CHAOS 2020

Proceedings

13th Chaotic Modeling and Simulation International Conference

Editor

Christos H. Skiadas

9-12 June, 2020

Imprint
Proceedings of the 13th Chaotic Modeling and Simulation International Conference (9-12 June, 2020)
Published by: ISAST: International Society for the Advancement of Science and Technology.
Editor: Christos H Skiadas

Preface

13th Chaotic Modeling and Simulation International Conference

9 – 12 June 2020

It is our pleasure to welcome the guests, participants and contributors to the 13th International Conference (CHAOS2020) on Chaotic Modeling, Simulation and Applications. We support the study of nonlinear systems and dynamics in an interdisciplinary research field and very interesting applications will be presented. We intend to provide a widely selected forum to exchange ideas, methods, and techniques in the field of Nonlinear Dynamics, Chaos, Fractals and their applications in General Science and in Engineering Sciences.

The principal aim of CHAOS2020 International Conference is to expand the development of the theories of the applied nonlinear field, the methods and the empirical data and computer techniques, and the best theoretical achievements of chaotic theory as well.

Chaotic Modeling and Simulation Conferences continue to grow considerably from year to year thus making a well established platform to present and disseminate new scientific findings and interesting applications.

We thank all the contributors to the success of this conference and especially the authors of this *Proceedings Volume.* Special thanks to the Plenary, Keynote and Invited Presentations, the Scientific Committee, the ISAST Committee, Yiannis Dimotikalis and Aris Meletiou and the web supporting team, the Conference Secretary Eleni Molfesi and all the members of the Secretariat.

November 2020

Christos H. Skiadas,
Conference Chair

Honorary Committee and Scientific Advisors

International Scientific Committee

Plenary – Keynote – Invited Speakers

Nail Akhmediev
Department of Theoretical Physics, Australian National University
Recent advances in rogue wave theory

Elena Babatsouli
Communicative Disorders Department, University of Louisiana at Lafayette, USA
Order in disordered speech data

Mark Edelman
Department of Physics, Stern College at Yeshiva University, New York, NY, USA
Courant Institute of Mathematical Sciences, New York University, New York, NY, USA
Department of Mathematics, BCC, CUNY, 2155 University Avenue, Bronx, New York, USA
Evolution of Systems with Power-Law Memory: Do We Have to Die?

Jean-Marc GINOUX
Laboratoire d'Informatique et des Systèmes, LIS. Université de Toulon, France
Archives Henri Poincaré, UMR CNRS 7117
Albert Einstein and the doubling of the deflection of light

Nikolay V . Kuznetsov
Department of Applied Cybernetics, Saint-Petersburg State University,
Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences
Theory of hidden oscillations (with a tribute to Gennady A. Leonov)

Wieslaw M. Macek
Faculty of Mathematics and Natural Sciences
Cardinal Stefan Wyszynski University, Warsaw, Poland and
Space Research Centre, Polish Academy of Sciences, Warsaw, Poland
Reconnection and Turbulence in Space Plasmas on Kinetic Scale

Riccardo Meucci
Istituto Nazionale di Ottica – CNR, Largo E. Fermi 6, 50125 Firenze, Italy
Recent Advances in Controlling Chaos

Contents

x

Higgs boson and Higgs field in fractal models of the Universe: active femtoobjects, new Hubble constants, solar wind, heliopause

Valeriy S. Abramov

Donetsk Institute for Physics and Engineering named after A.A. Galkin, Ukraine
(E-mail: vsabramov2018@gmail.com)

Abstract. Theoretically the relationship between the main parameters of active femtoobjects and the Higgs boson in fractal models of the Universe was investigated. To describe the structure of the solar wind, heliopause, new Hubble constants are proposed. Estimates of the main parameters are conformed with the experimental data obtained by the Planck space observatory (based on Fermi-LAT and Cerenkov telescopes), UTR-2 and URAN-2 radio telescopes, Parker Solar Probe, Voyager 2 and Voyager 1. Within the framework of the anisotropic model, a description of the main characteristics of the model femtoobject and its relationships with the parameters of the Higgs boson and the Higgs field was performed. To take into account the stochastic behavior of the parameters of a model femtoobject (an active object with dimensions of the order of the classical electron radius), random variables are introduced. Using the example of a hydrogen atom, we estimated the radius of a proton, its mean square deviation, and compared it with an experiment. Estimates of the anomalous contributions to the magnetic moments of leptons based on the lepton quantum number are obtained.
Keywords: model femtoobject, Higgs boson and Higgs field, fractal models of the Universe, Hubble constants, structure of the solar wind, heliopause, hydrogen atom, proton and electron radii, magnetic moments of leptons.

1 Introduction

To describe fractal cosmological objects (using binary black holes and neutron stars as an example), the model was proposed in [1, 2] that takes into account the relation between the parameters of the Higgs boson and relict photons, gravitons. Within the framework of this model, the possibility of radiation of gravitational waves from such cosmological objects in the superradiation regime is shown [2]. Higgs field accounting made it possible to propose an anisotropic model of fractal cosmology, within the framework of which it is possible to describe the effect of accelerated expansion of the Universe [3]. In this case, a transition to the description of atomic defects, active nanoobjects, and neutrinos is possible [4, 5]. Active objects in fractal quantum systems have their own characteristic features of behavior [6 - 8]. In this case, superradiative states of active objects may appear [7]. When describing various physical fields (gravitational, electromagnetic, neutrino, deformation, stress) in fractal quantum systems, it is necessary to take into account the ordering effect of the

