MATTEX 2018,
CONFERENCE PROCEEDING, v. 1, pp. 39- 41

MODELING TIME SERIES OF COUNTS UNDER CENSORING AND TRUNCATION*

ISABEL PEREIRA, MARIA EDUARDA SILVA

ABSTRACT: *Censored and truncated data are frequently encountered in diverse fields including environmental monitoring, medicine, economics and social sciences. Censoring occurs when observations are available only for a restricted range, e.g., due to a detection limit. Truncation, on the other hand, occurs if observations in some range are lost. This work considers the analysis of time series of counts under censoring and truncation based on first order integer autoregressive (INAR) models. The focus is on estimation and inference problems.*

KEYWORDS: *Censored count series, INAR model, Truncated time series*

1 Introduction

Observations collected over time or space are often autocorrelated rather than independent. Time series measurements are often observed with data irregularities, e.g., observations with a detection limit. For instance, a monitoring device usually has a detection limit and it records the limit value when the true value exceeds/precedes the detection limit. This is often called censoring. So, censoring occurs if observations Y_i are available for a restricted range of Y_i arising from aggregation or may be imposed by survey design. This is very common in various situations in physical sciences, business, and economics. However the dependent variables can be limited in their range as well by truncation. Truncation arises if observations Y_i in some range of Y_i are totally lost.However the dependent variables can be limited in their range by truncation. Truncation arises if observations in some range are totally lost. Examples of left truncation include the number of bus trips made per week in surveys taken on buses, the number of shopping trips made by individuals sampled at a mall, and the number of unemployment spells among a pool of unemployed. Right truncation occurs if high counts are not observed. The most common form of truncation in count models is left truncation at zero. There is an extensive literature on regression models in which dependent variables are censored or underlying distributions are truncated, see [3] for a comprehensive review including discrete dependent variables. Nevertheless, the analysis of time series under censoring or truncation has received little attention in the literature and regard continuous valued processes only, see [4] and [1].

The model considered in this work is a time series of counts $X_1,\dots,X_T$, generated according to the first order integer autoregressive process

(1) $$X_t = \alpha \diamond X_{t-1} + \varepsilon_t$$

where the arrival process ε_t is a sequence of independent and identically distributed non-negative integer valued random variables, independent of X_{t-1}, with mean μ_ε and finite variance σ_ε^2 and, conditional on X_{t-1}, $\alpha \diamond X_{t-1}$ is an integer valued random variable whose probability distribution, denoted by $g(\cdot|x;\alpha)$ depends on the parameter α which may be a vector of parameters. Thus, $'\diamond'$ denotes a random operator, usually called thinning operator, which always produces integer values and introduces serial dependence via the conditioning on X_{t-1}. For a review of thinning

*This work was supported in part by the Portuguese Foundation for Science and Technology (FCT-Fundação para a Ciência e a Tecnologia), through CIDMA - Center for Research and Development in Mathematics and Applications, within project UID/MAT/04106/2013.

operators see [2]. The specifications of the thinning operator and arrival process lead to INAR(1) models that ensure required distributional properties of the marginal distributions, namely that the marginal distribution of X_t is from the same family as ε_t. The transition probabilities are given by

$$\mathrm{p}(X_t|X_{t-1}) = P\left[X_t = k|X_{t-1} = l\right] = \sum_{j=0}^{\min\{k,l\}} g(j|l)P\left[\varepsilon_t = k-j\right]. \tag{2}$$

Considering time series of counts based on first order integer autoregressive models this work aims at giving a contribution towards this direction.

2 A truncated count model

For some reason we do not observe X_t but Y_t which results from left truncating X_t at a value N. This means that a subset of the population that generated the data is unobserved. Then we define a **left truncated at** N **model**, for Y_t as

$$\begin{aligned} X_t &= \alpha \diamond X_{t-1} + \varepsilon_t \\ Y_t &= X_t, \text{if } X_t \geq N \end{aligned} \tag{3}$$

The transition probabilities are now

$$P\left[Y_t = y_t|Y_{t-1} = y_{t-1}\right] = \frac{P\left[X_t = y_t|X_{t-1} = y_{t-1}\right]}{\sum_{i=N}^{\infty}\sum_{j=N}^{\infty} P\left[X_t = j|X_{t-1} = i\right] P^*\left[X_{t-1} = i\right]} \tag{4}$$

with $P^*\left[X_t = j\right] = \frac{P[X_t=j]}{\sum_{i=N}^{\infty} P[X_t=i]}$ and zero outside $y_t, y_{t-1} \in [N, \infty[$.

3 A censored count model

Censoring occurs if observations Y_t are available for a restricted range due to aggregation or detection limits. We can say that censored data are âĂŸâĂŸpiled upâĂŹâĂŹ at a censoring point. We define a **censored at** N **model** as

$$\begin{aligned} X_t &= \alpha \diamond (X_{t-1}) + \varepsilon_t \\ Y_t &= \min\{X_t, N\} = \begin{cases} X_t, & \text{if } X_t \leq N \\ N, & \text{if } X_t > N \end{cases} \end{aligned} \tag{5}$$

The transition probabilities are

$$P\left[Y_t = y_t | Y_{t-1} = y_{t-1}\right] =$$

$$(6) \qquad \begin{cases} P\left[X_t \geq N | X_{t-1} \geq N\right] = \\ \sum_{i=N}^{+\infty} \sum_{j=N}^{+\infty} P\left[X_t = j | X_{t-1} = i\right] P^*\left[X_{t-1} = i\right], & y_t, y_{t-1} = N \\ \\ P\left[X_t = y_t | X_{t-1} \geq N\right] = \\ \sum_{i=N}^{+\infty} P\left[X_t = y_t | X_{t-1} = i\right] P^*\left[X_{t-1} = i\right], & y_t < N, y_{t-1} = N \\ \\ P\left[X_t \geq N | X_{t-1} = y_{t-1}\right] = \\ \sum_{j=N}^{+\infty} P\left[X_t = j | X_{t-1} = y_{t-1}\right], & y_t = N, y_{t-1} < N \\ \\ \frac{P[X_t = y_t | X_{t-1} = y_{t-1}]}{P[X_t < N | X_{t-1} < N]} = \\ \frac{P[X_t = y_t | X_{t-1} = y_{t-1}]}{\sum_{i=N}^{+\infty} \sum_{j=0}^{N} P[X_t = j | X_{t-1} = i] P^*[X_{t-1} = i]}, & y_t, y_{t-1} < N \end{cases}$$

4 Estimation procedures

Neglecting censoring and truncation in the time series hinders meaningful statistical inference, leading to model misspecification, biased parameter estimation, and poor forecasts. We study the regression properties of the censored INAR(1) models and show that least squares estimation of the parameters is no longer appropriate. Likelihood analysis of the censored INAR(1) processes is developed, including maximum likelihood and maximum pseudo-likelihood estimations. Some problems related to the censored or truncated count data models are pointed out and illustrated.

REFERENCES:

[1] Choi, Seokwoo Jake and Portnoy, Stephen (2016). Quantile Autoregression for Censored Data. *J. Time. Ser. Anal.*,37, pp. 603âĂŞ623.

[2] Gouveia, Sónia and Scotto, Manuel and Weiss, Christian H. and Ferreira, Paulo J.S.G. (2016). Binary autoregressive geometric modelling in a DNA context. *Journal of the Royal Statistical Society, Series C*.

[3] Greene, W. (2005). Censored Data and Truncated Distributions, *Working Papers 05-08*, New York University, Leonard N. Stern School of Business, Department of Economics.

[4] Park, Jung Wook and Genton, Marc G. and Ghosh, Sujit K. (2007). Censored time series analysis with autoregressive moving average models. *The Canadian Journal of Statistics*, 35, pp. 151-168.

Isabel Pereira
University of Aveiro & CIDMA, Portugal
E-mail:isabel.pereira@ua.pt

Maria Eduarda Silva
University of Porto & CIDMA, Portugal
E-mail:me.pintosilva@gmail.com

ON THE COMPOSITE LOGNORMAL-PARETO DISTRIBUTION WITH UNCERTAIN THRESHOLD*

SEZEN MUTALI (CIORABAI)

ABSTRACT: *An important challenge related to composite distribution is the estimation of their unknown threshold. For this purpose, several methods have been proposed in the literature, yielding different results for the same data. Therefore, in this paper we go further on and assume that the threshold is uncertain, allowing it to take values in a closed interval. Based on the arithmetic of intervals, we study the resulting model under continuity and differentiability conditions.*

KEYWORDS: *Composite Lognormal-Pareto Distribution, Uncertain Threshold, Interval arithmetic.*

1 Introduction

A spliced distribution is piecewise built from different distributions defined on mutually disjoint intervals. In particular, a composite distribution is usually built from two components: for example, the composite lognormal-Pareto distribution consists of a lognormal density up to a certain threshold and of a Pareto density from that threshold on. We note that a composite distribution can also be interpreted as a two-component mixture of two distributions singly truncated from above and, respectively, from below (their truncation point being the threshold itself), with corresponding mixing weights. Such distributions are useful to model, e.g., loss data arising in insurance because they are typically positively skewed. Therefore, the usual lognormal-Pareto distribution is well studied in the actuarial literature, starting from the paper of Cooray and Ananda [2]. An important challenge is the estimation of the unknown threshold. For this purpose, several methods have been proposed in the literature, yielding different results for the same data, usually necessitating the examination of all possible positions of the threshold relatively to the data; see, e.g., Teodorescu and Vernic [7] or the R package of Nadarajah and Bakar [4] and the references therein. Moreover, Pigeon and Denuit [5] assumed that the threshold is the outcome of a random variable (r.v.). In this paper, we go further on and assume that the threshold is uncertain, allowing it to take values in a closed interval. Based on the arithmetic of intervals, we study the resulting model under continuity and differentiability conditions. Such an assumption has the purpose to better capture the uncertainty encountered in the data set.

Therefore, the paper is structured in two more sections: in Section 2, we recall the basics of interval arithmetic and the form of the composite lognormal-Pareto distribution with its main characteristics; in Section 3, we introduce the threshold of interval type and study its effect on the parameters, the density, the distribution function and the expected value of the resulting lognormal-Pareto model. We end with a conclusions and further study section.

2 Preliminaries

2.1 Basics of interval arithmetic

Measurement uncertainties can be directly incorporated into calculation by means of the interval arithmetic, i.e., by using closed intervals on the real line represented as $[\underline{x},\overline{x}] = \{x \in \mathbb{R} : \underline{x} \leq x \leq \overline{x}\}$. Consequently, an arithmetic for intervals and interval valued extensions of functions had

*Partially supported by Scientific Research Grant DDVU 02/91 of MON and Scientific Research Grant RD-08-75/2016 of Shumen University.

been defined with the purpose to provide rigorous enclosures of solutions in various applications (see, e.g., the books [1] and [3]). In the following, we recall several interval arithmetic operation:

1. addition: $[\underline{x},\overline{x}]+[\underline{y},\overline{y}]=[\underline{x}+\underline{y},\overline{x}+\overline{y}]$, $a+[\underline{x},\overline{x}]=[a+\underline{x},a+\overline{x}], a\in\mathbb{R}$;
2. subtraction: $[\underline{x},\overline{x}]-[\underline{y},\overline{y}]=[\underline{x}-\overline{y},\overline{x}-\underline{y}]$;
3. multiplication: $[\underline{x},\overline{x}]\cdot[\underline{y},\overline{y}]=[\underline{m},\overline{m}]$, where:

$$\underline{m}=\min\left(\underline{xy},\underline{x}\overline{y},\overline{x}\underline{y},\overline{xy}\right),\overline{m}=\max\left(\underline{xy},\underline{x}\overline{y},\overline{x}\underline{y},\overline{xy}\right);$$

$$a\cdot[\underline{x},\overline{x}]=\begin{cases}[a\underline{x},a\overline{x}] & ,a>0\\ [a\overline{x},a\underline{x}] & ,a<0\end{cases}.$$

2.2 Composite lognormal-Pareto distribution

We shall use the composite lognormal-Pareto density definition given by Scollnik [6]:

$$f(x;\mu,\sigma,a,\theta,r)=\begin{cases} r\frac{1}{x\sigma\sqrt{2\pi}\phi\left(\frac{\ln\theta-\mu}{\sigma}\right)}e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2} & ,0<x\le\theta\\ (1-r)\frac{a\theta^a}{x^{a+1}} & ,x>\theta,\end{cases} \tag{1}$$

where $\sigma,\theta,a>0$ and $\mu\in\mathbb{R}$. Also, φ and ϕ denote the standard normal density function and, respectively, distribution function. Imposing the continuity condition $f(\theta_-)=f(\theta_+)$ leads to:

$$r(\theta)=\left(1+\frac{e^{-\frac{1}{2}\left(\frac{\ln\theta-\mu}{\sigma}\right)^2}}{a\sigma\sqrt{2\pi}\phi\left(\frac{\ln\theta-\mu}{\sigma}\right)}\right)^{-1}.$$

We note that $r(\theta)$ is an increasing function in θ. Moreover, a differentiability condition is usually imposed, i.e., $f'(\theta_-)=f'(\theta_+)$, yielding:

$$\frac{\ln\theta-\mu}{\sigma}=a\sigma, \tag{2}$$

from where we obtain:

$$r=\left(1+\frac{\varphi(a\sigma)}{a\sigma\phi(a\sigma)}\right)^{-1}. \tag{3}$$

The distribution function of the composite lognormal-Pareto is given by:

$$F(x;\mu,\sigma,a,\theta,r)=\begin{cases} r\frac{\phi\left(\frac{\ln x-\mu}{\sigma}\right)}{\phi\left(\frac{\ln\theta-\mu}{\sigma}\right)} & ,0<x\le\theta\\ r+(1-r)\left(1-\left(\frac{\theta}{x}\right)^a\right) & ,x>\theta.\end{cases} \tag{4}$$

The expected value of this composite distribution exists only for $a>1$ and is equal to:

$$\mathbb{E}(f)=\frac{r}{\phi(a\sigma)}e^{\mu+\frac{\sigma^2}{2}}+(1-r)\frac{a\theta}{a-1},a>1. \tag{5}$$

Remark 1. *According to the remark in the introduction, we can rewrite model (1) as a two-component mixture, which, in terms of r.v.s, is expressed as*

$$Z = IX + (1-I)Y,$$

where I is a Bernoulli (r) random variable, X is a lognormal distributed r.v. truncated from above in θ, while Y is a Pareto r.v. defined for $x > \theta$.

3 Composite lognormal-Pareto with interval type threshold

We now assume that the threshold θ is of interval type, i.e., $\theta \in [\underline{\theta}, \overline{\theta}]$. From condition (2) we note that if θ is of interval type then at least one of the other parameters must also be of interval type. Considering the meaning of the parameters, we assume that μ, the lognormal parameter related to the mean, is also of interval type, while the parameters σ, a are single values. Since $\mu = \ln\theta - a\sigma^2$, the corresponding interval for μ results from the interval of θ as $\mu \in [\underline{\mu}, \overline{\mu}]$, where $\underline{\mu} = \ln\underline{\theta} - a\sigma^2, \overline{\mu} = \ln\overline{\theta} - a\sigma^2$; therefore, we shall omit μ from the notation. From formula (3), we also note that since a, σ are now constant values, r is also constant, so we also omit it from the notation.

We start by noting that when θ is of interval type, we no longer have a proper distribution, but a bundle (or fascicle) of composite distributions. Let us now have a look as some density components of this bundle.

In Figure 1, we plotted the density (1) for $a = 0.5, \sigma = 1$ and $\theta \in \{2; 2.3; 2.5; 2.7; 3\} \subset [2, 3]$. In this case, $r = 0.4955$, while $m \in [0.1932, 0.5986]$.

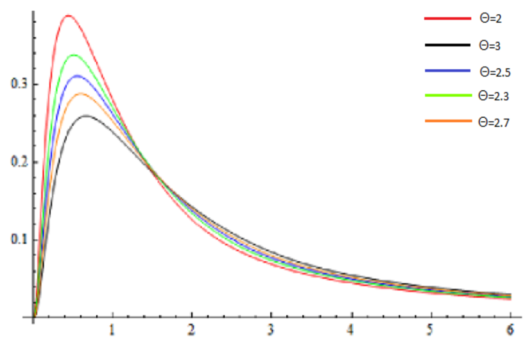

Figure 1: Density shapes for θ of interval type

We shall now see some properties of the resulting fascicle of distributions.

Proposition 1. *When $\theta \in [\underline{\theta}, \overline{\theta}]$ and $x > 0$ is fixed, it holds that*
i) The density $f(x;\sigma,a,\theta)$ is also of interval type;
ii) The distribution function $F(x;\sigma,a,\theta)$ is also of interval type;
iii) The expected value $\mathbb{E}(\sigma,a,\theta) = \mathbb{E}(f)$ is also of interval type.

Proof. *i)* In this case,

$$f(x;\sigma,a,\theta) = \begin{cases} r\dfrac{\varphi\left(\frac{\ln x - \ln\theta + a\sigma^2}{\sigma}\right)}{x\sigma\phi(a\sigma)} & ,0 < x \leq \theta \\ (1-r)\frac{a\theta^a}{x^{a+1}} & ,x > \theta. \end{cases} \tag{6}$$

Since x is fixed and we vary $\theta \in [\underline{\theta}, \overline{\theta}]$, we have three cases:
Case 1: When $x < \underline{\theta} \leq \theta \leq \overline{\theta}$, we use formula:

$$f(x;\sigma,a,\theta) = r\frac{\varphi(\frac{\ln x - \ln\theta + a\sigma^2}{\sigma})}{x\sigma\phi(a\sigma)},$$

which is clearly continuous in θ, hence $f(x;\sigma,a,[\underline{\theta},\overline{\theta}])$ is also an interval.
Case 2: When $\underline{\theta} \leq \theta \leq \overline{\theta} < x$, we use formula:

$$f(x;\sigma,a,\theta) = (1-r)\frac{a\theta^a}{x^{a+1}},$$

which immediately yields $f(x;\sigma,a,[\underline{\theta},\overline{\theta}]) = \left[(1-r)\frac{a\underline{\theta}^a}{x^{a+1}}, (1-r)\frac{a\overline{\theta}^a}{x^{a+1}}\right]$.
Case 3: When $\underline{\theta} \leq x \leq \overline{\theta}$, we have two subcases: if $\underline{\theta} \leq \theta \leq x$, then we use the second formula of equation (6) and obtain an interval, while if $x \leq \theta \leq \overline{\theta}$ we use the first formula of equation (6) and obtain another interval. We prove that these two intervals overlap. More precisely, we prove that we have equality at the limit $\theta = x$, i.e.,

$$r\frac{\varphi(\frac{\ln x - \ln x + a\sigma^2}{\sigma})}{x\sigma\phi(a\sigma)} = (1-r)\frac{ax^a}{x^{a+1}} \Leftrightarrow r\frac{\varphi(a\sigma)}{x\sigma\phi(a\sigma)} = (1-r)\frac{a}{x}. \tag{7}$$

To prove this we insert the formula (3) of r and obtain:

$$r\frac{\varphi(a\sigma)}{x\sigma\phi(\alpha\sigma)} = \frac{a\sigma\phi(a\sigma)}{a\sigma\phi(a\sigma)+\varphi(a\sigma)} \cdot \frac{\varphi(a\sigma)}{x\sigma\phi(\alpha\sigma)} = \frac{\varphi(a\sigma)}{a\sigma\phi(a\sigma)+\varphi(a\sigma)} \cdot \frac{a}{x},$$

while

$$(1-r)\frac{a}{x} = \left(1 - \frac{a\sigma\phi(a\sigma)}{a\sigma\phi(a\sigma)+\varphi(a\sigma)}\right)\frac{a}{x} = \frac{\varphi(a\sigma)}{a\sigma\phi(a\sigma)+\varphi(a\sigma)} \cdot \frac{a}{x},$$

i.e., the equality (7) holds, which completes the proof of *(i)*.
ii) Based on the distribution function's formula:

$$F(x;\sigma,a,\theta) = \begin{cases} r\frac{\phi\left(\frac{\ln x - \ln\theta + a\sigma^2}{\sigma}\right)}{\phi(a\sigma)} & ,0 < x \leq \theta \\ r + (1-r)\left(1 - \left(\frac{\theta}{x}\right)^a\right) & ,x > \theta, \end{cases}$$

the proof is very similar with the above one, hence we omit it.
iii) For $a > 1$, we rewrite:

$$\mathbb{E}(\sigma,a,\theta) = \frac{r}{\phi(a\sigma)}e^{\ln\theta - a\sigma^2 + \frac{\sigma^2}{2}} + (1-r)\frac{a\theta}{a-1},$$

which is continuous and increasing in θ and, clearly, $\mathbb{E}(\sigma,a,[\underline{\theta},\overline{\theta}]) = \left[\mathbb{E}(\sigma,a,\underline{\theta}), \mathbb{E}(\sigma,a,\overline{\theta})\right]$. This completes the proof. □

In Figure 2, we plotted the expected value for $\theta \in [1,3], \sigma = 1$ and various $a \in \{1.5, 2, 3\}$. In Figure 3, we plotted the expected value for $\theta \in [1,3], a = 2$ and various $\sigma \in \{0.5, 1, 1.3\}$. We note that both plots support the conclusion of *(iii)* in the above proposition.

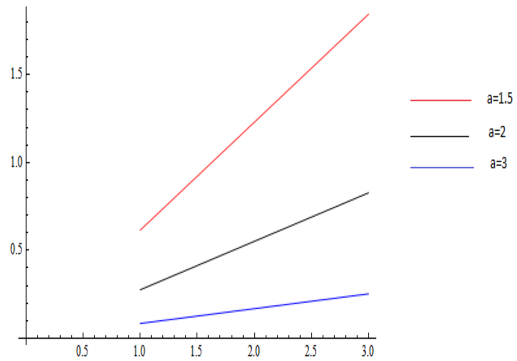

Figure 2: Expected value for θ of interval type and various a

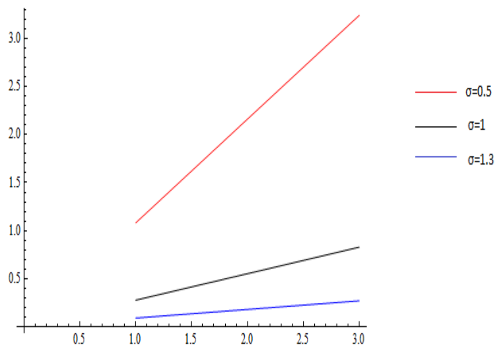

Figure 3: Expected value for θ of interval type and various σ

4 Conclusions

In this paper, we studied the effect of an interval type threshold on the composite lognormal-Pareto distribution and concluded that the density, the distribution function and the expected value also become of interval type. Because of the restrictions imposed by the continuity and differentiability conditions, we think that this study should be extended to the situation where the differentiability condition is relaxed. Moreover, as further work, we intend to study the same distribution under the assumption that the threshold is a fuzzy number and to apply the results on some real data.

Acknowledgements. The author is very grateful to the anonymous referee for the valuable comments.

REFERENCES:

[1] Alefeld, G. and Herzberger, J. (2012). Introduction to interval computation. Academic press.

[2] Cooray, K. and Ananda, M. M. (2005). Modeling actuarial data with a composite lognormal-Pareto model. Scandinavian Actuarial Journal, 2005(5), 321-334.

[3] Moore, R. E., Kearfott, R. B. and Cloud, M. J. (2009). Introduction to interval analysis. Siam.

[4] Nadarajah, S. and Bakar, S. (2013). CompLognormal: An R Package for Composite Lognormal Distributions. R journal, 5(2), 98-104.

[5] Pigeon, M. and Denuit, M. (2011). Composite Lognormal–Pareto model with random threshold. Scandinavian Actuarial Journal, 2011(3), 177-192.

[6] Scollnik, D. P. (2007). On composite lognormal-Pareto models. Scandinavian Actuarial Journal, 2007(1), 20-33.

[7] Teodorescu, S. and Vernic, R. (2013). On composite Pareto models. Mathematical Reports, 15(65), 11-29.

Sezen Mutali (Ciorabai),
Ovidius University of Constanta, Romania
E-mail: m.sezenn@yahoo.com

MERGING OF BIVARIATE COMPOUND POISSON RISK PROCESSES WITH SHOCKS*

PAVLINA K. JORDANOVA

ABSTRACT: *The paper considers an example of reducing a bivariate Compound Poisson risk process with common shocks to an univariate Cramer-Lundberg risk model. Then using the results about Compound Poisson risk model the main numerical characteristics and distributions related to the new model are obtained. Analogous approach can be applies for any number of compound Poisson processes. It can be useful also in renewal theory or queuing theory for simplifying systems or building stochastically equivalent models.*

KEYWORDS: *Multivariate risk models, Common shocks, Cramer-Lundberg risk model.*

1 Introduction

The theory of Compound Poisson risk processes origins since the works of Filip Lundberg [9] and Harald Cramer [3] and finds good applications nowadays. In 1967 Marshall and Olkin [10], [11] investigate a bivariate Poisson model with dependent coordinates. They lay the foundations of multivariate models with common shocks. Particular cases of such models are considered for example in Cossette and Marceau [2], or Wand [19]. In any of them the task for characterising of total claim amount distribution reduces to calculations of multivariate compound distributions. Recursive algorithms for bivariate compound distributions of Panjer's family are obtained e.g. in Panjer and Willmot[13, 14], Hesselager [12] among others. Sundt and Vernic [16, 17, 18] generalize these recursions to the multivariate case. Recently Jordanova and Stehlik [8] show that in many cases these multivariate risk models can be reduced to classical Cramer-Lundberg model. Here we consider an example of merging of Compound Poisson processes. More general results could be seen in Jordanova and Stehlik [8].

The organization of the paper is the following...

Along the paper $\perp$ means independent, $HPP(\lambda)$ is an abbreviation of "homogeneous Poisson process with intensity $\lambda > 0$", $g_\xi(x) = Ez^\xi$ is the probability generating function of the random variable ξ and $\sum_{i=1}^{0} = 0$. The independence between the summands and the number of summands in the following random sums can be relaxed without changing the final results if we assume that the compounding distributions describe the distribution of the summands, given the number of summands.

2 Description of the model

Consider a risk process R consisting of two dependent types of polices. The dependent counting processes M_1 and M_2 of both types are compound Poisson process $B_k = \{B_k(t) : t \geq 0\}$, $k = 1,2$ with common shocks, described by independent compound Poisson process $B_0 = \{B_0(t) : t \geq 0\}$. More precisely for all $t \geq 0$

(1) $$M_1(t) = B_1(t) + B_0(t), \quad M_2(t) = B_2(t) + B_0(t),$$

*The author is grateful to the bilateral projects Bulgaria - Austria, 2016-2019, "Feasible statistical modelling for extremes in ecology and finance", Contract number 01/8, 23/08/2017, WTZ Project No. BG 09/2017, and the scientific project RD-08-125/06.02.2018, Shumen University.

where the processes B_0, B_1 and B_2 are independent, and for $k = 0,1,2$,

$$B_k(t) = \sum_{i=1}^{N_k(t)} \eta_{ki},\ N_k \sim HPP(\lambda_k),\ g_k(z) = Ez^{\eta_{k1}}, z \geq 0, \quad \{\eta_{ki}\}_{i=1}^{\infty} \perp N_k, \tag{2}$$

$\eta_{k1}, \eta_{k2}, ...$ independent identically distributed (i.i.d.), and $\eta_{11} \perp \eta_{21} \perp \eta_{01}$.

$S = \{S(t) : t \geq 0\}$ is the total claim amount process. For all $t \geq 0$

$$S(t) = S_1(t) + S_2(t), \tag{3}$$

$$\begin{aligned} S_1(t) &= \sum_{i=1}^{B_1(t)} Y_{1i} + \sum_{i=1}^{B_0(t)} Y_{0i}, \quad \{Y_{ki}\}_{i=1}^{\infty} \perp B_i,\ k = 0,1, \\ S_2(t) &= \sum_{i=1}^{B_2(t)} Y_{2i} + \sum_{i=1}^{B_0(t)} Y_{0i}, \quad \{Y_{ki}\}_{i=1}^{\infty} \perp B_i,\ k = 0,2. \end{aligned} \tag{4}$$

$Y_{k1}, Y_{k2}, ...$ i.i.d., and $Y_{01} \perp Y_{11} \perp Y_{21}$.

$R = \{R(t) : t \geq 0\}$ is the risk reserve process and for all $t \geq 0$

$$R(t) = u + ct - S(t), \tag{5}$$

where $u \geq 0$ is the initial capital and $c > 0$ is the premium income rate.

3 Stochastic equivalent models

In this section first we obtain a stochastically equivalent presentation of the bivariate counting process $\{(M_1(t), M_2(t)) : t \geq 0\}$ which shows that it is a particular case of multivariate compound Poisson processes of type I, discussed in Sundt and Vernic [17, 18]. Then, in Theorem 2 we derive the stochastic equivalent presentation of the processes (S_1, S_2) as a compound Poisson process of Type I. Finally, in Theorem 3 we describe a stochastic equivalence presentation of the risk process R which allows us to state that this is a Cramer-Lundberg (C-L) risk process.

Denote by $\lambda = \lambda_1 + \lambda_2 + \lambda_0$. In order to formulate our main results let us define a random vector (I_1, I_2, I_0) with probability mass function (p.m.f.)

$$P(I_1 = 1, I_2 = 0, I_0 = 0) = \frac{\lambda_1}{\lambda},\ dP(I_1 = 0, I_2 = 1, I_0 = 0) = \frac{\lambda_2}{\lambda},\ P(I_1 = 0, I_2 = 0, I_0 = 1) = \frac{\lambda_0}{\lambda}, \tag{6}$$

and zero otherwise. Assume random vectors $(I_{11}, I_{21}, I_{01}), (I_{12}, I_{22}, I_{02}), ...$ are i.i.d. with p.m.f. (6) and $(\eta_{11}, \eta_{21}, \eta_{01}), (\eta_{12}, \eta_{22}, \eta_{02}), ...$ are i.i.d. with independent coordinates with probability generating functions (p.g.fs.) correspondingly g_1, g_2, g_0. Denote by $N \sim HPP(\lambda)$,

$$A_1(t) = \sum_{i=1}^{N(t)} (I_{1i}\eta_{1i} + I_{0i}\eta_{0i}), \tag{7}$$

$$A_2(t) = \sum_{i=1}^{N(t)} (I_{2i}\eta_{2i} + I_{0i}\eta_{0i}). \tag{8}$$

Assume N, $(I_{11}, I_{21}, I_{01}), (I_{12}, I_{22}, I_{02}), ...$ and $(\eta_{11}, \eta_{21}, \eta_{01}), (\eta_{12}, \eta_{22}, \eta_{02}), ...$ are independent.

Theorem 1. The bivariate processes (M_1, M_2) and (A_1, A_2) coincide in the sense of their finite dimensional distributions.

Proof: All processes (M_1, M_2) and (A_1, A_2) have homogeneous and independent additive increments and they start from the coordinate beginning, therefore in order to prove their stochastic equivalence it is enough to prove equality of their univariate time intersections. Due to the uniqueness of the correspondence between the probability laws and their p.g.fs. it is enough to derive equality between the p.g.fs. Consider $t > 0$, $z_1 > 0$ and $z_2 > 0$. The definition and the multiplicative property of p.g.fs., (1), (2), and the well known form of the p.g.f. of compound Poisson processes imply

$$\begin{aligned}
E\left[z_1^{M_1(t)} z_2^{M_2(t)}\right] &= E\left[z_1^{B_1(t)+B_0(t)} z_2^{B_2(t)+B_0(t)}\right] = E\left[z_1^{B_1(t)}\right] E\left[z_2^{B_2(t)}\right] E\left[(z_1 z_2)^{B_0(t)}\right] \\
&= E\left[z_1^{\sum_{i=1}^{N_1(t)} \eta_{1i}}\right] E\left[z_2^{\sum_{i=1}^{N_2(t)} \eta_{2i}}\right] E\left[(z_1 z_2)^{\sum_{i=1}^{N_0(t)} \eta_{0i}}\right] \\
&= exp\{-\lambda_1 t[1 - g_1(z_1)]\} exp\{-\lambda_2 t[1 - g_2(z_2)]\} exp\{-\lambda_0 t[1 - g_0(z_1 z_2)]\} \\
&= exp\{-(\lambda_1 + \lambda_2 + \lambda_0)t - \lambda_1 t g_1(z_1) - \lambda_2 t g_2(z_2) - \lambda_0 t g_0(z_1 z_2)\} \\
&= exp\left\{-(\lambda_1 + \lambda_2 + \lambda_0)t\left[1 - \frac{\lambda_1}{\lambda} g_1(z_1) - \frac{\lambda_2}{\lambda} g_2(z_2) - \frac{\lambda_0}{\lambda} g_0(z_1 z_2)\right]\right\}.
\end{aligned}$$

The definitions (7), (8), the formula for double expectations, the independence between N and the other components of the processes A_1 and A_2 entail

$$\begin{aligned}
E\left[z_1^{A_1(t)} z_2^{A_2(t)}\right] &= E\left[z_1^{\sum_{i=1}^{N(t)}(I_{1i}\eta_{1i}+I_{0i}\eta_{0i})} z_2^{\sum_{i=1}^{N(t)}(I_{2i}\eta_{2i}+I_{0i}\eta_{0i})}\right] \\
&= \sum_{n=0}^{\infty} E\left[z_1^{\sum_{i=1}^{n}(I_{1i}\eta_{1i}+I_{0i}\eta_{0i})} z_2^{\sum_{i=1}^{n}(I_{2i}\eta_{2i}+I_{0i}\eta_{0i})}\right] P(N(t) = n)
\end{aligned}$$

By the multiplicative property of p.g.fs. and (6),

$$E\left[z_1^{A_1(t)} z_2^{A_2(t)}\right] = \sum_{n=0}^{\infty} \left\{E\left[z_1^{I_{1i}\eta_{11}+I_{01}\eta_{01}} z_2^{I_{21}\eta_{21}+I_{01}\eta_{01}}\right]\right\}^n P(N(t) = n). \tag{9}$$

Now (6) and the total probability formula imply

$$E\left[z_1^{I_{1i}\eta_{11}+I_{01}\eta_{01}} z_2^{I_{21}\eta_{21}+I_{01}\eta_{01}}\right] = \frac{\lambda_1}{\lambda} g_1(z_1) + \frac{\lambda_2}{\lambda} g_2(z_2) + \frac{\lambda_0}{\lambda} g_0(z_1 z_2). \tag{10}$$

Therefore the fact that $N(t) \sim HPP(\lambda t)$ entails

$$\begin{aligned}
E\left[z_1^{A_1(t)} z_2^{A_2(t)}\right] &= \sum_{n=0}^{\infty} \left\{\frac{\lambda_1}{\lambda} g_1(z_1) + \frac{\lambda_2}{\lambda} g_2(z_2) + \frac{\lambda_0}{\lambda} g_0(z_1 z_2)\right\}^n P(N(t) = n) \\
&= exp\left\{-\lambda t\left[1 - \frac{\lambda_1}{\lambda} g_1(z_1) + \frac{\lambda_2}{\lambda} g_2(z_2) + \frac{\lambda_0}{\lambda} g_0(z_1 z_2)\right]\right\}
\end{aligned}$$

which is exactly $E\left[z_1^{M_1(t)} z_2^{M_2(t)}\right]$, and completes the proof.

Theorem 2. The bivariate total claim amount processes (S_1, S_2), described in (5) is a compound poisson Process of type I (which means with one and the same number of summands). It is stochastically equivalent to the process (S_3, S_4), where

$$S_3(t) = \sum_{n=1}^{N(t)} \left[I_{1n} \sum_{i=1}^{\eta_{1n}} Y_{1i} + I_{0n} \sum_{i=1}^{\eta_{0n}} Y_{0i} \right], \quad t \geq 0, \tag{11}$$

$$S_4(t) = \sum_{n=1}^{N(t)} \left[I_{2n} \sum_{i=1}^{\eta_{2n}} Y_{2i} + I_{0n} \sum_{i=1}^{\eta_{0n}} Y_{0i} \right], \quad t \geq 0, \tag{12}$$

where (I_{11}, I_{21}, I_{01}) is a vector with distribution (6) and the other random elements are the same as those described in (6), (7) and (8). The equivalence is in the sense of their finite dimensional distributions.

Proof: All processes have homogeneous and independent additive increments and they start from the coordinate beginning, therefore in order to prove their stochastic equivalence it is enough to prove equality of their univariate time intersections. Due to the uniqueness of the correspondence between the probability laws and their Laplace-Stieltjes transforms(LSTs), it is enough to derive equality between the (LSTs).

Consider $t \geq 0$, $z_1 \geq 0$ and $z_2 \geq 0$. The definition of LSTs., (5), and the multiplicative property of LSTs, the formula for relation between the LST of compound distribution and LST of summands and the number of summands, the well known form of the LST of compound Poisson processes, (2), (1) imply

$$\begin{aligned}
Ee^{-z_1 S_1(t) - z_2 S_2(t)} &= Ee^{-z_1 \sum_{i=1}^{B_1(t)} Y_{1i} - (z_1+z_2)\sum_{i=1}^{B_0(t)} Y_{0i} - z_2 \sum_{i=1}^{B_2(t)} Y_{2i}} \\
&= Ee^{-z_1 \sum_{i=1}^{B_1(t)} Y_{1i}} Ee^{-(z_1+z_2)\sum_{i=1}^{B_0(t)} Y_{0i}} Ee^{-z_2 \sum_{i=1}^{B_2(t)} Y_{2i}} \\
&= g_{B_1(t)}(Ee^{-z_1 Y_{11}}) g_{B_0}[Ee^{-(z_1+z_2)Y_{01}}] g_{B_2(t)}(Ee^{-Y_{21}}) \\
&= e^{-\lambda_1 t[1 - g_1(Ee^{-z_1 Y_{11}})]} e^{-\lambda_0 t\left\{1 - g_0[Ee^{-(z_1+z_2)Y_{01}}]\right\}} e^{-\lambda_2 t[1 - g_2(Ee^{-z_2 Y_{21}})]} \\
&= e^{-(\lambda_1+\lambda_2+\lambda_3)t\left\{1 - \left[\frac{\lambda_1}{\lambda} g_1(Ee^{-z_1 Y_{11}}) + \frac{\lambda_0}{\lambda} g_0[Ee^{-(z_1+z_2)Y_{01}}] + \frac{\lambda_2}{\lambda} g_2(Ee^{-z_2 Y_{21}})\right]\right\}}.
\end{aligned}$$

From the other side if we consider $z_3 \geq 0$ and $z_4 \geq 0$, the definitions (11), (12), the formula for double expectations, the independence between N and the other components of the processes S_3 and S_4 entail

$$\begin{aligned}
Ee^{-z_3 S_3(t) - z_4 S_4(t)} &= Ee^{-z_3 \sum_{n=1}^{N(t)} \left[I_{1n} \sum_{i=1}^{\eta_{1n}} Y_{1i} + I_{0n} \sum_{i=1}^{\eta_{0n}} Y_{0i}\right] - z_4 \sum_{n=1}^{N(t)} \left[I_{2n} \sum_{i=1}^{\eta_{2n}} Y_{2i} + I_{0n} \sum_{i=1}^{\eta_{0n}} Y_{0i}\right]} \\
&= \sum_{n=0}^{\infty} \left\{ Ee^{-z_3 I_{11} \sum_{i=1}^{\eta_{11}} Y_{1i} - (z_3+z_4) I_{01} \sum_{i=1}^{\eta_{01}} Y_{0i} - z_4 I_{21} \sum_{i=1}^{\eta_{21}} Y_{2i}} \right\}^n P(N(t) = n).
\end{aligned}$$

Now (6) and the total probability formula and the formula about the relation between the for LST of a compound, the LST of the summands and p.g.f. of the number of summands, imply

$$Eexp\left\{ -z_3 I_{11} \sum_{i=1}^{\eta_{11}} Y_{1i} - (z_3+z_4) I_{01} \sum_{i=1}^{\eta_{01}} Y_{0i} - z_4 I_{21} \sum_{i=1}^{\eta_{21}} Y_{2i} \right\} =$$

$$\begin{aligned}
&= \frac{\lambda_1}{\lambda} Ee^{-z_3 \sum_{i=1}^{\eta_{11}} Y_{1i}} + \frac{\lambda_0}{\lambda} Ee^{-(z_3+z_4)\sum_{i=1}^{\eta_{01}} Y_{0i}} + \frac{\lambda_2}{\lambda} Ee^{-z_4 \sum_{i=1}^{\eta_{21}} Y_{2i}} \\
&= \frac{\lambda_1}{\lambda} g_1(Ee^{-z_3 Y_{11}}) + \frac{\lambda_0}{\lambda} g_2(Ee^{-(z_3+z_4)Y_{01}}) + \frac{\lambda_2}{\lambda} g_2(Ee^{-z_4 Y_{21}})
\end{aligned}$$

Therefore the formula for p.g.f. of Poisson distribution entails

$$\begin{aligned} Ee^{-z_3 S_3(t) - z_4 S_4(t)} &= \sum_{n=0}^{\infty} \left\{ \frac{\lambda_1}{\lambda} g_1(Ee^{-z_3 Y_{1i}}) + \frac{\lambda_0}{\lambda} g_2(Ee^{-(z_3+z_4)Y_{0i}}) + \frac{\lambda_2}{\lambda} g_2(Ee^{-z_4 Y_{2i}}) \right\}^n P(N(t)=n) \\ &= e^{-\lambda t \left\{1 - \left[\frac{\lambda_1}{\lambda} g_1(Ee^{-z_3 Y_{11}}) + \frac{\lambda_0}{\lambda} g_0[Ee^{-(z_3+z_4)Y_{01}}] + \frac{\lambda_2}{\lambda} g_2(Ee^{-z_4 Y_{21}})\right]\right\}}. \end{aligned}$$

which is exactly $Ee^{-z_1 S_1(t) - z_2 S_2(t)}$. In this way the proof is completed.

Now we are ready to say that the risk process R, defined in (5) is a C-L risk process.

Theorem 3. The risk process, described in (5) is stochastically equivalent to the process

$$R(t) = u + ct - \sum_{n=1}^{N(t)} \left[I_{1n} \sum_{i=1}^{\eta_{1n}} Y_{1i} + 2 I_{0n} \sum_{i=1}^{\eta_{0n}} Y_{0i} + I_{2n} \sum_{i=1}^{\eta_{2n}} Y_{2i} \right], \quad t \geq 0.$$

The proof of Theorem 3 follows immediately from the definition (5) of the risk process R, the definitions of S, S_1 and S_2, (see 4, and 5) and Theorem 2.

4 The characteristics of the risk model R

In this section using the well known results about the C-L risk model we obtain the numerical characteristics, conditional distributions and probabilities for ruin of the risk model (5).

Denote by

$\tau(u) = inf\{t > 0 : R(t) < 0 | R(0) = u\}, u \geq 0, inf = \infty$, the time of ruin with initial capital $u \geq 0$;

$\psi(u) = P(\tau(u) < \infty | R(0) = u)$, the probability for ruin in infinite horizon;

$\delta(u) = 1 - \psi(u)$, the probability for survival in infinite horizon;

$\lambda = \lambda_1 + \lambda_2 + \lambda_3$ is the intensity of the merged $N(t) \sim HPP(\lambda)$. Then it is the parameter of the exponential distribution, which models the inter-arrival times in the merged risk process presentation.

The mixed claim size in the merged risk process presentation are

$$Y_s := I_{1s} \sum_{i=1}^{\eta_{1s}} Y_{1i} + 2 I_{0s} \sum_{i=1}^{\eta_{0s}} Y_{0i} + I_{2s} \sum_{i=1}^{\eta_{2s}} Y_{2i}, \quad s = 1, 2, ...$$

Note that these r.vs. are i.i.d.

In the next theorem using the results for the Cramer-Lundberg model (see e.g. Grandell [6], Gerber [5], Embrechts, Klueppelberg, Mikosch [4] Rolski, Schmidli, Schmidt, Teugels [15]) we characterise the risk process R.

Theorem 4. Consider the risk model R, defined in (5). Denote by

$$\mu = \lambda_1 E(\eta_{11}) E(Y_{11}) + 2\lambda_0 E(\eta_{01}) E(Y_{01}) + \lambda_2 E(\eta_{21}) E(Y_{21}).$$

Then

i) In case when the expectations are finite, the safety loading is $\rho = \frac{c}{\mu} - 1$ and the net profit condition is $c > \mu$.

ii) In case when the expectations are finite and the net profit is satisfied $\psi(0) = \frac{\mu}{c}$.

iii) The general formula for the Laplace - Stieltjes transform of $\delta(u)$ is

$$l_\delta(s) = \frac{s(c-\mu)}{cs - \lambda - \lambda_1 g_1(Ee^{-sY_{11}}) - \lambda_2 g_2(Ee^{-sY_{21}}) - \lambda_0 g_0(Ee^{-2sY_{01}})}.$$

iv) The distribution of the deficit at the time of ruin, given that ruin with initial capital zero happen has the following distribution

$\mathbb{P}(-R(\tau(0)+) \leq x | \tau(0) < \infty) =$

$$= \sum_{k=1}^{2} \left\{ \frac{\lambda_k}{\mu} \int_0^x P\left(\sum_{i=1}^{\eta_{k1}} Y_{k1} \geq y \right) dy \right\} + \frac{\lambda_0}{\mu} \int_0^x P\left(\sum_{i=1}^{\eta_{01}} Y_{01} \geq \frac{y}{2} \right) dy.$$

v) Given the second moments exist,

$\mathbb{E}(-R(\tau(0)+) | \tau(0) < \infty) =$

$$= \frac{1}{\mu} \left\{ \sum_{k=1}^{2} \lambda_k \left[E(\eta_{k1}) D(Y_{k1}) + E(\eta_{k1}^2)[E(Y_{k1})]^2 \right] + 4\lambda_0 \left[E(\eta_{01}) D(Y_{01}) + E(\eta_{01}^2)[E(Y_{01})]^2 \right] \right\}.$$

$\mathbb{E}(\tau(0) | \tau(0) < \infty) =$

$$= \frac{\sum_{k=1}^{2} \lambda_k \left[E(\eta_{k1}) D(Y_{k1}) + E(\eta_{k1}^2)[E(Y_{k1})]^2 \right] + 4\lambda_0 \left[E(\eta_{01}) D(Y_{01}) + E(\eta_{01}^2)[E(Y_{01})]^2 \right]}{\mu(c-\mu)}.$$

vi) The joint distribution of the severity of (deficit at) ruin and the risk surplus just before the ruin with initial capital zero is:

$\mathbb{P}(-R(\tau(0)+) > x, R(\tau(0)-) > y | \tau(0) < \infty) =$

$$= \sum_{k=1}^{2} \left\{ \frac{\lambda_k}{\mu} \int_0^{x+y} P\left(\sum_{i=1}^{\eta_{k1}} Y_{k1} \geq y \right) dy \right\} + \frac{\lambda_0}{\mu} \int_0^{x+y} P\left(\sum_{i=1}^{\eta_{01}} Y_{01} \geq \frac{y}{2} \right) dy$$

vii) The distribution of the claim causing ruin is

$\mathbb{P}(R(\tau(0)-) - R(\tau(0)+) \leq x | \tau(0) < \infty) =$

$$= \sum_{s=1}^{2} \frac{\lambda_s}{\mu} \int_0^x y dP(\sum_{i=0}^{\eta_{s1}} Y_{s1} < y) + \frac{2\lambda_0}{\mu} \int_0^{\frac{x}{2}} y dP(\sum_{i=0}^{\eta_{01}} Y_{01} < y).$$

viii) If the distribution of Y_1 is domilatetly varying which means that

$$\limsup_{x\to\infty} \frac{\lambda_1 P(\sum_{i=0}^{\eta_{11}} Y_{11} < \frac{x}{2}) + \lambda_2 P(\sum_{i=0}^{\eta_{21}} Y_{21} < \frac{x}{2}) + \lambda_0 P(\sum_{i=0}^{\eta_{01}} Y_{01} < \frac{x}{4})}{\lambda_1 P(\sum_{i=0}^{\eta_{11}} Y_{11} < x) + \lambda_2 P(\sum_{i=0}^{\eta_{21}} Y_{21} < x) + \lambda_0 P(\sum_{i=0}^{\eta_{01}} Y_{01} < \frac{x}{2})} < \infty,$$

then

$$\lim_{u\to\infty} \frac{\psi(u)}{\int_0^u \left[\sum_{k=1}^{2} \lambda_k P\left(\sum_{i=1}^{\eta_{k1}} Y_{k1} \geq y \right) \right] + \lambda_0 P\left(\sum_{i=1}^{\eta_{01}} Y_{01} \geq \frac{y}{2} \right) dy} = \frac{1}{c-\mu}.$$

ix) For this model, the small claim condition means that there exists the Cramer-Lundberg exponent ε which is the smallest possible solution of the equation

$$\lambda_1[1-g_1(Ee^{\varepsilon Y_{11}})]+\lambda_0[1-g_0(Ee^{2\varepsilon Y_{01}})]+\lambda_2[1-g_2(Ee^{\varepsilon Y_{21}})]=c\varepsilon.$$

In that case $\psi(u) \le e^{-\varepsilon u}$ and choosing $\alpha \in (0,1)$ we can find appropriate premium income rate c, in such a way that $\psi(u) \le \alpha$.

If additionally

$$\sum_{k=1}^{2}\left\{\frac{\lambda_k}{\mu}\int_0^\infty xe^{\varepsilon x}P\left(\sum_{i=1}^{\eta_{k1}}Y_{k1}\ge x\right)dx\right\}+\frac{\lambda_0}{\mu}\int_0^\infty xe^{\varepsilon x}P\left(\sum_{i=1}^{\eta_{01}}Y_{01}\ge\frac{x}{2}\right)dx<\infty,$$

then the following Cramer-Lundberg approximation of the probability of ruin holds

$$\lim_{u\to\infty}e^{\varepsilon u}\psi(u)=\frac{c-\mu}{\varepsilon\left\{\sum_{k=1}^{2}[\lambda_k\int_0^\infty xe^{\varepsilon x}P\left(\sum_{i=1}^{\eta_{k1}}Y_{k1}\ge x\right)dx]+\lambda_0\int_0^\infty xe^{\varepsilon x}P\left(\sum_{i=1}^{\eta_{01}}Y_{01}\ge\frac{x}{2}\right)dx\right\}}.$$

Scetch of the proof: The Total probability formula entails that the c.d.f. of Y_1 is

$$F_{Y_1}(x)=\frac{\lambda_1}{\lambda}P(\sum_{i=0}^{\eta_{11}}Y_{11}<x)+\frac{\lambda_2}{\lambda}P(\sum_{i=0}^{\eta_{21}}Y_{21}<x)+\frac{\lambda_0}{\lambda}P(\sum_{i=0}^{\eta_{01}}Y_{01}<\frac{x}{2}) \tag{13}$$

Using the double expectation formula we obtain their mean and LST

$$EY_1=\frac{\lambda_1}{\lambda}E(\eta_{1s})E(Y_{11})+\frac{2\lambda_0}{\lambda}E(\eta_{0s})E(Y_{01})+\frac{\lambda_2}{\lambda}E(\eta_{2s})E(Y_{21})=:\frac{\mu}{\lambda},$$

$$Ee^{-sY_1}=\frac{\lambda_1}{\lambda}g_1(Ee^{-sY_{11}})+\frac{\lambda_2}{\lambda}g_2(Ee^{-sY_{21}})+\frac{\lambda_0}{\lambda}g_0(Ee^{-2sY_{01}}).$$

In order to complete the proof we just replace these expressions in the well known formulae for the C-L model.

