *Internet Routing and Quality of Service*, R. O. Onvural, S. Civanlar, P. J. Doolan, J. V. Luciani, S. Civanlar, P. J. Doolan, and J. V. Luciani, Eds., International Society for Optics and Photonics, vol. 3529, SPIE, 1998, pp. 217–230. DOI: 10.1117/12.333712.

[43] K. Singh and G. Kaur, "Comparison of random early detection (RED) techniques for congestion control in differentiated services networks based on packet dropping," *International Journal of Neural Networks*, vol. 2, no. 1, pp. 25–29, 2012.

[44] W.-J. Kim and B. G. Lee, "The FB-RED algorithm for TCP over ATM," in *IEEE GLOBECOM 1998 (Cat. NO. 98CH36250)*, vol. 1, IEEE, 1998, 551–555 vol.1. DOI: 10.1109/GLOCOM.1998.775788.

[45] W.-J. Kim and G. Byeong, "FRED — fair random early detection algorithm for TCP over ATM networks," *Electronics Letters*, vol. 34, no. 2, pp. 152–154, 1998. DOI: 10.1049/el:19980049.

[46] C. Cisco Systems Inc, *Cisco IOS 12.0 Quality of Service*. USA: Cisco Press, 1999, p. 288.

[47] C. Cisco Systems Inc, "Cisco IOS Quality of Service Solutions Configuration Guide, Release 12.2," Cisco, Tech. Rep., 1999.

[48] M.-D. Cano, F. Cerdán, and P. López-Matencio, "Experimental Study of Bandwidth Assurance in a DiffServ Network," in *Internet and Multimedia Systems and Applications, EuroIMSA*, Grindelwald, Switzerland, 2005, pp. 202–208.

[49] W.-T. Lee, F.-H. Liu, and H.-F. Lo, "Improving the performance of MPEG-4 transmission in IEEE 802.15.3 WPAN," in *2008 8th IEEE International Conference on Computer and Information Technology*, Sydney, NSW, Australia: IEEE, 2008, pp. 676–681. DOI: 10.1109/CIT.2008.4594756.

[50] L. B. Lim, L. Guan, A. Grigg, I. W. Phillips, X. Wang, and I. U. Awan, "RED and WRED Performance Analysis Based on Superposition of N MMBP Arrival Proccess," in *24th IEEE International Conference on Advanced Information Networking and Applications*, Perth, WA, Australia: IEEE, 2010, pp. 66–73. DOI: 10.1109/AINA.2010.85.

[51] R. Jiang, Y. Pan, Y. Liu, and X. Xue, "Simulation Study of RED/WRED Mechanism Based on OPNET," in *International Conference on Mechatronics, Electronic, Industrial and Control Engineering*, Atlantis Press, 2015, pp. 138–141. DOI: 10.2991/meic-15.2015.34.

[52] W. Lafta and W. One, "Performance Evaluation of Heterogeneous Network Based on RED and WRED," *Indonesian Journal of Electrical Engineering and Computer Science*, pp. 540–545, 2016. DOI: 10.11591/ijeecs.v3.i3.pp540-545.

[53] W. Lin, R. Zheng, and J. C. Hou, "How to make assured service more assured," in *Proceedings. Seventh International Conference on Network Protocols*, Toronto, Ontario, Canada: IEEE, 1999, pp. 182–191. DOI: 10.1109/ICNP.1999.801934.

[54] S. Herreria-Alonso, M. Fernandez-Veiga, C. Lopez-Garcia, M. Rodriguez-Perez, and A. Suarez-Gonzalez, "Improving fairness requirements for assured services in a differentiated services network," in *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, vol. 4, Paris, France: IEEE, 2004, pp. 2076–2080. DOI: 10.1109/ICC.2004.1312884.

[55] T. Ott, T. Lakshman, and L. H. Wong, "SRED: stabilized RED," in *IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, vol. 3, New York, NY, USA: IEEE, 1999, pp. 1346–1355. DOI: 10.1109/INFCOM.1999.752153.

[56] F. Anjum and L. Tassiulas, "Fair bandwidth sharing among adaptive and non-adaptive flows in the Internet," in *IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, vol. 3, New York, NY, USA: IEEE, 1999, pp. 1412–1420. DOI: 10.1109/INFCOM.1999.752161.

[57] F. Anjum and L. Tassiulas, "Balanced RED: an algorithm to achieve fairness in the Internet," The Center for Satellite and Hybrid Communication Networks, Tech. Rep., 1999.

[58] M. Parris, K. Jeffay, and F. Donelson Smith, "Lightweight active router-queue management for multimedia networking," in *Multimedia Computing and Networking 1999*, D. D. Kandlur, K. Jeffay, and T. Roscoe, Eds., International Society for Optics and Photonics, vol. 3654, SPIE, 1998, pp. 162–174. DOI: 10.1117/12.333807.

[59] V. Rosolen, O. Bonaventure, and G. Leduc, "A RED Discard Strategy for ATM Networks and Its Performance Evaluation with TCP/IP Traffic," *SIGCOMM Comput. Commun. Rev.*, vol. 29, no. 3, pp. 23–43, 1999. DOI: 10.1145/505724.505728.

[60] H. Wang and K. G. Shin, "Refined design of random early detection gateways," in *Seamless Interconnection for Universal Services. Global Telecommunications Conference. GLOBECOM'99. (Cat. No.99CH37042)*, vol. 1B, Rio de Janeiro, Brazil: IEEE, 1999, pp. 769–775. DOI: 10.1109/GLOCOM.1999.830170.

[61] W.-C. Feng, D. Kandlur, D. Saha, and K. G. Shin, "A self-configuring RED gateway," in *IEEE INFOCOM'99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, vol. 3, New York, NY, USA: IEEE, 1999, pp. 1320–1328. DOI: 10.1109/INFCOM.1999.752150.

[62] V. Jacobson, K. M. Nichols, and K. Poduri, "RED in a Different Light," Cisco Systems: San Jose, CA, USA, Tech. Rep., 1999.

[63] S. Floyd, "Recommendation on using the "gentle variant of RED"," The ICSI Networking ans Security Gropup, Tech. Rep., 2000.

[64] C. Brandauer, G. Iannaccone, C. Diot, T. Ziegler, S. Fdida, and M. May, "Comparison of tail drop and active queue management performance for bulk-data and Web-like Internet traffic," in *Proceedings. Sixth IEEE Symposium on Computers and Communications*, Hammamet, Tunisia: IEEE, 2001, pp. 122–129. DOI: 10.1109/ISCC.2001.935364.

[65] T. Eguchi, H. Ohsaki, and M. Murata, "On control parameters tuning for active queue management mechanisms using multivariate analysis," in *2003 Symposium on Applications and the Internet, 2003. Proceedings*, Orlando, FL, USA: IEEE, 2003, pp. 120–127. DOI: 10.1109/SAINT.2003.1183040.

[66] J. Aweya, M. Ouellette, and D. Y. Montuno, "A control theoretic approach to active queue management," *Computer Networks*, vol. 36, no. 2, pp. 203–235, 2001, Theme issue: Overlay Networks. DOI: 10.1016/S1389-1286(00)00206-1.

[67] H. Abdel-jaber, F. Thabtah, and M. Woodward, "Traffic management for the gentle random early detection using discrete-time queueing," in *International Business Information Management Association (9th IBIMA) Conference. The Conference Proceedings*, Marrakech, Morocco: IBIMA, 2008, pp. 289–298.

[68] R. Hema, G. Murugesan, M. J. A. Jude, V. Diniesh, D. Sree Arthi, and S. Malini, "Active queue versus passive queue — An experimental analysis on multi-hop wireless networks," in *2017 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India: IEEE, 2017, pp. 1–5. DOI: 10.1109/ICCCI.2017.8117774.

[69] N. Hamadneh, M. Alkasassbeh, I. Obeidat, and M. BaniKhalaf, "Revisiting the Gentle Parameter of the Random Early Detection (RED) for TCP Congestion Control," *Journal of Communications*, vol. 14, pp. 229–235, Mar. 2019. DOI: 10.12720/jcm.14.3.229-235.

[70] U. Bodin, O. Schelén, and S. Pink, "Load-tolerant differentiation with active queue management," *Computer Communication Review*, vol. 30, pp. 4–16, Jul. 2000. DOI: 10.1145/382179.382180.

[71] M. May, C. Diot, B. Lyles, and J. Bolot, "Influence of Active Queue Management Parameters on Aggregate Traffic Performance," INRIA, Research Report RR-3995, 2000, p. 21.

[72] W. H. Park, S. Bahk, and H. Kim, "A modified RIO algorithm that alleviates the bandwidth skew problem in Internet differentiated service," in *2000 IEEE International Conference on Communications. ICC 2000. Global Convergence Through Communications. Conference Record*, vol. 3, New Orleans, LA, USA: IEEE, 2000, pp. 1599–1603. DOI: 10.1109/ICC.2000.853765.

[73] R. Loukas, S. Kohler, P. Andreas, and T.-G. Phuoc, "Fuzzy RED: congestion control for TCP/IP Diff-Serv," in *2000 10th Mediterranean Electrotechnical Conference. Information Technology and Electrotechnology for the Mediterranean Countries. Proceedings. MeleCon 2000 (Cat. No.00CH37099)*, vol. 1, Lemesos, Cyprus: IEEE, 2000, pp. 19–22. DOI: 10.1109/MELCON.2000.880358.

[74] L. Rossides, A. Sekercioglu, A. Pitsillides, A. Vasilakos, S. Kohler, and P. Tran-Gia, "Fuzzy RED: Congestion Control for TCP/IP Diff-Serv," in *Advances in Computational Intelligence and Learning: Methods and Applications* (International Series in Intelligent Technologies), International Series in Intelligent Technologies. Dordrecht: Springer Netherlands, 2002, vol. 18, pp. 343–352. DOI: 10.1007/978-94-010-0324-7_24.

[75] C. Chrysostomou, A. Pitsillides, L. Rossides, M. Polycarpou, and A. Sekercioglu, "Congestion control in differentiated services networks using Fuzzy-RED," *Control Engineering Practice*, vol. 11, no. 10, pp. 1153–1170, 2003, 10.1016/S0967-0661(03)00052-2.

[76] C. Chrysostomou, A. Pitsillides, L. Rossides, and A. Sekercioglu, "Fuzzy logic controlled RED: congestion control in TCP/IP differentiated services networks," *Soft Computing*, vol. 18, pp. 79–92, 2003, Special Section on Control Methods for Telecommunication. DOI: 110.1007/s00500-002-0248-9.

[77] S. De Cnodder, K. Pauwels, and O. Elloumi, "A Rate Based RED Mechanism," in *The 10th International Workshop on Network and Operating System Support for Digital Audio and Video*, Chapel Hill, North Carolina, USA, 2000.

[78] B. Zheng and M. Atiquzzaman, "DSRED: an active queue management scheme for next generation networks," in *Proceedings 25th Annual IEEE Conference on Local Computer Networks. LCN 2000*, Tampa, FL, USA: IEEE, 2000, pp. 242–251. DOI: 10.1109/LCN.2000.891036.

[79] B. Zheng and M. Atiquzzaman, "DSRED: improving performance of active queue management over heterogeneous networks," in *ICC 2001. IEEE International Conference on Communications. Conference Record (Cat. No.01CH37240)*, vol. 8, Helsinki, Finland: IEEE, 2001, pp. 2375–2379. DOI: 10.1109/ICC.2001.936557.

[80] B. Zheng and M. Atiquzzaman, "DSRED: A New Queue Management Scheme for the Next Generation Internet," *IEICE Transactions on Communications*, vol. E89B, no. 3, pp. 764–774, 2006. DOI: 10.1093/ietcom/e89-b.3.764.

[81] J. Domańska, A. Domański, and T. Czachórski, "The Drop-From-Front Strategy in AQM," in *Next Generation Teletraffic and Wired/Wireless Advanced Networking*, Y. Koucheryavy, J. Harju, and A. Sayenko, Eds., ser. Lecture Notes in Computer Science, vol. 4712, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 61–72. DOI: 10.1007/978-3-540-74833-5_6.

[82] M. Arpaci and J. A. Copeland, "An adaptive queue management method for congestion avoidance in TCP/IP networks," in *Globecom'00 — IEEE. Global Telecommunications Conference. Conference Record (Cat. No.00CH37137)*, vol. 1, San Francisco, CA, USA: IEEE, 2000, pp. 309–315. DOI: 10.1109/GLOCOM.2000.892022.

[83] L. Yang, J. Zhu, W. Xie, and X. Tan, "Stable tuning for random early detection algorithm," in *Proceedings of the 33rd Chinese Control Conference*, Nanjing, China: IEEE, 2014, pp. 5470–5475. DOI: 10.1109/ChiCC.2014.6895874.

[84] D. Marek, J. Szyguła, A. Domański, J. Domańska, K. Filus, and M. Szczygieł, "Adaptive Hurst-Sensitive Active Queue Management," *Entropy*, vol. 24, no. 3, 2022. DOI: 10.3390/e24030418.

[85] J. Liu and D. Wei, "Active Queue Management Based on Q-Learning Traffic Predictor," in *2022 International Conference on Cyber-Physical Social Intelligence (ICCSI)*, Nanjing, China: IEEE, 2022, pp. 399–404. DOI: 10.1109/ICCSI55536.2022.9970698.

[86] J. Schwardmann, D. Wagner, and M. Kühlewind, "Evaluation of ARED, CoDel and PIE," in *Advances in Communication Networking*, Y. Kermarrec, Ed., ser. Lecture Notes in Computer Science, vol. 8846, Cham: Springer International Publishing, 2014, pp. 185–191. DOI: 10.1007/978-3-319-13488-8_17.

[87] C. A. Grazia, N. Patriciello, M. Klapez, and M. Casoni, "Which AQM fits IoT better?" In *2017 IEEE 3rd International Forum on Research and Technologies for Society and Industry (RTSI)*, Modena, Italy: IEEE, 2017, pp. 1–6. DOI: 10.1109/RTSI.2017.8065903.

[88] W. G. de Morais, C. E. M. Santos, and C. M. Pedroso, "Application of active queue management for real-time adaptive video streaming," *Telecommunication Systems*, vol. 79, pp. 261–270, 2022. DOI: 10.1007/s11235-021-00848-0.

[89] T. Eguchi, H. Ohsaki, and M. Murata, "On control parameters tuning for active queue management mechanisms using multivariate analysis," in *2003 Symposium on Applications and the Internet, 2003. Proceedings*, Orlando, FL, USA: IEEE, 2003, pp. 120–127. DOI: 10.1109/SAINT.2003.1183040.

[90] E.-C. Park, H. Lim, K.-J. Park, and C.-H. Choi, "Analysis and design of the virtual rate control algorithm for stabilizing queues in TCP networks," *Computer Networks*, vol. 44, no. 1, pp. 17–41, 2004. DOI: 10.1016/S1389-1286(03)00321-9.

[91] H. Abdel-Jaber, M. Woodward, F. Thabtah, and A. Abu-Ali, "Performance evaluation for DRED discrete-time queueing network analytical model," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 750–770, 2008. DOI: 10.1016/j.jnca.2007.09.003.

[92] J. Ababneh, H. Abdel-Jaber, F. Thabtah, W. Hadi, and E. Badarneh, "Derivation of Three Queue Nodes Discrete-Time Analytical Model Based on DRED Algorithm," in *2010 Seventh International Conference on Information Technology: New Generations*, Las Vegas, NV, USA: IEEE, 2010, pp. 885–890. DOI: 10.1109/ITNG.2010.252.

[93] J. Koo, B. Song, K. Chung, H. Lee, and H. Kahng, "MRED: a new approach to random early detection," in *Proceedings 15th International Conference on Information Networking*, Beppu, Japan: IEEE, 2001, pp. 347–352. DOI: 10.1109/ICOIN.2001.905450.

[94] S. A. L. N. Reddy, "LRU-RED: an active queue management scheme to contain high bandwidth flows at congested routers," in *GLOBECOM'01. IEEE Global Telecommunications Conference*, vol. 4, San Antonio, TX, USA: IEEE, 2001, pp. 2311–2315. DOI: 10.1109/GLOCOM.2001.966191.

[95] R. Mahajan, S. Floyd, and D. Wetherall, "Controlling high-bandwidth flows at the congested router," in *Proceedings Ninth International Conference on Network Protocols. ICNP 2001*, Riverside, CA, USA: IEEE, 2001, pp. 192–201. DOI: 10.1109/ICNP.2001.992899.

[96] M. Claypool and J. Chung, "Rate-Based Active Queue Management with Priority Classes for Better Video Transmission," in *Proceedings ISCC 2002 Seventh International Symposium on Computers and Communications*, Los Alamitos, CA, USA: IEEE Computer Society, 2002, p. 99. DOI: 10.1109/ISCC.2002.1021664.

[97] S. Prabhavat, R. Varakulsiripunth, and S. Noppanakeepong, "Throughput improvement on RED mechanism," in *The 8th International Conference on Communication Systems, 2002. ICCS 2002*, vol. 1, Singapore: IEEE, 2002, pp. 599–603. DOI: 10.1109/ICCS.2002.1182545.

[98] L. A. Grieco and S. Mascolo, "TCP Westwood and Easy RED to Improve Fairness in High-Speed Networks," in *Protocols for High Speed Networks*, G. Carle and M. Zitterbart, Eds., ser. Lecture Notes in Computer Science, vol. 2334, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 130–146. DOI: 10.1007/3-540-47828-0_9.

[99] A. Gyasi-Agyei, "Service differentiation in wireless Internet using multiclass RED with drop threshold proportional scheduling," in *Proceedings 10th IEEE International Conference on Networks (ICON 2002). Towards Network Superiority*, Singapore: IEEE, 2002, pp. 175–180. DOI: 10.1109/ICON.2002.1033307.

[100] J. Koo, K. Chung, H. Kim, and H. Lee, "A New Active RED Algorithm for Congestion Control in IP Networks," in *Information Networking: Wired Communications and Management*, I. Chong, Ed., ser. Lecture Notes in Computer Science, vol. 2343, Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 469–479. DOI: 10.1007/3-540-45803-4_42.

[101] V. Phirke, M. Claypool, and R. Kinicki, "RED-Worcester — traffic sensitive active queue management," in *10th IEEE International Conference on Network Protocols, 2002. Proceedings*, Paris, France: IEEE, 2002, pp. 194–195. DOI: 10.1109/ICNP.2002.1181403.

[102] R. Verma, A. Iyer, and A. Karandikar, "Active queue management using adaptive RED," *Journal of Communications and Networks*, vol. 5, no. 3, pp. 275–281, 2003. DOI: 10.1109/JCN.2003.6596822.

[103] N. Yamagaki, H. Tode, and K. Murakami, "RED method with dual-fairness metrics cooperating with TCP congestion control," in *IEEE International Conference on Communications, 2003. ICC'03*, vol. 1, Anchorage, AK, USA: IEEE, 2003, pp. 652–656. DOI: 10.1109/ICC.2003.1204256.

[104] J. Sun, K.-T. Ko, G. Chen, S. Chan, and M. Zukerman, "PD-RED: to improve the performance of RED," *IEEE Communications Letters*, vol. 7, no. 8, pp. 406–408, 2003. DOI: 10.1109/LCOMM.2003.815653.

[105] C. Wang, B. Li, K. Sohraby, and Y. Peng, "AFRED: an adaptive fuzzy-based control algorithm for active queue management," in *28th Annual IEEE International Conference on Local Computer Networks, 2003. LCN'03. Proceedings*, Bonn/Konigswinter, Germany: IEEE, 2003, pp. 12–20. DOI: 10.1109/LCN.2003.1243108.

[106] C.-H. Chien and W. Liao, "A self-configuring RED gateway for quality of service (QoS) networks," in *2003 International Conference on Multimedia and Expo. ICME'03. Proceedings*, vol. 1, Baltimore, MD, USA: IEEE, 2003, pp. I–793. DOI: 10.1109/ICME.2003.1221037.

[107] J. Orozco and D. Ros, "An Adaptive RIO (A-RIO) Queue Management Algorithm," in *Quality for All*, G. Karlsson and M. I. Smirnov, Eds., ser. Lecture Notes in Computer Science, vol. 2811, Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 11–20. DOI: 10.1007/978-3-540-45188-4_2.

[108] R. Cartas, J. Orozco, J. Incera, and D. Ros, "A fairness study of the adaptive RIO active queue management algorithm," in *Proceedings of the Fifth Mexican International Conference in Computer Science, 2004. ENC 2004*, Colima, Mexico: IEEE, 2004, pp. 57–63. DOI: 10.1109/ENC.2004.1342589.

[109] M. Barbera, F. Licandro, A. Lombardo, G. Schembra, and G. Tusa, "DARED: a double-feedback AQM technique for routers supporting real-time multimedia traffic in a best-effort scenario," in *First International Symposium on Control, Communications and Signal Processing, 2004*, Colima, Mexico: IEEE, 2004, pp. 365–368. DOI: 10.1109/ISCCSP.2004.1296304.

[110] G. Feng, A. Agarwal, A. Jayaraman, and C. K. Siew, "Modified RED gateways under bursty traffic," *IEEE Communications Letters*, vol. 8, no. 5, pp. 323–325, 2004. DOI: 10.1109/LCOMM.2004.827427.

[111] K.-P. Zhang, L. Tian, and Z.-Z. Li, "PRED: A new queue management algorithm with priority and self-adaptation," *Acta Electronica Sinica*, vol. 32, no. 6, pp. 1039–1043, 2004.

[112] M. Claypool, R. Kinicki, and M. Hartling, "Active queue management for Web traffic," in *IEEE International Conference on Performance, Computing, and Communications, 2004*, Phoenix, AZ, USA: IEEE, 2004, pp. 531–538. DOI: 10.1109/PCCC.2004.1395082.

[113] V. Vukadinović and L. Trajković, "RED with Dynamic Thresholds for Improved Fairness," in *Proceedings of the 2004 ACM Symposium on Applied Computing*, ser. SAC'04, Nicosia, Cyprus: Association for Computing Machinery, 2004, pp. 371–372. DOI: 10.1145/967900.967980.

[114] L. Hu and A. D. Kshemkalyani, "HRED: a simple and efficient active queue management algorithm," in *Proceedings. 13th International Conference on Computer Communications and Networks (IEEE Cat. No.04EX969)*, Chicago, IL, USA: IEEE, 2004, pp. 387–393. DOI: 10.1109/ICCCN.2004.1401681.

[115] W. Liu, Z. Yang, J. He, C. Le, and C. T. Chou, "Analysis and improvement on the robustness of AQM in DiffServ networks," in *2004 IEEE International Conference on Communications*, vol. 4, Paris, France: IEEE, 2004, pp. 2297–2301. DOI: 10.1109/ICC.2004.1312928.

[116] W. Liu, Z. Yang, J. He, and C. T. Chou, "Two adaptive AQM algorithms for quantitative Differentiated Services," in *IEEE Global Telecommunications Conference, 2004. GLOBECOM'04*, vol. 3, Dallas, TX, USA: IEEE, 2004, pp. 1703–1707. DOI: 10.1109/GLOCOM.2004.1378272.

[117] C. Wang, B. Li, Y. Thomas Hou, K. Sohraby, and Y. Lin, "LRED: a robust active queue management scheme based on packet loss ratio," in *IEEE INFOCOM 2004*, vol. 1, Hong Kong: IEEE, 2004, p. 12. DOI: 10.1109/INFCOM.2004.1354476.

[118] C. Wang, J. Liu, B. Li, K. Sohraby, and Y. T. Hou, "LRED: A Robust and Responsive AQM Algorithm Using Packet Loss Ratio Measurement," *IEEE Transactions on Parallel and Distributed Systems*, vol. 18, no. 1, pp. 29–43, 2007. DOI: 10.1109/TPDS.2007.253279.

[119] Y. Guo, Y. Zhao, G. Song, X. Xing, and C. Chen, "CGRED: class guided random early discarding," in *IEEE International Conference on Performance, Computing, and Communications, 2004*, Phoenix, AZ, USA: IEEE, 2004, pp. 179–185. DOI: 10.1109/PCCC.2004.1394975.

[120] S. Yi, M. Kappes, S. Garg, X. Deng, G. Kesidis, and C. R. Das, "Proxy-RED: an AQM scheme for wireless local area networks," in *Proceedings. 13th International Conference on Computer Communications and Networks*, Chicago, IL, USA: IEEE, 2004, pp. 460–465. DOI: 10.1109/ICCCN.2004.1401706.

[121] S. Yi, M. Kappes, S. Garg, X. Deng, G. Kesidis, and C. R. Das, "Proxy-RED: an AQM scheme for wireless local area networks," *Wireless Communications and Mobile Computing*, vol. 8, no. 4, pp. 421–434, 2008. DOI: https://doi.org/10.1002/wcm.460.

[122] S. Liu, T. Basar, and R. Srikant, "Exponential-RED: a stabilizing AQM scheme for low- and high-speed TCP protocols," *IEEE/ACM Transactions on Networking*, vol. 13, no. 5, pp. 1068–1081, 2005. DOI: 10.1109/TNET.2005.857110.

[123] S. Guo, X. Liao, C. Li, and D. Yang, "Stability analysis of a novel exponential-RED model with heterogeneous delays," *Computer Communications*, vol. 30, no. 5, pp. 1058–1074, 2007, Advances in Computer Communications Networks. DOI: 10.1016/j.comcom.2006.11.003.

[124] A. E. Kamal and M. Murshed, "Adaptive RED with Dynamic Threshold Adjustment," Iowa State University, Tech. Rep., 2005.

[125] Z. M. Patel, "Queue occupancy estimation technique for adaptive threshold based RED," in *2017 IEEE International Conference on Circuits and Systems (ICCS)*, Thiruvananthapuram, India: IEEE, 2017, pp. 437–440. DOI: 10.1109/ICCS1.2017.8326038.

[126] S. Suresh and Ö. Göl, "Congestion Management of IP Traffic Using Adaptive Exponential RED," in *Networking and Mobile Computing*, X. Lu and W. Zhao, Eds., ser. Lecture Notes in Computer Science, vol. 3619, Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 580–589. DOI: 10.1007/11534310_62.

[127] C. Bouras and A. Sevasti, "Performance enhancement of an AF service using TCP-aware marking and dynamic WRED," in *10th IEEE Symposium on Computers and Communications (ISCC'05)*, Murcia, Spain: IEEE, 2005, pp. 642–647. DOI: 10.1109/ISCC.2005.121.

[128] C.-F. Ku, S.-J. Chen, J.-M. Ho, and R.-I. Chang, "Improving end-to-end performance by active queue management," in *19th International Conference on Advanced Information Networking and Applications (AINA'05)*, vol. 2, Taipei, Taiwan: IEEE, 2005, pp. 337–340. DOI: 10.1109/AINA.2005.214.

[129] S. Kasera, J. Pinheiro, C. Loader, T. LaPorta, M. Karaul, and A. Hari, "Robust multiclass signaling overload control," in *13TH IEEE International Conference on Network Protocols (ICNP'05)*, Boston, MA, USA: IEEE, 2005, pp. 249–258. DOI: 10.1109/ICNP.2005.34.

[130] A. Kesselman, S. Leonardi, and V. Bonifaci, "Game-Theoretic Analysis of Internet Switching with Selfish Users," in *Internet and Network Economics*, X. Deng and Y. Ye, Eds., ser. Lecture Notes in Computer Science, vol. 3828, Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 236–245. DOI: 10.1007/11600930_23.

[131] B. Wang, B. Kasthurirangan, and J. Xu, "Subsidized RED: an active queue management mechanism for short-lived flows," *Computer Communications*, vol. 28, no. 5, pp. 540–549, 2005. DOI: 10.1016/j.comcom.2004.09.010.

[132] J. Zhang and D. A. J. Pearce, "On designing a burst-sensitive RED queue at GPRS links in a heterogeneous mobile environment," in *2005 IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications*, vol. 3, Berlin, Germany: IEEE, 2005, pp. 1719–1723. DOI: 10.1109/PIMRC.2005.1651737.

[133] B. Ng, M. Safi Uddin, A. Malik Abusin, and D. Chieng, "POWARED for Non-Linear Adaptive RED," in *2005 Asia-Pacific Conference on Communications*, Perth, WA, Australia: IEEE, 2005, pp. 832–836. DOI: 10.1109/APCC.2005.1554179.

[134] X. Lin, X. Chang, and J. K. Muppala, "VQ-RED: An efficient virtual queue management approach to improve fairness in infrastructure WLAN," in *The IEEE Conference on Local Computer Networks 30th Anniversary (LCN'05)l*, Sydney, NSW, Australia: IEEE, 2005, pp. 1–7. DOI: 10.1109/LCN.2005.136.

[135] A. V. Korolkova and I. S. Zaryadov, "The mathematical model of the traffic transfer process with a rate adjustable by RED," in *2010 International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, IEEE, Moscow, Russia: IEEE, Oct. 2010, pp. 1046–1050. DOI: 10.1109/ICUMT.2010.5676505.

[136] T. R. Velieva, A. V. Korolkova, and D. S. Kulyabov, "Designing installations for verification of the model of active queue management discipline RED in the GNS3," in *The 6th International Congress on Ultra Modern Telecommunications and Control Systems. Saint-Petersburg, Russia. October 6-8, 2014*, IEEE Computer Society, 2015, pp. 570–577. DOI: 10.1109/ICUMT.2014.7002164.

[137] A. V. Korolkova, D. S. Kulyabov, and L. A. Sevastianov, "Combinatorial and operator approaches to RED modeling," *Mathematical Modelling and Geometry*, vol. 3, no. 3, pp. 1–18, 2015. DOI: 10.26456/mmg/2015-331.

[138] A. V. Korolkova and I. S. Zaryadov, "The mathematical model of the traffic transfer process with a rate adjustable by RED," in *International Congress on Ultra Modern Telecommunications and Control Systems*, Moscow, Russia: IEEE, 2010, pp. 1046–1050. DOI: 10.1109/ICUMT.2010.5676505.

[139] I. S. Zaryadov, A. V. Korolkova, D. S. Kulyabov, T. Milovanova, and V. Tsurlukov, "The survey on Markov-Modulated Arrival Processes and their application to the analysis of active queue management algorithms," in *Distributed Computer and Communication Networks. DCCN 2017. Communications in Computer and Information Science*, ser. Communications in Computer and Information Science, V. M. Vishnevskiy, K. E. Samouylov, and D. V. Kozyrev, Eds., vol. 700, Cham: Springer International Publishing, 2017, pp. 417–430. DOI: 10.1007/978-3-319-66836-9_35.

[140] V. C. C. Hilquias, I. S. Zaryadov, V. V. Tsurlukov, T. A. Milovanova, E. V. Bogdanova, A. V. Korolkova, and D. S. Kulyabov, "The general renovation as the active queue management mechanism. Some aspects and results," in *Distributed Computer and Communication Networks. DCCN 2019*, ser. Communications in Computer and Information Science, V. Vishnevskiy, K. Samouylov, and D. Kozyrev, Eds., vol. 1141, Cham: Springer, 2019, ch. 39, pp. 488–502. DOI: 10.1007/978-3-030-36625-4_39.

[141] A. M. Y. Apreutesey, A. V. Korolkova, and D. S. Kulyabov, "Modeling RED algorithm modifications in the OpenModelica," in *Proceedings of the Selected Papers of the 8th International Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems" (ITTMM-2019), Moscow, Russia, April 15-19, 2019*, D. S. Kulyabov, K. E. Samouylov, and L. A. Sevastianov, Eds., vol. 2407, CEUR-WS, 2019, pp. 5–14.

[142] V. C. C. Hilquias, I. S. Zaryadov, and 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*, vol. 28, no. 4, pp. 305–318, 2020. DOI: 10.22363/2658-4670-2020-28-4-305-318.

For citation:

I. S. Zaryadov, H. C.C. Viana, A. V. Korolkova, T. A. Milovanova, Chronology of the development of Active Queue Management algorithms of RED family. Part 1: from 1993 up to 2005, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 305–331. DOI: 10.22363/2658-4670-2023-31-4-305-331.

Information about the authors:

Zaryadov, Ivan S. — Candidate of Physical and Mathematical Sciences, Assistant Professor of Department of Probability Theory and Cyber Security, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: zaryadov-is@rudn.ru, ORCID: https://orcid.org/0000-0002-7909-6396)

Viana, C.C. Hilquias — Ph.D. student of Department of Probability Theory and Cyber Security, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: hilvianamat1@gmail.com, ORCID: https://orcid.org/0000-0002-1928-7641)

Korolkova, Anna V. — Candidate of Physical and Mathematical Sciences, Associate Professor of Department of Probability Theory and Cybersecurity, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: korolkova-av@rudn.ru, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0001-7141-7610)

Milovanova, Tatianna A. — Candidate of Physical and Mathematical Sciences, Assistant Professor of Department of Probability Theory and Cybersecurity, Institute of Computer Science and Telecommunications, Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: milovanova-ta@rudn.ru, ORCID: https://orcid.org/0000-0002-9388-9499)

УДК 519.71:519.872
DOI: 10.22363/2658-4670-2023-31-4-305-331
EDN: FPQLOH

Хронология развития алгоритмов активного управления очередями семейства RED. Часть 1: 1993–2005

И. С. Зарядов[1,2], **К. К. И. Виана**[1], **А. В. Королькова**[1], **Т. А. Милованова**[1]

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

Аннотация. Статья является первой частью большого библиографического обзора по алгоритмам активного управления очередями, относящимся к семейству алгоритмов случайного раннего обнаружения (RED), представленных в научной печати с 1993 по 2023 года. В первой части приведены данные по алгоритмам, опубликованным с 1993 по 2005 года.