13th CHAOS Conference Proceedings, 9 - 12 June 2020, Florence, Italy

corresponding operators [8]. Coherent laser spectroscopy methods and the modern development of nanotechnology make it possible to study active femtoobjects (protons, neutrons, atomic and muon hydrogens, leptons) in fractal quantum systems. Estimates of the characteristic sizes for the proton radius and Rydberg constant in atomic and muon hydrogens were obtained in [9–11]. Note that active femtoobjects such as leptons have anomalies in magnetic properties [12 - 14]. For neutrinos, the effect of oscillations (mutual transformations of the electron, muon neutrino and τ-neutrino into each other) is observed [13].
The relationships between the Higgs boson parameters and active nanoelements in fractal systems were studied in [15–17]. Features of the behavior of coupled states of a vortex–antivortex pair were considered in [16]. In [17], the description of the relations of the Higgs boson parameters with cosmological objects in the Universe was proposed. For the accelerated expansion of the Universe, within the framework of this model [17], the relationships of the Hubble constant (old value) with the parameters of the Higgs boson and relict radiation were obtained. The experimental data on the attenuation of gamma rays against an intergalactic background, obtained by the Planck space observatory (based on Fermi-LAT and Cerenkov telescopes), made it possible to determine new values of the Hubble constant and the density of matter in the Universe [18]. The authors explain these new values by the interaction of γ rays with relic photons. In this case, it becomes necessary to agreement the old and new values of the Hubble constants both within the framework of our model and with the cosmological model ΛCDM (plane cosmology). On the other hand, experimental data on the compound, structure, and behavior of the solar wind (flows of various particles) near the Sun [19 – 24], Earth [25] and in interstellar space (near the heliopause) [26 – 30] should also be associated with new values of the Hubble constant, the expansion rate, and the density of matter in the Universe.
The aim of this work is to describe the main characteristics of active femtoobjects, the solar wind, heliopause and their relationships with the parameters of the Higgs boson and the Higgs field in fractal models of the Universe.

2 Description of model femtoobject

The compound of the solar wind may include active nanoobjects [4 - 7] and femtoobjects. Based on the results of [1, 2, 4 - 7], we introduce the main parameters ξ_{2p}, Ω_{A0}, r_p of a model femtoobject

$$\xi_{2p} = \chi_0 / n'_F = 1/(N'_p - N)\,; \quad \Omega_{A0} = n_{A0} E_e / E_{H0}\,; \quad r_p = 2r_e / (z'_\mu n_F)\,, \qquad (1)$$

which are related with the known parameters of quantum electrodynamics

$$r_e = e^2/(m_e c_0^2)\,; \quad \hbar c_0 = e \cdot e_{\alpha 0}\,; \quad e_{\alpha 0} = e \cdot \alpha_0\,; \quad \alpha_0 = \hbar c_0 / e^2\,; \quad E_e = m_e c_0^2 = e^2 / r_e\,;$$

$$r_{0p} = m_e r_e / m_p = e^2 / E_p\,; \; E_p = m_p c_0^2 = e^2 / r_{0p}\,; \; \mu_B = e\hbar / 2m_e\,; \; \mu_N = e\hbar / 2m_p\,. \qquad (2)$$

Here r_e and r_{0p}, m_e and m_p, E_e and E_p are classical radii, rest masses, rest energies for electron and proton, respectively; c_0 is limited speed of light in vacuum; $\hbar$ is Planck's constant; e is electron charge; α_0 is fine structure constant; $e_{\alpha 0}$ is renormalized electron charge; μ_B is Bohr magneton; μ_N is nuclear magneton. Next we will use the numerical values $E_e = 0.51099907\,\text{MeV}$, $m_p/m_e = 1836.152701$, $E_p = 938.2723226\,\text{MeV}$, $r_e = 2.81794092\,\text{fm}$, $r_{0p} = 1.534698568\,\text{am}$. Note that in this work, model femtoobjects are active objects with sizes of the order of the classical electron radius r_e. Model attoobjects with sizes of the order of the classical proton radius r_{0p} describe the internal structure of nucleons (the presence of a core and scalar, vector clouds [12]). In fractal quantum systems (such as atomic and muon hydrogen), model attoobjects can lead to a change in the main parameters (1), anomalies in magnetic properties (2) and stochastic behavior [8] of model femtoobjects and leptons. In our model, the main parameters of the model femtoobject are related to the resting energy of the Higgs boson E_{H0}, the main parameter n_{A0} for black holes [1, 2], the number of quanta n_F, n'_F of the fermionic field ($n_F + n'_F = 1$) from the anisotropic model (taking into account the presence of the Higgs field) [3], and the cosmological redshift z'_μ [1, 2], the effective susceptibility χ_0 in the absence of the Higgs field [4–7] and the effective number N in the Dicke superradiation model [2]. The numerical values of these parameters are: $E_{H0} = 125.03238\,\text{GeV}$, $n_{A0} = 58.04663887$, $n_F = 0.945780069$, $n'_F = 0.054219931$, $z'_\mu = 7.18418108$, $\chi_0 = 0.257104198$, $N = 17.0073101$. Using formulas (1), we find the numerical values of the main parameters of the model femtoobject $\xi_{2p} = 4.741876161$, $\Omega_{A0} = 237.232775 \cdot 10^{-6}$, $r_p = 0.829458098\,\text{fm}$ and $N'_p = 17.21819709$.