5 Conclusive remarks

Due to their feasibility risk models with dependent classes of business have recently attracted much attention is scientific literature. Although in general multivariate case with common shocks the formulae for their characteristics are huge, here using the bivariate model we show that these models can be reduced to the C-L case. This allows us to use already known results which reasonably simplifies the calculations. In analogous way we can reduce similar multivariate (with more then two classes of business) risk models with arbitrary common shocks to C-L model. Many particular cases can investigated. For example the numbers of claims in different groups can be of any discrete integer probability type. η_{01}, η_{11}, η_{21} can be constants, Poisson, Geometric, Negative binomial, Truncated geometric, Logarithmic, Binomial ets. We mention these, because there are explicit formulae and Panjer's type recursions (see [15]) for their compounds. These three r. vs. are not obligatory identically distributed. In case when the number of summands is governed by some distribution which has no investigated compound then higher-order asymptotic expansions presented e.g. in [1] and the references there in can be very useful. The experience of the author shows that the Monte-Carlo method is flexible, easy to implement, not so time consuming, and gives good results. See e.g. [7].

REFERENCES:

[1] Albrecher, H., Hipp, Ch., and Kortschak, D. Higher-order expansions for compound distributions and ruin probabilities with subexponential claims. Scandinavian Actuarial Journal, **2010** (2) (2010), 105–135.

[2] Cossette, H., Marceau, H., The discrete-time risk model with correlated classes of business. Insurance: Mathematics and Economics, **26** (2) (2000), 133–149.

[3] Cramer, H., On the mathematical theory of risk, Skandia Jubilee Volume, Centraltryckeriet, Stockholm, Sweden, (1930).

[4] Embrechts, Paul, Klueppelberg, Cl., Mikosch, Th., Modelling extremal events: for insurance and finance, Springer, (1997).

[5] Gerber, Hans U., An introduction to mathematical risk theory, Huebner Foundation, (1979).

[6] Grandell, Jan, Aspects of risk theory, Springer-Verlag, New York (1991).

[7] Jordanova, P., Petkova M., Stehlik M., Compound power series distribution with negative multinomial siummands: characterisation amd risk process, accepted in RevStat (2017).

[8] Jordanova, P.K., Stehlik, M., On multivariate modifications of Cramer-Lundberg risk model with constant intensities. Accepted in Stochastic Analysis and Applications, (2018). https://doi.org/10.1080/07362994.2018.1471403

[9] Lundberg, F., Some supplementary researches on the collective risk theory, Scandinavian Actuarial Journal, Taylor and Francis, **15** (3) (1932), 137—158.

[10] Marshall, A.W., Olkin, I., A multivariate exponential distribution. Journal of the American Statistical Association, **62** (1967), 30–44.

[11] Marshall, A.W., Olkin, I., Families of multivariate distributions. Journal of the American Statistical Association, **83** (1988), 834–841.

[12] Hesselager, O., Recursions for certain bivariate counting distributions and their compound distributions. ASTIN Bulletin, **26** (1996), 35–52.

[13] Panjer, H.H., Recursive evaluation of a famfly ol compound distrlbutions, ASTIN Bulletin, **12**, (1992), 22–26.

[14] Panjer, H.H., Willmot, G. E. Recursions for compound distributions. ASTIN Bulletin: The Journal of the IAA **13** (1) (1982), 1–12.

[15] Rolski, T., Schmidli, H., Schmidt, V., Teugels, J., Stochastic processes for insurance and finance, John Wiley and Sons, **505**, (2009).

[16] Sundt, B., On some extentions of Panjer's, class of counting dlstrlbtmons. ASTIN Bulletin, **22**, (1992), 61–80.

[17] Sundt, B., Vernic, R., Recursions for convolutions and compound distributions with insurance applications. Springer Science and Business Media, (2009).

[18] Vernic, Raluca. On the Evaluation of the Distribution of a General Multivariate Collective Model: Recursions versus Fast Fourier Transform. Risks **6**(3) (2018), 87, 1–14.

[19] Wang, S., Aggregation of correlated risk portfolios: models and algorithms, Proceedings of the Casualty Actuarial society, **85** (163) (1998), 848–939.

Pavlina Jordanova
Faculty of Mathematics and Informatics, Konstantin Preslavski University,
Universitetska str. 115, 9700 Shumen, Bulgaria
E-mail: p.jordanova@shu.bg

A VISUALIZATION AND A SHAPE CHARACTERISATION OF A CLASS OF CYLINDRICAL HELICES *

GEORGI H. GEORGIEV, RADOSTINA P. ENCHEVA, CVETELINA L. DINKOVA

ABSTRACT: *In this paper we investigate parameterizations of cylindrical helices generated by unit speed plane curves. We also examine properties of their focal curves. Illustrative examples of cylindrical helices with periodical curvature and torsion are presented.*

KEYWORDS: *Cylindrical Helix, Curvature, Torsion, Focal Curve*

1 Introduction

Constructing a new curve from a given regular curve is an important task in the classical differential geometry. In this paper we examine cylindrical helices generated by unit speed plane curves of class C^3 and their focal curves. Any unit speed plane curve is uniquely determined up to Euclidean motion of $\mathbb{E}^2$ by its curvature function K. Every plane curve with a non-zero curvature is fully determined up to direct similarity by another function $\widetilde{K}$ called a shape curvature. For any regular parameterized curve $\boldsymbol{\alpha} = \boldsymbol{\alpha}(t) : I \longrightarrow \mathbb{E}^2$ the shape curvature function $\widetilde{K}(t)$ can be expressed by $\widetilde{K}(t) = \dfrac{d}{dt}\left(\dfrac{1}{K(t)}\right) / \left\| \dfrac{d\boldsymbol{\alpha}}{dt} \right\|$. Analogously, any unit speed space curve is uniquely determined up to Euclidean motion of $\mathbb{E}^3$ by its curvature $\kappa_1 \geq 0$ and its torsion κ_2. If $\kappa_1 > 0$ this curve is fully determined up to direct similarity by other functions $\widetilde{\kappa}_1$ and $\widetilde{\kappa}_2$ called a shape curvature and a shape torsion, respectively. Any regular parameterized space curve $\boldsymbol{\gamma} = \boldsymbol{\gamma}(t) : I \longrightarrow \mathbb{E}^3$ has a shape curvature and a shape torsion defined by functions

$$\text{(1)} \qquad \widetilde{\kappa}_1(t) = \frac{\frac{d}{dt}\left(\frac{1}{\kappa_1(t)}\right)}{\left\|\frac{d\boldsymbol{\gamma}(t)}{dt}\right\|} \quad \text{and} \quad \widetilde{\kappa}_2(t) = \frac{\kappa_2(t)}{\kappa_1(t)}$$

The geometry of curves with respect to an orientation-preserving Euclidean motion is completely presented in [6]. For instance, one important theorem proved in this book is

Theorem 1.1. *[6, p.137](Fundamental Theorem of Plane Curves) A unit-speed curve $\boldsymbol{\alpha} : I \to \mathbb{E}^2$ whose curvature is given piecewise-continuous function $K : I \to \mathbb{R}$ is parameterized by*

$$\text{(2)} \qquad \begin{cases} \boldsymbol{\alpha}(s) = \left(\int \cos\theta(s)ds + d_1, \int \sin\theta(s)ds + d_2\right), \\ \theta(s) = \int K(s)ds + \theta_0, \end{cases}$$

where d_1, d_2, θ_0 are constants of integration.

Elements of the theory of curves related to direct similarities are given in [2], [3] and [4]. There are special space curves called cylindrical helices with the property $\dfrac{\kappa_2(t)}{\kappa_1(t)} = const$. These curves are studied in [7] and [8]. There exists a well-known construction for obtaining a new space curve from a given space curve.

*Partially supported by Scientific Research Grant RD–08–111/05.02.2018 of Shumen University.

Definition 1.1. *[6, p.241] A **focal curve** (or an evolute) of a regular C^3-space curve $\boldsymbol{\gamma}: (a,b) \to \mathbb{E}^3$ with a nonzero curvature and a torsion is the curve given by*

$$\mathbf{C}_\gamma(t) = \boldsymbol{\gamma}(t) + c_1(t)\mathbf{n_1}(t) + c_2(t)\mathbf{n_2}(t), \tag{3}$$

where $\mathbf{n_1}$ is a principal unit normal vector field of $\boldsymbol{\gamma}$, $\mathbf{n_2}$ is a binormal unit vector field of $\boldsymbol{\gamma}$. The coefficients $c_1(t)$ and $c_2(t)$ are smooth functions called focal curvatures of $\boldsymbol{\gamma}$, given by

$$c_1(t) = \frac{1}{\kappa_1(t)}, \quad c_2(t) = -\frac{\frac{d}{dt}\kappa_1(t)}{\parallel \frac{d\boldsymbol{\gamma}(t)}{dt} \parallel \kappa_1(t)^2\kappa_2(t)} = \frac{\frac{dc_1(t)}{dt}}{\parallel \frac{d\boldsymbol{\gamma}(t)}{dt} \parallel \kappa_2(t)}, \tag{4}$$

where $\kappa_1(t)$ and $\kappa_2(t)$ are the curvature and the torsion of $\boldsymbol{\gamma}$.

In this paper we describe a construction which is divided into three steps. First, we obtain a cylindrical helix from a given unit speed curve in the *xy*-plane. Second, we determine the focal curve of the obtained cylindrical helix. Third, we study the orthogonal projection of the focal curve on the *xy*-plane. This scheme applied to plane curves with periodic signed curvature is presented.

2 Parameterizations of cylindrical helices

Let $\boldsymbol{\alpha} = \boldsymbol{\alpha}(s)$, $s \in I \subseteq \mathbb{R}$ be a unit speed regular C^3-curve in the Euclidean plane $\mathbb{E}^2 \equiv O\vec{\mathbf{e}}_1\vec{\mathbf{e}}_2$ with a vector parametric equation $\boldsymbol{\alpha}(s) = (x(s), y(s), 0)$ parameterized by an arc-length parameter $s \in I$. The signed curvature of $\boldsymbol{\alpha}$ is

$$K = <\boldsymbol{\alpha}'', J\boldsymbol{\alpha}'> = x'(s).y''(s) - x''(s).y'(s), \tag{5}$$

where J is the complex structure of $\mathbb{R}^2$ and "$< \cdot, \cdot >$" denotes the scalar (or dot) product of two vectors. We define a new space curve $\boldsymbol{\gamma} = \boldsymbol{\gamma}(s)$, $s \in I \subseteq \mathbb{R}$ in the Euclidean space $\mathbb{E}^3 \equiv O\vec{\mathbf{e}}_1\vec{\mathbf{e}}_2\vec{\mathbf{e}}_3$ with a parametric representation

$$\boldsymbol{\gamma}(s) = (x(s), y(s), as+b) = \boldsymbol{\alpha}(s) + (as+b)\vec{\mathbf{e}}_3, \ a, b = const \tag{6}$$

This curve lies on the generalized cylinder S_α with a base curve $\boldsymbol{\alpha}(s)$ and rulings parallel to *z*-axis. The curvature and the torsion of $\boldsymbol{\gamma}$ are expressed by the next lemma.

Lemma 2.1. *[5, p.5640] Let $\boldsymbol{\gamma}$ be a space curve with a parametrization* (6). *Then the Euclidean curvature κ_1 and the Euclidean torsion κ_2 of $\boldsymbol{\gamma}$ are given by*

$$\kappa_1 = \frac{|K|}{1+a^2}, \quad \kappa_2 = \frac{a.K}{1+a^2}, \tag{7}$$

where K is the signed curvature of $\boldsymbol{\alpha}$.

Observe that for the curve $\boldsymbol{\gamma}$ the condition characterizing cylindrical helixes $\kappa_2/\kappa_1 = const$ holds. More precisely, $\boldsymbol{\gamma}$ is a cylindrical helix with a constant curvature ratio $\kappa_2/\kappa_1 = \varepsilon a$, where $\varepsilon = sign(K)$. In what follows we will call ***$\boldsymbol{\gamma}$ a cylindrical helix generated by the plane curve $\boldsymbol{\alpha}$***. Moreover, this cylindrical helix is a geodesic curve on S_α. Analogous condition for the focal curvatures c_1 and c_2 of $\boldsymbol{\gamma}$ can be written in the form $c_1.c_1' - \varepsilon a\sqrt{1+a^2}c_2 \equiv 0$. The next statement gives a relation between focal curvatures of $\boldsymbol{\gamma}$ and the signed curvature of $\boldsymbol{\alpha}$.

Theorem 2.2. *Let $\boldsymbol{\alpha} = \boldsymbol{\alpha}(s)$, $s \in I$ be a unit-speed C^3-plane curve in $\mathbb{E}^2$ with a signed curvature $K \neq 0$. Let $\boldsymbol{\gamma}(s)$ be a space curve given by* (6)*. If c_1 and c_2 are the focal curvatures of $\boldsymbol{\gamma}$, then*

$$c_1 = \frac{1+a^2}{|K|} \text{ and } c_2 = -\frac{\sqrt{1+a^2}^{\,3}.K'}{a.|K|^3}, \tag{8}$$

where a is the constant slope of $\boldsymbol{\gamma}$ with respect to $\vec{\mathbf{e}}_3$.

Proof. From (4) and (7) we get

$$c_1(s) = \frac{1}{\kappa_1} = \frac{1+a^2}{|K|} \quad \text{and} \quad c_2(s) = -\frac{\frac{d}{ds}\kappa_1(s)}{\| \frac{d\boldsymbol{\gamma}(s)}{ds} \| \, \kappa_1(s)^2 \kappa_2(s)} = -\frac{\sqrt{1+a^2}^{3}.K'}{a.|K|^3}.$$

□

Now, we give a useful relation between the shape curvature and the shape torsion of the cylindrical helix and the shape curvature of the generating plane curve.

Corollary 2.2.1. *Relations between the shape curvature $\widetilde{\kappa}_1$ and the shape torsion $\widetilde{\kappa}_2$ of the cylindrical helix $\boldsymbol{\gamma}$ generated by a unit speed C^3-plane curve $\boldsymbol{\alpha}$ and the shape curvature $\widetilde{K}$ of $\boldsymbol{\alpha}$ are*

$$\widetilde{\kappa}_1 = \sqrt{1+a^2}|\widetilde{K}| \quad \text{and} \quad \widetilde{\kappa}_2 = \varepsilon a.$$

Proof. In [2, p.285] it was shown that the shape curvature of $\boldsymbol{\alpha}$ is $\widetilde{K} = (1/K)'$. From equations (7) and (1) we conclude that the shape curvature and the shape torsion of $\boldsymbol{\gamma}$ are given by

$$\widetilde{\kappa}_1 = \frac{\frac{d}{ds}(\frac{1}{\kappa_1})}{\|\frac{d\gamma}{ds}\|} = \frac{1}{\sqrt{1+a^2}}\frac{d}{ds}\left(\frac{1+a^2}{|K|}\right) = \sqrt{1+a^2}\left(\frac{1}{|K|}\right)' = \sqrt{1+a^2}|\widetilde{K}| \text{ and } \widetilde{\kappa}_2 = \frac{\kappa_2}{\kappa_1} = \varepsilon a.$$

□

Corollary 2.2.2. *A relation between the focal curvatures of the cylindrical helix $\boldsymbol{\gamma}$ with a parametric equation* (6) *and the shape curvature of the base unit-speed C^3-plane curve $\boldsymbol{\alpha}$ is*

$$\frac{c_2}{c_1} = \frac{\sqrt{1+a^2}}{a}\widetilde{K}. \tag{9}$$

Proof. From (8) the ratio of the focal curvatures is

$$\frac{c_2}{c_1} = -\frac{\sqrt{1+a^2}}{a}\frac{K'}{K^2} = \frac{\sqrt{1+a^2}}{a}\left(\frac{1}{K}\right)' = \frac{\sqrt{1+a^2}}{a}\widetilde{K}.$$

□

3 Focal curves of cylindrical helices generated by unit speed plane curves

Theorem 3.1. *Let $\boldsymbol{\alpha} = \boldsymbol{\alpha}(s)$, $s \in I$ be a unit-speed C^3-plane curve in $\mathbb{E}^2$ and let $K \neq 0$ be the signed curvature of $\boldsymbol{\alpha}$. Suppose that $\boldsymbol{\gamma}(s)$ is the cylindrical helix defined by* (6)*. Then the focal curve $\mathbf{C}_\gamma$ of $\boldsymbol{\gamma}$ possesses a parametrization*

$$\mathbf{C}_\gamma(s) = \boldsymbol{\alpha}(s) + \frac{(1+a^2)K'}{K^3}\mathbf{T} + \frac{1+a^2}{K}\mathbf{N} + \frac{(as+b)aK^3-(1+a^2)K'}{aK^3}\vec{\mathbf{e}}_3, \tag{10}$$

where $\mathbf{T}$ and $\mathbf{N}$ form the Frenet frame of $\boldsymbol{\alpha}$.

Proof. First, we find the Frenet frame $\mathbf{t}$, $\mathbf{n_1}$, $\mathbf{n_2}$ of $\boldsymbol{\gamma}$. From equation (6) and the structure equations of $\boldsymbol{\alpha}$ we obtain the cross vector products

$$\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2} = (\boldsymbol{\alpha}'+a.\vec{\mathbf{e}}_3)\times\boldsymbol{\alpha}'' = (\boldsymbol{\alpha}'+a.\vec{\mathbf{e}}_3)\times K.J\boldsymbol{\alpha}' = K(\boldsymbol{\alpha}'\times J\boldsymbol{\alpha}'+a.\vec{\mathbf{e}}_3\times J\boldsymbol{\alpha}') = K(\vec{\mathbf{e}}_3-a.\boldsymbol{\alpha}')$$

$$\left(\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}\right)\times\frac{d\boldsymbol{\gamma}}{ds} = K(\vec{\mathbf{e}}_3-a.\boldsymbol{\alpha}')\times(\boldsymbol{\alpha}'+a.\vec{\mathbf{e}}_3) = K(1+a^2)\vec{\mathbf{e}}_3\times\boldsymbol{\alpha}' = K(1+a^2)J\boldsymbol{\alpha}'.$$

Their norms are $\left\|\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}\right\| = |K|\sqrt{1+a^2}$ and $\left\|\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}\right\|.\left\|\frac{d\boldsymbol{\gamma}}{ds}\right\| = (1+a^2)|K|$. From the Frenet formulas with respect to an arbitrary parameter we get the unit principal normal vector

$$\mathbf{n_1} = \frac{\left(\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}\right)\times\frac{d\boldsymbol{\gamma}}{ds}}{\left\|\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}\right\|.\left\|\frac{d\boldsymbol{\gamma}}{ds}\right\|} = \frac{K(1+a^2)J\boldsymbol{\alpha}'}{(1+a^2)|K|} = \varepsilon J\boldsymbol{\alpha}', \tag{11}$$

and the unit binormal vector

$$\mathbf{n_2} = \frac{\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}}{\left\|\frac{d\boldsymbol{\gamma}}{ds} \times \frac{d^2\boldsymbol{\gamma}}{ds^2}\right\|} = \frac{K(\vec{\mathbf{e}}_3-a.\boldsymbol{\alpha}')}{|K|\sqrt{1+a^2}} = \frac{\varepsilon(\vec{\mathbf{e}}_3-a.\boldsymbol{\alpha}')}{\sqrt{1+a^2}} \tag{12}$$

Having in mind $\mathbf{T} = \boldsymbol{\alpha}'$ and $\mathbf{N} = J\boldsymbol{\alpha}'$ and using (3), (8), (11) and (12) we obtain (10). □

Definition 3.1. *The focal curve $\mathbf{C}_\gamma$ of $\boldsymbol{\gamma}$ with parametrization* (10) *is called* ***an associated space curve of*** *a unit-speed plane curve $\boldsymbol{\alpha}$ with a nonzero signed curvature.*

The curvature and the torsion of the associated space curve can be expressed in terms of the signed curvature and the shape curvature of $\boldsymbol{\alpha}$.

Theorem 3.2. *The curvature K_1 and the torsion K_2 of the associated curve $\mathbf{C}_\gamma$ of the unit speed C^3-plane curve $\boldsymbol{\alpha}$ with a signed curvature $K \neq 0$ and a shape curvature $\widetilde{K}$ are given by*

$$K_1 = \frac{a^2|K|}{(1+a^2)\left|a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right|}, K_2 = \frac{aK}{(1+a^2)\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)}. \tag{13}$$

Proof. Equation (10) can be written in the form

$$\mathbf{C}_\gamma(s)=\boldsymbol{\alpha}(s)-\frac{(1+a^2)\widetilde{K}}{K}\mathbf{T}+\frac{1+a^2}{K}\mathbf{N}+\left((as+b)+\frac{(1+a^2)\widetilde{K}}{aK}\right)\vec{\mathbf{e}}_3. \tag{14}$$

Then, its derivatives with respect to the arc-length parameter s of $\boldsymbol{\alpha}$ are

$$\frac{d\mathbf{C}_\gamma}{ds}=\left(a+\frac{1+a^2}{a}\left(\frac{\widetilde{K}}{K}\right)'\right)(\vec{\mathbf{e}}_3-a\mathbf{T})$$

$$\frac{d^2\mathbf{C}_\gamma}{ds^2}=\left(a+\frac{1+a^2}{a}\left(\frac{\widetilde{K}}{K}\right)'\right)(-aK.\mathbf{N})+\frac{1+a^2}{a}\left(\frac{\widetilde{K}}{K}\right)''(\vec{\mathbf{e}}_3-a\mathbf{T})$$

$$\frac{d^3\mathbf{C}_\gamma}{ds^3}=\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)(K^2.\mathbf{T}-K'\mathbf{N})-2(1+a^2)\left(\frac{\widetilde{K}}{K}\right)''K.\mathbf{N}+\frac{1+a^2}{a}\left(\frac{\widetilde{K}}{K}\right)'''(\vec{\mathbf{e}}_3-a\mathbf{T}).$$

The cross vector product and the triple scalar product are

$$\begin{aligned}&\frac{d\mathbf{C}_\gamma}{ds}\times\frac{d^2\mathbf{C}_\gamma}{ds^2}=\frac{K}{a}\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)^2(\mathbf{T}+a\vec{\mathbf{e}}_3),\\&\frac{d\mathbf{C}_\gamma}{ds}\frac{d^2\mathbf{C}_\gamma}{ds^2}\frac{d^3\mathbf{C}_\gamma}{ds^3}=<\frac{d\mathbf{C}_\gamma}{ds}\times\frac{d^2\mathbf{C}_\gamma}{ds^2},\frac{d^3\mathbf{C}_\gamma}{ds^3}>=\frac{K^3}{a}\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)^3\end{aligned} \tag{15}$$

and the norms of the first derivative and the cross vector product are

$$\left\|\frac{dC_\gamma}{ds}\right\|=\frac{\sqrt{1+a^2}}{a}\left|a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right|,\ \left\|\frac{d\mathbf{C}_\gamma}{ds}\times\frac{d^2\mathbf{C}_\gamma}{ds^2}\right\|=\frac{|K|}{a}\sqrt{1+a^2}\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)^2. \tag{16}$$

From $K_1=\dfrac{\left\|\frac{d\mathbf{C}_\gamma}{ds}\times\frac{d^2\mathbf{C}_\gamma}{ds^2}\right\|}{\left\|\frac{d\mathbf{C}_\gamma}{ds}\right\|^3}$, $K_2=\dfrac{\frac{d\mathbf{C}_\gamma}{ds}\frac{d^2\mathbf{C}_\gamma}{ds^2}\frac{d^3\mathbf{C}_\gamma}{ds^3}}{\left\|\frac{d\mathbf{C}_\gamma}{ds}\times\frac{d^2\mathbf{C}_\gamma}{ds^2}\right\|^2}$ and equations (15), (16) we get (13). □

Corollary 3.2.1. *The associated curve $\mathbf{C}_\gamma$ of the unit speed C^3-plane curve $\boldsymbol{\alpha}$ is a cylindrical helix.*

Proof. From (13) we have

$$\frac{K_2}{K_1}=\frac{\varepsilon\delta}{a}=const,$$

where $\varepsilon=sign(K),\delta=sign\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)$. Hence the associated curve is a cylindrical helix. □

Theorem 3.3. *The shape curvature $\widetilde{K}_1$ and the shape torsion $\widetilde{K}_2$ of the associated curve $\mathbf{C}_\gamma$ of the unit speed C^3-plane curve $\boldsymbol{\alpha}$ with a signed curvature $K\neq 0$ and a shape curvature $\widetilde{K}$ are given by*

$$\widetilde{K}_1=\frac{\sqrt{1+a^2}}{a}\left|\widetilde{K}+\frac{\frac{(1+a^2)}{K}\left(\frac{\widetilde{K}}{K}\right)''}{a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'}\right|\quad and\quad \widetilde{K}_2=\frac{\delta}{\varepsilon a}, \tag{17}$$

where $\varepsilon=sign(K)$ and $\delta=sign\left(a^2+(1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right)$.

Proof. From (13) we see that the first focal curvature of $\mathbf{C}_\gamma$ is equal to

$$C_1 = \frac{1}{K_1} = \frac{1+a^2}{a^2|K|}\left|a^2 + (1+a^2)\left(\frac{\widetilde{K}}{K}\right)'\right|, \tag{18}$$

where the shape curvature $\widetilde{K} = (1/K)'$ of $\boldsymbol{\alpha}$ and $(\widetilde{K}/K)' = (3K'^2 - KK'')/K^4$ are smooth functions depending on the arc-length parameter s of α. After differentiation of (18) with respect to s and some calculations we get

$$\frac{dC_1(s)}{ds} = \frac{1+a^2}{a^2}\left|\widetilde{K}\left((1+a^2)\left(\frac{\widetilde{K}}{K}\right)' + a^2\right) + \frac{1+a^2}{K}\left(\frac{\widetilde{K}}{K}\right)''\right| \tag{19}$$

From $\widetilde{K}_1 = \dfrac{dC_1(s)}{ds} / \left\|\dfrac{dC_\gamma}{ds}\right\|$, $\widetilde{K}_2 = K_2/K_1$ and (19), (16) we obtain (17). □

Definition 3.2. *The orthogonal projection of the associated curve $\mathbf{C}_\gamma$ on the plane $\mathbb{E}^2 = Oxy$ is called **a generalized focal curve** of the plane curve $\boldsymbol{\alpha}$ and denoted by $\boldsymbol{\beta}$.*

From equation (14) we deduce that the generalized focal curve of $\boldsymbol{\alpha}$ is determined by

$$\boldsymbol{\beta}(s) = \boldsymbol{\alpha}(s) - \frac{(1+a^2)\widetilde{K}}{K}\boldsymbol{\alpha}' + \frac{1+a^2}{K}J\boldsymbol{\alpha}'. \tag{20}$$

4 Applications

4.1 Generalized focal curve of cycloid

Let we consider the cycloid $\boldsymbol{\alpha}_0(t) = (c.(t+\sin t), c.(1+\cos t), 0)$, $\quad c = const > 0, t \in \mathbb{R}$. The arc-length parametrization of one hump of the cycloid $\boldsymbol{\alpha}_0$ is

$$\boldsymbol{\alpha}(s) = \left(\frac{s\sqrt{16c^2 - s^2}}{8c} + 2c\arcsin\left(\frac{s}{4c}\right), \frac{16c^2 - s^2}{8c}, 0\right), -4.c \le s \le 4.c \tag{21}$$

The natural equation of the above cycloid or the so-called Cesàro equation of that plane curve is $R^2 + s^2 = 16c^2$, where $R = 1/|K|$. Then, from (6) the corresponding cylindrical helix has a constant speed parametrization

$$\boldsymbol{\gamma}(s) = \left(\frac{s\sqrt{16c^2 - s^2}}{8c} + 2c\arcsin\left(\frac{s}{4c}\right), \frac{16c^2 - s^2}{8c}, as + b\right) \tag{22}$$

and from (10) the associated curve of $\boldsymbol{\alpha}(s)$ has constant speed parametrization

$$\mathbf{C}_\gamma(s) = \left(\frac{s\sqrt{16c^2 - s^2}}{8c} + 2c\arcsin\left(\frac{s}{4c}\right), \frac{16c^2 - s^2}{8c} - 4c(1+a^2), -\frac{s}{a} + b\right). \tag{23}$$

Finally, the generalized focal curve $\boldsymbol{\beta}$ of $\boldsymbol{\alpha}(s)$ is also a cycloid with an arc-length parametrization

$$\boldsymbol{\beta}(s) = \left(\frac{s\sqrt{16c^2 - s^2}}{8c} + 2c\arcsin\left(\frac{s}{4c}\right), \frac{16c^2 - s^2}{8c} - 4c(1+a^2), 0\right) \tag{24}$$

It is obvious that the curve $\boldsymbol{\beta}$ can be obtained from the curve $\boldsymbol{\alpha}$ via translation determined by the vector $\vec{\mathbf{p}} = (0, -4c(1+a^2), 0)$.

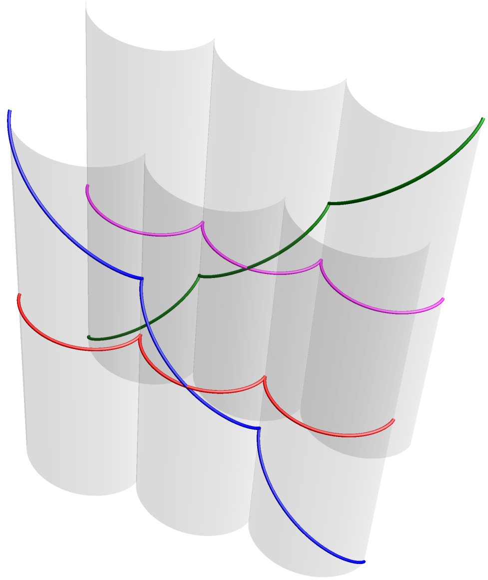

Fig. 1: Cycloid $\boldsymbol{\alpha}$ in red, generalized focal curve $\boldsymbol{\beta}$ in purple, cylindrical helix $\boldsymbol{\gamma}$ in blue, associated curve $\mathbf{C}_\gamma$ of $\boldsymbol{\alpha}$ in green

4.2 Generalized focal curve of closed plane curves with periodic signed curvature

Let $\boldsymbol{\alpha}=\boldsymbol{\alpha}(s)$ be a unit speed plane curve with a periodic curvature $K(s)=\frac{m}{n}-\sin s$, where $m\in\mathbb{Z}, n>1, n\in\mathbb{N}$ and $(m,n)=1$. Such a curve is introduced and studied in [1]. Then by (2) the function $\theta(s)$ becomes $\theta(s)=\int K(s)ds+\theta_0=\frac{\pi}{2}+\cos s+\frac{m}{n}s$ and a unit speed parametrisation of $\boldsymbol{\alpha}$ is

$$\boldsymbol{\alpha}(s)=\left(\int\cos\left(\frac{\pi}{2}+\cos s+\frac{m}{n}s\right)ds,\int\sin\left(\frac{\pi}{2}+\cos s+\frac{m}{n}s\right)ds,0\right).$$

From (6) it follows that the corresponding cylindrical helix has a constant speed parametrization

$$\boldsymbol{\gamma}(s)=\left(\int\cos\left(\frac{\pi}{2}+\cos s+\frac{m}{n}s\right)ds,\int\sin\left(\frac{\pi}{2}+\cos s+\frac{m}{n}s\right)ds,as+b\right)$$

and from (10) the associated curve of $\boldsymbol{\alpha}(s)$ has a parametrization

$$\mathbf{C}_\gamma(s)=\left(-\int\sin\left(\frac{ms}{n}+\cos s\right)ds+A(s),\int\cos\left(\frac{ms}{n}+\cos s\right)ds-B(s),as+b+C(s)\right),$$

where $A(s)=-\frac{\left(1+a^2\right)n\left(\sin\theta(s)(m-n\sin s)^2+n^2\cos s.\cos\theta(s)\right)}{(m-n\sin s)^3}$,
$B(s)=\frac{\left(1+a^2\right)n\left(n^2\cos s.\sin\theta(s)-(m-n\sin s)^2\cos\theta(s)\right)}{(m-n\sin s)^3}$ and $C(s)=\frac{\left(1+a^2\right)n^3\cos s}{a(m-n\sin s)^3}$.
Then the generalized focal curve $\boldsymbol{\beta}$ of $\boldsymbol{\alpha}(s)$ has a parametrization

$$\boldsymbol{\beta}(s)=\left(-\int\sin\left(\frac{ms}{n}+\cos s\right)ds+A(s),\int\cos\left(\frac{ms}{n}+\cos s\right)ds-B(s),0\right).$$

It is obvious that associated focal $\mathbf{C}_\gamma$ can be obtained from $\boldsymbol{\gamma}$ by matrix function $T_{\vec{\mathbf{r}}(s)}$ in homogeneous coordinates $(T_{\vec{\mathbf{r}}(s)}) = \begin{pmatrix} 1 & 0 & 0 & A(s) \\ 0 & 1 & 0 & B(s) \\ 0 & 0 & 1 & C(s) \\ 0 & 0 & 0 & 1 \end{pmatrix}$

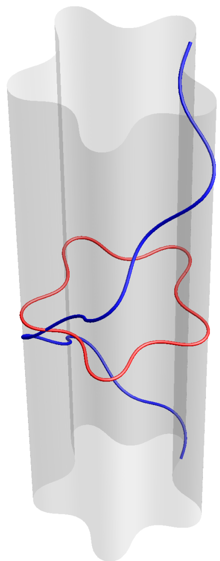

Fig. 2: Closed plane curve $\boldsymbol{\alpha}$ for $m = 1$ and $n = 5$ in red and its cylindrical helix $\boldsymbol{\gamma}$ in blue

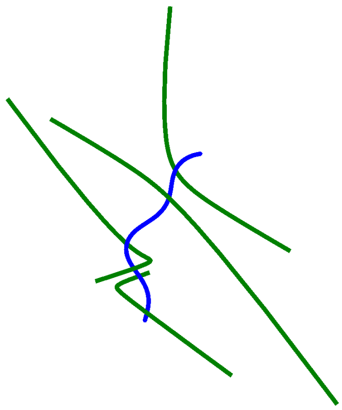

Fig. 3: Focal curve of $\boldsymbol{\gamma}$ in green

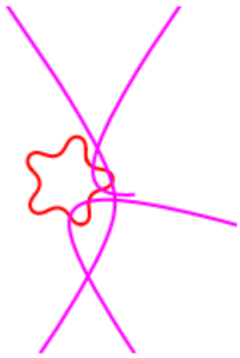

Fig. 4: Plane curve $\boldsymbol{\alpha}$ with a part of its generalized focal curve $\boldsymbol{\beta}$ in purple

REFERENCES:

[1] Arroyo, J., Garay, O.J., Menca, J.J., When is a peroidic function the curvature of a closed plane curve?, Am. Math. Mon. **115** (5), (2008), 405–414.

[2] Encheva, R.P., Georgiev, G.H., Curves on the shape spere, Result. Math. **44** (2003), 279–288.

[3] Encheva, R.P., Georgiev, G.H., Shapes of space curves, J. Geom. Graph. **7** (2003), 145–155.

[4] Encheva, R.P., Georgiev, G.H., Similar Frenet Curves, Result. Math., **55** (3-4) (2009), 359–372. http://www.springerlink.com/content/x6145691n7437r17/

[5] Georgiev, G.H., Encheva, R.P., Dinkova, Cv.L., Geometry of cylindrical curves over plane curves, Applied Mathematical Sciences,**113** (9) (2015), 5637–5649. http://dx.doi.org/10.12988/ams.2015.56456

[6] Gray, A., Abbena, E., Salamon, S., Modern Differential Geometry of Curves and Surfaces, Chapman Hall/CRC, (2006).

[7] Izumiya, S., Takeuchi, N., Generic properties of Helices and Bertrand curves, Journal of Geometry, **74** (2002), 97-109.

[8] Izumiya, S., Takeuchi, N., Special curves and ruled surfaces, Cotributions to Algebra and Geometry, **44** (2003), 203–212.

Georgi Georgiev, Radostina Encheva, Cvetelina Dinkova
Konstantin Preslavsky University of Shumen
E-mails: georgiev@shu.bg, r.encheva@shu.bg, c.dinkova@shu.bg

IMPROVED LOCAL CONVERGENCE ANALYSIS OF THE INVERSE WEIERSTRASS METHOD FOR SIMULTANEOUS APPROXIMATION OF POLYNOMIAL ZEROS*

GYURHAN H. NEDZHIBOV

ABSTRACT: *In this work we establish new local convergence result with a-priori and a-posteriori error estimates for the Inverse Weierstrass iterative method for simultaneous approximations of polynomial zeros. Our approach enlarges the convergence radius and improves the known local convergence results.*

KEYWORDS: *Polynomial zeros, Simultaneous method, Weierstrass method, Durand-Kerner method, Inverse Weierstrass method, Local convergence*

1 Introduction

Let $P(z)$ be a monic polynomial

(1) $$P(z) = a_0 + a_1 z + \ldots + a_{n-1} z^{n-1} + z^n\,,$$

of degree $n \geq 2$, with simple real or complex zeros $\alpha_1, \alpha_2, \ldots, \alpha_n$, and let $z_1^{(0)}, z_2^{(0)}, \ldots, z_n^{(0)}$ be distinct reasonable close approximations of these zeros.

In this study we consider a simultaneous iterative method defined by

(2) $$\mathbf{z}^{(k+1)} = \mathbf{G}\left(\mathbf{z}^{(k)}\right) = \mathbf{G}^{k+1}\left(\mathbf{z}^{(0)}\right)\,,\ \ k = 0, 1, 2, \ldots\,,$$

where $\mathbf{G} : \mathbf{C}^n \to \mathbf{C}^n$ is a vector valued function with components

(3) $$G_i = G_i(\mathbf{z}) = \frac{z_i^2}{z_i + W_i(\mathbf{z})}\,,\ \mathbf{z} = (z_1, \ldots, z_n)\,,\ i = 1, \ldots, n\,,$$

and the *Weierstrass' correction* $W_i : \mathscr{D} \subset \mathbf{C}^n \to \mathbf{C}$ is defined by

(4) $$W_i(\mathbf{z}) = \frac{P(z_i)}{\prod_{j \neq i}^{n} (z_i - z_j)}\,,\ \ (i = 1, \ldots, n)$$

where $\mathscr{D}$ is the set of all vectors in $\mathbf{C}^n$ with distinct components. Then we can define the operator $W : \mathscr{D} \subset \mathbf{C}^n \to \mathbf{C}^n$ by $W(z) = (W_1(z), \ldots, W_n(z))$.

The method (2)-(3) was firstly introduced in [1], and some recent results were obtained in [2, 3, 4, 5]. The obtained results in these works are modifications of local convergence results of classical Weierstrass iterative method presented in [6, 7, 8, 9, 10, 11, 12].

Throughout this paper, we will use the p-norm defined by

(5) $$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}$$

*Paper written with financial support of Shumen University under Grant RD 08-145/2018.

for some $1 \le p \le \infty$ and we will follow the usual convention that a summation over the empty set of indices equals 0, while a product over the same set equals 1. We use the function $d : \mathbf{C}^n \to \mathbf{R}_+$ defined by

$$d(x) = \min\{\delta(x), \gamma(x)\}, \tag{6}$$

where

$$\delta(x) = min_{i \neq j}|x_i - x_j| \text{ and } \gamma(x) = min_j |x_j| \quad (j = 1, \dots, n). \tag{7}$$

Further, for a number p such that $1 \le p \le \infty$ we denote by q the conjugate exponent of p, i.e. q is defined by means of

$$\frac{1}{p} + \frac{1}{q} = 1 \text{ and } 1 \le q \le \infty.$$

We study the local convergence of the *Inverse Weierstrass method* (2)-(3) with respect to the function of initial conditions $E : \mathscr{C}^n \to \mathscr{R}_+$ defined as follows

$$E(z) = \frac{\|z - \alpha\|_p}{d(\alpha)}. \tag{8}$$

In our previous work (see [4])we have proved the following convergence result.

Theorem 1.1. *Let $P \in \mathscr{C}[z]$ be a monic polynomial of degree $n \ge 2$, where $\alpha = \{\alpha \in \mathscr{C}^n : \alpha_i \neq 0$ and $\alpha_i \neq \alpha_j$ for $i, j = 1, \dots, n\}$ is the root vector of P, and let $1 \le p \le \infty$, $d = d(\alpha) = \min\{\delta, \gamma\}$, where $\delta = min_{j \neq i}|\alpha_i - \alpha_j|$ and $\gamma = min_i |\alpha_i|$ for $i, j = 1, \dots, n$ ($i \neq j$). Suppose $z^{(0)} \in \mathscr{C}^n$ is an initial guess satisfying*

$$E(z^{(0)}) = \left\| \frac{z^{(0)} - \alpha}{d(\alpha)} \right\|_p < R(n,p) = \frac{\theta^{\frac{1}{n-1}} - 1}{2^{\frac{1}{q}}(\theta^{\frac{1}{n-1}} - 1) + (n-1)^{-\frac{1}{p}}}, \tag{9}$$

where

$$\theta = \frac{3b - 2 + \sqrt{b^2 - 4b + 20}}{2(b+1)} \text{ and } b = 2^{\frac{1}{q}}. \tag{10}$$

Then the following statements hold true.
(i) CONVERGENCE. *The Inverse Weierstrass iteration (2)-(3) is well defined and converges quadratically to the root-vector α of P.*
(ii) A POSTERIORI ERROR ESTIMATE. *For all $k \ge 0$ we have the estimate*

$$\|z^{(k+1)} - \alpha\|_p \le \lambda^{2^k} \|z^{(k)} - \alpha\|_p, \tag{11}$$

(iii) A PRIORI ERROR ESTIMATE. *For all $k \ge 1$ we have the estimate*

$$\|z^{(k)} - \alpha\|_p \le \lambda^{2^k - 1} \|z^{(0)} - \alpha\|_p, \tag{12}$$

where $\lambda = E(z^{(0)})/R(n,p)$.

The main purpose of this work is to improve this theorem and obtain new local convergence result with larger radius of convergence.

2 Main Results

First, we introduce some auxiliary results. We will state following three known lemmas that we will use without proofs (the proofs may be found, e.g. in [12]).

Lemma 2.1. *Let $u \in \mathscr{C}^n$ and $1 \le p \le \infty$. Then*

$$|u_i| + |u_j| \le 2^{\frac{1}{q}} \|u\|_p \text{ for any } i, j = 1, 2, \ldots, n. \tag{13}$$

Lemma 2.2. *Let $u \in \mathscr{C}^n$ and $1 \le p \le \infty$. Then*

$$\left(\prod_{i=1}^{n} (1 + |u_i|) - 1 \right) \le \left(1 + \frac{\|u\|_p}{n^{1/p}} \right)^n - 1\,. \tag{14}$$

Lemma 2.3. *Let $n \in \mathscr{N}$, $t \ge 0$ and $0 \le \varphi \le 1$. Then*

$$(1 + \varphi t)^n - 1 \le \varphi \left((1 + t)^n - 1 \right). \tag{15}$$

Now, we will prove the following two lemmas.

Lemma 2.4. *Let $0 < t < 1/2^{1/q}$, $n \ge 2$ and*

$$\phi(t) = 1 + \frac{t}{(n-1)^{1/p}(1 - 2^{1/q} t)}\,. \tag{16}$$

Suppose that

$$\phi(t)^{n+1} < 2^{1/q}\,, \tag{17}$$

then it follows

$$\frac{1+t}{1-t} \phi(t)^{n-1} < 2\,. \tag{18}$$

Proof. It is easy to prove that

$$\tilde{t} = \frac{2^{\frac{1}{q(n+1)}} - 1}{2^{\frac{1}{q}} \left(2^{\frac{1}{q(n+1)}} - 1 \right) + (n-1)^{-\frac{1}{p}}} \tag{19}$$

is the unique root of the equation

$$\phi(t)^{n+1} = 2^{1/q}$$

in the interval $(0, 1/2^{1/q})$. Taking into account that the function $\phi(t)^{n+1}$ is continuous and strictly increasing on the interval $I = [0, \tilde{t}]$, we deduce that (17) is equivalent to

$$t < \tilde{t}\,.$$

From this and (19) it follows that

$$\frac{1+t}{1-t} < \frac{(2^{\frac{1}{q}} + 1)(2^{\frac{1}{q(n+1)}} - 1) + (n-1)^{-1/p}}{(2^{\frac{1}{q}} - 1)(2^{\frac{1}{q(n+1)}} - 1) + (n-1)^{-1/p}} \le \frac{3.2^{\frac{1}{n+1}} - 2}{2^{\frac{1}{n+1}}}\,. \tag{20}$$

From (17) it follows that

$$\phi(t)^{n-1} < 2^{\frac{n-1}{q(n+1)}}. \tag{21}$$

The last two relations (20) and (21) it follows that

$$\frac{1+t}{1-t}\phi(t)^{n-1} < \frac{3.2^{\frac{1}{n+1}}-2}{2^{\frac{1}{n+1}}}2^{\frac{n-1}{q(n+1)}} = 3.2^{\frac{n-1}{n+1}} - 2.2^{\frac{n-2}{n+1}} = \theta(n). \tag{22}$$

The function $\theta(n)$ is is continuous and strictly increasing for $n \geq 2$ and

$$\lim_{n\to\infty} \theta(n) = 2. \tag{23}$$

Now from (22) and (23) follows the statement (18). The lemma is proved.

The next lemma is due to our previous work [4] in the case when $\left\|\frac{z^{(k)}-\alpha}{d(\alpha)}\right\|_p$ is replaced by $\frac{\|z^{(k)}-\alpha\|_p}{d(\alpha)}$.

Lemma 2.5. *Let $P \in \mathscr{C}[z]$ be a monic polynomial of degree $n \geq 2$, where $\alpha = \{\alpha \in \mathscr{C}^n : \alpha_i \neq 0$ and $\alpha_i \neq \alpha_j$ for $i, j = 1, \dots, n\}$ is the root-vector of P, $1 \leq p \leq \infty$. Let for any $k \geq 0$*

$$E_k = E(z^{(k)}) = \frac{\|z^{(k)}-\alpha\|_p}{d} < \frac{1}{2^{\frac{1}{q}}}, \tag{24}$$

where $d = d(\alpha)$ is defined by (6).Then the iteration $z^{(k)}$ is well defined and it has distinct components. Besides,

$$\|z^{(k+1)}-\alpha\|_p \leq \sigma_k \|z^{(k)}-\alpha\|_p \tag{25}$$

and

$$E(z^{(k+1)}) \leq \sigma_k E(z^{(k)}), \tag{26}$$

where

$$\sigma_k = \sigma(E_k) = \frac{\frac{1}{1-E_k}\left(1+\frac{E_k}{(n-1)^{1/p}(1-2^{1/q}E_k)}\right)^{n-1} - 1}{1-\frac{E_k}{1-E_k}\left(1+\frac{E_k}{(n-1)^{1/p}(1-2^{1/q}E_k)}\right)^{n-1}}. \tag{27}$$

Proof. Using the triangle inequality, Lemma 2.1 and (24), we get for $i \neq j$

$$\begin{aligned} |z_i^{(k)} - z_j^{(k)}| &\geq |\alpha_i - \alpha_j| - |z_i^{(k)} - \alpha_i| - |z_j^{(k)} - \alpha_j| \\ &= \left(1 - \left|\frac{z_i^{(k)}-\alpha_i}{\alpha_i-\alpha_j}\right| - \left|\frac{z_j^{(k)}-\alpha_j}{\alpha_i-\alpha_j}\right|\right)|\alpha_i-\alpha_j| \\ &\geq \left(1 - 2^{\frac{1}{q}}\frac{\|z^{(k)}-\alpha\|_p}{d(\alpha)}\right)d(\alpha) = (1-2^{\frac{1}{q}}E_k)d > 0, \end{aligned} \tag{28}$$

which means that $z^{(k)}$ has distinct components, i.e. the iteration is well defined. Now we will prove the following estimates

$$|z_i^{(k+1)} - \alpha_i| \le \sigma_k |z_i^{(k)} - \alpha_i| \quad \text{for} \quad i = 1,2,\dots,n. \tag{29}$$

For easy of later comparisons, we will use the following equivalent form of (3)

$$z_i^{(k+1)} = z_i^{(k)} - \frac{W_i(z^{(k)})}{1 + \frac{W_i(z^{(k)})}{z_i^{(k)}}}, \quad i = 1,2,\dots,n, \tag{30}$$

which implies

$$z_i^{(k+1)} - \alpha_i = z_i^{(k)} - \alpha_i - \frac{W_i(z^{(k)})}{1 + \frac{W_i(z^{(k)})}{z_i^{(k)}}} = (z_i^{(k)} - \alpha_i)\left[1 - \frac{\prod_{j\ne i}^n \frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}}}{1 + \frac{W_i(z^{(k)})}{z_i^{(k)}}}\right]$$

and consequently

$$(z_i^{(k+1)} - \alpha_i) = (z_i^{(k)} - \alpha_i)\left[\frac{1 - \prod_{j\ne i}^n \frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}} + \frac{W_i(z^{(k)})}{z_i^{(k)}}}{1 + \frac{W_i(z^{(k)})}{z_i^{(k)}}}\right].$$

Therefore

$$|z_i^{(k+1)} - \alpha_i| = A_i^{(k)} |z_i^{(k)} - \alpha_i|, \tag{31}$$

where

$$A_i^{(k)} := \left|\frac{1 - \prod_{j\ne i}^n \frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}} + \frac{W_i(z^{(k)})}{z_i^{(k)}}}{1 + \frac{W_i(z^{(k)})}{z_i^{(k)}}}\right|.$$

Now, we can bound the amplification factor $A_i^{(k)}$ as follows

$$A_i^{(k)} \le \frac{\left|\prod_{j\ne i}^n \frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}} - 1\right| + \left|\frac{z_i^{(k)} - \alpha_i}{z_i^{(k)}}\right| \left|\prod_{j\ne i}^n \frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}}\right|}{1 - \left|\frac{z_i^{(k)} - \alpha_i}{z_i^{(k)}}\right| \left|\prod_{j\ne i}^n \frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}}\right|}. \tag{32}$$

We next establish the following inequalities

$$|z_i^{(k)}| \ge |\alpha_i| - |z_i^{(k)} - \alpha_i| \ge d - \|z^{(k)} - \alpha\|_p \ge (1 - E_k)d > 0. \tag{33}$$

It follows from (28) and the definition of $E(z^{(k)})$ that

$$\left|\frac{z_i^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}}\right| = \left|1 + \frac{z_j^{(k)} - \alpha_j}{z_i^{(k)} - z_j^{(k)}}\right| \le 1 + \frac{|z_j^{(k)} - \alpha_j|}{(1 - 2^{\frac{1}{q}} E_k) d(\alpha)}. \tag{34}$$

From (32), the last two inequalities (34) and (33), and Lemma 2.2, we obtain

$$A_i^{(k)} \le \frac{\left(1+\frac{E_k}{(n-1)^{\frac{1}{p}}(1-2^{\frac{1}{q}}E_k)}\right)^{n-1} - 1 + \frac{E_k}{1-E_k}\left(1+\frac{E_k}{(n-1)^{\frac{1}{p}}(1-2^{\frac{1}{q}}E_k)}\right)^{n-1}}{1-\frac{E_k}{1-E_k}\left(1+\frac{E_k}{(n-1)^{\frac{1}{p}}(1-2^{\frac{1}{q}}E_k)}\right)^{n-1}} \tag{35}$$

and consequently

$$A_i^{(k)} \le \frac{\frac{1}{1-E_k}\left(1+\frac{E_k}{(n-1)^{\frac{1}{p}}(1-2^{\frac{1}{q}}E_k)}\right)^{n-1} - 1}{1-\frac{E_k}{1-E_k}\left(1+\frac{E_k}{(n-1)^{\frac{1}{p}}(1-2^{\frac{1}{q}}E_k)}\right)^{n-1}} = \sigma_k\,.$$

Finally, from the last expression and (31) we obtain (29). Taking the p-norm in (29), we deduce the inequality in (25). We get the second inequality in (26) taking the p-norm and by dividing both sides of inequality (29) by $d(\alpha)$.

Now we are ready to state the main result of this paper which improves the previous results introduced in [4].