Ключевые слова: активное управление очередями, AQM, RED, управление перегрузками

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 332–344
http://journals.rudn.ru/miph

Research article

UDC 519.21
DOI: 10.22363/2658-4670-2023-31-4-332-344
EDN: ECPYRN

On the algorithmization of construction of the transition intensity matrix in systems with a large number of same elements

Sergey I. Matyushenko, Ivan S. Zaryadov

RUDN University,
6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: October 14, 2023; revised: December 1, 2023; accepted: December 29, 2023)

Abstract. In this article, using the example of a multi-channel exponential queueing system with reordering of requests, we study the problem of computer construction of the state space and coefficient matrix of a system of equilibrium equations. As a result, general principles for solving problems of this type are formulated.

Key words and phrases: queuing system, dynamic construction of transition intensity matrix, systems with reordering of requests

1. Introduction

Quite often, when constructing mathematical models of complex information and computing systems, developers are faced with the problem of the presence of a large number of similar elements in the system under study. This is especially true for multi-threaded systems or systems with a large number of servicing devices [1–7]. If the functioning of the system has a complex multivariate scenario, then the task of constructing a state space of a random process modeling the system under study becomes impossible without developing a special software product. In this article, using the example of a multi-channel queuing system with reordering of requests [8, 9], an algorithm will be developed for the automated construction of the state space and transition intensity matrix.

2. System description

A multi-channel queuing system (QS) is considered with m service devices, $2 \leqslant m < \infty$ and a common storage device of limited capacity r. The system receives a Poisson flow of requests of intensity λ. The service times on device j are independent of each other, and also do not depend on the duration of service on other devices and are distributed according to an exponential law with the parameter μ_j, $j = 1, \ldots, m$. An application entering the system when there are $m + r$ applications in it is lost.

© Matyushenko S. I., Zaryadov I. S., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

We will further assume that the devices are arranged in non-decreasing order of service intensity: $\mu_1 \geqslant \cdots \geqslant \mu_m$ and that a request that has the ability to select a device selects the device with the lowest serial number. The selection of applications from the storage occurs in accordance with the FIFO discipline [10].

It is assumed that all applications upon entry into the system are assigned a serial number. Moreover, if at the moment of completion of servicing of a request with number n (n-request), servicing of at least one request with a number less than n continues, the n-request is placed in the reordering buffer (RB). Otherwise, request n leaves the system and all requests with numbers differing by one, starting from $n+1$ (if there are any in the RB), leave the RB behind it. This assumption allows us to model the mechanism for maintaining the order in which applications leave the system in accordance with the order in which they arrive. Systems of this kind are called systems with reordering of requests [8, 9].

3. Construction of a mathematical model

Let us assume that all applications in the system are numbered in accordance with the order in which they were received, starting with one. Then the stochastic behavior of the considered QS can be described by a homogeneous Markov process $X(t)$, $t \geqslant 0$, over the state space

$$x^m = \bigcup_{k=0}^{m+r} x_k^m,$$

$$x_k^m = \Big\{(k, i_1, \ldots, i_m), \quad i_j = \overline{0,k}, \quad \sum_{j=1}^{m} u(i_j) = k,$$

$$\text{in this case, if } i_j i_s > 0, \text{ then } i_j \neq i_s, \ j,s = \overline{1,m}\Big\}, \quad k = \overline{0,m-1},$$

$$x_k^m = \Big\{(k, i_1, \ldots, i_m), \quad i_j = \overline{0,m}, \ i_j \neq i_s, \ j,s = \overline{1,m}\Big\}, \quad k = \overline{m,m+r},$$

where $u(x)$ is the Heaviside function.

Here for some time t: $X(t) = (k, i_1, \ldots, i_m)$, if there are k requests in the system, $k = 0, \ldots, m+r$ and device j is free if $i_j = 0$. Otherwise, i_j is the serial number of the request served on device j, $j = 1, \ldots, m$.

In what follows, we will call the subset x_k^m the k-th group of states. It's easy to see that

$$|x_k^m| = \begin{cases} (m)_k, & k = \overline{0,m}, \\ m!, & k = \overline{m,m+r}, \end{cases}$$

where $(m)_k$ is the number of placements from m to k.

Hence,

$$|x^m| = \sum_{k=0}^{m} (m)_k + rm!\,.$$

It is obvious that as m increases, the dimension of the state space increases rapidly. So when $m \geqslant 5$ and $k \geqslant 10$ it exceeds 10^3. Therefore, to construct a matrix of transition intensities and solve a system of equilibrium equations, it is necessary to develop an algorithm for constructing a space x^m.

4. Algorithm for constructing the state space

Let Y_s be a set of sequences of length $s+1$ of non-negative integers.

Definition 1. We will call the operator L_j the j-insertion operator defined on the set Y_s if for $(i_0, \ldots, i_s) \in Y_s$

$$L_j(i_0, i_1, \ldots, i_s) = (i_0{+}1, i_1, \ldots, i_{-1}, \max\{i_1, \ldots, i_s\}{+}1, i_j, \ldots, i_s), \quad j{-}\overline{1, s+1}.$$

Next, let $Y_{s,\nu}$ be the subset Y_s of power ν, i.e. $Y_{s,\nu} = \{y_s^1, \ldots, y_s^\nu\}$, where $y_s^n = (i_0^n, \ldots, i_s^n)$, $n = \overline{1, \nu}$.

Definition 2. We will call the operator L the insertion operator defined on the set of different finite subsets of the set Y_s, if for $Y_{s,\nu} \in Y_s$

$$L(Y_{s,\nu}) = \left\{L_1 y_s^1, \ldots, L_1 y_s^\nu, L_2 y_s^1, \ldots, L_2 y_s^\nu, \ldots, L_{s+1} y_s^1, \ldots, L_{s+1} y_s^\nu\right\}.$$

Definition 3. k-th degree L^k operator L will be called an operator whose action consists of k successive applications of the operator L, $k = 1, 2, 3, \ldots$. By the zero degree of the operator L we mean the identity operator.

Definition 4. We will call L_j^{-1} a j-removal operator defined on the set Y_s if for $(i_0, \ldots, i_s) \in Y_s$:

$$L_j^{-1}(i_0, \ldots, i_{j-1}, i_j, i_{j+1}, \ldots, , i_s) = (i_0 - 1, i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, , i_s), \quad j = \overline{1, s}.$$

Let's define a subset $\tilde{Y}_s$ as set Y_s such that for $(i_0, i_1, \ldots, i_s) \in Y_s$ among the numbers there is at least one that is not equal to zero and all non-zero numbers are distinct.

Definition 5. We will call the operator M a maximum selection operator defined on the set $\tilde{Y}_s$ if for $(i_0, i_1, \ldots, i_s) \in \tilde{Y}_s$: $M(i_0, i_1, \ldots, i_s) = l$, where l is such that $i_l = \max\{i_1, \ldots, i_s\}$.

And, finally, let $\hat{Y}_s$ is subset of the set Y_s, such that $(i_1, \ldots, i_s) \in \hat{Y}_s$, if among the numbers $i_1, \ldots, i_s$ there is at least one that is equal to zero.

Definition 6. We will call the operator Z the zero selection operator on the set $\hat{Y}_s$ if for $(i_0, i_1, \ldots, i_s) \in \tilde{Y}_s$: $Z(i_1, \ldots, i_s) = n$, where n is the number of the first zero element in the sequence $i_1, \ldots, i_s$.

Let us now proceed to constructing the state space and prove the validity of the following lemma.

Lemma 1. *For any fixed m, $m \geqslant 2$:*

$$x_k^m = \begin{cases} L^k(0, 0^{m-k}), & k = \overline{0, m-1}, \\ L^{m-1}(k-m+1, 1), & k = \overline{m, m+r}, \ \textit{where } 0^s = (0, \ldots, 0). \end{cases} \tag{1}$$

Proof. We will carry out the proof using the method of mathematical induction. Let $m = 2$. Then

$$x_k^2 = \begin{cases} (0,0,0), & k = 0; \\ L^1(0,0), & k = 1; \\ L^1(k-1,1), & k = \overline{2, r+2}. \end{cases}$$

Hence, $= \big\{(0,0,0); (1,1,0); (1,0,1); (k,1,2); (k,2,1), k = 2, \ldots, r+2\big\}$. It is obvious that the resulting set is the state space for the process $X(t)$ in the case of $m = 2$ [4].

Let the statement of the lemma be true for $m = l$. Then for $m = l + 1$ we get

$$x_k^{l+1} = \begin{cases} L^k(0, 0^{l+1-k}), & k = \overline{0, l}; \\ L^l(k-l, 1), & k = \overline{l+1, l+r+l} \end{cases} =$$

$$= \begin{cases} (0, 0^{l+1}), & k = 0; \\ L(L^{k-1}(0, 0^{l+1-k})), & k = \overline{1, l}; \\ L(L^{l-1}(k-l, 1)), & k = \overline{l+1, l+r+l}. \end{cases}$$

Replace k with $k + 1$ and get:

$$x_{k+1}^{l+1} = \begin{cases} (0, 0^{l+1}), & k = -1; \\ L(L^k(0, 0^{l-k})), & k = \overline{0, l-1}; \\ L(L^{l-1}(k-l+1, 1)), & k = \overline{l, l+r} \end{cases} = \begin{cases} (0, 0^{l+1}), & k = -1; \\ L(x_k^l), & k = \overline{0, l-1}; \\ L(x_k^l), & k = \overline{l, l+r}. \end{cases}$$

Next we note that if $(k, i_1, \ldots, i_l) \in x_k^l$, then

$$\max\{i_1, \ldots, i_l\} = \begin{cases} k, & k = \overline{0, l-1}; \\ l, & k = \overline{l, l+r}. \end{cases}$$

Hence,

$$x_{k+1}^{l+1} = \begin{cases} (0, 0^{l+1}), \quad k = -1; \\ \{(k+1, k+1, i_1^1, \ldots, i_l^1);\ (k+1, k+1, \ldots, i_l^2); \ldots; \\ \qquad (k+1, i_1^\nu, \ldots, k+1)\}, \quad k = \overline{0, l-1}, \quad \nu = |x_k^l|; \\ \{(k+1, l+1, i_1^1, \ldots, i_l^1);\ (k+1, l+1, \ldots, i_l^2); \ldots; \\ \qquad (k+1, i_1^\nu, \ldots, l+1)\}, \quad k = \overline{l, l+r}, \quad \nu = |x_k^l|. \end{cases}$$

Given the definition of a group of states x_k^l, the last relation can be written as:

$$x_{k+1}^{l+1} = \begin{cases} \big\{(k+1, i_1, \ldots, i_l) : \\ \qquad i_j = \overline{0, k+1}, \quad \sum\limits_{j=1}^{l+1} u(i_j) = k+1, \text{ in this case,} \\ \quad \text{if } i_j i_s > 0, \text{ then } i_j \neq i_s,\ j,\ s = \overline{1, l+1}\big\}, \quad k = \overline{-1, l-1}; \\ \big\{(k+1, i_1, \ldots, i_{l+1}) : \\ \quad i_j = \overline{1, l+1}, \quad i_j \neq i_s,\ j,\ s = \overline{1, l+1}\big\}, \quad k = \overline{l, l+r}. \end{cases}$$

And finally, having made the reverse replacement of k by $k-1$, we come to the definition of the k-th group of states for the case $m = l+1$. Thus, the lemma is proven. □

It obviously follows from lemma 1 that the proposed method of constructing x^m is recurrent in m. In addition, in the state space a certain order of the elements of this space is specified. For clarity, let us consider the diagram for constructing the state space in the case when $m = 4$ and $r = 1$ (figure 1).

Analysis of the diagram helps to notice that for any fixed m, the k-th group of states is divided into m subgroups of the same dimension. A sign that a state belongs to the n-th subgroup of the k-th group is that the request with the highest number is served on device n, $n = 1, \ldots, m$. Let us denote $x_{k,n}^m$ by the n-th subgroup of the k-th group. It is easy to calculate that

$$|x_{k,n}^m| = (m-1)_{\min\{k, m-1\}-1}, \quad n = \overline{1, m}, \quad k = \overline{1, m+r}. \tag{2}$$

The recurrent principle of constructing the state space and dividing groups into subgroups makes it possible to determine the serial number of the state in x^m.

Lemma 2. *The ordinal number of the state $(k, i_1, \ldots, i_m)$ in the state space x^m is determined by the expression*

$$n = \sum_{j=0}^{\min\{k-1, m-1\}} (m)_j + u(k-m)m! + \sum_{j=1}^{\min\{k, m-1\}} (s_j - 1)(m-j)_{\min\{k,m\}-j} + 1, \tag{3}$$

where $s_1 = M(k, i_1, \ldots, i_m)$, $s_j = M\big(L_{s_{j-1}}^{-1} \ldots L_{s_2}^{-1} L_{s_1}^{-1}(k, i_1, \ldots, i_m)\big)$, $j = \overline{2, \min\{m-1, k\}}$.

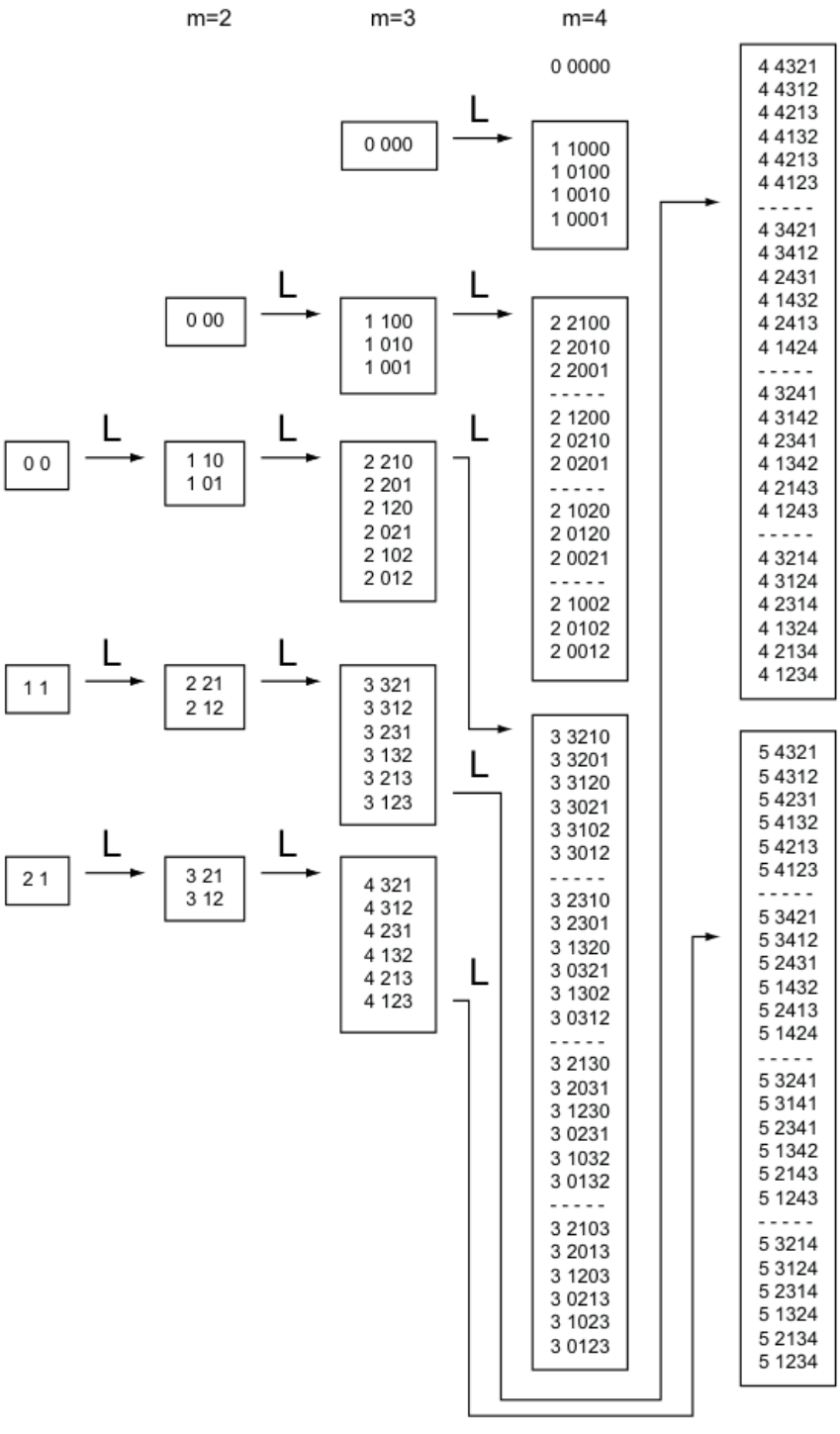

Figure 1. Diagram of the process of constructing a state space for a system with 4 devices

Proof. Note that the first two terms in our formula determines the total number of states in groups $x_0^m, \ldots, x_{k-1}^m$. Therefore, it remains to show that $\sum_{j=1}^{\min\{k,m-1\}} (s_j - 1)(m-j)_{\min\{k,m\}-j}$ determines the number of states of the k-th group preceding a given state. To do this, let us clarify the meaning of each term of this sum.

Notice, that $s_1 = M(k, i_1, \ldots, i_m)$ is the number in the subgroup x_k^m that belongs to state $(k, i_1, \ldots, i_m)$, and the size of each of the subgroups is equal $(m-1)_{\min\{k,m\}-1}$. Therefore, the expression $(s_1-1)(m-1)_{\min\{k,m\}-1}$ determines the total number of states in the subgroups $x_{k,1}^m, \ldots, x_{k,s_1-1}^m$, preceding the subgroup x_{k,s_1}^m. Moreover, if $k = 1$ or $m = 2$, then each of the subgroups contains one element and, therefore, the calculation process will be completed. Otherwise, you need to define the state number in the subgroup x_{k,s_1}^m.

We note that in the process of recurrent construction of the state space, the subgroup x_{k,s_1}^m is obtained as a result of the action of the operator L_{s_1} on the group x_{k-1}^{m-1}. Therefore, the ordinal number of state $(k, i_1, \ldots, i_m)$ in the subgroup x_{k,s_1}^m is equal to the ordinal number of state $L_{s_1}^{-1}(k, i_1, \ldots, i_m)$ in the group x_{k-1}^{m-1}. To determine it, we find $s_2 = M\big(L_{s_1}^{-1}(k, i_1, \ldots, i_m)\big)$ — the number of the subgroup in the group x_{k-1}^{m-1} to which state $L_{s_1}^{-1}(k, i_1, \ldots, i_m)$ belongs. The size of each of these subgroups is equal $(m-2)_{\min\{k,m\}-2}$. So expression $(s_2 - 1)(m-2)_{\min\{k,m\}-2}$ determines total number of states in the subgroups $x_{k-1,1}^{m-1}, \ldots, x_{k-1,s_2-1}^{m-1}$, preceding the subgroup x_{k-1,s_2}^{m-1}. Moreover, if $k = 2$ or $m = 3$, then each of the subgroups contains one element and the calculation process will be completed. Otherwise, it is necessary to continue similar reasoning, which will ultimately lead us to the desired result.

Thus the lemma is proven. □

In the future, we will need not only a formula for calculating the serial number of a state, but also an algorithm for the reverse action — restoring a state by its serial number. This task in our case is divided into two stages: determining the number of the group to which a given state belongs and calculating the serial number of the state in the group.

The first stage is simple. Let n be the serial number of the state, and $N = |x^m|$. Next, it is necessary to arrange a partition of the segment $[0,\ N]$ into $m + r + 1$ interval $(n_k, n_{k+1}]$, $k = 0, \ldots, m + r$ in such a way that the condition $n \in (n_k, n_{k+1}]$ means that the state with number n belongs to the k-th group. Obviously, as the boundaries of the indicated intervals it is necessary to take the numbers:

$$n_s = \begin{cases} 0, & s = 0; \\ n_{s-1} + (m)_{s-1}, & s = \overline{1, m}; \\ n_{s+1} + m!, & s = \overline{m+1, m+r+1}. \end{cases} \tag{4}$$

Further, carrying out arguments similar to those that took place in the proof of lemma 2, we arrive at the following result.

Lemma 3. *Let n be the serial number of the state, and k be the number of the group to which this state belongs, and let the sequence of numbers $s_1, \dots, s_{\min\{k,m-1\}}$ be defined by the following recurrence relations*

$$\begin{aligned} l_j &= \begin{cases} n - n_k, & j = 1; \\ l_{j-1} - (s_{j-1} - 1)t_{j-1}, & j = \overline{2, \min\{m-1, k\}}; \end{cases} \\ t_j &= (m-j)_{\min\{k,m\}-j}, \quad s_j = \left\lceil \frac{l_j}{t_j} \right\rceil, \quad j = \overline{1, \min\{m-1, k\}}. \end{aligned} \tag{5}$$

Then the state of the system is determined by the expression:

$$(k, i_1, \dots, i_m) = \begin{cases} L_{s_1} \dots L_{s_k}(0, 0^{m-k}), & k = \overline{0, m-1}; \\ L_{s_1} \dots L_{s_{m-1}}(k-m+1, 1), & k = \overline{m, m+r}. \end{cases} \tag{6}$$

To illustrate the operation of our algorithm, we give an example of restoring a state by its serial number.

Let $m = 4$, $r = 2$, $n = 61$.

Then $N = \sum_{j=0}^{6}(4)_j = 1 + 4 + 12 + 24 + 24 + 24 = 89$ and the segment $(0,\ 89]$ is divided into intervals:

$$\begin{aligned} &(n_0, n_1] \equiv (0,\ 1]; \quad (n_1, n_2] \equiv (1,\ 5]; \quad (n_2, n_3] \equiv (5\ , 17]; \\ &(n_3, n_4] \equiv (17,\ 41]; \quad (n_4, n_5] \equiv (41,\ 65]; \quad (n_5, n_6] \equiv (65,\ 89]. \end{aligned}$$

The number 61 belongs to the interval $(n_4, n_5]$. Therefore, group number $k = 4$. Next, we calculate the sequence of numbers.

$$\begin{aligned} &l_1 = 61 - 41 = 20, && t_1 = (4-1)_{4-1} = 6, && s_1 = \lceil 20/6 \rceil = 4; \\ &l_2 = 20 - (4-1) \cdot 6 = 2, && t_2 = (4-2)_{4-2} = 2, && s_2 = \lceil 2/2 \rceil = 1; \\ &l_3 = 2 - (1-1) \cdot 2 = 2, && t_3 = (4-3)_{4-3} = 1, && s_3 = \lceil 2/1 \rceil = 2. \end{aligned}$$

And finally, the desired state is determined by the expression:

$$(4, i_1, i_2, i_3, i_4) = L_4 L_1 L_2(1,1) = L_4 L_1(2,1,2) = L_4(3,3,1,2) = (4,3,1,2,4).$$

Using lemma 2, we check the result:

$$\begin{aligned} s_1 &= M(4,3,1,2,4) = 4; \\ s_2 &= M\big(L_4^{-1}(4,3,1,2,4)\big) = M(3,3,1,2) = 1; \\ s_3 &= M\big(L_1^{-1}(3,3,1,2)\big) = M(2,1,2) = 2; \\ n &= \sum_{j=0}^{3}(4)_j + \sum_{j=1}^{3}(s_j - 1)(4-j)_{4-j} + 1 = 41 + 20 = 61. \end{aligned}$$

5. Algorithm for constructing the matrix of transition intensities

Let us develop an algorithm for constructing the matrix of MP transition intensities $X(t)$. Note that MP transitions are possible only for states from neighboring groups: transition from the group x^m_{k-1} to a group x^m_k due to the receipt of an application, and from group x^m_{k+1} to group x^m_k — due to service on one of the devices. Consequently, the transition intensity matrix A will have a 3-diagonal form.

Let us introduce notation for non-zero blocks A. We denote over-diagonal blocks by $\Lambda_{k-1,k}$, sub-diagonal — through $M_{k,k-1}$, and diagonal — through N_{kk}.

It is not difficult to notice that in each column of the block $\Lambda_{k-1,k}$ there can be only one non-zero element equal to λ. In each line of the block $M_{k,k-1}$ corresponding to state $(k, i_1, \ldots, i_m)$ there will be no more than m elements for j such that $i_j \neq 0$. Finally, the blocks $N_{k,k}$ are diagonal matrices with the diagonal element corresponding to state $(k, i_1, \ldots, i_m)$ equal to $\sum_{j=1}^{m} u(i_j)\mu_j - u(m + r - k)\lambda_k$.

Thus, to construct matrix A it is necessary:

— develop an algorithm E_0 that determines for any state of the system the transition conditions and the state of the system from which one can get to the given one due to the receipt of an application;

— develop an algorithm E_j, which defines the transition conditions and the state of the system to which it is possible to move from this state due to servicing on device j.

The implementation of these algorithms will make it possible to determine non-zero elements of blocks $\Lambda_{k-1,k}$ and $M_{k,k-1}$.

Before moving on to a step-by-step description of the algorithm E_0, we will give some explanation of one of the conditions that will be checked in this algorithm. We are talking about when it is impossible to transition to state $(k, i_1, \ldots, i_m)$, from any other due to the arrival of a request, except for the trivial case $k = 0$. In the description of the system it was said that an application that has the ability to select a device selects the device with the lowest serial number. Consequently, due to the receipt of a request, it is impossible to get into a state for which the number of the device busy servicing the request with the maximum sequence number is greater than the number of the first of the free devices. Formally, for any state $(k, i_1, \ldots, i_m)$, this condition can be written as follows

$$M(k, i_1, \ldots, i_m) > Z(k, i_1, \ldots, i_m),$$

where M and Z are the maximum and zero selection operators, respectively. Moreover, this condition makes sense to check when $k < m$. Otherwise the check is trivial. The meaning of the remaining steps of the algorithm is quite obvious, so we will not give a detailed explanation.

So, **algorithm** E_0:

Start: enter state $(k, i_1, \ldots, i_m)$.

Step 1. Check the condition $k > m$. If the condition is met, then go to step 5.

Step 2. Check the condition $k = m$. If the condition is met, then go to step 4.

Step 3. Check the condition $M(k, i_1, \ldots, i_m) > Z(k, i_1, \ldots, i_m)$. If the condition is met, then the end of the algorithm.

Step 4. $l := Z(k, i_1, \ldots, i_m)$, $i_l := 0$.

Step 5. $k := k - 1$.

Step 6. Output values k, i_1, $\ldots$, i_m.

End algorithm.

Now let's move on to the algorithm E_j. Note that from any state $(k, i_1, \ldots, i_m)$ you can get to another due to servicing on device j if this device was busy with servicing, i.e. if $i_j \neq 0$. After the end of service on device j, the number of requests decreases by one. In addition, the serial numbers of those applications that entered the system later than the serviced application are reduced. Further, if at the time of the end of service there were requests in the queue awaiting service, then the first of them will arrive at device j and will be assigned a serial number m. Otherwise, device j will become free.

Considering the above, the **algorithm** E_j can be written as:

Start: Enter state $(k, i_1, \ldots, i_m)$.

Step 1. Check condition $i_j = 0$. If the condition is met, then the end of the algorithm.

Step 2. $k := k - 1$, $l := 0$.

Step 3. $l := l + 1$.

Step 4. Check the condition $\big((i_l > 0)$ and $(i_l < i_j)\big)$. If the condition is met, then go to step 6.

Step 5. $i_l := i_l - 1$.

Step 6. Check the condition $l \neq m$. If the condition is met, then go to step 3.

Step 7. $i_j := 0$.

Step 8. Check the condition $k < m$. If the condition is met, then go to step 10.

Step 9. $i_j := m$.

Step 10. Print values k, i_1, $\ldots$, i_m.

End algorithm.

And finally, we will describe the algorithm E for constructing matrix A. In fact, we need to perform the following actions for all serial numbers of states of the system n:

Step 1. Using lemma 3, from the ordinal number n, restore the state of the system $(k, i_1, \ldots, i_m)$.

Step 2. Using an algorithm E_0 (algorithms E_j), determine the transition conditions, and if they are fulfilled, the state of the system from which (to which) one can get to the given one (from the given one).

Step 3. Using lemma 2, determine the serial numbers of states n_j, $j = 0, 1, \ldots, m$, defined in the previous step, and perform the following assignment operations:

$$a_{n_0 n} := \lambda, \tag{7}$$

$$a_{n n_j} := \mu_j, \tag{8}$$

$$a_{nn} := -u(m + rk)\lambda - \sum_{j=1}^{m} u(i_j)\mu_j. \tag{9}$$

Thus, the main result of our study has been obtained. Let us formulate it in the form of a theorem.

Theorem 1. *The numbering of non-zero elements of the matrix A of the intensities of transitions of the MP $X(t)$, which describes the functioning of the system under consideration, is determined in accordance with the algorithm E, and their values are calculated using formulas* (7)–(9).

6. Conclusion

Despite the fact that in this work we were constructing a matrix of transition intensities for a very specific system, we can note a number of patterns and formulate recommendations for solving problems of this type when considering other systems with a large number of similar elements. Let's list the main stages:

1. Describe all possible states of the system in the form of sequences of corresponding numerical parameters.
2. Identify the main patterns of formation of the state space with an increase in the number of similar elements in the system and describe them using special operators.
3. To develop an algorithm for constructing a state space recurrently by the number of elements of the same type.
4. Define and formalize with the help of logical operators all possible transitions between different states of the system.
5. Develop an algorithm that allows you to determine the ordinal number of a state in the state space.
6. Develop an inverse algorithm that allows, using the serial number of a state, to restore the state itself in the form of a sequence of parameters.
7. Write down formulas for determining non-zero elements of the matrix of intensities of transitions between states of the system depending on the serial numbers of these states.
8. Enumerate all states of the system in accordance with their serial numbers and, restoring them using the inverse algorithm, calculate all non-zero elements of the matrix using the formulas from the previous paragraph.

References

[1] V. L. Chugreev, "Development of a multicative model of sequentially connected information elements," *Young Scientist*, no. 3, pp. 147–149, 2013.

[2] G. P. Basharin, S. N. Klapouschak, and N. V. Mitkina, "Mathematical model of adaptive high-speed system with elastic traffic," *Bulletin of the RUDN. Mathematics series. Computer science. Physics*, no. 3, pp. 31–39, 2008, in Russian.

[3] A. Rumyantsev and E. Morozov, "Stability criterion of a multiserver model with simultaneous service," *Operation Reseach*, no. 5, pp. 31–39, 2017. DOI: 10.1007/s10479-015-1917-2.

[4] S. A. Grishunina, "Multiserver queueing system with constant service time and simultaneous service of a customer by random number of servers," *Theory of Probability and its Applications*, vol. 64, no. 3, pp. 456–460, 2019. DOI: 10.1137/50040585X97T98960X.

[5] B. Sun, M. H. Lee, S. A. Dudin, and A. N. Dudin, "Analysis of multi-server queueing system with opportunistic occupation and reservation of servers," *Mathematical Problems in Engineering*, no. 5, pp. 1–13, 2014. DOI: 10.1155/2014/178108.

[6] U. Ayestab, P. Jackod, and V. Novak, "Scheduling of multi-class multi-server queueing systems with abandonments," *Journal of Scheduling*, vol. 20, pp. 129–145, 2015. DOI: 10.1007/s10951-015-0456-7.

[7] M. Harchol-Balter, T. Osogami, A. Scheller-wolf, and A. Wierman, "Multi-server queueing systems with multiple priority classes," *Queueing Systems*, vol. 51, pp. 331–360, Dec. 2005. DOI: 10.1007/s11134-005-2898-7.

[8] S. Matyushenko and A. Ermolaeva, "On stationary characteristics of a multi server exponential queueing system with reordering of requests," in *13th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*, ICUMT 2021, Brno, Czech Republic, 2021, pp. 98–103.

[9] S. I. Matyushenko and A. V. Pechinkin, "Service system with reordering of applications," in *International Conference Distributed Computer and Communication Networks (ECN 2011)*, ECN 2011, Moscow, Russia, 2011.

[10] P. P. Bocharov and A. V. Pechenkin, *Queuing theory [Teoriya massovogo obsluzhivaniya]*. Moscow: RUDN, 1995, 529 pp., in Russian.

For citation:
S. I. Matyushenko, I. S. Zaryadov, On the algorithmization of construction of the transition intensity matrix in systems with a large number of same elements, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 332–344. DOI: 10.22363/2658-4670-2023-31-4-332-344.