To take into account the stochastic behavior of the parameters of the model femtoobject, we introduce a random variable $\hat{\xi}_{rp}$ with two possible values ξ_{1p}, ξ_{2p} and their corresponding probabilities P_{1p}, P_{2p}, and expected value $M(\hat{\xi}_{rp}) = 1$. Based on the parameters ξ_{2p}, Ω_{A0} from (1) we find the probabilities P_{1p}, P_{2p}, possible value ξ_{1p}, variance $D(\hat{\xi}_{rp})$, standard deviation $\sigma(\hat{\xi}_{rp})$

$$P_{1p} = \xi_{2p} / (\xi_{2p} + \Omega_{A0}); \quad P_{2p} = \Omega_{A0} / (\xi_{2p} + \Omega_{A0}); \quad P_{1p} + P_{2p} = 1;$$

$$\xi_{1p} = (1 - \xi_{2p} P_{2p}) / P_{1p}; \quad D(\hat{\xi}_{rp}) = (\xi_{2p} - \xi_{1p})^2 P_{1p} P_{2p}; \quad \sigma(\hat{\xi}_{rp}) = D^{1/2}(\hat{\xi}_{rp}). \quad (3)$$

The values of these parameters from (3) are equal: $P_{1p} = 0.999949973$, $P_{2p} = 50.027 \cdot 10^{-6}$, $\xi_{1p} = 0.999812796$, $D(\hat{\xi}_{rp}) = 700.495 \cdot 10^{-6}$, $\sigma(\hat{\xi}_{rp}) = 0.026466865$.

Next, we introduce a random variable $\hat{r}_p = r_p \cdot \hat{\xi}_{rp}$ with two possible values r_p^*, r_e^* and their corresponding probabilities P_{1p}, P_{2p}. If r_p is a constant value, then the possible values r_p^*, r_e^*, expected value $M(\hat{r}_p)$, variance $D(\hat{r}_p)$, standard deviation $\sigma(\hat{r}_p)$ are found by the formulas

$$r_p^* = \xi_{1p}\, r_p\,; \quad r_e^* = \xi_{2p}\, r_p\,; \quad M(\hat{r}_p) = r_p^* P_{1p} + r_e^* P_{2p} = r_p\,;$$

$$D(\hat{r}_p) = (r_e^* - r_p^*)^2 P_{1p} P_{2p}\,; \quad \sigma(\hat{r}_p) = D^{1/2}(\hat{r}_p)\,. \tag{4}$$

The numerical values are equal: $r_p^* = 0.82930282\,\text{fm}$, $r_e^* = 3.933187582\,\text{fm}$, $D(\hat{r}_p) = 481.936 \cdot 10^{-6}(\text{fm})^2$, $\sigma(\hat{r}_p) = 0.021953046\,\text{fm}$. Our calculated value of the proton radius r_p^* almost coincides with the new experimental value of 0.8293 fm for the proton radius in the hydrogen atom, obtained by 2S-4P spectroscopy (based on quantum interference) [11].

Based on the anisotropic model [1, 2, 4], we find the relationship of the radii r_p, r_p^* with other characteristic parameters r_p', x_p, y_p, $r_{p\tau}$, $r_{p\tau}^*$, $r_{p\mu}$

$$r_p' - r_p = x_p\,; \quad x_p = r_p\, \text{sn}(u_\mu;k_\mu)\,; \quad y_p = r_p\, \text{cn}(u_\mu;k_\mu)\,; \quad x_p^2 + y_p^2 = r_p^2\,;$$

$$r_p - y_p = 3(r_p' - r_{p\tau})\,; \quad r_{p\tau}^* = r_{p\tau}\, n_{F\tau}\,; \quad 2r_{p\mu} = r_p'(1 - S_{1u} - S_{2u}) - 4(r_p - r_p^*)\,. \tag{5}$$

The parameter $\text{sn}(u_\mu;k_\mu) = \sin\varphi_\mu = 0.057234291$ is related to the angle φ_μ [1, 2]; quantum numbers $n_{F\tau} = 0.950987889$, $n_{F\tau}' = 1 - n_{F\tau}$ are related with the lepton quantum number $\Omega_{\tau L} = (n_{F\tau}')^2 = 0.002402187$ from [5]; parameters $|S_{1u}| = 0.046741575$, $S_{2u} = 0.033051284$ defined in [4]. Further, based on expressions (5), we find the numerical values of the characteristic parameters: $r_p' = 0.876931544\,\text{fm}$, $x_p = 0.047473446\,\text{fm}$, $y_p = 0.828098429\,\text{fm}$, $r_{p\tau} = 0.876478321\,\text{fm}$, $r_{p\tau}^* = 0.833520268\,\text{fm}$, $r_{p\mu} = 0.841841587\,\text{fm}$. Our calculated values r_p' and $r_{p\mu}$ practically coincide with the values of 0.8768 fm (the CODATA value) and 0.84184 fm (determined on the basis of fine and ultrafine splitting in the framework of quantum electrodynamics) [9], respectively. Our calculated value $r_{p\tau}^*$ practically coincides with the value of 0.8335 fm for muonic hydrogen [10]. Our anisotropic model [1, 2, 4] also makes it possible to estimate the measurement error δr_p, $\delta r_p'$ using the formulas

$$\delta r_p = \chi_{32}\, r_p' = r_{p\chi}\, \text{sn}(u_\mu;k_\mu)[1 + \text{sn}(u_\mu;k_\mu)]\,; \quad r_{p\chi} = 2r_e\, \chi_{11} / (z_\mu'\, n_F)\,;$$

$$\delta r_p' = r_{d\tau} S_{2u}\,; \quad r_{d\tau} = |\chi_{ef}|\, r_{F\tau}\,; \quad r_{F\tau} = n_{F\tau}\, r_p'\,. \tag{6}$$