Theorem 2.6. *Let $P \in \mathscr{C}[z]$ be a monic polynomial of degree $n \ge 2$, where $\alpha = \{\alpha \in \mathscr{C}^n : \alpha_i \ne 0$ and $\alpha_i \ne \alpha_j$ for $i, j = 1, \dots, n\}$ is the root vector of P, and let $1 \le p \le \infty$. Suppose $z^{(0)} \in \mathscr{C}^n$ is an initial guess satisfying*

$$E(z^{(0)}) = \frac{\|z^{(0)}-\alpha\|_p}{d(\alpha)} < R(n,p) = \frac{2^{\frac{1}{q(n+1)}}-1}{2^{\frac{1}{q}}(2^{\frac{1}{q(n+1)}}-1)+(n-1)^{-\frac{1}{p}}}\,. \tag{36}$$

Then the following statements hold true.
(i) CONVERGENCE. *The Inverse Weierstrass iteration (2)-(3) is well defined and converges quadratically to the root-vector α of P.*
(ii) A POSTERIORI ERROR ESTIMATE. *For all $k \ge 0$ we have the estimate*

$$\|z^{(k+1)}-\alpha\|_p \le \lambda^{2^k}\|z^{(k)}-\alpha\|_p\,, \tag{37}$$

(iii) A PRIORI ERROR ESTIMATE. *For all $k \ge 1$ we have the estimate*

$$\|z^{(k)}-\alpha\|_p \le \lambda^{2^k-1}\|z^{(0)}-\alpha\|_p\,, \tag{38}$$

where $\lambda = E(z^{(0)})/R(n,p)$.

Proof. (i) From the Lemma 2.4 it follows that $R = R(n,p)$ is the unique solution of the equation $\phi(t)^{n+1} = 2^{1/q}$ in the interval $(0, 1/2^{1/q})$, where $\phi(t)$ is defined by (16). First, we will prove that for any $k \ge 0$ the iteration $z^{(k)}$ in (2)-(3) is well defined and

$$E(z^{(k)}) \le R\lambda^{2^k}\,.$$

We shall use mathematical induction to prove the statement. First, we confirm that the base case $k = 0$ is true due to definition of λ. Using the assumption $n \ge 2$ and that $(n-1)^{-\frac{1}{p}} > 0$ for any

$1 \le p < \infty$, it can be shown that $R < 1/2^{1/q}$. Then the initial assumption $E(z^{(0)}) < R$ implies $E(z^{(0)}) < 1/2^{1/q}$. From this and Lemma 2.5 we deduce that the iteration $z^{(0)}$ is well defined.

Now, we will prove that

$$\sigma(E_0) \le \lambda\,.$$

From (27) it follows that $\sigma(E_k)$ can be written in the following equivalent form

$$\sigma(E_k) = \frac{\left(1+\frac{E_k}{(n-1)^{1/p}(1-2^{1/q}E_k)}\right)^{n-1} - 1 + E_k}{1 - E_k - E_k\left(1+\frac{E_k}{(n-1)^{1/p}(1-2^{1/q}E_k)}\right)^{n-1}}\,, \tag{39}$$

which implies

$$\sigma(E_0) \le \frac{\left(\left(1+\frac{\lambda R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1} - 1\right) + \lambda R}{1 - R - R\left(1+\frac{R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1}}\,.$$

From Lemma 2.3, where $\varphi = \lambda$ and $t = \frac{R}{(n-1)^{1/p}(1-2^{1/q}R)}$ we deduce

$$\sigma(E_0) \le \frac{\lambda\left(\left(1+\frac{R}{(n-1)^{1/p}(1-2^{1/q}\lambda R)}\right)^{n-1} - 1\right) + \lambda R}{1 - R - R\left(1+\frac{R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1}} \le \lambda\sigma(R)\,, \tag{40}$$

where

$$\sigma(R) = \frac{\left(\left(1+\frac{R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1} - 1\right) + R}{1 - R - R\left(1+\frac{R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1}}\,.$$

It is easy to show that the assumption $\sigma(R) < 1$ is equivalent to the assumption defined by (18), where $t = R$. Therefore, from (40) using Lemma 2.3 and the definition of R we deduce that $\sigma(E_0) \le \lambda$.

Suppose that for any $k \ge 0$ is fulfilled

$$E(z^{(k)}) \le R\lambda^{2^k} \tag{41}$$

and we will prove that

$$E(z^{(k+1)}) \le R\lambda^{2^{k+1}}\,.$$

From (39) and the assumption by induction we obtain

$$\sigma(E_k) \le \frac{\left(\left(1+\frac{\lambda^{2^k}R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1} - 1\right) + \lambda^{2^k}R}{1 - R - R\left(1+\frac{R}{(n-1)^{1/p}(1-2^{1/q}R)}\right)^{n-1}} \le \lambda^{2^k}\sigma(R) < \lambda^{2^k}\,. \tag{42}$$

Therefore, from (42), the assumption (41) and the estimate (26) it follows that

$$E(z^{(k+1)}) \le \sigma_k E(z^{(k)}) \le \lambda^{2^k} R\lambda^{2^k} = \lambda^{2^{k+1}}R\,.$$

Using the inequality $R < 1/2^{1/q}$ it follows that $E(z^{(k+1)}) < 1/2^{1/q}$, which implies by Lemma 2.5 that the iteration $z^{(k+1)}$ is well defined.

(ii) From (42) and Lemma 2.5 estimate (25) it is trivial to prove that

$$\|z^{(k+1)} - \alpha\|_p \leq \lambda^{2^k} \|z^{(k)} - \alpha\|_p .$$

(iii) From the assertion (ii) and the sum of geometric progression, we obtain

$$\begin{aligned} \|z^{(k)} - \alpha\|_p &\leq \lambda^{2^{k-1}} \|z^{(k-1)} - \alpha\|_p \leq \lambda^{2^{k-1}} \lambda^{2^{k-2}} \|z^{(k-2)} - \alpha\|_p \leq \dots \\ &\leq \lambda^{2^{k-1}} \lambda^{2^{k-2}} \dots \lambda^{2^0} \|z^{(0)} - \alpha\|_p = \lambda^{2^k - 1} \|z^{(0)} - \alpha\|_p . \end{aligned}$$

which implies (38). The theorem is proved.

In the case of use the maximum vector norm, i.e. if $p = \infty$, we obtain the following corollary of Theorem 2.6.

Theorem 2.7. *Let $P \in \mathscr{C}[z]$ be a monic polynomial of degree $n \geq 2$, where $\alpha = \{\alpha \in \mathscr{C}^n : \alpha_i \neq 0$ and $\alpha_i \neq \alpha_j$ for $i, j = 1, \dots, n\}$ is the root vector of P, and let $p = \infty$. Suppose $z^{(0)} \in \mathscr{C}^n$ is an initial guess satisfying*

$$E(z^{(0)}) = \frac{\|z^{(0)} - \alpha\|_\infty}{d(\alpha)} < \frac{2^{\frac{1}{n+1}} - 1}{2.2^{\frac{1}{n+1}} - 1} . \tag{43}$$

Then the Inverse Weierstrass iteration (2)-(3) is well defined, converges quadratically to the root-vector α of P and the error estimates (37) and (38) are hold true.

3 Conclusion

In this work we investigate convergence analysis of the Inverse Weierstrass iterative method for simultaneous approximation of polynomial zeros. Our goal was to obtain new local convergence analysis results and to improve the existing ones. We establish a new convergence theorem with larger radius of convergence.

REFERENCES:

[1] Nedzhibov, G.H., Similarity transformations between some companion matrices, *AIP Conf. Proc.* **1631**, (2014), 375–382.

[2] Nedzhibov, G.H., Inverse Weierstrass-Durand-Kerner Iterative Method, *International Journal of Applied Mathematics, ISSN:2051-5227*, **28**, Issue.2, (2013), 1258–1264.

[3] Nedzhibov, G.H., On local convergence analysys of the Inverse WDK method, MATHTECH 2016, *Proceedings of the international conference***1**, (2016), 118–126.

[4] Nedzhibov, G.H., Local convergence of the Inverse Weierstrass method for simultaneous approximation of polynomial zeros, *International Journal of Mathematical Analysis* **10**, No. 26, (2016), 1295–1304.

[5] Nedzhibov, G.H., Convergence of the modified inverse Weierstrass method for simultaneous approximation of polynomial zeros, *Communications in Numerical Analysis* **No. 1**, (2016), 74–80.

[6] Kjurkchiev, N.V., Markov, S.M., Two interval methods for algebraic equations with real roots, *Pliska Stud. Math. Bulg.* **5**, (1983), 118–131.

[7] Zheng, S.M., On convergence of the Durand-Kerners method for finding all roots of a polynomial simultaneously, *Kexue Tongbao* **27**, **1982**, 1262–1265.

[8] Petković, M., On initial conditions for the convergence of simultaneous root-finding methods, *Computing* **57**, (1996), 163–177.

[9] Han, D.F., The convergence of the Durand-Kerner method for simultaneously finding all zeros of a polynomial, *J. Comput. Math.* **18**, (2000), 567–570.

[10] Niel, A.M., The simultaneous approximation of polynomial roots, *Computers and Mathematics with applications*, **41**, (2001), 1–14.

[11] Proinov, P.D., A new semilocal convergence theorem for the Weierstrass method from data at one point, *C. R. Acad. Bulg. Sci.* **59**, No 2, (2006), 131–136.

[12] Proinov, P.D., Petkova, M.D., A new semilocal convergence theorem for the Weierstrass method for finding zeros of a polynomial simultaneously, *Journal of complexity*, **30**, Issue.3, (2014), 366-380.

Gyurhan Nedzhibov
Konstantin Preslavsky University of Shumen
E-mail: gyurhan@shu.bg

COUPLED FIXED POINTS AND COUPLED BEST PROXIMITY POINTS FOR CYCLIC KANNAN TYPE CONTRACTION MAPS IN MODULAR FUNCTION SPACES*

ATANAS V. ILCHEV, BOYAN G. ZLATANOV

ABSTRACT: *We have found sufficient conditions for the existence and uniqueness of coupled fixed points and coupled best proximity points for Kannan type contraction maps in modular function spaces.*

KEYWORDS: *Coupled fixed points, Coupled best proximity points, Modular function space, Kannan maps*

1 Introduction

A fundamental result in fixed point theory is the Banach Contraction Principle in Banach spaces or in complete metric spaces. Fixed point theory is an important tool for solving equations $Tx = x$ for mappings T defined on subsets of metric or normed spaces. It is widely applied to nonlinear integral equations and differential equations.

The concept of coupled fixed point theorem is introduced in [3]. Later on Bhaskar and Lakshmikantham [1] introduced the notions of a mixed monotone mapping, studied the problems of uniqueness of a coupled fixed point in partially ordered metric spaces and applied their theorems to problems of the existence of solution for a periodic boundary value problem. Harjani, López and Sadarangani obtained in [5] some coupled fixed point theorems for a mixed monotone operator in a complete metric space endowed with a partial order by using altering distance functions. They applied their results to the study of the existence and uniqueness of a nonlinear integral equation. Recent results about the application of coupled fixed point theorems for solving integral equations are obtained in [9].

Other kind of a generalization of the Banach Contraction Principle is the notion of cyclic maps [13]. Because a non-self mapping $T : A \to B$ does not necessarily have a fixed point, one often attempts to find an element x which is in some sense closest to Tx. Best proximity point theorems are relevant in this perspective. The notion of best proximity point is introduced in [2]. This definition is more general than the notion of cyclic maps [13], in sense that if the sets intersect then every best proximity point is a fixed point. A sufficient condition for the existence and uniqueness of the best proximity points in uniformly convex Banach spaces is given in [2].

First results in the approximation of the sequence of successive iterations, which converges to the best proximity point for cyclic contractions is obtained in [26] and for coupled best proximity points in [6].

Besides the idea of defining a norm and considering a Banach space, another direction of generalization of the Banach Contraction Principle is based on considering an abstractly given functional defined on a linear space, which controls the growth of the members of the space. This functional is usually called modular and it defines a modular space. The theory of modular spaces was initiated by Nakano [21] in connection with the theory of ordered spaces, which was further generalized by Musielak and Orlicz [20]. Modular function spaces are subclass of the modular spaces. The study of the geometry of modular function spaces was initiated by Kozlowski [16, 14, 15]. Fixed point results in modular function spaces were obtained first by Khamsi, Kozlowski

*The first author is partially supported by fund MU17-FMI-007 of University of Plovdiv Paisii Hilendarski.

and Reich [11]. Further development of the theory of fixed points in modular function spaces can be found in the exhaustive references of the survey article [17] and in the book [10]. Kozlowski has contributed a lot towards the study of modular function spaces both on his own and with his collaborators. First results about best proximity points in modular function spaces are obtained in [7, 25].

A genarization of the idea of coupled fixed points and coupled best proximity points in modular function spaces was done in [8]. We have tried to generalize the idea of coupled fixed points and coupled best proximity points in modular function spaces for Kannan contraction maps.

2 Preliminaries

Following [17, 10] we will recall some basic notions and facts about modular function spaces.

Let Ω be a nonempty set and Σ be a nontrivial σ–algebra of subsets of Ω. Let $\mathscr{P}$ be a δ–ring of subsets of Ω, such that $E \cap A \in \mathscr{P}$ for any $E \in \mathscr{P}$ and $A \in \Sigma$. let us assume that there exists an increasing sequence of sets $K_n \in \mathscr{P}$ such that $\Omega = \cup K_n$. By $\mathscr{E}$ we denote the linear space of all simple functions with supports from $\mathscr{P}$. By $\mathscr{M}_\infty$ we will denote the space of all extended measurable functions, i.e. all functions $f : \Omega \to [-\infty, \infty]$ such that there exists a sequence $\{g_n\} \subset \mathscr{E}$, $|g_n| \leq |f|$ and $g_n(\omega) \to f(\omega)$ for all $\omega \in \Omega$. By $\mathbf{1}_A$ we denote the characteristic function of the set A.

Definition 1. *Let $\rho : \mathscr{M}_\infty \to [0, \infty]$ be a nontrivial convex and even function. We say that ρ is a regular convex function pseudomodular if:*

(i) $\rho(0) = 0$;

(ii) *ρ is monotone, i.e., $|f(\omega)| \leq |g(\omega)|$ for all $\omega \in \Omega$ implies $\rho(f) \leq \rho(g)$, where $f, g \in \mathscr{M}_\infty$;*

(iii) *ρ is orthogonaly subadditive, i.e., $\rho(f\mathbf{1}_{A\cup B}) \leq \rho(f\mathbf{1}_A) + \rho(f\mathbf{1}_B)$, where $A, B \in \Sigma$ such that $A \cap B \neq \emptyset$, $f \in \mathscr{M}_\infty$;*

(iv) *ρ has the Fatou property, i.e., $|f_n(\omega)| \uparrow |f(\omega)|$ for all $\omega \in \Omega$ implies $\rho(f_n) \uparrow \rho(f)$, where $f \in \mathscr{M}_\infty$;*

(v) *ρ is order continuous in $\mathscr{E}$, i.e., $g_n \in \mathscr{E}$ and $|g_n(\omega)| \downarrow 0$ implies $\rho(g_n) \downarrow 0$.*

Similarly as in the case of measure spaces, we say that a set $A \in \Sigma$ is ρ–null if $\rho(g\mathbf{1}_A) = 0$ for every $g \in \mathscr{E}$. We say that a property holds ρ–almost everywhere if the exceptional set is ρ–null. As usual we identify any pair of measurable sets whose symmetric difference is ρ–null as well as any pair of measurable functions differing only on a ρ–null set. With this in mind we define

$$\mathscr{M}(\Omega, \sigma, \mathscr{P}, \rho) = \{f \in \mathscr{M}_\infty; |f(\omega)| < \infty \rho \ - \ a.e.\},$$

where each $f \in \mathscr{M}(\Omega, \sigma, \mathscr{P}, \rho)$ is actually an equivalence class of functions equal ρ a.e. rather than an individual function. Where no confusion exists we will write $\mathscr{M}$ instead of $\mathscr{M}(\Omega, \sigma, \mathscr{P}, \rho)$.

Definition 2. *Let ρ be a regular convex function pseudomodular.*

(1) *We say that $\rho(0)$ is a regular convex function semimodular if $\rho(\alpha f) = 0$ for every $\alpha > 0$ implies $f = 0$ ρ–a.e.;*

(2) *We say that ρ is a regular convex function modular if $\rho(f) = 0$ implies $f = 0$ ρ–a.e.*

The class of all nonzero regular convex function modular defined on Ω will be denoted by $\mathfrak{R}$.

Let us denote $\rho(f,E) = \rho(f1_E)$ for $f \in \mathscr{M}$, $E \in \Sigma$. It is easy to prove that $\rho(f,E)$ is a function pseudomodular in the sense of Definition 2.1.1 in [16] (more precisely, it is a function pseudomodular with the Fatou property). Therefore, we can use all results of the standard theory of modular function spaces as per the framework defined by Kozlowski [16, 14, 15], see also Musielk [20] for the basics of the general modular theory.

Definition 3. *Let ρ be a convex function modular.*

(a) A modular function space is the vector space $L_\rho(\Omega,\Sigma)$, or briefly L_ρ, defined by

$$L_\rho = \{f \in \mathscr{M} : \rho(\lambda f) \to 0 \text{ as } \lambda \to 0\};$$

(b) The following formula defines a norm in L_ρ (frequently called Luxemburg norm):

$$\|f\|_\rho = \inf\left\{\alpha > 0 : \rho\left(\frac{f}{\alpha}\right) \leq 1\right\}.$$

For the rest of the article if we state someting about a norm we will mean Luxemburg norm $\|\cdot\|_\rho$, which is generated by the modular ρ.

In this way, Lebesgue, Orlicz, Musielak–Orlicz, Lorentz, Orlicz–Lorentz are examples of modular function spaces.

In the following theorem, we recall some of the basic properties of modular function spaces.

Theorem 4. *Let $\rho \in \mathfrak{R}$.*

(1) $(L_\rho, \|f\|_\rho)$ is a complete and the norm $\|\cdot\|_\rho$ is a monotone w.r.t the natural order in $\mathscr{M}$.

(2) $\|f_n\|_\rho \to 0$ iff $\rho(\alpha f_n) \to 0$ for every $\alpha > 0$.

(3) If $\rho(\alpha f_n) \to 0$ for an α, then there exists a subsequence $\{g_n\}$ of $\{f_n\}$ such that $g_n \to 0$ ρ–a.e.

(4) If $\{f_n\}$ converges uniformly to f on a set $E \in \mathscr{P}$, then $\rho(\alpha(f_n - f),E) \to 0$ for every $\alpha > 0$.

(5) Let $f_n \to f$ ρ–a.e. There exists a nondecreasing sequence of sets $H_k \in \mathscr{P}$ such that $H_k \uparrow \Omega$ and f_n converges uniformly to f on every H_k (Egoroff Theorem).

(6) $\rho(f) \leq \liminf \rho(f_n)$ whenever $f_n \to f$ ρ–a.e. (Note that this property is equivalent to the Fatou property).

(7) Define $L^0_\rho\{f \in L_\rho : \rho(f,\cdot)$ is order continuous$\}$ and $E_\rho = \{f \in L_\rho : \lambda f \in L^0_\rho$ for every $\lambda > 0\}$ we have

(a) $L_\rho \supset L^0_\rho \supset E_\rho$;

(b) E_ρ has the Lebesgue property, i.e. $\rho(\alpha f, D_k) \to 0$ for $\alpha > 0$, $f \in E_\rho$ and $D_k \downarrow \emptyset$;

(c) E_ρ is the closure of $\mathscr{E}$ (in the sense of $\|\cdot\|_\rho$).

The next definition gives generalizations of the classical notions for normed spaces in the context of modular function spaces.

Definition 5. *Let $L_\rho \in \mathfrak{R}$.*

(a) We say that $\{f_n\}$ is a ρ–convergent to f and we write $f_n \to f(\rho)$ if and only if $\rho(f_n - f) \to 0$;

(b) A sequence $\{f_n\}_{n=1}^{\infty} \subset L_\rho$ is called ρ–Cauchy if for any $\varepsilon > 0$ there exists $N \in \mathbb{N}$, such that for any $m > n \geq N$ there holds the inequality $\rho(f_m - f_n) < \varepsilon$;

(c) The modular function space L_ρ is called ρ–complete if any ρ–Cauchy sequence is ρ–convergent.

(d) A set $B \subset L_\rho$ is called ρ–closed if for any sequence of $f_n \in B$, the convergence $f_n \to f(\rho)$ implies that f belongs to B;

(e) A set $B \subset L_\rho$ is called ρ–bounded if its ρ–diameter $\delta_\rho(B) = \sup\{\rho(f-g) : f, g \in B\} < \infty$;

(f) A set $B \subset L_\rho$ is called ρ–a.e. closed if for any $\{f_n\}$ in C which ρ–a.e. converges to some f, then we must have $f \in C$;

(g) Let $A, B \subset L_\rho$. We define the ρ–distance between the sets A and B by

$$d_\rho(A,B) = \inf\{\rho(f-g) : f \in A, g \in B\}$$

and $d_\rho(f,B) = \inf\{\rho(f-g) : g \in B\}$ if A consists of a single element f;

(h) We say that a function modular ρ has the Δ_2–property if $\sup_{n\in\mathbb{N}} \rho(2f_n, D_k) \to 0$, whenever $D_k \downarrow \emptyset$ and $\sup_{n\in\mathbb{N}} \rho(f_n, D_k) \to 0$.

(i) ([10], p.116) A function modular $\rho \in \mathfrak{R}$ is called uniformly continuous if for any $L > 0$ and $\varepsilon > 0$ there exists $\delta = \delta(L,\varepsilon) > 0$ such that if $\rho(x) \leq L$ and $\rho(y) < \delta$ there holds the inequality $|\rho(x+y) - \rho(x)| < \varepsilon$.

Theorem 6. *Let $\rho \in \mathfrak{R}$. Then L_ρ is ρ–complete.*

Theorem 7. *Let $\rho \in \mathfrak{R}$. The following conditions are equivalent*

(a) ρ has Δ_2;

(b) L_ρ^0 is a linear subspace of L_ρ;

(c) $L_\rho = L_\rho^0 = E_\rho$;

(d) If $\rho(f_n) \to 0$, then $\rho(2f_n) \to 0$;

(e) If $\rho(\alpha f_n) \to 0$ for an $\alpha > 0$, then $\|f_n\|_\rho \to 0$, i.e., the modular convergence is equivalent to the norm convergence.

Let us mention that the ρ–convergence do not imply ρ–Cauchy, since ρ– does not satisfy the triangle inequality. If ρ has Δ_2–property then ρ–convergence imply ρ–Cauchy.

Generalization of convexity properties for Banach spaces are investigated for modular function spaces in [12]. As demonstrated in [17] one concept of uniform convexity for Banach spaces generates several different types of uniform convexity in modular function spaces. This is due primarily to the fact that in general the modular function is not homogeneous.

Definition 8. *Let $\rho \in \mathfrak{R}$ and $i \in \{1,2\}$. Let $r > 0$, $\varepsilon > 0$. Define*

$$D_i(r,\varepsilon) = \{(f,g) : f,g \in L_\rho, \rho(f) \le r, \rho(g) \le r, \rho\left(\frac{f-g}{i}\right) \ge \varepsilon r\}.$$

Let $\delta_i(r,\varepsilon) = \inf\left\{1 - \frac{1}{r}\rho\left(\frac{f+g}{2}\right) : (f,g) \in D_i(r,\varepsilon)\right\} > 0$ *if* $D_i(r,\varepsilon) \neq \emptyset$ *and* $\delta_i(r,\varepsilon) = 1$ *if* $D_i(r,\varepsilon) = \emptyset$.

(i) We say that ρ satisfies (UCi) if for any $r > 0$, $\varepsilon > 0$ there holds the inequality $\delta_i(r,s) > 0$.

(ii) We say that ρ satisfies (UUCi) if for every $s \ge 0$, $\varepsilon > 0$ there exists $\eta_i(s,\varepsilon) > 0$, depending on s and ε such that

$$\delta_i(r,s) > \eta_i(s,\varepsilon) > 0 \text{ for } r > s.$$

If ρ is (UC1) we obtain that the inequality

$$\rho\left(\frac{x+y}{2}\right) \le r\left(1 - \delta_1(r,\varepsilon)\right) \tag{1}$$

holds for every $\rho(x), \rho(y) \le r$ and $\rho(x-y) \ge r\varepsilon$.

Proposition 9. *The following conditions characterize relationship between the notions, that are defined in Definition 8*

(1) (UUCi) implies (UCi) for $i \in 1,2$;

(2) $\delta_1(r,\varepsilon) \le \delta_2(r,\varepsilon)$;

(3) (UC1) implies (UC2);

(4) (UUC1) implies (UUC2);

(5) If $\rho \in \mathfrak{R}$, then (UUC1) and (UUC2) are equivalent;

(6) I ρ is homogeneous (e.g. is a norm) then all conditions (UC1), (UC2), (UUC1) and (UUC2) are equivalent.

Lemma 10. *([25]) Let $\rho \in \mathfrak{R}$. Let ρ be (UC1), has the Δ_2–property, $A \subset L_\rho$ be a ρ–closed and convex subset, $B \subset L_\rho$ be ρ–closed subset and $A \cup B$ be ρ–bounded. If the sequences $\{x_n\}_{n=1}^\infty, \{z_n\}_{n=1}^\infty \subset A$ and $\{y_n\}_{n=1}^\infty \subset B$ be such that:*

(i) $\lim_{n\to\infty} \rho(z_n - y_n) = d_\rho$;

(ii) for every $\varepsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that for every $m > n \ge N_0$ there holds the inequality $\rho(x_m - y_n) \le d_\rho + \varepsilon$.

Then for every $\varepsilon > 0$ there exists $N_1 \in \mathbb{N}$ such that for every $m > n \ge N_1$ there holds the inequality $\rho(x_m - z_n) < \varepsilon$.

Lemma 11. *([25]) Let $\rho \in \mathfrak{R}$. Let ρ be (UC1), has the Δ_2–property, A be a ρ–closed and convex subset of L_ρ, B be ρ–closed subset of L_ρ and $A \cup B$ be ρ–bounded. If the sequences $\{x_n\}_{n=1}^\infty, \{z_n\}_{n=1}^\infty \subset A$ and $\{y_n\}_{n=1}^\infty \subset B$ be such that:*

(i) $\lim_{n\to\infty}\rho(z_n-y_n)=d_\rho$;

(ii) $\lim_{n\to\infty}\rho(x_n-y_n)=d_\rho$.

Then $\lim_{n\to\infty}\rho(x_n-z_n)=0$.

Lemma 12. *([25]) Let $\rho\in\mathfrak{R}$. Let ρ has the Δ_2–property, be uniformly continuous, $A,B\subset L_\rho$ be subsets and $A\cup B$ be ρ–bounded. If the sequences $\{x_n\}_{n=1}^\infty,\{z_n\}_{n=1}^\infty\subset A$ and $\{y_n\}_{n=1}^\infty\subset B$ be such that:*

(i) $\lim_{n\to\infty}\rho(z_n-x_n)=0$;

(ii) $\lim_{n\to\infty}\rho(z_n-y_n)=d_\rho$.

Then $\lim_{n\to\infty}\rho(x_n-y_n)=d_\rho$.

We recall that M is called an Orlicz function, provided M is even, convex, continuous non-decreasing in $[0,\infty)$ function with $M(0)=0$, $M(t)>0$ for any $t\neq 0$. Let M be an Orlicz function and let (Ω,Σ,μ) be a measure space. Let us consider the space $L^0(\Omega)$ consisiting of all measurable real–velued functions on Ω and define for every $f\in L^0(\Omega)$ the Orlicz function modular $\widetilde{M}(f)=\int_\Omega M(f(t))d\mu(t)$.

Definition 13. *The Orlicz space $L_M(\Omega,\Sigma,\mu)$ is the space of all classes of equivalent μ–measurable functions $f:\Omega\to\mathbb{R}$ over the measure space (Ω,Σ,μ) such that $\widetilde{M}(\lambda f)\to 0$ as $\lambda\to 0$ or equivalently $\widetilde{M}\left(\frac{f}{\lambda}\right)<\infty$ for some $\lambda>0$.*

The function $\widetilde{M}$ is a regular convex function modular and it is called Orlicz function modular. An extensive study of Orlicz spaces can be found in [18, 19, 22, 23].

If $M(t)=|t|^p$, $p\geq 1$ we obtain the space $L_p(\Omega,\Sigma,\mu)$. The most common examples of Orlicz spaces are the sequence spaces ℓ_M, the function spaces $L_M(0,1)$ and $L_M(0,\infty)$ that correspond to the cases: Ω countable union of atoms of equal mass, $\Omega=[0,1]$ and $\Omega=(0,\infty)$, μ the usual Lebesgue measure.

We say that M satisfies the Δ_2–condition if there exist constants $C,t_0>0$, such that $M(2t)\leq CM(t)$ for any $t\geq t_0$. It is easy to observe that if M satisfies the Δ_2–condition, then the Orlicz function modular $\widetilde{M}$ has the Δ_2 property.

If we restrict to the Orlicz space $L_M(0,1)$, then the Orlicz function modular is defined by $\widetilde{M}(f)=\displaystyle\int_0^1 M(f(s))d\mu(s)$. We will denote the corresponding modular function space by $L_{\widetilde{M}}(0,1)$. When $M=|t|^p$ we will denote $L_{\widetilde{M}}(0,1)$ by $L_{\widetilde{p}}(0,1)$.

3 Main Results

Coupled best proximity points and coupled fixed points in metric spaces are defined in [24] and [3]. Following [24, 3] we will give definitions for coupled best proximity points and coupled fixed points in modular function spaces.

Just to simplify the notations we will put $d=d_\rho(A,B)$.

Definition 14. *Let A and B be nonempty subsets of a modular function space L_ρ, $F:A\times A\to B$. An ordered pair $(x,y)\in A\times A$ is called a coupled best proximity point of F if*

$$\rho(x-F(x,y))=\rho(y-F(y,x))=d.$$

Definition 15. *Let A be nonempty subset of a modular function space X, $F : A\times A \to A$. An ordered pair $(x,y) \in A\times A$ is said to be a coupled fixed point of F in A if $x = F(x,y)$ and $y = F(y,x)$.*

It is easy to see that if $A = B$ in Definition 14, then a coupled best proximity point reduces to a coupled fixed point.

Iterated sequences for investigation of existence of coupled best proximity points and coupled fixed points in metric spaces is defined in [3, 24]. Following [24, 3] we will give definitions for these iterated sequences in modular functional spaces.

Definition 16. *([24])Let A and B be nonempty subsets of a modular function space L_ρ. Let $F : A\times A \to B$ and $G : B\times B \to A$. For any pair $(x,y) \in A\times A$ we define the sequences $\{x_n\}_{n=0}^{\infty}$ and $\{y_n\}_{n=0}^{\infty}$ by $x_0 = x$, $y_0 = y$ and*

$$\begin{array}{ll} x_{2n+1} = F(x_{2n},y_{2n}), & y_{2n+1} = F(y_{2n},x_{2n}) \\ x_{2n+2} = G(x_{2n+1},y_{2n+1}), & y_{2n+2} = G(y_{2n+1},x_{2n+1}) \end{array}$$

for all $n \geq 0$.

We will generalize the notion of cyclic contraction pair (F,G) of maps in metric spaces [4, 24] for modular function spaces.

Definition 17. *Let A and B be nonempty subsets of a modular function space X, $F : A\times A \to B$ and $G : B\times B \to A$. The ordered pair (F,G) is said to be a cyclic ρ–Kannan contraction pair if there exist non-negative number α, such that $\alpha < 1/2$ and there holds the inequality*

$$\rho(F(x,y)-G(u,v)) \leq \alpha\left(\rho(x-F(x,y))+\rho(u-G(u,v))\right)+(1-2\alpha)d(A,B)$$

for all $(x,y) \in A\times A$ and $(u,v) \in B\times B$.

Definition 18. *Let A be nonempty subsets of a modular function space X, $F : A\times A \to A$ is said to be a cyclic ρ-Kannan contraction map if there exist non-negative numbers α, such that $\alpha < 1/2$ and there holds the inequality*

$$\rho(F(x,y)-F(u,v)) \leq \alpha\left(\rho(x-F(x,y))+\rho(u-F(u,v))\right)$$

for all $x,y,u,v \in A$.

Theorem 19. *Let $\rho \in \Re$. Let $A \subset L_\rho$ be nonempty, ρ–closed and ρ–bounded. Let $F : A\times A \to A$ be a cyclic ρ-Kannan contraction map. Then F has unique coupled fixed points $(x,y) \in A$. Moreover for any $(x_0,y_0) \in A$ the sequences $\{x_n\}$, $\{y_n\}$ defined by the equations:*

$$\begin{array}{c} x_1 = F(x_0,y_0),\ y_1 = F(y_0,x_0), \\ x_{n+1} = F(x_n,y_n),\ y_{n+1} = F(y_n,x_n), \end{array} \tag{2}$$

$n = 1,2,\ldots$, converge to the unique coupled fixed points $(x,y) \in A$.

Proof. From the definition of (2) we get

$$\rho(x_{n+1}-x_n) = \rho(F(x_n,y_n)-F(x_{n-1},y_{n-1})) \leq \alpha\rho(x_n-x_{n+1})+\alpha\rho(x_{n-1}-x_n)$$

and
$$\rho(y_{n+1}-y_n)=\rho(F(y_n,x_n)-F(y_{n-1},x_{n-1}))\leq \alpha\rho(y_n-y_{n+1})+\alpha\rho(y_{n-1}-y_n).$$
And we get
$$\rho(x_{n+1}-x_n)\leq \frac{\alpha}{1-\alpha}\rho(x_{n-1}-x_n)$$
and
$$\rho(y_{n+1}-y_n)\leq \frac{\alpha}{1-\alpha}\rho(y_{n-1}-y_n).$$
Therefore after summing the above inequalities we get
$$\begin{aligned}\rho(x_{n+1}-x_n)+\rho(y_{n+1}-y_n) &\leq \frac{\alpha}{1-\alpha}(\rho(x_n-x_{n-1})+\rho(y_n-y_{n-1}))\\ &\leq \left(\frac{\alpha}{1-\alpha}\right)^2(\rho(x_{n-1}-x_{n-1})+\rho(y_{n-1}-y_{n-2})\\ &\leq \cdots\leq \left(\frac{\alpha}{1-\alpha}\right)^n(\rho(x_1-x_0)+\rho(y_1-y_0)).\end{aligned}$$
From the last chain of inequalities we get that for any $n,p\in\mathbb{N}$ there holds
$$\rho(x_{n+p}-x_n)+\rho(y_{n+p}-y_n)\leq \left(\frac{\alpha}{1-\alpha}\right)^n(\rho(x_p-x_0)+\rho(y_p-y_0)).$$
From the assumption that A is a ρ–bounded set it follows that existence of $M>0$, such that $\rho(u-v)\leq M$ for any $u,v\in A$. Consequently
$$\rho(x_{n+p}-x_n)+\rho(y_{n+p}-y_n)\leq \left(\frac{\alpha}{1-\alpha}\right)^n(\rho(x_p-x_0)+\rho(y_p-y_0))\leq 2M\left(\frac{\alpha}{1-\alpha}\right)^n.$$
From $\frac{\alpha}{1-\alpha}\in[0,1)$ it follows that for any $\varepsilon>0$ there is N_0 such that for any $n\geq N_0$, there holds $\left(\frac{\alpha}{1-\alpha}\right)^n<\frac{\varepsilon}{2M}$. Then for any $n_0\geq\mathbb{N}$ and any $p\geq 1$ the inequality $\max\{\rho(x_{n+p}-x_n),\rho(y_{n+p}-y_n)\}\leq 2M\left(\frac{\alpha}{1-\alpha}\right)^n<\varepsilon$ holds. Thus for every $\varepsilon>0$ there exists N_0 such that for any $n\geq N_0$ the inequality $\rho(z_{n+p}-z_n)<\varepsilon$ holds, for $p\in\mathbb{N}$ and $z=x$ or $z=y$. Therefore $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences. From the completeness of L_ρ and the closeness of A it follows that there are $x,y\in A$, such that $\lim_{n\to\infty}\rho(x_n-x)=0$ and $\lim_{n\to\infty}\rho(y_n-y)=0$.

We will proof that $F(x,y)=x$ and $F(y,x)=y$.
$$\begin{aligned}\rho\left(x-F(x,y)\right) &= \lim_{n\to\infty}\rho\left(x_{n+1}-F(x,y)\right)\leq \lim_{n\to\infty}\rho\left(F(x_n,y_n)-F(x,y)\right)\\ &\leq \lim_{n\to\infty}\alpha\rho(x_n-x_{n+1})+\alpha\rho(x-F(x,y)).\end{aligned}$$
Thus we get the inequality $(1-\alpha)\rho\left(x-F(x,y)\right)\leq\lim_{n\to\infty}\alpha\rho(x_n-x_{n+1})=0$, i.e. $x=F(x,y)$. In a similar fashion we get
$$\begin{aligned}\rho\left(y-F(y,x)\right) &= \lim_{n\to\infty}\rho\left(y_{n+1}-F(y,x)\right)\leq \lim_{n\to\infty}\rho\left(F(y_n,x_n)-F(y,x)\right)\\ &\leq \lim_{n\to\infty}\alpha\rho(y_n-y_{n+1})+\alpha\rho(y-F(y,x)).\end{aligned}$$
Thus we get the inequality $(1-\alpha)\rho\left(y-F(y,x)\right)\leq\lim_{n\to\infty}\alpha\rho(y_n-y_{n+1})=0$, i.e. $x=F(x,y)$. Therefore (x,y) is coupled fixed points of F in A.

Suppose (u,v) is another coupled fixed points of F in A i.e. $u = F(u,v)$ and $v = F(v,u)$. From the inequalities

$$\rho(x-u) = \rho(F(x,y) - F(u,v) \leq \alpha\rho(x - F(x,y)) + \alpha\rho(u - F(u,v))$$

and

$$\rho(y-v) = \rho(F(y,x) - F(v,u) \leq \alpha\rho(y - F(y,x)) + \alpha\rho(v - F(v,u))$$

we get that $\rho(x-u) = \rho(y-v) = 0$, which is a contradiction. So the assumption that there exists second coupled fixed points is not true.

By what we have just proved for any initial guess (u_0, v_0) the sequences $\{u_n\}_{n=1}^{\infty}$ and $\{v_n\}_{n=1}^{\infty}$ defined in (2) converge to a coupled fixed point (u,v) of F in A, which is unique. Thus $\{u_n\}_{n=1}^{\infty}$ and $\{v_n\}_{n=1}^{\infty}$ converge to the coupled fixed point (x,y) □

Theorem 20. *Let $\rho \in \Re$. Assume that ρ satisfies (UC1), has the Δ_2–property and be uniformly continuous. Let $A,B \subseteq L_\rho$ be ρ–closed, ρ–bounded, convex subsets, $F : A\times A \to B$ and $G : B\times B \to A$ and the ordered pair (F,G) be an cyclic ρ–Kannan contraction pair. Then there exists a unique order pair $(x,y) \in A\times A$ such that (x,y) is a coupled ρ–best proximity points of F in A (i.e. $\rho(x - F(x,y)) + \rho(y - F(y,x)) = 2d(A,B)$). There holds $x = G(F(x,y),F(y,x))$, $y = G(F(y,x),F(x,y))$ the order pair $(F(y,x),F(x,y))$ is a coupled ρ–best proximity points of G in B. More over for any initial guess $(x_0,y_0) \in A\times A$ the iterated sequences $\{x_n\}$, $\{y_n\}$ defined by*

$$\begin{array}{l} x_{2n+1} = F(x_{2n},y_{2n}),\ y_{2n+1} = F(y_{2n},x_{2n}), \\ x_{2n+2} = G(x_{2n+1},y_{2n+1}), y_{2n+2} = G(y_{2n+1},x_{2n+1}), \\ n = 0,1,2\ldots \end{array} \tag{3}$$

satisfied

$$\lim_{n\to\infty}\rho(x_{2n}-x) = 0,\ \lim_{n\to\infty}\rho(y_{2n}-y) = 0,\ \lim_{n\to\infty}\rho(x_{2n+1} - F(x,y)) = 0,\ \lim_{n\to\infty}\rho(y_{2n+1} - F(y,x)) = 0.$$

Lemma 21. *Let $\rho \in \Re$. Assume that ρ satisfies (UC1), has the Δ_2–property and be uniformly continuous. Let $A,B \subset L_\rho$ be nonempty, ρ–closed and ρ–bounded, convex subsets, $F : A\times A \to B$ and $G : B\times B \to A$ and the ordered pair (F,G) be a cyclic ρ–Kannan contraction pair. If $(x_0,y_0) \in A$ and we define:*

$$\begin{array}{ll} x_{2n+1} = F(x_{2n},y_{2n}), & y_{2n+1} = F(y_{2n},x_{2n}) \\ x_{2n+2} = G(x_{2n+1},y_{2n+1}), & y_{2n+2} = G(y_{2n+1},x_{2n+1}) \end{array}$$

for all $n \in \mathbb{N}\cup\{0\}$, then $\lim_{n\to\infty}\rho(x_{2n} - x_{2n+1}) = d$, $\lim_{n\to\infty}\rho(y_{2n} - y_{2n+1}) = d$.

Proof. From the assumption that (F,G) is a cyclic ρ–Kannan contraction pair we get

$$\begin{array}{rcl} \rho(x_{2n+1} - x_{2n}) & = & \rho(F(x_{2n},y_{2n}) - G(x_{2n-1},y_{2n-1})) \\ & \leq & \alpha\rho(x_{2n} - x_{2n+1}) + \alpha\rho(x_{2n-1} - x_{2n}) + (1-2\alpha)d \end{array}$$

and

$$\begin{array}{rcl} \rho(y_{2n+1} - y_{2n}) & = & \rho(F(y_{2n},x_{2n}) - G(y_{2n-1},x_{2n-1})) \\ & \leq & \alpha\rho(y_{2n} - y_{2n+1}) + \alpha\rho(y_{2n-1} - y_{2n})) + (1-2\alpha)d. \end{array}$$

Therefore

$$\rho(x_{2n+1} - x_{2n}) \leq \frac{\alpha}{1-\alpha}\rho(x_{2n} - x_{2n-1}) + \left(1 - \frac{\alpha}{1-\alpha}\right)d$$

and

$$\rho(y_{2n+1}-y_{2n}) \leq \frac{\alpha}{1-\alpha}\rho(y_{2n}-y_{2n-1})+\left(1-\frac{\alpha}{1-\alpha}\right)d$$

Thus

$$\begin{array}{rcl}
\rho(x_{2n+1}-x_{2n})-d & \leq & \left(\frac{\alpha}{1-\alpha}\right)(\rho(x_{2n}-x_{2n-1})-d)\\
& \leq & \cdots\\
& \leq & \left(\frac{\alpha}{1-\alpha}\right)^{2n}(\rho(x_1-x_0)-d)\\
\rho(y_{2n+1}-y_{2n})-d & \leq & \left(\frac{\alpha}{1-\alpha}\right)(\rho(y_{2n}-y_{2n-1})-d)\\
& \leq & \cdots\\
& \leq & \left(\frac{\alpha}{1-\alpha}\right)^{2n}(\rho(y_1-y_0)-d)
\end{array}$$

Therefore $\lim_{n\to\infty}\rho(x_{2n}-x_{2n+1})=d$ and $\lim_{n\to\infty}=\rho(y_{2n}-y_{2n+1})=d$. □

Lemma 22. *Let $\rho\in\mathfrak{R}$. Assume that ρ satisfies (UC1), has the Δ_2–property and be uniformly continuous. Let $A,B\subseteq L_\rho$ be ρ–closed, ρ–bounded, convex subsets, $F:A\times A\to B$ and $G:B\times B\to A$ and the ordered pair (F,G) be a cyclic ρ–Kannan contraction pair. If $(x_0,y_0)\in A\times A$ and we define:*

$$\begin{array}{ll}
x_{2n+1}=F(x_{2n},y_{2n}), & y_{2n+1}=F(y_{2n},x_{2n})\\
x_{2n+2}=G(x_{2n+1},y_{2n+1}), & y_{2n+2}=G(y_{2n+1},x_{2n+1})
\end{array}$$

for all $n\in\mathbb{N}\cup\{0\}$, then for $\varepsilon>0$, there exists a positive integer N_0 such that for all $m>n\geq N_0$ such that $n+m$ is an odd number we have $\rho(x_m-x_n)+\rho(y_m-x_n)<2d+\varepsilon$.

Proof. From the assumption that (F,G) is a cyclic ρ–Kannan contraction pair we get

$$\begin{array}{rcl}
\rho(x_{2m}-x_{2n+1}) & = & \rho(G(x_{2m-1},y_{2m-1})-F(x_{2n},y_{2n}))\\
& \leq & \alpha\rho(x_{2m-1}-x_{2m})+\alpha\rho(x_{2n}-x_{2n+1})+(1-2\alpha)d.
\end{array}$$

Thus

$$\begin{array}{rcl}
\rho(x_{2m}-x_{2n+1})-d & \leq & \alpha(\rho(x_{2m-1}-x_{2m})-d)+\alpha(\rho(x_{2n}-x_{2n+1})-d)\\
& \leq & \left(\frac{\alpha}{1-\alpha}\right)^{2m-1}\alpha(\rho(x_0-x_1)-d)+\left(\frac{\alpha}{1-\alpha}\right)^{2n}\alpha(\rho(x_0-x_1)-d)
\end{array}$$

and

$$\begin{array}{rcl}
\rho(y_{2m}-y_{2n+1}) & = & \rho(G(y_{2m-1},x_{2m-1})-F(y_{2n},x_{2n}))\\
& \leq & \alpha\rho(y_{2m-1}-y_{2m})+\alpha\rho(x_{2n}-x_{2n+1})+(1-2\alpha)d
\end{array}$$

Thus

$$\begin{array}{rcl}
\rho(y_{2m}-y_{2n+1})-d & \leq & \alpha(\rho(y_{2m-1}-y_{2m})-d)+\alpha(\rho(y_{2n}-y_{2n+1})-d)\\
& \leq & \left(\frac{\alpha}{1-\alpha}\right)^{2m-1}\alpha(\rho(y_0-y_1)-d)+\left(\frac{\alpha}{1-\alpha}\right)^{2n}\alpha(\rho(y_0-y_1)-d)
\end{array}$$

Therefore for any $\varepsilon>0$, there exists a positive integer N_0 such that for all $m>n\geq N_0$ such that $n+m$ is an odd number we have $\rho(x_m-x_n)+\rho(y_m-x_n)<2d+\varepsilon$. □

Lemma 23. *Let $\rho\in\mathfrak{R}$. Assume that ρ satisfies (UC1), has the Δ_2–property and be uniformly continuous. Let $A,B\subseteq L_\rho$ be ρ–closed, ρ–bounded, convex subsets, $F:A\times A\to B$ and $G:B\times B\to A$ and the ordered pair (F,G) be a cyclic ρ–Kannan contraction pair. If $(x_0,y_0)\in A\times A$ and we define:*

$$\begin{array}{ll}
x_{2n+1}=F(x_{2n},y_{2n}), & y_{2n+1}=F(y_{2n},x_{2n})\\
x_{2n+2}=G(x_{2n+1},y_{2n+1}), & y_{2n+2}=G(y_{2n+1},x_{2n+1}).
\end{array}$$

for all $n\in\mathbb{N}\cup\{0\}$. If the sequences $\{x_{2n}\}_{n=0}^\infty$ and $\{y_{2n}\}_{n=0}^\infty$ are ρ–convergent to x and y respectively, then $\rho(x-F(x,y))=d$ and $\rho(y-F(y,x))=d$.

Proof. From Lemma 21 we have that there hold the equalities $\lim_{n\to\infty}\rho(x_{2n}-x_{2n+1})=d$ and $\lim_{n\to\infty}\rho(x_{2n+2}-x_{2n+1})=d$. Thus from Lemma 11 we obtain $\lim_{n\to\infty}\rho(x_{2n}-x_{2n+2})=0$. From $\lim_{n\to\infty}\rho(x_{2n}-x_{2n-1})=d$ and $\lim_{n\to\infty}\rho(x_{2n}-x)=0$ applying Lemma 12 we get $\lim_{n\to\infty}\rho(x_{2n-1}-x)=d$.

From Lemma 21 we have that there hold the equalities $\lim_{n\to\infty}\rho(y_{2n}-y_{2n+1})=d$ and $\lim_{n\to\infty}\rho(y_{2n+2}-y_{2n+1})=d$. Thus from Lemma 11 we obtain $\lim_{n\to\infty}\rho(y_{2n}-y_{2n+2})=0$. From $\lim_{n\to\infty}\rho(y_{2n}-y_{2n-1})=d$ and $\lim_{n\to\infty}\rho(y_{2n}-y)=0$ applying Lemma 12 we get $\lim_{n\to\infty}\rho(y_{2n-1}-y)=d$.

Since

$$\begin{aligned}\rho(F(x,y)-x) &\leq \lim_{n\to\infty}\rho(F(x,y)-x_{2n})=\lim_{n\to\infty}\rho(F(x,y)-G(x_{2n-1},y_{2n-1}))\\ &\leq \lim_{n\to\infty}[\alpha\rho(x-F(x,y))+\alpha\rho(x_{2n-1}-x_{2n})+(1-2\alpha)d]=d.\end{aligned}$$

Thus

$$(1-\alpha)d\leq(1-\alpha)\rho(F(x,y)-x)\leq\lim_{n\to\infty}\alpha\rho(x_{2n-1}-x_{2n})+(1-2\alpha)d=(1-\alpha)d.$$

Consequently $\rho(F(x,y)-x)=d$.

Since

$$\begin{aligned}\rho(F(y,x)-y) &\leq \lim_{n\to\infty}\rho(F(y,x)-y_{2n})=\lim_{n\to\infty}\rho(F(y,x)-G(y_{2n-1},x_{2n-1}))\\ &\leq \lim_{n\to\infty}[\alpha\rho(y-F(y,x))+\alpha\rho(y_{2n-1}-y_{2n})+(1-2\alpha)d]=d.\end{aligned}$$

Thus

$$(1-\alpha)d\leq(1-\alpha)\rho(F(y,x)-y)\leq\lim_{n\to\infty}\alpha\rho(y_{2n-1}-y_{2n})+(1-2\alpha)d=(1-\alpha)d.$$

Consequently $\rho(F(y,x)-y)=d$. □

Proof of Theorem 20. For and pair $(x_0,y_0)\in(A\times A)$ we consider the sequences $\{x_n\}$, $\{y_n\}$. From Lemma 21 we have that $\lim_{n\to\infty}\rho(x_{2n}-x_{2n+1})=d$. By Lemma 22 we have that for every $\varepsilon>0$ there exists $N_0\in\mathbb{N}$, such that there holds the inequality $\rho(x_{2m}-x_{2n+1})<d+\varepsilon$ for every $m,n\geq N_0$. From Lemma 10 by setting $z_n=x_{2n}$, $x_m=x_{2m}$ and $y_n=x_{2n+1}$ it follows there exists $N_1\in\mathbb{N}$, such that the inequality $\rho(T^{2m}x-T^{2n}x)<\varepsilon$ holds for every $m>n\geq N_1$. Therefore the sequence $\{x_{2n}\}_{n=1}^{\infty}$ is a ρ–Cauchy sequence. The proof that $\{y_{2n}\}_{n=1}^{\infty}$ is a ρ–Cauchy sequence can be made in a similar fashion.

From the completeness of L_ρ and closeness of A is follows that there are $x,y\in A$, such that $\lim_{n\to\infty}\rho(x_{2n}-x)=0$ and $\lim_{n\to\infty}\rho(y_{2n}-y)=0$. By Lemma 23 it follows that (x,y) is a coupled ρ–best proximity point.