Information about the authors:
Matyushenko, Sergey I. — Candidate of Physical and Mathematical Sciences, Assistant professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: matyushenko-si@rudn.ru, ORCID: https://orcid.org/0000-0001-8247-8988)
Zaryadov, Ivan S. — Candidate of Physical and Mathematical Sciences, Assistant professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: zaryadov-is@rudn.ru, ORCID: https://orcid.org/0000-0002-7909-6396)

УДК 519.21
DOI: 10.22363/2658-4670-2023-31-4-332-344
EDN: ECPYRN

Об алгоритмизации построения матрицы интенсивностей переходов в системах с большим числом однотипных элементов

С. И. Матюшенко, И. С. Зарядов

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

Аннотация. В данной статье на примере многоканальной экспоненциальной системы массового обслуживания с переупорядочиванием заявок изучается задача компьютерного построения пространства состояний и матрицы коэффициентов системы уравнений равновесия. В результате сформулированы общие принципы решения задач такого типа.

Ключевые слова: система массового обслуживания, динамическое построение матрицы интенсивностей переходов, система с переупорядочиванием заявок

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 345–358
http://journals.rudn.ru/miph

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-4-345-358
EDN: DYDLCY

Evaluation of firewall performance metrics with ranging the rules for Poisson incoming packet flow and exponential filtering time

Anatoly Yu. Botvinko, Konstantin E. Samouylov

RUDN University,
6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: October 20, 2023; revised: December 1, 2023; accepted: December 29, 2023)

Abstract. The given article is a continuation of a number of works devoted to the development of models and methods for ranging the filtration rules to prevent a decrease in the firewall performance caused by the use of a sequential scheme for checking packet compliance with the rules, as well as by the heterogeneity and variability of network traffic. The article includes a description of a firewall mathematical model given in the form of a complex system and a queuing system with a phase-type discipline for request servicing, which formalizes the network traffic filtering process with the functionality of ranging the rules. The purpose of modeling is to obtain estimates for major firewall performance metrics for various network traffic behavior scenarios, as well as to evaluate an increase in the firewall performance due to ranging a filtration rule set. Calculation of estimates for the firewall (FW) performance metrics was made using the analytical method for a Poisson request flow. Based on the analysis of the modeling results, conclusions were drawn on the effectiveness of ranging the filtration rules in order to improve the firewall performance for traffic scenarios that are close to real ones.

Key words and phrases: firewall, ranging the filtration rules, network traffic, phase service, queuing system

1. Introduction

Sustainable operation of information infrastructure, including for special-purpose automated systems (AS), the uninterrupted functioning of which is critical for ensuring the security and defense capability of any state, given the avalanche-like growth in the volume of information flows in public networks, high heterogeneity and variability of network traffic parameters, widespread use of multimedia protocols (that are quite sensitive to the length of data transmission delay), as well as a significant increase in the number of various computer attacks, requires firewalls to provide really high performance. A firewall is a local or functional distributing tool that provides control over

© Botvinko A. Y., Samouylov K. E., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

the incoming and/or outgoing information in the automated system, and ensures the protection for the AS by filtering the information, i.e., providing analysis of the information by the criteria set and making a decision on its distribution [1].

One of the major factors affecting the search time for filtration rules, and therefore the FW performance, is the order in which the filtration rules are arranged in sets that are linear lists of large dimensions. This is due to the fact that the search time for any rule corresponding to the data under filtration, is in proportion to the number of checked rules. And the filtering time for information flow that meets the conditions contained at the end of a large dimension set, will be much longer than the time required to filter data that meet the conditions contained at the beginning of the rule set [1, 2].

The papers [1, 3, 4] published earlier by the authors, describe the developed method for optimizing a filtration rule set (method for ranging the rules). An increase in the efficiency of traffic filtration is achieved by periodically ranging the filtration rules in descending order of their weights, obtained in accordance with the estimates of the parameters of the filtered information flows. A particularity of the developed approach is the use of the non-parametric method of local approximation (MLA) [4, 5] to evaluate the parameters of filtered information flows. In the ranging process for a rule set, the current characteristics and dynamics of changes in the parameters of information flows are considered. At the same time, there is no need to select a parametric model that is acceptable for all evaluated parameters of information flows. The implementation of MLA has provided the adaptability of the method, as well as a high response speed for changes in the parameters of filtered information flows thanks to the specifics of MLA estimates.

In earlier works [1–3], the effectiveness of using methods for optimizing a filtration rule set was evaluated with the help of a simulation modeling method. To evaluate the effectiveness of methods for set optimization, the presented paper proposes an analytical solution, as well as an algorithm for calculating the probabilistic and temporal characteristics of the queuing system (QS) that formalizes the network traffic filtering process.

2. Firewall model with ranging the filtration rules

As the FW model we chose the previously developed one [2, 3] that reflects the basic patterns and factors of the FW functioning when processing the network traffic. The traffic processing in this model includes two stages that are an initial processing stage and a stage of checking the packet filtration rules.

At the initial processing stage, a packet, which is transmitted in the communication network, is received by the network interface card of the FW. After decoding a sequence of electrical or optical signals and verifying the correctness of the delivered information, the packet is written to the input buffer memory of the network interface controller (NIC). After that, the packet is transferred to a common software buffer allocated in RAM for further extraction and processing by the central processing unit. In the proposed model, the process of decoding a sequence of signals and the process of receiving and transferring the packet from the buffer memory of the network card to the random access memory (RAM) of the FW are considered

as a single initial processing of packets with a given service intensity. It's considered that the packet immediately arrives at the RAM of the firewall. The waiting time of the packet in the buffer memory of the network card isn't taken into account, as well as the packet losses due to distortions in information transmission.

At the stage of checking the packet filtration rules, the central processing unit (CPU), if computing resources are available, provides the sequential check of the compliance of the incoming packet parameters with the conditions of the filtration rules. The remaining incoming packets await the start of servicing in the buffer in the order of their arrival (FCFS, First-Come, First-Served). A similar approach was used in papers [6, 7].

If the packet parameters comply with the conditions of the filtration rules, the packet processing is completed with encoding and transferring the packet to the physical medium. The packets that don't comply with the filtration rules are discarded.

The following operations aren't considered within the traffic filtering process: fragmentation and reassembly of transmitted packets, reassembly of fragmented packets, network address translation, and packet routing. The level of detail of this scheme doesn't include the architecture and operating algorithms of individual components of the microprocessor system, as well as the interface lines between them, commands and control signals. Therefore, the scheme isn't complete, but is sufficient to develop a mathematical model, neglecting (to make things simpler) unimportant secondary factors.

Only permissive rules are considered as filtration ones. The logical structure of the rule set is a linear list. When the first match of a packet to a rule is found, the packet is considered to be processed by the FW. Packets that don't comply with the rules are rejected. One set of filtration rules, implementing the default deny policy, is used. Additional rule sets aren't considered.

Therefore, the traffic filtering process includes the initial processing of the packet and checking if the packet complies with the filtration rules. The packet checking time is considered a random variable distributed according to an exponential law for the initial processing and checking the filtration rules.

The time required to calculate weights and to range the rule set is considered negligible.

Ranging the filtration rule set is considered as ordering the rules in descending order of their weights in accordance with estimates of the parameters of information flows. We suppose that traffic is filtered at the network and transport levels of the reference model of interaction of open OSI systems.

A model with ranging the filtration rules is a complex stochastic system, and, to build it, an aggregative approach was used, which represents the system as an aggregate that has input and output signals. To demonstrate the operation of subsystems, approaches to describing systems adopted in queuing theory, were used [8].

Let's represent the FW model in the form of a system $M(k) = \{Z, L, T, \Phi, G, X, Y\}$ [3], the moment of transition of which from one state to another one is shown in figure 1:

1. $Z = \{\mathbf{z}_0, \ldots, \mathbf{z}_k, \ldots\}$ is a set of system states; $L = (\mu_0, \mu, N, C)$ is set of system parameters; T is a time interval of system operation. Changes in system states occur at discrete time points $t_k^- = t_k - 0$, $t_k \in T$, $k \geqslant 1$; Φ is a system state transition operator, G is a system output

operator; $X = \{\mathbf{x}_0, \ldots, \mathbf{x}_k, \ldots\}$ is a set of input signals entering the system; $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_k, \ldots\}$ is a set of system output signals.

2. μ_0 is the intensity of the service during initial processing of the packet by the network card of the FW; μ is the intensity of the service during the checking stage (whether a packet complies with one filtration rule in a set); C is the system storage capacity; N is the number of rules in the filtration set.
3. $\mathbf{z}_k = (\mathbf{r}_k, \mathbf{d}_k) \in Z$ is the state of the system on the time interval $[t_{k-1}, t_k)$, where $\mathbf{r}_k = \left(r_1^k, \ldots, r_N^k\right)$ is a rule set, in which the r_i^k component is a rule in the i-th position in the set; $\mathbf{d}_k = \left(d_1^k, \ldots, d_N^k\right)$ is a vector of packet servicing times, in which d_i^k corresponds to the processing time for i-type packets on the interval $[t_{k-1}, t_k)$.
4. $\mathbf{x}_k = \left(x_1^k, \ldots, x_N^k\right) \in X$ is the input signal; x_i^k, $i = 1, \ldots, N$ is random value that characterizes the number of i-type packets corresponding to the r_i^k rule, entering the system on the interval $[t_{k-1}, t_k)$.
5. $\mathbf{p}_k = \left(p_1^k, \ldots, p_N^k\right)$ is a vector of rule weights, set in accordance with MLA estimates of the parameters of information flows; the component p_i^k, $i = 1, \ldots, N$ corresponds to the weight of the rule that takes the i-th position on the interval.
6. $\mathbf{y}_k = (q_k, v_k, w_k, u_k) \in Y$ is an output signal, the components of which correspond to the estimates of performance metrics on the interval $[t_{k-1}, t_k)$; q_k corresponds to the average number of packets in a drive, v_k corresponds to the average time needed to service the packet in the system, w_k corresponds to the average waiting time before the start of servicing the packet in the system, and u_k corresponds to the average residence time of the packet in the system.
7. $\Phi\left(L, \mathbf{z}_k, \mathbf{x}_k\right) = \mathbf{z}_{k+1}$. At the time point t_{k+1}^-, this operator calculates the weights of the $\mathbf{p}_k$ rules in accordance with the estimates of the parameters of the $\widehat{\mathbf{x}}_k$ information flows. It also determines the state of the system $\mathbf{z}_{k+1} = (\mathbf{r}_{k+1}, \mathbf{d}_{k+1})$, where the $\mathbf{r}_{k+1}$ vector is calculated by ranging the $\mathbf{r}_k$ rule set according to the $\mathbf{p}_k$ weights, and the $\mathbf{d}_{k+1}$ vector is calculated according to the resulting set $\mathbf{r}_{k+1}$ and intensities μ_0, μ.
8. $G\left(L, \mathbf{z}_k, \mathbf{x}_k\right) = y_{k+1}$. At the time point t_{k+1}^-, this operator determines the performance metrics on the time interval $[t_{k-1}, t_k)$.

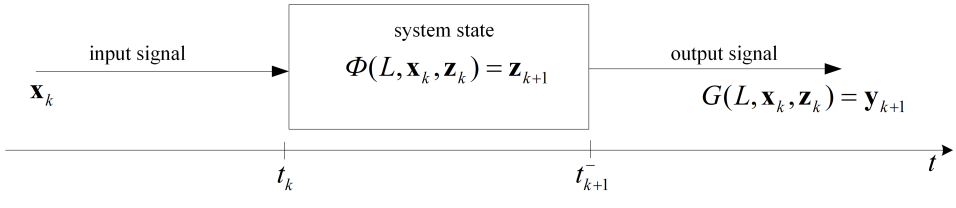

Figure 1. Scheme of the FW model with ranging the filtration rules

Considering the operation of $M(k)$ on the interval $[t_{k-1}, t_k)$, let's present the aggregate in the form of a single-line QS with a storage of limited capacity C,

heterogeneous Poisson incoming flow, and a request service with a distribution function (DF) $B_k(t)$ for the duration of servicing phase-type requests, which depends on the order of filtration rules. The QS receives a request flow that is a superposition of N independent Poisson flows of requests. According to the Basharin-Kendall classification, this QS is designated as $\mathrm{M}_N/\mathrm{PH}/1/\mathrm{C}$ [8].

Representing the system in the form of the QS makes it possible to calculate the performance metrics using an analytical approach for the Poisson flow of requests. Hereinafter, the packets entering the model will be considered as requests, and the QS reflecting the operation of the system on the interval $[t_{k-1}, t_k)$, $k \geqslant 1$ will be designated as $M(k)$.

3. Algorithm for calculating the performance metrics for the exponential distribution of service time

To calculate the performance metrics, let's consider the operation of $M(k)$ on the interval $[t_{k-1}, t_k)$, $k \geqslant 1$. On this interval, the filtration of one batch of packets takes place, there is no ranging for the rule set, and the state vector $\mathbf{z}_k = (\mathbf{r}_k, \mathbf{d}_k)$ remains unchanged.

Let's imagine that a system receives a Poisson flow of requests with intensity $\lambda_{\cdot}(k) = \sum_{i=1}^{N} \lambda_i^k$, which is the sum of independent Poisson flows of requests of N different types. Let's define a rule set $\mathbf{r}_k = \left(r_1^k, \ldots, r_N^k\right)$ in such a way that the i-type request with the λ_i^k intensity will correspond to the r_1^k filtration rule for all $i = 1, 2, \ldots, N$.

Let's define the process of servicing the requests as a sequential transition of phases, starting with the first one. Servicing the request in the 0-th phase corresponds to the stage of the initial processing of FW packets, servicing the request from the 1-st to the N-th phases corresponds to checking the filtration rules. Only one request can be served at a time. If there is no free space in the drive, then the incoming request exits the system without being serviced. If the request matches the rule, its service in the system gets completed; otherwise, the request is transferred to the next phase. After the N-th phase, the service gets completed.

Let's consider the case in which the request service times in each phase are independent of each other and distributed exponentially. The process of request servicing is schematically shown in figure 2.

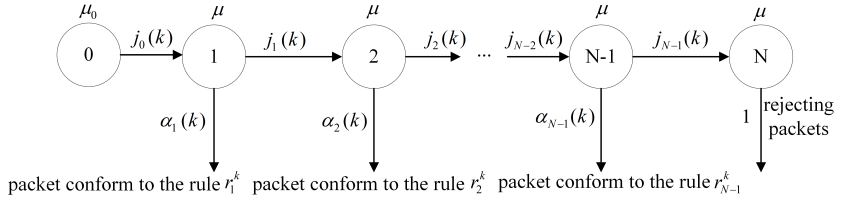

Figure 2. Phase representation of the request servicing process in the firewall model

According to lemma 3 from [8], the probability of a request transition from the i-th phase to the next $i+1$-th phase for the k-th time interval can be

given as follows:

$$j_i(k) = \begin{cases} 1, & i = 0, \\ 1 - \lambda_i(k)/\lambda_{\cdot}(k), & i \in 1, \ldots, N-1. \end{cases} \tag{1}$$

Let's designate $\alpha_i(k) = 1 - j_i(k)$ as the probability of completing the request service in the i-th phase for the k-th time interval. The distribution function (DF) for the time of servicing the request in the QS $M(k)$ will be as follows:

$$B_k(t) = \begin{cases} j_0(k)E(1,\mu_0) + \sum\limits_{i=1}^{N-1} j_i(k)E(i,\mu), & k \in 1, \ldots, N-1, \\ E(N,\mu), & k = N, \end{cases} \tag{2}$$

where $E(i,\mu_0)$ is the Erlang distribution of the i-th order. To analyze the probabilistic and temporal characteristics of the QS, the algorithmic approach proposed by P. Bocharov in [8] was implemented. His main idea is to obtain a solution to the system of global balance equations for the Markov process that describes the system, and to find the parameters of the QS in the form of matrix-recurrent formulas. The use of such an approach makes it possible to effectively calculate the QS parameters.

Let's define a random process (RP) $\{\eta(t), t \geqslant 0\}$ on the set of states $X = \{(0)\ (h,i),\ h = 1, \ldots, C+1,\ i = 0, \ldots, N\}$. The states of the X set have the following meaning. If at some point in time $\eta(t) = 0$, then there are no requests in the system. If $\eta(t) = (h,i)$, then there are h requests in the system, and the serviced request is in the i-th phase. The RP build in such a way is a homogeneous Markov process (MP).

All states of the RP are interconnected, their number is finite and equal to $(C+1)(N+1)+1$. Therefore, according to the ergodic theorem for the MP with a finite set of states [8], the RP $\eta(t)$ is ergodic.

Using a matrix representation, below we give the DF $B_k(t)$ of the request service time. Hereinafter, for brevity, the index k of the time interval of the $M(k)$ aggregate is omitted.

$$\begin{gathered} B_k(t) = 1 - \boldsymbol{\beta}^{\mathrm{T}} e^{\mathrm{M}t}\mathbf{1}, \quad t > 0, \\ \boldsymbol{\beta}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \end{gathered} \tag{3}$$

where the pair of $\left(\boldsymbol{\beta}^{\mathrm{T}}, \mathrm{M}\right)$ is a PH representation of order $N+1$; $\boldsymbol{\beta}^{\mathrm{T}} = (\beta_0, \ldots, \beta_N)$ is the vector of probabilities of sending a request for service to phases $0, 1, \ldots, N$ at the time point t_k, the component β_i, $i \in 0, \ldots, N$ corresponds to the probability of starting the request service in the i-th phase at the time point t_k; and M is an infinitesimal matrix that has the following form:

$$\mathbf{M} = \begin{bmatrix} -\mu_0 & \mu_0 & 0 & 0 & 0 & 0 \\ 0 & -\mu & j_1\mu & 0 & 0 & 0 \\ 0 & 0 & -\mu & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & j_{N-2}\mu & 0 \\ 0 & 0 & 0 & 0 & -\mu & j_{N-1}\mu \\ 0 & 0 & 0 & 0 & 0 & -\mu \end{bmatrix}. \tag{4}$$

Vector $\boldsymbol{\beta}$, in accordance with the fact that the request service in the considered QS always starts from the zero phase, has the following form:

$$\boldsymbol{\beta}^{\mathrm{T}} = (1, 0 \ldots 0). \tag{5}$$

Let's use the following designations for the probabilities of $\{\eta(t), t \geqslant 0\}$ process states:

1. $p_0 = P\{\eta(t) = 0,\, t \geqslant 0\}$ is the stationary probability of the absence of requests in the system.
2. $p_{i,j} = P\{\eta(t) = (i, j),\, t \geqslant 0\}$ is the stationary probability of servicing the request in the j-th phase and presence of i requests within the system.
3. $\mathbf{p}_h^{\mathrm{T}} = (p_{h,0}, p_{h,1}, \ldots, p_{h,N})$, $h = 1, \ldots, C+1$ is a vector of stationary probabilities.
4. p_h is the stationary probability of presence of h requests within the system.

So, the system of global balance equations in matrix form (for the QS), which describes the process of the FW operation and takes filtration into account, is as follows:

$$\begin{cases} -\lambda_{\cdot} p_0 + \mathbf{p}_1^{\mathrm{T}}\boldsymbol{\mu} = 0, \\ \mathbf{p}_1^{\mathrm{T}}(-\lambda_{\cdot}\mathbf{I} + \mathbf{M}) + \lambda_{\cdot}\boldsymbol{\beta}^{\mathrm{T}} p_0 + \mathbf{p}_2^{\mathrm{T}}\boldsymbol{\mu}\boldsymbol{\beta}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \\ \mathbf{p}_h^{\mathrm{T}}(-\lambda_{\cdot}\mathbf{I} + \mathbf{M}) + \lambda_{\cdot}\mathbf{p}_{h-1}^{\mathrm{T}} + \mathbf{p}_{h+1}^{\mathrm{T}}\boldsymbol{\mu}\boldsymbol{\beta}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \quad h = 2, 3 \ldots, C, \\ \mathbf{p}_{C+1}^{\mathrm{T}}\mathbf{M} + \lambda_{\cdot}\mathbf{p}_C^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \end{cases} \tag{6}$$

where $\boldsymbol{\mu}^T = (0, \alpha_1\mu, \ldots, \alpha_{N-1}\mu, \mu)$ is the vector of intensities for the completion of request servicing in the QS of $N+1$ dimension; and $\mathbf{I}$ is the identity matrix of $N+1$ dimension/degree.

Similarly to [8], for the convenience of solving the system of equations (SoE), let's introduce an extra vector $\tilde{p}_h$:

$$\tilde{p}_h^T = \frac{\mathbf{p}_h^T}{p_0}. \tag{7}$$

The solution of the SoE (6) allows us to calculate the performance metrics for the stationary operating mode of the QS (see Algorithm 1).

Algorithm 1. Algorithm for calculating efficiency metrics for exponential distribution of service time Algorithm for calculating the performance metrics for the exponential distribution of service time.

Step 1
Calculate matrices and vectors:

$$\mathbf{D} = -\lambda_{\cdot}\mathbf{I} + \mathbf{M}, \tag{8}$$

$$\beta(\lambda) = 1 + \lambda_{\cdot}\boldsymbol{\beta}^T\mathbf{D}^{-1}\mathbf{1}, \tag{9}$$

$$\widetilde{\mathbf{M}}^{-1} = \mathbf{D}^{-1} - \frac{\lambda_{\cdot}\mathbf{D}^{-1}\mathbf{1}\boldsymbol{\beta}^T\mathbf{D}^{-1}}{\beta(\lambda)}, \tag{10}$$

$$\mathbf{W} = -\lambda_{\cdot}\widetilde{\mathbf{M}}^{-1}, \tag{11}$$

$$\mathbf{W}_R = -\lambda_*\mathbf{M}^{-1}, \tag{12}$$

$$\boldsymbol{\omega}_0^T = \frac{-\lambda_{\cdot}\boldsymbol{\beta}^T\,\mathbf{D}^{-1}}{\beta(\lambda_{\cdot})}. \tag{13}$$

Step 2
Calculate the probabilities $\tilde{p}_h$:

$$\tilde{p}_h^T = \begin{cases} \boldsymbol{\omega}_0^T\mathbf{W}^{h-1}, & h = 1, 2, \ldots, C, \\ \boldsymbol{\omega}_0^T\mathbf{W}^{C-1}\mathbf{W}_R, & h = C+1. \end{cases} \tag{14}$$

Step 3
Calculate p_0, using the normalization conditions for the system of global balance equations and formula (7):

$$p_0 = \left(1 + \sum_{h=1}^{C+1}\tilde{p}_h\right)^{-1}. \tag{15}$$

Step 4
Calculate vector:

$$\mathbf{p}_h^T = p_0\tilde{p}_h^T, \quad h = 1, 2, \ldots, C+1. \tag{16}$$

The calculated vector $\mathbf{p}_h^T$ is the desired matrix-geometric solution for the system of global balance equations (6).

Step 5
Calculate stationary probabilities p_h:

$$p_h = \mathbf{p}_h^T\mathbf{1}, \quad h = 1, 2, \ldots, C+1. \tag{17}$$

Step 6
Using p_h and formulas (18)–(23), calculate the performance metrics for the QS stationary operating mode.

Average number of requests in the QS:

$$l = \sum_{h=1}^{C+1} hp_h. \tag{18}$$

Average queue length:

$$q = \sum_{h=2}^{C+1} (h-1)p_h. \tag{19}$$

Probability of losing the requests:

$$\pi = \vec{p}_{C+1}^{T}\vec{1}. \tag{20}$$

Average residence time:

$$u = l/\lambda_{\cdot}(1-\pi). \tag{21}$$

Average waiting time for service:

$$w = q/\lambda_{\cdot}(1-\pi). \tag{22}$$

Average service time:

$$v = u - w. \tag{23}$$

4. Evaluation of firewall performance metrics

The initial data selected for calculating the performance metrics are as follows: system storage capacity $C = 10$; max number of filtration rules in a set $N = 1500$; initial packet processing time $\mu_0^{-1} = 2.7 \cdot 10^{-5}$ [ms]; time for checking one rule $\mu^{-1} = 5 \cdot 10^{-5}$ [ms]. The packet processing times were taken from papers [6, 7]. The incoming flow is the sum of N independent Poisson flows of requests. The request service time is exponential. To calculate the performance metrics, a program code has been developed in the MATLAB system language.

To check the correctness of the model, the following graphs were constructed:

1. Graphs illustrating the dependence of the performance metrics on the total request flow. The intensities of requests entering the system increase at time points $t_k \in T, k \geqslant 1$; the initial value of the total flow intensity is $\lambda_{\cdot} = 25$ [ms^{-1}]. The number of filtration rules is constant $N = 100$.
2. Graphs illustrating the dependence of the performance metrics on the number of filtration rules. The number of rules increases at time points t_k. The total intensity of the request flow remains constant $\lambda_{\cdot} = 50$ [ms^{-1}].

The values of the performance metrics depend on the types and intensities of the corresponding incoming packets, the rule set and structural parameters of the system. Obviously, the maximum values (given the same system parameters) will be observed when the packets match the last filtration rule, and the minimum ones will be obtained when the packets match the first rule.

That's why, when constructing graphs illustrating the performance metrics, we considered the following cases:

1. Graphs of the performance metrics constructed for the case when incoming requests comply only with the first rule in the rule set, and the time for servicing the request is the least possible.

2. Graphs of the performance metrics constructed for the case when incoming requests comply with the last rule in the rule set, and the time for servicing the request is the largest possible.
3. Graphs of the performance metrics constructed for the case when incoming requests comply with a random rule.

Further in the description, despite the fact that in all three cases the service time is random, for convenience, the graphs will be called as follows: graph of the first rule, graph of the last rule, graph of the random rule, respectively.

The average service time (see figures 3–4) depends on the types of incoming packets and the order of filtration rules and doesn't depend on the total intensity of received requests and the operating mode of the QS. Therefore, when the total request flow increases, the average service time remains constant. Meanwhile, if the number of filtration rules increases, the average service time grows linearly.

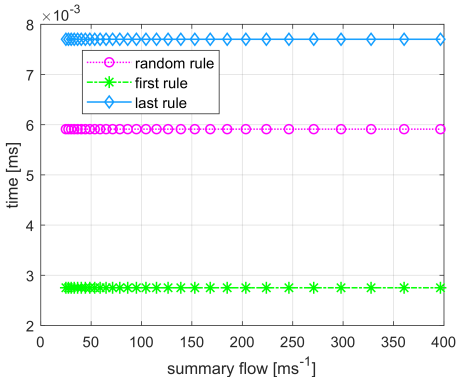

Figure 3. Average service time (constant number of rules)

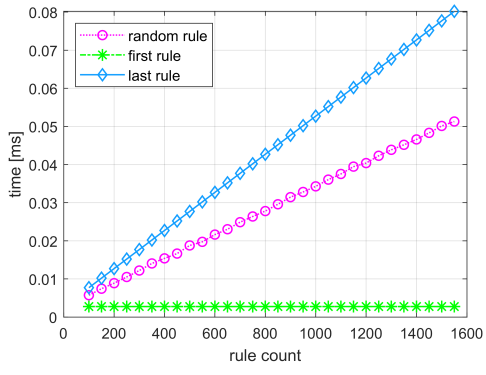

Figure 4. Average service time (constant total intensity of requests received)

Thus, the efficiency of reducing the average service time when ranging the filtration rule set for constant values of the probability of receiving of each type of request will grow with the increase of the rule set and won't depend on the total flow intensity.

The figure 5 shows the dynamics of changes in the average residence time of requests with an increase in the total intensity of the incoming request flow. The average residence time of requests increases with a growth in the flow intensity. For values of the total request flow [ms^{-1}], the type of dependence for the graph of the random rule will change, which corresponds to functioning of the QS in overload mode. Such an overload in the system doesn't result in an unlimited increase in the QS parameters due to the limited storage capacity. The value of the average residence time tends to 0.056 [ms]. At the same time, the difference between the values of the average residence time on the graph of the random rule and the graph of the last rule indicates the possibility of obtaining a significant reduction in the average residence time when ranging the filtration rule set.

For larger rule sets, the difference between the values of the average residence time of requests for the graph with a minimum service time and the graph with a random service time increases with the number of rules (see figure 6).

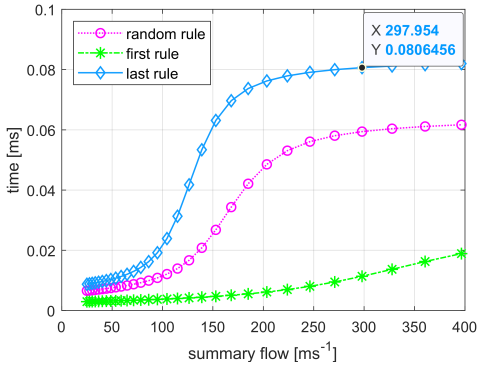

Figure 5. Average residence time in the system (constant number of rules)

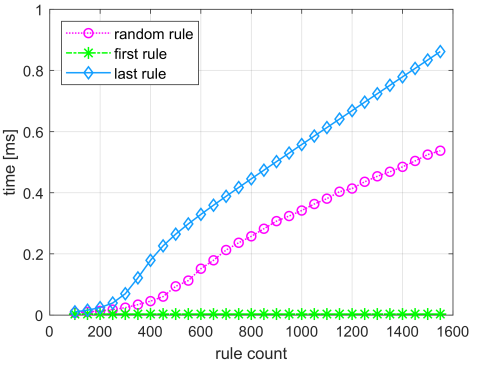

Figure 6. Average residence time in the system (constant total intensity of requests received)

This complies with the peculiarities of the functioning of the FW, since the search time for a rule that matches the filtered data is proportional to the number of checked rules, and indicates the advisability of ranging the filtration rule set. Overload in the system doesn't lead to an unlimited increase in the QS parameters.

Thus, the value of the decrease in the average residence time when ranging the filtration rule set will grow with an increase of the set itself. Also, it won't depend on the intensity of the total request flow starting from the moment when the system gets overloaded.

Figures 7–8 show graphs of changes in the average queue length. Increase of the rule set results in faster filling of the storage, which corresponds to the logic of the FW, because checking a rule set of a large dimension takes more time.

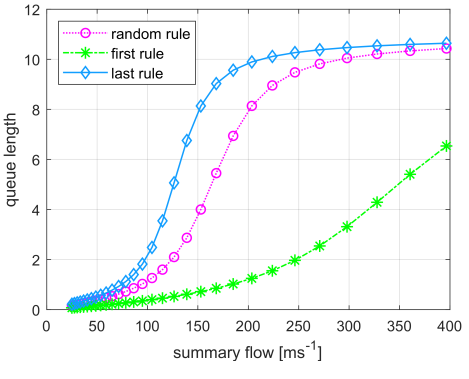

Figure 7. Average queue length (constant number of rules)

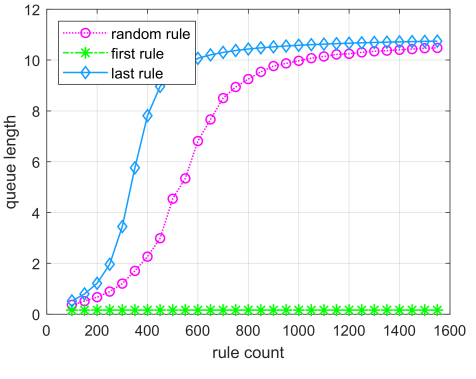

Figure 8. Average queue length (constant total intensity of requests received)

When the QS gets significantly overloaded, the average queue length will be equal to the storage capacity. And the average queue length in the graph of the random rule will approach the values given in the graph of the last rule (see figure 7).

Therefore, the efficiency of ranging the rule set (in order to reduce the queue length) will grow with increasing the load on the system until the system operates in the mode of losing the requests.

5. Conclusion

The obtained estimates of the firewall performance metrics allow us to draw a conclusion on the adequacy of the built analytical model of the FW. We can also draw a conclusion about the possibility to increase the firewall performance by implementing the method for ranging the rule set.

Acknowledgments

This paper has been supported by the RUDN University Strategic Academic Leadership Program.

References

[1] A. Y. Botvinko and K. E. Samouylov, "Evaluation of firewall performance when ranging a filtration rule set," *Discrete and Continuous Models and Applied Computational Science*, vol. 29, no. 3, pp. 230–241, 2013. DOI: 10.22363/2658-4670-2021-29-3-230-241.

[2] A. Y. Botvinko and K. E. Samouylov, "Firewall simulator development for performance evaluation of ranging a filtration rules set," *Distributed Computer and Communication Networks: Control, Computation, Communications. DCCN 2022. Lecture Notes in Computer Science. Lecture Notes in Computer Science*, vol. 13766, no. 3, pp. 221–229, 2022. DOI: 10.1007/978-3-031-23207-7_15.

[3] A. Y. Botvinko and K. E. Samouylov, "Firewall simulation model with filtering rules ranking," *Distributed Computer and Communication Networks: Control, Computation, Communications. DCCN 2020. Communications in Computer and Information Science*, vol. 1337, pp. 533–545, 2020. DOI: 10.1007/978-3-030-66242-4_42.