Taking into account χ_{11}=0.181800122, χ_{32}=0.010405201, $|\chi_{ef}|$=0.250425279 from [1, 2] and expressions (6) we find estimates of measurement errors δr_p= 0.009124649 fm, $\delta r'_p$= 0.006902512 fm, which do not disagree the experimental estimates of 0.0091 fm from [11], 0.0069 fm from [9], respectively. In this case, the calculated value of the radius $r_{d\tau}$= 0.208842481 fm in our model is near the mean square radius of the electric charge distribution in the core of nucleons equal to 0.21 fm [12]. The radius $r_{F\tau}$= 0.833951278 fm is related with the characteristic radii $r'_{F\tau}$, $r_{\tau L}$ and the value $\Omega'_{\tau L}$= 0.97597813 by the expressions

$$r'_{F\tau}= n'_{F\tau} r'_p ; \quad (r'_{F\tau})^2+(r_{\tau L})^2=(r'_p)^2 ;$$

$$r^2_{\tau L}= \Omega'_{\tau L}(r'_p)^2 ; \quad \Omega'_{\tau L}=1-\Omega_{\tau L}= n_{F\tau}(1+n'_{F\tau}) . \qquad (7)$$

The values of these radii are equal: $r'_{F\tau}$= 0.042980266 fm, $r_{\tau L}$= 0.866334751 fm. Anomalies in the magnetic moments of leptons can be determined by the influence of CMB radiation. In this case, relict radiation can lead to effects of renormalization of the initial parameters: fine structure constant α_0, electron charge e, limiting speed of photon propagation in vacuum c_0; rest masses m_e, m_μ, m_τ and magnetons μ_B, $\mu_\mu= e\hbar / 2m_\mu$, $\mu_\tau= e\hbar / 2m_\tau$ for electron, muon, τ-lepton, respectively. The magnetic moments of leptons $< \hat{\mu}_e >$, $< \hat{\mu}_\mu >$, $< \hat{\mu}_\tau >$ for an electron, muon, τ-lepton, respectively, are determined by the expressions

$$2 <\hat{\mu}_e>= (2+\Omega_{\mu e})\mu_B ; \quad 2 <\hat{\mu}_\mu>= (2+\Omega_{\mu\mu})\mu_\mu ; \quad 2 <\hat{\mu}_\tau>= (2+\Omega_{\mu\tau})\mu_\tau . \quad (8)$$

Anomalous contributions to magnetic moments and renormalization effects are described by parameters $\Omega_{\mu e}$, $\Omega_{\mu\mu}$, $\Omega_{\mu\tau}$ for electron, muon, τ-lepton, respectively, based on the lepton number $\Omega_{\tau L}$

$$\Omega_{\mu e}= \Omega_{\tau L}-\Omega_{HL} ; \quad \Omega_{HL}=E_{HL}/E_{H0} ; \quad E_{HL}= n'_{H3}E_e ; \quad N' = 17.21088699 ; \quad (9)$$

$$\Omega_{\mu\mu}= \Omega_{\tau L}-\Omega'_{NL} ; \quad \Omega'_{NL}= E'_{NL}/E_{H0} ; \quad E'_{NL}=N'E_e ; \quad (N'-N)\cdot\chi_0= n'_{\mu F} ; \quad (10)$$

$$\Omega_{\mu\tau}= \Omega_{\tau L}-0.5(\Omega_{HL}+\Omega_{GL}) ; \quad \Omega_{GL}=E_{GL}/E_{H0} ; \quad E_{GL}= n_G E_e . \qquad (11)$$

Additional contributions Ω_{HL}, Ω'_{NL}, Ω_{GL} are determined based on the energies E_{HL}, E'_{NL}, E_{GL} and the resting energy of the Higgs boson E_{H0}. From (9) - (11) it follows that these additional energies are determined by the numbers of quanta n'_{H3}, N', n_G and the rest energy of the electron E_e. Wherein

$$n'_{H3} = n_{H3}/(1+\Omega_{0\nu}) ; \quad 1+\Omega_{0\nu} = 1+(n'_F)^2 = 1+(N'_p-N)^2\cdot\chi_0^2 ; \qquad (12)$$

$$n_{H3} = Q_{H3} n_{h2} = 0.5 Q_{H3} n_{A0} ; \quad n_{A0} = z'_\mu(z'_\mu+1) - n_Q / n_g ; \quad n_Q = 2n_G . \qquad (13)$$

Here $n_g = 8$, $n_Q = 6$, $n_G =< \hat{c}_G \hat{c}^+_G >= 3$ and $n'_G =< \hat{c}^+_G \hat{c}_G >= 2$ can be

interpreted as the numbers of quanta of the gluon, quark, excited, and ground states of the gravitational fields, respectively; neutrino density $\Omega_{0\nu} = 0.002939801$ [4]. Based on (13) we find $n_{H3} = 20.33926863$. Further, taking into account (12), (10), we obtain $n'_{H3} = 20.27965049$, $n'_{\mu F} = 0.052340473$. Based on equations (9) - (11) we find the energies $E_{HL} = 10.36288254\,\text{MeV}$, $E'_{NL} = 8.794747246\,\text{MeV}$, $E_{GL} = 1.53299721\,\text{MeV}$; additional contributions $\Omega_{HL} = 82.88159067 \cdot 10^{-6}$, $\Omega'_{NL} = 70.33975716 \cdot 10^{-6}$, $\Omega_{GL} = 12.26080164 \cdot 10^{-6}$. The found parameters $\Omega_{\mu e}/2 = 1159.652705 \cdot 10^{-6}$, $\Omega_{\mu\mu}/2 = 1165.923621 \cdot 10^{-6}$, $\Omega_{\mu\tau}/2 = 1177.307902 \cdot 10^{-6}$ coincide with the data [14] for anomalies of the magnetic moments of leptons.