First we will prove that if (x,y) is a coupled ρ–best proximity point then there holds

$$x=G(F(x,y),F(y,x))\ \text{ and }\ y=G(F(y,x),F(x,y)).$$

From the inequality

$$\begin{aligned}S_1 &= \rho(F(x,y)-G(F(x,y),F(y,x)))\\ &\leq \alpha\rho(x-F(x,y))+\alpha\rho(F(x,y)-G(F(x,y),F(y,x)))+(1-2\alpha)d\end{aligned}$$

we get

$$\begin{aligned}(1-\alpha)d &\le (1-\alpha)\rho(F(x,y)-G(F(x,y),F(y,x)))\\ &\le \alpha\rho(x-F(x,y))+(1-2\alpha)d=(1-\alpha)d\end{aligned}$$

and thus it follows that $\rho(F(x,y)-G(F(x,y),F(y,x)))=d$. From the equality $\rho(F(x,y)-x)=d$ and Lemma 11 it follows that $x=G(F(x,y),F(y,x))$. By similar arguments from the inequality

$$\begin{aligned}S_2 &= \rho(F(y,x)-G(F(y,x),F(x,y)))\\ &\le \alpha\rho(y-F(y,x))+\alpha\rho(F(y,x)-G(F(y,x),F(x,y)))+(1-2\alpha)d\end{aligned}$$

we get

$$\begin{aligned}(1-\alpha)d &\le (1-\alpha)\rho(F(y,x)-G(F(y,x),F(x,y)))\\ &\le \alpha\rho(y-F(y,x))+(1-2\alpha)d(A,B)=(1-\alpha)d\end{aligned}$$

and thus it follows that $\rho(F(y,x)-G(F(y,x),F(x,y)))=d$. From the equality $\rho(F(y,x)-y)=d$ and Lemma 11 it follows that $y=G(F(y,x),F(x,y))$.

Consequently $(F(x,y),F(y,x))$ is a coupled best proximity point of G in B.

It remains to prove that the coupled best proximity points is unique. Let us suppose the contrary, i.e. there exists $(u,v)\in A\times A$ such that $\rho(x-u)+\rho(y-v)>0$. We can write the inequalities

$$\begin{aligned}d &= \rho(F(x,y)-u)=\rho(F(x,y)-G(F(u,v),F(v,u)))\\ &\le \alpha\rho(x-F(x,y))+\alpha\rho(F(u,v)-G(F(u,v),F(v,u)))+(1-2\alpha)d\\ &= \alpha\rho(x-F(x,y))+\alpha\rho(F(u,v)-u)+(1-2\alpha)d=d,\end{aligned}\tag{4}$$

$$\begin{aligned}d &= \rho(F(y,x)-v)=\rho(F(y,x)-G(F(v,u),F(u,v)))\\ &\le \alpha\rho(y-F(y,x))+\alpha\rho(F(y,x)-G(F(v,u),F(u,v)))+(1-2\alpha)d\\ &= \alpha\rho(y-F(y,x))+\alpha\rho(F(y,x)-y)+(1-2\alpha)d=d,\end{aligned}\tag{5}$$

$$\begin{aligned}S_3 &= \rho(F(u,v)-x)=\rho(F(u,v)-G(F(x,y),F(y,x)))\\ &\le \alpha\rho(u-F(u,v))+\alpha\rho(F(x,y)-G(F(x,y),F(y,x)))+(1-2\alpha)d,\end{aligned}\tag{6}$$

$$\begin{aligned}S_4 &= \rho(F(v,u)-y)=\rho(F(v,u)-G(F(y,x),F(x,y)))\\ &\le \alpha\rho(v-F(v,u))+\alpha\rho(F(y,x)-G(F(y,x),F(x,y)))+(1-2\alpha)d.\end{aligned}\tag{7}$$

Therefore

$$\rho(F(x,y)-u)=\rho(F(y,x)-v)=\rho(F(u,v)-x)=\rho(F(v,u)-y)=d.$$

From

$$\rho(F(x,y)-x)=\rho(F(y,x)-y)=\rho(F(u,v)-u)=\rho(F(v,u)-v)=d.$$

and Lemma 11 it follows that $x=u$ and $y=v$.

From Lemma 23 it follows that for every $\varepsilon>0$ there exists N, such that for any $n,m\ge N$ there holds $\rho(x_{2n}-x_{2m+1})\le d+\varepsilon$. Therefore applying Lemma 10 and $\lim_{n\to\infty}\rho(x_{2n}-x_{2n+1})=d$ it follows that $\{x_{2n+1}\}_{n=1}^{\infty}$ is a Cauchy sequence. From the completeness of L_ρ and closeness of B it follows that there is $\overline{x}\in B$ such that $\lim_{n\to\infty}\rho(x_{2n+1}-\overline{x})=0$.

By similar arguments we get from Lemma 23 it follows that for every $\varepsilon > 0$ there exists N, such that for any $n,m \geq N$ there holds $\rho(y_{2n} - y_{2m+1}) \leq d + \varepsilon$. Therefore applying Lemma 10 and $\lim_{n\to\infty} \rho(y_{2n} - y_{2n+1}) = d$ it follows that $\{y_{2n+1}\}_{n=1}^{\infty}$ is a Cauchy sequence. From the completeness of L_ρ and closeness of B it follows that there is $\bar{y} \in B$ such that $\lim\limits_{n\to\infty} \rho(y_{2n+1} - \bar{y}) = 0$.

We will show that $\bar{x} = F(x,y)$ and $\bar{y} = F(y,x)$.

From $\lim_{n\to\infty} \rho(x_{2n+1} - \bar{x}) = 0$, $\lim_{n\to\infty} \rho(x_{2n+1} - x_{2n}) = d$ and Lemma 12 it follows that $\lim_{n\to\infty} \rho(x_{2n} - \bar{x}) = d$. Using the uniform continuity of ρ we get $\rho(x - \bar{x}) = \lim_{n\to\infty} \rho(x_{2n} - \bar{x}) = d$. By $\rho(x - F(x,y)) = d$ and Lemma 10 it follows that $\bar{x} = F(x,y)$.

From $\lim_{n\to\infty} \rho(y_{2n+1} - \bar{y}) = 0$, $\lim_{n\to\infty} \rho(y_{2n+1} - y_{2n}) = d$ and Lemma 12 it follows that $\lim_{n\to\infty} \rho(y_{2n} - \bar{y}) = d$. Using the uniform continuity of ρ we get $\rho(y - \bar{y}) = \lim_{n\to\infty} \rho(y_{2n} - \bar{y}) = d$. By $\rho(y - F(y,x)) = d$ and Lemma 10 it follows that $\bar{y} = F(y,x)$.

It remains to show that any iterated sequences $\{x_n\}_{n=1}^{\infty}$ and $\{y_n\}_{n=1}^{\infty}$ are ρ–convergent towards the uniques coupled best proximity points (x,y) of F in A

Indeed by taking of an arbitrary initial guess $(u_0, v_0) \in A \times A$ we get, that the iterated sequences $\{u_n\}_{n=1}^{\infty}$ and $\{v_n\}_{n=1}^{\infty}$ are ρ–convergent towards a coupled best proximity points (u,v) of F in A. From the uniqueness coupled best proximity points (x,y) of F in A it follows that $\{u_n\}_{n=1}^{\infty}$ and $\{v_n\}_{n=1}^{\infty}$ are ρ–convergent towards the unique coupled best proximity points (x,y) of F in A.□

REFERENCES:

[1] Bhaskar, T.G., V. Lakshmikantham, V. Fixed Point Theorems in Partially Ordered Metric Spaces and Applications. Nonlinear Anal., **65** (7) (2006), 1379–1393.

[2] Eldred, A. Veeramani, P. Existence and Convergence of Best Proximity Points. J. Math. Anal. Appl., **323** (2) (2006), 1001–1006. http://dx.doi.org/10.1016/j.jmaa.2005.10.081.

[3] Guo, D., Lakshmikantham, V. Coupled Fixed Points of Nonlinear Operators with Applications. Nonlinear Anal., **11** (5) (1987), 623–632. http://dx.doi.org/10.1016/0362-546X(87)90077-0.

[4] Animesh Gupta, Rajput, S.S., Kaurav, P.S. Coupled Best Proximity Point Theorem in Metric Spaces. Int. J. Anal. Appl., **4** (2) (2014), 201–215.

[5] Harjani, J., López, B., Sadarangani, K. Fixed Point Theorems for Mixed Monotone Operators and Applications to Integral Equations. Nonlinear Anal., **74** (5) (2011), 1749âĂŞ-1760.

[6] Ilchev, A., Zlatanov, B. Error Estimates for Approximation of Coupled Best Proximity Points for Cyclic Contractive Maps. Applied Mathematics and Computation, **290** (2016), 412âĂŞ-425.

[7] Ilchev, A., Zlatanov, B. Fixed and Best Proximity Points for Kannan Cyclic Contractions in Modular Function Spaces. Journal of Fixed Point Theory Applications, **19** (4) (2017), 2873âĂŞ-2893.

[8] Ilchev, A., Zlatanov, B. Coupled Fixed Points and Coupled best Proximity Points in Modular Function Spaces. International Journal of Pure and Applied Mathematics, **118** (4) (2018), 957-977.

[9] Işik, H., Türkoğlu, D. Coupled Fixed Point Theorems for New Contractive Mixed Monotone Mappings and Applications to Integral Equations. Filomat, **28** (6) (2014), 1253-âĂŞ1264.

[10] Khamsi, M.A., Kozlowski, W.M. Fixed Point Theory in Modular Function Spaces. Birhäuser Heidelberg New York Dordrecht London (2015).

[11] Khamsi, M.A., Kozlowski, W.M., Reich, S. Fixed Point Theory in Modular Function Spaces. Nonlinear Anal., **14** (11) (1990), 935–953. http://dx.doi.org/10.1016/0362-546X(90)90111-S.

[12] Khamsi, M.A., Kozlowski, W.M., Chen Shutao. Some Geometrical Properties and Fixed Point Theorems in Orlicz Spaces. J. Math. Anal. Appl., **155** (2) (1991), 393–412.

[13] Kirk, W.A., Srinivasan, P., Veeramani, P. Fixed Points for Mappings Satisfying Cyclical Contractive Conditions. Fixed Point Theory, **4** (1) (2003), 79–89.

[14] Kozlowski, W.M., Notes on modular function spaces I. Comment. Math., **28** (1) (1988), 87–100.

[15] Kozlowski, W.M., Notes on modular function spaces II. Comment. Math., **28** (1) (1988), 101–116.

[16] Kozlowski, W.M., Modular Function Spaces. Marcel Dekker Inc. New York and Basel (1988).

[17] Kozlowski, W.M., Advancements in Fixed Point Theory in Modular Function Spaces. Arabian Journal of Mathematics, **1** (4) (2012), 477–494. http://dx.doi.org/10.1007/s40065-012-0051-0.

[18] Lindenstrauss, J., Tzafriri, L., Classical Banach Spaces I. Sequence Spaces. SpringerâĂŞ-Verlag Berlin Heidelberg New York, (1977).

[19] Lindenstrauss, J., Tzafriri, L., Classical Banach Spaces II. Sequence Spaces. SpringerâĂŞ-Verlag Berlin Heidelberg New York, (1979).

[20] Musielak, J., Orlicz, W., On Modular Spaces. Stud. Math., **18** (1) (1959), 49–65.

[21] Nakano, H. Modular Semi–Ordered Spaces. Maruzen Tokyo (1950).

[22] Rao, M., Ren, Z., Theory of Orlicz Spaces. Marcel Dekker Inc. New York Basel Hong Kong (1991).

[23] Rao, M., Ren, Z., Applications of Orlicz Spaces. Marcel Dekker Inc. New York Basel (2002).

[24] Sintunavarat, W., Kumam, P., Coupled Best Proximity Point Theorem in Metric Spaces. Fixed Point Theory Appl., **2012/1/93** (2012).

[25] Zlatanov, B., Best Proximity Points in Modular Function Spaces. Arabian Journal of Mathematics, **4** (3) (2015), 215–227.

[26] Zlatanov, B., Error Estimates for Approximating of Best Proximity Points for Cyclic Contractive Maps. Carpathian J. Math., **32** (2) (2016), 265–270.

Boyan Zlatanov
Plovdiv University "Paisii Hilendarski"
E-mail: bzlatanov@gmail.com

Atanas Ilchev
Plovdiv University "Paisii Hilendarski"
E-mail: atanasilchev1@gmail.com

MATTEX 2018,
CONFERENCE PROCEEDING, v. 1, pp. 89- 99

SPECTRAL STABILITY FOR PERIODIC STANDING WAVES OF THE KLEIN-GORDON SYSTEM*

SEVDZHAN A. HAKKAEV, TURHAN M. SYULEYMANOV

ABSTRACT: *We consider spectral stability for periodic wave solutions for the coupled Klein-Gordon equation. We find conditions on parameters of the waves which imply stability and instabilty of periodic standing waves. This is achieved via the abstract stability criteria developed by [1]*

KEYWORDS: *periodic standing waves, Klein-Gordon equation*

1 Introduction

Consider the following coupled Klein-Gordon equations

$$\left| \begin{array}{l} u_{tt} - u_{xx} + u - (|u|^2 + \beta|v|^2)u = 0 \\ \\ v_{tt} - v_{xx} + v - (\beta|u|^2 + |v|^2)v = 0, \end{array} \right. \tag{1}$$

where u and v are complex valued functions, and β is real parameter. This system arise as a model of the interaction of two fields [10]. System (1) also can be interpreted as a coupled version of the Klein-Gordon equation

$$u_{tt} - u_{xx} + u - |u|^2 u = 0. \tag{2}$$

The existence and stability properties of standing wave $u(t,x) = e^{iwt}\varphi(x)$ is an important question both from theoretical and practical point of view. For the classical nonlinear Klein-Gordon equation

$$u_{tt} - \Delta u + u - |u|^p u = 0 \tag{3}$$

Shatah [8] proved the orbital stability of standing waves for $d \geq 3$ and $1 < p < 1 + \frac{4}{d}$, $w_{p,d} < |w| < 1$. Sufficient conditions for instability are given in [3, 5, 6]. The orbital instability for $d \geq 3$ and $1 + \frac{4}{d} < p < 1 + \frac{4}{d-2}$ is established in [9]. Complete characterizations of the linear stability for all values of $p > 1$, all dimensions $d \geq 1$ and all values of $w \in (-1, 1)$ is given in [12].

The stability of periodic waves have been studied extensively in the last decade. The existence and orbital stability of periodic standing waves in one dimensional case for the equation (3) was considered in [7]. The stability and instability of dnoidal and cnoidal type periodic standing waves is obtained by using the theory developed in [5, 6].

General abstract framework of spectral stability for second order Hamiltonian systems, recently developed in [1, 11, 12]. In [11, 12] is studied the stability problem of second order in time nonlinear differential equations on the hole line. In these papers, the authors applied the abstract results to the Boussinesq, Klein-Gordon, and Klein-Gordon-Zakharov equations. In [1], the authors developed the instability index theory for quadratic operator pencils. Using the abstract results of [1] in [2] is studied the spectral stability of two parametric periodic traveling-standing wave solution of the equation (3) in one dimensional case.

In this paper we are interested in spectral stability of periodic standing waves. It is well known that the spectra of linearized equation depends on the choice of function space. In the space

*This work is partially supported by Scientific Grant RD-08-119/2018 of Shumen University

of periodic functions spectrum consists of isolated eigenvalues, while in the space of bounded functions the spectrum is continuous.

Using the theory developed in [1, 11, 12] for the spectral stability of waves for the second order in time nonlinear equations, we give a complete characterization of the linear stability of dnoidal periodic standing waves for the equation (1) with respect to the perturbation of the same period.

The paper is organized as follows. In Section 2, we present the construction of our main object of study - the periodic standing waves. This is not a new material by any means, but we do it in order to single out the solutions of interest. In Section 3, we setup the spectral stability problem and we outline the theory for linearized stability for second order PDE in [1], and point out to the relevant spectral theoretic results about their linearized operators. In Section 4, we prove the main results.

2 Periodic standing waves

In this section we will construct the periodic standings waves for the system (1) in the form $u(t,x) = e^{iwt}\varphi(x)$, $v(t,x) = e^{iwt}\psi(x)$, where φ and ψ are periodic functions with period $2T$. Plugging in the system , we obtain

$$\left| \begin{array}{l} -\varphi'' + \sigma\varphi - (\varphi^2 + \beta\psi^2)\varphi = 0 \\ \\ -\psi'' + \sigma\psi - (\beta\varphi^2 + \psi^2)\psi = 0, \end{array} \right. \tag{4}$$

We now looking for periodic standing wave solutions of (4) in two cases of interest.

2.1 case $(\varphi, 0)$

In this case we look for periodic waves for the equation (4) when $\psi = 0$. Now equation (4) is reduced to the equation

$$-\varphi'' + (1 - w^2)\varphi - \varphi^3 = 0. \tag{5}$$

Integrating once the above equation, we get

$$\varphi'^2 = -\frac{1}{2}\varphi^4 + \sigma\varphi^2 + \frac{a}{2}, \tag{6}$$

or

$$\varphi'^2 = \frac{1}{2}\left(-\varphi^4 + 2\sigma\varphi^2 + a\right), \tag{7}$$

where $\sigma = 1 - w^2$ and a is a constant of integration. Let $a < 0$ and $\sigma > 0$. Denote by $\varphi_0 > \varphi_1 > 0$ the positive solutions of $-\rho^4 + 2\sigma\rho^2 + a = 0$. Then $\varphi_1 \le \varphi \le \varphi_0$ and the solution φ

$$\varphi(x) = \varphi_0 dn(\alpha x, \kappa), \tag{8}$$

where

$$\varphi_0^2 + \varphi_1^2 = 2\sigma, \;\; \alpha = \sqrt{\frac{1}{2}}\varphi_0, \;\; \kappa^2 = \frac{\varphi_0^2 - \varphi_1^2}{\varphi_0^2} = \frac{2\varphi_0^2 - 2\sigma}{\varphi_0^2}. \tag{9}$$

Since dn has a fundamental period $2K(\kappa)$, then the fundamental period of solution (8) is

$$2T = \frac{2K(\kappa)}{\alpha},\ T \in I = \left(\frac{\sqrt{2}\pi}{\sqrt{\sigma}}, \infty\right). \tag{10}$$

Here and below $K(k)$ and $E(k)$ are, as usual, the complete elliptic integrals of the first and second kind in a Legendre form.

2.2 case (φ,φ)

We consider the case $\varphi = \psi$. For the periodic function φ, we have the following ODE

$$-\varphi'' + \sigma\varphi - (\beta+1)\varphi^3 = 0. \tag{11}$$

Multiplying the equation (11) by φ' and integrating, we obtain the equation

$$\varphi'^2 = \frac{\beta+1}{2}\left[-\varphi^4 + \frac{2\sigma}{\beta+1}\varphi^2 + a\right], \tag{12}$$

where a is a constant of integration. Let $\varphi_0 > \varphi_1 > 0$ are positive roots of the polynomial $-\rho^4 + \frac{2\sigma}{\beta+1}\rho^2 + a$. Then up to translation the solution of (12) is given by

$$\varphi(x) = \varphi_0 dn(\alpha x, \kappa), \tag{13}$$

where

$$\varphi_0^2 + \varphi_1^2 = \frac{2\sigma}{\beta+1},\ \alpha = \sqrt{\frac{\beta+1}{2}}\varphi_0,\ \kappa^2 = \frac{\varphi_0^2 - \varphi_1^2}{\varphi_0^2}. \tag{14}$$

3 Linearized Equation

In this section, we set up the linearized problem for (1). We show that it reduced to the study of quadratic operator pencil, for which the general theory is developed in [1, 11, 12]. We take the perturbation in the form

$$u(t,x) = e^{iwt}(\varphi(x) + p(t,x)),\ u(t,x) = e^{iwt}(\varphi(x) + q(t,x)),$$

where $p(t,x)$ and $q(t,x)$ are complex valued functions. Plugging this ansatz in (1) and ignoring all quadratic and higher order terms yields the following *linear* equation for (p,q)

$$\begin{cases} p_{tt} + 2iwp_t - p_{xx} + \sigma p - (\beta+1)\varphi^2 p - 2\varphi^2 Rep - 2\beta\varphi^2 Req = 0 \\ q_{tt} + 2iwq_t - q_{xx} + \sigma q - (\beta+1)\varphi^2 q - 2\varphi^2 Req - 2\beta\varphi^2 Rep = 0 \end{cases} \tag{15}$$

Splitting the real and imaginary parts $p = F + iG,\ q = R + iS$ allows us to rewrite the linearized problem as the following system

$$\vec{U}_{tt} + J\vec{U}_t + \mathscr{H}\vec{U} = 0, \tag{16}$$

where $\vec{U} = (F,R,G,S)$ and

$$J = \begin{pmatrix} 0 & 0 & -2w & 0 \\ 0 & 0 & 0 & -2w \\ 2w & 0 & 0 & 0 \\ 0 & 2w & 0 & 0 \end{pmatrix} \quad \mathscr{H} = \begin{pmatrix} H_1 & -2\beta\varphi\psi & 0 & 0 \\ -2\beta\varphi\psi & H_2 & 0 & 0 \\ 0 & 0 & H_3 & 0 \\ 0 & 0 & 0 & H_4 \end{pmatrix},$$

$$\begin{aligned} H_1 &= -\partial_x^2 + \sigma - 3\varphi^2 - \beta\psi^2 \\ H_2 &= -\partial_x^2 + \sigma - \beta\varphi^2 - 3\psi^2 \\ H_3 &= -\partial_x^2 + \sigma - \varphi^2 - \beta\psi^2 \\ H_4 &= -\partial_x^2 + \sigma - \beta\varphi^2 - \psi^2. \end{aligned}$$

Note that $\mathscr{H}^* = \mathscr{H}$, while $J^* = -J$. If we consider the eigenvalue problem associated with (16), that is $\vec{U} = e^{\lambda t}\vec{V}$, we arrive at

$$\lambda^2\vec{V} + \lambda J\vec{V} + \mathscr{H}\vec{V} = 0 \tag{17}$$

Definition 1. *We say that the quadratic pencil given by the couple (J,H) is scpectrally unstable, if there exists an T periodic function $\vec{V} \in D(\mathscr{H})$ and $\lambda : \Re\lambda > 0$, so that*

$$\lambda^2\vec{V} + \lambda J\vec{V} + H\vec{V} = 0.$$

Otherwise, we say that the quadratic pencil (J,H) is stable.

3.1 Stability of quadratic pencils

We now present the theory developed in [1] for the stability of quadratic pencils, which we will be able to apply to our problem (17). We only present a corollary of the results therein, which fits our purposes. We start with some notations.

For a self-adjoint operator H, define the number of the negative eigenvalues

$$n(H) = \#\{\lambda \in (-\infty,0) \cap \sigma(H)$$

Next, for a subspace M denote $P_M : M \to M$ to be the orthogonal projection onto the subspace M and $P_M^{\perp} := Id - P_M$. For an operator T, denote $T_M := P_M T P_M$. In particular, if M is finite dimensional, with orthogonal basis $\{\chi_j\}_{j=1}^n$, P_M is given in a matrix form by $\{\langle T\chi_i, \chi_j\rangle\}_{i,j=1}^n$.

Theorem 1. *Let $H = H^*$, so that $n(H) < \infty$ and $dim(Ker(H)) < \infty$. Assume that*

1. $$(P_H^{\perp} H P_H^{\perp})^{-1}, (P_H^{\perp} H P_H^{\perp})^{-1} P_H^{\perp} J P_H^{\perp}, \tag{18}$$

   are both compact operators on L^2.
2. *There exists a positive operator S, so that $X_S = \{u : \langle Su, Su\rangle < \infty\}$ is dense in L^2.*
3. *$S^{-1}JS^{-1}, S^{-1}$ are compact, whereas $S(H+\lambda)^{-1}S$ is compact for $\lambda >> 1$.*
4. *$J : Ker(H) \to Ker(H)^{\perp}$*

5. $(Id - JH^{-1}J)|_{Ker(H)}$ *is invertible*†

Then,

$$k_r + k_c + k_- = n(H) - n((Id - JH^{-1}J)|_{Ker(H)}). \tag{19}$$

where k_r *is the number of positive solutions* λ *of* (17)*,* k_- *is the number of solutions* λ *of* (17) *with positive real part, whereas* k_c *is the total Krein index for the quadratic pencil.*

Remark: The non-negative integers k_c, k_- are even. As a consequence, if $n(H) = 1$, it follows from (19) that $k_c = k_- = 0$ and

$$k_r = 1 - n((Id - JH^{-1}J)|_{Ker(H)}). \tag{20}$$

If the right side of (19) is odd number, then $k_r \geq 1$ and hence we have instability.

4 Main Results.

We are interested in the spectral stability of periodic standing waves for the system (1). We apply the theory developed in [1] for the stability of quadratic operator pencils. We review and state main results regarding the spectral theory of Hill operators arising in the linearization around corresponding standing waves. First, we establish the spectral stability of periodic standing wave solutions $(\varphi, 0)$.

Theorem 2. *Let* $0 < \beta < 1$. *The dnoidal solution* $(\varphi, 0)$*, where* φ *described in (8), is scpectrally stable for* $|w| > \sqrt{\frac{1}{1+M(\kappa)}}$*, and unstable for* $|w| < \sqrt{\frac{1}{1+M(\kappa)}}$ *where*

$$M(\kappa) := \frac{(2-\kappa^2)[E^2(\kappa) - (1-\kappa^2)K^2(\kappa)]}{(2-\kappa^2)E^2(\kappa) - 2(1-\kappa^2)E(\kappa)K(\kappa)}.$$

For $1 < \beta < 3$ *and* $|w| > \sqrt{\frac{1}{1+M(\kappa)}}$*, the solution are unstable.*

We now give a different formulation of the main result. Let $T > \sqrt{2}\pi$. Then, the waves described in (8) are one parameter family of waves, having a fundamental period $2T$, which can be parametrized by $w : |w| < \sqrt{1 - \frac{2\pi^2}{T^2}}$, (note that w, κ are in one-to-one relation given by (10)). Now, Theorem 2 asserts that the stable waves in this family are exactly those with $|w| \geq w_T$, where $w_T \in \left(0, \sqrt{1 - \frac{2\pi^2}{T^2}}\right)$ is determined as follows. Let κ_T be the unique solution of

$$\left(1 - \frac{K(\kappa)(2-\kappa^2)}{T^2}\right)(1 + M(\kappa)) = T$$

Then, $w_T = \sqrt{\frac{1}{1+M(\kappa_T)}}$.

Proof of Theorem 2. First, we will verify that the pencil given by $(J, \mathscr{H})$ (or equivalently the eigenvalue problem (17)) satisfy the requirements 1-3 of Theorem 1. The compactness of the

†Note that this is well-defined since $J : Ker(H) \to Ker(H)^{\perp}$ and hence $JH^{-1}J$ is well-defined.

operators in (18) follows from the compactness of the embeddings $H^2[-L,L] \hookrightarrow H^1[-L,L] \hookrightarrow L^2[-L,L]$. We take the operator

$$S := \begin{pmatrix} (1+|\partial_x|)^{3/4} & 0 & 0 & 0 \\ 0 & (1+|\partial_x|)^{3/4} & 0 & 0 \\ 0 & 0 & (1+|\partial_x|)^{3/4} & 0 \\ 0 & 0 & 0 & (1+|\partial_x|)^{3/4} \end{pmatrix}.$$

Clearly $X_S = H^{3/4}[-L,L] \times H^{3/4}[-L,L] \times H^{3/4}[-L,L] \times H^{3/4}[-L,L]$ is dense in $L^2 \times L^2 \times L^2 \times L^2$. In addition, $S^{-1}JS^{-1}$ is smoothing of order $1/2$, S^{-1} smoothing of order $3/2$, whereas $S(H+\lambda)^{-1}S, \lambda >> 1$ is smoothing of order $1/2$. Thus, all the operators in question are compact in their action on $L^2 \times L^2 \times L^2 \times L^2$.

Next, we will consider the spectral properties of the operator of linearization. Note that in this case the operator of linearization $\mathscr{H}$ is in the form

$$\mathscr{H} = \begin{pmatrix} Q_1 & 0 & 0 & 0 \\ 0 & Q_2 & 0 & 0 \\ 0 & 0 & Q_3 & 0 \\ 0 & 0 & 0 & Q_2 \end{pmatrix},$$

where

$$Q_1 = -\partial_x^2 + \sigma - 3\varphi^2$$
$$Q_2 = -\partial_x^2 + \sigma - \beta\varphi^2$$
$$Q_3 = -\partial_x^2 + \sigma - \varphi^2.$$

These operators are self-adjoint acting on $L^2_{per}[0,2T]$ with domain $H^2_{per}[0,2T]$. From the Floquet theory, it follows that its spectrum is purely discrete

$$\nu_0 < \nu_1 \leq \nu_2 < \nu_3 \leq \nu_4 < \dots$$

where ν_0 is always a simple eigenvalue. If $\phi_n(x)$ is the eigenfunction corresponding to ν_n, then

$$\begin{array}{l} \phi_0 \text{ has no zeroes in } [0,2T]; \\ \phi_{2n+1},\ \phi_{2n+2} \text{ have each just } 2n+2 \text{ zeroes in } [0,2T). \end{array}$$

Using (8) and (9), and taking $y = \alpha x$ as an independent variable in Q_1, one obtains $Q_1 = \alpha^2 \Lambda_1$ with an operator Λ_1 in $[0,2K(k)]$ given by

$$\Lambda_1 = -\frac{d^2}{dy^2} + 6k^2 sn^2(y;k) - 4 - \kappa^2,$$

where we have used the relation $dn^2(x;\kappa) + \kappa^2 sn^2(x,\kappa) = 1$. The operator Λ_1 is Hill's operator with Lamé potential. It is well-known that the first five eigenvalues of $\Lambda_2 = -\partial_y^2 + 6k^2 sn^2(y,k)$, with periodic boundary conditions on $[0,4K(k)]$ are simple. These eigenvalues and corresponding eigenfunctions are:

$$\begin{array}{ll} \nu_0 = 2 + 2k^2 - 2\sqrt{1-k^2+k^4}, & \phi_0(y) = 1 - (1+k^2 - \sqrt{1-k^2+k^4}) sn^2(y,k), \\ \nu_1 = 1 + k^2, & \phi_1(y) = cn(y,k)dn(y,k) = sn'(y,k), \\ \nu_2 = 1 + 4k^2, & \phi_2(y) = sn(y,k)dn(y,k) = -cn'(y,k), \\ \nu_3 = 4 + k^2, & \phi_3(y) = sn(y,k)cn(y,k) = -k^{-2}dn'(y,k), \\ \nu_4 = 2 + 2k^2 + 2\sqrt{1-k^2+k^4}, & \phi_4(y) = 1 - (1+k^2 + \sqrt{1-k^2+k^4}) sn^2(y,k). \end{array}$$

Since the eigenvalues of Q_1 and Λ_1 are related by $\lambda_n = \alpha^2 \nu_n$, it follows that the first three eigenvalues of the operator Q_1, equipped with periodic boundary condition on $[0,2K(k)]$ are simple and $\lambda_0 < 0, \lambda_1 = 0, \lambda_2 > 0$. The corresponding eigenfunctions are $\phi_0(\alpha x), \phi_1(\alpha x) = const.\varphi'$ and $\phi_2(\alpha x)$.

From (11), we have $Q_3\varphi = 0$ and since $\varphi > 0$ it follows that zero is the first eigenvalue of operator Q_3, which is simple.

First we consider the case $0 < \beta < 1$. We have $Q_2 > Q_3$, and therefore operator Q_2 is strong positive and $KerQ_2 = \{\varnothing\}$.

Hence, $n(\mathscr{H}) = 1$, $dimKer\mathscr{H} = 2$, and $\vec{\psi}_0 = \frac{1}{||\varphi||}(0,0,\varphi,0)$, $\vec{\psi}_1 = \frac{1}{||\varphi'||}(\varphi',0,0,0) \in Ker\mathscr{H}$.

The anti-selfadjointness of J yields $\langle J\psi_i, \psi_i\rangle = 0, i = 1,2$. By direct computations $\langle J\psi_i, \psi_j\rangle = 0,\ \ i,j = 1,2$. With this, we have verified that $J : Ker(\mathscr{H}) \to Ker(\mathscr{H})^{\perp}$. Thus, in order to determine the stability one needs to compute the index $n((Id - J\mathscr{H}^{-1}J)|_{Ker(\mathscr{H})})$. For the $(Id - J\mathscr{H}^{-1}J)|_{Ker(\mathscr{H})}$ we have the following matrix representation

$$U := \begin{pmatrix} \langle (Id - J\mathscr{H}^{-1}J)\vec{\psi}_0, \psi_0\rangle & \langle (Id - J\mathscr{H}^{-1}J)\vec{\psi}_0, \psi_1\rangle \\ \langle (Id - J\mathscr{H}^{-1}J)\vec{\psi}_1, \psi_0\rangle & \langle (Id - J\mathscr{H}^{-1}J)\vec{\psi}_1, \psi_1\rangle \end{pmatrix}$$

We have

$$\mathscr{H}^{-1}J\vec{\psi}_0 = \frac{1}{||\varphi||}\begin{pmatrix} -2wQ_1^{-1}\varphi \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathscr{H}^{-1}J\vec{\psi}_1 = \frac{1}{||\varphi'||}\begin{pmatrix} 0 \\ 0 \\ Q_3^{-1}\varphi' \\ 0 \end{pmatrix}.$$

Thus

$$U = \begin{pmatrix} 1 + \frac{4w^2}{||\varphi||^2}\langle Q_1^{-1}\varphi, \varphi\rangle & 0 \\ 0 & 1 + \frac{4w^2}{||\varphi'||^2}\langle Q_3^{-1}\varphi', \varphi'\rangle \end{pmatrix}.$$

Since the operator Q_3 is non-negative, then $\langle Q_3^{-1}\varphi', \varphi'\rangle \geq 0$.

Hence, the stability criteria is as follows

- Stability, if $1 + \frac{4w^2}{||\varphi||^2}\langle Q_1^{-1}\varphi, \varphi\rangle < 0$ (one negative and positive eigenvalues of U, $n((Id - J\mathscr{H}^{-1}J)|_{Ker(\mathscr{H})}) = 1$)

-Instability, if $1 + \frac{4w^2}{||\varphi||^2}\langle Q_1^{-1}\varphi, \varphi\rangle > 0$ (two positive eigenvalues of U, $n((Id - J\mathscr{H}^{-1}J)|_{Ker(\mathscr{H})}) = 0$)

Thus it remains to compute $\langle Q_1^{-1}\varphi, \varphi\rangle$. We have $Q_1\varphi' = 0$. The function

$$\psi(x) = \varphi'(x)\int^x \frac{1}{\varphi'^2(s)}ds, \quad \begin{vmatrix} \varphi' & \psi \\ \varphi'' & \psi' \end{vmatrix} = 1$$

is also solution of $Q_1\psi = 0$. Formally, since φ' has zeros using the identities

$$\frac{1}{cn^2(y,\kappa)} = \frac{1}{dn(y,\kappa)}\frac{\partial}{\partial_y}\frac{sn(x,\kappa)}{cn(y,\kappa)}, \quad \frac{1}{sn^2(y,\kappa)} = -\frac{1}{dn(y,\kappa)}\frac{\partial}{\partial_y}\frac{cn(x,\kappa)}{sn(y,\kappa)}$$

and integrating by parts we get

$$\psi(x) = \frac{1}{\alpha^2\kappa^2\varphi_0}\left[\frac{1 - 2sn^2(\alpha x,\kappa)}{dn(\alpha x,\kappa)} - \alpha\kappa^2 sn(\alpha x,\kappa)cn(\alpha x,\kappa)\int_0^x \frac{1 - 2sn^2(\alpha s,\kappa)}{dn^2(\alpha s,\kappa)}ds\right]$$

Thus, we may construct Green function

$$Q_1^{-1} f = \varphi' \int_0^x \psi(s) f(s) ds - \psi(s) \int_0^x \varphi'(s) f(s) s + C_f \psi(x),$$

where C_f is chosen such that $Q_1^{-1} f$ is periodic with same period as $\varphi(x)$.

After integrating by parts, we get

$$\langle Q_1^{-1}\varphi, \varphi \rangle = -\langle \varphi^3, \psi \rangle + \frac{\varphi^2(T) + \varphi(0)^2}{2} \langle \varphi, \psi \rangle + C_\varphi \langle \varphi, \psi \rangle, \tag{21}$$

We have

$$\begin{aligned} &\langle \varphi, \psi \rangle = \tfrac{1}{\alpha^3 \kappa^2} [E(\kappa) - K(\kappa)] \\ &\langle \varphi^3, \psi \rangle = \tfrac{\varphi_0^2}{2\alpha^3 \kappa^2} [(2-\kappa^2)E(\kappa) - 2(1-\kappa^2)K(\kappa)] \\ &C_\varphi = -\tfrac{\varphi''(T)}{2\psi'(T)} \langle \varphi, \psi \rangle + \tfrac{\varphi^2(T) - \varphi^2(0)}{2}. \end{aligned} \tag{22}$$

With this finally we get

$$\langle Q_1^{-1}\varphi, \varphi \rangle = \frac{\varphi_0^2}{2\alpha^3\kappa^2} \frac{E^2(\kappa) - (1-\kappa^2)K^2(\kappa)}{2(1-\kappa^2)K(\kappa) - (2-\kappa^2)E(\kappa)} < 0.$$

We have

$$||\varphi||^2 = \int_0^T \varphi_0^2 dn^2(\alpha x, \kappa) dx = \frac{2\varphi_0^2}{\alpha} \int_0^{K(\kappa)} dn^2(y,\kappa) dy = \frac{2\varphi_0^2}{\alpha} E(\kappa)$$

Hence

$$1 + \frac{4w^2}{||\varphi||^2} \langle Q_1^{-1}\varphi, \varphi \rangle = 1 - \frac{w^2}{1-w^2} \frac{(2-\kappa^2)[E^2(\kappa) - (1-\kappa^2)K^2(\kappa)]}{(2-\kappa^2)E^2(\kappa) - 2(1-\kappa^2)E(\kappa)K(\kappa)}. \tag{23}$$

Thus, the standing waves are stable if $|w| > \frac{1}{\sqrt{(1+M(\kappa)}}$ and unstable if $|w| < \frac{1}{\sqrt{(1+M(\kappa)}}$.

Now if $1 < \beta < 3$, then $Q_1 < Q_2 < Q_3$. From Comparison Theorem, $n(Q_2) = 1$ and $KerQ_2 = \{\emptyset\}$. Hence, $n(\mathscr{H}) = 2$ and for $|w| < \frac{1}{\sqrt{(1+M(\kappa)}}$, $n(\mathscr{H}) - n(U) = 1$. $\square$

Now we will consider the spectral stability of periodic standing waves of the form (φ, φ), where φ is given by (13). We have the following result.

Theorem 3. *Let $\beta > 1$. The two parameter family of dnoidal solutions (φ, φ), where φ described in (13), are spectrally stable if $|w| > \sqrt{\frac{1}{1+M(\kappa)}}$, and unstable if $|w| < \sqrt{\frac{1}{1+M(\kappa)}}$, where*

$$M(\kappa) := \frac{(2-\kappa^2)[E^2(\kappa) - (1-\kappa^2)K^2(\kappa)]}{(2-\kappa^2)E^2(\kappa) - 2(1-\kappa^2)E(\kappa)K(\kappa)}.$$

For $0 < \beta < 1$ and $|w| > \sqrt{\frac{1}{1+M(\kappa)}}$, the solutions (φ, φ) are unstable.

Proof. Similarly as in the Theorem 2, we have that the pencil given by $(J,\mathscr{H})$ satisfy the requirements 1-3 of Theorem 1.

Note that in the case $\varphi=\psi$ the operator $\mathscr{H}$ can be diagonalized. Let

$$B=\begin{pmatrix} 1 & 1 & 0 & 0\\ -\frac{1}{2} & \frac{1}{2} & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{pmatrix}.$$

Then

$$\mathscr{H}_D=B\mathscr{H}B^{-1}=\begin{pmatrix} L_1 & 0 & 0 & 0\\ 0 & L_2 & 0 & 0\\ 0 & 0 & L_3 & 0\\ 0 & 0 & 0 & L_3\end{pmatrix},$$

where

$$\begin{aligned} L_1&=-\partial_x^2+\sigma-3(\beta+1)\varphi^2\\ L_2&=-\partial_x^2+\sigma-(3-\beta)\varphi^2\\ L_3&=-\partial_x^2+\sigma-(\beta+1)\varphi^2.\end{aligned}$$

Next, we list the spectral properties of the operator $\mathscr{H}$ and $\mathscr{H}_D$. Since $\mathscr{H}_D$ is a diagonal operator, with entries $L_i,\ i=1,2,3$, we have that $\sigma(\mathscr{H}_D)=\sigma(L_1)\cup\sigma(L_2)\cup\sigma(L_3)$. The operators $L_i,\ i=1,2,3$ are clearly Hill operators, so that the standard Floquet theory is applicable to them.

Using that $\kappa^2sn^2(y)+dn^2(y)=1$ and formulas (14), we obtain

$$\begin{aligned} L_1\ &=-\partial_x^2+\sigma-3(\beta+1)\varphi_0^2dn^2(\alpha x,\kappa)\\ &=\alpha^2\left[-\partial_y^2+6\kappa^2sn^2(y,\kappa)-(4+\kappa^2)\right],\end{aligned}$$

where $y=\alpha x$.

It follows that the first three eigenvalues of the operator L_1, equipped with periodic boundary condition on $[0,2K(k)]$ are simple.

From (11), we have that $L_3\varphi=0$ and since $\varphi>0$, then zero is the first eigenvalue, which is simple with corresponding eigenfunction φ. Moreover, $L_2=L_3+2(\beta-1)\varphi^2$ and for $\beta>1$, the operator L_2 is strong positive and $KerL_2=\{\emptyset\}$.

Hence, $n(\mathscr{H})=1$, $dimKer\mathscr{H}=3$, and $\Psi_0=\frac{1}{\sqrt{2}||\varphi'||}(\varphi',\varphi',0,0), \Psi_1=\frac{1}{||\varphi||}(0,0,\varphi,0), \Psi_2=\frac{1}{||\varphi||}(0,0,0,\varphi)\in Ker\mathscr{H}$.

The anti-selfadjointness of J yields $\langle J\psi_i,\psi_i\rangle=0, i=1,2$. Moreover, by direct computations $\langle J\psi_i,\psi_j\rangle=0,\ i,j=1,2$. With this, we have verified that $J:Ker(\mathscr{H})\to Ker(\mathscr{H})^{\perp}$.

No we will estimate the index $n((Id-J\mathscr{H}^{-1}J)|_{Ker(H)})$. For the $(Id-J\mathscr{H}^{-1}J)|_{Ker(H)}$ we have the following matrix representation

$$U:=\begin{pmatrix} \langle(Id-J\mathscr{H}^{-1}J)\Psi_0,\Psi_0\rangle & \langle(Id-J\mathscr{H}^{-1}J)\Psi_0,\Psi_1\rangle & \langle(Id-J\mathscr{H}^{-1}J)\Psi_0,\Psi_2\rangle\\ \langle(Id-J\mathscr{H}^{-1}J)\Psi_1,\Psi_0\rangle & \langle(Id-J\mathscr{H}^{-1}J)\Psi_1,\Psi_1\rangle & \langle(Id-J\mathscr{H}^{-1}J)\Psi_1,\Psi_2\rangle\\ \langle(Id-J\mathscr{H}^{-1}J)\Psi_2,\Psi_0\rangle & \langle(Id-J\mathscr{H}^{-1}J)\Psi_2,\Psi_1\rangle & \langle(Id-J\mathscr{H}^{-1}J)\Psi_2,\Psi_2\rangle\end{pmatrix}$$

We have

$$\mathscr{H}^{-1}J\Psi_0=\frac{\sqrt{2}w}{||\varphi||}\begin{pmatrix} 0\\ 0\\ L_3^{-1}\varphi'\\ L_3^{-1}\varphi'\end{pmatrix}$$

$$\mathscr{H}^{-1}J\Psi_1 = -\frac{w}{||\varphi'||}\begin{pmatrix} L_1^{-1}\varphi + L_2^{-1}\varphi \\ L_1^{-1}\varphi - L_2^{-1}\varphi \\ 0 \\ 0 \end{pmatrix},$$

$$\mathscr{H}^{-1}J\Psi_2 = -\frac{w}{||\varphi||}\begin{pmatrix} L_1^{-1}\varphi - L_2^{-1}\varphi \\ L_1^{-1}\varphi + L_2^{-1}\varphi \\ 0 \\ 0 \end{pmatrix}.$$

Thus

$$U = \begin{pmatrix} 1+C & 0 & 0 \\ 0 & 1+A+B & A-B \\ 0 & A-B & 1+A+B \end{pmatrix},$$

where

$$A = \frac{2w^2}{||\varphi||^2}\langle L_1^{-1}\varphi, \varphi\rangle, \ B = \frac{2w^2}{||\varphi||^2}\langle L_2^{-1}\varphi, \varphi\rangle, \ C = \frac{2w^2}{||\varphi'||^2}\langle L_3^{-1}\varphi', \varphi'\rangle$$

The eigenvalues of the matrix U are $\rho_1 = 1+2A$, $\rho_2 = 1+2B$, $\rho_3 = 1+C$. Since, L_2 is non-negative and L_3 is strong positive, then $\rho_2 = 1+2B = 1+\frac{4w^2}{||\varphi||^2}\langle L_2^{-1}\varphi, \varphi\rangle > 0$ and $\rho_3 = 1+C = 1+\frac{2w^2}{||\varphi'||^2}\langle L_3^{-1}\varphi', \varphi'\rangle > 0$.

Hence, the stability criteria is as follows

- Stability, if $1+2A = 1+\frac{4w^2}{||\varphi||^2}\langle L_1^{-1}\varphi, \varphi\rangle < 0$ (one negative and tow positives eigenvalues of U, $n((Id - J\mathscr{H}^{-1}J)|_{Ker(\mathscr{H})}) = 1$)

-Instability, if $1+2A = 1+\frac{4w^2}{||\varphi||^2}\langle L_1^{-1}\varphi, \varphi\rangle > 0$ (three positive eigenvalues of U, $n((Id - J\mathscr{H}^{-1}J)|_{Ker(\mathscr{H})}) = 0$)

Similarly as in the previous case, we get

$$\rho_1 = 1 - \frac{w^2}{1-w^2}\frac{(2-\kappa^2)E^2(\kappa) - (1-\kappa^2)(2-\kappa^2)K^2(\kappa)}{(2-\kappa^2)E^2(\kappa) - 2(1-\kappa^2)E(\kappa)K(\kappa)}.$$

Thus, the standing waves are stable if $|w| > \sqrt{\frac{1}{1+M(\kappa)}}$ and unstable if $|w| < \sqrt{\frac{1}{1+M(\kappa)}}$.

Now if $0 < \beta < 1$, then $L_1 < L_2 < L_3$. From the Comparison Theorem, we have $n(L_2) = 1$ and $KerL_2 = \{\emptyset\}$, and $n(\mathscr{H}) - n(U) = 1$. $\square$

REFERENCES:

[1] J. Bronski, M. Johnson, T. Kapitula, *An instability index theory for quadratic pencils and applications*, to appear, Comm. Math. Phys. (2014).

[2] A. Demirkaya, S. Hakkaev, M. Stanislavova, A. Stefanov, *On the spectral stability of periodic waves of the Klein-Gordon equation*, Differential and Integral Equations, **28** (2015), 431-454

[3] M. Grillakis, *Linearized instability for nonlinear Schrodinger and Klein-Gordon equations*, Comm. Pure Appl. Math., **41**(1988), no.6, 747-774.

[4] M. Grillakis, *Analysis of the linearization around a critical point of an infinite-dimensional Hamiltonian system*, Comm. Pure Appl. Math. **43** (1990), p. 299–333.

[5] M. Grillakis, J. Shatah, W. Strauss, *Stability of solitary waves in the presence of symmetry* I, J. Funct. Anal. **74** (1987), 160–197.

[6] M. Grillakis, J. Shatah, W. Strauss, *Stability of solitary waves in the presence of symmetry* II, J. Funct. Anal. **94** (1990), 308–348.

[7] F. Natali, A. Pastor, *Stability and instability of periodic standing wave solutions for some Klein-Gordon equations*, J. Math. Anal. Appl., **347**(2008), p. 428-441

[8] J. Shatah, *Stable standing waves of nonlinear Klein-Gordon equations*, Comm. Math. Phys., **91**(1983), no.3, p. 313–327.

[9] J. Shatah, *Unstable ground states on nonlinear Klein-Gordon equations*, Trans. Amer. Math. Soc., **290**(1985), no.2, 701–710.

[10] I. Segal, *Nonlinear partial differential equations in Quantum Field Theory*, Pros. Symp. Appl. Math. A.M.S. **17**(1965), 210-226

[11] M. Stanislavova, A. Stefanov, *Linear stability analysis for traveling waves of second order in time PDE'S*, Nonlinearity, **25** (2012), p. 2625–2654.

[12] M. Stanislavova, A. Stefanov, *Spectral stability analysis for special solutions of second order in time PDE's: the higher dimensional case*, Physica D, **262** (2013), p. 1–13.

Sevdzhan Hakkaev
Konstantin Preslavsky University of Shumen
E-mail: s.hakkaev@shu.bg

Turhan Syuleymanov
Konstantin Preslavsky University of Shumen
E-mail: turhan@abv.bg

PROPERTIES OF MULTIPLICATIVE INTEGRALS III*

GALINA S. BORISOVA, RADOSTINA V. POPOVA, NINA S. SHUMAN, GERGANA A. TODOROVA

ABSTRACT: *In this paper we present some properties of special kind of multiplicative integrals playing an essential role in obtaining of the asymptotics of nondissipative curves, generated by unbounded nondissipative operators A with different domains of A and its adjoint A^* in a Hilbert space.*

KEYWORDS: *Nonselfadjoint operator, unbounded operator, dissipative operator, operator colligation, triangular model, coupling, multiplicative integral*

1 Introduction

In this paper we continue the presentation of some properties of special kind of multiplicative integrals and this paper is a continuation of the papres [2], [3]. These properties of multiplcative integrals play an important role in the further development of the investigations of nonselfadjoint unbounded operators (with finite dimensional imaginary parts) based on the theory of the characteristic operator functions and the triangular models of M.S. Livšic. The presented properties are inequalities concerning the multiplicative integrals which are so-called limit values of multiplicative integrals connected with nonselfadjoint unbounded K^r- operators A in a Hilbert space H with differen domains of A and its adjoint A^* and presented as a regular coupling of dissipative and antidissipative operators with real absolutely continuous spectra. The triangular model of the regular couplings of this class of nonselfadjoint operators has been introduced and investigated by K.P. Kirchev and G.S. Borisova in [6, 7, 8]. In the course of investigations of these operators A (the characteristic operator functions, the resolvent of A, the asymptotic behaviour of the corresponding continuous curves) we use the properties of the multiplicative integrals from the form

$$\overset{\rightarrow}{\int_a^b} e^{-i\frac{1+\lambda\alpha(v)}{\alpha(v)-\lambda}T(v)dv}, \tag{1.1}$$

($\lambda \in \mathbb{C}$, $\lambda \neq \alpha(v)$ for $v \in [a,b]$), where $\alpha(v)$ is a nondecreasing real function in $[a,b]$, $T(v)$ is a measurable nonnegative $m\times m$ matrix function, satisfying the conditions

$$\int_a^b trT(v)dv < +\infty; \quad \int_a^b ||T(v)||dv < +\infty. \tag{1.2}$$

The integral (1.1) is the multiplicative Stieltjes integral, defined as

$$\begin{aligned}\overset{\rightarrow}{\int_a^b} e^{f(t)G(t)dt} &= \lim_{\max\triangle\theta_k\to 0} \overset{\rightarrow}{\prod_{k=1}^{n}} e^{f(\tau_k)(E(\theta_k)-E(\theta_{k-1}))} = \\ &= e^{f(\tau_1)(E(\theta_1)-E(\theta_0))}e^{f(\tau_2)(E(\theta_2)-E(\theta_1))}\dots e^{f(\tau_n)(E(\theta_n)-E(\theta_{n-1}))},\end{aligned} \tag{1.3}$$

where $E(\theta) = \int_a^\theta G(t)dt$ and the limit in (1.3) is taken over all the partitions $a = \theta_0 < \theta_1 < \ldots < \theta_n = b$ of the interval $[a,b]$ and all the choices of intermediate points τ_k such that $\theta_{k-1} \le \tau_k \le \theta_k$ $(k = 1,2,\ldots,n)$, $G(\theta)$ is integrable matrix function on $[a,b]$ and $||E(\theta') - E(\theta'')|| \le |\theta' - \theta''|$.