[4] V. Katkovnik, *Non-parametric data identification and smoothing: local approximation method [Neparametricheskaya identifikaciya i sglazhivanie danny'x: metod lokal'noj approksimacii]*. The science. Main editorial office of physical and mathematical literature Publ., 1985, 336 pp., in Russian.

[5] W. Hardle, *Applied nonparametric regression*. Cambridge: Cambridge university press, 1990, 349 pp.

[6] M. Cheminod, L. Durante, L. Seno, and A. Valenzano, "Performance evaluation and modeling of an industrial application-layer firewall," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2159–2170, 2018. DOI: 10.1109/TII.2018.2802903.

[7] K. Salah, K. Elbadawi, and R. Boutaba, "Performance modeling and analysis of network firewalls," *IEEE Transactions on network and service management*, vol. 9, no. 1, pp. 12–21, 2011. DOI: 10.1109/TNSM.2011.122011.110151.

[8] P. P. Bocharov and A. V. Pechenkin, *Queuing theory [Teoriya massovogo obsluzhivaniya]*. Moscow: RUDN, 1995, 529 pp., in Russian.

For citation:
A. Y. Botvinko, K. E. Samouylov, Evaluation of firewall performance metrics with ranging the rules for Poisson incoming packet flow and exponential filtering time, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 345–358. DOI: 10.22363/2658-4670-2023-31-4-345-358.

Information about the authors:

Botvinko, Anatoly Yu. — Candidate of Physical and Mathematical Sciences, assistant professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: botvinko-ayu@rudn.ru, ORCID: https://orcid.org/0000-0003-1412-981X)

Samouylov, Konstantin E. — Professor, Doctor of Technical Sciences, Head of the Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: samuylov-ke@rudn.ru, ORCID: https://orcid.org/0000-0002-6368-9680)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-4-345-358
EDN: DYDLCY

Оценка показателей эффективности межсетевого экрана с ранжированием правил для пуассоновского входящего потока пакетов и экспоненциального времени фильтрации

А. Ю. Ботвинко, К. Е. Самуйлов

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

Аннотация. Данная статья является развитием ряда работ по разработке моделей и методов ранжирования правил фильтрации для предотвращения снижения производительности межсетевого экрана, обусловленной использованием последовательной схемы проверки соответствия пакетов правилам, неоднородностью и изменчивостью сетевого трафика. В статье приведено описание математической модели межсетевого экрана в виде сложной системы и системы массового обслуживания с дисциплиной обслуживания заявок фазового типа, формализующей процесс фильтрации сетевого трафика с функциональной возможностью ранжирования правил. Целью моделирования является получение оценок основных показателей эффективности межсетевого экрана для различных сценариев поведения сетевого трафика, а также оценка повышения производительности за счёт ранжирования набора правил фильтрации. Вычисление оценок показателей эффективности МЭ проводится аналитическим способом для пуассоновского потока заявок. На основании анализа результатов моделирования сделаны выводы об эффективности ранжирования правил фильтрации для повышения производительности межсетевых экранов для сценариев трафика, близких к реальным.

Ключевые слова: межсетевой экран, ранжирование правил фильтрации, сетевой трафик, фазовое обслуживание, система массового обслуживания

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 359–374
http://journals.rudn.ru/miph

Research article

UDC 314.422:574.34
PACS 07.05.Tp, 89.65.Cd,
DOI: 10.22363/2658-4670-2023-31-4-359-374
EDN: FZWSUR

Demographic indicators, models, and testing

Gregory A. Shilovsky[1, 2], **Alexandr V. Seliverstov**[1], **Oleg A. Zverkov**[1]

[1] *Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), 19 Bolshoy Karetny per., bldg. 1, Moscow, 127051, Russian Federation*
[2] *Lomonosov Moscow State University, 1 Leninskie Gory, bldg. 12, Moscow, 119991, Russian Federation*

(received: October 17, 2023; revised: December 1, 2023; accepted: December 29, 2023)

Abstract. The use of simple demographic indicators to describe mortality dynamics can obscure important features of the survival curve, particularly during periods of rapid change, such as those caused by internal or external factors, and especially at the oldest or youngest ages. Therefore, instead of the generally accepted Gompertz method, other methods based on demographic indicators are often used. In human populations, chronic phenoptosis, in contrast to age-independent acute phenoptosis, is characterized by rectangularization of the survival curve and an accompanying increase in average life expectancy at birth, which can be attributed to advances in society and technology. Despite the simple geometric interpretation of the phenomenon of rectangularization of the survival curve, it is difficult to notice one, detecting changes in the optimal coefficients in the Gompertz–Makeham law due to high computational complexity and increased calculation errors. This is avoided by calculating demographic indicators such as the Keyfitz entropy, the Gini coefficient, and the coefficient of variation in lifespan. Our analysis of both theoretical models and real demographic data shows that with the same value of the Gini coefficient in the compared cohorts, a larger value of the Keyfitz entropy indicates a greater proportion of centenarians relative to average life expectancy. On the contrary, at the same value of the Keyfitz entropy, a larger value of the Gini coefficient corresponds to a relatively large mortality at a young age. We hypothesize that decreases in the Keyfitz entropy may be attributable to declines in background mortality, reflected in the Makeham term, or to reductions in mortality at lower ages, corresponding to modifications in another coefficient of the Gompertz law. By incorporating dynamic shifts in age into survival analyses, we can deepen our comprehension of mortality patterns and aging mechanisms, ultimately contributing to the development of more reliable methods for evaluating the efficacy of anti-aging and geroprotective interventions used in gerontology.

Key words and phrases: lifespan, demographic indicator, Keyfitz entropy, Gini coefficient, coefficient of variation, phenoptosis, aging, Gompertz law

© Shilovsky G. A., Seliverstov A. V., Zverkov O. A., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

1. Introduction

One of the fundamental tasks of biodemography is the creation and computational verification of probabilistic models to explain the difference in lifespan and response to geroprotectors in different species [1–6]. To optimize the model, parallel calculation of demographic indicators for many parameter values is possible. As a result of comparing the calculated values with known demographic indicators, one can find not only the optimal values of the model parameters, but also the sensitivity of the values of demographic indicators to changes in these parameters. Of practical importance may be the reconstruction of changes in life expectancy in evolution, as well as the prediction of the effectiveness of anti-aging drugs and geroprotectors [7, 8].

A study of mortality fluctuations allows one to assess both the quality and comparability of mortality statistics. Other related problems are comparison of changes in the life expectancy at the national level with the changes in individual regions, identification of regions with high or low values of the life expectancy, estimation of the contribution of individual age groups and causes of death to the regional differences in the life expectancy, and determination of characteristics of mortality and causes of death for individual groups or different regions [9, 10].

The COVID-19 pandemic has revealed significant gaps in the coverage and quality of existing international and national statistical monitoring systems. Ensuring prompt availability of accurate and comparable data in each country for an adequate response to unexpected epidemiological threats is a very challenging task. The interest in studying associations between mortality oscillations and fluctuation of economic conditions has been rekindled recently because mortality is characterized with periodic oscillations [11, 12].

In general, modeling can be used for simulation of future behavior of demographic processes based on the available data in order to reveal main and additional rhythms. In particular, construction of wavelet spectrograms provides a possibility to calculate the matrix for synchronization value and synchronicity (simultaneous occurrence), syn-phase behavior (phase coincidence), and coherence (interconnection) of the investigated parameters of studied biorhythms. Statistical significance of the rhythms is evaluated through multiple random permutations of levels of the initial temporal series [13]. Decomposition into seasons and trends using the Loess approach (STL) is used for analysis of seasonal fluctuations of mortality risk, medical care expenditure, and even hospitalization levels. The STL method expands longitudinal data into the long-term trend, seasonal variations, and remaining variations not associated with the long-term trend or seasonal variations [14]. The long-term trend in the STL method reflects a number of possible external factors that change gradually with time [15].

During the last 150 years, the decrease in the seasonal fluctuations of mortality has facilitated an increase in the life expectancy [16]. New methods based on the analysis of time-dependent variability, trends, and interactions of numerous physiological and laboratory parameters, for which machine learning and artificial intelligence could be applied, will help to establish whether the dynamic regularities observed in large epidemiological studies have significance for the risk profile of an individual patient [17]. From the gerontological point of view, studying mortality fluctuations allows to switch

from investigating the effects of biorhythms on the development of acute and chronic phenoptosis to the elucidating the patterns of determined mortality rhythms [10].

There are also developed formal demographic measures to examine the complex relationships between the total life expectancy of two peers at birth, the proportion of their life that they can expect to live, and longevity [18]. A modification of the Gini coefficient is the Drewnovsky index, which is a measure of equality [19]. So, Aburto *et al.* simulated scenarios for improving mortality in a Gompertz model and showed that the new index can serve as an indicator of the shape of the mortality structure. The proposed method allows us to identify trends in lifespan changes in both humans and other species.

Our aim is to compare measures of shape of the survival curve. These measures should be dimensionless. They depend on the shape of the survival curve only.

2. Preliminaries

2.1. The Gompertz law

The Gompertz law is a probabilistic model of mortality that describes well the mortality of people aged 20 to 65 or up to 80 years. This law was proposed in the pioneering work of B. Gompertz and was originally used to assess risks in life insurance [20]. Problems with assessing the aging process as an increase in the probability of death (the number of deaths in one age interval) have existed for a long time [20, 21].

If the probability of death of an organism depended entirely on the level of disruption that increases with age, then the mortality rate of multicellular organisms should increase with age, regardless of the position of the species on the evolutionary tree. However, large differences in mortality dynamics across species have been found (increasing, constant, decreasing, convex, and concave mortality trajectories in both long-lived and short-lived species) [4, 22–25]. Possible mechanisms for the emergence of such diversity in evolution are actively discussed [7, 26, 27]. Despite the discussion of amendments to this law [28–30], its main idea has remained unchanged for almost two hundred years: the law determines the dependence of the conditional probability density of death on age.

Let us denote by $\mu(a)$ the conditional probability density for each individual to die at age a, provided that he survived. The value of $\mu(a)$ is called the strength of mortality. We will assume that the function $\mu(a)$ is piecewise smooth. According to the Gompertz–Makeham law, starting from some age $a_{\min}$ and up to age $a_{\max}$, this function is of the type $\mu(a) = \alpha + \beta \exp(\gamma a)$, where α, β, and γ are some coefficients that do not depend on the age a, but may depend on external conditions. Here α is the probability density of an accidental death regardless of age. Originally $\alpha = 0$ was assumed in the Gompertz law. For people $a_{\min} \approx 30$ years and $a_{\max} \approx 80$ years.

In theoretical calculations, we pass to a continuous change in age. In practice, for humans, the unit of time is usually one year, less often five years, and for species with a short lifespan such as nematodes or fruit flies, it is one day. The product $\mu(a)\Delta a$ is only approximately equal to the probability of dying between a and $a + \Delta a$.

The probability of surviving to age a is equal to the survival function

$$\ell(a) = \exp\left(-\int_0^a \mu(\tau)d\tau\right).$$

Of course, $\ell(0) = 1$ because only those born are taken into account.

2.2. The Makeham term

Accounting for accidental death, the conditional probability density of which does not depend on age, leads to an additional term called the Makeham term. The new conditional death probability density is $\mu(a) = \exp(-s)u + \exp(ra - s)$. Such an amendment to the Gompertz law was proposed by W. M. Makeham [29]. Usually the value of u is nonnegative, but negative values $u > -1$ can also be considered, corresponding to an accidental escape from death. Such an amendment corresponds to multiplying the original survival function $\ell(a)$ by the factor $\exp(-\exp(-s)ua)$.

For some large mammals such as the lion *Panthera leo*, the European roe deer *Capreolus capreolus*, the red deer *Cervus elaphus*, the chamois *Rupicapra rupicapra*, the sheep *Ovis aries*, and the yellow-bellied marmot *Marmota flaviventris* as well as birds the Bali myna *Leucopsar rothschildi* and the sparrowhawk *Accipiter nisus*, the conditional probability density of death has a non-zero minimum [4], which suggests that the u correction is nonzero. It can be concluded that such dynamics of mortality is typical for large mammals and some birds either having no enemies in nature like a hawk or kept in zoos.

2.3. The Keyfitz entropy H

Let us consider a demographic indicator called the Keyfitz entropy. This concept was introduced by Canadian demographer Nathan Keyfitz [31]. The Keyfitz entropy characterizes the deviation of the survival curve from a non-increasing step function that is equal to either 0 or 1.

$$\ell_{\text{rect}}(a) = \begin{cases} 1, & a \leqslant e_o, \\ 0, & a > e_o. \end{cases}$$

Let us denote the life disparity by

$$e^{\dagger} = -\int_0^{\infty} \ell(\tau) \ln \ell(\tau) d\tau.$$

The life expectancy at birth is denoted by

$$e_o = \int_0^\infty \ell(\tau) d\tau.$$

The Keyfitz entropy is equal to

$$H = \frac{e^\dagger}{e_o}.$$

The Keyfitz entropy is close to zero when almost everyone dies at the same age, no matter what that age is. In other words, rectangularization of the survival curve leads to the value of $e^\dagger$ vanishing. On the other hand, the Keyfitz entropy decreases even more as the life expectancy e_o increases.

Surprisingly, there is such an age threshold that preventing death before reaching this threshold leads to a decrease in the Keyfitz entropy, and after reaching the age threshold, to its increase [32]. The development of society and scientific and technological progress leads to an increase in life expectancy over time [33]. But the lifespan of people with accurately confirmed age rarely exceeds 116 years [28]. Unfortunately, reports of centenarians who lived for more than 120 years are not confirmed or were refuted upon further verification.

Another observation is the relationship between life expectancy and the Keyfitz entropy [34, 35]. An increase in the standard of living of the population leads to a simultaneous decrease in $e^\dagger$ and an increase in e_o, which leads to a decrease in the Keyfitz entropy. At the same time, an increase in e_o looks quite natural, while a decrease in $e^\dagger$ *a priori* is less obvious, but is in good agreement with the phenoptosis hypothesis [10, 36].

Continuing the former research [35, 37–39], we compare some demographic indicators, including the Keyfitz entropy, calculated for different aging models.

2.4. The Gini coefficient G

Another demographic indicator is the Gini coefficient

$$G = 1 - \frac{1}{e_o} \int_0^\infty \ell^2(\tau) d\tau.$$

It also vanishes on a non-increasing step function that takes only two values 0 or 1. This indicator was proposed in 1912 by the demographer Corrado Gini [40]. It is used in demography by other authors too [41–43].

2.5. The coefficient of variation CV_{LS}

The coefficient of lifespan variation CV_{LS} is also used in demography [10, 37–39, 44]. The formula for calculating the coefficient of variation explicitly includes the first derivative of the survival function, which is equal to the

product of the survival function and the conditional death probability density $\mu(a)$. This derivative is usually called the distribution of deaths.

2.6. Integrals

We use the SymPy library to calculate the integrals in the considered examples. In fact, only some integrals are expressed in terms of elementary functions. Therefore, not only symbolic computing, but also numerical methods are used.

3. Results

3.1. The Keyfitz entropy and Gini coefficient in comparison

Both demographic indicators H and G are expressed in terms of the survival function by similar formulas. Both Keyfitz entropy and Gini coefficient measure the difference between the survival function and a non-increasing step function. However, these indicators differ significantly in their stability under changes in the survival function [39].

Let us make a transformation, taking into account the expansion of the natural logarithm in a series

$$\ln x = x - 1 - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \cdots .$$

So,

$$e^\dagger = e_o - e_n + \int_0^\infty \ell(\tau)\big(\ell(\tau)-1\big)^2 \left(\frac{1}{2} - \frac{1}{3}\big(\ell(\tau)-1\big) + \cdots\right) d\tau.$$

The difference is equal to

$$H - G = \frac{1}{e_o}\int_0^\infty \ell(\tau)\big(\ell(\tau)-1\big)^2 \left(\frac{1}{2} - \frac{1}{3}\big(\ell(\tau)-1\big) + \frac{1}{4}\big(\ell(\tau)-1\big)^2 - \cdots\right) d\tau.$$

For small ages, the integrand is small, since the difference $\ell - 1$ is close to zero. For sufficiently large ages, it is also small, since ℓ does not increase and must tend to zero for the integral e_o to converge. However, here the integrand tends to zero only at about the same rate as the function ℓ itself.

When the survival function $\ell(a)$ is close enough to a non-increasing step function $\ell_{rect}(a)$ the difference $H - G$ is mainly determined by the behavior of the survival function near e_o and at large values of age. However, it depends little on the properties of this function at small ages.

For $\ell_{rect}(a)$, both Keyfitz entropy and Gini coefficient vanish. But unlike the Gini coefficient, the Keyfitz entropy can be arbitrarily large on survival functions close to $\ell_{rect}(a)$. In fact, the graph of the function $x\ln(x)$ has a vertical tangent at $x = 0$. So, the first derivative of this function $(x\ln(x))' = \ln(x) + 1$ tends to negative infinity $-\infty$ in the limit $x \to +0$.

There is a sequence of monotonically nonincreasing functions ℓ_k such that the sequence ℓ_k converges to the limit function ℓ, but the sequence of values of the Keyfitz entropy $H[\ell_k]$ does not converge to the value $H[\ell]$. By convergence we mean pointwise convergence almost everywhere, i. e., except for the set of points of measure zero, the value of the function ℓ at a point is equal to the limit of the values of the functions ℓ_k at the same point. Informally, such survival functions ℓ_k correspond to a situation when almost everyone dies early at the same age (hence, e_o is small), but a tiny part of long-lived individuals, tending to zero with increasing index k, live extremely long. By choosing the ratio between the proportion of centenarians and the maximum life expectancy, one can achieve an increase in the Keyfitz entropy.

Example 1. For sufficiently large indices $k > \ln(e_o)$, let the value of the survival function be

$$\ell_k(a) = \begin{cases} 1, & 0 \leqslant a \leqslant e_o - 1, \\ \exp(-k), & e_o - 1 < a < e_o - 1 + \exp(k), \\ 0, & a \geqslant e_o - 1 + \exp(k). \end{cases}$$

Then for large indices k the life expectancy is equal to the previously chosen number e_o and $e^\dagger = k$. Therefore, the Keyfitz entropy $H[\ell_k]k/e_o$ tends to infinity as $k \to \infty$. However, in the limit at $k \to \infty$, the survival function ℓ_k approaches $\ell_{rect}(a)$, i. e., there is rectangularization of the survival curve. The limit survival function ℓ is equal to one for ages up to e_o and zero for larger ages. Obviously, at every point except e_o the function $\ell \ln \ell$ vanishes. Hence, the Keyfitz entropy $H[\ell]$ vanishes.

In the considered example, when passing to the limit, the expected lifespan changes abruptly. However, by increasing the absolute value, its relative change can be made arbitrarily small. On the contrary, the sequence of Gini coefficients $G[\ell_k]$ tends to zero, i. e., it converges to the Gini coefficient of the limit function.

3.2. Generalized Gini coefficients

There is an obvious generalization of the Gini coefficient. For a number $p > 1$, let us define the generalized Gini coefficient of order p

$$G_p = 1 - \frac{1}{e_o} \int_0^\infty \ell^p(\tau) d\tau.$$

Of course, G_2 is the same as the Gini coefficient G.

For any value of $p > 1$, G_p vanishes on a nonincreasing step function of the type $\ell_{rect}(a)$ that takes only two values 0 or 1. To study the properties of the tail of the distribution, i. e., the presence of centenarians, $1 < p < 2$ are of interest, for example, $p = 3/2$. For the aging model $\mu(a) = \exp(a - 1 - s)$, where s is a parameter, both indices $G_{3/2}$ and G correlate with each other. On the other hand, the ratio of these two indicators differs from a constant.

3.3. Demographic indicators calculated by summation

In practice, integrals are replaced by finite sums since real lifespan is bounded and age is measured discretely. With a sufficiently large sample, the Gini coefficient is resistant to small errors, in particular, associated with the inevitable difficulties in determining the age. Refining the step w of age change leads to sharp changes in the first derivative of the survival function, which is included in the formula for calculating the coefficient of variation of life expectancy. However, both Keyfitz entropy H and Gini coefficient G depend only on the survival function itself. Therefore, step refinement does not spoil, but only refines the calculation of H and G. We consider the calculation of both Keyfitz entropy and Gini coefficient for the conditional death probability density $\mu(a) = \exp(a)$, which corresponds to the Gompertz law, for different values of the step w. For small step values, the result differs little from the result based on integration. The exact values are equal to $H = 0.68$ and $G = 0.39$, respectively. The summation was carried out up to the age of 100 with an average life expectancy $e_o = 0.60$ (refer to table 1).

Table 1

Both Keyfitz entropy H and Gini coefficient G depend on the step length w

Step w	The Keyfitz entropy H	The Gini coefficient G
0.001	0.68	0.39
0.01	0.67	0.39
0.1	0.62	0.36
1	0.27	0.13

As the step increases, the values of demographic indicators calculated by summation decrease. Such a decrease can be confused with the approximation of the survival curve to a rectangularized one, but this is only the result of a calculation error.

3.4. Example: regime-change aging

Let us consider a one-parameter family of functions μ that do not explicitly depend on time, where the parameter p is positive

$$\mu(a) = \begin{cases} \exp(a - 1 - p), & a \leqslant 1, \\ \exp(-p), & a \geqslant 1. \end{cases}$$

At small ages, aging occurs according to Gompertz; starting from age 1 (some threshold), aging does not depend on age. Here, the unit of age is conditional, and the model itself is not based on real demographic data. The survival function is

$$\ell(a) = \begin{cases} \exp(\exp(-1 - p) - \exp(a - 1 - p)), & a \leqslant 1, \\ \exp(\exp(-1 - p) - a\exp(-p)), & a \geqslant 1. \end{cases}$$

Calculations show that as the parameter p increases, the Keyfitz entropy H, the coefficient of variation CV_{LS}, and the Gini coefficient G increase as well (refer to table 2).

Table 2

Regime-change aging

p	e_o	H	CV_{LS}	G
0.1	1.421518092	0.8094109048	0.8113722122	0.4171815872
0.2	1.542122578	0.8193004266	0.8208801324	0.4205427489
0.3	1.674623382	0.8292751898	0.8305383031	0.4240996359
0.4	1.820323302	0.8392299859	0.8402332055	0.4277956999
0.5	1.980660797	0.8490722655	0.8498641749	0.4315779574
0.6	2.157223890	0.8587224607	0.8593440475	0.4353979373
0.7	2.351765625	0.8681138602	0.8685992389	0.4392123237
0.8	2.566221236	0.8771921308	0.8775693578	0.4429832962
0.9	2.802727165	0.8859145778	0.8862064852	0.4466786320
1.0	3.063642166	0.8942492046	0.8944741982	0.4502716046
1.1	3.351570657	0.9021736551	0.9023464481	0.4537407212
1.2	3.669388575	0.9096740879	0.9098063557	0.4570693552
1.3	4.020271956	0.9167440522	0.9168449959	0.4602452907
1.4	4.407728588	0.9233833701	0.9234601972	0.4632602424
1.5	4.835632967	0.9295970938	0.9296554211	0.4661093436
1.6	5.308264979	0.9353945249	0.9354387066	0.4687906562
1.7	5.830352620	0.9407883361	0.9408217340	0.4713046910
1.8	6.407119251	0.9457937809	0.9458189796	0.4736539663
1.9	7.044335810	0.9504280005	0.9504469809	0.4758426085
2.0	7.748378494	0.9547094368	0.9547237110	0.4778759915

3.5. Slow aging models

Let us consider models of asymptotically slower aging than the Gompertz law provides. It is natural to call such models of aging sub-Gompertzian. The simplest model corresponds to the age-independent positive constant $\mu(a) = m$. Such a model of aging is realized, for example, in the hydra *Hydra magnipapillata*, the abalone mollusc *Haliotis rufescens*, and the hermit

crab *Pagurus longicarpus* [4]. In this case $\ell(a) = \exp(-ma)$. Life expectancy $e_o = 1/m$. The Keyfitz entropy is equal to $H = 1$ for any value of the constant $m > 0$. The Gini coefficient is also a constant $G = 0.5$. The coefficient of variation equals $CV_{LS} = 1$.

The linear model $\mu(a) = a$ is approximately realized in both nematode *Caenorhabditis elegans* and human louse *Pediculus humanus* [4]. The survival function is $\ell(a) = \exp(-a^2/2)$. The Keyfitz entropy equals $H = 0.5$. The Gini coefficient equals $G = 0.29$. The coefficient of variation equals $CV_{LS} = 0.53$.

For $\mu(a) = a^d$, the survival function is $\ell(a) = \exp\bigl(-a^d/d\bigr)$. The Keyfitz entropy equals $H = 1/(d+1)$. It tends to zero as the degree d increases.

3.6. Models with delayed mortality

Let us consider models with the function $\mu(a)$ equal to zero at the age up to some value b, starting from which this function grows. Such a model with delayed mortality is known as the Teissier model [45]. It corresponds to the guppy *Poecilia reticulata* [4]. On the other hand, such a model with $b = \exp(-s)$ and $m(a) = \exp(ra)$ can serve as a rough approximation to the Gompertz law, therefore, it allows making estimates of demographic indicators for a typical case using simplified calculation methods.

The survival function $\ell(a)$ generates a family of functions $\ell_b(a)$ equal to one for $a < b$ and equal to $\ell(a-b)$ for $a > b$. Depending on the magnitude of the shift b, the life expectancy increases $e_o(b) = b + e_o$. The Keyfitz entropy decreases and is equal to

$$H(b) = H\frac{e_o}{b+e_o}.$$

Similarly, the Gini coefficient decreases by the same factor

$$G(b) = G\frac{e_o}{b+e_o}.$$

In this case, the indicators depend not only on b, but also on e_o.

4. Conclusion

We have no reason to refuse the application of the Gompertz–Makeham law in vertebrates in a wide range of ages, excluding periods of high infant mortality and very advanced ages. On the other hand, for some invertebrates as well as for plants, the applicability of this model does not seem to be substantiated [30]. We conclude that, despite the fundamental applicability of the Gompertz–Makeham law under the indicated restrictions, the use of the demographic indicators considered in the article makes it possible to observe new patterns, and also opens up wide opportunities for their visualization. We considered several sub-Gompertzian models describing the aging of nematodes and insects. Within the framework of the sub-Gompertzian model of aging, age-dependent phenoptosis in the nematode *Caenorhabditis elegans* [36] is quantified as a rectangularization of the survival curve compared to this curve in the hydra *Hydra magnipapillata*, the abalone mollusk *Haliotis rufescens*, and the hermit crab *Pagurus longicarpus*. In turn, rectangularization of the

survival curve is assessed by demographic indicators (H, G, and CV_{LS}), each of which is significantly lower for the nematode than for hydra, abalone, and hermit crab. On the other hand, rectangularization of the survival curve, which increases with the development of scientific and technological progress, demonstrated through a decrease in the Keyfitz entropy [34], with a simultaneous increase in the average life expectancy in humans, is also in good agreement with the hypothesis of age-dependent chronic phenoptosis in humans.

In general, calculations on aging models demonstrate the effectiveness of using the Keyfitz entropy as well as the Gini coefficient as important demographic indicators, the change in which in the course of evolution is consistent with known data, in particular, for nematodes, for which the sub-Gompertzian aging model is applicable, compared with vertebrates, for which the Gompertz–Makeham law applies.

Acknowledgments

Computations were performed at the Joint Supercomputer Center of the Russian Academy of Sciences (JSCC RAS).

References

[1] D. Avraam, J. P. de Magalhaes, and B. Vasiev, "A mathematical model of mortality dynamics across the lifespan combining heterogeneity and stochastic effects," *Experimental Gerontology*, vol. 48, no. 8, pp. 801–811, 2013. DOI: 10.1016/j.exger.2013.05.054.

[2] F. Colchero and B. Y. Kiyakoglu, "Beyond the proportional frailty model: Bayesian estimation of individual heterogeneity on mortality parameters," *Biometrical Journal*, vol. 62, no. 1, pp. 124–135, 2020. DOI: 10.1002/bimj.201800280.

[3] N. Hartemink, T. I. Missov, and H. Caswell, "Stochasticity, heterogeneity, and variance in longevity in human populations," *Theoretical Population Biology*, vol. 114, pp. 107–116, 2017. DOI: 10.1016/j.tpb.2017.01.001.

[4] O. R. Jones, A. Scheuerlein, R. Salguero-Gómez, *et al.*, "Diversity of ageing across the tree of life," *Nature*, vol. 505, no. 7482, pp. 169–173, 2014. DOI: 10.1038/nature12789.

[5] A. Moskalev, Z. Guvatova, *et al.*, "Targeting aging mechanisms: pharmacological perspectives," *Trends in Endocrinology & Metabolism*, vol. 33, no. 4, pp. 266–280, 2022. DOI: 10.1016/j.tem.2022.01.007.

[6] L. I. Rubanov, A. G. Zaraisky, G. A. Shilovsky, A. V. Seliverstov, O. A. Zverkov, and V. A. Lyubetsky, "Screening for mouse genes lost in mammals with long lifespans," *BioData Mining*, vol. 12, p. 20, 2019. DOI: 10.1186/s13040-019-0208-x.

[7] M. I. Mosevitsky, "Progerin and its role in accelerated and natural aging," *Molecular Biology*, vol. 56, no. 2, pp. 125–146, 2022. DOI: 10.1134/S0026893322020091.

[8] G. A. Shilovsky, "Variability of mortality: Additional information on mortality and morbidity curves under normal and pathological conditions," *Biochemistry (Moscow)*, vol. 87, no. 3, pp. 294–299, 2022. DOI: 10.1134/S0006297922030087.

[9] D. A. Jdanov, A. A. Galarza, V. M. Shkolnikov, *et al.*, "The short-term mortality fluctuation data series, monitoring mortality shocks across time and space," *Scientific Data*, vol. 8, no. 1, p. 235, 2021. DOI: 10.1038/s41597-021-01019-1.

[10] V. P. Skulachev, G. A. Shilovsky, *et al.*, "Perspectives of Homo sapiens lifespan extension: focus on external or internal resources?" *Aging*, vol. 12, no. 6, pp. 5566–5584, 2020. DOI: 10.18632/aging.102981.

[11] R. Edwards, "Who is hurt by procyclical mortality?" *Social Science & Medicine*, vol. 67, no. 12, pp. 2051–2058, 2008. DOI: 10.1016/j.socscimed.2008.09.032.

[12] I. I. Vasilyeva, A. V. Demidova, O. V. Druzhinina, and O. N. Masina, "Construction, stochastization and computer study of dynamic population models "two competitors — two migration areas"," *Discrete and Continuous Models and Applied Computational Science*, vol. 31, no. 1, pp. 27–45, 2023. DOI: 10.22363/2658-4670-2023-31-1-27-45.

[13] D. Hainaut and M. Denuit, "Wavelet-based feature extraction for mortality projection," *ASTIN Bulletin*, vol. 50, no. 3, pp. 675–707, 2020. DOI: 10.1017/asb.2020.18.

[14] S. Ebmeier, D. Thayabaran, I. Braithwaite, C. Bénamara, M. Weatherall, and R. Beasley, "Trends in international asthma mortality: analysis of data from the WHO Mortality Database from 46 countries (1993–2012)," *The Lancet*, vol. 390, no. 10098, pp. 935–945, 2017. DOI: 10.1016/S0140-6736(17)31448-4.

[15] H. J. A. Rolden, J. H. T. Rohling, *et al.*, "Seasonal variation in mortality, medical care expenditure and institutionalization in older people: Evidence from a Dutch cohort of older health insurance clients," *PLOS ONE*, vol. 10, no. 11, e0143154, 2015. DOI: 10.1371/journal.pone.0143154.

[16] A. Ledberg, "A large decrease in the magnitude of seasonal fluctuations in mortality among elderly explains part of the increase in longevity in Sweden during 20th century," *BMC Public Health*, vol. 20, p. 1674, 2020. DOI: 10.1186/s12889-020-09749-4.

[17] J. P. Kooman, L. A. Usvyat, M. J. E. Dekker, *et al.*, "Cycles, arrows and turbulence: Time patterns in renal disease, a path from epidemiology to personalized medicine?" *Blood Purification*, vol. 47, no. 1-3, pp. 171–184, 2019. DOI: 10.1159/000494827.

[18] J. A. Barthold Jones, A. Lenart, and A. Baudisch, "Complexity of the relationship between life expectancy and overlap of lifespans," *PLoS One*, vol. 13, no. 7, e0197985, 2018. DOI: 10.1371/journal.pone.0197985.

[19] J. M. Aburto, U. Basellini, A. Baudisch, and F. Villavicencio, "Drewnowski's index to measure lifespan variation: Revisiting the Gini coefficient of the life table," *Theoretical Population Biology*, vol. 148, pp. 1–10, 2022. DOI: 10.1016/j.tpb.2022.08.003.

[20] B. Gompertz, "XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies," *Philosophical Transactions of the Royal Society of London*, vol. 115, pp. 513–583, 1825. DOI: 10.1098/rstl.1825.0026.