3 New Hubble Constants

The parameters of active nanoobjects and femtoobjects are related with cosmological parameters. To describe accelerated expansion of the Universe in model I [17] and the anisotropic model [1, 2, 4], the Hubble constants H_{01}, H_{02}, H_0, characteristic distances L_{01}, L_{02}, L_0, speeds υ_{01}, υ_{02}, υ_0 were introduced

$$H_{01} = c_0 / L_{01} = \upsilon_{01} / L_0; \quad H_{02} = c_0 / L_{02} = \upsilon_{02} / L_0; \quad H_0 = \upsilon_0 / L_0. \qquad (14)$$

The values $L_0 = 1\text{Mpc}$, $H_{01} = 73.2\,\text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$, $L_{01} = 4.0954948\,\text{Gpc}$ (distance to supernova type 1a), $\upsilon_{01} = 73.2\,\text{km}\cdot\text{s}^{-1}$ and $L_{02} = 4.2574359\,\text{Gpc}$ (event horizon), $H_{02} = 70.415674\,\text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$, $\upsilon_{02} = 70.415674\,\text{km}\cdot\text{s}^{-1}$ were obtained on the basis of the analysis of supernova type 1a [3] and measurements by Cepheids, respectively. The Hubble constant $H_0 = 67.83540245\,\text{km}\cdot\text{s}^{-1}\cdot\text{Mpc}^{-1}$, velocity $\upsilon_0 = 67.83540245\,\text{km}\cdot\text{s}^{-1}$ were introduced in [1, 2, 4] to describe the radiation of gravitational waves, relict photons from binary black holes, neutron stars based on the expression

$$\upsilon_0 = \upsilon_{01}/\Omega_{tH}; \quad \Omega_{tH} = Q_{H0} + |S'_{01}|; \quad Q_{H0} = \upsilon_{01}/\upsilon_{02} = H_{01}/H_{02} = L_{02}/L_{01}. \qquad (15)$$

Here are $Q_{H0} = 1.039541282$, $|S'_{01}| = 0.039541282$. New experimental data on the attenuation of γ-rays against an intergalactic background [18] make it possible to introduce a new Hubble constant H_0^*, velocity υ_0^*, and matter density Ω_m based on expressions

$$H_0^* = \upsilon_0^* / L_0; \quad \upsilon_0^* = \upsilon_{01} / \Omega_{tH}^*; \quad \Omega_{tH}^* - \Omega_{tH} = S_{012}; \quad S_{012} = |S'_{01}| - S'_{02}. \qquad (16)$$

Here is $S'_{02} = 0.03409$. The numerical values of $H_0^* = 67.49443576\,\text{kms}^{-1}\cdot\text{Mpc}^{-1}$, $\upsilon_0^* = 67.49443576\,\text{km}\cdot\text{s}^{-1}$, $\Omega_m = (n'_F + \Omega'_{c1})/2 = 0.141145722$ (the parameter $\Omega'_{c1} = 0.228071512$ is related to the gap in the energy spectrum of relict

photons) are close to the experimental data from [18]. From [16] it follows the connection of parameters H_{01}, υ_{01} for the accelerated expansion of the Universe with new parameters H_0^*, υ_0^*. Our parameters H_0, υ_0 and new parameters H_0^*, υ_0^* are close to the main parameters H_0', υ_0' of the model ΛCDM (plane cosmology). In our model H_0', υ_0' are defined by expressions

$$H_0' = \upsilon_0' / L_0; \quad \upsilon_0' = \upsilon_{01}/\Omega_{tH}'; \quad \Omega_{tH}' = \Omega_{tH}^* + \Omega_{0\nu} + n_g\Omega_{A0}/n_{A0}. \qquad (17)$$

Values H_0'=67.30995226 km·s^{-1}·Mpc^{-1}, υ_0'=67.30995226 km·s^{-1} are close to the parameters of the planar cosmology model.

4 Solar wind and heliopause

The Sun is the source of solar wind (flows of photons and various particles) [19]. Photons achieve the Earth after 8 min, and high-energy particles arrive with a delay of 100 min [20]. To estimate the characteristic distances and times, we use

$$L_{ES}' = L_{ES}/Q_{H0} = c_0 t_{ES} = \upsilon_{H0}\, t_{ES}'; \quad n_{H0} = Q_{H0}^2 = (1+|S_{01}'|)^2 = \upsilon_{01}^2/\upsilon_{02}^2, \quad (18)$$

where $\upsilon_{H0}^2 = c_0^2/n_{H0}$. Taking into account the numerical values of the distance from the Earth to the Sun $L_{ES} = 1\,\text{au} = 1.495995288\cdot 10^8$ km, the limiting speed of light in a vacuum $c_0 = 2.99792458\cdot 10^5$ km s^{-1}, we find estimates of the refractive index of the medium $n_{H0} = 1.080646077$, the speed of photon propagation in the medium υ_{H0}=2.883891801·10^5 km s^{-1}, the distance L_{ES}'=0.961962759 au, and the times of arrival of photons to the Earth from the Sun in vacuum $t_{ES} = 480.0293392$ s and in the medium $t_{ES}' = 499.0103147$ s.