*Partially supported by Scientific Research Grant RD-08-119/2018 of Shumen University.

The limits of the multiplicative integrals from the form (in the sense of a strong limit)

$$\begin{aligned} s-\lim_{\delta\to 0}\overset{\overset{b}{\rightarrow}}{\underset{a}{\int}} e^{\frac{-iT(v)}{v-(x\pm i\delta)}dv} &= \\ = s-\lim_{\delta\to 0}\overset{\overset{x-\delta}{\rightarrow}}{\underset{a}{\int}} e^{\frac{-iT(v)}{v-x}dv} e^{\pm\pi T(x)} \overset{\overset{b}{\rightarrow}}{\underset{x+\delta}{\int}} e^{\frac{-iT(v)}{v-x}dv} \end{aligned} \tag{1.4}$$

are used essentially in the case of bounded dissipative operators [10], the case of bounded nonselfadjoint operators, presented as a coupling of dissipative and antidisipative operators [5], the case of unbounded nonselfadjoint operators, presented as a coupling of dissipative and antidissipative operators and with equal domains of the operator and its adjoint [7]. The equality (1.4) is proved by L.A. Sakhnovich in [10] and it is an analogue for the multiplicative integrals of the well-known Privalov's theorem [9] for the limit values for the integral

$$f(\lambda)=\int_a^b \frac{p(t)}{t-\lambda}dt$$

in the scalar case.

The limits from the form (1.4) have been used for obtaining of the asymptotics of the so-called continuous curves (or the processes) $e^{itA}f$ as $t\to\pm\infty$, where A is a nonselfadjoint operator in a Hilbert space H, $f\in H$. For a dissipative operator A (i.e. the imaginary part of the operator A is nonnegative operator) immediately follows that there exists the limit $\lim\limits_{t\to+\infty}(e^{itA}f,e^{itA}f)$ ($f\in H$). The explicit form of these limitis and the strong limits $s-\lim\limits_{t\to+\infty} e^{-itA^*}e^{itA}$ has been obtained (with the help of the limits (1.4)) in [5], [4], [7] in the the case of a dissipative bounded operator A, in the case of bounded nonselfadjoint operators A, presented as a coupling of dissipative and antidisipative operators [5], in the case of unbounded nonselfadjoint operators A, presented as a coupling of dissipative and antidissipative operators and with equal domains of the operator and its adjoint.

In the case of considered unbounded K^r- operators in the course of obtaining the asymptotics of the corresponding nondissipative curves we essentially use the existence and the form of the limit values of the multiplicative integrals (in the sense of a strong limit) for almost all $x\in[a,b]$

$$s-\lim_{\delta\to 0}\overset{\overset{b}{\rightarrow}}{\underset{a}{\int}} e^{-i\frac{1+(x\pm i\delta)v}{v-(x\pm i\delta)}T(v)dv},\quad \delta>0. \tag{1.5}$$

2 Preliminary results

At first we will remind a proposition, obtained in [10], and the form of the limit (1.5), obtained and presented in [1], [7], [8], which we will use in this paper.

Theorem 2.1. *([10]) Let $m\times m$ matrix functions ($m\le\infty$) $T_1(t)$ and $T_2(t)$ are integrable in $[a,b]$ and for almost all t in $[a,b]$ satisfy the inequalities*

$$\frac{T_k(t)-T_k^*(t)}{i}\le 0,\qquad k=1,2.$$

Then

$$\left\| \overset{\curvearrowright}{\int_a^b} e^{-iT_1(t)dt} - \overset{\curvearrowright}{\int_a^b} e^{-iT_2(t)dt} \right\| \le \int_a^b ||T_1(t) - T_2(t)||dt.$$

Theorem 2.2. *([1], [7], [8]) Let the matrix function $T(x)$ is integrable and nonnegative in $[a,b]$. Then for almost all $x \in \mathbb{R}$ there exist the limits (1.5) and they have the form*

$$\begin{gathered} s - \lim_{\delta \to 0} \overset{\rightarrow}{\int_a^b} e^{-i\frac{1+(x\pm i\delta)v}{v-(x\pm i\delta)}T(v)dv} = \\ = s - \lim_{\varepsilon \to 0} \overset{\rightarrow}{\int_a^{x-\varepsilon}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{\pm\pi(1+x^2)T(x)} \overset{\rightarrow}{\int_{x+\varepsilon}^{b}} e^{-i\frac{1+vx}{v-x}T(v)dv} \end{gathered} \tag{2.1}$$

($\delta > 0$, $\varepsilon > 0$).

Let now $\alpha(x)$ be a nondecreasing unbounded real function, defined in (a,b) $(-\infty \le a < b \le +\infty)$. Let $\Pi(x)$ be a measurable $n \times m$ $(1 \le n \le m, r \le m)$ matrix function whose rows are linearly independent on each point of a set with a positive measure and satisfying the conditions

$$\int_a^b tr\, B(x)dx < +\infty, \quad \int_a^b ||\Pi(x)||^2 dx < +\infty, \tag{2.2}$$

where $B(x) = \Pi^*(x)\Pi(x)$.

Let $L : \mathbb{C}^m \longrightarrow \mathbb{C}^m$, $L^* = L$, $\det L \neq 0$. Without loss of generality we can suppose that L has the representation

$$L = J_1 - J_2 + S + S^*, \tag{2.3}$$

where $J_1, J_2, S, S^* : \mathbb{C}^m \longrightarrow \mathbb{C}^m$,

$$J_1 = \begin{pmatrix} I_{r_1} & 0 \\ 0 & 0 \end{pmatrix}, \; J_2 = \begin{pmatrix} 0 & 0 \\ 0 & I_{m-r_1} \end{pmatrix}, \; S = \begin{pmatrix} 0 & 0 \\ \widehat{S} & 0 \end{pmatrix}, \tag{2.4}$$

I_k is the identity matrix in $\mathbb{C}^k$ $(k = r_1, m - r_1)$, $\widehat{S}$ is a $(m - r_1) \times r_1$ matrix, r_1 is the number of the positive eigenvalues and $m - r_1$ is the number of the negative eigenvalues of the matrix L.

Let the matrix function $B(x)$ satisfy also the conditions

$$B(x)J_1 = J_1 B(x)$$

and $\alpha(x)B(x)J_2$ be an integrable matrix function on (a,b) $(-\infty \le a < b \le +\infty)$. Let us consider a Hilbert space $\mathbf{L}^2(a,b;\mathbb{C}^n)$, whose elements are vector functions $f(x)$ from the form

$$f(x) = (f_1(x), f_2(x), \dots, f_n(x)), \quad (f_k \in \mathbf{L}^2(a,b)).$$

The scalar product in $\mathbf{L}^2(a,b;\mathbb{C}^n)$ is defined by the formula

$$(f,g) = \int_a^b f(x)g^*(x)dx.$$

We will denote by $\| \ \|$ the norm of an operator function in $\mathbb{C}^n$ and by $\| \ \|_{\mathbf{L}^2}$- the norm in $\mathbf{L}^2(a,b;\mathbb{C}^n)$.

Let $Q(x)$ be a measurable matrix function on (a,b) satisfying the condition

$$\Pi(x)Q(x) = I \tag{2.5}$$

for almost all $x \in (a,b)$. Then the operators P_1 and P_2, defined by the equalities

$$P_1 f(x) = f(x)\Pi(x)J_1 Q(x), \quad P_2 f(x) = f(x)\Pi(x)J_2 Q(x)$$

onto $\mathbf{L}^2(a,b;\mathbb{C}^n)$, are orthogonal projectors in $\mathbf{L}^2(a,b;\mathbb{C}^n)$.

The model A, describing the class of K^r-operators presented as a coupling of dissipative and antidissipative operators with real absolutely continuous spectra and with different domains of A and A^* has been introduced in [6] and has the form

$$\begin{aligned} Af(x) = AGg(x) &= \alpha(x)g(x) + \\ &+ i\int_a^x g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_1 \overset{\rightarrow}{\int_\xi^x} e^{i\alpha(v)B_1(v)dv} J_1\Pi^*(x)(\alpha(x)-i)d\xi - \\ &- i\int_a^x g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_2 \overset{\rightarrow}{\int_\xi^x} e^{-i\alpha(v)B_2(v)dv} J_2\Pi^*(x)(\alpha(x)-i)d\xi + \\ &+ \int_a^b g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_2 \overset{\rightarrow}{\int_\xi^b} e^{-i\alpha(v)B_2(v)dv} d\xi S \overset{\rightarrow}{\int_a^x} e^{B_1(v)dv} J_1\Pi^*(x), \end{aligned} \tag{2.6}$$

where G is an invertible operator in $\mathbf{L}^2(\mathbb{R};\mathbb{C}^n)$ and

$$G = I + P_1KP_2, \tag{2.7}$$

$$\begin{aligned} KP_2 g(x) &= \\ &= -i\int_a^b g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_2 \overset{\rightarrow}{\int_\xi^b} e^{-i\alpha(v)B_2(v)dv} d\xi S \overset{\rightarrow}{\int_a^x} e^{B_1(v)dv} J_1\Pi^*(x), \end{aligned} \tag{2.8}$$

$B_1(x) = B(x)J_1$, $B_2(x) = B(x)J_2$, for each $f(x) = Gg(x) \in \mathbf{L}^2(\mathbb{R};\mathbb{C}^n)$ such that $Af \in \mathbf{L}^2(\mathbb{R};\mathbb{C}^n)$. Using the form of the projectors P_1, P_2 the model A, defined by (2.6), takes the the form

$$Af(x) = AGg(x) = P_1AP_1g(x) + P_2AP_2g(x) + iKP_2g(x), \tag{2.9}$$

where K is defined by (2.8) and

$$\begin{aligned} P_1AP_1g(x) &= \alpha(x)g(x)\Pi(x)J_1Q(x) + \\ &+ i\int_a^x g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_1 \overset{\rightarrow}{\int_\xi^x} e^{i\alpha(v)B_1(v)dv} J_1\Pi^*(x)(\alpha(x)-i)d\xi, \\ P_2AP_2g(x) &= \alpha(x)g(x)\Pi(x)J_2Q(x) - \\ &- i\int_a^x g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_2 \overset{\rightarrow}{\int_\xi^x} e^{-i\alpha(v)B_2(v)dv} J_2\Pi^*(x)(\alpha(x)-i)d\xi, \\ P_1AP_2g(x) &= \\ &= \int_a^b g(\xi)(\alpha(\xi)+i)\Pi(\xi)J_2 \overset{\rightarrow}{\int_\xi^b} e^{-i\alpha(v)B_2(v)dv} d\xi S \overset{\rightarrow}{\int_a^x} e^{B_1(v)dv} J_1\Pi^*(x), \\ P_2AP_1g(x) &= 0 \end{aligned} \tag{2.10}$$

and $D_A = G(D_{A_1} \oplus D_{A_2}) \subset \mathbf{L}^2(\mathbb{R};\mathbb{C}^n)$. The representation (2.10) and straightforward calculations show that $A_1 = P_1 A$ is a dissipative operator onto $P_1 G^{-1} D_A$ and $A_2 = P_2 A$ is an antidissipative operator onto $P_2 G^{-1} D_A = P_2 D_A = D_{A_2}$. In other words

$$\begin{aligned} &Im\,(A_1 P_1 G^{-1} f(x), P_1 G^{-1} f(x)) \geq 0, \quad (f \in D_A),\\ &Im\,(A_2 P_2 G^{-1} f(x), P_2 G^{-1} f(x)) \leq 0, \quad (f \in D_A). \end{aligned}$$

The representation (2.9) implies that A is a regular coupling of a dissipative operator and an antidissipative one, i.e.

$$A = A_1 P_1 G^{-1} + A_2 P_2 + iKP_2$$

and $A = A_1 \vee A_2$.

The model A, defined by (2.6), is a closed densely defined operator as a coupling of a dissipative operator and an antidissipative one with real spectra.

For our further considerations we need the resolvent of the coupling of the model A, defined by (2.6). It turns out that for each $\lambda : Im\,\lambda \neq 0$ the operator $A - \lambda I$ is invertible and the explicit form of the resolvent can be obtained. In the case $\alpha(x) = x$ when $\lambda \neq i$ the resolvent $(A - \lambda I)^{-1}$ is given by the next theorem.

Theorem 2.3. *([8]) The model A, defined by (2.6), has the resolvent*

$$\begin{aligned} (A-\lambda I)^{-1} f(x) = {} & \tfrac{f(x)}{\alpha(x)-\lambda} - \\ & - i \int_{-\infty}^{x} \tfrac{\alpha(\xi)+i}{\alpha(\xi)-\lambda} f(\xi)\Pi(\xi) J_1 \overset{\curvearrowright}{\int_{\xi}^{x}} e^{-i\frac{1+\lambda\alpha(v)}{\alpha(v)-\lambda} B_1(v)dv} d\xi J_1 \Pi^*(x) \tfrac{\alpha(x)-i}{\alpha(x)-\lambda} + \\ & + i \int_{-\infty}^{x} \tfrac{\alpha(\xi)+i}{\alpha(\xi)-\lambda} f(\xi)\Pi(\xi) J_2 \overset{\curvearrowright}{\int_{\xi}^{x}} e^{i\frac{1+\lambda\alpha(v)}{\alpha(v)-\lambda} B_2(v)dv} d\xi J_2 \Pi^*(x) \tfrac{\alpha(x)-i}{\alpha(x)-\lambda} - \\ & - \tfrac{i}{\lambda-i} \int_{-\infty}^{+\infty} \tfrac{\alpha(\xi)+i}{\alpha(\xi)-\lambda} f(\xi)\Pi(\xi) J_2 \overset{\curvearrowright}{\int_{\xi}^{+\infty}} e^{i\frac{1+\lambda\alpha(v)}{\alpha(v)-\lambda} B_2(v)dv} d\xi S. \\ & . \overset{\curvearrowright}{\int_{-\infty}^{x}} e^{-i\frac{1+\lambda\alpha(v)}{\alpha(v)-\lambda} B_1(v)dv} J_1 \Pi^*(x) \tfrac{\alpha(x)-i}{\alpha(x)-\lambda} \end{aligned} \tag{2.11}$$

for each $\lambda : Im\,\lambda \neq 0$, $\lambda \neq i$, for each $f \in \mathbf{L}^2(\mathbb{R};\mathbb{C}^n)$ and the resolvent is a bounded operator in $\mathbf{L}^2(\mathbb{R};\mathbb{C}^n)$.

The model (2.6) generates semigroups of operators $\{T_t\}_{t\leq 0}$ and $\{T_t\}_{t\geq 0}$ from the class (C_0) with generators iA (see [8], [7]), defined by the equality

$$\begin{aligned} T_t f(x) = {} & -\tfrac{1}{2\pi i} \int_{-\infty}^{+\infty} e^{it(\xi - i\delta)} (A - (\xi - i\delta)I)^{-1} f(x) d\xi + \\ & + \tfrac{1}{2\pi i} \int_{-\infty}^{+\infty} e^{it(\xi + i\delta)} (A - (\xi + i\delta)I)^{-1} f(x) d\xi \end{aligned} \tag{2.12}$$

in the sense of a principal value where $f = (A - \lambda_0 I)^{-1} g$ for all $g \in D_1$, λ_0 is an arbitrary fixed number with $Im\,\lambda_0 > 0$, δ is an arbitrary number with $0 < \delta < Im\,\lambda_0$ when $t > 0$ and $Im\,\lambda_0 < 0$, $0 < \delta < -Im\,\lambda_0$ when $t < 0$.

The explicit obtaining of the asymptotics of the corresponding nondissipative processes $T_t f$ as $t \to \pm\infty$ allows to construct the scattering theory (as in the bounded case of the model A in [5]) for the couple (A^*, A): in other words to obtain the wave operators $W_\pm(A^*, A)$, the scattering operator and the similarity of A and the operator $\mathscr{Q}$ of multiplication by the independent variable. All results are obtained explicitly in [6], [7]), [8] using the multiplicative integrals, their properties and matrix generalization of the classical gamma-function, introduced in [5].

3 Main results

The presented properties of the multiplicative integrals in this paper and in [2], [3] play an important role for obtaining the asymptotics of the corresponding nondissipative processes $T_t f$ as $t \to \pm\infty$.

Let $m \times m$ matrix function $T(x)$, defined on $\mathbb{R}$, satisfy the conditions:

(i) $||T(x)|| \le C$, $||xT(x)|| \le C \ \forall x \in \mathbb{R}$;

(ii) $T(x) \in C_{\alpha_1}(\mathbb{R})$, $xT(x) \in C_{\alpha_2}(\mathbb{R})$ $(0 < \alpha_1 \le 1, 0 < \alpha_2 \le 1)$ (i.e. $||T(x_1) - T(x_2)|| \le C|x_1 - x_2|^{\alpha_1}$, $||x_1 T(x_1) - x_2 T(x_2)|| \le C|x_1 - x_2|^{\alpha_2} \ \forall x_1, x_2 \in \mathbb{R}$).

Here C is a constant and we denote by $||\ ||$ the norm in $\mathbb{C}^m$. Let now $\alpha = \min\{\alpha_1, \alpha_2\}$.

Further we will introduce some appropriate denotations. Let us denote the next operators using the multiplicative integrals and the limit values from the form (2.1):

$$\widetilde{T}(x) = (1+x^2)T(x), \tag{3.1}$$

$$\mathscr{U}_{2w}(x) = s - \lim_{\delta \to 0} \overset{\curvearrowright}{\int_w^{x-\delta}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} e^{-iT(x)x(x-\delta-w)}, \tag{3.2}$$

for all w, u, x such that $-\infty \le w < x < u \le +\infty$.

Then the next theorem is true.

Theorem 3.1. *Let the nonnegative or nonpositive matrix function $T(x)$ be integrable matrix function on $\mathbb{R}$ and satisfy the conditions **(i)** and **(ii)**. Then*

$$||\mathscr{U}_{2w}(x) - \mathscr{U}_{2w}(\xi)|| \le \widetilde{C}(1+|x|)\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'} \tag{3.3}$$

for some constant $\widetilde{C} > 0$, for all $w, \xi, x : w < \xi < x$, $0 < x - w < 1$ and $\alpha' = \alpha/(1+\alpha)$.

Proof. From the form (3.2) of the operator function $\mathscr{U}_{2w}(x)$ we have

$$||\mathscr{U}_{2w}(x) - \mathscr{U}_{2w}(\xi)|| = \Bigg\| \overset{\curvearrowright}{\int_w^{x-\delta}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} e^{-iT(x)x(x-\delta-w)} -$$
$$- \overset{\curvearrowright}{\int_w^{\xi-\delta}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} e^{-iT(\xi)\xi(\xi-\delta-w)} \Bigg\| \le$$

$$\le \Bigg\| \overset{\curvearrowright}{\int_w^{x-\delta}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\curvearrowright}{\int_w^{\xi-\delta}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \Bigg\|.$$

$$(3.4) \qquad \cdot\left\|e^{-iT(x)x(x-\delta-w)}\right\|+$$

$$+\left\|\overset{\xi-\delta}{\underset{w}{\overset{\rightarrow}{\int}}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv}e^{iT(\xi)\int\limits_{w}^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv}\right\|\cdot\left\|e^{-iT(x)x(x-\delta-w)}-e^{-iT(\xi)\xi(\xi-\delta-w)}\right\|\leq$$

$$\leq\left\|\overset{x-\delta}{\underset{w}{\overset{\rightarrow}{\int}}} e^{-i\frac{1+vx}{v-x}T(v)dv}e^{iT(x)\int\limits_{w}^{x-\delta}\frac{1+vx}{v-x}dv}-\overset{\xi-\delta}{\underset{w}{\overset{\rightarrow}{\int}}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv}e^{iT(\xi)\int\limits_{w}^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv}\right\|+$$

$$+\left\|e^{-iT(x)x(x-\delta-w)}-e^{-iT(\xi)\xi(\xi-\delta-w)}\right\|.$$

In the last inequality in (3.4) we have used that the matrix function $T(v)$ is selfadjoint on $\mathbb{R}$. Now we consider the second addend in the right hand side of the last inequality in (3.4) and we obtain that

$$\left\|e^{-iT(x)x(x-\delta-w)}-e^{-iT(\xi)\xi(\xi-\delta-w)}\right\|=$$
$$=\left\|e^{-iT(x)x(x-\delta-w)\int\limits_{1}^{e}\frac{1}{v}dv}-e^{-iT(\xi)\xi(\xi-\delta-w)\int\limits_{1}^{e}\frac{1}{v}dv}\right\|\leq$$
$$\leq\int\limits_{1}^{e}\frac{||T(x)x(x-\delta-w)-T(\xi)\xi(\xi-\delta-w)||}{v}dv\leq||T(x)x(x-\delta-w)-T(\xi)\xi(\xi-\delta-w)||=$$
$$=||T(x)x(x-\delta-w)-T(\xi)\xi(z-\delta-w)+T(\xi)\xi(x-\delta-w)-T(\xi)\xi(\xi-\delta-w)||\leq$$
$$\leq||T(x)x-T(\xi)\xi||.|x-\delta-w|+||T(\xi)\xi||.|(x-\delta-w)-(\xi-\delta-w)|\leq$$
$$\leq C(x-\xi)^{\alpha_1}+C(x-\xi)=C(x-\xi)^{\alpha_1}+C(x-\xi)^{\alpha_1}(x-\xi)^{1-\alpha_1}\leq$$
$$\leq C_1(x-\xi)^{\alpha_1}\leq C_1(x-\xi)^{\alpha}\leq C_1(x-\xi)^{\alpha'}$$

(for some appropriate constant $C_1>0$, for all $\alpha':0<\alpha'<\alpha<1$), where we have used Theorem 2.1. Consequently we have obtained the inequality

$$(3.5) \qquad \left\|e^{-iT(x)x(x-\delta-w)}-e^{-iT(\xi)\xi(\xi-\delta-w)}\right\|\leq C_1(x-\xi)^{\alpha'} \quad (\forall\alpha':0<\alpha'<\alpha<1).$$

On the other side in the case, when $\frac{x-\xi}{\xi-w}\geq 1$, we have

$$(3.6) \qquad \left\|e^{-iT(x)x(x-\delta-w)}-e^{-iT(\xi)\xi(\xi-\delta-w)}\right\|\leq 2\leq 2\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'}.$$

In the case when $\frac{x-\xi}{\xi-w}<1$, using (3.5), we obtain

$$(3.7) \qquad \begin{gathered}\left\|e^{-iT(x)x(x-\delta-w)}-e^{-iT(\xi)\xi(\xi-\delta-w)}\right\|\leq\\ \leq C_1(x-\xi)^{\alpha'}=C_1\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'}(\xi-w)^{1-\alpha'}\leq C_1\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'}\end{gathered}$$

for some constant $C_1>0$ and $\forall\alpha':0<\alpha'\leq\alpha<1$.

Now we consider the first addend in the right hand side of the last inequality in (3.4). Then

$$\left\| \overset{\overrightarrow{x-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| \leq$$

$$\leq \left\| \overset{\overrightarrow{x-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{\xi-\delta}\frac{1+vx}{v-x}dv} \right\| + \tag{3.8}$$

$$+ \left\| \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{\xi-\delta}\frac{1+vx}{v-x}dv} - \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\|.$$

Next we obtain that

$$\left\| \overset{\overrightarrow{x-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{\xi-\delta}\frac{1+vx}{v-x}dv} \right\| \leq$$

$$\leq \left\| \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} \right\| . \left\| \overset{\overrightarrow{x-\delta}}{\underset{\xi-\delta}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} - e^{-iT(x)\int_{\xi-\delta}^{x-\delta}\frac{1+vx}{v-x}dv} \right\| . \left\| e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} \right\| \leq$$

$$\leq \left\| \overset{\overrightarrow{x-\delta}}{\underset{\xi-\delta}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} - e^{-iT(x)\int_{\xi-\delta}^{x-\delta}\frac{1+vx}{v-x}dv} \right\| \leq \int_{\xi-\delta}^{x-\delta} \left\| \frac{1+vx}{v-x}T(v) - \frac{1+vx}{v-x}T(x) \right\| dv =$$

$$= \int_{\xi-\delta}^{x-\delta} \left| \frac{1+vx}{v-x} \right| ||t(v) - T(x)||dv \leq$$

$$\leq \int_{\xi-\delta}^{x-\delta} \frac{||T(v)-T(x)||}{|v-x|}dv + \int_{\xi-\delta}^{x-\delta} \frac{||xvT(x)-vxT(v)||}{|v-x|}dv \leq$$

$$\leq C\int_{\xi-\delta}^{x-\delta} \frac{(x-v)^{\alpha_1}}{x-v}dv + \int_{\xi-\delta}^{x-\delta} \frac{||xvT(x)-x^2T(x)+x^2T(x)-vxT(v)||}{|v-x|}dv \leq$$

$$\leq C\int_{\xi-\delta}^{x-\delta} \frac{(x-v)^{\alpha}}{x-v}dv + \int_{\xi-\delta}^{x-\delta} \frac{|v-x|.||xT(x)||+|x|.||xT(x)-vT(v)||}{|v-x|}dv \leq$$

$$\leq C\int_{\xi-\delta}^{x-\delta} (x-v)^{\alpha-1}dv + C\int_{\xi-\delta}^{x-\delta} dv + C|x|\int_{\xi-\delta}^{x-\delta} \frac{(x-v)^{\alpha_2}}{v-x}dv \leq$$

$$\leq C(1+|x|)\int_{\xi-\delta}^{x-\delta} (x-v)^{\alpha-1}dv + C(x-\xi) \leq C_2(x-\xi)^{\alpha} \leq$$

$$\leq C_2(1+|x|)\left(\frac{x-\xi}{\xi-w}\right)^{\alpha} \leq C_2(1+|x|)\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'}$$

for some appropriate constant $C_2 > 0$ and $\forall\alpha' : 0 < \alpha' \leq \alpha < 1$. Consequently for the first addend in the right hand side of the inequality (3.8) we have

$$\left\| \overset{\overrightarrow{x-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\overrightarrow{\xi-\delta}}{\underset{w}{\int}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{\xi-\delta}\frac{1+vx}{v-x}dv} \right\| \leq \tag{3.9}$$

$$\leq C_2(1+|x|)\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'}$$

for some appropriate constant $C_2 > 0$ and $\forall \alpha' : 0 < \alpha' \leq \alpha < 1$.

Let us choose the number γ such that $0 < \gamma < 1$ and $(x-\xi)^{\gamma} > \delta$ (i.e. $\xi - (x-\xi)^{\gamma} < x - \delta$) for an arbitrary sufficiently small fixed number $\delta > 0$. Now we consider the second addend in the right hand side of the inequality (3.8):

(3.10)

$$\left\| \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-\delta} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int\limits_{w}^{\xi-\delta}\frac{1+vx}{v-x}dv} - \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-\delta} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int\limits_{w}^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| \leq$$

$$\leq \left\| \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-(x-\xi)^{\gamma}} e^{-i\frac{1+vx}{v-x}T(v)dv} - \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-(x-\xi)^{\gamma}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} \right\| . \left\| \mathop{\overset{\rightarrow}{\int}}\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta} e^{-i\frac{1+vx}{v-x}T(v)dv} \right\| . \left\| e^{iT(x)\int\limits_{w}^{\xi-\delta}\frac{1+vx}{v-x}dv} \right\| +$$

$$+ \left\| \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-(x-\xi)^{\gamma}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} \right\| . \left\| \mathop{\overset{\rightarrow}{\int}}\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta} e^{-i\frac{1+vx}{v-x}T(v)dv} - \mathop{\overset{\rightarrow}{\int}}\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta} e^{-i\frac{1+vx}{v-x}T(x)dv} \right\| . \left\| e^{iT(x)\int\limits_{w}^{\xi-\delta}\frac{1+vx}{v-x}dv} \right\| +$$

$$+ \left\| \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-\delta} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} \right\| . \left\| \mathop{\overset{\leftarrow}{\int}}\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta} e^{i\frac{1+v\xi}{v-\xi}T(v)dv} - e^{iT(\xi)\int\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| . \left\| e^{iT(x)\int\limits_{w}^{\xi-\delta}\frac{1+vx}{v-x}dv} \right\| +$$

$$+ \left\| \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-\delta} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} \right\| . \left\| e^{iT(\xi)\int\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| . \left\| e^{iT(x)\int\limits_{w}^{\xi-(x-\xi)^{\gamma}}\frac{1+vx}{v-x}dv} - e^{iT(\xi)\int\limits_{w}^{\xi-(x-\xi)^{\gamma}}\frac{1+vx}{v-x}dv} \right\| \leq$$

$$\leq \int\limits_{w}^{\xi-(x-\xi)^{\gamma}} \left\| \left(\frac{1+vx}{v-x} - \frac{1+v\xi}{v-\xi} \right) T(v) \right\| dv + \int\limits_{w}^{\xi-(x-\xi)^{\gamma}} \left| \frac{1+vx}{v-x} \right| .||T(v) - T(\xi)||dv +$$

$$+ \int\limits_{w}^{\xi-(x-\xi)^{\gamma}} \left| \frac{1+v\xi}{v-\xi} \right| .||T(v) - T(\xi)||dv + \int\limits_{w}^{\xi-(x-\xi)^{\gamma}} \left\| \frac{1+vx}{v-x}T(x) - \frac{1+v\xi}{v-\xi}T(\xi) \right\| dv.$$

Using the following denotations the relations (3.10) can be written in the form

$$\left\| \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-\delta} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int\limits_{w}^{\xi-\delta}\frac{1+vx}{v-x}dv} - \mathop{\overset{\rightarrow}{\int}}\limits_{w}^{\xi-\delta} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int\limits_{w}^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| \leq$$

$$\leq \int\limits_{w}^{\xi-(x-\xi)^{\gamma}} \left\| \left(\frac{1+vx}{v-x} - \frac{1+v\xi}{v-\xi} \right) T(v) \right\| dv + \int\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta} \left| \frac{1+vx}{v-x} \right| .||T(v) - T(\xi)||dv + \tag{3.11}$$

$$+ \int\limits_{\xi-(x-\xi)^{\gamma}}^{\xi-\delta} \left| \frac{1+v\xi}{v-\xi} \right| .||T(v) - T(\xi)||dv + \int\limits_{w}^{\xi-(x-\xi)^{\gamma}} \left\| \frac{1+vx}{v-x}T(x) - \frac{1+v\xi}{v-\xi}T(\xi) \right\| dv =$$

$$= I_1 + I_2 + I_3 + I_4.$$

At first we consider the right hand side of (3.11) in the case, when

$$\frac{x-\xi}{\xi-w} \leq 1.$$

Then for I_1 in (3.11) we obtain

$$I_1 = \int\limits_{w}^{\xi-(x-\xi)^\gamma} \left\| \left(\tfrac{1+vx}{v-x} - \tfrac{1+v\xi}{v-\xi} \right) T(v) \right\| dv \le$$
$$\le \int\limits_{w}^{\xi-(x-\xi)^\gamma} ||T(v)|| \tfrac{x-\xi}{(x-v)(\xi-v)} dv + \int\limits_{w}^{\xi-(x-\xi)^\gamma} \tfrac{|vx-v\xi}{|v-x|.|v-\xi|} ||vT(v)|| dv \le$$
$$\le C(x-\xi)\tfrac{1}{(x-\xi)^\gamma} \int\limits_{w}^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv + C(1+|x|)(x-\xi) \int\limits_{w}^{\xi-(x-\xi)^\gamma} \tfrac{(x-v)^{\alpha_1}}{(x-v)(\xi-v)} dv +$$
$$+ C(1+|x|)(x-\xi) \int\limits_{w}^{\xi-(x-\xi)^\gamma} \tfrac{||xT(x)||}{(x-v)(\xi-v)} dv \le$$
$$\le C_3(1+|x|)(x-\xi)^{1-\gamma}(-\gamma \ln(x-\xi) + (x-\xi)^{\alpha\gamma} + 1) \le$$
$$\le C_4(1+|x|)(x-\xi)^{1-\gamma}(|\ln(x-\xi)| + (x-\xi)^{\alpha\gamma} + 1),$$

where C_3 and C_4 are appropriate constants. The last inequalities imply that

$$I_1 \le C_4(1+|x|)(x-\xi)^{1-\gamma}(|\ln(x-\xi)| + (x-\xi)^{\alpha\gamma} + 1). \tag{3.12}$$

Next for I_2 from (3.11) it follows that

$$I_2 = \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \left| \tfrac{1+vx}{v-x} \right| . ||T(v) - T(\xi)|| dv \le$$
$$\le \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{||T(v)-T(x)||}{x-v} dv + \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{|x|.||vT(v)-vT(x)||}{x-v} dv \le$$
$$\le C \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (x-v)^{\alpha-1} dv + |x| \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{||vT(v)-xT(x)||}{x-v} dv + |x| \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{(x-v)T(x)}{x-v} dv \le$$
$$\le C \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (x-v)^{\alpha-1} dv + C|x| \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (x-v)^{\alpha-1} dv + C|x| \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (x-v)^{\alpha-1} dv \le$$
$$\le C_5(1+|x|) \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (x-v)^{\alpha-1} dv \le C_6(1+|x|)(x-\xi)^{\alpha\gamma}$$

for appropriate constants C_5 and C_6. Consequetly

$$I_2 \le C_6(1+|x|)(x-\xi)^{\alpha\gamma}. \tag{3.13}$$

Now for I_3 from (3.11) we obtain

$$I_3 = \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \left| \tfrac{1+v\xi}{v-\xi} \right| . ||T(v) - T(\xi)|| dv \le$$
$$\le \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{||T(v)-T(\xi)||}{\xi-v} dv + |\xi| \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{||vT(v)-\xi T(\xi)+\xi T(\xi)-vT(\xi)||}{\xi-v} dv \le$$
$$\le C \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (\xi-v)^{\alpha_1-1} dv + (|\xi-x|+|x|) \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} \tfrac{C(\xi-v)^{\alpha_2}+(\xi-v)||T(\xi||}{\xi-v} dv \le$$
$$\le C \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (\xi-v)^{\alpha-1} dv + 2C(1+|x|) \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (\xi-v)^{\alpha-1} dv \le$$
$$\le C_7(1+|x|) \int\limits_{\xi-(x-\xi)^\gamma}^{\xi-\delta} (\xi-v)^{\alpha-1} dv = C_8(1+|x|)^{\alpha\gamma},$$

i.e.

$$I_3 \le C_8(1+|x|)^{\alpha\gamma} \tag{3.14}$$

for some constants $C_7, C_8 > 0$.

For I_4 in the inequality (3.11) after straightforward calculations we obtain

$$I_4 = \int\limits_w^{\xi-(x-\xi)^\gamma} \left\| \tfrac{1+vx}{v-x} T(x) - \tfrac{1+v\xi}{v-\xi} T(\xi) \right\| dv \le$$

$$\le \int\limits_w^{\xi-(x-\xi)^\gamma} \left(\tfrac{||T(x)-T(\xi)||}{|v-\xi|} + ||T(\xi)|| \left| \tfrac{1}{v-x} - \tfrac{1}{v-\xi} \right| + \tfrac{|v|||xT(x)-\xi T(\xi)||}{|v-\xi|} + ||v\xi T(\xi)|| \left| \tfrac{1}{v-x} - \tfrac{1}{v-\xi} \right| \right) dv \le$$

$$\le C(x-\xi)^{\alpha_1} \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv + C(x-\xi) \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{(x-v)(\xi-v)} dv +$$

$$+ C(x-\xi)^{\alpha_2} \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{|v-x|+|x|}{x-v} dv + C(x-\xi) \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{|v-x|+|x|}{(x-v)(\xi-v)} dv \le$$

$$\le C(x-\xi)^{\alpha} \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv + C(x-\xi)^{1-\gamma} \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv +$$

$$+ C(x-\xi)^{\alpha}(1+|x|) \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv + C(x-\xi)^{1-\gamma}(1+|x|) \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv \le$$

$$\le C_9(1+|x|)((x-\xi)^\alpha + (x-\xi)^{1-\gamma}) \int\limits_w^{\xi-(x-\xi)^\gamma} \tfrac{1}{x-v} dv =$$

$$= C_9(1+|x|)((x-\xi)^\alpha + (x-\xi)^{1-\gamma}) \left(\gamma|\ln(x-\xi)| + \ln(\xi-w) + \ln\left(1+\tfrac{x-\xi}{\xi-w}\right)\right) \le$$

$$\le C_{10}(1+|x|)((x-\xi)^\alpha + (x-\xi)^{1-\gamma}) \left(|\ln(x-\xi)| + \tfrac{x-\xi}{\xi-w}\right)$$

for some constants C_9, C_{10}. Hence

$$I_4 \le C_{10}(1+|x|)((x-\xi)^\alpha + (x-\xi)^{1-\gamma}) \left(|\ln(x-\xi)| + \frac{x-\xi}{\xi-w}\right). \tag{3.15}$$

Let us choose $\alpha' = \min\{\alpha\gamma, 1-\gamma\}$. For example, if $\gamma = \frac{1}{1+\alpha}$, then $\alpha' = \frac{\alpha}{1+\alpha}$. The inequalities (3.10) and (3.11) together with (3.12), (3.13), (3.14), (3.15) and the choice $\alpha' = \frac{\alpha}{1+\alpha}$ imply that

$$\left\| \overset{\xi-\delta}{\underset{w}{\overrightarrow{\int}}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int\limits_w^{\xi-\delta}\frac{1+vx}{v-x}dv} - \overset{\xi-\delta}{\underset{w}{\overrightarrow{\int}}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int\limits_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| \le \tag{3.16}$$

$$\le C_{11}(1+|x|)\left(\tfrac{x-\xi}{\xi-w}\right)^{\alpha'}.$$

($C_{11} > 0$ is an appropriate constant). In the course of proving the inequality (3.16) we have used the inequality

$$|\ln y| \le \frac{1}{\beta e} y^{-\beta} \quad for\ \ \beta > 0,\ \ 0 < y < 1.$$

Now in the case when $\frac{x-\xi}{\xi-w} < 1$ from the relations (3.4), (3.6) and (3.16) for $\alpha' = \frac{\alpha}{1+\alpha}$ it follows that

$$\left\| \overset{x-\delta}{\underset{w}{\overrightarrow{\int}}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int\limits_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\xi-\delta}{\underset{w}{\overrightarrow{\int}}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int\limits_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| \le \tag{3.17}$$

$$\le \widehat{C}(1+|x|)\left(\tfrac{x-\xi}{\xi-w}\right)^{\alpha'}.$$

In the case $\frac{x-\xi}{\xi-w} \geq 1$ from the properties of the multiplicative integrals when $T(x)$ is selfadjoint matrix fnction it follows that

$$\left\| \overset{x-\delta}{\underset{w}{\overrightarrow{\int}}} e^{-i\frac{1+vx}{v-x}T(v)dv} e^{iT(x)\int_w^{x-\delta}\frac{1+vx}{v-x}dv} - \overset{\xi-\delta}{\underset{w}{\overrightarrow{\int}}} e^{-i\frac{1+v\xi}{v-\xi}T(v)dv} e^{iT(\xi)\int_w^{\xi-\delta}\frac{1+v\xi}{v-\xi}dv} \right\| \leq 2 \leq 2\left(\frac{x-\xi}{\xi-w}\right)^{\alpha'}. \tag{3.18}$$

The inequalities (3.17) and (3.18) imply that there exists a constant $\widetilde{C} > 0$ such that the inequality (3.3) is true for $\alpha' = \frac{\alpha}{1+\alpha}$, $w < \xi < x$, $x - w < 1$. The theorem is proved. □

The inequality (3.3) together with other similar nequalities, presented in [2], [3], play an essential role in the process of obtaining the asymptotics of nondissipative curves generated by unbounded operators from the class K^r with different domains of the operator and its adjoint and presented as couplings of dissipative and antidissipative operators with real absolutely continuous spectra.

Finally, it has to be mentioned that (3.3) and all presented inequalities in [2] and [3] are satisfied when $T(x)$ is nonnegative or nonpositive operator function in infinite dimensional space.

REFERENCES:

[1] Borisova, G.S., Limit values of multiplicative integrals, Annuals of Konstantin Preslavski University, Faculty of mathematics and Informatics, v. **XVII C** (2016), 3–18.

[2] Borisova, G.S., Properties of multiplicative integrals I, CONFERENCE PROCEEDINGS, MATTEX 2016, v. **1**, (2016) 13–20.

[3] Borisova, G.S., Properties of multiplicative integrals II, CONFERENCE PROCEEDINGS, MATTEX 2016, v. **1**, (2016) 21–30.

[4] Borisova, G.S., The operators $A_\gamma = \gamma A + \overline{\gamma} A^*$ for a class of nondissipative operators A with a limit of the corresponding correlationfunction, Serdica Math. J., **29**, (2003), 109–140.

[5] Kirchev, K.P., Borisova, G.S., Nondissipative Curves in Hilbert Spaces Having a Limit of the Corresponding Correlation Function, Integral Equations Operator Theory, **40** (2001), 309–341.

[6] Kirchev, K., G. Borisova, G., A triangular model of regular couplings of dissipative and anridissipative operators, Comptes rendus de l'Acadèmie bulgare des sciences, **58** (5) (2005), 481–486.

[7] Kirchev, K., Borisova, G., Triangular models and asymptotics of continuous curves with bounded and unbounded semigroup generators, Serdica Math. J., **31** (2005), 95–174.

[8] Kirchev, K.P., Borisova, G.S., Regular Couplings of Dissipative and Anti-Dissipative Unbounded Operators, Asymptotics of the Corresponding Non-Dissipative Processes and the Scattering Theory, Integral Equations Operator Theory, **57** (2007), 339–379.

[9] Privalov, I.I., Boundary characteristics of analytic functions, Moscow, 1950 (Russian).

[10] Sakhnovich, L.A., Dissipative operators with an absolutely continuous spectrum, Works Mosk. Mat. Soc., vol. **19** (1968), 211–270 (Russian).

Galina Borisova
Faculty of Mathematics and Informatics, Shumen University
9712 Shumen, Bulgaria
E-mail: `g.borisova@shu.bg`

MATTEX 2018,
CONFERENCE PROCEEDING, v. 1, pp. 113- 120

ON THE COMPOUND BINOMIAL DISTRIBUTION*

KRASIMIRA Y. KOSTADINOVA

ABSTRACT: *In this paper a compound binomial distribution with geometric compounding distribution is analyzed. We derive some of its basic properties, recursion formulas and probability mass function. Then, the method of moments estimation for parameters of a compound binomial distribution is applied.*

KEYWORDS: *Compound distributions, geometric distribution, binomial distribution, moment estimation method.*

1 Introduction

In the present paper we consider a compound probability distribution, which has the form

(1) $$N = X_1 + X_2 + \ldots + X_Z.$$

Here the random variable Z belongs to the class of the Generalized Powers Series Distributions. The binomial, negative binomial, logarithmic series and Poisson distributions belong to this class, see [10]. The random variables X_i are independent, identically distributed as random variable X, which has a geometric distribution with parameter $\gamma \in (0,1)$. In [3] is defined a compound Generalized Powers Series Distributions, where the compounding random variable X has a shifted geometric distribution with parameter $1-\gamma$, denoted by $X \sim Ge_1(1-\gamma),\ \gamma \in [0,1)$.

In the insurance risk theory, some of the Generalized Powers Series Distributions are used for the number of claims during a given time period. The random sum N is interpreted as the aggregate claim amount in the risk model and we say that the random variable N has a compound distribution. When Z has a Poisson distribution, then the random variable N has a Pólya-Aeppli distribution, see [2]. When Z has a negative binomial distribution, then the random variable N has a compound Pólya distribution, see [4].

In the literature there are many generalizations of the classical risk model. One such generalization is by compounding of the counting process, which is the homogeneous Poisson process in the classical risk model. For example in [7] is defined Pólya-Aeppli process. Another such generalization is by mixing, see for example [1], [5] and [6].

The rest of this article is organized as follows. In Section 2, we consider a compound binomial distribution and derive its probability mass function, some properties and recursion formulas. In Section 3, the method of moments estimation for parameters of a compound binomial distribution is given.

2 Compound binomial distribution

In this section we consider the random sum (1), where the random variables X_i are independent, identically distributed as random variable X. In this paper, we suppose that the random variable X has a geometric distribution with parameter $\gamma \in (0,1)$, denoted by $X \sim Ge(\gamma)$. The probability mass function and probability generating function of the random variable X are given by

(2) $$q_i = P(X = i) = \gamma(1-\gamma)^i, \ \ i = 0,1,\ldots$$

*Supported by Scientific Research Grant RD-08-145/2018 of Shumen University.

and

$$\psi_1(s) = \frac{\gamma}{1-(1-\gamma)s}, \ |s| < \frac{1}{1-\gamma}. \tag{3}$$

Let the random variable Z is independent of the random variables $X_i, \ i = 1,2,\ldots$ We suppose that the random variable Z has a binomial distribution with parameters $n \in \mathbf{N}$ and $\theta \in (0,1)$, denoted by $Z \sim Bi(n,\theta)$. Then the random variable N has a compound binomial distribution with compounding random variable X.

The probability mass function and probability generating function of the random variable Z are given by

$$P(Z=i) = \binom{n}{i}\theta^i(1-\theta)^{n-i}, \ \ i = 0,1,\ldots,n \tag{4}$$

and

$$\psi_Z(s) = (1-\theta+\theta s)^n. \tag{5}$$

Then the probability generating function of N is given by

$$\psi(s) = \psi_N(s) = (1-\theta+\theta\psi_1(s))^n, \tag{6}$$

where $\psi_1(s)$ is the probability generating function of the compounding distribution, given by (3).

Definition 2.1. *The probability distribution of the random variable N, defined by the probability generating function (6) and compounding distribution, given by (2) and (3) is called a compound binomial distribution with notation $N \sim compBi(n,\theta,\gamma)$.*

Remark 2.1. *The mean and the variance of the compound binomial distribution are given by*

$$E(N) = \frac{n\theta(1-\gamma)}{\gamma},$$

$$Var(N) = \frac{(1-\gamma)((1-\gamma)(2-\theta)+\gamma)}{\gamma^2}n\theta.$$

For the Fisher index of dispersion we obtain

$$FI(N) = \frac{Var(N)}{E(N)} = 1 + \frac{(1-\gamma)(2-\theta)}{\gamma} > 1,$$

i.e. the compound binomial distribution is over-dispersed.

It is known that when $FI(N) < 1$, the distribution is called under-dispersed. When $FI(N) = 1$, the distribution is equi-dispersed and when $FI(N) > 1$, the distribution is over-dispersed, see [8] and [12].

From Remark 2.1 follows that the compound binomial distribution is over-dispersed related to the Poisson distribution, which $FI(N) = 1$. This makes the compound binomial distribution suitable for financial data.

2.1 The Probability Mass Function

The probability function of the random variable N is given by expanding the probability generating function $\psi(s)$ in powers of s. Denote by $f(i) = P(N = i)$, $i = 0, 1, \dots$, the probability mass function of the random variable N. We rewrite the probability generating function of (6) in the form

$$\psi(s) = \sum_{m=0}^{n} \binom{n}{m} \left(\frac{\theta\gamma}{1-(1-\gamma)s} \right)^m (1-\theta)^{n-m}. \tag{7}$$

Denote by $\psi^{(i)}(s) = \dfrac{\partial^{(i)}\psi(s)}{\partial s^i}$, for $i = 0, 1, \dots$, the derivatives of $\psi(s)$. From (7) we get the following

$$\psi^{(i)}(s) = (1-\gamma)^i \sum_{m=1}^{n} \binom{n}{m} \frac{m(m+1)\dots(m+i-1)}{(1-(1-\gamma)s)^{m+i}} (\theta\gamma)^m (1-\theta)^{n-m}. \tag{8}$$

From [2], it is known that

$$f(i) = \left. \frac{\psi^{(i)}(s)}{i!} \right|_{s=0}. \tag{9}$$

Theorem 2.1. *The probability mass function of the compound binomial distribution, denoted by $N \sim compBi(n, \theta, \gamma)$, is given by*

$$\begin{aligned} f(0) &= (1-\theta+\theta\gamma)^n, \\ f(i) &= (1-\gamma)^i \sum_{m=1}^{n} \tbinom{n}{m} \tbinom{m+i-1}{i} (\theta\gamma)^m (1-\theta)^{n-m}, \quad i = 1, 2, \dots \end{aligned} \tag{10}$$

Proof. The initial value $f(0) = (1-\theta+\theta\gamma)^n$ follows simply from the probability generating function $\psi(0) = f(0)$. Then (10) follows from (8) and (9).

□

On Figure 1 is given a graphic of the probability mass function of the binomial distribution, denoted by $Z \sim Bi(n, \theta)$. On Figure 2 is given a graphic of the probability mass function of the compound binomial distribution, denoted by $N \sim compBi(n, \theta, \gamma)$. A graphic of the probability mass function of $Z \sim Bi(n, \theta)$ and $N \sim compBi(n, \theta, \gamma)$ is given on Figure 3. All these graphics are made for $n = 20$, $\theta = 0.7$ and $\gamma = 0.5$.

From Figure 2 we see that the tail distribution is longer than the tail distribution on Figure 1. From Figure 3, one can see that the tail of the compound binomial distribution becomes heavier than the tail of the binomial distribution.

The following proposition gives an extension of the Panjer recursion formulas (see [9]).

Proposition 2.1. *The probability mass function of the compound binomial distribution satisfies the following recursions*

$$(1-\theta+\theta\gamma) i f(i) = (1-\gamma) \left[(1-\theta)(i-1) f(i-1) + n\theta \sum_{j=0}^{i-1} q_j f(i-j-1) \right], \; i = 2, 3, \dots \tag{11}$$

and $f(1) = \frac{n\theta(1-\gamma)}{1-\theta+\theta\gamma} q_0 f(0)$ with $f(0) = (1-\theta+\theta\gamma)^n$.