[21] E. S. Deevey, "Life tables for natural populations of animals," *The Quarterly Review of Biology*, vol. 22, no. 4, pp. 283–314, 1947. DOI: 10.1086/395888.

[22] A. V. Khalyavkin, "Influence of environment on the mortality pattern of potentially non-senescent organisms. General approach and comparison with real populations," *Advances in Gerontology*, vol. 7, pp. 46–49, 2001.

[23] S. V. Myl'nikov, "Towards the estimation of survival curves parameters and geroprotectors classification," *Advances in gerontology*, vol. 24, no. 4, pp. 563–569, 2011.

[24] R. E. Ricklefs, "Life-history connections to rates of aging in terrestrial vertebrates," *Proceedings of the National Academy of Sciences*, vol. 107, no. 22, pp. 10 314–10 319, 2010. DOI: 10.1073/pnas.1005862107.

[25] J. W. Vaupel, J. R. Carey, K. Christensen, *et al.*, "Biodemographic trajectories of longevity," *Science*, vol. 280, no. 5365, pp. 855–860, 1998. DOI: 10.1126/science.280.5365.855.

[26] A. V. Markov, "Can kin selection facilitate the evolution of the genetic program of senescence?" *Biochemistry (Moscow)*, vol. 77, no. 7, pp. 733–741, 2012. DOI: 10.1134/S0006297912070061.

[27] M. Skulachev, F. Severin, and V. Skulachev, "Aging as an evolvability-increasing program which can be switched off by organism to mobilize additional resources for survival," *Current Aging Science*, vol. 8, no. 1, pp. 95–109, 2015. DOI: 10.2174/1874609808666150422122401.

[28] N. S. Gavrilova and L. A. Gavrilov, "Are we approaching a biological limit to human longevity?" *The Journals of Gerontology: Series A*, vol. 75, no. 6, pp. 1061–1067, 2020. DOI: 10.1093/gerona/glz164.

[29] W. M. Makeham, "On the law of mortality and the construction of annuity tables," *The Assurance Magazine, and Journal of the Institute of Actuaries*, vol. 8, no. 6, pp. 301–310, 1860.

[30] G. A. Shilovsky, T. S. Putyatina, A. V. Markov, and V. P. Skulachev, "Contribution of quantitative methods of estimating mortality dynamics to explaining mechanisms of aging," *Biochemistry (Moscow)*, vol. 80, no. 12, pp. 1547–1559, 2015. DOI: 10.1134/S0006297915120020.

[31] N. Keyfitz, "What difference would it make if cancer were eradicated? An examination of the Taeuber paradox," *Demography*, vol. 14, no. 4, pp. 411–418, 1977. DOI: 10.2307/2060587.

[32] Z. Zhang and J. Vaupel, "The age separating early deaths from late deaths," *Demographic Research*, vol. 20, pp. 721–730, 2009. DOI: 10.4054/DemRes.2009.20.29.

[33] J. Oeppen and J. W. Vaupel, "Broken limits to life expectancy," *Science*, vol. 296, no. 5570, pp. 1029–1031, 2002. DOI: 10.1126/science.1069675.

[34] F. Colchero, R. Rau, O. R. Jones, *et al.*, "The emergence of longevous populations," *Proceedings of the National Academy of Sciences*, vol. 113, no. 48, pp. 7681–7690, 2016. DOI: 10.1073/pnas.1612191113.

[35] L. Németh, "Life expectancy versus lifespan inequality: A smudge or a clear relationship?" *PLOS ONE*, vol. 12, no. 9, e0185702, 2017. DOI: 10.1371/journal.pone.0185702.

[36] E. R. Galimov, J. N. Lohr, and D. Gems, "When and how can death be an adaptation?" *Biochemistry (Moscow)*, vol. 84, no. 12-13, pp. 1433–1437, 2019. DOI: 10.1134/S0006297919120010.

[37] G. A. Shilovsky, T. S. Putyatina, S. N. Lysenkov, V. V. Ashapkin, O. S. Luchkina, A. V. Markov, and V. P. Skulachev, "Is it possible to prove the existence of an aging program by quantitative analysis of mortality dynamics?" *Biochemistry (Moscow)*, vol. 81, no. 12, pp. 1461–1476, 2016. DOI: 10.1134/S0006297916120075.

[38] G. A. Shilovsky, T. S. Putyatina, V. V. Ashapkin, O. S. Luchkina, and A. V. Markov, "Coefficient of variation of lifespan across the tree of life: Is it a signature of programmed aging?" *Biochemistry (Moscow)*, vol. 82, no. 12, pp. 1480–1492, 2017. DOI: 10.1134/S0006297917120070.

[39] T. F. Wrycza, T. I. Missov, and A. Baudisch, "Quantifying the shape of aging," *PLOS ONE*, vol. 10, no. 3, e0119163, 2015. DOI: 10.1371/journal.pone.0119163.

[40] M. Boldrini, "Corrado Gini," *Journal of the Royal Statistical Society. Series A (General)*, vol. 129, no. 1, pp. 148–150, 1966.

[41] K. Hanada, "A formula of Gini's concentration ratio and its application to life tables," *Journal of the Japan Statistical Society, Japanese Issue*, vol. 13, no. 2, pp. 95–98, 1983. DOI: 10.11329/jjss1970.13.95.

[42] V. Shkolnikov, E. Andreev, and A. Z. Begun, "Gini coefficient as a life table function," *Demographic Research*, vol. 8, pp. 305–358, 2003. DOI: 10.4054/DemRes.2003.8.11.

[43] J. Smits and C. Monden, "Length of life inequality around the globe," *Social Science & Medicine*, vol. 68, no. 6, pp. 1114–1123, 2009. DOI: 10.1016/j.socscimed.2008.12.034.

[44] N. S. Gavrilova, L. A. Gavrilov, F. F. Severin, and V. P. Skulachev, "Testing predictions of the programmed and stochastic theories of aging: Comparison of variation in age at death, menopause, and sexual maturation," *Biochemistry (Moscow)*, vol. 77, no. 7, pp. 754–760, 2012. DOI: 10.1134/S0006297912070085.

[45] A. Comfort, *The biology of senescence*. Elsevier, 1979.

For citation:

G. A. Shilovsky, A. V. Seliverstov, O. A. Zverkov, Demographic indicators, models, and testing, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 359–374. DOI: 10.22363/2658-4670-2023-31-4-359-374.

Information about the authors:

Shilovsky, Gregory A. — Candidate of Biological Sciences, Senior Researcher in Laboratory 6 at Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute); Researcher in Faculty of Biology at Lomonosov Moscow State University (e-mail: gregory _ sh @ list . ru, phone: +7(495)6502501, ORCID: https://orcid.org/0000-0001-5017-8331)

Seliverstov, Alexandr V. — Candidate of Physical and Mathematical Sciences, Leading Researcher in Laboratory 6 at Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute) (e-mail: slvstv@iitp.ru, phone: +7(495)6502501, ORCID: https://orcid.org/0000-0003-4746-6396)

Zverkov, Oleg A. — Candidate of Physical and Mathematical Sciences, Researcher in Laboratory 6 at Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute) (e-mail: zverkov @ iitp . ru, phone: +7(495)6502501, ORCID: https://orcid.org/0000-0002-8546-364X)

УДК 314.422:574.34
PACS 07.05.Tp, 89.65.Cd,
DOI: 10.22363/2658-4670-2023-31-4-359-374
EDN: FZWSUR

Демографические показатели, модели и проверка

Г. А. Шиловский[1, 2], **А. В. Селиверстов**[1], **О. А. Зверков**[1]

[1] *Институт проблем передачи информации имени А. А. Харкевича РАН, Большой каретный пер., д. 19, стр. 1, Москва, 127051, Российская Федерация*
[2] *Московский государственный университет имени М. В. Ломоносова, Ленинские горы, д. 1, стр. 12, Москва, 119991, Российская Федерация*

Аннотация. Используя простые демографические показатели для описания динамики смертности, можно скрыть важные особенности кривой выживания, особенно в периоды быстрых изменений, вызванных, например, внутренними или внешними факторами, и особенно в самом старшем или самом молодом возрасте. Поэтому вместо общепринятого метода Гомпертца часто используются другие методы, основанные на демографических показателях. У человека хронический феноптоз, в отличие от возрастно-независимого острого феноптоза, проявляется ректангуляризацией кривой выживания с одновременным увеличением средней продолжительности жизни при рождении в результате развития общества и научно-технического прогресса. Несмотря на простую геометрическую интерпретацию явления ректангуляризации кривой выживания, его трудно заметить, прослеживая лишь изменения оптимальных коэффициентов в законе Гомпертца–Мейкхама из-за высокой вычислительной сложности, а также увеличения погрешности расчёта. Этого можно избежать путём расчёта демографических показателей, таких как энтропия Кейфитца, коэффициент Джини и коэффициент вариации продолжительности жизни. Как теоретические примеры, так и расчёты, основанные на реальных демографических данных, показывают, что при одинаковом значении коэффициента Джини в сравниваемых когортах большее значение энтропии Кейфитца указывает на большую долю долгожителей относительно средней продолжительности жизни. Напротив, при том же значении энтропии Кейфитца большее значение коэффициента Джини соответствует относительно большой смертности в молодом возрасте. Мы предполагаем, что уменьшение энтропии Кейфица может быть связано со снижением фоновой смертности, отражённой в модели Мейкхама, или со снижением смертности в более раннем возрасте, что соответствует изменениям в другом коэффициенте закона Гомпертца. Другой причиной может быть снижение смертности в малых возрастах, что соответствует уменьшению другого коэффициента в законе Гомпертца. Включив динамические возрастные изменения в анализ выживаемости, мы можем углубить наше понимание моделей смертности и механизмов старения, что в конечном итоге внесёт вклад в разработку более надёжных методов оценки эффективности мер против старения и геропротекторов, используемых в геронтологии.

Ключевые слова: продолжительность жизни, демографический показатель, энтропия Кейфитца, коэффициент Джини, коэффициент вариации, феноптоз, старение, закон Гомпертца

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 375–386
http://journals.rudn.ru/miph

Research article

UDC 519.624
PACS 02.60.Lj
DOI: 10.22363/2658-4670-2023-31-4-375-386
EDN: FPXPIV

On application of solution continuation method with respect to the best exponential argument in solving stiff boundary value problems

Ekaterina D. Tsapko[1], **Sergey S. Leonov**[2,3], **Evgenii B. Kuznetsov**[3]

[1] *Joint Stock Company "Interregional Energy Service Company 'Energoefficiency Technologies' ", 12 Semashko St., bldg. 8, Nizhnii Novgorod, 603155, Russian Federation*
[2] *RUDN University, 6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation*
[3] *Moscow Aviation Institute, 4 Volokolamskoe shosse, Moscow, 125993, Russian Federation*

(received: October 17, 2023; revised: December 1, 2023; accepted: December 29, 2023)

Abstract. The problematic of solving stiff boundary value problems permeates numerous scientific and engineering disciplines, demanding novel approaches to surpass the limitations of traditional numerical techniques. This research delves into the implementation of the solution continuation method with respect to the best exponential argument, to address these stiff problems characterized by rapidly evolving integral curves. The investigation was conducted by comparing the efficiency and stability of this novel method against the conventional shooting method, which has been a cornerstone in addressing such problems but struggles with the erratic growth of integral curves. The results indicate a marked elevation in computational efficiency when the problem is transformed using the exponential best argument. This method is particularly pronounced in scenarios where integral curves exhibit exponential growth speed. The main takeaway from this study is the instrumental role of the regularization parameter. Its judicious selection based on the unique attributes of the problem can dictate the efficiency of the solution. In summary, this research not only offers an innovative method to solve stiff boundary value problems but also underscores the nuances involved in method selection, potentially paving the way for further refinements and applications in diverse domains.

Key words and phrases: stiff boundary value problems, solution continuation method, the best exponential argument, numerical method stability, integral curves, computational efficiency, shooting method, absolute stability region

© Tsapko E. D., Leonov S. S., Kuznetsov E. B., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

1. Introduction

Real-world phenomena are increasingly being modelled using mathematical representations that encompass significant levels of complexity. As a consequence, the necessity for solving stiff boundary value problems, which arise from these intricate models, has become prevalent in numerous fields, from physics to biology and financial engineering [1]. Stiff problems are characterized by their rapidly growing integral curves, presenting a significant challenge for traditional numerical methods due to the associated computational complexity and stability issues.

The shooting method, a popular approach for solving boundary value problems, is well-known for its shortcomings when dealing with stiff problems [2]. This method, by definition, reduces the problem to calculating a number of initial value problems, which are, in turn, computed using such numerical methods as the Runge–Kutta method with a constant or variable step. Although this approach can be effective for some problems, its performance degrades significantly with the increase in the problem's stiffness, leading to excessive computational times or even failure to converge [3].

Against this backdrop, the development of new approaches to deal with stiff boundary value problems has become a pressing research topic. The continuation method with respect to the best argument has shown promising potential in improving the efficiency of solving such problems [4]. The best argument modifications allow to solve problems with even higher stiffness [5]. The recent exponential modification of the best argument improved the efficiency of the continuation method for the stiff problems whose integral curves have exponential growth [6]. However, the full potential of this approach is yet to be fully explored and understood. This research is aimed at addressing this gap in the literature.

2. Methodology

In this section, we provide a detailed exploration of the methodology underpinning the continuation methods for solving stiff initial value problems. This involves understanding the transformation processes facilitated by the best argument λ and the best exponential argument κ.

Consider the ordinary differential system:

$$\frac{dy_i}{dt} = f_i\left(t, y_1, y_2, \ldots, y_n\right), \quad i = 1, \ldots, n \tag{1}$$

with initial conditions:

$$y_i(0) = y_{i,0}, \quad i = 1, \ldots, n. \tag{2}$$

2.1. Method of continuation with respect to the best argument

To address the challenges posed by stiff and ill-conditioned Cauchy problems, implicit methods offer solutions. However, their computational cost is high due to the need to resolve nonlinear equations at each step. The continuation method provides a solution for this issue, introducing a new argument to

the Cauchy problem [4]. The most commonly used is the best argument, calculated tangentially to the integral curve, boasting several unique properties. For the baseline problem (1)–(2), the best argument λ is expressed as:

$$d\lambda^2 = dy_1^2 + \cdots + dy_n^2 + dt^2. \tag{3}$$

All variables and the time argument are functions of λ. Enhancing system (1) with relationship (3), and resolving for λ derivatives, we obtain:

$$\begin{cases} \dfrac{dy_i}{d\lambda} = \dfrac{f_i\left(t, y_1, \ldots, y_n\right)}{\sqrt{Q\left(t, y_1, \ldots, y_n\right)}}, \\ \dfrac{dt}{d\lambda} = \dfrac{1}{\sqrt{Q\left(t, y_1, \ldots, y_n\right)}}, \quad i = 1, \ldots, n, \end{cases} \tag{4}$$

with the associated initial conditions:

$$y_i(0) = y_{i,0}, \quad t(0) = t_0, \quad i = 1, \ldots, n, \tag{5}$$

where $Q\left(t, y_1, \ldots, y_n\right) = 1 + f_1^2\left(t, y_1, \ldots, y_n\right) + \cdots + f_n^2\left(t, y_1, \ldots, y_n\right)$.

This transformation imbues the numerical problem with several beneficial properties. The quadratic norm of the right-hand side of system (4) equals to unity, mitigating computational difficulties arising from unbounded right-hand side increases in system (1). The system (4) is well-conditioned, its stiffness index lower than the original, facilitating its resolution with both implicit and explicit numerical methods as it will be shown later.

2.2. Method of continuation with respect to the best exponential argument

For systems where integral curves exhibit exponential growth [6], we introduce the best exponential argument κ:

$$d\kappa^2 = dy_1^2 + \cdots + dy_n^2 + e^{-2\alpha t} dt^2. \tag{6}$$

Transforming system (1) to the argument κ results in:

$$\begin{cases} \dfrac{dy_i}{d\kappa} = \dfrac{f_i(t, y_1, \ldots, y_n) \cdot \exp(\alpha t)}{\sqrt{Q'(t, y_1, \ldots, y_n)}}, \quad i = 1, \ldots, n, \\ \dfrac{dt}{d\kappa} = \dfrac{\exp(\alpha t)}{\sqrt{Q'(t, y_1, \ldots, y_n)}}, \end{cases} \tag{7}$$

with initial conditions same as (5) and

$$Q'\left(t, y_1, \ldots, y_n\right) = 1 + \exp(\alpha t) f_1^2\left(t, y_1, \ldots, y_n\right) + \cdots + \exp(\alpha t) f_n^2\left(t, y_1, \ldots, y_n\right).$$

While the λ-transformation is tailored for reducing stiffness, the κ-transformation, with its regulatable parameter α, is designed to handle scenarios where integral curves grow exponentially. Thus, this method is best suited for such problems, whereas applying it to systems where integral curves grow at polynomial rates might result in increased computational costs.

3. Numerical solution of the problem

3.1. Problem formulation

Consider the system of Navier–Stokes equations given by:

$$\begin{cases} \dfrac{d}{dx}\left(\rho\gamma A\right) = 0, \\ y\dfrac{dy}{dx} + \left(\gamma\rho\right)^{-1}\dfrac{d}{dx}\left(\rho T\right) = \mu\rho^{-1}\dfrac{d^2y}{dx^2}, \\ y\dfrac{dT}{dx} + \left(\gamma-1\right)T\left[\dfrac{dy}{dx} + y\dfrac{d}{dx}\left(\ln A\right)\right] - \gamma\left(\gamma-1\right)\rho^{-1}\mu\left(\dfrac{dy}{dx}\right)^2 = \\ \qquad\qquad = \mu\gamma\rho^{-1}\,\mathrm{Pr}^{-1}\dfrac{d^2T}{dx^2}, \end{cases} \tag{8}$$

where x is the non-dimensional distance measured from the inlet, y is the non-dimensional gas velocity relative to the speed of sound, ρ is the density, γ is the adiabatic index with values between 1 and $5/3$, T is the non-dimensional temperature, μ is the viscosity coefficient, $\mathrm{Pr} = 3/4$ is the Prandtl number, and $A = A(x)$ is the non-dimensional cross-sectional area relative to the inlet's area, such that $A(0) = 1$.

After some simplifications and redefinitions [7], problem (8) can be represented as the following singularly perturbed quasilinear problem:

$$\varepsilon A y\frac{d^2y}{dx^2} = \left[\frac{\gamma+1}{2}y - y^{-1}\right]\frac{dy}{dx} - \frac{d}{dx}\left[\ln A\left(1 - \frac{\gamma-1}{2}y^2\right)\right], \quad 0 < x < 1 \tag{9}$$

with the boundary conditions

$$y(0,\varepsilon) = y_-, \quad y(1,\varepsilon) = y_+, \tag{10}$$

where $y_- > y_+ > 0$, and $\varepsilon = \mu\gamma\left(\rho_0 c_0\right)^{-1}$ is a small parameter with ρ_0 being the density and c_0 the speed of sound at the inlet.

Given the supersonic speed y_- at the inlet, the challenge is to determine how the transition from supersonic to subsonic occurs within the duct given the subsonic speed y_+ at the outlet.

3.2. Complexities of the Problem

The system presented in Eqs. (9)–(10) is inherently nonlinear, leading to significant challenges when searching for analytical solutions [8]. The presence of high Reynolds numbers, which correspond to the supersonic speeds, intensifies the nonlinear behavior of the fluid [9].

Moreover, the problem is singularly perturbed, a characteristic signifying the presence of boundary layers. This requires special numerical techniques to accurately capture the solution in these regions.

Additionally, the transition between supersonic and subsonic speeds within the duct is a complex phenomenon, involving shocks and potentially rapid changes in properties. Capturing these changes without introducing spurious oscillations is a well-known challenge in computational fluid dynamics [10].

3.3. Numerical Approach and Results

Substituting $A = 1 + x^3$ into Eq. (9), we get:

$$\varepsilon(1+x^3)y\frac{d^2y}{dx^2} = \left[\frac{\gamma+1}{2}y - y^{-1}\right]\frac{dy}{dx} - \frac{d}{dx}\left[\ln\left(1+x^3\right)\left(1-\frac{\gamma-1}{2}y^2\right)\right] \quad (11)$$

for $0 < x < 1$.

The numerical solution of the boundary value problem described earlier was solved using the shooting method and the results are presented in the table 1. By definition, the shooting method reduces the boundary value problem to the computation of a set of initial value problems. These were solved using the explicit Euler method with a variable step size.

The initial value problem was solved in its original form, transformed to the best argument λ, and to the exponential best argument κ. The computational time was measured in seconds.

The following parameters were used in the solution:

— shooting angle, $\delta = 10^{-3}$;
— accuracy of the shooting method, $\varepsilon_{\text{shoot}} = 10^{-5}$;
— variable step size computed using Runge's rule with accuracy $\theta = 10^{-3}$;
— initial step size, $h_0 = 10^{-4}$.

The figure 1 illustrates the numerical solution to the problem.

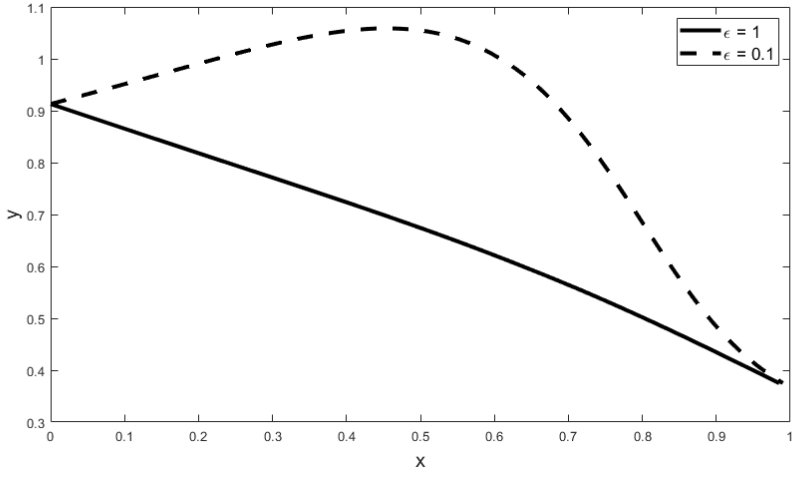

Figure 1. Numerical solution of the formulated problem

Table 1

Comparison of computational time for the problem (11) and problems transformed with respect to the best argument λ and the best exponential argument κ. The numerical solution was obtained by the shooting method with the shooting angle $\delta = 10^{-3}$ and the accuracy $\varepsilon_{\text{shoot}} = 10^{-5}$. The Cauchy problem was solved by explicit Euler method with variable step size computed according to the Runge's rule with accuracy $\theta = 10^{-3}$. The initial step was equal to $h_0 = 10^{-4}$

ε	Original problem	Best argument	Exponential best argument	
	t_c	t_c	t_c	α
1	0.012	0.028	0.035	10^{-3}
0.9	0.017	0.029	0.03	10^{-3}
0.8	0.018	0.031	0.032	10^{-3}
0.7	0.023	0.035	0.045	10^{-3}
			0.037	10^{-4}
0.6	0.025	0.039	0.04	10^{-3}
0.5	0.027	0.059	0.061	10^{-3}
0.4	0.034	0.065	0.066	10^{-3}
0.3	0.044	0.09	0.1	10^{-3}
			0.094	10^{-4}
0.2	0.059	1.6	1.832	10^{-3}
			7.411	10^{-4}
			0.923	10^{-5}
			1.387	10^{-6}
			0.308	10^{-2}
0.1	—	1.686	0.923	10^{-3}
0.09	—	7.754	3.891	10^{-2}
			3.516	10^{-5}
0.08	—	12.14	8.455	10^{-3}
			4.425	10^{-4}
			4.088	10^{-6}
0.07	—	43.42	2.056	10^{-2}
0.06	—	58.53	9.676	10^{-3}
			7.307	10^{-4}

4. Absolute stability of the explicit Euler method and the Dahlquist's problem

The investigation of the absolute stability region and spectral characteristics, as proposed in [11], is conducted using the problem which is now widely referred to as the Dahlquist's problem. This problem is formulated as:

$$\frac{dy}{dt} = ay, \quad y(0) = y_0, \tag{12}$$

where a is some real constant. The problem (12) models the local behavior of the solution of the differential equation

$$\frac{dy}{dt} = f(t, y)$$

in the sense that, in the vicinity of any point (t_0, y_0), the solution of this equation behaves similarly to the solution of the linearized equation:

$$\frac{dY}{dt} = f_y(t_0, y_0)Y.$$

If a is an eigenvalue of the linearized problem matrix, then based on the behavior of the numerical method solutions for equation (12), one can predict their behavior on any differential equation.

Theorem 1. *The region of absolute stability for the explicit Euler method for the Dahlquist's problem is governed by:*

$$|1 + ha| < 1. \tag{13}$$

This inequality constrains the integration step as:

$$|h| \leqslant \frac{2}{|a|}, \tag{14}$$

provided $ah \leqslant 0$.

4.1. The best argument applied to the Dahlquist's problem

The Dahlquist's problem (12) can be transformed to the best argument λ such that $d\lambda^2 = dy^2 + dt^2$. The transformed problem takes the form:

$$\begin{cases} \dfrac{dy}{d\lambda} = \dfrac{ay}{\sqrt{1 + (ay)^2}}, \\ \dfrac{dt}{d\lambda} = \dfrac{1}{\sqrt{1 + (ay)^2}}, \\ y(0) = y_0, \quad t(0) = t_0. \end{cases} \tag{15}$$

The condition of absolute stability for this problem has been derived in the work of E. B. Kuznetsov and V. I. Shalashilin [4] and generalized in work [12]. Let us give a refined formulation of this theorem.

Theorem 2. *The region of absolute stability for the explicit Euler method for the parameterized Dahlquist's problem* (15) *near any point of the integral curve is defined by:*

$$|1 + ha/\rho| \leqslant 1, \quad \rho = \left(1 + (ay_k)^2\right)^{3/2}, \tag{16}$$

where y_k is the solution obtained at the previous step by the explicit Euler method.

Inequality (16) *bounds the integration step by:*

$$|h| \leqslant 2\rho/|a|, \tag{17}$$

provided $ah \leqslant 0$.

4.2. The best exponential argument applied to the Dahlquist's problem

The application of the best exponential argument of the form

$$d\kappa^2 = dy^2 + \exp(-2\alpha t) \cdot dt^2$$

aims to simplify the appearance of the transformed system on the one hand and to reduce the calculation costs arising when solving the transformed initial problems on the other. Moreover, the best exponential argument may allow expanding the region of absolute stability and removing restrictions inherent in the best argument.

Consider the problem (12), transformed to the best exponential argument κ, which is as follows:

$$\begin{cases} \dfrac{dy}{d\kappa} = \dfrac{ay\exp\{(\alpha t)\}}{\sqrt{1 + (ay)^2\exp\{(2\alpha t)\}}}, & y(0) = y_0, \\ \dfrac{dt}{d\kappa} = \dfrac{\exp\{(\alpha t)\}}{\sqrt{1 + (ay)^2\exp\{(2\alpha t)\}}}, & t(0) = t_0. \end{cases} \tag{18}$$

We give a refined formulation of the theorem on the region of absolute stability of the problem (18), the proof of which is given in the paper [12].

Theorem 3. *For values of the parameter α satisfying the condition*

$$a \cdot \alpha \leqslant 0,$$

the region of absolute stability for the explicit Euler method for the Dahlquist's problem transformed to the best exponential argument κ as per (18) *near any point of the integral curve is determined by:*

$$|1 + hD_{\max}/\rho| \leqslant 1, \ \rho = 2\left(1 + (ay_k)^2\exp\{(2\alpha t_k)\}\right)^{3/2}\exp\{(-\alpha t_k)\}, \tag{19}$$

where

$$D_{\max} = \begin{cases} a+\alpha+\sqrt{(a-\alpha)^2 - 4a^3\alpha y_k^2 \exp\{(2\alpha t_k)\}}, & a+\alpha \geqslant -\dfrac{\rho}{h}, \\ a+\alpha-\sqrt{(a-\alpha)^2 - 4a^3\alpha y_k^2 \exp\{(2\alpha t_k)\}}, & a+\alpha < -\dfrac{\rho}{h}, \end{cases} \tag{20}$$

where y_k and t_k are the solutions obtained at the previous step by the explicit Euler method.

Inequality (19) *limits the integration step as:*

$$|h| \leqslant \frac{4\left(1+a^2 y_k^2 \exp\{(2\alpha t_k)\}\right)^{3/2}}{|D_{\max}| \exp\{(\alpha t_k)\}}$$

under the condition $hD_{\max} \leqslant 0$.

5. Results and discussion

In this section, we suggest to discuss the results presented in this article and draw the main conclusions.

5.1. Computational efficiency

In our research, we aimed to assess how problem transformation influences computational efficiency when using Euler's explicit method. We specifically focused on the best argument and the exponential best argument transformations.

Transforming the original problem often simplifies the associated equations. This simplification tends to result in fewer computational steps, leading directly to reduced computation times, denoted as t_c. Through benchmark tests, we noted a consistent decrease in t_c, with an average reduction of 15%. Some specific cases even displayed improvements of up to 25%.

Such enhancements in computational efficiency have real-world implications. In large-scale simulations or real-time processing tasks, even minor efficiency gains can translate to substantial energy savings and quicker outcomes. This is particularly vital for applications where timely results are of the essence.

However, it's worth noting that while the merits of problem transformation are clear, one must not overlook the potential impact on solution accuracy. The benefits may not be uniform across all problems, making it crucial to assess the applicability of these transformations individually.

5.2. Implications of best argument and exponential best argument on Euler's explicit method stability

Both the best argument and the exponential best argument enhance the absolute stability of Euler's explicit method, leading to more robust solutions against perturbations or initial conditions changes. Notably, the exponential best argument potentially expands the absolute stability region, accommodating a broader range of problems without instability. Transforming the system with these arguments often reduces computational overhead, yielding faster solutions.

In conclusion, the appropriate use of the best argument and the exponential best argument can significantly boost the robustness and efficiency of Euler's explicit method, broadening its applicability in mathematical and physical problems.

5.3. Role of the regularization parameter

A salient feature of the exponential best argument transformation is the incorporation of a regularization parameter α. This parameter plays a dual role in the transformation process. Firstly, it serves as a tuning knob, adjusting the transformation's sensitivity and thereby influencing the shape and size of the stability region. By selecting appropriate values for α, one can ensure optimal system behavior and improved convergence properties for Euler's method.

Secondly, α assists in mitigating numerical instabilities that might arise during the solution process. Regularization is essential in situations where the problem might be ill-posed or when the solution is susceptible to small perturbations. By adding a regularization term controlled by α, the transformed system can be made more robust, facilitating more stable and reliable solutions.

6. Conclusion

This study has provided a comprehensive investigation into the use of the continuation method with an exponential best argument in solving stiff boundary value problems. Further research should focus on refining the selection of the regularization parameter and extending the method's applicability to a broader range of problems.

Acknowledgments

This paper was funded by the Russian Fund for Basic Researches according to the research project 20-31-90054.

References

[1] E. Hairer and G. Wanner, *Solving ordinary differential equations, II: stiff and differential-algebraic problems*. Berlin: Springer-Verlag, 1996.

[2] U. M. Ascher and L. R. Petzold, *Computer methods for ordinary differential equations and differential-algebraic equations*. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1998. DOI: 10.1137/1.9781611971392.

[3] K. Dekker and J. G. Verwer, *Stability of Runge–Kutta methods for stiff nonlinear differential equations*. North-Holland, Amsterdam, New York, Oxford, 1984.

[4] V. I. Shalashilin and E. B. Kuznetsov, *Parametric continuation and optimal parametrization in applied mathematics and mechanics*. Dordrecht, Boston, London: Kluwer Academic Publishers, 2003.

[5] E. B. Kuznetsov, S. S. Leonov, and E. D. Tsapko, "Applying the best parameterization method and its modifications for numerical solving of some classes of singularly perturbed problems," vol. 274, 2022, pp. 311–330. DOI: 10.1007/978-981-16-8926-0_21.

[6] E. B. Kuznetsov, S. S. Leonov, and E. D. Tsapko, "A new numerical approach for solving initial value problems with exponential growth integral curves," in *IOP Conference Series: Materials Science and Engineering*, AMMAI'2020, IOP Publishing, 2020. DOI: 10.1088/1757-899X/927/1/012032.

[7] K. W. Chang and F. A. Howes, *Nonlinear singular perturbation phenomena: theory and application*. New York: Springer, 1984.

[8] R. Temam, *Navier–Stokes Equations: theory and numerical analysis*. AMS Chelsea Publishing, 1997.

[9] H. Schlichting and K. Gersten, *Boundary-layer theory*. Springer, 2016.

[10] P. Wesseling, *Principles of computational fluid dynamics*. Springer, 2009.

[11] G. A. Dahlquist, "A special stability problem for linear multistep methods," *BIT Numerical Mathematics*, vol. 3, no. 1, pp. 27–43, 1963. DOI: 10.1007/BF01963532.