To estimate the delay time t_{0m} of particles, arriving on the Earth from the Sun, we use the expressions

$$2t_{0m} = \tau_{0\gamma}\ln N_{0m}; \quad \tau_{0\gamma} = \tau_{0\alpha}/n_{0\alpha}; \quad \tau_{0\alpha} = \nu_{0\alpha}^{-1}; \quad n_{0\alpha} = 1.5 + |\xi_{0H}|^2;$$

$$\ln N_{0m} = 2n_{0\alpha}\ln N_{0\alpha}; \quad Q_{H2}N_{0\alpha} = 0.5 + \Omega_{c1}' + n_{F\tau}'; \quad \nu_{0\alpha} = \nu_{H0}/N_{0A}. \quad (19)$$

Expressions (19) were obtained in the framework of the Dicke theory of superradiance and describe the main parameters $\tau_{0\gamma}$, t_{0m} of the superradiance pulse in a medium from a state with the number of particles N_{0m}.

Based on the numerical values N_{0A}=3.557716045·10^5, ν_{H0}=50.182731 Hz, $|\xi_{0H}|^2$=0.181800122, Q_{H2}=1/3, $n_{F\tau}'$=0.049012111 we find estimates of the frequency $\nu_{0\alpha}$=141.0532217 μHz, relaxation time $\tau_{0\alpha}$=118.1587096 min, fractal parameter $n_{0\alpha}$=1.681800122, coherent spontaneous relaxation time $\tau_{0\gamma}$=70.25728449 min, effective numbers of active particles $N_{0\alpha}$=2.331250869

and $N_{0m}=17.23047995$, delay time $t_{0m}=100.0101199\,\text{min}$.

To estimate the characteristic parameters for the region near the boundary of the heliopause, we first find the relationships between the rest energies E_{0E} and E_{H0}, rest masses M_E and m_{H0}, the gravitational radii of Schwarzschild R_{GE} and R_{H0} for the Earth and the Higgs boson, respectively, by the formulas

$$E_{0E}/N_a E_{H0}=M_E/N_a m_{H0}=R_{GE}/N_a R_{H0}=n_{0E};\quad E_{H0}=c_0^2 m_{H0};$$

$$m_{H0}=c_0^2 R_{H0}/2GN_a;\quad R_{GE}=A_G E_{0E};\quad A_G=R_{H0}/E_{H0}=2GN_a/c_0^4;$$

$$R_{GE}=N_{GE}L_{ES}=n_{0E}N_a R_{H0};\quad M_E=5.977\cdot 10^{27}\,\text{g}. \qquad (20)$$

Based on (20) we find the parameters of the theory $A_G=0.960836162\,\text{fm}\,(\text{eV})^{-1}$, $n_{0E}=73.87419814$, $R_{GE}=5.347530124\cdot 10^{18}\,\text{km}$, $N_{GE}=3.574563481\cdot 10^{10}$.

Taking into account (18) in the framework of the anisotropic model [4] we find the characteristic velocities υ_{hS}, υ'_{hS}, distances L_{hS}, L'_{hS}, time of arrival of the signal from the heliopause to the Earth t_{hS} from the expressions

$$\upsilon'_{hS}=Q_{H0}\upsilon_{hS}=|\chi_{ef}|\upsilon_{01};\quad L_{hS}=N_{hS}L_{ES};\quad L'_{hS}=N'_{hS}L_{ES};\quad L^*_{hS}=L_{hS}/Q_{H0};$$

$$N'_{hS}=n_{H0}N_{hS};\quad L_{hS}/R_{GE}=\upsilon_{hS}^2/c_0^2;\quad L_{hS}/L_{ES}=t_{hS}/t_{ES}. \qquad (21)$$

Based on (18) - (21), the values $|\chi_{ef}|=0.250425279$ from [4], we find the estimates $\upsilon_{hS}=|\chi_{ef}|\upsilon_{02}=17.63386481\,\text{km s}^{-1}$, $N_{hS}=123.6734916$, $N'_{hS}=133.6472735$, $L_{hS}=1.850149607\cdot 10^{10}\,\text{km}$, $t_{hS}=t_{ES}N_{hS}=16.49080679\,\text{hour}$. The speed υ_{hS} is close to the speed $\upsilon_{\text{V2}}=17.5\,\text{km s}^{-1}$ of the V2 probe; the distance $L^*_{hS}=118.9692932\,\text{au}$ is near the distance to the heliopause boundary $L_{\text{V2}}=119\,\text{au}$ from [26].

To describe the transition region near the boundary of the heliopause, we introduce the times t_1, t_2, t_3, distances L_1, L_2, L_3. Next, we find the characteristic time intervals t_{31}, t_{21}, t_{32} by the formulas

$$t_{31}=t_3-t_1=1/\nu_{31};\quad \nu_{31}=(1-\psi_{02})\nu_{H0}S_{2u}/N_{0A};\quad t_{21}=t_2-t_1=t_{31}P_\tau;$$

$$t_{32}=t_3-t_2=t_{31}P'_\tau;\quad P_\tau+P'_\tau=1;\quad P'_\tau=1/(2+S'_{03}). \qquad (22)$$

Using the parameters $\psi_{02}=0.984494334$, $S'_{03}=0.460458718$ from [4], we obtain numerical values: frequency $\nu_{31}=0.072287263\,\mu\text{Hz}$; probabilities $P_\tau=0.593571722$, $P'_\tau=0.406428278$; time intervals $t_{31}=160.1122188\,\text{day}$, $t_{21}=95.03808539\,\text{day}$, $t_{32}=65.07413336\,\text{day}$. The obtained values of the intervals t_{21} and t_{32} practically coincide with the time intervals of 95 days and 65 days for the transition region near the heliopause boundary from [26, Fig. 1a].