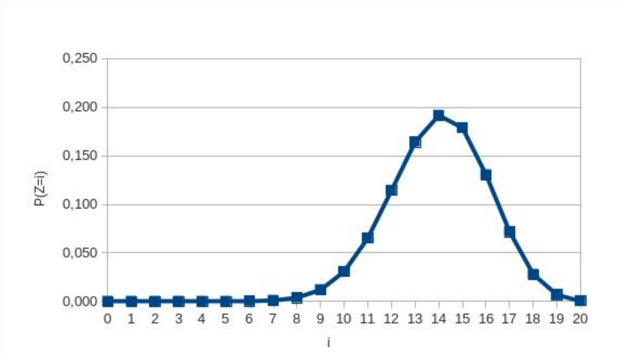

Figure 1. The PMF of Z~Bi(20, 0.7)

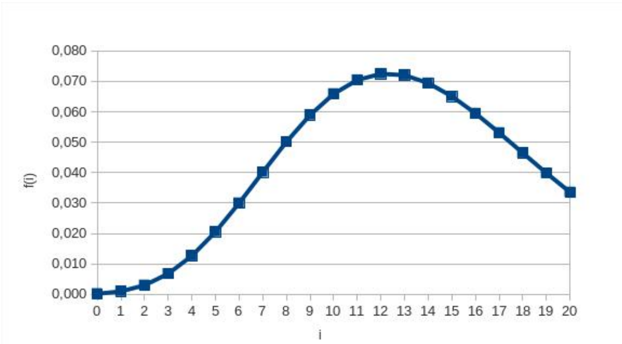

Figure 2. The PMF of N~compBi(20, 0.7, 0.5)

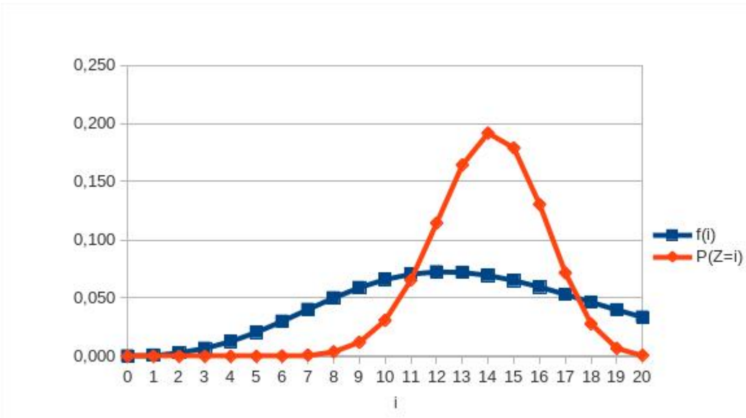

Figure 3. The PMF of Z~Bi(20, 0.7) and N~compBi(20, 0.7, 0.5)

Proof. Differentiation in (6) leads to

$$\frac{\partial \psi(s)}{\partial s} = \frac{n\theta(1-\gamma)}{(1-\theta)(1-(1-\gamma)s)+\theta\gamma}\psi_1(s)\psi(s), \tag{12}$$

where $\psi(s) = \sum_{i=0}^{n} f(i)s^i$, $\frac{\partial \psi(s)}{\partial s} = \sum_{i=0}^{n}(i+1)f(i+1)s^i$, and $\psi_1(s) = \sum_{j=0}^{\infty} q_j s^j$. The equation (12) has the form

$$[(1-\theta)(1-(1-\gamma)s)+\theta\gamma]\sum_{i=0}^{n}(i+1)f(i+1)s^i \quad = n\theta(1-\gamma)\sum_{i=0}^{n}f(i)s^i\sum_{j=0}^{\infty}q_js^j.$$

Changing the variable from $i+j=l \Rightarrow i=l-j$, yields

$$[(1-\theta)(1-(1-\gamma)s)+\theta\gamma]\sum_{i=0}^{n}(i+1)f(i+1)s^i \quad = n\theta(1-\gamma)\sum_{j=0}^{\infty}q_j\sum_{l=j}^{n+j}f(l-j)s^l.$$

Interchanging the order of summing in the double sum and equivalent transformations results in

$$\begin{aligned}[1-\theta+\theta\gamma]\sum_{i=0}^{n}(i+1)f(i+1)s^i \quad &= (1-\theta)(1-\gamma)\sum_{i=1}^{n+1}if(i)s^i \\ &+n\theta(1-\gamma)\left[\sum_{i=0}^{n}\sum_{j=0}^{i}q_jf(i-j)+\sum_{i=n}^{\infty}\sum_{j=i-n}^{i}q_jf(i-j)\right]s^i.\end{aligned}$$

The recursions (11) are obtained by equating the coefficients of s^i on both sides for fixed $i=0,1,2,\dots$ For $i=0$ follows that

$$f(1)=\frac{n\theta(1-\gamma)}{1-\theta+\theta\gamma}q_0f(0).$$

For $i=1,\dots,n$

$$(1-\theta+\theta\gamma)(i+1)f(i+1)=(1-\gamma)\left[(1-\theta)if(i)+n\theta\sum_{j=0}^{i}q_jf(i-j)\right],$$

and hence (11).

□

In the next proposition we give an alternative recursion formulas.

Proposition 2.2. *The probability mass function of the compound binomial distribution satisfies the recursions*

$$\begin{aligned}(1-\theta+\theta\gamma)if(i) \quad &= (1-\gamma)\left[(i-1)(2-2\theta+\theta\gamma)+n\theta\gamma\right]f(i-1) \\ &-(1-\gamma)^2(1-\theta)(i-2)f(i-2),\ i=2,3,\dots\end{aligned} \tag{13}$$

and $f(1)=\frac{n\theta\gamma(1-\gamma)}{1-\theta+\theta\gamma}f(0)$ *with* $f(0)=(1-\theta+\theta\gamma)^n$.

Proof. Differentiation in (6) leads to

$$\psi'(s)=\frac{n\theta}{1-\theta+\theta\psi_1(s)}\psi_1'(s)\psi(s), \tag{14}$$

where $\psi(s)=\sum_{i=0}^{n}f(i)s^i$, $\frac{\partial\psi(s)}{\partial s}=\sum_{i=0}^{n}(i+1)f(i+1)s^i$, and

$$\psi_1'(s)=\frac{(1-\gamma)\gamma}{(1-(1-\gamma)s)^2}$$

is the derivative of (3). So, the equation (14) has the form

$$[(1-\theta)(1-(1-\gamma)s)+\theta\gamma](1-(1-\gamma)s)\sum_{i=0}^{n}(i+1)f(i+1)s^i \quad = n\theta\gamma(1-\gamma)\sum_{i=0}^{n}f(i)s^i,$$

or equivalently

$$\begin{aligned}(1-\theta+\theta\gamma)\sum_{i=0}^{n}(i+1)f(i+1)s^i \quad &= (2-2\theta+\theta\gamma)(1-\gamma)\sum_{i=1}^{n+1}if(i)s^i \\ &-(1-\gamma)^2(1-\theta)\sum_{i=2}^{n+2}(i-1)f(i-1)s^i+n\theta\gamma(1-\gamma)\sum_{i=0}^{n}f(i)s^i.\end{aligned}$$

The recursions are obtained by equating the coefficients of s^i on both sides for fixed $i=0,1,2,\ldots$ □

3 Method of moments

3.1 The Procedure

Let $X_1, X_2, \ldots,$ are independent, identically distributed random variables, which have some distribution. Then the kth moment of the distribution is defined by

$$\mu_k = E(X^k).$$

For example $\mu_1 = E(X)$ and $\mu_2 = Var(X)+(E(X))^2$.

The procedure follows these four steps (see for example [11]):

1) If the model has m parameters, we compute m moments, $\mu_1, \mu_2, \ldots, \mu_m$ and obtain m equations with m unknowns.

2) Then we solve so that these m parameters as a function of the moments, i.e. every parameter we express by $\mu_i,\ i=1,2,\ldots,m$.

3) After that, based on the data $X=(X_1,X_2,\ldots,X_n)$, we compute the first m sample moments, $\overline{X^m} = \frac{1}{n}\sum_{i=1}^{n}X_i^m$.

4) We replace the distributional moments μ_m by the sample moments $\overline{X^m}$.

3.2 Example - compound binomial distribution

Let us denote by $\mu_k = E(N^k)$, the kth moment of the random variable N. Using that

$$\mu_1 = E(N) = \frac{n\theta(1-\gamma)}{\gamma},$$

$$\mu_2 = E(N^2) = \mu_1\left[1+\mu_1+\frac{(1-\gamma)(2-\theta)}{\gamma}\right],$$

$$\mu_3 = E(N^3) = \frac{\mu_1^3}{(n\theta)^2}\left[(n-1)(n-2)\theta^2+6(n-1)\theta+6\right]+3\mu_2-2\mu_1,$$

by the method of moments for the parameters n, θ and γ of the compound binomial distribution, we obtain

$$\widehat{n} = \frac{(\overline{X})^2[2\overline{X}\overline{X^2}+2\overline{X^2}(\overline{X})^2-2(\overline{X})^3-(\overline{X^2})^2-(\overline{X})^2-(\overline{X})^4 \mp |u|v]}{u(v^2 \pm |u|v)},$$

$$\widehat{\theta} = \frac{2\overline{X}\,\overline{X^3} - 3(\overline{X^2})^2 + (\overline{X})^2 + (\overline{X})^4 \pm |u|v}{\overline{X}[\overline{X^2} + \overline{X^3} + \overline{X}\,\overline{X^2} - (\overline{X})^2] - 2(\overline{X^2})^2},$$

$$\widehat{\gamma} = \frac{\overline{X}[\overline{X}\,\overline{X^2} + 3\overline{X^2}(\overline{X})^2 - (\overline{X})^3 - (\overline{X^2})^2 - (\overline{X})^4 + w \mp |u|v]}{\overline{X}[2(\overline{X^2})^2 + \overline{X^2}.\overline{X^3} + 3\overline{X}(\overline{X^2})^2 - \overline{X}\,\overline{X^3} - \overline{X^3}(\overline{X})^2 - \overline{X^2}(\overline{X})^3 + w \mp |u|v] - 2(\overline{X^2})^3},$$

where $u = \overline{X^2} - \overline{X} - (\overline{X})^2$, $v = \sqrt{2\overline{X}\,\overline{X^3} - 3(\overline{X^2})^2 + (\overline{X})^2 + (\overline{X})^4}$, $w = \overline{X}\,\overline{X^2} - (\overline{X})^2 - (\overline{X})^3 - \overline{X^2}(\overline{X})^2$ and $\overline{X^k} = \frac{1}{n}\sum_{i=1}^{n} X_i^k$, $k = 1,2,3$.

Table 1.

i	*Observed*
0	370410
1	43000
2	3930
3	300
4	27
5	3
≥ 6	0

Taking the data from the Table 1 we obtain the following moment estimates for the parameters n, θ and γ : $\widehat{n} = 2$, $\widehat{\theta} = 0.972$ and $\widehat{\gamma} = 0.937$.

Concluding remarks

In this paper we have introduced a compound binomial distribution as a compound binomial distribution with geometric compounding distribution. Also, we find the moments, the recursion formulas and probability mass function and then the method of moments estimation for its parameters is applied.

REFERENCES:

[1] Grandell, J., Mixed Poisson Process. Chapman & Hall, London (1997).

[2] Johnson, N.L., Kemp, A.W., Kotz, S., Univariate Discrete Distributions. Third edition, John Wiley & Sons, Hoboken, New Jersey (2005).

[3] Kolev, N., Minkova, L., Neytchev P., Inflated-Parameter Family of Generalized Power Series Distributions and Their Application in Analysis of Overdispersed Insurance Data. ARCH Research Clearing House, **2** (2000), 295–320.

[4] Kostadinova, K.Y., On the compound Pólya distribution. MATTEX 2016, vol. **1** (2016), 51–58.

[5] Lazarova, M.D., Minkova, L.D., I-Delaporte process and applications. Mathematics and Computers in Simulation, vol. **133** (2017), 135–141.

[6] Minkova, L.D., I-Pólya Process and Applications. Commun. Statist.-Theory and Methods, vol. **40** (2011), 2847–2855.

[7] Minkova, L.D., The Pólya-Aeppli process and ruin problems. J. Appl. Math. Stoch. Analysis, **3** (2004), 221–234.

[8] Minkova, L.D., Balakrishnan, N., Compound weighted Poisson distributions. Metrika, **76(4)** (2013), 543–558.

[9] Panjer, H., Recursive evaluation of a family of compound discrete distributions. ASTIN Bull., vol. **12** (1) (1981), 22–26.

[10] Patil, G.P., Certain properties of the generalized power distributions. Ann. Inst. Stat. Math., vol. **14** (1962), 179–182.

[11] Rice, J.A., Mathematical Statistics and Data Analysis. Duxbury Press, Belmont, CA, third edition (2007).

[12] Xekalaki, E., Under- and overdispersion. John Wiley & Sons, Hoboken, New Jersey. In: Teugels, J.L., Sundt, B. (eds) Encyclopedia of actuarial science, vol. **3** (2006), 1700–1705.

Krasimira Yankova Kostadinova,
Faculty of Mathematics and Informatics, Shumen University
E-mail: kostadinova@shu.bg

MATTEX 2018, SECTION **MATHEMATICS AND PHYSICS**
CONFERENCE PROCEEDING, v. 1, pp. 121 - 135

NUMERICAL COMPARATIVE ANALYSIS ON THE CLASSICAL VASICEK MODEL FOR DETERMING THE ZERO COUPON BOND'S PRICE

MEGLENA D. LAZAROVA, SYLVI-MARIA T. GUROVA

ABSTRACT: *In finance theory different models aiming to prognose the price of a given bond are considered. The finance modeling constructs an abstract introduction of a model that can describe the real world financial condition. This is a mathematical model that can be used to introduce a simplified version of the extension of the financial assets or a portfolio of business project or another investment. Usually the finance modeling is concerned with the exercising of the assets' prices or cooperative finances of a quantative nature. In the present paper we consider the classical Vasicek model without a risk factor and the classical Vasicek model with a risk factor. We give three numerical experiments for determing the zero coupon bond's price. The first one takes into consideration that the market price of risk is not changing. The other experiments are made for different values of the market price of risk.*

KEYWORDS: *finance modeling, interest rates, bond's price, zero coupon, stochastic differential equations, partial differential equations*

1 Introduction

In finance Vasicek's model is a mathematical model which is describing the random movement of the interest rates. The most universal model is the single-factor short-term model which is describing the random movement of the interest rate influenced by one source of a market risk. Oldrich Vasicek, see Vasicek (1977), [7] introduced this model as a stochastic investment model. In the classical Vasicek model the movement of the interest rate is given by a stochastic differential equation. i.e it is defined by a diffusion process. The assumption that is used in the model is that the bond market is a non-arbitrary market and that there exists a risk-free portfolio with a quantity *W*. The arbitrary argument is similar to that used in the Black-Scholes model for determing the put and call option's pricing. The main purpose in the classical Vasicek model is to determine the zero coupon bond's price at time *t* with maturity *T*. As this price can changes in time quite accidently it can be interpreted as a stochastic process for which the Ito's lemma can be applied.

In this paper we consider two financial models. The first one is the classical Vasicek model without a risk factor and the second one is the classical Vasicek model with a risk factor. For these two models we consider the zero coupon bond's price which is determined by using the apparatus of the stochastic differential equations and the partial differential equations. As there is a connection between the stochastic differential equations and the partial differential equations and it is obvious from the Fayman-Kac's theorem, see Oksendal (1998), [4] it is a good way to determine the final bond's price. We introduce a numerical comparative analysis and give some essential results. The analysis is made thanks to the partial differential equations given in the models with a specific boundary condition.

In section 2 we consider the classical Vasicek model. For this model we pay a great attention on the market price of risk and the value of the zero coupon bond. A numerical analysis and some comparative characteristics are made in Section 3. Some concluding remarks are given in Section 4.

2 A classical Vasicek model

The main application of the classical Vasicek model, see Nazil M. N. (2009), [3] in finance is to determine the zero coupon bond's price using the fact that the interest rate is following a stochastic differential equation. Usually this model is used for a valuation of interest-rate derivatives and it is also adapted to the credit markets. It is one of the earliest and the simplest term structure models based on increasingly realistic assumptions about the random movement of the interest rates.

Let r_0, , $\alpha > 0$, $\sigma > 0$ and μ are real constants. In the classical Vasicek model the interest rate is given by:

(1) $dr(t) = \alpha(\mu - r)dt + \sigma dW_t$ with an initial condition $r(0) = r_0$,

where

r - the instantaneous short rate interest

α - the speed of mean reversion

μ - the long-run expected value for the interest rate r

σ - the instantaneous standard deviation of the interest rate r

W_t - a standard Wiener process with mean 0 and standard deviation 1

The stochastic process used by Vasicek is known as the Ornstein-Uhlenbesk process. The first term in the stochastic process proposed by Vasicek pulls the short rate *r* back towards μ. So μ can be thought of as the long-run level of the short rate. When the short rate *r* is above μ, the first term trends to pull *r* downward since α is assumed to be positive. When the short rate *r* is below μ then *r* trends to drift upward. The impact of the mean reversion is to create realistic interest rate cycles with the level of α determing the length and the violence of rises and falls in interest rates.

The expression $\alpha(\mu - r)$ in the classical Vasicek model is called a drift. It is the expected moment change in the interest rate in time *t*. The constant α is called a rate of drift adjustment and the constant μ a level of drifting. The definition a level of drifting comes from the fact that the drift $\alpha(\mu - r)$ depends on that if $r > \mu$ or $r < \mu$. If the inequality $r > \mu$ holds then the drift is negative as the constant α is always a non-negative constant. In this situation the trend of the interest rate is moving down to the equilibrium. Conversely if the inequality $r > \mu$ holds then the drift is positive the trend of the interest rate is moving up to the equilibrium.

Let $P(r,t,T)$ be the price value of a zero coupon bond at time *t* and maturity *T*. The price depends from the interest rate movement and also the market value of the price reflects to the market expectation of the interest rate in the future. By Ito's lemma we obtain the following movement in the price of a zero coupon bond:

(2) $$dP(r,t,T) = P_r dr + \frac{1}{2} P_{rr}(dr)^2 + P_t dt,$$

where $P_r = \dfrac{\partial P(r,t,T)}{\partial r}$, $P_t = \dfrac{\partial P(r,t,T)}{\partial t}$ and $P_{rr} = \dfrac{\partial^2 P(r,t,T)}{\partial r^2}$.

Taking into account that $dr(t) = \alpha(\mu - r)dt + \sigma dW_t$ and substituting in (2) we obtain the following equation for the zero coupon bond's price:

(3) $$dP = P_r[\alpha(\mu - r)dt + \sigma dW_t] + \frac{1}{2}P_{rr}\sigma^2 dt + P_t dt$$

The calculations are obtained thanks to the following differential table:

X	dW_t	dt
dW_t	dt	0
dt	0	0

i.e $(dW_t)^2 = dt,\ dW_t.dt = 0$ and $dt.dt = 0$, see Stoyanov (1978) [6].

In the further exposition of the model is extremely important to find a formula for the zero coupon bond's price $P(r,t,T)$. The realization of this task is achieved by the assumption of that the bond market is non-arbitrary. Let on this market an investor forms a portfolio with a quantity *W* which includes a unit quantity of a zero coupon bond 1 with maturity T_1 and a quantity *w* of a zero coupon bond 2 with a maturity T_2. Then the portfolio's value is given by:

(4) $$W = P_1 + w.P_2,$$

where P_1 is the zero coupon bond's price of the first portfolio's bond and P_2 is the zero coupon bond's price of the second portfolio's bond. Then the change of the portfolio's value in time is given by:

(5) $$dW = dP_1 + w.dP_2,$$

We write the Ito's formula separately for the differentials dP_1, dP_2 and obtain the following partial differential equations:

$$dP_1 = P_{1_r}dr + \frac{1}{2}P_{1_{rr}}(dr)^2 + P_{1_t}dt = P_{1_r}[\alpha(\mu - r)dt + \sigma dW_t] + \frac{1}{2}P_{1_{rr}}\sigma^2 dt + P_{1_t}dt$$

$$dP_2 = P_{2_r}dr + \frac{1}{2}P_{2_{rr}}(dr)^2 + P_{2_t}dt = P_{2_r}\left[\alpha(\mu - r)dt + \sigma dW_t\right] + \frac{1}{2}P_{2_{rr}}\sigma^2 dt + P_{2_t}dt$$

We substitute the obtained values of dP_1 and dP_2 in the stochastic differential equation (5) and after that we obtain the following equation:

(6) $$dW = \left[\frac{1}{2}P_{1_{rr}}\sigma^2 + P_{1_r}\alpha(\mu - r) + P_{1_t} + wP_{2_t}\right]dt + w\left[\frac{1}{2}P_{2_{rr}}\sigma^2 + P_{2_r}\alpha(\mu - r)\right]dt + \left(P_{1_r} + wP_{2_r}\right)\sigma dW_t$$

If the coefficient in front of the differential dW_t is a zero i.e. $P_{1_r} + w.P_{2_r} = 0$ or $w = -\frac{P_{1_r}}{P_{2_r}}$, then we say that on the market is forming a riskless portfolio. Once the risk has been removed the change of the portfolio's value in time is given by:

(7) $$dW = \left(rP_1 + rwP_2\right)dt$$

Equating the right sides of the equations (6) and (7) and substituting by $w = -\frac{P_{1_r}}{P_{2_r}}$ we obtain the following equation:

(8) $$rP_1 - \frac{P_{1_r}}{P_{2_r}}P_2 r = \frac{1}{2}P_{1_{rr}}\sigma^2 + P_{1_r}\alpha(\mu - r) + P_{1_t} - \frac{P_{1_r}}{P_{2_r}}.P_{2_t} - \frac{1}{2}\sigma^2\frac{P_{1_r}}{P_{2_r}}.P_{2_{rr}} - \frac{P_{1_r}}{P_{2_r}}.P_{2_r}\alpha(\mu - r)$$

After simplifying and dividing by $P_{1_r}.\sigma$ we obtain the following equality:

(9) $$\frac{\frac{1}{2}\sigma^2 P_{1_{rr}} + \alpha(\mu - r)P_{1_r} + P_{1_t} - rP_1}{\sigma P_{1_r}} = \frac{\frac{1}{2}\sigma^2 P_{2_{rr}} + \alpha(\mu - r)P_{2_r} + P_{2_t} - rP_2}{\sigma P_{2_r}} = -\lambda$$

i.e. the market price of risk is given by:

$$-\lambda = \frac{\frac{1}{2}\sigma^2 P_{rr} + \alpha(\mu - r)P_r + P_t - rP}{\sigma P_r}$$

The market price of risk λ is assumed to be constant. From the last equation we obtain the partial differential equation of the zero coupon bond's price for the classical Vasicek model of the form:

(10) $\frac{1}{2}\sigma^2 P_{rr} + [\alpha(\mu - r) + \lambda\sigma]P_r + P_t - rP = 0$ with boundary condition

(11) $P(r,T,T) = 1$

i.e the zero coupon bond's price at maturity equals its principal amount, 1.

The price value of a zero coupon bond in the classical Vasicek model is analytically found, see Donald R. van Deventer et. al (1997), [1] and is given by the following formula:

(12) $$P(r,t,T) = P(r,\tau) = e^{-rF(\tau) - G(\tau)},$$

where $F(t,T)=F(\tau)=\frac{1}{\alpha}\left(1-e^{-\alpha\tau}\right)$ and $G(t,T)=G(\tau)=\left[\mu+\frac{\lambda\sigma}{\alpha}-\frac{\sigma^2}{2\alpha^2}\right]\left[\tau-F(\tau)\right]+\frac{\sigma^2}{4\alpha}F^2(\tau)$ are functions of the remaining time to maturity $T-t=\tau$.

The stochastic process proposed by Vasicek, see Donald R. van Deventer et al. (1997), [1] allows to calculate the expected value and the variance of the short rate at any time in the future s from the perspective of the current time t. Denoting the short rate at time t by $r(t)$, the expected value of the short rate at future time s is as follows:

$$(13)\qquad E_t\left[r(s)\right]=\mu+\left[r(t)-\mu\right]e^{-\alpha(s-t)}$$

The standard deviation of the potential values of the interest rate r around this mean value is:

$$(14)\qquad \text{Standard deviation}_t\left[r(s)\right]=\sqrt{\frac{\sigma^2}{2\alpha}\left[1-e^{-2\alpha(s-t)}\right]}$$

Of course there are some financial problems that are interested also in the constant $\sigma^2/2\alpha$ which is called a long term variance. The meaning of this constant is that the future expected values for the interest rate r may be regrouped in a long-term period of time with this value of the variance. Because the interest rate $r(s)$ at future time s is normally distributed there is a positive probability that $r(s)$ can be negative. This is inconsistent with a no-arbitrage economy in the special sense that consumers hold an option to hold cash instead of investing at negative interest rates. The magnitude of this theoretical problem with the Vasicek model depends on the level of the interest rates and the chosen parameters.

3 Numerical analysis on the classical Vasicek model and some comparative characteristics

In section 2 we consider the classical Vasicek model and the analytical obtaining of the zero coupon bond's price. The equation (10) given in the model with the boundary condition (11) is a parabolic partial differential equation and we are interested in the market price of risk λ. As we have the condition that the market is a non-arbitrary and the portfolio is riskless then we decide to take $\lambda=0$. By this decision we eliminate the market price of risk and obtain a model without a risk factor. The parabolic partial differential equation for this model is given by:

$$(15)\qquad \frac{1}{2}\sigma^2P_{rr}+\alpha(\mu-r)P_r+P_t-rP=0 \text{ with boundary condition (11)}$$

This is the first model used in our numerical analysis. We compare it with the classical Vasicek model for given different values of the constant λ. When $\lambda\neq0$ we can also use another probabilistic numerical method for valuating options, see Sobhani and Milev (2018), [5]. When λ is a function of time, see Gzyl et. al (2017), [2] then the interest rate and the volatility parameters depend on time and such model reflects more realistic on the random movement of the asset prices in financial market.

Starting from the parabolic partial differential equations (10) and (15) and using the software Mathematica 11 we give a numerical solution for the zero coupon bond's prices for the two considered models. The experiments are as follows:

3.1 First numerical experiment

This numerical experiment uses the data given in the table below. It is seen that for each year the market price of risk λ is a constant. Also the interest rate r is taking different values for the period of 1 year. The other parameters are fixed. Based on these data we obtain the zero coupon bond's prices without a risk and with a risk factor given in Figure 1 and Figure 2.

	α	μ	σ	λ	r	t (days)	T (year)	**Price**
Vasicek model without a risk factor	**0.10**	**0.01**	**0.01**		**0.010**	**0**	**1**	**0.1296**
	0.10	**0.01**	**0.01**		**0.015**	**73**	**1**	**0.1843**
	0.10	**0.01**	**0.01**		**0.020**	**146**	**1**	**0.2618**
	0.10	**0.01**	**0.01**		**0.025**	**219**	**1**	**0.3719**
	0.10	**0.01**	**0.01**		**0.030**	**292**	**1**	**0.5288**
	0.10	**0.01**	**0.01**		**0.019**	**365**	**1**	**1.00**
Vasichek model with a risk factor	**0.10**	**0.01**	**0.01**	**0.10**	**0.010**	**0**	**1**	**0.0050**
	0.10	**0.01**	**0.01**	**0.10**	**0.015**	**73**	**1**	**0.0140**
	0.10	**0.01**	**0.01**	**0.10**	**0.020**	**146**	**1**	**0.0390**
	0.10	**0.01**	**0.01**	**0.10**	**0.025**	**219**	**1**	**0.1073**
	0.10	**0.01**	**0.01**	**0.10**	**0.030**	**292**	**1**	**0.2945**
	0.10	**0.01**	**0.01**	**0.10**	**0.019**	**365**	**1**	**1.00**

Table 1
Specific values determing the bond's price
in Vasicek models

For fixed values of the parameters α, μ, σ and λ and changing the value of r and t we obtain the following figures.

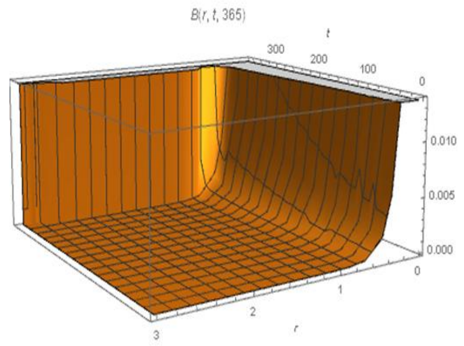

Figure 1
The zero coupon bond's price
for the Vasicek model without a risk factor
$\alpha = 0.10,\ \ \mu = 0.01,\ \ \sigma = 0.01$

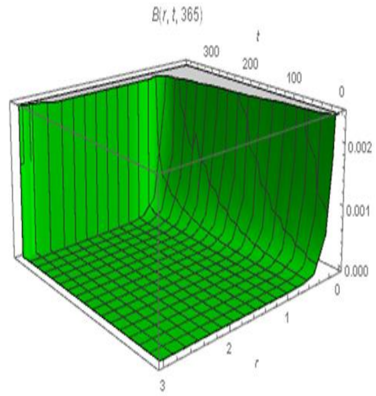

Figure 2
The zero coupon bond's price
for the Vasicek model with a risk factor
$\alpha = 0.10,\ \ \mu = 0.01,\ \ \sigma = 0.01,\ \ \lambda = 0.10$

At first we cannot find an essential difference between Figure 1 and Figure 2. When we compare them we obtain the following graphic:

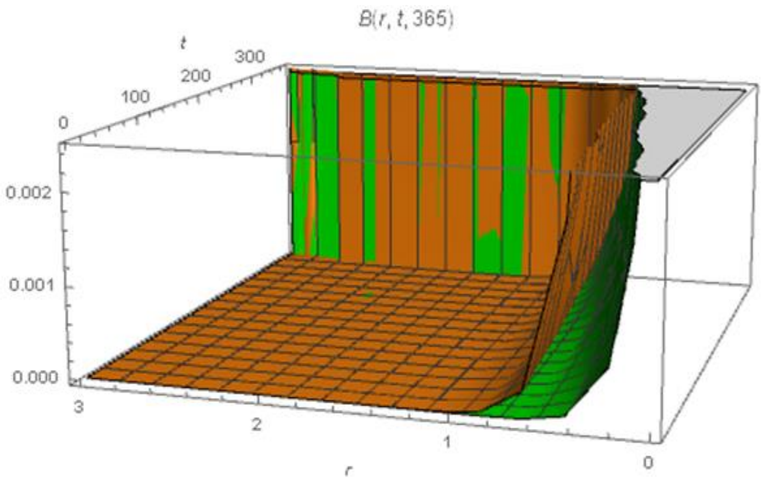

Figure 3
A comparative graphic characteristic showing
the difference in zero coupon bond's prices between
the model with risk and the model without a risk for
given parameters $\alpha = 0.10,\ \ \mu = 0.01, \sigma = 0.01, \lambda = 0.10$

The graphic and the table below show the difference in the comparative characteristic.

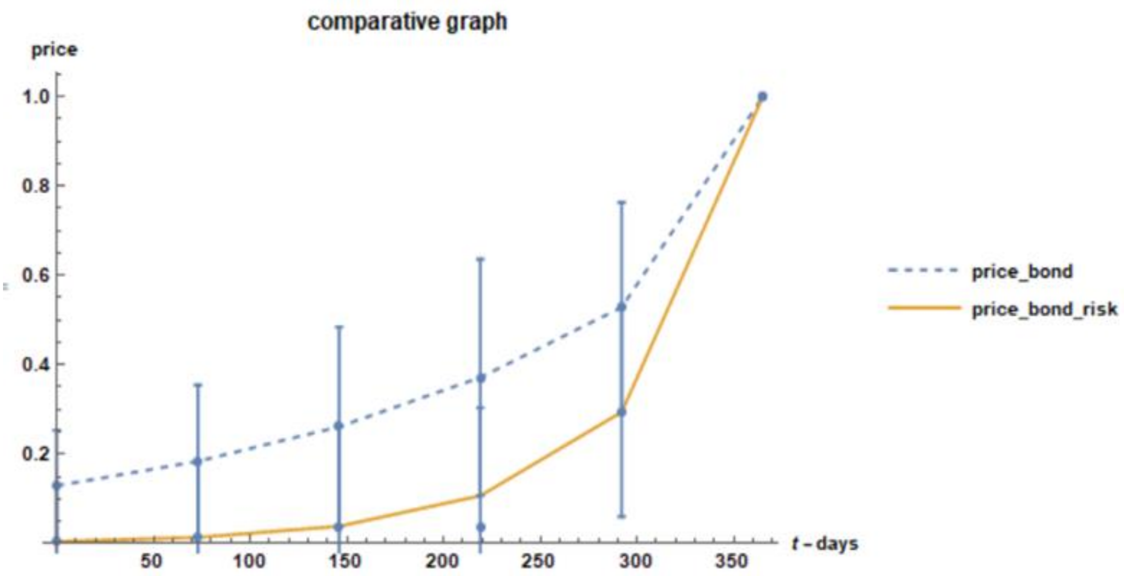

Figure 4
A comparative graph with given deviations for the first experiment

t days	0	73	146	219	292	365=1 year
error	0.1246	0.1703	0.2228	0.390	0.2343	0

Table 2
The deviation of the value of the bond's price for different periods of time

Conclusion:

The lines in Figure 4 represent the change in the value of the zero coupon bond's price over a period of one year. It can be seen that they are the contour of Figure 3. Based on the two tables and the graphics we can conclude that with a risk factor the zero coupon bond's price is decreasing in time.

3.2 Second numerical experiment

This numerical experiment uses the data given in the table below. It is seen that for each year the market price of risk λ is increasing. Also the interest rate r is taking different values in the time period of 5 years. The other parameters are fixed. Based on these data we obtain the zero coupon bond's prices without a risk and with a risk factor given in Figure 5 and Figure 6.

	α	μ	σ	λ	r	t (year)	T (year)	Price
Vasicek model without a risk factor	0.20	0.10	0.05		0.010	0	5	0.8275
	0.20	0.10	0.05		0.015	1	5	0.8601
	0.20	0.10	0.05		0.020	2	5	0.8938
	0.20	0.10	0.05		0.025	3	5	0.9287
	0.20	0.10	0.05		0.030	4	5	0.9644
	0.20	0.10	0.05		0.019	5	5	1.00
Vasichek model with a risk factor	0.20	0.10	0.05	0.05	0.010	0	5	0.8087
	0.20	0.10	0.05	0.10	0.015	1	5	0.8337
	0.20	0.10	0.05	0.15	0.020	2	5	0.8693
	0.20	0.10	0.05	0.20	0.025	3	5	0.9125
	0.10	0.10	0.05	0.25	0.030	4	5	0.9588
	0.10	0.10	0.05	0.30	0.019	5	5	1.00

Table 3
Another values determing the bond's price
in Vasicek models

The Table 3 is constructed based on the increasing values of the parameters α, μ, σ and λ in comparison with Table 1. Again the values of r and t are changing and the other parameters are fixed. This leads to the following figures.

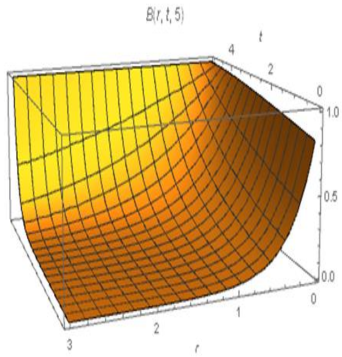

Figure 5
The zero coupon bond's price
for the Vasicek model without a risk factor
with given parameters
$\alpha = 0.20,\ \mu = 0.10,\ \sigma = 0.05$

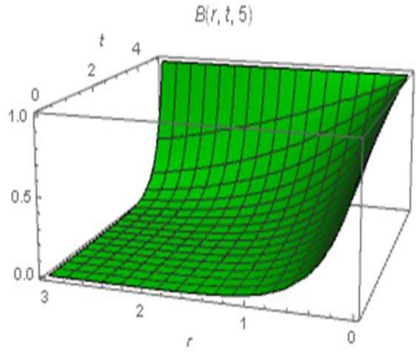

Figure 6
The zero coupon bond's price
for the Vasicek model with a risk factor
with given parameters
$\alpha = 0.20,\ \sigma = 0.05,\ \lambda = 0.20,\ \mu = 0.10$

At first we cannot find an essential difference between Figure 5 and Figure 6. When we compare them we obtain the following graphic:

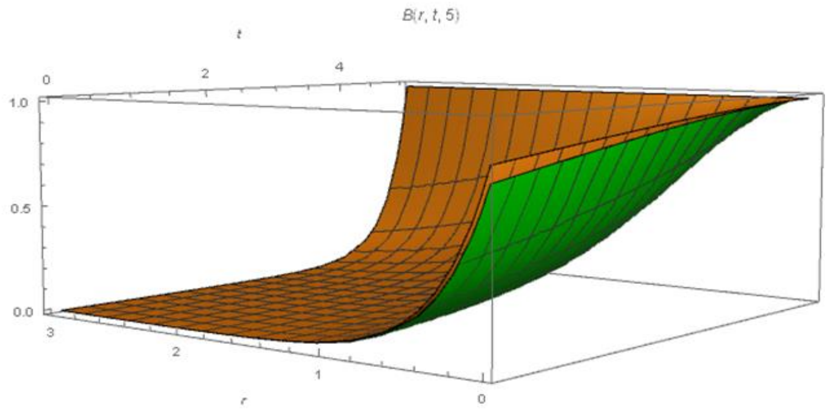

Figure 7
A comparative graphic characteristic showing
the difference in zero coupon bond's prices between
the model with risk and the model without a risk for
given parameters $\alpha = 0.20,\ \mu = 0.10,\ \sigma = 0.05,\ \lambda = 0.20$

The graphic and the table below show the difference in the comparative characteristic.

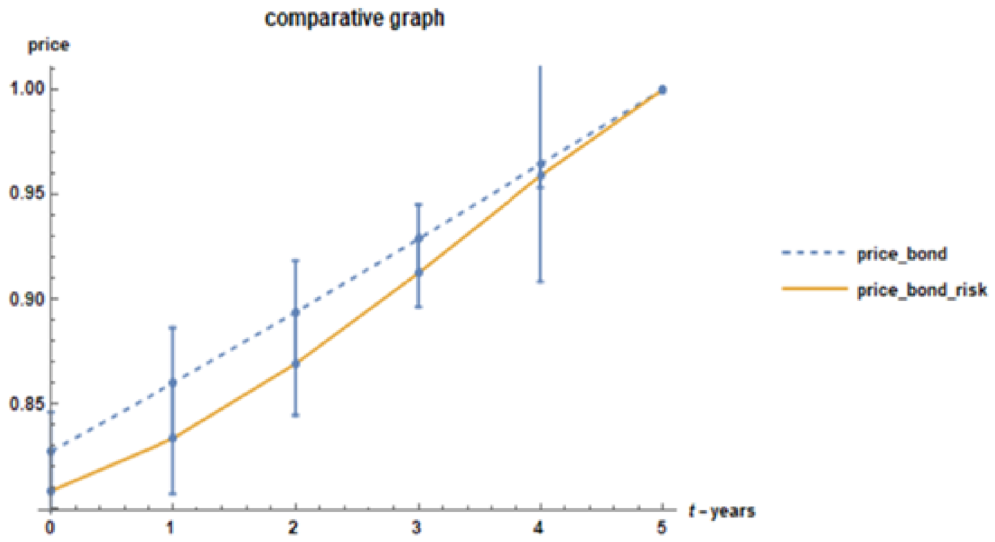

Figure 8
A comparative graph with given deviations for the second experiment

t years	0	1	2	3	4	5
error	0.0188	0.0264	0.0245	0.0162	0.0056	0

Table 4
The deviation of the value of the bond's price for different periods of time

Conclusion:

The lines in Figure 8 represent the change in the value of the zero coupon bond's price over a period of five years. Obviously they are the contour of Figure 7. Based on the two tables and the graphics we can conclude once again that with a risk factor the zero coupon bond's price is slightly decreasing in time.

3.3 Third numerical experiment

This numerical experiment uses the data given in the table below. It is seen that for each year the market price of risk λ is increasing. Also the interest rate r is taking different values in the time period of 10 years. The other parameters are fixed. Based on these data we obtain the zero coupon bond's prices without a risk and with a risk factor given in Figure 7 and Figure 8.

	α	μ	σ	λ	r	t (year)	T (year)	**Price**
Vasicek model without a risk factor	**0.30**	**0.25**	**0.10**		**0.010**	**0**	**10**	**0.2357**
	0.30	**0.25**	**0.10**		**0.015**	**2**	**10**	**0.3366**
	0.30	**0.25**	**0.10**		**0.020**	**4**	**10**	**0.4742**
	0.30	**0.25**	**0.10**		**0.025**	**6**	**10**	**0.6516**
	0.30	**0.25**	**0.10**		**0.030**	**8**	**10**	**0.8517**
	0.30	**0.25**	**0.10**		**0.019**	**10**	**10**	**1.00**
Vasichek model with a risk factor	**0.30**	**0.25**	**0.10**	**0.05**	**0.010**	**0**	**10**	**0.2103**
	0.30	**0.25**	**0.10**	**0.10**	**0.015**	**2**	**10**	**0.2852**
	0.30	**0.25**	**0.10**	**0.15**	**0.020**	**4**	**10**	**0.4038**
	0.30	**0.25**	**0.10**	**0.20**	**0.025**	**6**	**10**	**0.5829**
	0.30	**0.25**	**0.10**	**0.25**	**0.030**	**8**	**10**	**0.8172**
	0.30	**0.25**	**0.10**	**0.30**	**0.019**	**10**	**10**	**1.00**

Table 5
Values determing the bond's price
in the last experiment Vasicek models

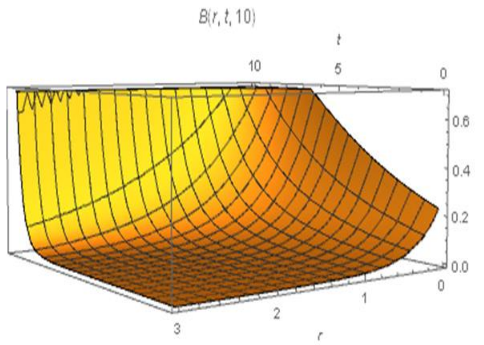

Figure 7
The zero coupon bond's price
for the Vasicek model without a risk factor
with given parameters:
$\alpha = 0.30,\ \ \mu = 0.25,\ \ \sigma = 0.10$

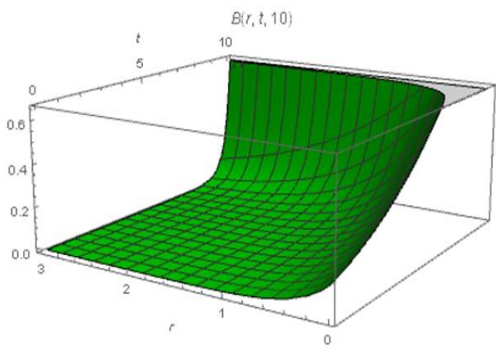

Figure 8
The zero coupon bond's price
for the Vasicek model with a risk factor
with given parameters:
$\alpha = 0.30,\ \ \mu = 0.25,\ \ \sigma = 0.10,\ \ \lambda = 0.20$

The comparative graphic is:

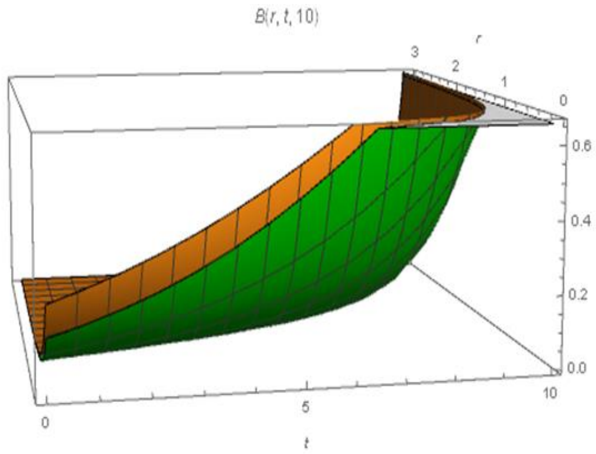

Figure 9
A comparative graphic characteristic showing
the difference in zero coupon bond's prices between
the model with risk and the model without a risk for
given parameters $\alpha = 0.30,\ \ \mu = 0.25, \sigma = 0.10, \lambda = 0.20$

The graphic and the table below show the difference in the comparative characteristic.

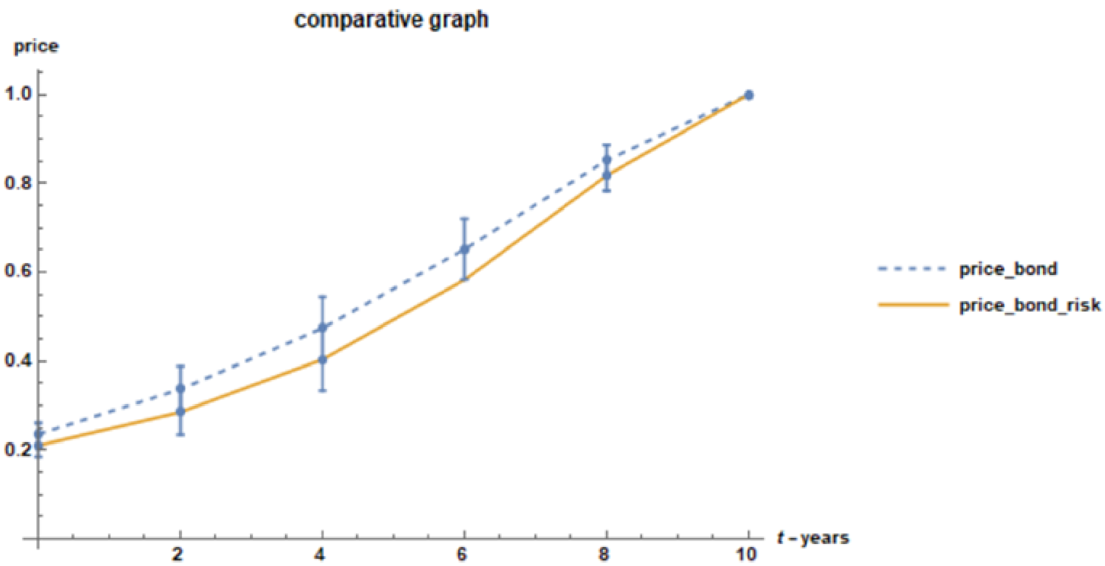

Figure 10
A comparative graph with given deviations for the third experiment

t years	0	2	4	6	8	10
error	0.0254	0.0514	0.0704	0.0687	0.0345	0

Table 6
The deviation of the value of the bond's price for different periods of time for the last experiment

Conclusion:

The lines in Figure 10 represent the change in the value of the zero coupon bond's price over a period of ten years. Based on the two tables and the graphics we can conclude once again that with a risk factor the zero coupon bond's price is strongly decreasing in time.

4 Concluding remarks

In this paper we introduced the zero coupon bond's price in the classical Vasicek model. We considered the influence of the market price of a risk λ over the bond's price and called this model – Vasicek model with a risk factor. Analogously we eliminate the influence of the market price of a risk and called the model - Vasicek model without a risk factor. Taking different values of the risk factor and the other parameters we made a numerical comparative analysis for the changing of the bond's prices in selected periods of time. Finally, we gave three numerical experiments and made the conclusions that the zero coupon bond's price with risk factor is decreasing in time.

5 Acknowledgement

The first author is partially supported by project "Stochastic project and simulation models in the field of medicine, social sciences and dynamic systems" founded by the National Science Fund of the Ministry of Education and Science of Bulgaria (Grant No DN 12/11/20 Dec. 2017).

The second author is partially supported by the financial funds allocated to the Sofia University "St. Kliment Ohridski" (Grant No 80-10-222/2018) and by the EC under H2020 Project VI-SEEM (Contract Number 675121), as well as by Operational Programme "Science and Education for Smart Growth", project name "CoE on Informatics and ICT" (Contract No: BG05M2OP001-1.001-0003-001).

REFERENCES:

[1] Donald van Deventer, R., Imai, K., Financial Risk Analytics – A term structure model approach for banking, insurance and investment management. IRWIN professional Publishing, USA (1997).

[2] Gzyl H., Milev M., Tagliani A. Discontinuous payoff option pricing by Mellin transform: A probabilistic approach. Finance Research Letters, **20,** 281-288 (2017).

[3] Nazir, M. N., Short rates and bond prices in one-factor models, U.U.D.M. Project Report (2009)

[4] Oksendal, B., Stochastic differential equations – an addition with application. Springer-Verlag. Fifth edition, Berlin Heidelberg (1998).

[5] Sobhani A., Milev M., A numerical method for pricing discrete double barrier option by Legendre multiwavelet. Journal of Computational and Applied Mathematics, **328**, 355-364 (2018).

[6] Stoyanov, J. Stochstic processes – Theory and Applications, Sofia (1998).

[7] Vasicek, O., An equilibrium characterization of the term structure. J. Financial Economics. **5** (1977).

Meglena Delcheva Lazarova
Technical University of Sofia, Faculty of Applied Mathematics and Informatics,
8 Sv. Kliment Ohridski Blvd., Sofia
E-mail: meglena.laz@tu-sofia.bg

Sylvi-Maria Todorova Gurova
First Affiliation: Institute of Information and Communication Technologies – BAS,
Acad. G. Bonchev, St., bl. 25A
E-mail: smgurova@parallel.bas.bg
Second Affiliation: Sofia University "St. Kliment Ohridski", Faculty of Mathematics and Informatics,
Master's Programme "Computational Mathematics and Mathematical modeling

MATRIX GAME, EQUILIBRIUM AND COMPUTER REALIZATION OF WOLFRAM MATHEMATICA

YANITSA M. MANDZHUKOVA, RADOSLAVA S. TERZIEVA, SNEZHANA G. HRISTOVA

ABSTRACT: *The Game Theory is a relatively new branch in mathematics and can be used as an apparatus for modeling and solving real processes and phenomena.*

Models of real life situations that are adequately described by games are studied. The models are solved using the Wolfram Mathematica software package.

KEYWORDS: *Game theory, matrix games, equilibrium.*

МАТРИЧНИ ИГРИ, РАВНОВЕСИЕ И КОМПЮТЪРНА РЕАЛИЗАЦИЯ НА WOLFRAM MATHEMATICA*

ЯНИЦА М. МАНДЖУКОВА, РАДОСЛАВА С. ТЕРЗИЕВА, СНЕЖАНА Г. ХРИСТОВА

АБСТРАКТ: *Теория на игрите е сравнително нов клон в математиката и се използва като апарат за моделиране и решаване на редица реални процеси и явления в икономиката, бизнеса, политиката, биологията и пр.*

В тази статия са предложени и изследвани някои модели на реални ситуации, описани адекватно чрез игри. Моделите са решени с помощта на софтуерния пакет Wolfram Mathematica.

1 Въведение

Теория на игрите е дял от приложната математика. С помощта на методите от теорията на игрите стратегически се изучават математически модели, като се взимат решения в конфликтни ситуации. Конфликтни са ситуациите, при които има две взаимодействащи си страни с противоположни цели. При това, резултатът от всяко действие на едната страна зависи от начина на действие на противоположната страна.

Всяка една ситуация, при която всеки участник е зависим от действията на останалите участници, може да се моделира като игра ([1], [3], [4]). Това обуславя популярността и приложимостта на теорията на игрите. Реално тази теория може да помогне за вземането на оптималното решение във всяка една житейска ситуация. Тя постепенно излиза от математиката и икономиката и навлиза, както в психологията и социологията- за да се опита да обясни етичното поведение, обществения и социалния избор, така и в политологията, където е полезна при изследване на ескалиращо напрежение и международни военни конфликти.

В настоящата статия е показан начин за изследване на матрични игри, които са използване за модели на конкретни реални ситуации. Написани са кодове на Wolfram Mathematica и са намерени доминиращи стратегии, равновесие на съответните модели.

* Настоящата работа е частично финансирана по проект МУ17ФМИ007 и СП17ФМИ005 при НПД при Пловдивски Университет „П.Хилендарски“.

2 Основни теоретични сведения за матрични игри

Една безкоалиционна игра G(И,X,P) с k играча се нарича игра с нулева сума, ако за всеки изход на играта $x = (x_1, x_2, \ldots, x_n) \in X = X_1 \times X_2 \times \ldots \times X_n$, сумата от плащанията на всички играчи е нулева, т.е. е изпълнено $\sum_{j=1}^{k} P_j(x) = 0$.

Безкоалиционните игри с двама играчи с нулева сума се наричат антагонистични. При антагонистичните игри за всеки изход на играта е изпълнено, че плащанията на двамата играчи са равни по абсолютна стойност, но с противоположни знаци. Това означава, че интересите на двамата играчи са антагонистични, противоположни.

Антагонистичните игри могат да бъдат крайни или безкрайни. Крайните антагонистични игри се наричат матрични игри.

Доминиращи стратегии

Стратегията x_i е доминираща, ако елементите на i-тия ред в матрицата M са по-големи или равни от съответните елементи във всички останали редове.

Стратегията y_j е доминираща, ако елементите на j-тия стълб в матрицата M са по-малки или равни от съответните елементи във всички останали стълбове.

Седлова точка и цена на играта

Изходът на матрична игра, при който нейното плащане е по-малко или равно на всяко плащане в съответния й ред и е по-голямо или равно на всяко плащане в съответния й стълб се нарича седлова точка.

Ако съществува число V, такова, че

- за И1 има стратегия, която му гарантира печалба ПОНЕ V,
- за И2 има стратегия, която му гарантира печалба НЕ ПОВЕЧЕ от V, то V се нарича цена на играта.

Ако една матрична игра има седлова точка, то нейното плащане е цена на играта.

3 Моделиране на някои конкретни реални ситуации чрез игри.

Ще разгледаме примери, с които ще илюстрираме основните понятия и методи в теория на игрите.

Пример 1. Две големи компютърни фирми предлагат определен вид лаптоп. Маркетингово проучване показва печалбата на двете фирми, при различните цени, измежду които те могат да изберат (Таблица 1).

Таблица 1.

Печалбата (в млн. лв) на двете фирми.