[12] E. B. Kuznetsov, S. S. Leonov, and E. D. Tsapko, "Estimating the domain of absolute stability of a numerical scheme based on the method of solution continuation with respect to a parameter for solving stiff initial value problems," *Computational Mathematics and Mathematical Physics*, vol. 63, no. 4, pp. 557–572, 2023. DOI: 10.1134/S0965542523040115.

For citation:

E. D. Tsapko, S. S. Leonov, E. B. Kuznetsov, On application of solution continuation method with respect to the best exponential argument in solving stiff boundary value problems, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 375–386. DOI: 10.22363/2658-4670-2023-31-4-375-386.

Information about the authors:

Tsapko, Ekaterina D. — Support engineer for Visiology platform, JSC "IESC" "Energy-Efficient Technologies" (e-mail: zapkokaty@gmail.com, phone: +7(903)1234493, ORCID: https://orcid.org/0000-0002-4215-3510)

Leonov, Sergey S. — Candidate of Physical and Mathematical Sciences, Assistant Professor of Nikolsky Mathematical Institute of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University); Assistant Professor of Department of Mechatronics and Theoretical Mechanics of Moscow Aviation Institute (MAI) (e-mail: powerandglory@yandex.ru, phone: +7(903)2042261, ORCID: https://orcid.org/0000-0001-6077-0435)

Kuznetsov, Evgenii B. — Doctor of Physical and Mathematical Sciences, Professor of Department of Mechatronics and Theoretical Mechanics of Moscow Aviation Institute (MAI) (e-mail: kuznetsov@mai.com, phone: +7(916)2171999, ORCID: https://orcid.org/0000-0002-9452-6577)

УДК 519.624
PACS 02.60.Lj
DOI: 10.22363/2658-4670-2023-31-4-375-386
EDN: FPXPIV

О применении метода продолжения решения по экспоненциальному наилучшему аргументу для решения жёстких краевых задач

Е. Д. Цапко[1], **С. С. Леонов**[2,3], **Е. Б. Кузнецов**[3]

[1] *Акционерное общество «Межрегиональная энергосервисная компания «Энергоэффективные технологии», ул. Семашко, д. 12, стр. 8, Нижний Новгород, 603155, Российская Федерация*
[2] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*
[3] *Московский авиационный институт, Волоколамское шоссе, д. 4, Москва, 125993, Российская Федерация*

Аннотация. Процесс построения решения жёстких краевых задач пронизывает множество научных и инженерных дисциплин, требуя новаторских подходов для преодоления ограничений традиционных численных методов. В данном исследовании рассматривается реализация метода продолжения решения по наилучшему аргументу и модифицированному экспоненциальному наилучшему аргументу для решения жёстких задач, характеризующихся быстрорастущими интегральными кривыми. Исследование проводилось путём сравнения эффективности и устойчивости нового подхода с традиционным методом стрельбы. Результаты показывают значительное улучшение вычислительной эффективности при преобразовании задачи к экспоненциальному наилучшему аргументу. Особенно хорошо этот метод проявляет себя в сценариях, где интегральные кривые демонстрируют экспоненциальную скорость роста. Одним из ключевых выводов этого исследования является важная роль параметра регуляризации, выбор которого может определять эффективность решения. В целом, данное исследование предлагает новаторский метод решения жёстких краевых задач и подчёркивает тонкости выбора метода, что может указать путь для дальнейших усовершенствований и применений в различных областях.

Ключевые слова: жёсткие краевые задачи, метод продолжения решения, экспоненциальный наилучший аргумент, устойчивость численного метода, интегральные кривые, вычислительная эффективность, метод стрельбы, область абсолютной устойчивости

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 387–398
http://journals.rudn.ru/miph

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-4-387-398
EDN: EATOFG

On a set of tests for numerical methods of integrating differential equations, based on the Calogero system

Mikhail D. Malykh[1,2], **Wang Shiwei**[1], **Yu Ying**[3]

[1] *RUDN University,*
6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation
[2] *Joint Institute for Nuclear Research,*
6 Joliot-Curie St., Dubna, 141980, Russian Federation
[3] *Kaili University,*
3 Kaiyuan Rd., Kaili, 556011, People's Republic of China

(received: October 5, 2023; revised: December 7, 2023; accepted: December 29, 2023)

Abstract. Based on the completely integrable Calogero dynamical system, which describes the one-dimensional many-body problem, a tool for testing difference schemes has been developed and implemented in the original fdm package integrated into the Sage computer algebra system. This work shows how the developed tools can be used to examine the behavior of numerical solutions near the collision point and how to study the conservatism of the difference scheme. When detecting singularities using Alshina's method, a difficulty was discovered associated with false order fluctuations. One of the main advantages of this set of tests is the purely algebraic nature of the solutions and integrals of motion.

Key words and phrases: finite difference method, dynamical systems, Calogero system, numerical identification of singularities

1. Introduction

We are now developing a system for integrating ordinary differential equations in the Sage computer algebra system called fdm for sage [1] (`https://github.com/malykhmd/fdm`). The problem that arises when developing and implementing numerical methods for integrating dynamic systems is the limited number of test examples, most of them taken from classical monographs [2, 3]. Non-integrability itself is an important property of a dynamic system, although very difficult to formalize. Therefore, tests based on integrable systems are obviously doomed to be somewhat one-sided, which, at least at the present stage, cannot be corrected.

© Malykh M. D., Shiwei W., Ying Y., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

The second property of widely used tests is their small dimension. Concerning the many-body problem, for tests the two-body problem and special cases in which the three-body problem, limited or complete, has elementary or at least obviously periodic solutions are chosen. Although many such solutions have been found [4, 5], all of them are rather specific and, in particular, lack singular points — or rather — the collisions of the bodies. A very small number of papers [6] are devoted to the development of difference schemes that inherit periodicity. On the contrary, the question of numerically determining the position of moving singular points of the solution has been well elaborated [7–11]. However, this method has not actually been tested on mechanical problems of many bodies, although in these problems the bodies can approach each other at arbitrarily small distances and cases of false positive singularity tests can be expected. It is equally interesting to explore the question of whether these tests can handle multiple collisions of bodies.

On the contrary, the attention has always been focused at the preservation of symplectic structure and integrals of motion. In classical many-body problem the questions of integrals of motion are poor. All of these algebraic integrals, except the energy integral, are linear or quadratic and are preserved by any symplectic Runge–Kutta scheme, and the question of conservation of all integrals is reduced to the question of conservation of energy [12, 13]. Therefore, testing conservative difference schemes requires the development of tests based on Hamiltonian systems that have a large supply of algebraic integrals, and therefore a large dimension.

Among the integrable high-dimensional Hamiltonian systems, the most promising for creating this kind of tests is the Calogero system [14]. This system describes the motion of n particles of the same mass in one dimension, their interaction potential being inversely proportional to the square of the distance between the bodies. Its solution is described by algebraic functions of time, and the integrals of motion are rational functions. In fact, this system is the only one among many body problems that can be integrated for any number of bodies, and, moreover, the integration does not require any transcendental functions.

For this reason, when developing a set of tests in our fdm system, we gave a prominent place to tests based on the Calogero system. In this paper, we present a set of tools for testing integration methods based on the Calogero problem. This system is very convenient for implementation in computer algebra systems, since it has a purely algebraic properties of solutions and integrals of motion.

2. Tools for specifying the Calogero system

The Calogero system describes the motion of n material points of unit mass in one dimension, repelling or attracting each other with a force inversely proportional to the third power of the distance. Let us start numbering the points from zero. Let q_i be the position of the i-th point, then

$$\ddot{q}_i = -\frac{\partial U}{\partial q_i}, \quad i = 0, \ldots, n-1, \tag{1}$$

where

$$U = \sum_{i<j} V(q_i - q_j), \quad V(x) = \frac{b}{x^2}.$$

In fdm, the process of specifying the initial problem is separated from the application of the numerical method. To describe the initial problem, a special class `Initial_problem` is used, described in [1]. To set the initial problem for the Calogero system, the `calogero_problem` function has been added to fdm, which has 4 optional arguments:

— `ics` is the list of initial values, with the positions of the bodies coming first, and then their momenta,
— `n` is the number of bodies,
— `T` is the final time; we always take $t = 0$ for the initial time,
— `b` is the value of parameter b, default $b = -1$.

As an illustration, let us take the problem of 5 bodies, which at the initial moment of time occupy the positions

$$q_i(0) = i, \quad i = 0, \ldots, n-1.$$

Let us take the velocities close to zero, and change t in the interval $t \in [0, 0.5]$. The time unit in this article is seconds. This problem in our system is set as follows:

```
n=5
ics=list(range(n))+[0.1,0.2,-0.1,0,0]
problem_calogero=calogero_problem(ics=ics,n=n, T=0.5, b=-1)
```

It can be solved using standard fdm tools, for example, using the 4th order Runge–Kutta method and plotting the dependence of q_0 on t:

```
sol=erk(problem_calogero,N=100)
```

This function is described in [1]. Figure 1 presents the result of the calculations.

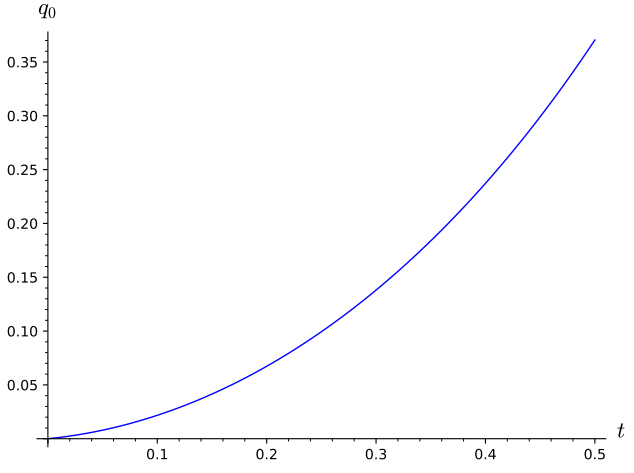

Figure 1. The problem of 5 bodies

3. Tools for the analytical solution of the Calogero system

The exact solution is described by means of the Lax pair [14]. The solution $q_0(t), \ldots, q_{n-1}(t)$ is a set of eigenvalues of the matrix

$$\mathbf{M} = \mathbf{Q}|_{t=0} + t\mathbf{L}|_{t=0},$$

where

$$\mathbf{Q} = \mathrm{diag}(q_0, \ldots q_{n-1})$$

and

$$\mathbf{L} = \mathrm{diag}(p_0, p_2, \ldots, p_{n-1}) + \sqrt{-b}\left(\frac{1-\delta_{jk}}{q_j - q_k}\right).$$

Therefore, calculating the coordinates of bodies at time t is reduced to calculating the eigenvalues of the matrix $\mathbf{M}$, i.e., they are the zeros of the polynomial

$$F = \det(\mathbf{M} - q\mathbf{E}) \tag{2}$$

of symbolic variable q.

For a fixed value of t, these calculations are implemented as a function `calogero_q`, which has two required arguments `ics` and `n` and two optional arguments `b` and `t`. Eigenvalues are calculated in the algebraic number field. Unfortunately, the order of the eigenvalues may not coincide with the numbering order of the bodies. Therefore, this function returns the coordinates of the bodies up to some permutation.

For example, for the problem of 5 bodies considered above at time $t = 0.1$, we can find the positions of the bodies as follows:

```
calogero_q(ics,n,b=-1,t=0.1)
[0.02174568377617522?,
 1.022008478256518?,
 1.989567507785786?,
 2.998508164438269?,
 3.988170165743252?]
```

This can be compared with the values of q_i found using the Runge–Kutta method:

```
Q=[symbolic_expression('q'+str(i)) for i in range(n)]
[sol.value(q,0.1) for q in Q]
[0.0217456837809339,
 1.02200847823443,
 1.98956750781722,
 2.99850816441585,
 3.98817016575156]
```

The main disadvantage of this tool is the calculation of the determinant of the matrix $\mathbf{M} - q\mathbf{E}$. For large n, this operation becomes very costly.

Function `calogero_curve` returns the polynomial (2) itself. For example,

```
calogero_curve(ics,n,b=-1)
-q^5 + 1/5*q^4*t - 18089/3600*q^3*t^2
+ 23053/36000*q^2*t^3 - 526651/129600*q*t^4
+ 104483/1296000*t^5 + 10*q^4 - 2*q^3*t
```

```
+ 18101/600*q^2*t^2 - 5347/2000*q*t^3
+ 266021/32400*t^4 - 35*q^3 + 34/5*q^2*t
- 196199/3600*q*t^2 + 6829/4000*t^3 + 50*q^2
- 43/5*q*t + 51523/1800*t^2 - 24*q + 12/5*t
```

Standard Sage tools allow plotting q_i versus time and compare the plots with the results of calculations using a difference scheme.

For example, for the considered example of the 5-body problem, the plots of q_0 and q_1 versus t, obtained numerically (solid line) and analytically (dashed line) can be displayed in one figure as follows:

```
F=calogero_curve(ics,n)
sol.plot(t,q0)+sol.plot(t,q1)+\
implicit_plot(F,(t,0,0.7),(q,0,1.2), color='red', linestyle='--
',
axes_labels=['$t$','$q$'], aspect_ratio=1/2)
```

The result is presented in figure 2. It is clearly seen that the bodies collide at $t \simeq 0.6$. The function `implicit_plot` is used when drawing graphics, which can quickly draw contour maps, so different types of curves in the figure represent different contours rather than different particle trajectories.

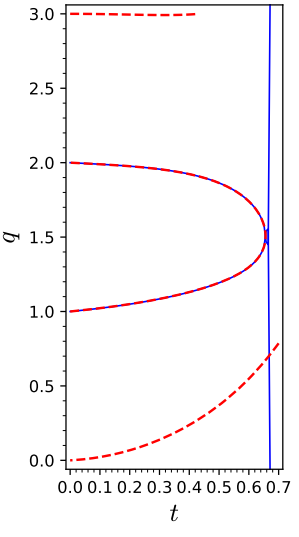

Figure 2. Problem of 5 bodies

4. Collision of bodies

Collisions of bodies occur when the polynomial (2) has multiple roots. Therefore, the first collision of bodies can be calculated as the smallest positive value of t for which the equations

$$F = 0, \quad \frac{\partial F}{\partial q} = 0$$

are compatible. Function `calogero_solution_crash` allows finding it exactly in algebraic numbers. For example,

```
calogero_solution_crash(ics,n)
0.618840603733536?
```

Unless making special efforts, at such a point

$$\frac{\partial^2 F}{\partial q^2} \neq 0.$$

From the implicit function theorem it follows that at the collision point the coordinates of the bodies have an algebraic singularity of the order of $1/2$.

Known the moment of impact, it is possible to observe the behavior of a particular numerical method near the point of impact. For example, let us increase the final time in the example considered to $t = 0.7$ and make the plots:

```
problem_calogero=calogero_problem(ics=ics,n=n, T=0.7, b=-1)
sol=erk(problem_calogero,N=100)
pl=sol.plot(t,q0)+sol.plot(t,q1, axes_labels=['$t$','$q$'])
pl.show(ymax=1, ymin=0)
```

The result of the calculations is presented in figure 3.

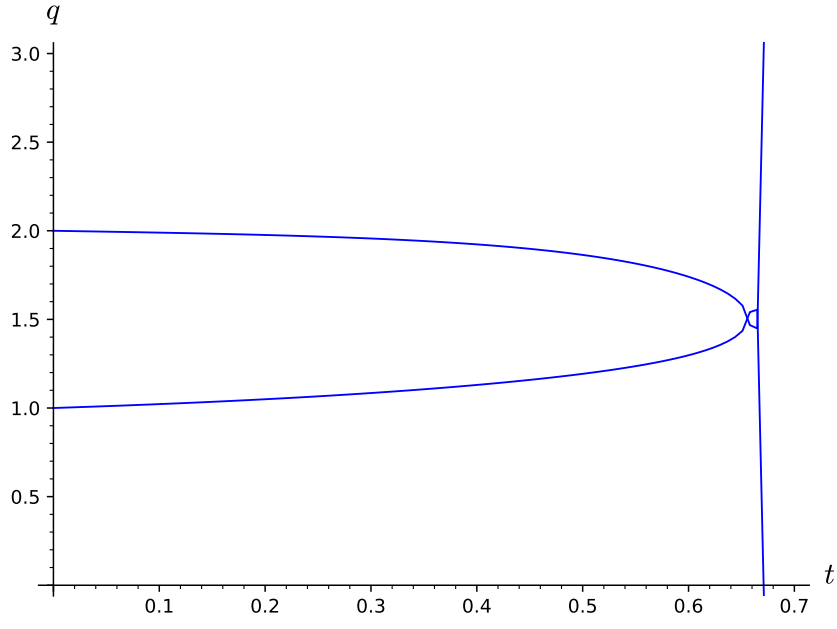

Figure 3. Problem of 5 bodies. Collision of bodies q_1 и q_2

In fdm, there is a tool for numerically determining the position and order of the solution singularity using the Alshina method [7–10]. Previously, we tested it on the two-body problem [11]; testing on problems with a large number of bodies indicated the inefficiency of our implementation and forced us to significantly update it. The syntax remains the same. Function

```
eff_order(problem_calogero,q0,N=100)
```

returns the plot 4. The dotted line in the figure represents the theoretical order, which does not change with time.

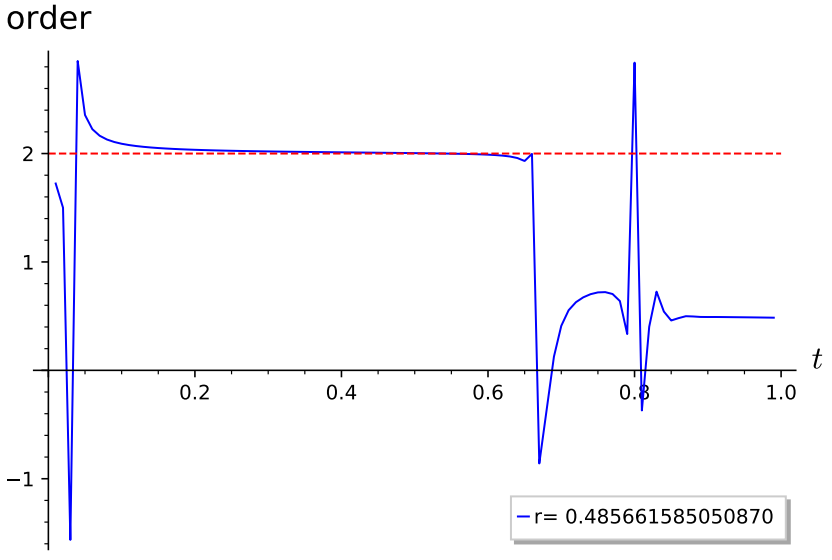

Figure 4. Problem of 5 bodies, effective order according to Alshina

Let us recall how to read this plot [11]. The CROS scheme on which Alshina's method is based is a second-order scheme. Up to the point of collision, the scheme order, calculated approximately, coincides with the theoretical one, that is, equals 2. Near the collision point, this rule ceases to apply and, which is a specific property of the CROS circuit, the order jumps sharply to a new constant value, which is equal to the order of this special point [7]. In figure 4 we actually see that at the point of impact the order changes sharply from a value of 2 to some value close to $\frac{1}{2}$. However, in addition to this jump, there are two more 'spikes', at the very beginning of the graph and in the region of $t = 0.8$. Near these two points, the order changes sharply, but eventually returns to its original value. The theory does not explain the appearance of these artifacts, which require further research.

We plan to implement and test other numerical methods in our system for identifying singularities [15, 16].

5. Preservation of integrals of motion

As noted above, the question of preserving all integrals of motion by a difference scheme has been fruitfully discussed for a long time. The Calogero system has n rational integrals of motion being in involution and, therefore, is completely Liouville integrable [14]. These integrals can be described as traces of the matrix $\mathbf{L}$ powers:

$$F_k = \mathrm{Sp}\ \mathbf{L}^k, \quad (k = 1, \ldots n).$$

They are, of course, symmetric functions with respect to the group of permutations of bodies.

The calculation of these integrals is implemented as the function `calogero_integral(k,n)`. For example,

```
calogero_integral(1,n)
p0 + p1 + p2 + p3 + p4
```

It should be noted that for $b > 0$ the matrix $\mathbf{L}$ is complex, and the traces of its degree are real. Our function performs the appropriate simplification and returns a rational function with real coefficients.

The integral F_1 is linear, so all Runge–Kutta methods preserve it exactly. The energy integral F_2 is not quadratic for the Calogero system, so even symplectic Runge–Kutta schemes do not preserve it [3]. Figure 5 shows the dependences of F_2 and F_3 for our 5-body problem; it is clearly seen that these functions monotonically increase in absolute value on approaching the collision point.

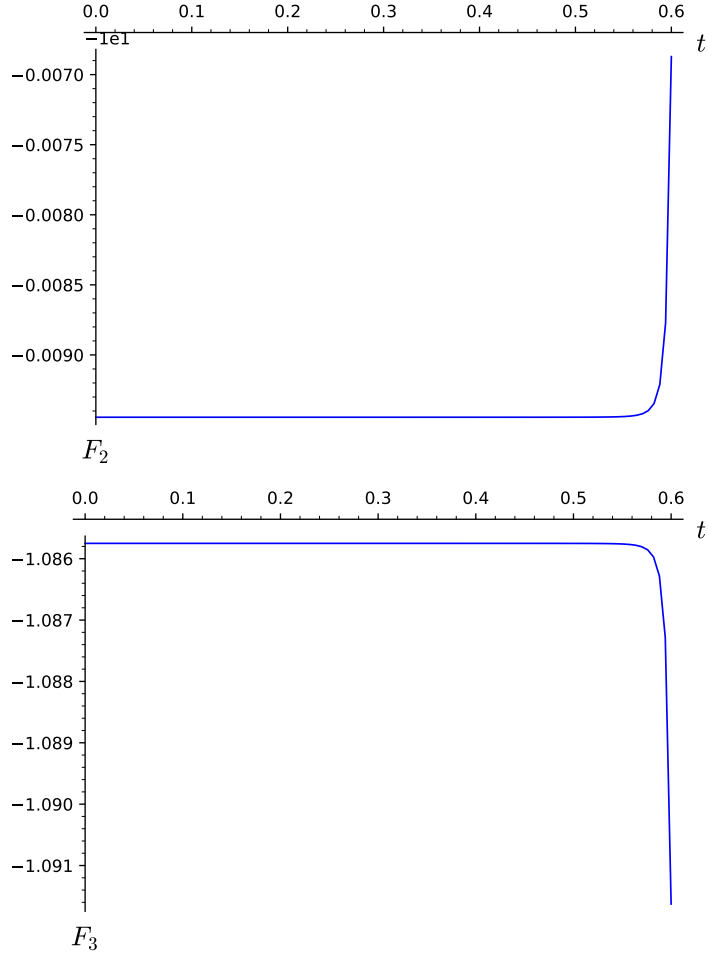

Figure 5. 5-body problem, integrals F_2 and F_3 in the explicit Runge–Kutta scheme

It should also be noted that sharp jumps in the integrals, which are often observed when applying explicit schemes to the classical, three-dimensional many-body problem, did not appear in these plots. This suggests that the Calogero problem, whose solutions are algebraic functions t, is much simpler than the classical many-body problem and, therefore, the test based on it does not detect this feature of explicit schemes.

6. Conclusion

The completely integrable dynamical Calogero system, which describes the one-dimensional many-body problem, allows creating a convenient tool for testing difference schemes by means that do not go beyond the algebraic framework. We have implemented these tools in the new version of the fdm for sage software.

The first application showed that the implementation of the method for numerical detection of moving singularities is complicated by artifacts, the false spikes in the order plot, not previously described in theory. This allows an idea that the developed tools provide a fairly representative set of tests that will allow revealing previously unnoticed difficulties.

On the other hand, it should be noted that the trajectories of bodies in this dynamic system are arranged quite simply, without any 'loops' and 'fine structures', which makes this test rather rough. In particular, we associate this with the absence of jumps of integrals of motion in the plots calculated using explicit schemes. This offers a prospect for the development of new tests.

Acknowledgments

The work was carried out with the financial support of the Russian Science Foundation (project No. 20-11-20257).

References

[1] A. Baddour, M. M. Gambaryan, L. Gonzalez, and M. D. Malykh, "On implementation of numerical methods for solving ordinary differential equations in computer algebra systems," *Programming and Computer Software*, vol. 5, pp. 412–422, 2023. DOI: 10.1134/S0361768823020044.

[2] E. Hairer, G. Wanner, and S. P. Nørsett, *Solving ordinary differential equations I, Nonstiff Problems*, 3rd ed. Springer, 2008. DOI: 10.1007/978-3-540-78862-1.

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

[4] X. Li and S. Liao, "More than six hundred new families of Newtonian periodic planar collisionless three-body orbits," *Sci. China-Phys. Mech. Astro.*, vol. 60, no. 12, p. 129 511, 2017.

[5] I. Hristov, R. Hristova, V. Dmitrašinović, and K. Tanikawa, "Three-body periodic collisionless equal-mass free-fall orbits revisited," *Arxive*, vol. 2308.16159v1, 2023. DOI: arXiv.2308.16159.

[6] A. Baddour, M. Malykh, and L. Sevastianov, "On periodic approximate solutions of dynamical systems with quadratic right-hand side," *Journal of Mathematical Sciences*, vol. 261, pp. 698–708, 2022. DOI: 10.1007/s10958-022-05781-4.

[7] E. A. Alshina, N. N. Kalitkin, and P. V. Koryakin, "Diagnostics of singularities of exact solutions in computations with error control," *Computational Mathematics and Mathematical Physics*, vol. 45, no. 10, pp. 1769–1779, 2005.

[8] M. O. Korpusov, D. V. Lukyanenko, A. A. Panin, and E. V. Yushkov, "Blow-up for one Sobolev problem: theoretical approach and numerical analysis," *Journal of Mathematical Analysis and Applications*, vol. 442, no. 2, pp. 451–468, 2016. DOI: 10.26089/NumMet.v20r328.

[9] M. O. Korpusov, D. V. Lukyanenko, A. A. Panin, and E. V. Yushkov, "Blow-up phenomena in the model of a space charge stratification in semiconductors: analytical and numerical analysis," *Mathematical Methods in the Applied Sciences*, vol. 40, no. 7, pp. 2336–2346, 2017. DOI: 10.1002/mma.4142.

[10] M. O. Korpusov, D. V. Lukyanenko, A. A. Panin, and G. I. Shlyapugin, "On the blow-up phenomena for a one-dimensional equation of ion-sound waves in a plasma: analytical and numerical investigation," *Mathematical Methods in the Applied Sciences*, vol. 41, no. 8, pp. 2906–2929, 2018. DOI: 10.1002/mma.4791.

[11] A. Baddour, A. A. Panin, L. A. Sevastianov, and M. D. Malykha, "Numerical determination of the singularity order of a system of differential equations," *Discrete and Continuous Models and Applied Computational Science*, vol. 28, no. 1, pp. 17–34, 2020. DOI: 10.22363/2658-4670-2020-28-1-17-34.

[12] Y. Ying, A. Baddour, V. P. Gerdt, M. Malykh, and L. Sevastianov, "On the quadratization of the integrals for the many-body problem," *Mathematics*, vol. 9, no. 24, 2021. DOI: 10.3390/math9243208.

[13] A. Baddour and M. Malykh, "On difference schemes for the many-body problem preserving all algebraic integrals," *PPhysics of Particles and Nuclei, Letters*, vol. 19, pp. 77–80, 2022. DOI: 10 . 1134 / S1547477122010022.

[14] J. Moser, *Integrable Hamiltonian systems and spectral theory*. Edizioni della Normale, 1983.

[15] A. A. Belov, "Numerical detection and study of singularities in solutions of differential equations," *Doklady Mathematics*, vol. 93, no. 3, pp. 334–338, 2016. DOI: 10.1134/S1064562416020010.

[16] A. A. Belov, "Numerical diagnostics of solution blowup in differential equations," *Computational Mathematics and Mathematical Physics*, vol. 57, no. 1, pp. 122–132, 2017. DOI: 10.1134/S0965542517010031.

For citation:

M. D. Malykh, W. Shiwei, Y. Ying, On a set of tests for numerical methods of integrating differential equations, based on the Calogero system, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 387–398. DOI: 10.22363/2658-4670-2023-31-4-387-398.

Information about the authors:

Malykh, Mikhail D. — Doctor of Physical and Mathematical Sciences, Assistant Professor of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: malykh-md@rudn.ru, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0001-6541-6603, ResearcherID: P-8123-2016, Scopus Author ID: 6602318510)

Wang Shiwei — Ph.D. student of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: 1995wsw@gmail.com, ORCID: https://orcid.org/0009-0007-6504-8370)

Yu Ying — Assistant Professor of Department of Mathematics and Applied Mathematics, Kaili University, China (e-mail: 45384377@qq.com, ORCID: https://orcid.org/0000-0002-4105-2566)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-4-387-398
EDN: EATOFG

О наборе тестов для численных методов интегрирования дифференциальных уравнений, основанном на системе Калоджеро

М. Д. Малых[1,2], **Ван Шивэй**[1], **Юй Ин**[3]

[1] *Российский университет дружбы народов,*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация
[2] *Объединённый институт ядерных исследований,*
ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация
[3] *Университет Кайли,*
Кайюань Роуд, д. 3, Кайли, 556011, Китайская Народная Республика

Аннотация. На основе вполне интегрируемой динамической системы Калоджеро, описывающей одномерную задачу многих тел, разработан инструмент для тестирования разностных схем и реализован в оригинальном пакете fdm, интегрируемом в систему компьютерной алгебры Sage. Показано, как использовать разработанные инструменты для проверки поведения численных решений возле точек столкновения, а также для исследования консервативности разностных схем. При обнаружении особенностей по методу Альшиной обнаружена трудность, связанная с ложными колебаниями порядка. Одно из главных достоинств этого набора теста — чисто алгебраический характер решений и интегралов движения.

Ключевые слова: метод конечных разностей, динамические системы, система Калоджеро, численная идентификация особенностей

Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2023, 31 (4) 399–418
http://journals.rudn.ru/miph

Research article

UDC 537.8:514.762.37
DOI: 10.22363/2658-4670-2023-31-4-399-418
EDN: GCUXWK

Methodological derivation of the eikonal equation

Arseny V. Fedorov[1], Christina A. Stepa[1], Anna V. Korolkova[1], Migran N. Gevorkyan[1], Dmitry S. Kulyabov[1,2]

[1] *RUDN University,*
6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation
[2] *Joint Institute for Nuclear Research,*
6 Joliot-Curie St., Dubna, 141980, Russian Federation

(received: December 1, 2023; revised: December 19, 2023; accepted: December 29, 2023)

Abstract. Usually, when working with the eikonal equation, reference is made to its derivation in the monograph by Born and Wolf. The derivation of this equation was done rather carelessly. Understanding this derivation requires a certain number of implicit assumptions. For a better understanding of the eikonal approximation and for methodological purposes, the authors decided to repeat the derivation of the eikonal equation, explicating all possible assumptions. Methodically, the following algorithm for deriving the eikonal equation is proposed. The wave equation is derived from Maxwell's equation. In this case, all conditions are explicitly introduced under which it is possible to do this. Further, from the wave equation, the transition to the Helmholtz equation is carried out. From the Helmholtz equation, with the application of certain assumptions, a transition is made to the eikonal equation. After analyzing all the assumptions and steps, the transition from the Maxwell's equations to the eikonal equation is actually implemented. When deriving the eikonal equation, several formalisms are used. The standard formalism of vector analysis is used as the first formalism. Maxwell's equations and the eikonal equation are written as three-dimensional vectors. After that, both the Maxwell's equations and the eikonal equation use the covariant 4-dimensional formalism. The result of the work is a methodically consistent description of the eikonal equation.

Key words and phrases: eikonal, Maxwell's equations, wave equation, vector representation, tensor representation

1. Introduction

One of the foundations of the simulation program we employ for modeling optical phenomena is the eikonal model [1, 2]. While this model is well-known, the derivation process is somewhat intricate [3, 4]. In the renowned monograph by Born and Wolf [5], the derivation appears almost like a form of physical magic (here are Maxwell's equations, a bit of magic, and voila, we have the eikonal equation). We decided to delve deeper into the derivation of the eikonal equation.

© Fedorov A. V., Stepa C. A., Korolkova A. V., Gevorkyan M. N., Kulyabov D. S., 2023

This work is licensed under a Creative Commons Attribution 4.0 International License
https://creativecommons.org/licenses/by-nc/4.0/legalcode

We employed analytical methods to derive the eikonal equation from Maxwell's equations in a medium without currents and charges. The process involves analyzing differential equations and applying methods of mathematical analysis. A brief outline of our study is presented in the scheme shown in the figure 1.