The characteristic distance L_3 for interstellar space (outside the heliopause at

$L_3 > L_2$) is determined from the expressions

$$L_3 = N_{L3} L_{ES}\,; \quad N_{L3} = (1 - \Omega_{hL} - S_{2u}) N_{hS}\,. \tag{23}$$

Using the parameters $\Omega_{hL} = 0.000118617$ from [4, 5], N_{hS} from (21), we find the value $N_{L3} = 119.5712542$ and the estimate of the distance $L_3 = 119.5712542\,\text{au}$. To estimate the distance L_1 (inside the heliosphere for $L_1 < L_2$), we use the characteristic distances $L_{\mu e}$, $L_{\mu\mu}$, $L_{\mu\tau}$ for e, μ, τ-leptons, respectively, determined by the expressions

$$\begin{aligned}
&L_{\mu e} = N_{\mu e} L_{ES}\,; \quad N_{\mu e} = n_{\mu e} N_{hS}\,; \quad n_{\mu e} = (2 + \Omega_{\mu e}) - (1 + S_{1u})\,; \\
&L_{\mu\mu} = N_{\mu\mu} L_{ES}\,; \quad N_{\mu\mu} = n_{\mu\mu} N_{hS}\,; \quad n_{\mu\mu} = (2 + \Omega_{\mu\mu}) - (1 + S_{1u})\,; \\
&L_{\mu\tau} = N_{\mu\tau} L_{ES}\,; \quad N_{\mu\tau} = n_{\mu\tau} N_{hS}\,; \quad n_{\mu\tau} = (2 + \Omega_{\mu\tau}) - (1 + S_{1u})\,.
\end{aligned} \tag{24}$$

Using the parameters $\Omega_{\mu e}$, $\Omega_{\mu\mu}$, $\Omega_{\mu\tau}$ from (9) - (11), based on (24) we find the estimates of distances $L_{\mu e} = 118.1796344\,\text{au}$, $L_{\mu\mu} = 118.1811855\,\text{au}$, $L_{\mu\tau} = 118.1840014\,\text{au}$. For search of the characteristic distance L_2 (as the heliopause boundary), we consider a random variable $\hat{L}_2$ with two possible values L_3 from (23), $L_1 = L_{\mu e}$ from (24) and their corresponding probabilities $P_{\psi 01}$, $P'_{\psi 01}$. For expected value $M(\hat{L}_2)$, variance $D(\hat{L}_2)$, deviation $\sigma(\hat{L}_2)$, we have

$$\begin{aligned}
&M(\hat{L}_2) = P_{\psi 01} L_3 + P'_{\psi 01} L_{\mu e} = L_2\,; \quad D(\hat{L}_2) = (L_3 - L_{\mu e})^2 P_{\psi 01} P'_{\psi 01}\,; \quad \sigma(\hat{L}_2) = D^{1/2}(\hat{L}_2)\,; \\
&P_{\psi 01} + P'_{\psi 01} = 1\,; \quad P'_{\psi 01} = \psi_{01} / (1 + S'_{03} + \psi_{01})\,; \quad \psi_{01} = 1.015268884\,.
\end{aligned} \tag{25}$$

The numerical values of the distance $L_2 = 119.0005661\,\text{au}$ and space intervals $L_{32} = L_3 - L_2 = 0.57068813\,\text{au}$, $L_{21} = L_2 - L_{\mu e} = 0.8209317\,\text{au}$ practically coincide with the characteristic values of 119 au, 0.57 au, 0.82 au, respectively, from [26, Fig. 1a]. Based on (22), (25), we find the average values of the velocities υ_{21} (inside the heliosphere), υ_{32} (outside the heliopause), the jump in velocities $\delta\upsilon_{21}$ (at the heliopause) and the ratio of velocities $\upsilon_{32} / \upsilon_{21}$

$$\begin{aligned}
&\upsilon_{21} = L_{21} / t_{21} = L_{31} P_{\psi 01} / t_{31} P_\tau\,; \quad \upsilon_{32} = L_{32} / t_{32} = L_{31} P'_{\psi 01} / t_{31} P'_\tau\,; \quad L_{31} = L_3 - L_1\,; \\
&\delta\upsilon_{21} = \upsilon_{32} - \upsilon_{21}\,; \quad \upsilon_{32} / \upsilon_{21} = \psi_{01} = \varepsilon_{01} / E_{H0} = \nu_{01} / \nu_{H0}\,.
\end{aligned} \tag{26}$$

The numerical values are equal: $\upsilon_{21} = 14.95635805\,\text{km s}^{-1}$, $\upsilon_{32} = 15.18472495\,\text{km s}^{-1}$, $\delta\upsilon_{21} = 228.366896\,\text{m s}^{-1}$. We note, that the probabilities $P_{\psi 01}$ and P_τ are coupled through a conditional probability $P_{\psi\tau}$, and the ratio of the velocities and the jump in velocities allow us to introduce probabilities P_ψ, P'_ψ using expressions of the type

$$P_{\psi 01} = P_\tau P_{\psi\tau}\,; \quad P_{\psi\tau} = (2 + S'_{03}) / (1 + S'_{03} + \psi_{01}) = 1/(1 + n_{01})\,; \quad P_\psi + P'_\psi = 1\,;$$

$$P_\psi = 1/\psi_{01} = \upsilon_{21}/\upsilon_{32}\,; \quad P'_\psi = \delta\upsilon_{21}/\upsilon_{32}\,; \quad n_{01} = (\psi_{01} - 1)/(2 + S'_{03})\,. \qquad (27)$$