	A	***B***	**C**
X	$5, -5$	$-1,1$	$1, -1$
Y	$10, -10$	$4, -4$	$3, -2$
Z	$6, -6$	$3, -3$	$2, -3$

От таблица 1 се вижда, че играчите имат краен брой стратегии, като играч 1 (И1) има стратегии $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$, а И2 - $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, както и че двамата играчи са с антагонистични интереси (всеки изход на играта има нулева сума). В този случай играта се нарича матрична и е напълно дефинирана само с матрицата от плащанията на първия играч. Следователно можем да представим матрицата на плащанията (М) за двамата играчи по следния начин:

Таблица 2.

Матрица на плащанията (М)

	A	$\boldsymbol{B}$	C
$\boldsymbol{X}$	5	−1	1
$\boldsymbol{Y}$	10	4	3
Z	6	3	2

Тъй като интересите на двамата играчи са антагонистични/противоположни, то платежната матрица на И2 е – М.

От Таблица 2 се вижда, че елементите във втория ред са по-големи от съответните елементи в останалите редове. Следователно, можем да твърдим, че стратегията $\boldsymbol{Y}$ е доминираща стратегия за И1. Докато елементите в нито един от стълбовете не са по-малки или равни от съответните елементи в останалите стълбове. Следователно, няма доминираща стратегия за И2.

$$\text{И1}: \begin{array}{c} \boldsymbol{X} \\ \boldsymbol{Y} \\ \boldsymbol{Z} \\ \\ \end{array} \begin{array}{c} \begin{pmatrix} 5 & -1 & 1 \\ \mathbf{10} & \mathbf{4} & \mathbf{3} \\ 6 & 3 & 2 \end{pmatrix} \\ \begin{matrix} \boldsymbol{A} & \quad \boldsymbol{B} & \quad \boldsymbol{C} \end{matrix} \end{array} \qquad \text{И2}: \begin{pmatrix} 5 & -1 & 1 \\ 10 & 4 & 3 \\ 6 & 3 & 2 \end{pmatrix}$$

Фигура 1. Намиране на доминиращи стратегии.

За намиране на седловата точка с кръгче отбелязваме **min** по ред, с квадратче **max** по стълб (Таблица 3).

Таблица 3.

Намиране на седлова точка

	A	$\boldsymbol{B}$	C
$\boldsymbol{X}$	5	−1	1
$\boldsymbol{Y}$	$\boxed{10}$	$\boxed{4}$	$\boxed{3}$
Z	6	3	2

Разглежданата игра има седлова точка $(\boldsymbol{Y}, \boldsymbol{C})$ и цената на играта е 3.

Реализираме решението чрез Wolfram Mathematica:

- Изготвяне матрица на плащанията:

```
a1 = 5; a2 = -1; a3 = 1; a4 = 10; a5 = 4; a6 = 3; a7 = 6; a8 = 3; a9 = 2;
Grid[{{" ", "A", "B", "C"}, {"X", a1, a2, a3}, {"Y", a4, a5, a6}, {"Z", a7, a8, a9}}, Frame → All, Background → LightGray]
```

- Изготвяне на код, намиращ доминиращите стратегии за двамата играчи:

```
If[a1 > a4 && a2 > a5 && a3 > a6 && a1 > a7 && a2 > a8 && a3 > a9, "Стратегия X е доминираща стратегия за първия играч.",
 If[a4 > a1 && a5 > a2 && a6 > a3 && a4 > a7 && a6 > a3 && a6 > a9, "Стратегия Y е доминираща стратегия за първия играч.",
  If[a7 > a1 && a8 > a2 && a9 > a3 && a7 > a4 && a8 > a5 && a9 > a6, "Стратегия Z е доминираща стратегия за първия играч.", "Няма доминираща стратегия за първия играч."]]]
If[a1 < a2 && a1 < a3 && a4 < a5 && a4 < a6 && a7 < a8 && a7 < a9, "Стратегия A е доминираща стратегия за втория играч.",
 If[a2 < a1 && a2 < a3 && a6 < a4 && a6 < a7 && a8 < a7 && a8 < a9, "Стратегия B е доминираща стратегия за втория играч.",
  If[a3 < a1 && a3 < a2 && a6 < a4 && a6 < a5 && a9 < a7 && a9 < a8, "Стратегия C е доминираща стратегия за втория играч.", "Няма доминираща стратегия за втория играч."]]]
```

- Търсене на седлова точка:

```
max = Max[Min[a1, a2, a3], Min[a4, a5, a6], Min[a7, a8, a9]];
min = Min[Max[a1, a4, a7], Max[a2, a5, a7], Max[a3, a6, a9]];
If[min == max, "Има седлова точка.", "Няма седлова точка."]
Print["Цената на играта е равна на ", min, "."]
```

В изпълнение на този код, получаваме доминиращите стратегии за двата играчи (ако има такива), седлова точка и цена на играта.

Резултати:

```
Стратегия Y е доминираща стратегия за първия играч.
Няма доминираща стратегия за втория играч.
Има седлова точка.
Цената на играта е равна на 3.
```

Пример 2. Рибари, които работят на един риболовен кораб имат възможност да сложат мрежите за риба или в открито море или в залива, като таксата им за залива е 10 лв, а за открито море - 20 лв. В открито море има повече и по-хубава риба, но има възможност и от вълнения на водата, които биха унищожили мрежите им. Докато в залива, те ще спечелят от улова на риба 35 (x100) лв. Ако изберат да отидат в открито море и има вълнение, то те ще спечеля само 15 (x100) лв, а ако няма вълнение – 75 (x100) лв.

Кое място да изберат рибарите? Колко ще спечелят при този избор?

Първоначално определяме кои са стратегиите в дадената игра – **Открито, Закрито, Вълнение, Без вълнение.**

Как да попълним таблицата?

<u>Открито</u> <u>(- 20 лв.)</u>

С вълнение (+ 15 лв.) – 20 + 15= - 5

Без вълнение (+ 75 лв.) – 20 + 75= 5

<u>Закрито</u> <u>(- 10 лв.)</u>

С вълнение (+ 35 лв.) – 10 + 35= 25

Без вълнение (+ 35 лв.) – 10 + 35= 25

Сега изготвяме матрицата на плащанията и попълваме плащанията в нея.

Таблица 4

Матрица на плащанията за Пример 2

	Вълнение	*Без вълнение*
Открито	−5	55
Закрито	25	25

Използваме написания по-горе алгоритъм за решаване на такъв тип задачи чрез Wolfram Mathematica.

```
a1 = -5; a2 = 55; a3 = 25; a4 = 25;
Grid[{{" ", "Вълнение", "Без вълнение"}, {"Открито", a1, a2}, {"Закрито", a3, a4}}, Frame → All, Background → LightGray]
```

	Вълнение	Без вълнение
Открито	-5	55
Закрито	25	25

```
max = Max[Min[a1, a2], Min[a3, a4]];
min = Min[Max[a1, a3], Max[a2, a4]];
If[min == max, "Има седлова точка.", "Няма седлова точка."]
Print["Седловата точка е равна на ", min, "."]
```

Решение:

Има седлова точка.

Цената на играта е равна на 25.

Извод : Има седлова точка , която е (**Закрито, Вълнение**), а цената на играта е 25, което означава, че рибарите трябва да отидат в залива без значение дали има или няма вълнение. Те не са заинтересовани да се отклоняват от този избор. При този избор ще спечелят 25(x100) лв.

Заключение

В настоящата работа, благодарение на теория на игрите са моделирани и изследвани две коренно различни реални ситуации. Разгледани са матрични игри, които са решени с помощта на математическия софтуер Wolfram Mathematica.

Благодарности

Изследванията са частично извършени по проект МУ17ФМИ007 и СП17ФМИ 005 към НПД при ПУ „П. Хилендарски“.

ЛИТЕРАТУРА:

[1] Иванов, И., *Теория на игрите с икономически приложения*, Унив. изд-во „Кл. Охридски“, 2009.

[2] Гочева–Илиева, С., *Въведение в системата Mathematica*, Екс-Прес, 2009.

[3] Diaz, J., Jurado, I., Janiero, M., *An Introductory Course on Mathematical Game Theory*, AMS, 2010.

[4] Julmi, C., *Introduction to Game Theory*, BookBoon, 2012.

Яница Маринова Манджукова, докторант Радослава Сашкова Терзиева,
Снежана Христова
Пловдивски университет „Паисий Хилендарски“, катедра: Приложна математика и моделиране
E-mail: radoslavaterzieva@abv.bg (автор за кореспонденция)

ON THE FOUR-POINTS BOUNDS FOR THE EFFECTIVE CONDUCTIVITY OF MULTI-PHASE DISPERSIONS*

Krasimir D. Tsvyatkov

ABSTRACT: *Bounds on the effective conductivity derived by using the Hashin-Shtrikman variational principle are considered for dispersions of homogeneous spheres, which can have different, random conductivity σ_p. Generally, the bounds contain two statistical parameters $I_\sigma^{(3)}$ and $I_\sigma^{(4)}$ involving three- and four-points moments of conductivity field $\sigma(\boldsymbol{x})$, respectively. The structure of this parameters is investigated for such dispersions. It is turned out that $I_\sigma^{(3)}$ is represented by the well-known geometrical tree-points parameter ζ_2 and another two-points geometrical parameter, but $I_\sigma^{(4)}$ is represented with two two-points and with one three- and one four-points geometrical parameters. For 2D dispersions the four-points geometrical parameter is expressed by the parameter ζ_2.*

KEYWORDS: *effective conductivity, random dispersions, variational bounds*

ВЪРХУ ЧЕТИРИТОЧКОВИТЕ ГРАНИЦИ ЗА ЕФЕКТИВНАТА ПРОВОДИМОСТ НА МНОГОФАЗНИ ДИСПЕРСИИ

Красимир Д. Цвятков

1 Увод

Разглеждаме дисперсия от хомогенни непресичащи се сфери със случайна проводимост σ_p, разпределени случайно в неограничена матрица с проводимост σ_m, заемаща цялото d-мерно евклидово пространство $\mathbb{R}^n$, $d = 2, 3$. Връзката между плътността на електрическия ток $\boldsymbol{J}(\boldsymbol{x})$ и електрическото поле $\boldsymbol{E}(\boldsymbol{x})$ в средата се дава от закона на Ом:

$$\boldsymbol{J}(\boldsymbol{x}) = \sigma(\boldsymbol{x})\boldsymbol{E}(\boldsymbol{x}), \tag{1}$$

където $\sigma(\boldsymbol{x})$ е случайното поле на проводимостта на средата. Предполагаме, че тя е статистически хомогенна и изотропна. Тогава *ефективната проводимост* σ_e на средата обикновено се дефинира чрез равенството

$$\langle \boldsymbol{J}(\boldsymbol{x})\rangle = \sigma_e \langle \boldsymbol{E}(\boldsymbol{x})\rangle, \tag{2}$$

където скобите $\langle\ \rangle$ означават усреднение по ансамбъла от реализации на средата. Намирането на σ_e е основна задача в теорията на композитните материали и е поставена в трудове на класиците на съвременната наука, вж. например трактата на Максуел [1].

В общия случай ефективната проводимост σ_e зависи от пълното статистическо описание на средата. На практика, обаче, разполагаме само с информацията, давана от първите няколко корелационни функции за средата. Ето защо единственото, което може да се направи строго и последователно, е получаването на граници за σ_e. Разглеждайки класическия вариационен принцип

$$W\big[\widehat{\boldsymbol{E}}(\cdot)\big] = \frac{1}{2}\langle \sigma(\boldsymbol{x})\widehat{\boldsymbol{E}}(\boldsymbol{x}) \cdot \widehat{\boldsymbol{E}}(\boldsymbol{x})\rangle \longrightarrow \min \tag{3}$$

*Partially supported by Scientific Research Grant RD-08-119/2018 of Shumen University.

върху пробни полета от вида

$$\widehat{\boldsymbol{E}}(\boldsymbol{x}) = \langle \boldsymbol{E}\rangle + \lambda \langle \boldsymbol{E}\rangle \cdot \int \nabla\nabla G(\boldsymbol{x}-\boldsymbol{y})\sigma'(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \tag{4}$$

където λ е скаларен параметър, Беран [2] получава горна граница за σ_e при $d = 3$, включваща статистическия параметър

$$I_\sigma^{(3)} = \frac{1}{\langle \sigma'^3\rangle} \iint G_{,ij}(\boldsymbol{z})G_{,ji}(\boldsymbol{w})M_3^\sigma(\boldsymbol{z},\boldsymbol{w})\,\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{w}, \tag{5}$$

зависещ от триточковия момент $M_3^\sigma(\boldsymbol{z},\boldsymbol{w}) = \langle \sigma'(\boldsymbol{0})\sigma'(\boldsymbol{z})\sigma'(\boldsymbol{w})\rangle$ на флуктуацията $\sigma'(\boldsymbol{x}) = \sigma(\boldsymbol{x}) - \langle\sigma\rangle$ на полето $\sigma(\boldsymbol{x})$; в (4) и (5) $G(\boldsymbol{x})$ е фундаменталното решение на уравнението на Лаплас, т.е. $G(\boldsymbol{x}) = 1/4\pi|\boldsymbol{x}|$ при $d = 3$ и $G(\boldsymbol{x}) = -\ln|\boldsymbol{x}|/2\pi$ при $d = 2$. Ще отбележим също, че тук и по-нататък участващите градиенти на $G(\boldsymbol{x})$ се разглеждат като обобщени функции. Полетата (4) при $\lambda = 1/\langle\sigma\rangle$ представляват първите два члена от пертубационното разлагане на $\boldsymbol{E}(\boldsymbol{x})$ за слабо хетерогенна среда. При $d = 2$ границите от този тип са получени от Силнуцер [3].

Заедно с хетерогенната среда да разгледаме хомогенна неограничена среда с проводимост σ_0 и да въведем индуцираната поляризация

$$\boldsymbol{P}(\boldsymbol{x}) = (\sigma(\boldsymbol{x}) - \sigma_0)\boldsymbol{E}(\boldsymbol{x}) \tag{6}$$

относно тази среда за сравнение. Полето $\boldsymbol{P}(\boldsymbol{x})$ се подчинява на вариационния принцип на Хашин и Щтрикман [4], съгласно който функционалът

$$U\big[\widehat{\boldsymbol{P}}(\cdot)\big] = \sigma_0 \langle \boldsymbol{E}\rangle^2 + \langle 2\widehat{\boldsymbol{P}}(\boldsymbol{x})\cdot\langle \boldsymbol{E}\rangle + \widehat{\boldsymbol{P}}(\boldsymbol{x})\cdot\widehat{\boldsymbol{E}}'(\boldsymbol{x}) - \widehat{\boldsymbol{P}}(\boldsymbol{x})\cdot\widehat{\boldsymbol{P}}(\boldsymbol{x})/(\sigma(\boldsymbol{x})-\sigma_0)\rangle \tag{7}$$

е стационарен за истинското поле $\boldsymbol{P}(\boldsymbol{x})$ в хетерогенната среда и неговата стационарна стойност U_{st} е $U_{\mathrm{st}} = \sigma_e \langle \boldsymbol{E}\rangle^2$. При това U_{st} е минимум при $\sigma(\boldsymbol{x}) \le \sigma_0$ и максимум при $\sigma(\boldsymbol{x}) \ge \sigma_0$. В (7) $\langle \boldsymbol{E}\rangle^2 = \langle \boldsymbol{E}\rangle \cdot \langle \boldsymbol{E}\rangle$ и

$$\widehat{\boldsymbol{E}}'(\boldsymbol{x}) = \frac{1}{\sigma_0}\int \nabla\nabla G(\boldsymbol{x}-\boldsymbol{y})\cdot\widehat{\boldsymbol{P}}'(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}, \tag{8}$$

където $\widehat{\boldsymbol{P}}'(\boldsymbol{y})$ е флуктуацията на пробното поле $\widehat{\boldsymbol{P}}(\boldsymbol{y})$.

Според (6) аналогът на пробните полета (4) на Беран са полетата

$$\boldsymbol{P}(\boldsymbol{x}) = (\sigma(\boldsymbol{x}) - \sigma_0)\left\{\lambda_1\langle \boldsymbol{E}\rangle + \lambda_2\langle \boldsymbol{E}\rangle \cdot \int \nabla\nabla G(\boldsymbol{x}-\boldsymbol{y})\sigma'(\boldsymbol{y})\,\mathrm{d}\boldsymbol{y}\right\}, \tag{9}$$

като сега вариационният принцип на Хашин-Щтрикман позволява да въведем скаларен параметър и пред $\langle \boldsymbol{E}\rangle$. При $\lambda_2 = 0$ получаваме пробните полета $\boldsymbol{P}(\boldsymbol{x}) = \lambda(\sigma(\boldsymbol{x}) - \sigma_0)\langle \boldsymbol{E}\rangle$, които при екстремизиране на U относно λ водят до получаване на граници за σ_e, които в случая на двуфазна среда съвпадат с границите на Хашин и Щтрикман [4]. За N-фазна среда при $N \ge 3$ тези автори получават по-добри граници за σ_e, разглеждайки λ като функция на проводимостите $\sigma_1, \ldots, \sigma_N$ на компонентите на средата. Разглеждането сега на функционала U върху класа от пробни полета (9) при едновременно вариране на параметрите λ_1 и

λ_2 ще доведе до намиране на граници за σ_e, включващи още статистически параметър $I_\sigma^{(4)}$, зависещ от четвъртия момент

$$M_4^\sigma(\boldsymbol{z},\boldsymbol{v},\boldsymbol{w}) = \langle \sigma'(\mathbf{0})\sigma'(\boldsymbol{z})\sigma'(\boldsymbol{v})\sigma'(\boldsymbol{w})\rangle. \tag{10}$$

В случая $d=3$ тези граници бяха изведени от автора в [5, 6] и разглеждани за дисперсия от сфери с една и съща, неслучайна проводимост σ_p до ред ϕ_p^2, където ϕ_p е обемната концентрация на сферите в дисперсията. За дисперсии от сфери със случайна проводимост σ_p параметрите $I_\sigma^{(3)}$ и $I_\sigma^{(4)}$ съдържат в сложна интегрална форма както статистиката на разпределение на центровете на сферите (геометрията на средата), така и статистиката на разпределение на проводимостта σ_p (материалните параметри на средата). Тук ще изследваме структурата на параметрите $I_\sigma^{(3)}$ и $I_\sigma^{(4)}$, отделяйки една от друга геометричната и материалната статистика. За параметъра $I_\sigma^{(3)}$ това вече по същество е направено от автора в [7, 8], където границите на Беран и кластерните граници на Торкуато [9] се разглеждат под общ чадър. Изследването на $I_\sigma^{(3)}$ се свежда до разглеждането на два геометрични параметъра, единият от които съдържа дву- и три-точковата, а другият само дву-точковата функции на разпределение на центровете на сферите. Както ще видим по-долу, положението с параметъра $I_\sigma^{(4)}$ е значително по-сложно.

2 Четириточкови граници за ефективната проводимост

Рестрикцията $U(\lambda_1,\lambda_2)$ на функционала (7) върху класа от пробни полета (9) е квадратична форма на λ_1 и λ_2 :

$$U(\lambda_1,\lambda_2) = \left(a_0 + 2a_1\lambda_1 + 2a_2\lambda_2 + a_{11}\lambda_1^2 + 2a_{12}\lambda_1\lambda_2 + a_{22}\lambda_2^2\right)\langle \boldsymbol{E}\rangle^2, \tag{11}$$

където

$$a_0=\sigma_0,\quad a_1=\langle\sigma\rangle-\sigma_0,\quad a_2=-\frac{1}{d}\langle\sigma'^2\rangle,$$

$$a_{11} = -\frac{1}{d}\frac{\langle\sigma'^2\rangle}{\sigma_0} - (\langle\sigma\rangle-\sigma_0),\quad a_{12}=\frac{1}{d}\left\{\frac{1}{\sigma_0}\left[\langle\sigma'^3\rangle I_\sigma^{(3)} + (\langle\sigma\rangle-\sigma_0)\langle\sigma'^2\rangle\right] + \langle\sigma'^2\rangle\right\}, \tag{12}$$

$$\begin{aligned} a_{22} = \frac{1}{d}\Bigg\{\frac{1}{\sigma_0}\left[\langle\sigma'^4\rangle I_\sigma^{(4)} - 2\,(\langle\sigma\rangle-\sigma_0)\langle\sigma'^3\rangle I_\sigma^{(3)} - (\langle\sigma\rangle-\sigma_0)^2\langle\sigma'^2\rangle + \frac{1}{d^2}\langle\sigma'^2\rangle^2\right] \\ - \left[\langle\sigma'^3\rangle I_\sigma^{(3)} + (\langle\sigma\rangle-\sigma_0)\langle\sigma'^2\rangle\right]\Bigg\}; \end{aligned}$$

тук $I_\sigma^{(3)}$ е статистическият параметър (5), появяващ се в горната граница на Беран, а

$$I_\sigma^{(4)} = \frac{1}{\langle\sigma'^4\rangle}\iiint G_{,ij}(\boldsymbol{z})G_{,jk}(\boldsymbol{w})G_{,ki}(\boldsymbol{z}-\boldsymbol{v})M_4^\sigma(\boldsymbol{z},\boldsymbol{v},\boldsymbol{w})\mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{w} \tag{13}$$

е нов статистически параметър, който накратко коментирахме по-горе.

Екстримизирайки функцията $U(\lambda_1,\lambda_2)$, получаваме

$$\operatorname{extr} U(\lambda_1,\lambda_2) = \sigma_e(\sigma_0)\langle\boldsymbol{E}\rangle^2, \tag{14}$$

където

$$\sigma_{\text{HS}}^{(4)}(\sigma_0) = \sigma_0 + \frac{a_1 a_2 a_{12} - a_1^2 a_{22} - a_2^2 a_{11}}{a_{11} a_{22} - a_{12}^2}. \tag{15}$$

Съгласно вариационния принцип на Хашин и Щрикман величината $\sigma_{\text{HS}}^{(4)}(\sigma_0)$ представлява граница за ефективната проводимост σ_e : долна при $\sigma(\boldsymbol{x}) > \sigma_0$ и горна при $\sigma(\boldsymbol{x}) < \sigma_0$.

За двуфазна среда границите (15) са същите, получени от Фан-Тиен и Милтон [9], които използват разлагането в ред на Фурие на пробните полета (9), разглеждайки среда с периодична вътрешна структура при $d = 3$. За N-фазна среда при $N \geq 3$ техните граници са по-добри, тъй като по същество те разглеждат λ_1 и λ_2 като функции на проводимостите $\sigma_1, \ldots, \sigma_N$ на компонентите на средата, въвеждайки общо $2N$ вариращи се параметъра. По този начин те получават граници за σ_e, съдържащи $N^2(N-1)^2/4$ четириточкови и $N(N-1)^2/2$ триточкови геометрични параметъра. Макар и по-слабо ограничителни, разглежданите тук граници (15) се отнасят за произволна, не непременно дискретна случайна среда. Както границите на Фан-Тиен и Милтон, така и тези граници (даже границите при $\lambda_1 = 1$ и $\lambda_2 = 1/\langle\sigma\rangle$) са граници от четвърти ред в смисъл, че за N-фазна среда те съвпадат до ред $\delta\sigma_a \delta\sigma_b \delta\sigma_c \delta\sigma_d$, където $\delta\sigma_s = \sigma_s - \sigma_q$ при фиксирано q, $s = 1, \ldots, N$. В частност, за двуфазна среда от границите (15) получаваме

$$\frac{\sigma_e}{\sigma_2} = 1 + \sum_{k=1}^{4} \frac{1}{k!} R_k \left(\frac{\sigma_1 - \sigma_2}{\sigma_2}\right)^k + o\left(\left(\frac{\sigma_1 - \sigma_2}{\sigma_2}\right)^4\right), \tag{16}$$

с коефициенти $R_1 = \phi_1$, $R_2 = -2\,\phi_1\phi 2/d$ и

$$R_3 = \frac{6}{d}\,(\phi_1^2\phi_2 - A), \qquad R_4 = -\frac{24}{d^3}\left[\phi_1^2\phi_2(d^2\phi_1 - \phi_2) - d^2(Y_4 + 2\,\phi_1 A)\right], \tag{17}$$

където A и Y_4 са чисто геометрични параметри на средата, които са свързани с $I_\sigma^{(3)}$ и $I_\sigma^{(4)}$ чрез равенствата

$$\langle\sigma'^3\rangle\, I_\sigma^{(3)} = (\sigma_2 - \sigma_1)^3 A, \qquad \langle\sigma'^4\rangle\, I_\sigma^{(4)} = (\sigma_2 - \sigma_1)^4\, Y_4 \tag{18}$$

и се изразяват чрез моментите на индикаторната функция $\mathcal{I}_2(\boldsymbol{x})$ за фаза 2 (равна на 1, ако $\boldsymbol{x}$ лежи във фаза 2 и на 0, ако $\boldsymbol{x}$ лежи във фаза 1) чрез интегралите

$$A = \iint G_{,ij}(\boldsymbol{z}) G_{,ji}(\boldsymbol{w}) \langle\mathcal{I}_2(\boldsymbol{0})\mathcal{I}_2(\boldsymbol{z})\mathcal{I}_2(\boldsymbol{w})\rangle\, \mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{w}, \tag{19}$$

$$Y_4 = \iiint G_{,ij}(\boldsymbol{z}) G_{,jk}(\boldsymbol{w}) G_{,ki}(\boldsymbol{z} - \boldsymbol{v}) \langle\mathcal{I}_2(\boldsymbol{0})\mathcal{I}_2(\boldsymbol{z})\mathcal{I}_2(\boldsymbol{v})\mathcal{I}_2(\boldsymbol{w})\rangle\, \mathrm{d}\boldsymbol{z}\mathrm{d}\boldsymbol{v}\mathrm{d}\boldsymbol{w}; \tag{20}$$

в (16) и (17) ϕ_1 и $\phi_2 = 1 - \phi_1$ са обемните концентрации на двете фази с проводимости съответно σ_1 и σ_2.

За двуфазна среда границите от четвърти ред са изведени от Милтон [10, 11] чрез използване на аналитичен метод при условие, че са известни R_3 и R_4. Коефициентът R_3 е известен от триточковите граници на Беран:

$$R_3 = \frac{6}{d^2}\,\phi_1\phi_2\left[\phi_2 + (d-1)\zeta_1\right], \tag{21}$$

където ζ_1 и ζ_2 са добре изучени триточкови геометрични параметри, за които $\zeta_1 = 1 - \zeta_2$ и $0 < \zeta_2 < 1$, вж. например [12]. Разглеждайки σ_e като функция на σ_1 и σ_2 и позовавайки се на свойството $\sigma_e(\sigma_1, \sigma_2)\, \sigma_e(\sigma_2, \sigma_1) = \sigma_1\sigma_2$, което е в сила само при $d = 2$, Милтон [10, 11] показва, че

$$R_4 = 3\,\phi_1\phi_2^2\,(1 + \phi_2) - 2R_3(1 + 2\phi_2) \tag{22}$$

и по този начин получава в явна форма граници от четвърти ред за σ_e в двумерния случай, изразяващи се само чрез ζ_1 и ζ_2. Позовавайки се на апроксимацията на Паде за реда на Браун, Торкуато [13] извежда в друга форма същите тези граници за двуфазна среда.

Веднага се вижда, че от изразите за R_3 в (17) и (21) можем да изразим геометричния параметър A чрез ζ_2:

$$A = \frac{1}{d}\,\phi_1\phi_2\,[(d-1)(\zeta_2 - \phi_2) + \phi_1 - \phi_2]\,. \tag{23}$$

Имайки предвид изразите за R_3 и R_4 в (17), с помощта на равенството (22) можем да изразим също четириточковия параметър Y_4 чрез ζ_2 при $d = 2$:

$$Y_4 = \frac{1}{4}\phi_1\phi_2\,[(\phi_1 - \phi_2)(4\phi_2 - 3\zeta_2) - \phi 1]\,. \tag{24}$$

Очевидно практическото приложение на границите (15) за конкретна случайна структура зависи от възможността за пресмятане на статистическите параметри $I_\sigma^{(3)}$ и $I_\sigma^{(4)}$. Докато за параметъра $I_\sigma^{(3)}$ има съществен прогрес при пресмятането му за дисперсии от непресичащи се сфери с неслучайна проводимост (вж. например [12]), за параметъра $I_\sigma^{(4)}$ няма опити за неговото пресмятане за такива дисперсии при $d = 3$.

3 Статистическо описание на дисперсията

На всека сфера с център $\boldsymbol{x}_j$ да съпоставим неговата проводимост σ_j. Така получаваме системата $\{\boldsymbol{x}_j, \sigma_j\}$ от маркирани случайни точки $\boldsymbol{x}_j$, която може да се разглежда като система от точки, случайно разпределени в областта $\mathbb{R}^d \times Q$, където $Q = [0, +\infty)$ е интервалът на изменение на проводимостите σ_j. Да дефинираме случайното поле на плътността

$$\psi(\boldsymbol{x}, \sigma) = \sum_j \delta(\boldsymbol{x} - \boldsymbol{x}_j)\delta(\sigma - \sigma_j), \tag{25}$$

породено от системата $\{\boldsymbol{x}_j, \sigma_j\}$. Тази функция се дефинира чрез δ-функцията на Дирак и е въведена от Стратонович [14] за немаркирана система от точки $\{x_j\}$, случайно разпределени върху права. Нейното обобщение за маркирани случайни точки $\{\boldsymbol{x}_j\}$ с маркер радиуса a_j на сферата с център $\boldsymbol{x}_j$ при дисперсии от сфери със случайни радиуси се въвежда от Христов [15]. За дисперсии от сфери със случайни проводимости тази функция е въведена от автора в [5, 16]. С помощта на функцията $\psi(\boldsymbol{x}, \sigma)$ случайното поле на проводимостта $\sigma(\boldsymbol{x})$ може да се представи във вида

$$\sigma(\boldsymbol{x}) = \sigma_m + \widetilde{\sigma}(\boldsymbol{x}), \tag{26}$$

$$\widetilde{\sigma}(\boldsymbol{x}) = \iint (\sigma - \sigma_m) h(\boldsymbol{x} - \boldsymbol{y})\psi(\boldsymbol{y}, \sigma)\, \mathrm{d}\boldsymbol{y}\mathrm{d}\sigma, \tag{27}$$

където $h(\boldsymbol{x})$ е характеристична функция на сфера с център в координатното начало.

Случайната функция на плътността $\psi(\boldsymbol{y},\sigma)$ напълно определя дисперсията. Нейните моменти могат да се изразят чрез съвместните многоточкови вероятностни плътности $F_n(\boldsymbol{y}_1,\dots,\boldsymbol{y}_n;\sigma_1,\dots,\sigma_n)$ на разпределение на центровете $\boldsymbol{x}_j$ на сферите и техните проводимости σ_j, $n=1,2,\dots$. По-долу ще ни бъдат необходими изразите за първите четири момента, които се дават от формулите

$$\langle\psi(\boldsymbol{y},\sigma)\rangle = F_1(\boldsymbol{y};\sigma),$$
$$\langle\psi(\boldsymbol{y}_1,\sigma_1),\psi(\boldsymbol{y}_2,\sigma_2)\rangle = F_1(\boldsymbol{y}_1;\sigma_1)\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2}) + F_2(\boldsymbol{y}_1,\boldsymbol{y}_2;\sigma_1,\sigma_2),$$

$$\begin{aligned}\langle\psi(\boldsymbol{y}_1,\sigma_1),\psi(\boldsymbol{y}_2,\sigma_2,\psi(\boldsymbol{y}_3,\sigma_3)\rangle &= F_1(\boldsymbol{y}_1;\sigma_1)\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2})\delta(\boldsymbol{y}_{1,3})\delta(\sigma_{1,3})\\ &+ 3\left\{\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2})F_2(\boldsymbol{y}_2,\boldsymbol{y}_3;\sigma_2,\sigma_3)\right\}_s + F_3(\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{y}_3;\sigma_1,\sigma_2,\sigma_3),\end{aligned} \tag{28}$$

$$\begin{aligned}&\langle\psi(\boldsymbol{y}_1,\sigma_1),\psi(\boldsymbol{y}_2,\sigma_2,\psi(\boldsymbol{y}_3,\sigma_3),\psi(\boldsymbol{y}_4,\sigma_4)\rangle\\ &\quad= F_1(\boldsymbol{y}_1;\sigma_1)\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2})\delta(\boldsymbol{y}_{1,3})\delta(\sigma_{1,3})\delta(\boldsymbol{y}_{1,4})\delta(\sigma_{1,4})\\ &\quad+ 4\left\{\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2})\delta(\boldsymbol{y}_{1,3})\delta(\sigma_{1,3})F_2(\boldsymbol{y}_1,\boldsymbol{y}_4;\sigma_1,\sigma_4)\right\}_s\\ &\quad+ 3\left\{\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2})\delta(\boldsymbol{y}_{3,4})\delta(\sigma_{3,4})F_2(\boldsymbol{y}_1,\boldsymbol{y}_3;\sigma_1,\sigma_3)\right\}_s\\ &\quad+ 6\left\{\delta(\boldsymbol{y}_{1,2})\delta(\sigma_{1,2})F_3(\boldsymbol{y}_1,\boldsymbol{y}_3,\boldsymbol{y}_4;\sigma_1,\sigma_3,\sigma_4)\right\}_s + F_4(\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{y}_3,\boldsymbol{y}_4;\sigma_1,\sigma_2,\sigma_3,\sigma_4),\end{aligned}$$

където $m\{\,\}_s$ означава симетризация относно всичките m различни комбинации от индекси в скобите, $\boldsymbol{y}_{i,j}=\boldsymbol{y}_j-\boldsymbol{y}_i$ и $\sigma_{i,j}=\sigma_j-\sigma_i$, вж. [14, 15, 16].

Приемаме, че в дисперсията няма пространствени участъци, които да имат избирателност към сфери с различни проводимости. Това означава, че статистиката на разпределение на проводимостите σ_j на сферите е независима от тази на положенията им $\boldsymbol{x}_j$, т.е.

$$F_n(\boldsymbol{y}_1,\dots,\boldsymbol{y}_n;\sigma_1,\dots,\sigma_n) = f_n(\boldsymbol{y}_1,\dots,\boldsymbol{y}_n)P_n(\sigma_1,\dots,\sigma_n), \tag{29}$$

където $f_n(\boldsymbol{y}_1,\dots,\boldsymbol{y}_n)$ са съвместните многоточкови вероятностни плътности на разпределение на центровете $\boldsymbol{x}_j$ на сферите, а $P_n(\sigma_1,\dots,\sigma_n)$ са съвместните вероятностни плътности на разпределение на техните проводимости σ_j, $n=1,2,\dots$. Ще приемем също, че тези проводимости са статистически независими, т.е.

$$P_n(\sigma_1,\dots,\sigma_n) = P(\sigma_1)\dots P(\sigma_n), \tag{30}$$

където $P(\sigma)$ е вероятностната плътност на проводимостта σ на сфера от дисперсията. Ще отбележим още, че от предположената статистическа хомогенност на дисперсията следва, че $f_1(\boldsymbol{y})=n$, $f_k(\boldsymbol{y}_1,\boldsymbol{y}_2,\dots,\boldsymbol{y}_k)=n^k g_k(\boldsymbol{y}_2-\boldsymbol{y}_1,\dots,\boldsymbol{y}_k-\boldsymbol{y}_1)$, където n е средният брой центрове на сфери в единица обем и $g_k(\boldsymbol{y}_{1,2},\dots,\boldsymbol{y}_{1,k})$ са k-точковите функции на разпределение, $k=2,3,\dots$.

4 Статистическите параметри за дисперсията

Имайки предвид (26), да изразим четвъртия момент $M_4^\sigma(\boldsymbol{z},\boldsymbol{v},\boldsymbol{w})$ от (10) с моментите на полето $\widetilde{\sigma}(\boldsymbol{x})$. Така получаваме

$$\begin{aligned}M_4^\sigma(\boldsymbol{z},\boldsymbol{v},\boldsymbol{w}) &= \langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{z})\widetilde{\sigma}(\boldsymbol{v})\widetilde{\sigma}(\boldsymbol{w})\rangle\\ &- \langle\widetilde{\sigma}\rangle\left[\langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{z})\widetilde{\sigma}(\boldsymbol{v})\rangle + \langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{z})\widetilde{\sigma}(\boldsymbol{w})\rangle + \langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{v})\widetilde{\sigma}(\boldsymbol{w})\rangle + \langle\widetilde{\sigma}(\boldsymbol{z})\widetilde{\sigma}(\boldsymbol{v})\widetilde{\sigma}(\boldsymbol{w})\rangle\right]\\ &+ \langle\widetilde{\sigma}\rangle^2\left[\langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{z})\rangle + \langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{v})\rangle + \langle\widetilde{\sigma}(\mathbf{0})\widetilde{\sigma}(\boldsymbol{w})\rangle\right.\\ &\left.+ \langle\widetilde{\sigma}(\boldsymbol{z})\widetilde{\sigma}(\boldsymbol{v})\rangle + \langle\widetilde{\sigma}(\boldsymbol{z})\widetilde{\sigma}(\boldsymbol{w})\rangle + \langle\widetilde{\sigma}(\boldsymbol{v})\widetilde{\sigma}(\boldsymbol{w})\rangle\right] - 3\langle\widetilde{\sigma}\rangle^4,\end{aligned} \tag{31}$$

където $\langle\widetilde{\sigma}\rangle = (\langle\sigma_p\rangle - \sigma_m)\phi_p$. Позовавайки се сега на интегралната форма (27) на полето $\widetilde{\sigma}(\boldsymbol{x})$, изразяваме неговите моменти чрез моментите на полето $\psi(\boldsymbol{y},\sigma)$ и след това, използвайки формулите (28) за тези моменти при предположенията (29) и (30), за статистическия параметър $I_\sigma^{(4)}$ получаваме

$$\begin{aligned}\langle\sigma'^4\rangle I_\sigma^{(4)} &= (\langle\sigma_p\rangle-\sigma_m)^4\left[X_4^{(4)}-X_3^{(3)}+X_2^{(2)}-3X_1^{(1)}\right]\phi_p^4\\&+(\langle\sigma_p\rangle-\sigma_m)^2\langle(\sigma_p-\sigma_m)^2\rangle\left[X_3^{(4)}-X_2^{(3)}+X_1^{(2)}\right]\phi_p^3\\&+(\langle\sigma_p\rangle-\sigma_m)\langle(\sigma_p-\sigma_m)^3\rangle\left[X_2^{(4)}-X_1^{(3)}\right]\phi_p^2\\&+\langle(\sigma_p-\sigma_m)^2\rangle^2\,\widetilde{X}_2^{(4)}\phi_p^2\,+\,\langle(\sigma_p-\sigma_m)^4\rangle\,X_1^{(4)}\phi_p\,,\end{aligned}\tag{32}$$

където $X_n^{(m)}$ е изразът, съдържащ в интегрална форма n-тата корелационна функция $g_n(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_n)$, който се получава при заместване на полето $\widetilde{\sigma}(\boldsymbol{x})$ в m-тите му моменти в (31), $m,n=1,\ldots,4$. Същият смисъл има и означението на израза $\widetilde{X}_2^{(4)}$. Използвайки свойствата на функцията на Грийн $G(\boldsymbol{x})$ и нютоновия потенциал $\varphi(\boldsymbol{x})=\int h(\boldsymbol{x}-\boldsymbol{y})G(\boldsymbol{y})\mathrm{d}\boldsymbol{y}$, за изразите $X_1^{(m)}$ лесно намираме

$$X_1^{(1)}=X_1^{(4)}=-\frac{1}{d^2},\quad X_1^{(2)}=-\frac{d^2+2d+3}{d^2},\quad X_1^{(3)}=-\frac{2(d+1)}{d^2}.\tag{33}$$

За изследването на границите (15) е удобно да подредим $\langle\sigma'^4\rangle\,I_\sigma^{(4)}$ по степените на $(\langle\sigma_p\rangle-\sigma_m)$ и моментите на случайната проводимост σ_p на сферите. Така получаваме представянето

$$\begin{aligned}\langle\sigma'^4\rangle\,I_\sigma^{(4)} &= (\langle\sigma_p\rangle-\sigma_m)^4\,Y_4+(\langle\sigma_p\rangle-\sigma_m)^2\langle\sigma_p'^2\rangle\,Y_3+(\langle\sigma_p\rangle-\sigma_m)\langle\sigma_p'^3\rangle\,Y_2\\&+\langle\sigma_p'^2\rangle^2\,\widetilde{Y}_2-\frac{1}{d^2}\langle\sigma_p'^4\rangle\phi_p,\end{aligned}\tag{34}$$

където

$$\begin{aligned}Y_4 &= \left(X_4^{(4)}-X_3^{(3)}+X_2^{(2)}-3X_1^{(1)}\right)\phi_p^4+\left(X_3^{(4)}-X_2^{(3)}+X_1^{(2)}\right)\phi_p^3\\&\quad+\left(X_2^{(4)}+\widetilde{X}_2^{(4)}-X_1^{(3)}\right)\phi_p^2+X_1^{(4)}\phi_p,\\ Y_3 &= \left(X_3^{(4)}-X_2^{(3)}+X_1^{(2)}\right)\phi_p^3+\left[3\left(X_2^{(4)}-X_1^{(3)}\right)+2\,\widetilde{X}_2^{(4)}\right]\phi_p^2+6\,X_1^{(4)}\phi_p,\\ Y_2 &= \left(X_2^{(4)}-X_1^{(3)}\right)\phi_p^2+4\,X_1^{(4)}\phi_p,\quad \widetilde{Y}_2=\widetilde{X}_2^{(4)}\phi_p^2.\end{aligned}\tag{35}$$

От (34) веднага забелязваме, че за дисперсия с неслучайна проводимост σ_p на сферите е в сила равенството $\langle\sigma'^4\rangle I_\sigma^{(4)} = (\langle\sigma_p\rangle-\sigma_m)^4\,Y_4$, което сравнено с второто равенство в (18) при $\sigma_p=\sigma_2$ и $\sigma_m=\sigma_1$ показва, че тук Y_4 е същият четириточков геометричен параметър, представен в (20). Така установихме, че пресмятането на параметъра $I_\sigma^{(4)}$ се свежда до пресмятането на четири геометрични параметъра: четириточковия Y_4, триточковия Y_3 и двата двуточкови Y_2 и $\widetilde{Y}_2$. Според (24) в двумерният случай на дисперсия от цилиндрични влакна параметърът Y_4 се изразява чрез триточковия параметър ζ_2. Ще отбележим само, че даже и тогава оставащите параметри Y_2, $\widetilde{Y}_2$ и Y_3 се изразяват в по-сложна интегрална форма от съответните им геометрични параметри за параметъра $I_\sigma^{(3)}$. Ето защо се налага обмислянето на числена процедура за тяхното пресмятане.

ЛИТЕРАТУРА:

[1] Maxwell, J.C., A Treatise on Electricity and Magnetism. Vol.1, Ch. 9, article No 310-315, pp. 435-441, 1st edn. Oxford, United Kingdom: Clarendon Press (1873).

[2] Beran, M.J., Use of the variational approach to determine bounds for the effective permittivity in random media. Nuovo Cimento **38** (1965), 771–782.

[3] Silnutzer, N.R., Effective constants of statistically homogeneous materials, Ph.D. thesis, University of Pennsylvania (1972).

[4] Hashin, Z., Shtrikman, S., A variational approach to the theory of the effective magnetic permeability of multiphase materials. J. Appl. Phys., **33** (1962), 3125–3131.

[5] Tsvyatkov, Kr.D., On the method of functional series for evaluating the properties of random structured media, Ph.D. thesis, University of Sofia (1993) (In Bulgarian).

[6] Markov, K.Z., Zvyatkov, Kr.D., Functional series and Hashin-Shtrikman type bounds on the effective properties of random media, In: Recent Advances in Mathematical Modelling of Composite Materials, K. Z. Markov., Ed., World Sci. (1994), 59–106.

[7] Tsvyatkov, Kr.D., On the effective conductivity of multi-phase dispersions, Proc. of the Conference MATHTEX 2012, University of Shumen, Fac. Math. Inf., **1** (2013), 85–90 (in Bulgarian).

[8] Tsvyatkov, Kr.D., On the effective bulk modulus of multi-phase dispersions, Annuaire Univ. Shumen, Fac. Math. Inf., **XVIII** C (2017), 25–43 (in Bulgarian).

[9] Phan-Thien, N., Milton, G.W., New bounds on the effective thermal conductivity of N-phase materials. Proc. R. Soc. Lond. A **380** (1982), 333–348.

[10] Milton, G.W., Bounds on the transport and optical properties of a two-component composite material. J. Appl. Phys. **52** (1981), 5294–5304.

[11] Milton, G.W., Bounds on the elastic and transport properties of two-component composites. J. Mech. Phys. Solids **30** (1982), 177-191.

[12] Torquato, S., Random Heterogeneous Materials: Microstructure and Macroscopic Properties. Springer-Verlag (2002).

[13] Torquato, S., Effective electrical conductivity of two-phase disordered composite media. J. of Appl. Phys., **58** (1985), 3790–3797.

[14] Stratonovich, R.L., Topics in theory of random noises. Vol. 1, New York, Gordon and Breach (1967).

[15] Christov, C.I., A further development of the concept of random density function with application to Volterra-Wiener expansions. C. R. Acad. bulg. sci. **38** (1) (1985), 35–38.

[16] Zvyatkov, Kr.D., On the effective conductivity of a class of random dispersions, Annuaire Univ. Sofia, Fac. Math. Inf., Livre 2, **89** (1995), 217–235.

Красимир Димитров Цвятков
Шуменски университет “Еп. К. Преславски”, ФМИ, 9700 Шумен
E-mail: k.tsvyatkov@shu.bg

MATTEX 2018,
CONFERENCE PROCEEDING, v. 1, pp. 151 - 159

ITERATIVE METHODS FOR FINDING THE STABILIZING SOLUTION IN LQ THREE-PLAYER GAMES FOR POSITIVE SYSTEMS

NELI L. NEDELCHEVA-BAEVA

ABSTRACT: *We consider linear quadratic differential games for positive linear systems with an open loop information structure. We improved the existing methods for finding the Nash equilibrium, such as the Newton method and the Sylvester methods to obtain the stabilizing solution of Riccati equation to a three-player game. The sufficient conditions for convergence are derived. The proposed algorithms are illustrated on some numerical examples.*

KEYWORDS: *nonsymmetric algebraic Riccati equations, iteration methods, open loop Nash equilibrium, positive systems, generalized Riccati equation, stabilizing nonnegative solution*

ИТЕРАЦИОННИ МЕТОДИ ЗА ТЪРСЕНЕ НА СТАБИЛИЗИРАЩО РЕШЕНИЕ В ЛИНЕЙНОКВАДРАТИЧНИ ИГРИ С ТРИМА ИГРАЧИ ЗА ПОЛОЖИТЕЛНИ СИСТЕМИ

НЕЛИ Л. НЕДЕЛЧЕВА-БАЕВА

АБСТРАКТ: *Разглеждаме линейноквадратични игри за положителни системи с отворена информационна структура. Ние подобряваме съществуващите методи за търсене на равновесие на Наш, като методите на Нютон и на Силвестър за намиране на стабилизиращо решение на уравнението на Рикати за игра с трима играчи. Доказани са достатъчни условия за сходимост на методите. Предложените алгоритми са илюстрирани с някои изчислителни примери.*

1 Въведение

Разглеждаме линейноквадратична игра с отворена информационна структура за положителни системи като стратегиите на играчите се представят чрез стабилизиращо решение на свързано несиметрично уравнение на Рикати. Концепцията за равновесие на Наш в игрите, отчитащи различни информационни структури е въведена от W. van den Broek в [4, 5].

Приложенията на положителните системи се срещат естествено в екологичните и икономическите системи, много биологични модели и др.

Системи от вида

$$\dot{x} = Ax + \sum_{j=1}^{N} B_j u_j\,,\ \ j = 1,2,3\,,\quad x(0) = x_0\,, \tag{1}$$

се наричат положителни, ако за всички неотрицателни начални състояния x_0 и неотрицателни функции на управление u_j, то векторът на състоянието $x(t)$ е неотрицателен във всеки един момент. За управление на положителни системи от горния вид е необходимо да се реши матричното рикатиево уравнение от вида:

(2) $$-\begin{pmatrix} A^T & 0 & 0 \\ 0 & A^T & 0 \\ 0 & 0 & A^T \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}-\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}A-\begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \end{pmatrix}+\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}\begin{pmatrix} S_1 & S_2 & S_3 \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}=0,$$

където $(-A)$ е n x n Z-матрица, $S_j = B_j R_{jj}^{-1} B_j^T$ $(S_j = S_j{}^T)$ е неположителна матрица за $j = 1,2,3$, Q_j – симетрична квадратна неотрицателна матрица с размерност n, R_{jj} – симетрична квадратна отрицателно определена матрица от съответна размерност, B_j – неотрицателна матрица с размерност n x m_j, за $j = 1,2,3$, а X_1, X_2 и X_3 са неизвестни матрици. Ние ще използваме матрици от различни редове.

Ще използваме следните твърдения:

Теорема 1. (Теорема 1 в [1]) За Z-матрицата A са еквивалентни следните твърдения:

(i) A е M -матрица.

(ii) $A^{-1} \geq 0$.

(iii) $Av > 0$ за всеки вектор $v > 0$.

(iv) Всички собствени стойности на матрицата А имат положителни реални части, т.е. матрицата $-A$ е устойчива матрица.

Лема 1. Нека $-A \in R^{n\times n}$ е неособена M-матрица, β е отрицателно реално число $(-\beta > 0)$ и $C = -\beta I_n - A$, тогава C е неособена M-матрица. (виж Лема 2.4 [3]).

Дефиниция 1. Тройката (u_1^*, u_2^*, u_3^*) се нарича равновесна Наш стратегия за положителната система (1), ако $J_1(u_1^*, u_2^*, u_3^*) \geq J_1(u_1, u_2^*, u_3^*)$, $J_2(u_1^*, u_2^*, u_3^*) \geq J_2(u_1^*, u_2, u_3^*)$ и $J_3(u_1^*, u_2^*, u_3^*) \geq J_3(u_1^*, u_2^*, u_3)$ за всички неотрицателни начални състояния x_0 и всички допустими стратегии u_j, $j = 1,2,3$.

2 Метод на Нютон с параметри за трима играчи

Използваме итерационна формула (8) от [2], записана по следния начин:

$$-K^{(i+1)}\left(A - SK^{(i)}\right) - \left(D - K^{(i)}S\right)K^{(i+1)} = Q + K^{(i)}SK^{(i)}, i = 0,1,2 \ldots,$$

където $D = \begin{pmatrix} A^T & 0 & 0 \\ 0 & A^T & 0 \\ 0 & 0 & A^T \end{pmatrix}$, $Q = \begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \end{pmatrix}$, $S = (S_1 \quad S_2 \quad S_3)$, $K = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}$ и на база новите идеи от [3] извеждаме подобрена итерационна формула

(3) $$K^{(i+1)}\left(\beta I_n + A - SK^{(i)}\right) + \left(\alpha I_{3n} + D - K^{(i)}S\right)K^{(i+1)} = -Q - K^{(i)}SK^{(i)} + (\alpha+\beta)K^{(i)},$$

$i = 0,1,2$ ...., където α и β са отрицателни реални числа. Въвеждаме разходен функционал

(4) $$J_i(u_1, u_2, u_3) = \int_0^\infty \left(x^T Q_i x + \sum_{j=1}^3 u_j^T R_{ij} u_j\right) dt, \text{ за } i = 1 \div 3,$$

който се минимизира, а входната функция u_i е стратегията на i-тия играч.

Матриците X_1, X_2 и X_3 дефинират равновесна Наш стратегия (u_1^*, u_2^*, u_3^*) по следния начин $u_i^* = -R_{ii}^{-1} B_i^T X_i x^*$ за $i = 1,2,3$ като x^* е решение на системата (1).

Въвеждаме матрична функция $P(\boldsymbol{X}) = -D\boldsymbol{X} - \boldsymbol{X}A - Q + \boldsymbol{X}S\boldsymbol{X}$.