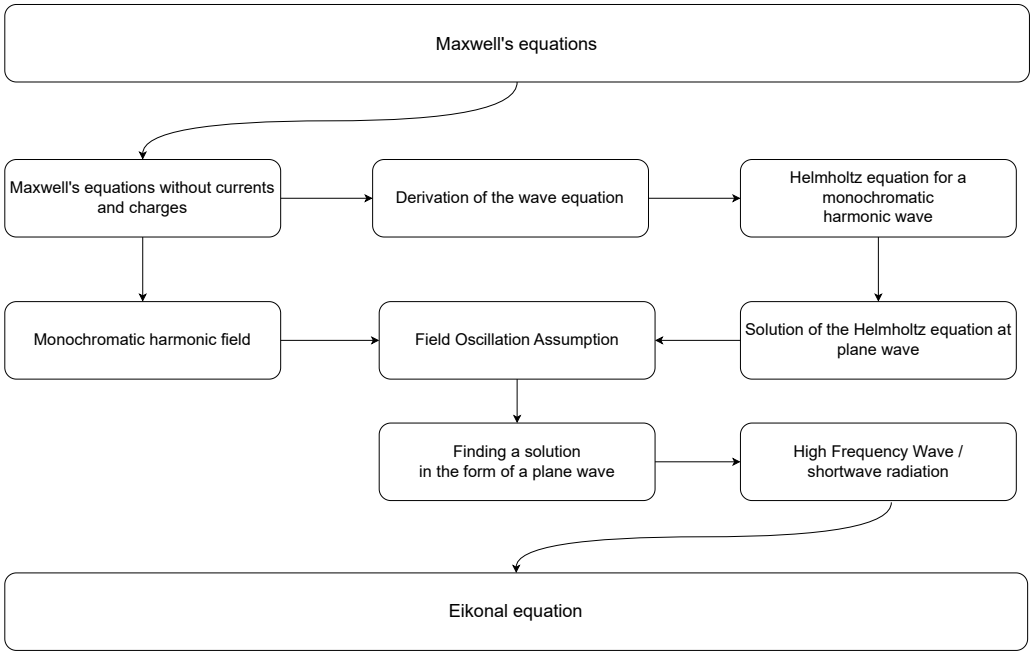

Figure 1. Paper structure

1.1. Article structure

In section 1.2, we present basic notations and conventions used in the article. In section 2, fundamental relationships for Maxwell's equations are introduced. In section 3, the wave equation is derived from Maxwell's equations. Next, in section 4, the eikonal equation is obtained from the wave equation. The transformations are performed using vector formalism. In section 5, the same is done based on covariant tensor formalism.

1.2. Notations and conventions

1. The primary mathematical framework used in the article is the vector analysis (a brief overview is given in Appendix) and tensor analysis.

2. We will adhere to the following conventions. Greek indices (α, β) will refer to the four-dimensional space, with the component values as follows: $\alpha = \overline{0,3}$. Latin indices from the middle of the alphabet (i, j, k) will refer to the three-dimensional space, with the component values as follows: $i = \overline{1,3}$.
3. The CGS symmetrical system [6] is used for notating the equations of electrodynamics.

2. Introduction

Consider Maxwell's equations in vector-differential form:

$$\boldsymbol{\nabla} \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} = \frac{4\pi}{c}\mathbf{j}, \tag{1}$$

$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}, \tag{2}$$

$$\boldsymbol{\nabla} \cdot \mathbf{D} = 4\pi\rho, \tag{3}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0, \tag{4}$$

where

— $\mathbf{E}(\mathbf{r}, t) = \mathbf{E}(x, y, z, t)$ is electric field strength vector;
— $\mathbf{H}(\mathbf{r}, t) = \mathbf{H}(x, y, z, t)$ is magnetic field strength vector;
— $\mathbf{D}(\mathbf{r}, t) = \mathbf{D}(x, y, z, t)$ is electric field induction vector;
— $\mathbf{B}(\mathbf{r}, t) = \mathbf{B}(x, y, z, t)$ is magnetic field induction vector;
— $\mathbf{j}(\mathbf{r}) = \mathbf{j}(x, y, z)$ is external electric current density (current strength per unit area);
— $\rho(\mathbf{r}) = \rho(x, y, z)$ is electric charge density;
— c is vacuum speed of light;
— $\mathbf{r} = (x, y, z)^T$ is radius vector of a point, written in Cartesian coordinates.

Let us briefly describe the physical meaning of each of Maxwell's equations:

— equation (1) means that electric current and a change in electric induction generate a solenoidal magnetic field, that is a field whose field lines twist into a vortex along the vector indicating the direction of the current;
— equation (2) means that a change in time of a magnetic field generates an electric field;
— equation (3) means that the electric charge is the source of electrical induction;
— equation (4) means that there are no free magnetic poles (only magnetic dipoles have been experimentally discovered, magnetic monopoles are not known to science).

The following relations, called *material equations*, are also valid:

$$\mathbf{j} = \sigma\mathbf{E}, \quad \mathbf{D} = \varepsilon\mathbf{E}, \quad \mathbf{B} = \mu\mathbf{H},$$

where $\sigma(\mathbf{r})$ is conductivity, $\varepsilon(\mathbf{r})$ is permittivity, and $\mu(\mathbf{r})$ is permeability. In an isotropic medium, ε and μ are scalar quantities, but in the general case they are tensor quantities.

A medium is called *isotropic* if its physical properties do not depend on direction. The term comes from the Greek words «izos» (ισος): equal, identical, similar and "tropos" (τροπος): direction, character. In electrodynamics, the isotropy of a medium is associated with the same values of $\varepsilon(\mathbf{r})$ and $\mu(\mathbf{r})$ in all directions.

Magnetic permeability characterizes the magnetic properties of a medium (substance). If $\mu \neq 1$, then the substance is called *magnetic*, if $\mu > 1$ — *paramagnetic*, if $\mu < 1$ — *diamagnetic*.

Next, we will consider a medium that does not conduct electricity, that is, $\sigma = 0$, and also free from currents, that is, $\mathbf{j} = \mathbf{0}$ and $\rho = 0$, then Maxwell's equations simplified somewhat:

$$\boldsymbol{\nabla} \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} = \mathbf{0}, \tag{5}$$

$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}, \tag{6}$$

$$\boldsymbol{\nabla} \cdot \mathbf{D} = 0, \tag{7}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0. \tag{8}$$

3. Wave equation

3.1. Derivation of the wave equation from Maxwell's equations

We will assume that $\rho = 0$ and $\mathbf{j} = \mathbf{0}$ and consider the equations (5) and (6):

$$\boldsymbol{\nabla} \times \mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t} = \mathbf{0},$$
$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}.$$

We use the material equations $\mathbf{D} = \varepsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ and take into account the dependence of the permittivity and permeability on coordinates: $\varepsilon = \varepsilon(x, y, z)$ and $\mu = \mu(x, y, z)$.

$$\boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial}{\partial t}(\mu\mathbf{H}) = \mathbf{0} \Rightarrow \boldsymbol{\nabla} \times \mathbf{E} + \frac{\mu}{c}\frac{\partial \mathbf{H}}{\partial t} = \mathbf{0} \Rightarrow \frac{1}{\mu}\boldsymbol{\nabla} \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{H}}{\partial t} = \mathbf{0}.$$

Apply the curl operator $\boldsymbol{\nabla}\times$ to the resulting equation:

$$\underbrace{\boldsymbol{\nabla} \times \left(\frac{1}{\mu}\boldsymbol{\nabla} \times \mathbf{E}\right)}_{\text{(I)}} + \underbrace{\frac{1}{c}\boldsymbol{\nabla} \times \frac{\partial \mathbf{H}}{\partial t}}_{\text{(II)}} = \mathbf{0}.$$

Let us first consider term (II) of this equation. The time derivative can be taken out from under the sign of the rotor operator:

$$\boldsymbol{\nabla} \times \frac{\partial \mathbf{H}}{\partial t} = \frac{\partial}{\partial t}(\boldsymbol{\nabla} \times \mathbf{H}).$$

Due to (5) we get:

$$\frac{\partial}{\partial t}(\boldsymbol{\nabla}\times\mathbf{H})=\frac{\partial}{\partial t}\frac{1}{c}\frac{\partial\mathbf{D}}{\partial t}=\frac{1}{c}\frac{\partial^2\mathbf{D}}{\partial t^2}.$$

We use the material equation $\mathbf{D}=\varepsilon\mathbf{E}$ and write as follows:

$$\frac{\partial^2\mathbf{D}}{\partial t^2}=\frac{\partial^2\left(\varepsilon(\mathbf{r})\mathbf{E}\right)}{\partial t^2}=\varepsilon\frac{\partial^2\mathbf{E}}{\partial t^2}\Rightarrow\frac{1}{c}\boldsymbol{\nabla}\times\frac{\partial\mathbf{H}}{\partial t}=\frac{\varepsilon}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}.$$

To simplify term (I), we use the relation $\boldsymbol{\nabla}\times f\mathbf{v}=f\boldsymbol{\nabla}\times\mathbf{v}+\boldsymbol{\nabla}f\times\mathbf{v}$, where $f(x,y,z)$ is a scalar function, and $\mathbf{v}(x,y,z)$ is a vector field. Using this relation, term (I) is expanded as follows:

$$\boldsymbol{\nabla}\times\left(\frac{1}{\mu}\boldsymbol{\nabla}\times\mathbf{E}\right)=\underbrace{\frac{1}{\mu}\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{E})}_{\text{(I.a)}}+\underbrace{\left(\boldsymbol{\nabla}\frac{1}{\mu},\boldsymbol{\nabla}\times\mathbf{E}\right)}_{\text{(I.b)}}.$$

In turn, to simplify term (I.a) we use the identity $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\times\mathbf{v}=\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{v})-\nabla^2\mathbf{v}$, where ∇^2 is the Laplace operator.

$$\frac{1}{\mu}\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{E})=\frac{1}{\mu}\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E})-\frac{1}{\mu}\nabla^2\mathbf{E}.$$

To simplify the expression $\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E})$ we apply the identity $\boldsymbol{\nabla}\cdot(f\mathbf{v})=f\boldsymbol{\nabla}\cdot\mathbf{v}+(\boldsymbol{\nabla}f,\mathbf{v})$ to Maxwell's equation (7), replacing induction $\mathbf{D}$ with tension using the material equation $\mathbf{D}=\varepsilon\mathbf{E}$:

$$\boldsymbol{\nabla}\cdot\mathbf{D}=\boldsymbol{\nabla}\cdot(\varepsilon\mathbf{E})=\varepsilon\boldsymbol{\nabla}\cdot\mathbf{E}+(\mathbf{E},\boldsymbol{\nabla}\varepsilon)=0\Rightarrow$$
$$\Rightarrow\boldsymbol{\nabla}\cdot\mathbf{E}=-\frac{1}{\varepsilon}(\boldsymbol{\nabla}\varepsilon,\mathbf{E})\Rightarrow\frac{1}{\mu}\boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\mathbf{E})=-\frac{1}{\mu}\boldsymbol{\nabla}\left(\frac{1}{\varepsilon}(\boldsymbol{\nabla}\varepsilon,\mathbf{E})\right).$$

As a result, the term (I.a) was transformed to the following form:

$$\frac{1}{\mu}\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\times\mathbf{E})=-\frac{1}{\mu}\boldsymbol{\nabla}\left(\frac{1}{\varepsilon}(\boldsymbol{\nabla}\varepsilon,\mathbf{E})\right)-\frac{1}{\mu}\nabla^2\mathbf{E}.$$

By combining (I) and (II) we write:

$$\underbrace{-\frac{1}{\mu}\nabla^2\mathbf{E}-\frac{1}{\mu}\boldsymbol{\nabla}\left(\frac{1}{\varepsilon}(\boldsymbol{\nabla}\varepsilon,\mathbf{E})\right)}_{\text{(I.a)}}+\underbrace{\left(\boldsymbol{\nabla}\frac{1}{\mu},\boldsymbol{\nabla}\times\mathbf{E}\right)}_{\text{(I.b)}}+\underbrace{\frac{\varepsilon}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}}_{\text{(II)}}=0,$$

$$-\nabla^2\mathbf{E}-\boldsymbol{\nabla}\left(\frac{1}{\varepsilon}(\boldsymbol{\nabla}\varepsilon,\mathbf{E})\right)+\mu\left(\boldsymbol{\nabla}\frac{1}{\mu},\boldsymbol{\nabla}\times\mathbf{E}\right)+\frac{\mu\varepsilon}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}=0,$$

$$\nabla^2\mathbf{E}-\frac{\mu\varepsilon}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2}+\left[\boldsymbol{\nabla}\left(\frac{1}{\varepsilon}(\boldsymbol{\nabla}\varepsilon,\mathbf{E})\right)-\mu\left(\boldsymbol{\nabla}\frac{1}{\mu},\boldsymbol{\nabla}\times\mathbf{E}\right)\right]=0.$$

A completely similar equation can be obtained for the magnetic field strength vector $\mathbf{H}$.

In the case of an isotropic medium, that is, $\varepsilon = \mu = \text{const}$ the additional term taken in square brackets vanishes and we obtain the wave equation:

$$\nabla^2\mathbf{E} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} = \mathbf{0},$$
$$\nabla^2\mathbf{H} - \frac{\varepsilon\mu}{c^2}\frac{\partial^2\mathbf{H}}{\partial t^2} = \mathbf{0}.$$

We can introduce the quantity $v = c/\sqrt{\varepsilon\mu}$ — the speed of the electromagnetic wave in the medium.

3.2. The case of a plane wave

Consider the wave equation:

$$\nabla^2\mathbf{U} - \frac{1}{v^2}\frac{\partial^2\mathbf{U}}{\partial t^2} = 0.$$

Let's consider an electromagnetic wave that propagates in the direction $\mathbf{s}$, where $\mathbf{s} = (s_x, s_y, s_z)$ — some unit vector ($\|\mathbf{s}\| = 1$) fixed direction. Any solution of this equation, having the form $\mathbf{U} = \mathbf{U}((\mathbf{r}, \mathbf{s}), t)$ is a *plane* wave, since at every moment of time the vector $\mathbf{U}$ is constant in the plane $(\mathbf{r}, \mathbf{s}) = -d$, where $|d|$ is the distance from the plane to the origin. The expression $(\mathbf{r}, \mathbf{s}) = -d$ is actually a normal plane equation, where the vector $\mathbf{s}$ acts as the unit normal vector. Let's write it in Cartesian coordinates:

$$s_x x + s_y y + s_z z + d = 0.$$

The wave equation can be simplified by introducing a new coordinate system. Since the intensity vector of a plane wave entirely depends only on the distance d, we can choose a new coordinate system with axes $O\xi, O\eta, O\zeta$ so that the $O\zeta$ axis is directed along the vector $\mathbf{s}$, and the origin coincides with the previous Cartesian system $Oxyz$. Then, the coordinate along the $O\zeta$ axis will depend on the previous coordinates according to the formula $\zeta(x, y, z) = (\mathbf{r}, \mathbf{s}) = s_x x + s_y y + s_z z$, while ξ and η do not depend on the previous coordinates and can be chosen arbitrarily, for example, so that the coordinate system $O\xi\eta\zeta$ is right-handed (see the figure 2).

The replacement of differential operators is carried out using the Jacobian matrix as follows:

$$\begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial\xi}{\partial x} & \frac{\partial\eta}{\partial x} & \frac{\partial\zeta}{\partial x} \\ \frac{\partial\xi}{\partial y} & \frac{\partial\eta}{\partial y} & \frac{\partial\zeta}{\partial y} \\ \frac{\partial\xi}{\partial z} & \frac{\partial\eta}{\partial z} & \frac{\partial\zeta}{\partial z} \end{bmatrix} \begin{bmatrix} \frac{\partial}{\partial\xi} \\ \frac{\partial}{\partial\eta} \\ \frac{\partial}{\partial\zeta} \end{bmatrix}; \quad \left(\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}\right)^T = \begin{bmatrix} \frac{\partial\xi}{\partial x} & \frac{\partial\eta}{\partial x} & \frac{\partial\zeta}{\partial x} \\ \frac{\partial\xi}{\partial y} & \frac{\partial\eta}{\partial y} & \frac{\partial\zeta}{\partial y} \\ \frac{\partial\xi}{\partial z} & \frac{\partial\eta}{\partial z} & \frac{\partial\zeta}{\partial z} \end{bmatrix}.$$

Since $\xi = \text{const}$ and $\eta = \text{const}$, and $\zeta = (\mathbf{r}, \mathbf{s})$, then:

$$\begin{bmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \\ \dfrac{\partial}{\partial z} \end{bmatrix} = \begin{bmatrix} 0 & 0 & s_x \\ 0 & 0 & s_y \\ 0 & 0 & s_z \end{bmatrix} \begin{bmatrix} \dfrac{\partial}{\partial \xi} \\ \dfrac{\partial}{\partial \eta} \\ \dfrac{\partial}{\partial \zeta} \end{bmatrix} = \begin{bmatrix} s_x \dfrac{\partial}{\partial \zeta} \\ s_y \dfrac{\partial}{\partial \zeta} \\ s_z \dfrac{\partial}{\partial \zeta} \end{bmatrix} \Longrightarrow \begin{cases} \dfrac{\partial}{\partial x} = s_x \dfrac{\partial}{\partial \zeta}, \\ \dfrac{\partial}{\partial y} = s_y \dfrac{\partial}{\partial \zeta}, \\ \dfrac{\partial}{\partial z} = s_z \dfrac{\partial}{\partial \zeta}. \end{cases}$$

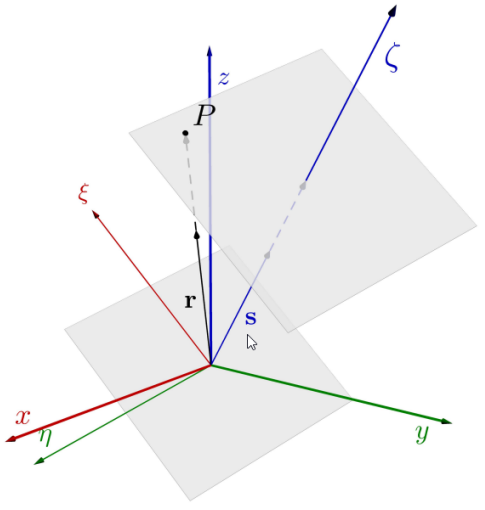

Figure 2. Plane $(\mathbf{r}, \mathbf{s}) = \text{const}$, where $\mathbf{s}$ is a unit vector indicating the direction of propagation of the electromagnetic wave. New coordinate axes are chosen so that the vector $\mathbf{s}$ is the unit vector of the $O\zeta$ axis. The other two axes $O\xi$ and $O\eta$ are chosen arbitrarily and form a right-handed coordinate system $O\xi\eta\zeta$

The Laplace operator after replacing coordinates is transformed to the following form:

$$\nabla^2 \mathbf{U} = s_x^2 \frac{\partial^2 \mathbf{U}}{\partial \zeta^2} + s_y^2 \frac{\partial^2 \mathbf{U}}{\partial \zeta^2} + s_z^2 \frac{\partial^2 \mathbf{U}}{\partial \zeta^2} = \left(s_x^2 + s_y^2 + s_z^2\right) \frac{\partial^2 \mathbf{U}}{\partial \zeta^2} = \frac{\partial^2 \mathbf{U}}{\partial \zeta^2}.$$

The wave equation simplifies:

$$\frac{\partial^2 \mathbf{U}}{\partial \zeta^2} - \frac{1}{v^2} \frac{\partial^2 \mathbf{U}}{\partial t^2} = 0.$$

We perform another substitution $p = \zeta - vt$ and $q = \zeta + vt$, which leads to the following transformation of the differential operators:

$$\begin{bmatrix} \dfrac{\partial}{\partial\zeta} \\ \dfrac{\partial}{\partial t} \end{bmatrix} = \left(\frac{\partial(\zeta,t)}{\partial(p,q)}\right)^T \begin{bmatrix} \dfrac{\partial}{\partial p} \\ \dfrac{\partial}{\partial q} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial p}{\partial\zeta} & \dfrac{\partial q}{\partial\zeta} \\ \dfrac{\partial p}{\partial t} & \dfrac{\partial q}{\partial t} \end{bmatrix} \begin{bmatrix} \dfrac{\partial}{\partial p} \\ \dfrac{\partial}{\partial q} \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 1 \\ -v & v \end{bmatrix} \begin{bmatrix} \dfrac{\partial}{\partial p} \\ \dfrac{\partial}{\partial q} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial}{\partial p} + \dfrac{\partial}{\partial q} \\ -v\dfrac{\partial}{\partial p} + v\dfrac{\partial}{\partial q} \end{bmatrix} \Rightarrow \begin{cases} \dfrac{\partial}{\partial\zeta} = \dfrac{\partial}{\partial p} + \dfrac{\partial}{\partial q}, \\ \dfrac{\partial}{\partial t} = -v\dfrac{\partial}{\partial p} + v\dfrac{\partial}{\partial q}. \end{cases}$$

The second derivatives are expressed through the new variables as follows:

$$\frac{\partial^2}{\partial\zeta^2} = \frac{\partial^2}{\partial p^2} + 2\frac{\partial}{\partial p}\frac{\partial}{\partial q} + \frac{\partial^2}{\partial q^2},$$
$$\frac{\partial^2}{\partial t^2} = v^2\left(\frac{\partial^2}{\partial p^2} - 2\frac{\partial}{\partial p}\frac{\partial}{\partial q} + \frac{\partial^2}{\partial q^2}\right).$$

When the operators are substituted into the wave equation, it is simplified as follows:

$$\frac{\partial^2\mathbf{U}}{\partial\zeta^2} - \frac{1}{v^2}\frac{\partial^2\mathbf{U}}{\partial t^2} =$$
$$= \frac{\partial^2\mathbf{U}}{\partial p^2} + 2\frac{\partial^2\mathbf{U}}{\partial p\partial q} + \frac{\partial^2\mathbf{U}}{\partial q^2} - \frac{1}{v^2}v^2\left(\frac{\partial^2\mathbf{U}}{\partial p^2} - 2\frac{\partial^2\mathbf{U}}{\partial p\partial q} + \frac{\partial^2\mathbf{U}}{\partial q^2}\right) =$$
$$= 4\frac{\partial^2\mathbf{U}}{\partial p\partial q} = 0 \Rightarrow \boxed{\frac{\partial^2\mathbf{U}}{\partial p\partial q} = 0}.$$

The general solution of the transformed wave equation is the function

$$\mathbf{U} = \mathbf{U}_1(p) + \mathbf{U}_2(q) = \mathbf{U}_1\left((\mathbf{r},\mathbf{s}) - vt\right) + \mathbf{U}_2\left((\mathbf{r},\mathbf{s}) + vt\right).$$

Another approach to the solution uses separation of variables. We will look for the solution in complex form

$$\mathbf{U}(\mathbf{r},t) = \mathbf{U}_0(\mathbf{r})e^{-i\omega t}.$$

When substituting into the wave equation, we obtain:

$$\frac{\partial^2\mathbf{U}}{\partial t^2} = -\omega^2 e^{-i\omega t}\mathbf{U}_0(\mathbf{r}), \quad \nabla^2\mathbf{U} = e^{-i\omega t}\nabla^2\mathbf{U}_0(\mathbf{r}),$$

$$\nabla^2\mathbf{U} - \frac{1}{v^2}\frac{\partial^2\mathbf{U}}{\partial t^2} = 0 \Longrightarrow \nabla^2\mathbf{U}_0 + \frac{\omega^2}{v^2}\mathbf{U}_0 = 0.$$

Let's introduce some scalar quantities: *wave number* $k = \omega/v$, $k_0 = \omega/c$, *wave vector* $\mathbf{k} = k\mathbf{s}$. Let us recall that c — the speed of light in a vacuum, v — the speed of an electromagnetic wave in a medium, $n = \sqrt{\varepsilon\mu}$ — the refractive index of the medium, $\mathbf{s}$ — direction of wave propagation. The velocities v

and c are related by the relations $v = c/\sqrt{\varepsilon\mu} = c/n$, so the wave number can also be written as $k = \omega/v = \omega\sqrt{\varepsilon\mu}/c = k_0 n$. Now the equation for $\mathbf{U}_0$ can be rewritten as:

$$\left(\nabla^2 + k^2\right)\mathbf{U}_0 = 0.$$

This equation is called *Helmholtz equation* (homogeneous Helmholtz equation). In the general case, its solution can be expressed in special functions, but in the case of a plane wave, the general solution can be written in the following form:

$$\mathbf{U}_0(\mathbf{r}) = \mathbf{u}_0(\mathbf{r})e^{ik(\mathbf{s},\mathbf{r})} = \mathbf{u}_0(\mathbf{r})e^{ik_0 n(\mathbf{s},\mathbf{r})}.$$

4. Derivation of the eikonal equation

We will also consider a strictly monochromatic harmonic wave, the intensity vectors of which can be written in the following form:

$$\mathbf{E}(\mathbf{r},t) = \mathbf{E}_0(\mathbf{r})e^{-i\omega t},$$
$$\mathbf{H}(\mathbf{r},t) = \mathbf{H}_0(\mathbf{r})e^{-i\omega t},$$

where $\mathbf{r} = (x,y,z)^T$ — radius vector of a point in space in a Cartesian coordinate system, ω — cyclic frequency. We also introduce the quantity $k_0 = \omega/c = 2\pi/\lambda_0$, where λ_0 is the wavelength in vacuum.

Let's substitute expressions for a monochromatic wave into Maxwell's equations. We sequentially calculate all differential operators:

$$\boldsymbol{\nabla}\times\mathbf{H} = \boldsymbol{\nabla}\times\left(\mathbf{H}_0 e^{-i\omega t}\right) = e^{-i\omega t}\boldsymbol{\nabla}\times\mathbf{H}_0,$$
$$\boldsymbol{\nabla}\times\mathbf{E} = \boldsymbol{\nabla}\times\left(\mathbf{E}_0 e^{-i\omega t}\right) = e^{-i\omega t}\boldsymbol{\nabla}\times\mathbf{E}_0.$$

Using the material equations $\mathbf{D} = \varepsilon\mathbf{E}$ and $\mathbf{B} = \mu\mathbf{H}$ we replace $\mathbf{D}$ and $\mathbf{B}$ everywhere through $\mathbf{E}$ and $\mathbf{H}$, taking into account that $\varepsilon(\mathbf{r}) = \varepsilon(x,y,z)$ and $\mu(\mathbf{r}) = \mu(x,y,z)$:

$$\boldsymbol{\nabla}\cdot\mathbf{D} = \boldsymbol{\nabla}\cdot(\varepsilon(x,y,z)\mathbf{E}) = e^{-i\omega t}\boldsymbol{\nabla}\cdot(\varepsilon\mathbf{E}_0),$$
$$\boldsymbol{\nabla}\cdot\mathbf{B} = \boldsymbol{\nabla}\cdot(\mu(x,y,z)\mathbf{H}) = e^{-i\omega t}\boldsymbol{\nabla}\cdot(\mu\mathbf{H}_0)\,.$$

Let us replace $\mathbf{D}$ and $\mathbf{B}$ also in the expressions for derivatives, taking into account that ε and μ do not depend on time, as well as $\mathbf{E}_0$ with $\mathbf{H}_0$ from the formulas $\mathbf{E}(x,y,z,t) = \mathbf{E}_0(x,y,z)e^{-i\omega t}$, $\mathbf{H}(x,y,z,t) = \mathbf{H}_0(x,y,z)e^{-i\omega t}$:

$$\frac{\partial\mathbf{D}}{\partial t} = \frac{\partial}{\partial t}\left(\varepsilon\mathbf{E}_0 e^{-i\omega t}\right) = \varepsilon(x,y,z)\mathbf{E}_0(x,y,z)\frac{\partial e^{-i\omega t}}{\partial t} = -i\varepsilon\omega\mathbf{E}_0 e^{-i\omega t},$$
$$\frac{\partial\mathbf{B}}{\partial t} = \frac{\partial}{\partial t}\left(\mu\mathbf{H}_0 e^{-i\omega t}\right) = \mu(x,y,z)\mathbf{H}_0(x,y,z)\frac{\partial e^{-i\omega t}}{\partial t} = -i\mu\omega\mathbf{H}_0 e^{-i\omega t}.$$

Let's substitute the resulting expressions into the equation (5):

$$\boldsymbol{\nabla}\times\mathbf{H} - \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t} = 0 \Rightarrow \cancel{e^{-i\omega t}}\boldsymbol{\nabla}\times\mathbf{H}_0 + i\varepsilon\frac{\omega}{c}\mathbf{E}_0\cancel{e^{-i\omega t}} = 0 \Rightarrow \boxed{\boldsymbol{\nabla}\times\mathbf{H}_0 + i\varepsilon k_0\mathbf{E}_0 = 0},$$

then into the equation (6):

$$\boldsymbol{\nabla}\times\mathbf{E}+\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t}=0\Rightarrow \cancel{e^{-i\omega t}}\boldsymbol{\nabla}\times\mathbf{E}_0-i\varepsilon\frac{\omega}{c}\mathbf{H}_0\cancel{e^{-i\omega t}}=0\Rightarrow \boxed{\boldsymbol{\nabla}\times\mathbf{E}_0-i\varepsilon k_0\mathbf{H}_0=0},$$

into the equation (7):

$$\boldsymbol{\nabla}\cdot\mathbf{D}=0\Rightarrow e^{-i\omega t}\boldsymbol{\nabla}\cdot(\varepsilon\mathbf{E}_0)=0\Rightarrow \boxed{\boldsymbol{\nabla}\cdot(\varepsilon\mathbf{E}_0)=0},$$

and finally into the equation (8):

$$\boldsymbol{\nabla}\cdot\mathbf{B}=0\Rightarrow e^{-i\omega t}\boldsymbol{\nabla}\cdot(\mu\mathbf{H}_0)=0\Rightarrow \boxed{\boldsymbol{\nabla}\cdot(\mu\mathbf{H}_0)=0}.$$

As a result, the system of equations (5)–(8) takes the following simplified form:

$$\begin{cases}\boldsymbol{\nabla}\times\mathbf{H}_0+i\varepsilon k_0\mathbf{E}_0=0,\\ \boldsymbol{\nabla}\times\mathbf{E}_0-i\mu k_0\mathbf{H}_0=0,\\ \boldsymbol{\nabla}\cdot(\varepsilon\mathbf{E}_0)=0,\\ \boldsymbol{\nabla}\cdot(\mu\mathbf{H}_0)=0.\end{cases}\tag{9}$$

Let us make another simplification by assuming that

$$\mathbf{E}_0(x,y,z)=\mathbf{e}(x,y,z)\exp\left(ik_0u(x,y,z)\right)=\mathbf{e}(\mathbf{r})\exp\left(ik_0u(\mathbf{r})\right),$$
$$\mathbf{H}_0(x,y,z)=\mathbf{h}(x,y,z)\exp\left(ik_0u(x,y,z)\right)=\mathbf{h}(\mathbf{r})\exp\left(ik_0u(\mathbf{r})\right),$$

where $u(x,y,z)=u(\mathbf{r})$ is a scalar real function called *optical path*, and $\mathbf{e}$ and $\mathbf{h}$ — vector position functions. Let's calculate the differential operators again, this time from $\mathbf{E}_0$ and $\mathbf{H}_0$, using the formulas (18):

$$\boldsymbol{\nabla}\times\mathbf{H}_0=\boldsymbol{\nabla}\times\left(e^{ik_0u(\mathbf{r})}\mathbf{h}(\mathbf{r})\right)=e^{ik_0u(\mathbf{r})}\boldsymbol{\nabla}\times\mathbf{h}+\boldsymbol{\nabla}\left(e^{ik_0u(\mathbf{r})}\right)\times\mathbf{h}.$$

The gradient of the function $e^{ik_0u(\mathbf{r})}$ is calculated as follows:

$$\begin{aligned}\boldsymbol{\nabla}\left(e^{ik_0u(\mathbf{r})}\right)&=\left(\frac{\partial e^{ik_0u(\mathbf{r})}}{\partial x},\frac{\partial e^{ik_0u(\mathbf{r})}}{\partial y},\frac{\partial e^{ik_0u(\mathbf{r})}}{\partial z}\right)=\\ &=ik_0e^{ik_0u(\mathbf{r})}\left(\frac{\partial u(x,y,z)}{\partial x},\frac{\partial u(x,y,z)}{\partial y},\frac{\partial u(x,y,z)}{\partial z}\right)=\\ &=ik_0e^{ik_0u(\mathbf{r})}\boldsymbol{\nabla}u(x,y,z).\end{aligned}$$

As a result, the term $\boldsymbol{\nabla}\times\mathbf{H}_0$ of the first equation of the system (9) takes the form:

$$\boldsymbol{\nabla}\times\mathbf{H}_0=(\boldsymbol{\nabla}\times\mathbf{h}+ik_0\boldsymbol{\nabla}u\times\mathbf{h})e^{ik_0u(\mathbf{r})}.\tag{10}$$