Доказваме, че методът е сходящ със следната теорема:

Теорема 2. Нека в положителната система (1) матрицата $-A$ е M-матрица. За матриците Q и S във функционалите на разходите (4) се предполага, че $Q \geq 0$ и $S \leq 0$. Предполагаме още, че съществува такова $\widehat{\boldsymbol{X}} = \begin{pmatrix} \widehat{\boldsymbol{K_1}} \\ \widehat{\boldsymbol{K_2}} \\ \widehat{\boldsymbol{K_3}} \end{pmatrix} \geq 0$, че $P(\widehat{\boldsymbol{X}}) > 0$, а редицата на

Нютон $\left(K^{(i)}\right)_{i\in N}$, $K^{(0)}=0$, е монотонно намаляваща и клони към решението $K\geq 0$. Решението K е най-малкото решение от множеството на всички неотрицателни решения.

Доказателство: $P(\widehat{\boldsymbol{X}})>0\ \Rightarrow\ -D\widehat{\boldsymbol{X}}-\widehat{\boldsymbol{X}}A-Q+\widehat{\boldsymbol{X}}S\widehat{\boldsymbol{X}}>0\ \Rightarrow\ \widehat{\boldsymbol{X}}S\widehat{\boldsymbol{X}}>\widehat{\boldsymbol{X}}A+D\widehat{\boldsymbol{X}}+Q$.

Началната матрица е $K^{(0)}=0$, като първият елемент $K^{(1)}$ се пресмята от уравнението на Силвестър $-K^{(i+1)}(\beta I_n+A)-(\alpha I_{3n}+D)K^{(i+1)}=Q$, което може да се запише като линейна система уравнения във вида

$[(-\beta I_n-A)^T\otimes I_{3n}+I_n\otimes(-\alpha I_{3n}-D)]vecK^{(1)}=vecQ$, където $vec: R^{i\mathrm{x}j}\to R^{ij\mathrm{x}1}$.

Означаваме $L^{(0)}=(-\beta I_n-A)^T\otimes I_{3n}+I_n\otimes(-\alpha I_{3n}-D)$ и записваме уравнението

$$L^{(0)}\ vecK^{(1)}=vecQ.$$

Тъй като $-A$ е M-матрица ($(-A)^{-1}\geq 0$), то от **Лема 1** следва, че $-\beta I_n-A$ е M-матрица. Матрицата $\left(-D\right)^{-1}\geq 0$, съгласно **Теорема 1.** следва, че $-D$ е M-матрица. Тъй като $\alpha<0$, съгласно **Лема 1** следва, че $-\alpha I_{3n}-D$ е M-матрица. Следователно $L^{(0)}$ също е M-матрица, което е еквивалентно на $\left(L^{(0)}\right)^{-1}\geq 0$, и правим извода, че $K^{(1)}\geq 0$.

Нека $L^{(i)}:=\left[\left(-\beta I_n-A+SK^{(i)}\right)^T\otimes I_{3n}+I_n\otimes\left(-\alpha I_{3n}-D+K^{(i)}S\right)\right]$, $\quad i=0,1,2\ ....$, където α и β са отрицателни реални числа.

При така направените предположения с метода на математическата индукция ще докажем, че са изпълнени следните свойства: **1**. $K^{(i)}\leq K^{(i+1)}$, **2**. $K^{(i)}\leq\widehat{\boldsymbol{X}}$ и **3**. $L^{(i)}$ е M-матрица. Ще докажем тези свойства при $i=0$:

Свойство **1**. $K^{(0)}=0$ и доказахме, че $K^{(1)}\geq 0$, следователно $0=K^{(0)}\leq K^{(1)}$.

Свойство **2**. $K^{(0)}=0$, а по допускане $\widehat{\boldsymbol{X}}\geq 0$, следователно $0=K^{(0)}\leq\widehat{\boldsymbol{X}}$.

Свойство **3**. По-горе доказахме, че $L^{(0)}$ е M-матрица.

Допускаме, че свойствата $1\div 3$ са изпълнени за $i>0$.Ще докажем, че са изпълнени за $i+1$. За да докажем свойство 2. за $i+1$ ще използваме (3) и $P(\widehat{\boldsymbol{X}})>0$. Разглеждаме матрицата

$$\begin{aligned}&\left(\widehat{\boldsymbol{X}}-K^{(i+1)}\right)\left(\beta I_n+A-SK^{(i)}\right)+\left(\alpha I_{3n}+D-K^{(i)}S\right)\left(\widehat{\boldsymbol{X}}-K^{(i+1)}\right)\\&<\widehat{\boldsymbol{X}}S\widehat{\boldsymbol{X}}+(\alpha+\beta)(\widehat{\boldsymbol{X}}-K^{(i)})-\widehat{\boldsymbol{X}}SK^{(i)}-K^{(i)}S\widehat{\boldsymbol{X}}+K^{(i)}SK^{(i)}\\&=\left(\widehat{\boldsymbol{X}}-K^{(i)}\right)S\left(\widehat{\boldsymbol{X}}-K^{(i)}\right)+(\alpha+\beta)(\widehat{\boldsymbol{X}}-K^{(i)})\leq 0,\end{aligned}$$

по допускане $\widehat{\boldsymbol{X}}-K^{(i)}\geq 0$, а $S\leq 0$ и $\alpha+\beta<0$. Тъй като $L^{(i)}$ е M-матрица заключаваме, че $\widehat{\boldsymbol{X}}-K^{(i+1)}\geq 0$, следователно $K^{(i+1)}\leq\widehat{\boldsymbol{X}}$.

Ще докажем свойство 3. за $i+1$ като използваме (3), $P(\widehat{\boldsymbol{X}})>0$ и следното равенство:

$$\begin{aligned}&-K^{(i+1)}\left(\beta I_n+A-SK^{(i+1)}\right)-\left(\alpha I_{3n}+D-K^{(i+1)}S\right)K^{(i+1)}\\&=\left(K^{(i+1)}-K^{(i)}\right)S\left(K^{(i+1)}-K^{(i)}\right)+Q-(\alpha+\beta)K^{(i)}+K^{(i+1)}SK^{(i+1)}\end{aligned}\tag{5}$$

Разглеждаме матрицата

$\left(\widehat{\boldsymbol{X}}-K^{(i+1)}\right)\left(\beta I_n+A-SK^{(i+1)}\right)+\left(\alpha I_{3n}+D-K^{(i+1)}S\right)\left(\widehat{\boldsymbol{X}}-K^{(i+1)}\right)$,

използваме (5) и получаваме

$$\begin{aligned}&=\beta\widehat{\boldsymbol{X}}+\widehat{\boldsymbol{X}}A-\widehat{\boldsymbol{X}}SK^{(i+1)}+\alpha\widehat{\boldsymbol{X}}+D\widehat{\boldsymbol{X}}-K^{(i+1)}S\widehat{\boldsymbol{X}}\\&\quad+\left(K^{(i+1)}-K^{(i)}\right)S\left(K^{(i+1)}-K^{(i)}\right)+Q-(\alpha+\beta)K^{(i)}+K^{(i+1)}SK^{(i+1)}\\&<\widehat{\boldsymbol{X}}S\widehat{\boldsymbol{X}}+(\alpha+\beta)\widehat{\boldsymbol{X}}-\widehat{\boldsymbol{X}}SK^{(i+1)}-K^{(i+1)}S\widehat{\boldsymbol{X}}\\&\quad+\left(K^{(i+1)}-K^{(i)}\right)S\left(K^{(i+1)}-K^{(i)}\right)-(\alpha+\beta)K^{(i)}+K^{(i+1)}SK^{(i+1)}\\&=\left(K^{(i+1)}-K^{(i)}\right)S\left(K^{(i+1)}-K^{(i)}\right)+\left(\widehat{\boldsymbol{X}}-K^{(i+1)}\right)S\left(\widehat{\boldsymbol{X}}-K^{(i+1)}\right)+(\alpha+\beta)\left(\widehat{\boldsymbol{X}}-K^{(i)}\right)\leq 0,\end{aligned}$$

тъй като $\widehat{\boldsymbol{X}} - K^{(i+1)} \geq 0$ от свойство 2. и по допускане $\alpha + \beta < 0$ и $K^{(i)} \leq K^{(i+1)}$. Правим извода, че $L^{(i+1)} \; vec\big(K^{(i+1)} - \widehat{\boldsymbol{X}}\big) < 0$ и $L^{(i+1)}$ е Z-матрица, а следователно и M-матрица.

Сега ще докажем свойство 1. за $i + 1$. Разглеждаме матрицата
$\big(K^{(i+2)} - K^{(i+1)}\big)\big(\beta I_n + A - SK^{(i+1)}\big) + \big(\alpha I_{3n} + D - K^{(i+1)}S\big)\big(K^{(i+2)} - K^{(i+1)}\big)$, използваме (5) и получаваме
$= \big(K^{(i+1)} - K^{(i)}\big)S\big(K^{(i+1)} - K^{(i)}\big) + (\alpha + \beta)\big(K^{(i+1)} - K^{(i)}\big) \leq 0,$
тъй като по допускане $K^{(i+1)} - K^{(i)} \geq 0$, а $S \leq 0$ и $\alpha + \beta < 0$. Следователно $K^{(i+2)} - K^{(i+1)} \geq 0$, т.е. $K^{(i+2)} \geq K^{(i+1)}$.

Редицата на Нютон е монотонно растяща и ограничена отгоре от $\widehat{\boldsymbol{X}}$, следователно има граница и я означаваме с K, т.е. $\lim\limits_{i\to\infty} K^{(i)} = K \geq 0$.

3 Метод на Силвестър II с параметри за трима играчи

Използваме следната итерационна формула:

$$-(A^T - X_k^{(i)}\, S_k)\, X_k^{(i+1)} - X_k^{(i+1)}\Big(A - \textstyle\sum_{j=1}^N S_j X_j^{(i)}\Big) = Q_k + X_k^{(i)} S_k X_k^{(i)}, \text{ за } k = 1,2,3$$

и извеждаме подобрени итерационни формули

(6) $-(\alpha I + A^T - X_k^{(i)}\, S_k)\, X_k^{(i+1)} - X_k^{(i+1)}\Big(\beta I + A - \sum_{j=1}^N S_j X_j^{(i)}\Big) = Q_k + X_k^{(i)} S_k X_k^{(i)} - (\alpha + \beta) X_k^{(i)}$,
$k = 1,2,3.$

За изучаване на свойствата на метода (6) въвеждаме следните матрични функции:

$$P_k(X_1, X_2, X_3) \coloneqq -(\alpha I + A^T - X_k\, S_k)\, X_k - X_k\big(\beta I + A - \textstyle\sum_{j=1}^3 S_j X_j\big) \\ -Q_k - X_k\, S_k X_k + (\alpha + \beta) X_k, \qquad k = 1,2,3. \tag{7}$$

Чрез алгебрични преобразувания се показва, че за матричните функции $P_k(X_1, X_2, X_3)$, $k = 1,2,3$ е вярно следното представяне

$$P_k(Y_1, Y_2, Y_3, X_1, X_2, X_3) = -(\alpha I + A^T - Y_k\, S_k)\; X_k - (Y_k - X_k)\, S_k\, X_k - Q_k \\ -X_k\, S_k\, Y_k - X_k\big(\beta I + A - \textstyle\sum_{j=1}^3 S_j Y_j\big) - \sum_{j\neq k} X_k\, S_j\, \big(Y_j - X_j\big) + (\alpha + \beta) X_k, \tag{8}$$

$k = 1,2,3$, като Y_1, Y_2, Y_3 са симетрични матрици от съответната размерност.

Сходимостта на метода доказваме със следната теорема:

Теорема 3. Предполагаме съществуването на две матрици $\widehat{\boldsymbol{X}} = diag(\hat{X}_1, \hat{X}_2, \hat{X}_3)$ и $\boldsymbol{X}^{(0)} = diag(X_1^{(0)}, X_2^{(0)}, X_3^{(0)})$, за които $0 \leq \boldsymbol{X}^{(0)} \leq \widehat{\boldsymbol{X}}$, $P_k\Big(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\Big) \leq 0$ и $P_k\big(\hat{X}_1, \hat{X}_2, \hat{X}_3\big) \geq 0$, $k = 1,2,3$. Началната матрица $\boldsymbol{X}^{(0)}$ избираме, така че $\Big(-\Big(\alpha I + A^T - X_k^{(0)} S_k\Big)\Big)$ за $k = 1,2,3$ и $\Big(-\Big(\beta I + A^T - \sum_{j=1}^3 X_j^{(0)} S_j\Big)\Big)$ са М-матрици, където α и β са отрицателни реални числа, а I е единична матрица от ред n. Тогава матричните редици $\{X_1^{(i)}, X_2^{(i)}, X_3^{(i)}\}_{i=0}^\infty$, дефинирани чрез (6), притежават следните свойства:

(i) $\boldsymbol{X}^{(i+1)} \geq \boldsymbol{X}^{(i)}$, i=0,1,2,… ;

(ii) $\boldsymbol{X}^{(i+1)} \leq \widehat{\boldsymbol{X}}$ за i=0,1,2,… ;

(iii) $\Big(-\Big(\alpha I + A^T - X_k^{(i+1)} S_k\Big)\Big)$, $\Big(-\Big(\beta I + A^T - \sum_{j=1}^3 X_j^{(i+1)} S_j\Big)\Big)$, $k = 1,2,3$, са М-матрици за i=0,1,2,… ;

(iv) Матричните редици $\{X_1^{(i)}, X_2^{(i)}, X_3^{(i)}\}_{i=0}^\infty$ са сходящи към неотрицателно решение $\widetilde{\boldsymbol{X}} = (\tilde{X}_1, \tilde{X}_2, \tilde{X}_3)$ на уравнението (2), за което $\widetilde{\boldsymbol{X}} \leq \widehat{\boldsymbol{X}}$.

Доказателство. От
$P_1\Big(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\Big) = -\Big(\alpha I + A^T - X_1^{(0)} S_1\Big) X_1^{(0)}$

$$-X_1^{(0)}\left(\beta I + A - S_1X_1^{(0)} - S_2X_2^{(0)} - S_3X_3^{(0)}\right) - Q_1 - X_1^{(0)}S_1X_1^{(0)} + (\alpha+\beta)X_1^{(0)} \le 0$$

следва, че $P_1\left(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\right) = F_1^{(0)} \le 0$. Началната матрица $X^{(0)} = diag\left(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\right)$ избираме, така че $\left(-\left(\alpha I + A^T - X_j^{(0)}S_j\right)\right)$ за $j = 1,2,3$ и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(0)}S_j\right)\right)$ са М-матрици. Ще докажем, че $X_1^{(1)} - X_1^{(0)}$ е неотрицателна матрица. Образуваме разликата между итерационна формула (6) при k=1, i=0 и $P_1\left(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\right)$.

$0 - P_1\left(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\right)$

$= -(\alpha I + A^T - X_1^{(0)}S_1)X_1^{(1)} - X_1^{(1)}\left(\beta I + A - \sum_{j=1}^{3} S_jX_j^{(0)}\right) - Q_1 - X_1^{(0)}S_1X_1^{(0)} + (\alpha+\beta)X_1^{(0)}$

$-[-\left(\alpha I + A^T - X_1^{(0)}S_1\right)X_1^{(0)} - X_1^{(0)}\left(\beta I + A - \sum_{j=1}^{3} S_jX_j^{(0)}\right) - Q_1 - X_1^{(0)}S_1\, X_1^{(0)} + (\alpha+\beta)X_1^{(0)}]$

т.е.

$-F_1^{(0)} = -(\alpha I + A^T - X_1^{(0)}\, S_1)\;(X_1^{(1)} - X_1^{(0)} - (X_1^{(1)} - X_1^{(0)})(\beta I + A - \sum_{j=1}^{3} S_jX_j^{(0)})$.

Получихме линейно матрично уравнение на Силвестър спрямо неизвестната матрица $(X_1^{(1)} - X_1^{(0)})$, на което лявата страна е неотрицателна. Решението намираме като решим линейната система уравнения:

$$\left[\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(0)}S_j\right)\right)\otimes I_n + I_n\otimes\left(-\left(\alpha I + A^T - X_1^{(0)}S_1\right)\right)\right] vec\left(X_1^{(1)} - X_1^{(0)}\right) = -vec\, F_1^{(0)}.$$

Означаваме

(9) $L_k^{(i)} = -\left[\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(i)}S_j\right)\otimes I_n + I_n\otimes\left(\alpha I + A^T - X_k^{(i)}S_k\right)\right]$, k=1,2,3.

Следователно уравнението добива вида $L_1^{(0)}\;\, vec\left(X_1^{(1)} - X_1^{(0)}\right) = -vec\, F_1^{(0)}$.

Матрицата $\boldsymbol{X}^{(0)} = diag\left(X_1^{(0)}, X_2^{(0)}, X_3^{(0)}\right)$ е неотрицателна матрица, а $\left(-\left(\alpha I + A^T - X_1^{(0)}S_1\right)\right)$ и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(0)}S_j\right)\right)$ са М-матрици. Тогава получаваме, че матрицата $L_1^{(0)}$ е М-матрица, следователно $\left(L_1^{(0)}\right)^{-1} \ge 0$ и тъй като $-\,vec\, F_1^{(0)} \ge 0$, то

$$vec\left(X_1^{(1)} - X_1^{(0)}\right) = -\left(L_1^{(0)}\right)^{-1} vec\, F_1^{(0)} \ge 0,$$

което означава $X_1^{(1)} - X_1^{(0)} \ge 0$, или $X_1^{(1)} \ge X_1^{(0)}$. Аналогично могат да се докажат неравенствата $X_2^{(1)} \ge X_2^{(0)}$ и $X_3^{(1)} \ge X_3^{(0)}$. Следователно $\boldsymbol{X}^{(1)} \ge \boldsymbol{X}^{(0)}$. С това доказахме свойство (i) при i=0.

Ще докажем свойство (ii) при i=0. Знаем, че матриците $\left(-\left(\alpha I + A^T - X_1^{(0)}S_1\right)\right)$ и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(0)}S_j\right)\right)$ са М-матрици. Образуваме матрицата

(10) $\left(\alpha I + A^T - X_1^{(0)}S_1\right)\left(\hat{X}_1 - X_1^{(1)}\right) + \left(\hat{X}_1 - X_1^{(1)}\right)\left(\beta I + A - \sum_{j=1}^{3} S_jX_j^{(0)}\right)$.

Използваме итерационна формула (6) и за (10) пресмятаме

$= \alpha\hat{X}_1 + A^T\hat{X}_1 - X_1^{(0)}S_1\hat{X}_1 + \beta\hat{X}_1 + \hat{X}_1A - \sum_{j=1}^{3}\hat{X}_1S_jX_j^{(0)} + Q_1 + X_1^{(0)}S_1X_1^{(0)} - (\alpha+\beta)X_1^{(0)}$.

Използваме формула (7) за $P_1(\hat{X})$ и изразяваме

(11) $A^T\hat{X}_1 + \hat{X}_1A = -P_1\left(\hat{X}_1, \hat{X}_2, \hat{X}_3\right) + \sum_{j=1}^{3}\hat{X}_1S_j\hat{X}_j - Q_1$.

Така за матрицата (10) получаваме

$= -P_1\left(\hat{X}_1, \hat{X}_2, \hat{X}_3\right) + \sum_{j=1}^{3}\hat{X}_1S_j\hat{X}_j - Q_1$

$$+\alpha\hat{X}_1 - X_1^{(0)}S_1\hat{X}_1 + \beta\hat{X}_1 - \sum_{j=1}^{3}\hat{X}_1 S_j X_j^{(0)} + Q_1 + X_1^{(0)}S_1X_1^{(0)} - (\alpha+\beta)X_1^{(0)}$$
$$= -P_1\left(\hat{X}_1,\hat{X}_2,\hat{X}_3\right) + \left(\hat{X}_1 - X_1^{(0)}\right)S_1\left(\hat{X}_1 - X_1^{(0)}\right) + \sum_{j\neq 1}\hat{X}_1 S_j\left(\hat{X}_j - X_j^{(0)}\right) + (\alpha+\beta)\left(\hat{X}_1 - X_1^{(0)}\right)$$
$$= Z_1^{(1)}.$$

По допускане $P_1\left(\hat{X}_1,\hat{X}_2,\hat{X}_3\right) \geq 0$, S_1, S_2, S_3 са неположителни, $\hat{X}_1$, $\hat{X}_2$, $\hat{X}_3$ са неотрицателни, а също $\alpha+\beta<0$ и $\hat{X}_k \geq X_k^{(0)}$ за k=1,2,3. Следователно последната матрица е неположителна, $Z_1^{(1)} \leq 0$.

Тогава матрицата (10) задава следното матрично уравнение

$$-\left(\alpha I + A^T - X_1^{(0)}S_1\right)\left(\hat{X}_1 - X_1^{(1)}\right) - \left(\hat{X}_1 - X_1^{(1)}\right)\left(\beta I + A - \sum_{j=1}^{3}S_j X_j^{(0)}\right) = -Z_1^{(1)} \geq 0$$

или $L_1^{(0)}\ vec\left(\hat{X}_1 - X_1^{(1)}\right) = -vec\, Z_1^{(1)}$.

Следователно $\hat{X}_1 - X_1^{(1)} \geq 0$, или $\hat{X}_1 \geq X_1^{(1)}$. Аналогично могат да се докажат неравенствата $\hat{X}_2 \geq X_2^{(1)}$ и $\hat{X}_3 \geq X_3^{(1)}$. Следователно $\hat{\boldsymbol{X}} \geq \boldsymbol{X}^{(1)}$. С това доказахме свойство (ii) при i=0.

Ще докажем свойство (iii) при i=0. Разглеждаме матрицата

$$(12)\qquad \left(\alpha I + A^T - X_1^{(1)}S_1\right)\left(\hat{X}_1 - X_1^{(1)}\right) + \left(\hat{X}_1 - X_1^{(1)}\right)\left(\beta I + A - \sum_{j=1}^{3}S_j X_j^{(1)}\right)$$

От итерационна формула (6) при i=0 имаме

$$-(\alpha I + A^T - X_1^{(0)}\,S_1)\,X_1^{(1)} - X_1^{(1)}\left(\beta I + A - \sum_{j=1}^{3}S_j X_j^{(0)}\right) = Q_1 + X_1^{(0)}S_1X_1^{(0)} - (\alpha+\beta)X_1^{(0)}$$

и изразяваме

$$(13)\ -A^T X_1^{(1)} - X_1^{(1)}A = Q_1 - X_1^{(0)}S_1\left(X_1^{(1)} - X_1^{(0)}\right) - \sum_{j=1}^{3}X_1^{(1)}S_jX_j^{(0)} + (\alpha+\beta)\left(X_1^{(1)} - X_1^{(0)}\right).$$

Използваме (11) и (13) и за матрицата (12) получаваме

$$\left(\alpha I + A^T - X_1^{(1)}S_1\right)\left(\hat{X}_1 - X_1^{(1)}\right) + \left(\hat{X}_1 - X_1^{(1)}\right)\left(\beta I + A - \sum_{j=1}^{3}S_j X_j^{(1)}\right)$$
$$= -P_1\left(\hat{X}_1,\hat{X}_2,\hat{X}_3\right) + \left(\hat{X}_1 - X_1^{(1)}\right)S_1(\hat{X}_1 - X_1^{(1)}) + \sum_{j\neq 1}\hat{X}_1 S_j\left(\hat{X}_j - X_j^{(1)}\right)$$
$$+(X_1^{(1)} - X_1^{(0)})S_1(X_1^{(1)} - X_1^{(0)}) + \sum_{j\neq 1}X_1^{(1)}S_j\left(X_j^{(1)} - X_j^{(0)}\right) + (\alpha+\beta)(\hat{X}_1 - X_1^{(0)})$$
$$= W_1^{(1)}.$$

Така стигаме до равенството

$$(14)\qquad \left(\alpha I + A^T - X_1^{(1)}S_1\right)\left(\hat{X}_1 - X_1^{(1)}\right) + \left(\beta I + \hat{X}_1 - X_1^{(1)}\right)\left(A - \sum_{j=1}^{3}S_j X_j^{(1)}\right) = W_1^{(1)}.$$

Тъй като $\hat{X}_k \geq 0$, $P_k\left(\hat{X}_1,\hat{X}_2,\hat{X}_3\right) \geq 0$, $\hat{X}_k - X_k^{(1)} \geq 0$ за k=1,2,3, а също и $X_k^{(1)} \geq 0$, $X_k^{(1)} - X_k^{(0)} \geq 0$, $S_k \leq 0$ за k=1,2,3 и $\alpha+\beta<0$, то $W_1^{(1)} \leq 0$.

Разглеждаме матрицата $\hat{X}_1 - X_1^{(1)}$ като неотрицателно решение на уравнението на Силвестър във вида

$$-\left(\alpha I + A^T - X_1^{(1)}S_1\right)\left(\hat{X}_1 - X_1^{(1)}\right) - \left(\hat{X}_1 - X_1^{(1)}\right)\left(\beta I + A - \sum_{j=1}^{3}S_j X_j^{(1)}\right) = -W_1^{(1)}.$$

Представяме решението: $0 \leq \mathrm{vec}\,(\hat{X}_1 - X_1^{(1)}) = -\left(L_1^{(1)}\right)^{-1} vec\, W_1^{(1)}$.

Последното означава, че $\left(L_1^{(1)}\right)^{-1}$ е неотрицателна матрица, т.е. матриците $\left(-\left(\alpha I + A^T - X_1^{(1)}S_1\right)\right)$ и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3}X_j^{(1)}S_j\right)\right)$ са M-матрици.

Аналогично може да се докаже, че $\left(L_2^{(1)}\right)^{-1}$ и $\left(L_3^{(1)}\right)^{-1}$са неотрицателни матрици, т.е. матриците $\left(-\left(\alpha I + A^T - X_k^{(1)} S_k\right)\right)$, k=2,3 и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(1)} S_j\right)\right)$ са М-матрици.

С това доказахме свойство (iii) при i=0.

Допускаме, че за i=r е изпълнено:

(a_r) $\boldsymbol{X}^{(r+1)} \geq \boldsymbol{X}^{(r)} \geq 0$,

(b_r) $\boldsymbol{X}^{(r+1)} \leq \widehat{\boldsymbol{X}}$,

$(c_r)\left(-\left(\alpha I + A^T - X_k^{(r+1)} S_k\right)\right)$, k=1,2,3 и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(r+1)} S_j\right)\right)$, са М-матрици.

Трябва да докажем, че е изпълнено и при i=r+1

(a_{r+1}) $\boldsymbol{X}^{(r+2)} \geq \boldsymbol{X}^{(r+1)} \geq 0$,

(b_{r+1}) $\boldsymbol{X}^{(r+2)} \leq \widehat{\boldsymbol{X}}$,

(c_{r+1}) $\left(-\left(\alpha I + A^T - X_k^{(r+2)} S_k\right)\right)$, k=1,2,3 и $\left(-\left(\beta I + A^T - \sum_{j=1}^{3} X_j^{(r+2)} S_j\right)\right)$ са М-матрици.

Пресмятаме матриците $X_1^{(r+2)}$, $X_2^{(r+2)}$ и $X_3^{(r+2)}$ по итерационни формули (6). Доказателството се извършва по аналогичен начин като първо доказваме, че $P_k\left(X_1^{(r+1)}, X_2^{(r+1)}, X_3^{(r+1)}\right)$ са неположителни матрици чрез представянето (8).

По метода на математическата индукция следва, че свойства (i), (ii) и (iii) са изпълнени при i=0,1,2,…

От свойствата на членовете на матричните редици $\{ X_1^{(i)}, X_2^{(i)}, X_3^{(i)}\}_{i=0}^{\infty}$ правим извода, че тези редици от неотрицателни матрици са монотонни и ограничени и следователно сходящи. Означаваме границите на редиците съответно $\tilde{X}_1, \tilde{X}_2, \tilde{X}_3$. След граничен преход в итерационни формули (6) се установява, че неотрицателните матрици $\tilde{X}_1, \tilde{X}_2, \tilde{X}_3$ са решения на уравнението (2).

4 Числени експерименти

Ще направим експерименти с цел сравняване изчислителните качества на двата метода: метод на Нютон с параметри (МНП) и метод на Силвестър II ускорен с параметри (МС2УП) за трима играчи.

На всяка стъпка формираме нови матрични коефициенти и след това с двата метода едновременно търсим решението. Пресмятанията се извършват с точност $\varepsilon = 10e-11$ на компютър с 64-битова операционна система Windows 7 Professional и хардуерни параметри: процесор – Intel Core i3 CPU M380 @ 2.53 GHz, RAM памет – 4 GB. Матричните коефициенти A, B_i, Q_i и R_{ii} за $i = 1,2,3$ са дефинирани чрез MatLab *(R2015a)*. Като резултат за всяка стойност на n и всеки метод ние отбелязваме следните параметри: "Max It" – най-големият брой итерации, "Av It" – средният брой итерации и „CPU time“ – процесорното време, в секунди, за работа на съответния метод за съответните повторения.

Началните матрици избираме $X_1^{(0)} = X_2^{(0)} = X_3^{(0)} = 0$, т.е. $K^{(0)} = 0$ и след това пресмятаме $K^{(1)}$, $K^{(2)}$, …, $K^{(i)}$, … по итерационна формула (3) за метода на Нютон. Редицата е сходяща и намереното решение изпълнява условията на **Теорема 2**. За метода на Силвестър II ускорен $X_1^{(0)} = X_2^{(0)} = X_3^{(0)} = 0$, $X_1^{(1)}$се пресмята по итерационна формула

(6), $X_1^{(1)}$ се използва за пресмятане на $X_2^{(1)}$ и съответно $X_1^{(1)}$ и $X_2^{(1)}$ се използват за пресмятане на $X_3^{(1)}$, което води до ускоряване на метода. В резултат на тези пресмятания получената редица $X^{(1)}$, $X^{(2)}$, …, $X^{(i)}$, … е сходяща и изпълнява условията на **Теорема 3**. Участващите матрици са М-матрици и според **Теорема 1**. притежават свойството стабилност, което означава, че намерените решения са стабилизиращи решения.

Пример 1.

```
Alpha=-0.001; Beta=-0.002; k=1:100
A=abs(randn(n))/100;  s=max(abs(eig(A)))+5;
for i=1:n, A(i,i)=-(A(i,i))-s; end
B1 = zeros(n,1);    B1(1)=abs(randn(1,1))/2;
B2=eye(n,n);        B2(n,n)=n/3;
B3=B2;
Q1=zeros(n,n); Q1(1,1)=n; Q1(n,n)=1;
R11=-1;
Q2=Q1; Q3=Q2;
R22= eye(n,n); R22(1,1)=-40;  R22(n,n)=-40;
R33=R22;
```

Таблица 1. **При 100 повторения**

n	МНП			МС2УП		
	Max It	Av It	CPU time	Max It	Av It	CPU time
4	5	4,05	**0,391s**	5	4,06	0,889s
5	5	4,09	**0,264s**	5	4,13	1,109s
6	8	4,10	**0,360s**	11	4,26	1,217s
7	5	4,12	**0,357s**	5	4,29	1,389s
8	5	4,21	**0,968s**	6	4,66	1,685s

В пример 1. МНП е по-бърз от МС2УП.

Пример 2.

```
Alpha=-0.001; Beta=-0.002; k=1:200
A=abs(randn(n))/10;  s=max(abs(eig(A)))+5;
for i=1:n, A(i,i)=-(A(i,i))-s; end
B1 = zeros(n,1);    B1(1)=abs(randn(1,1));
B2=0.5*eye(n,n);        B2(n,n)=sqrt(n);
B3=B2;
Q1=0.25*eye(n,n); Q1(1,1)=n/50; Q1(n,n)=1/50;
R11=-0.75;
Q2=0.05*eye(n,n);
Q3=0.01*eye(n,n);
R22=-10*eye(n,n); R22(1,1)=-50;  R22(n,n)=-50;
R33=R22;
```

Таблица 2. **При 200 повторения**

n	МНП			МС2УП		
	Max It	Av It	CPU time	Max It	Av It	CPU time
100	4	3,405	**60,642s**	4	3,405	89,405s
120	4	3,360	**99,341s**	4	3,360	127,043s
140	4	3,395	**176,009s**	4	3,395	184,196s
160	4	3,315	270,605s	4	3,315	**257,840s**
180	4	3,290	387,325s	4	3,290	**327,692s**
200	4	3,290	526,239s	4	3,290	**410,850s**

В пример 2. МС2УП става по-бърз от МНП при размерност на матричните коефициенти $n \geq 160$.

5 Изводи

Направихме експерименти за изчисляване на стабилизиращо решение на несиметричното уравнение на Рикати (2) и сравнихме резултатите за двата предложени метода. МНП е по-бърз от МС2УП за малки размерности на матричните коефициенти. МС2УП изпреварва и става по-бърз от МНП за $n \geq 160$. ПМС2УП решава три уравнения на Силвестър на всяка итерационна стъпка. МНП решава линейни уравнения с големи размерности. Блочната структура при МНП го забавя при големи стойности на n. МС2УП има предимство в случай на решаване на три независими линейни матрични уравнения. По този начин МС2УП е ефективна алтернатива на МНП.

ЛИТЕРАТУРА:

[1] Azevedo-Perdicoulis, T., Jank, G. Linear quadratic Nash games on positive linear systems, European Journal of Control, 11, (2005),1-13.

[2] Jank, G., Kremer, D., Open loop Nash games and positive systems – solvability conditions for nonsymmetric Riccati equations. Proceedings of MTNS, Katolieke Universiteit, Leuven, Belgium, (2004).

[3] Ma, Ch., Lu, H., Numerical study on nonsymmetric algebraic Riccati equations, Mediterranean Journal of Mathematics, Vol. 13, Issue 6, (2016), 4961-4973.

[4] van den Broek, W. Uncertainty in Diferential Games. PhD-thesis Univ. Tilburg, Netherlands, (2001).

[5] van den Broek, W., Engwerda, J., Schumacher, J. Robust Equilibria in Indefinite Linear Quadratic Diferential Games, Journal of Optimization Theory and Applications, 119, 3, (2003), 565-595.

Нели Баева
E-mail: nelilnb@abv.bg

FOR MATRIX EQUATION $X - A^*XA - B^*X^{-1}B = I^*$

Aynur A. Ali

ABSTRACT: *In this paper we studi the matrix equation* $X - A^*XA - B^*X^{-1}B = I$. *Sufficient conditions for the existence of a positive definite solution have been obtained. An iterative method is proposed to find a positive definite solution of the equation under consideration. It is proved a convergence of the considered method. The theoretical results are illustrated by numerical examples.*

KEYWORDS: *nonlinear matrix equation, fixed point iteration, inversion free iteration, convergence rate*

ЗА МАТРИЧНОТО УРАВНЕНИЕ $X - A^*XA - B^*X^{-1}B = I$

Айнур А. Али

АБСТРАКТ: *В тази работа изучаваме матричното уравнение* $X - A^*XA - B^*X^{-1}B = I$. *Получени са достатъчни условия за съществуване на положително определено решение. Предложен е итерационен метод за намиране на положително определно решение на разглежданото уравнение. Доказана е сходимостта на този итерационен процес. Теоретичните резултати са илюстрирани с числени примери.*

1 Въведение

В тази статия разглеждаме матричното уравнение

$$X - A^*XA - B^*X^{-1}B = I, \tag{1}$$

където A и B са квадратни матрици, а I е единичната матрица.

В случай на $B = 0$, уравнението (1) се свежда до познатото уравнение на Стейн $X - A^*XA = I$, а при $A = 0$, се свежда до добре изученото уравнение $X - B^*X^{-1}B = I$ [2, 3, 4, 5, 6] и тяхната литература. За него е доказано, че винаги има положително определено решение.

D.Gao [1] изследват уравнението

$$X - A^*X^pA - B^*X^{-q}B = I \quad (0 < p, q < 1). \tag{2}$$

В [1] е доказано, че уравнението (2) винаги има единствено положително определено решение и са предложени два итерацинни метода за получаване на това решение. През последните години обект на изследване са и уравненията $X + A^*X^{-1}A - B^*X^{-1}B = I$ [13, 14, 18], $X + A^*X^{-1}A + B^*X^{-1}B = I$ [7, 8, 12], $X + \sum_{i=1}^{m} A_i^*X^{-1}A_i = I$ [10, 11, 15, 16].

Мотивирани от работите [1, 3, 14, 18], в тази работа изследваме уравнението (1). Получени са достатъчни условия за съществуването на положително определено решение на това уравнение. Предложен е итерационен метод за намиране на положително определено решение и е изследвана нейната сходимост. Накрая, с някои числени експерименти е илюстрирано теоретичното съдържание.

*Partially supported by Scientific Research Grant RD-08-145/2018 of Shumen University.

Използвани са следните означения: $A > 0(A \geq 0)$, означава че A е ермитова положително определена (полуопределена) матрица; за $N \geq M > 0$, с $[M, N]$ означаваме множеството от матрици $\{X : M \leq X \leq N\}$; $\lambda(C)$ е собствена стойност на $n \times n$ ермитова матрица C; $\rho(A)$ е спектралния радиус; и $\|A\|$ е спектралната норма ($\|A\| = \sqrt{\rho(A^*A)}$).

2 Основни резултати

Преди да изкажем основните резултати ще въведем една помощна лема, която ще използваме.

Лема 2.1. ***[17, Lemma 2.2.]*** *Нека C и W са квадратни матрици.*

1. *ако $\rho(C) < 1$, тогава уравнението на Стейн $X - C^*XC = W$ има единствено решение P и $P \geq 0(P > 0)$, когато $W \geq 0(W > 0)$*
2. *ако има някаква матрица $P > 0$, такава, че $P - C^*PC$ е положително определена(полуопределена) матрица, то $\rho(C) < 1(\rho(C) \leq 1)$*

Нека въведем и докажем първия основен резултат в следната теорема

Теорема 2.1. *Нека уравнението (1) има положително определено решение X. Тогава*

(i) *$\rho(A) < 1$,*

(ii) *$\rho(X^{-1}B) < 1$,*

(iii) *$X \geq M$, където M е решение на уравнението $X - A^*XA = I$.*

Доказателство:

Нека уравнение (1) има положително определено решение X. Уравнение (1) записваме по следния начин

$$X - A^*XA = I + B^*X^{-1}B. \tag{3}$$

За дясната страна на уравнение (3) при $X > 0$ имаме $I + B^*X^{-1}B > 0$. От условие 2. на Лема 2.1 за уравнението (3) следва, че $\rho(A) < 1$. При доказателството на (ii) използваме аналогични разсъждения на горното.

От (1) получаваме

$$X - (BX^{-1})^*X(X^{-1}B) = I + A^*XA. \tag{4}$$

За дясната страна на уравнението (4) имаме $I + A^*XA > 0$. Отново от Лема 2.1 следва, че

$$\rho(X^{-1}B) < 1.$$

При наличие на положително определено решение на уравнение (1) доказахме, че $\rho(A) < 1$. Съгласно Лема 2.1 точка 1 при $\rho(A) < 1$ уравнението $X - A^*XA = I$ има единствено решение $M > 0$, което се задава с формулата $M = \sum_{i=0}^{\infty}(A^*)^iA^i$ [17]. Разглеждаме

$$X - A^*XA = I + B^*X^{-1}B$$

и

$$M - A^*MA = I.$$

След почленно изваждане на горните уравнения получаваме

$$(X - M) - A^*(X - M)A = B^*XB. \tag{5}$$

За дясната страна на уравнението (5) при $X > 0$ имаме $B^*XB \geq 0$. Съгласно Лема 2.1 точка 1 имаме $X \geq M$. $\square$

Сега разглеждаме следния итерационен процес

$$\left\{ \begin{array}{l} X_0 = I, \;\; Y_0 = \beta I, \;\; \beta > 1 \\ X_{k+1} = I + A^*X_kA + B^*Y_k^{-1}B, \quad k = 0, 1, \ldots \\ Y_{k+1} = I + A^*Y_kA + B^*X_k^{-1}B, \end{array} \right. \tag{6}$$

Теорема 2.2. *Ако $\|A\|^2 + \|B\|^2 < 1$, то уравнението (1) има единствено положително определено решение X. Освен това итерационния процес (6) за*

$$\beta \geq \frac{1 + \|B\|^2}{1 - \|A\|^2}$$

е сходящ към единственото положително определено решение $X \in [I, \beta I]$.

Доказателство: За редиците $\{X_k\}$ и $\{Y_k\}$, генерирани от итерационния процес (6) имаме

$$X_1 = I + A^*X_0A + B^*Y_0^{-1} = I + A^*A + \frac{1}{\beta}B^*B \geq I = X_0$$

и

$$Y_1 = I + A^*Y_0A + B^*X_0^{-1}B = I + \beta A^*A + B^*B \leq \beta I = Y_0.$$

Разглеждаме разликата на Y_1 и X_1,

$$\begin{array}{rcl} Y_1 - X_1 & = & \beta A^*A + B^*B - A^*A - \dfrac{1}{\beta}B^*B \\ & = & (\beta - 1)A^*A + (1 - \dfrac{1}{\beta})B^*B \geq 0 \end{array}$$

Следователно $X_0 \leq X_1 \leq Y_1 \leq Y_0$. Нека $X_{k-1} \leq X_k \leq Y_k \leq Y_{k-1}$. Ще докажем, че $X_k \leq X_{k+1} \leq Y_{k+1} \leq Y_k$. Разглеждаме

$$\begin{array}{rcl} X_{k+1} - X_k & = & A^*X_kA + B^*Y_k^{-1}B - A^*X_{k-1}A + B^*Y_{k-1}^{-1}B \\ & = & A^*(X_k - X_{k-1})A + B^*(Y_k^{-1} - Y_{k-1}^{-1})B \\ & \geq & 0, \end{array}$$

$$\begin{array}{rcl} Y_k - Y_{k+1} & = & A^*Y_{k-1}A + B^*X_{k-1}^{-1}B - A^*Y_kA - B^*X_k^{-1}B \\ & = & A^*(Y_{k-1} - Y_k)A + B^*(X_{k-1}^{-1} - X_k^{-1})B \\ & \geq & 0, \end{array}$$

$$\begin{aligned} Y_{k+1} - X_{k+1} &= A^*Y_kA + B^*X_k^{-1}B - A^*X_kA - B^*Y_k^{-1}B \\ &= A^*(Y_k - X_k)A + B^*(X_k^{-1} - Y_k^{-1})B \\ &\geq 0. \end{aligned}$$

Съгласно метода на математическата индукция редиците $\{X_k\}$ и $\{Y_k\}$ са сходящи. Нека $\lim_{n\to\infty} X_n = X$ и $\lim_{n\to\infty} Y_n = Y$. От (6) имаме

$$\begin{aligned} X &= I + A^*XA + B^*Y^{-1}B \\ Y &= I + A^*YA + B^*X^{-1}B. \end{aligned}$$

След почленно изваждане на двете уравнения, получаваме

$$Y - X = A^*(Y - X)A + B^*X^{-1}(Y - X)Y^{-1}B$$

и

$$\begin{aligned} \|X - Y\| &\leq \|A\|^2\|Y - X\| + \|X^{-1}B\|\|Y^{-1}B\|\|Y - X\| \\ &\leq \|A\|^2\|Y - X\| + \|X^{-1}\|\|B\|\|Y^{-1}\|\|B\|\|Y - X\| \\ &\leq (\|A\|^2 + \|B\|^2)\|Y - X\|. \end{aligned}$$

Тъй като $(\|A\|^2 + \|B\|^2) < 1$, $X \equiv Y$. Следователно уравнението има решение.

Нека X и $\widetilde{X}$ са две различни положително определени решения на уравнението (1). Тогава

$$\begin{aligned} X &= I + A^*XA + B^*X^{-1}B, \\ \widetilde{X} &= I + A^*\widetilde{X}A + B^*\widetilde{X}^{-1}B, \\ \widetilde{X} - X &= A^*(\widetilde{X} - X)A - B^*X^{-1}(\widetilde{X} - X)\widetilde{X}^{-1}B. \end{aligned}$$

От тук

$$\|\widetilde{X} - X\| \leq (\|A\|^2 + \|B\|^2)\|\widetilde{X} - X\|. \tag{7}$$

Тъй като $(\|A\|^2 + \|B\|^2) < 1$, $\widetilde{X} \equiv X$.

С това теоремата е доказана. □

3 Числени експерименти

В тази секция даваме някои примери, които илюстрират, че матричните редици $\{X_k\}$ и $\{Y_k\}$, дефинирани от (6), са сходящи към единственото положително определено решение $\widehat{X}$ на уравнението (1).

Разглеждаме

$$R(X_k) = \|X_k - A^*X_kA - B^*X_k^{-1}B - I\|.$$

За стоп критерий използваме

$$\|Y_k - X_k\| \leq 10^{-10}.$$

Пример 3.1. *Разглеждаме матричното уравнение (1) с*

$$A = \frac{1}{56}\begin{pmatrix} 1 & 5 & 3 & 2 \\ -1 & -6 & 3 & 4 \\ -4 & 3 & 7 & 5 \\ 1 & 8 & 2 & 1 \end{pmatrix}, \quad B = \frac{1}{70}\begin{pmatrix} 7 & 9 & 6 & 8 \\ 7 & 5 & 8 & 3 \\ 9 & 8 & 6 & 7 \\ 11 & 5 & 9 & 3 \end{pmatrix}.$$

Разглеждаме итерационния процес (6) с $X_0 = I$ и $Y_0 = \frac{1+\|B\|^2}{1-\|A\|^2} = 1.2256$. За матриците A и B имаме $(\|A\|^2 + \|B\|^2) = 0.2141 < 1$. След 11 итерации по итерационния метод (6) получаваме

$$X_{11} \approx Y_{11} \approx \begin{pmatrix} 1.0592 & 0.0411 & 0.0349 & 0.0227 \\ 0.0411 & 1.0788 & 0.0450 & 0.0320 \\ 0.0349 & 0.0450 & 1.0641 & 0.0466 \\ 0.0227 & 0.0320 & 0.0466 & 1.0403 \end{pmatrix}$$

$$R(X_{11}) = 4.7877e^{-11}, \quad R(Y_{11}) = 4.8721e^{-11}$$

$$R_{11} = 4.5543e^{-13}.$$

Пример 3.2. *Разглеждаме уравнение (1) с*

$$A = \frac{1}{820}\begin{pmatrix} 41 & 15 & 23 & 35 & 66 \\ 25 & 12 & 27 & 45 & 21 \\ 23 & 27 & 28 & 16 & 24 \\ 15 & 45 & 16 & 52 & 65 \\ 66 & 21 & 24 & 65 & 35 \end{pmatrix}, \quad B = \frac{1}{830}\begin{pmatrix} 23 & 21 & 23 & 25 & 32 \\ 21 & 45 & 60 & 42 & 33 \\ 23 & 24 & 34 & 18 & 17 \\ 13 & 42 & 18 & 44 & 30 \\ 32 & 33 & 26 & 30 & 26 \end{pmatrix}.$$

Разглеждаме итерационния процес (6) с $X_0 = I$ и $Y_0 = \frac{1+\|B\|^2}{1-\|A\|^2} = 1.1454$. За матриците A и B е изпълнено $(\|A\|^2 + \|B\|^2) = 0.1387 < 1$. След 9 итерации по итерационния процес (6) получаваме

$$X_9 \approx Y_9 \approx \begin{pmatrix} 1.0154 & 0.0107 & 0.0177 & 0.0174 & 0.0151 \\ 0.0107 & 1.0137 & 0.0173 & 0.0161 & 0.0148 \\ 0.0177 & 0.0173 & 1.0826 & 0.0216 & 0.0218 \\ 0.0174 & 0.0161 & 0.0216 & 1.0242 & 0.0209 \\ 0.0151 & 0.0148 & 0.0218 & 0.0209 & 1.0224 \end{pmatrix}$$

$$R(X_9) = 8.2086e^{-12}, \quad R(Y_9) = 8.2086e^{-12},$$

$$R_9 = 2.2219e^{-16}.$$

ЛИТЕРАТУРА:

[1] Gao, D., Iterative Methods for Solving the Nonlinear Matrix Equation $X - A^*X^pA - B^*X^{-1}B = I(0 < p, q < 1)$, Advances in Linear Algebra and Matrix Theory, 72-78, September 27, 2017

[2] Ferrante, A., Levy, B., Hermitian solution of the equation $X = Q + NX^{-1}N^*$, Linear Algebra Appl., 247:359–373, 1996.

[3] Guo, CH., Lancaster, P., Iterative solution of two matrix equations, Math Comput, 68:1589–1603, 1999.

[4] B. Meini, B., Efficient Computation of the Extreme Solutions of $X + A^*X^{-1}A = Q$ and $X - A^*X^{-1}A = Q$, Math. Comp., 71:1189–1204, 2002.

[5] Hasanov, V.I., Ivanov, I.G., On two Perturbation Estimates of the Extreme Solutions to the Matrix Equations $X \pm A^*X^{-1}A = Q$, Linear Algebra Appl., 413:81–92, 2006.

[6] Hasanov, V.I., Notes on two perturbation estimates of the extreme solutions to the equations $X \pm A^*X^{-1}A = Q$, Appl. Math. Comp., 216:1355–1362, 2010.

[7] Long, J., Hu, X., Zhang, L., On the Hermitian positive definite solution of the matrix equation $X + A^*X^{-1}A + B^*X^{-1}B = I$, Bull. Braz. Math. Soc., 39:371–386, 2008.

[8] Vaezzadeh, S., Vaezpour, S., Saadati, R., Park, C., The iterative methods for solving nonlinear matrix equation $X + A^*X^{-1}A + B^*X^{-1}B = Q$, Adcances in Difference Equations 2013, 2013:229.

[9] Ran, A.C.M., Reurings, M.C.B., On the nonlinear matrix equation $X + A^*\mathcal{F}(X)A = Q$: solution and perturbation theory, Linear Algebra Appl., 346:15–26, 2002.

[10] Konstantinov, M., Petkov, P., Popchev, I., Angelova, V., Sensitivity of the matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i^* X^{p_i} A_i = 0$, $\sigma_i = \pm 1$., Appl. Comput. Math., 10:409-427, 2011.

[11] Popchev, I., Konstantinov, M., Petkov, P., Angelova, V., Norm-wise, mixed and component-wise condition numbers of matrix equation $A_0 + \sum_{i=1}^{k} \sigma_i A_i^* X^{p_i} A_i = 0$, $\sigma_i = \pm 1$., Appl. Comput. Math., 14:18-30, 2014.

[12] Duan, X.F., Wang, Q.W., Li, Ch.M., Positive definite solution of a class of nonlinear matrix equation, Linear and Multilinear Algebra, 62(6):839-852, 2014.

[13] Berzig, M., Duan, X., Samet, B., Positive definite solution of the matrix equation $X = Q - A^*X^{-1}A + B^*X^{-1}B$ via Bhaskar-Lakshmikantham fixed point theorem, Mathematical Sciences, 2012, 6:27, 2012.

[14] Ali, A., Hasanov, V., On some sufficient conditions for the existence of a positive definite solution of the matrix equation $X + A^*X^{-1}A - B^*X^{-1}B = I$, AIP Conference Proceedings, V. 1690, 2015, Article number 060001

[15] Y. He, J. Long, On the Hermitian positive definite solution of the nonlinear matrix equation $X + \sum_{i=1}^{m} A_i^* X^{-1} A_i = I$, *Appl. Math. Comput.*, 216:3480–3485, 2010.

[16] X. Duan, C. Li, A. Liao, Solutions and perturbation analysis for the nonlinear matrix equation $X + \sum_{i=1}^{m} A_i^* X^{-1} A_i = I$, *Appl. Math. Comput.*, 218:4458–4466, 2011.

[17] Lancaster, P., Tismenetsky, M., The Theory of matrices. 2nd ed., San Diego(CA):Academic Press, 1985.

[18] Hasanov, V., On the matrix equation $X + A^*X^{-1}A - B^*X^{-1}B = I$, Linear and Multilinear Algebra, 66(9):1783-1798, 2018

Айнур Али,
Факултет по математика и информатика,
Шуменски университет "Еп. К. Преславски"
E-mail: `ainur_80@mail.bg`