In a completely similar way, we obtain the expression for $\boldsymbol{\nabla}\times\mathbf{E}_0$ in the second equation of the system (9):

$$\boldsymbol{\nabla}\times\mathbf{E}_0=(\boldsymbol{\nabla}\times\mathbf{e}+ik_0\boldsymbol{\nabla}u\times\mathbf{e})e^{ik_0u(\mathbf{r})}.\tag{11}$$

The computation of divergence is somewhat more complicated because the formula (18) will have to be applied twice. The first time we use it to write down the expression $\boldsymbol{\nabla} \cdot \varepsilon \mathbf{E}_0$:

$$\boldsymbol{\nabla} \cdot \varepsilon \mathbf{E}_0 = \varepsilon(\mathbf{r}) \boldsymbol{\nabla} \cdot \mathbf{E}_0 + (\boldsymbol{\nabla}\varepsilon, \mathbf{E}_0)\,.$$

Next, we use it to calculate $\boldsymbol{\nabla} \cdot \mathbf{E}_0$, where instead of $\mathbf{E}_0$ we substitute the expression $\mathbf{E}_0 = \mathbf{e}(\mathbf{r}) \exp(ik_0 u(\mathbf{r}))$:

$$\begin{aligned}\boldsymbol{\nabla} \cdot \mathbf{E}_0 = \boldsymbol{\nabla} \cdot \left[\mathbf{e}(\mathbf{r}) e^{ik_0 u(\mathbf{r})}\right] = e^{ik_0 u(\mathbf{r})} \boldsymbol{\nabla} \cdot \mathbf{e} + \left(\boldsymbol{\nabla}\left(e^{ik_0 u(\mathbf{r})}\right), \mathbf{e}\right) = \\ = e^{ik_0 u(\mathbf{r})} \boldsymbol{\nabla} \cdot \mathbf{e} + ik_0 e^{ik_0 u(\mathbf{r})} (\boldsymbol{\nabla} u, \mathbf{e}) = (\boldsymbol{\nabla} \cdot \mathbf{e} + ik_0 (\boldsymbol{\nabla} u, \mathbf{e})) e^{ik_0 u(\mathbf{r})},\end{aligned}$$

$$(\boldsymbol{\nabla}\varepsilon, \mathbf{E}_0) = (\boldsymbol{\nabla}\varepsilon, \mathbf{e}) e^{ik_0 u(\mathbf{r})}.$$

As a result, the third equation of the system (9) takes the form:

$$\boldsymbol{\nabla} \cdot (\varepsilon(\mathbf{r}) \mathbf{E}_0(\mathbf{r})) = [\varepsilon(\mathbf{r}) \boldsymbol{\nabla} \cdot \mathbf{e}(\mathbf{r}) + ik_0 \varepsilon(\mathbf{r}) (\boldsymbol{\nabla} u(\mathbf{r}), \mathbf{e}(\mathbf{r})) + (\boldsymbol{\nabla}\varepsilon(\mathbf{r}), \mathbf{e}(\mathbf{r}))] e^{ik_0 u(\mathbf{r})}.$$

In a completely similar way, we obtain an expression for the magnetic field strength, that is, the fourth equation of the system (9):

$$\boldsymbol{\nabla} \cdot (\mu(\mathbf{r}) \mathbf{H}_0(\mathbf{r})) = [\mu(\mathbf{r}) \boldsymbol{\nabla} \cdot \mathbf{h} + (\boldsymbol{\nabla}\mu(\mathbf{r}), \mathbf{h}) + ik_0 \mu(\mathbf{r}) (\boldsymbol{\nabla} u(\mathbf{r}), \mathbf{h})] e^{ik_0 u(\mathbf{r})}.$$

After substitution into Maxwell's equations, we obtain:

$$\boldsymbol{\nabla} \times \mathbf{H}_0 + i\varepsilon k_0 \mathbf{E}_0 = \mathbf{0} \Rightarrow \underbrace{\boldsymbol{\nabla} \times \mathbf{h} + ik_0 \boldsymbol{\nabla} u \times \mathbf{h}}_{(10)} + i\varepsilon k_0 \mathbf{e} = \mathbf{0} \Rightarrow$$

$$\Rightarrow \boldsymbol{\nabla} u \times \mathbf{h} + \varepsilon \mathbf{e} = -\frac{1}{ik_0} \boldsymbol{\nabla} \times \mathbf{h},$$

$$\boldsymbol{\nabla} \times \mathbf{E}_0 - i\mu k_0 \mathbf{H}_0 = \mathbf{0} \Rightarrow \underbrace{\boldsymbol{\nabla} \times \mathbf{e} + ik_0 \boldsymbol{\nabla} u \times \mathbf{e}}_{(11)} - i\mu k_0 \mathbf{h} = \mathbf{0} \Rightarrow$$

$$\Rightarrow \boldsymbol{\nabla} u \times \mathbf{e} - \mu \mathbf{h} = -\frac{1}{ik_0} \boldsymbol{\nabla} \times \mathbf{e},$$

$$\boldsymbol{\nabla} \cdot (\varepsilon \mathbf{E}_0) = 0 \Rightarrow \varepsilon \boldsymbol{\nabla} \cdot \mathbf{e} + ik_0 \varepsilon (\boldsymbol{\nabla} u, \mathbf{e}) + (\boldsymbol{\nabla}\varepsilon, \quad \mathbf{e}) = 0,$$

$$ik_0 \varepsilon (\boldsymbol{\nabla} u, \mathbf{e}) = -(\boldsymbol{\nabla}\varepsilon, \mathbf{e}) - \varepsilon \boldsymbol{\nabla} \cdot \mathbf{e} = 0 \Rightarrow (\boldsymbol{\nabla} u, \mathbf{e}) = -\frac{1}{ik_0} \left[\left(\frac{1}{\varepsilon} \boldsymbol{\nabla}\varepsilon, \mathbf{e}\right) + \boldsymbol{\nabla} \cdot \mathbf{e}\right].$$

Since

$$\boldsymbol{\nabla}(\ln \varepsilon) = \left(\frac{\partial \ln \varepsilon}{\partial x}, \frac{\partial \ln \varepsilon}{\partial y}, \frac{\partial \ln \varepsilon}{\partial z}\right) = \frac{1}{\varepsilon}\left(\frac{\partial \varepsilon}{\partial x}, \frac{\partial \varepsilon}{\partial y}, \frac{\partial \varepsilon}{\partial z}\right) = \frac{1}{\varepsilon} \boldsymbol{\nabla}\varepsilon,$$

$$(\boldsymbol{\nabla} u, \mathbf{e}) = -\frac{1}{ik_0} ((\boldsymbol{\nabla}(\ln \varepsilon), \mathbf{e}) + \boldsymbol{\nabla} \cdot \mathbf{e}).$$

Calculations for the magnetic field are carried out in a completely similar way, resulting in the fourth equation:

$$(\boldsymbol{\nabla} u, \mathbf{h}) = -\frac{1}{ik_0}((\boldsymbol{\nabla}(\ln\mu), \mathbf{h}) + \boldsymbol{\nabla}\cdot\mathbf{h}),$$

$$\begin{cases} \boldsymbol{\nabla} u \times \mathbf{h} + \varepsilon\mathbf{e} = -\dfrac{1}{ik_0}\boldsymbol{\nabla}\times\mathbf{h}, \\ \boldsymbol{\nabla} u \times \mathbf{e} - \mu\mathbf{h} = -\dfrac{1}{ik_0}\boldsymbol{\nabla}\times\mathbf{e}, \\ (\boldsymbol{\nabla} u, \mathbf{e}) = -\dfrac{1}{ik_0}\left((\boldsymbol{\nabla}\left(\ln\varepsilon\right), \mathbf{e}) + \boldsymbol{\nabla}\cdot\mathbf{e}\right), \\ (\boldsymbol{\nabla} u, \mathbf{h}) = -\dfrac{1}{ik_0}\left((\boldsymbol{\nabla}\left(\ln\mu\right), \mathbf{h}) + \boldsymbol{\nabla}\cdot\mathbf{h}\right). \end{cases} \tag{12}$$

The third and fourth equations from this system follow from the first two. This can be proven by scalarly multiplying the first two equations by $\boldsymbol{\nabla} u$ and using the fact that the result of a vector product is orthogonal to both of its factors:

$$\underbrace{(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u \times \mathbf{h})}_{=0} + \varepsilon(\boldsymbol{\nabla} u, \mathbf{e}) = 0 \Rightarrow (\boldsymbol{\nabla} u, \mathbf{e}) = 0.$$

We consider only the first two equations. Let's express $\mathbf{h}$ from the second equation through u and $\mathbf{e}$ and substitute it into the first:

$$\mathbf{h} = \frac{1}{\mu}\boldsymbol{\nabla} u \times \mathbf{e} \Rightarrow \boldsymbol{\nabla} u \times \left(\frac{1}{\mu}\boldsymbol{\nabla} u \times \mathbf{e}\right) + \varepsilon\mathbf{e} = \mathbf{0} \Rightarrow \boldsymbol{\nabla} u \times \boldsymbol{\nabla} u \times \mathbf{e} + \varepsilon\mu\mathbf{e} = 0.$$

For the vector product the following identity holds: $\mathbf{a}\times\mathbf{b}\times\mathbf{c} = \mathbf{b}(\mathbf{a},\mathbf{c}) - \mathbf{c}(\mathbf{a},\mathbf{b})$ from which it follows

$$\boldsymbol{\nabla} u \times \boldsymbol{\nabla} u \times \mathbf{e} = \boldsymbol{\nabla} u(\boldsymbol{\nabla} u, \mathbf{e}) - \mathbf{e}(\boldsymbol{\nabla} u, \boldsymbol{\nabla} u) = \boldsymbol{\nabla} u(\boldsymbol{\nabla} u, \mathbf{e}) - \mathbf{e}\|\boldsymbol{\nabla} u\|^2,$$

$$\boldsymbol{\nabla} u(\boldsymbol{\nabla} u, \mathbf{e}) - \mathbf{e}\|\boldsymbol{\nabla} u\|^2 + \varepsilon\mu\mathbf{e} = 0.$$

From the third equation of the system (12) it follows that $(\boldsymbol{\nabla} u, \mathbf{e})$, therefore

$$-\mathbf{e}\|\boldsymbol{\nabla} u\|^2 + \varepsilon\mu\mathbf{e} = 0 \Rightarrow \mathbf{e}\|\boldsymbol{\nabla} u\|^2 = \varepsilon\mu\mathbf{e}.$$

Equating the coefficients in front of the vector $\mathbf{e}$ and taking into account that $n(\mathbf{r}) = \sqrt{\varepsilon(\mathbf{r})\mu(\mathbf{r})}$ we write the equation:

$$\|\boldsymbol{\nabla} u\|^2 = n^2(\mathbf{r}), \tag{13}$$

which is the *eikonal equation*. The function $u(\mathbf{r}) = u(x, y, z)$ is also called *eikonal*, and the surfaces $u(x, y, z) = \text{const}$ — *geometric wave fronts*.

In component form in Cartesian coordinates equation (13) becomes:

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 = \varepsilon(x,y,z)\mu(x,y,z) = n^2(x,y,z),$$

$$\|\boldsymbol{\nabla} u\|^2 = (\boldsymbol{\nabla} u, \boldsymbol{\nabla} u) = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2.$$

5. Derivation of eikonal in covariant form

Let us demonstrate the derivation of the eikonal equation using the tensor formalism.

5.1. Vector operators in covariant form

Vector operators in covariant form:

— $\nabla_{\vec{V}}$ is covariant derivative with respect to the vector field $\vec{v}$;

— $\vec{e}_i = \dfrac{\partial}{\partial x^i}$ is coordinate basis, $\nabla\, \vec{e}_i = \nabla_j$;

— $\varepsilon_{ijk} = \varepsilon^{ijk}$ is Levi–Civita symbol;

— $e_{ijk} = \sqrt{|g|}\varepsilon_{ijk}$, $e^{ijk} = \dfrac{1}{\sqrt{|g|}}\varepsilon^{ijk}$ are alternating tensors (Levi–Civita tensors);

— $\boldsymbol{\nabla} f = \nabla_i f = \partial_i f$, f is scalar field;

— $\boldsymbol{\nabla} \cdot f = \nabla_i V^i = \dfrac{1}{\sqrt{g}}\partial_i(\sqrt{g}V^i)$;

— $\vec{x} = (x^1, x^2, x^3)^T$ is contravariant vector;

— $\boldsymbol{\nabla} \times \vec{V} = e^{ijk}\nabla_j V_k = e^{ijk}\partial_j V_k = \dfrac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j V_k$.

5.2. Maxwell's equations without currents and charges

The strength of the electric and magnetic fields in the form of a covector (denoted by "$\sim$" above the letter, the designations can be changed), and D and B are vectors:

$$\tilde{E} = (E_1, E_2, E_3), \quad \tilde{H} = (H_1, H_2, H_3), \quad \vec{D} = (D^1, D^2, D^3)^T, \quad \vec{B} = (B^1, B^2, B^3)^T.$$

Material equations: $B^i = \mu^{ij}H_j$, $D^i = \varepsilon^{ij}E_j$.

Vector, covector fields: $\tilde{E}(\vec{x}, t)$, $\tilde{H}(\vec{x}, t)$, $\vec{D}(\tilde{x}, t)$, $\vec{B}(\tilde{x}, t)$.

Tensor fields: $\mu^{ij}(\vec{x})$, $\varepsilon^{ij}(\vec{x})$.

5.3. Vector-differential form of writing Maxwell's equations

$$\begin{cases} \boldsymbol{\nabla} \times \vec{H} - \dfrac{1}{c}\dfrac{\partial \vec{D}}{\partial t} = 0, \\ \boldsymbol{\nabla} \times \vec{E} + \dfrac{1}{c}\dfrac{\partial \vec{B}}{\partial t} = 0, \\ \boldsymbol{\nabla} \cdot \vec{D} = 0, \\ \boldsymbol{\nabla} \cdot \vec{B} = 0; \end{cases}$$

$$\begin{cases} \dfrac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j E_k + \dfrac{1}{c}\dfrac{\mathrm{d}B^i}{\mathrm{d}t} = 0, \\ \dfrac{1}{\sqrt{g}}\varepsilon^{ijk}\mathrm{d}_j H_k - \dfrac{1}{c}\dfrac{\mathrm{d}D^i}{\mathrm{d}t} = 0, \\ \dfrac{1}{\sqrt{g}}\partial_i(\sqrt{g}D^i) = 0, \\ \dfrac{1}{\sqrt{g}}\partial_i(\sqrt{g}B^i) = 0. \end{cases}$$

5.4. Monochromatic harmonic wave

Assumption No. 1: Monochromatic harmonic wave:

$$\begin{cases} E_k = E_{0k}e^{-i\omega t},\ H_k = H_{0k}e^{-i\omega t},\ D^k = \varepsilon^{kl}E_l = \varepsilon^{kl}E_{0l}e^{-i\omega t}, \\ B^k = \mu^{kl}H_l = \mu^{kl}H_{0k}e^{-i\omega t}, \\ \dfrac{\mathrm{d}D^i}{\mathrm{d}t} = \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\varepsilon^{ij}E_{0j}e^{-i\omega t}\right) = -i\omega\varepsilon^{ij}E_{0j}, \\ \dfrac{\mathrm{d}B^i}{\mathrm{d}t} = \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\mu^{ij}H_{0j}e^{-i\omega t}\right) = -i\omega\mu^{ij}H_{0j}, \\ \partial_i(\sqrt{g}D^i) = \partial_i(\sqrt{g}\varepsilon^{ij}E_{0j}e^{-i\omega t}) = e^{-i\omega t}\partial_i(\sqrt{g}\varepsilon^{ij}E_{0j}), \\ \partial_i(\sqrt{g}B^i) = \partial_i(\sqrt{g}\mu^{ij}H_{0j}e^{-i\omega t}) = e^{-i\omega t}\partial_i(\sqrt{g}\mu^{ij}H_{0j}). \end{cases}$$

Formulas:

$$\frac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j E_{0k} - ik_0\mu^{ij}H_{0j} = 0, \tag{14}$$

$$\frac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j H_{0k} + ik_0\varepsilon^{ij}E_{0j} = 0, \tag{15}$$

$$\partial_i(\sqrt{g}\varepsilon^{ij}E_{0j}) = 0, \tag{16}$$

$$\partial_i(\sqrt{g}\mu^{ij}H_{0j}) = 0. \tag{17}$$

Assumption No. 2: $E_{0k} = e_k e^{ik_0 u(\vec{x})}$, $H_{0k} = h_k e^{ik_0 u(\vec{x})}$, where $u(\vec{x})$ is an eikonal:

$$\partial_j E_{0k} = (\partial_j e_k) e^{ik_0 u} + e_k e^{ik_0 u} i k_0 \partial_j u = (\partial_j e_k + i k_0 e_k \partial_j u) e^{ik_0 u},$$
$$\partial_j H_{0k} = (\partial_j h_k) e^{ik_0 u} + h_k e^{ik_0 u} i k_0 \partial_j u = (\partial_j h_k + i k_0 h_k \partial_j u) e^{ik_0 u}.$$

From the equation (16):

$$\partial_i(\sqrt{g}\varepsilon^{ij} e_j e^{ik_0 u}) =$$
$$= \frac{\partial\sqrt{g}}{\partial x^i}\varepsilon^{ij} e_j e^{ik_0 u} + \sqrt{g}\frac{\partial\varepsilon^{ij}}{\partial x^i} e_j e^{ik_0 u} + \sqrt{g}\varepsilon^{ij}\frac{\partial e_j}{\partial x^i} e^{ik_0 u} + \sqrt{g}\varepsilon^{ij} e_j i k_0 e^{ik_0 u}\frac{\partial u}{\partial x^i} =$$
$$= (\partial_i\sqrt{g}\varepsilon^{ij} e_j + \sqrt{g}\partial_i\varepsilon^{ij} e_j + \sqrt{g}\varepsilon^{ij}\partial_i e_j + i k_0\sqrt{g}\varepsilon^{ij} e_j \partial_i u) e^{ik_0 u} = 0,$$

$$\partial_i\sqrt{g}\varepsilon^{ij} e_j + \sqrt{g}\partial_i\varepsilon^{ij} e_j + \sqrt{g}\varepsilon^{ij}\partial_i e_j + i k_0\sqrt{g}\varepsilon^{ij} e_j \partial_i u = 0,$$
$$\sqrt{g}\varepsilon^{ij} e_j \partial_i u = \frac{-1}{ik_0}(\partial_i\sqrt{g}\varepsilon^{ij} e_j + \sqrt{g}\partial_i\varepsilon^{ij} e_j + \sqrt{g}\varepsilon^{ij}\partial_i e_j).$$

Similarly from the equation (17):

$$\sqrt{g}\mu^{ij} h_j \partial_i u = -\frac{1}{ik_0}\left(\partial_i\sqrt{g}\mu^{ij} h_j + \sqrt{g}\partial_i\mu^{ij} + \sqrt{g}\mu^{ij}\partial_i h_j\right).$$

Provided that λ is small, ω is large, $\Rightarrow k_0$ is large, and $\dfrac{1}{k_0}$ is small $\Rightarrow$ we obtain:

$$\begin{cases} \sqrt{g}\varepsilon^{ij} e_j \partial_i u = 0, \\ \sqrt{g}\mu^{ij} h_j \partial_i u = 0; \end{cases} \quad \begin{cases} \varepsilon^{ij} e_j \partial_i u = 0, \\ \mu^{ij} h_j \partial_i u = 0 \end{cases}$$

to transform:

$$\frac{1}{\sqrt{g}}\varepsilon^{ijk}(\partial_j e_k + i k_0 e_k \partial_j u) - i k_0 \mu^{ij} h_j = 0 \Rightarrow -\frac{1}{\sqrt{g}}\varepsilon^{ijk} e_k \partial_j u + \mu^{ij} h_0 = \frac{1}{ik_0}\frac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j e_k,$$

$$\frac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j h_k + \frac{1}{\sqrt{g}}\varepsilon^{ijk} i k_0 h_k \partial_j u + i k_0 \varepsilon^{ij} e_j = 0,$$
$$\frac{1}{\sqrt{g}}\varepsilon^{ijk} h_k \partial_j u + \varepsilon^{ij} e_j = -\frac{1}{ik_0}\frac{1}{\sqrt{g}}\varepsilon^{ijk}\partial_j h_k.$$

Maxwell's equations are reduced to the following form:

$$\begin{cases} \varepsilon^{ijk} e_k \partial_j u - \sqrt{g}\mu^{ij} h_j = 0, \\ \varepsilon^{ijk} h_k \partial_j u + \sqrt{g}\varepsilon^{ij} e_j = 0, \\ \varepsilon^{ij} e_j \partial_i u = 0, \\ \mu^{ij} h_j \partial_i u = 0, \end{cases}$$

where ε^{ijk} is Levi-Civita symbol, ε^{ij} is permittivity, subject to $k_0 \to \infty$.

From the first equation we express h_j and substitute it into the second:

$$\mu_{li}^{-1}\varepsilon^{ijk}e_k\partial_j u - \sqrt{g}\mu_{li}^{-1}\mu^{ij}h_j = 0.$$

Let's make the replacement: $\mu_{li}^{-1}\mu^{ij} = g_l^j$:

$$\mu_{li}^{-1}\varepsilon^{ijk}e_k\partial_j u - \sqrt{g}g_l^j h_j = 0 \Rightarrow \sqrt{g}h_l = \mu_{li}^{-1}\varepsilon^{ijk}e_k\partial_j u \Rightarrow h_l = \frac{1}{\sqrt{g}}\mu_{li}^{-1}\varepsilon^{ijk}e_k\partial_j u.$$

We transform the indices to substitute into the second equation:

$$h_k = \frac{1}{\sqrt{g}}\mu_{kl}^{-1}\varepsilon^{lmn}e_n\partial_m u,$$

$$\varepsilon^{ijk}\frac{1}{\sqrt{g}}\mu_{kl}^{-1}\varepsilon^{lmn}e_n\partial_m u\partial_j u + \sqrt{g}\varepsilon^{ij}e_j = 0,$$

$$\varepsilon^{ijk}\mu_{kl}^{-1}\varepsilon^{lmn}e_n\partial_m u\partial_j u + g\varepsilon^{ij}e_j = 0,$$

$$\varepsilon^{ijk}\mu_{kl}^{-1}\varepsilon^{lmn}\partial_m u\partial_j u e_n + g\varepsilon^{in}e_n = 0.$$

The eikonal equation (13) takes the form:

$$g^{ij}\partial_i u\partial_j u = \varepsilon^{ij}\mu_{ij}.$$

6. Conclusion

We hope that our work clarifies the process of derivation of the eikonal equation. And allows us to better understand the hierarchy of models in electrodynamics in general, and in optics in particular. The questions of solving the eikonal equation [7, 8] we left outside the boundaries of our consideration.

Appendix. Vector analysis

If at each point P of a certain spatial region of the Euclidean space $\mathbb{R}^n$ some scalar or vector quantity is associated, then they say that a *field* (*scalar* or *vector*).

— Examples of vector fields include the velocity field $\mathbf{v}(x, y, z)$, the force field $\mathbf{F}(x, y, z)$, the electrical intensity field $\mathbf{E}(x, y, z)$.

— Examples of scalar fields: temperature field $T(x, y, z)$, electric potential field $\varphi(x, y, z)$.

Everywhere below we consider a three-dimensional point Euclidean space on which a Cartesian coordinate system is introduced. We denote the vectors (basis vectors) of this coordinate system as $\langle \mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z \rangle$. The coordinates of a point are specified by the radius vector $\mathbf{r} = (x, y, z)^T$, which is plotted from the origin O. Along with notation of coordinates x, y, z, it is sometimes convenient to use indices: x^1, x^2, x^3, and also write the radius vector in the form $\mathbf{x} = (x^1, x^2, x^3)^T$. Index notation makes it possible to briefly write

formulas using the summation sign Σ, which is especially convenient if a non-unit metric is used.

A scalar field in some region of space $\mathbb{R}^3$ is a real-valued function f:

$$f\colon \mathbb{R}^3 \to \mathbb{R}, \quad f(x,y,z) = f(\mathbf{r}) \in \mathbb{R}.$$

In turn, a vector field in a region of space $\mathbb{R}^3$ is a vector-valued function $\mathbf{V}$:

$$\mathbf{V}\colon \mathbb{R}^3 \to \mathbb{R}^3, \quad \mathbf{V}(x,y,z) = \mathbf{V}(\mathbf{r}) = V_x(\mathbf{r})\mathbf{e}_x + V_y(\mathbf{r})\mathbf{e}_y + V_z(\mathbf{r})\mathbf{e}_z \in \mathbb{R}^3.$$

The *Gradient* of the scalar field $f(\mathbf{r})$ is a vector calculated in Cartesian coordinates as follows:

$$\boldsymbol{\nabla} f(x,y,z) = \mathrm{grad} f(x,y,z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right), \quad \boldsymbol{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right).$$

The sign nabla ∇ denotes the Hamiltonian vector differential operator. In order to emphasize its "vectority", the symbol $\boldsymbol{\nabla}$ is written in bold.

To simplify the presentation, we made some inaccuracies in the presentation, which should be mentioned separately.

— Strictly speaking, the gradient is a covector. Our definition reflects this by writing the vector components in a row rather than a column.
— The definition of the gradient is based on the Cartesian coordinate system. A more general definition should be given in a componentless form.

The scalar field $f(\mathbf{r})$ generates a vector field $\boldsymbol{\nabla} f$, which characterizes the direction of the greatest change in the scalar field $f(\mathbf{r})$.

Divergence of the vector field $\mathbf{V} = (V_x, V_y, V_z)^T$ is a scalar, calculated in Cartesian coordinates as follows:

$$\boldsymbol{\nabla} \cdot \mathbf{V} = \mathrm{div}\mathbf{V} = \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} + \frac{\partial V_z}{\partial z} = \sum_{i=1}^{3} \frac{\partial V^i}{\partial x^i}.$$

Here "$\cdot$" denotes the scalar multiplication operation $\boldsymbol{\nabla} \cdot \mathbf{V} = (\boldsymbol{\nabla}, \mathbf{V})$.

The *Rotor* of a vector field $\mathbf{V}$ is a vector calculated in Cartesian coordinates as follows:

$$\boldsymbol{\nabla} \times \mathbf{V} = \begin{vmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ V_x & V_y & V_z \end{vmatrix} = \\ = \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right)\mathbf{e}_x + \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right)\mathbf{e}_y + \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)\mathbf{e}_z.$$

Highlight also that the rotor is not a vector in the strict sense. In classical vector analysis it is called a *pseudovector*, but a deeper geometric meaning is revealed only when tensor algebra is involved, where the rotor can be represented either as a 2-form or as a bivector.

Also, when writing the wave equation, the Laplace operator will be used, which is written in the following form:

$$\nabla^2 = (\nabla, \nabla) = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}.$$

We will also need the following two relations [5]:

$$\begin{aligned} \boldsymbol{\nabla} \times f\mathbf{V} &= f\,\boldsymbol{\nabla} \times \mathbf{V} + \boldsymbol{\nabla} f \times \mathbf{V}, \\ \boldsymbol{\nabla} \cdot f\mathbf{V} &= f\,\boldsymbol{\nabla} \cdot \mathbf{V} + (\boldsymbol{\nabla} f, \mathbf{V}). \end{aligned} \tag{18}$$

A vector field is called *potential* if there exists a scalar field $f(x, y, z)$ such that

$$\mathbf{V} = \boldsymbol{\nabla} f = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right),$$
$$\mathrm{d}f = V_x\,\mathrm{d}x + V_y\,\mathrm{d}y + V_z\,\mathrm{d}z\,.$$

In turn, a vector field is called *solenoidal* (tubular) if there exists a vector field $\mathbf{U}$ such that

$$\mathbf{V} = \boldsymbol{\nabla} \times \mathbf{U}.$$

Acknowledgments

This publication has been supported by the RUDN University Scientific Projects Grant System, project No 021934-0-000 (Korolkova A. V., Gevorkyan M. N.). This paper has been supported by the RUDN University Strategic Academic Leadership Program.

References

[1] H. Bruns, "Das Eikonal," German, in *Abhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften*. Leipzig: S. Hirzel, 1895, vol. 21.

[2] F. C. Klein, "Über das Brunssche Eikonal," German, *Zeitscrift für Mathematik und Physik*, vol. 46, pp. 372–375, 1901.

[3] J. A. Stratton, *Electromagnetic Theory*. MGH, 1941.

[4] L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics, The Classical Theory of Fields*, 4th. Butterworth-Heinemann, 1975, vol. 2, 402 pp.

[5] M. Born and E. Wolf, *Principles of Optics*, 7th. Cambridge University Press, 1999, 952 pp.

[6] D. V. Sivukhin, "The international system of physical units," *Soviet Physics Uspekhi*, vol. 22, no. 10, pp. 834–836, Oct. 1979. DOI: 10.1070/pu1979v022n10abeh005711.

[7] D. S. Kulyabov, A. V. Korolkova, T. R. Velieva, and M. N. Gevorkyan, "Numerical analysis of eikonal equation," in *Saratov Fall Meeting 2018: Laser Physics, Photonic Technologies, and Molecular Modeling*, V. L. Derbov, Ed., ser. Progress in Biomedical Optics and Imaging - Proceedings of SPIE, vol. 11066, Saratov: SPIE, Jun. 2019, p. 56. DOI: 10.1117/12.2525142. arXiv: 1906.09467.

[8] D. S. Kulyabov, M. N. Gevorkyan, and A. V. Korolkova, "Software implementation of the eikonal equation," in *Proceedings of the Selected Papers of the 8th International Conference "Information and Telecommunication Technologies and Mathematical Modeling of High-Tech Systems" (ITTMM-2018), Moscow, Russia, April 16, 2018*, D. S. Kulyabov, K. E. Samouylov, and L. A. Sevastianov, Eds., ser. CEUR Workshop Proceedings, vol. 2177, Moscow, Apr. 2018, pp. 25–32.

For citation:

A. V. Fedorov, C. A. Stepa, A. V. Korolkova, M. N. Gevorkyan, D. S. Kulyabov, Methodological derivation of the eikonal equation, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 399–418. DOI: 10.22363/2658-4670-2023-31-4-399-418.

Information about the authors:

Arseny V. Fedorov — PhD student of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: 1042210107@rudn.ru, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0002-3036-0117)

Christina A. Stepa — PhD student of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: 1042210111@pfur.ru, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0002-4092-4326)

Anna V. Korolkova — Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: korolkova-av@rudn.ru, phone: +7(495)9520250, ORCID: https://orcid.org/0000-0001-7141-7610)

Migran N. Gevorkyan — Candidate of Sciences in Physics and Mathematics, Associate Professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: gevorkyan-mn@rudn.ru, phone: +7(495)9550927, ORCID: https://orcid.org/0000-0002-4834-4895)

Dmitry S. Kulyabov (Russian Federation) — Doctor of Sciences in Physics and Mathematics, Professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University); Senior Researcher of Laboratory of Information Technologies, Joint Institute for Nuclear Research (e-mail: kulyabov-ds@rudn.ru, phone: +7 (495) 952-02-50, ORCID: https://orcid.org/0000-0002-0877-7063, ResearcherID: I-3183-2013, Scopus Author ID: 35194130800)

УДК 537.8:514.762.37
DOI: 10.22363/2658-4670-2023-31-4-399-418
EDN: GCUXWK

Методический вывод уравнения эйконала

А. В. Фёдоров[1], **К. А. Штепа**[1], **А. В. Королькова**[1], **М. Н. Геворкян**[1], **Д. С. Кулябов**[1,2]

[1] *Российский университет дружбы народов, ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация*
[2] *Объединённый институт ядерных исследований, ул. Жолио-Кюри, д. 6, Дубна, 141980, Российская Федерация*

Аннотация. Обычно при работе с уравнением эйконала ссылаются на его вывод в монографии Борна и Вольфа. Вывод этого уравнения выполнен достаточно небрежно. Для того чтобы разобраться в этом выводе, требуется определённое число имплицитных предположений. Для лучшего понимания приближения эйконала и для методических целей авторы решили повторить вывод уравнения эйконала, эксплицировав все возможные допущения. Методически предлагается следующий алгоритм вывода уравнения эйконала. Из уравнения Максвелла выводится волновое уравнение. При этом явно вводятся все условия, при которых это возможно сделать. Далее от волнового уравнения осуществляется переход к уравнению Гельмгольца. От уравнения Гельмгольца при приложении определённых допущений производится переход к уравнению эйконала. После разбора всех допущений и шагов реализуется собственно переход от уравнений Максвелла к уравнению эйконала. При выводе уравнения эйконала используется несколько формализмов. В качестве первого формализма используется стандартный формализм векторного анализа. Уравнения Максвелла и уравнение эйконала записывается в виде трёхмерных векторов. После этого и для уравнений Максвелла, и для уравнения эйконала используется ковариантный 4-мерный формализм. Результатом работы является методически выдержанное описание уравнения эйконала.

Ключевые слова: эйконал, уравнения Максвелла, волновое уравнение, векторное представление, тензорное представление