From (27) it follows, that n_{01} is a function of two arguments ψ_{01} and S'_{03}. If the Higgs field is absent ($\psi_{01} = 1$), then from (27) we obtain: $n_{01} = 0$; probabilities $P_{\psi\tau} = 1$, $P_{\psi 01} = P_\tau$, $P_\psi = 1$, $P'_\psi = 0$; jump in speed $\delta\upsilon_{21} = 0$ and equality of speeds $\upsilon_{21} = \upsilon_{32}$. The presence of the Higgs field ($\psi_{01} \neq 1$) leads to the appearance of a velocity jump, when crossing the heliopause boundary. Replacing the parameter S'_{03} in (27) with other parameters S'_{0x}, S_{xu} ($x = 1, 2, 3, 4$) of the energy (frequency) spectra leads to a change in the probabilities and stochastic behavior of the velocities υ_{21}, υ_{32}.

The anisotropic model [4] and expressions (1), (4) allow us to obtain relationships of velocities υ_{32}, υ_{21} with characteristic velocities $\upsilon_{\psi u}$, υ_{eu} (active nanoobjects, femtoobjects that are part of the solar and galactic wind) of the type

$$\upsilon_{32} = n'_F\,\upsilon_{\psi u} = \chi_0\,\upsilon_{eu} = \psi_{01}\upsilon_{21}\,; \quad \upsilon_{\psi u} = \xi_{2p}\,\upsilon_{eu}\,; \quad \xi_{2p} = r_e^*/r_p\,. \qquad (28)$$

Based on (28) we find the velocity estimates $\upsilon_{eu} = 59.04358906\,\mathrm{km\,s^{-1}}$, $\upsilon_{\psi u} = 279.9773874\,\mathrm{km\,s^{-1}}$. On the other hand, the characteristic solar wind velocity $\upsilon_{\psi u}$ is related to the Hubble constants H_{01} and H_{02}, H_0, H_0^*, H'_0, velocities υ_{01} and υ_{02}, υ_0, υ_0^*, υ'_0 for models from (14), (15), (16), (17), respectively, by expressions of the type

$$0.5\upsilon_{\psi u} = 2\upsilon_{02} - \upsilon_{0A} = \upsilon_W - \upsilon_q - \upsilon_{0A}\,; \quad \upsilon_q = \upsilon_{01} - \upsilon_{02} = \upsilon_W - 2\upsilon_{02}\,;$$

$$\upsilon_W = \upsilon_{01} + \upsilon_{02} = \upsilon_0\,\Omega_{tH} + \upsilon_{02} = \upsilon_0^*\,\Omega_{tH}^* + \upsilon_{02} = \upsilon'_0\,\Omega'_{tH} + \upsilon_{02}\ ; \quad \upsilon_{0A} = c_0/N_{0A}\,. \qquad (29)$$

Values of speeds are equal: $\upsilon_{0A} = 0.84265426\,\mathrm{km\,s^{-1}}$, $\upsilon_W = 143.615674\,\mathrm{km\,s^{-1}}$, $\upsilon_q = 2.784326\,\mathrm{km\,s^{-1}}$.

The velocity υ_{hS} from (21) is related to the characteristic velocities of relict photons υ_{ra}, υ_{ra}^* and the velocities υ_{02}, υ_0^*, $\upsilon_{0\rho}$, υ_W, $\upsilon_{h\rho}$ by expressions of the type

$$2\upsilon_{hS}\,\upsilon_{ra} = \upsilon_{ra}^*\,\upsilon_{02}\,; \quad \upsilon_{ra} = c_0/N_{ra}\,; \quad \upsilon_{ra}^* = 2\,|\,\chi_{ef}\,|\,\upsilon_{ra}\,;$$

$$\upsilon_{ra}^*\,\upsilon_{0\rho} = \upsilon_{ra}\,\upsilon_0^*\ ; \quad \upsilon_W^2 = \upsilon_{0\rho}^2 + \upsilon_{h\rho}^2\,; \quad N_{ra} = 1041.293475\,. \qquad (30)$$

Values of speeds are equal: $\upsilon_{ra} = 287.9039053\,\mathrm{km\,s^{-1}}$, $\upsilon_{ra}^* = 144.1968316\,\mathrm{km\,s^{-1}}$, $\upsilon_{0\rho} = 134.7596298\,\mathrm{km\,s^{-1}}$, $\upsilon_{h\rho} = 49.65182785\,\mathrm{km\,s^{-1}}$.

The experimental data obtained by the Wind probe (the interval of solar wind speed changes of 600–300 km s^{-1}, Fig. 6 from [25]), on the UTR-2, URAN-2 radio

telescopes (Fig. 5 from [25]) showed, that the solar wind in orbit and beyond the Earth's orbit consists of a set of particle flows with different velocities and densities. The structure of these flows depends on time and solar activity [19, 20]. An analysis [25] of intermode (intramode) interactions of particles of different flows was performed by the interplanetary scintillation method based on the behavior of space and time correlation functions for radiation intensity. The velocities $2\upsilon_{0\rho}$, $\upsilon_{\psi u}$ and υ_{ra} are close to the characteristic velocities of 270, 280 and 290 km s^{-1} of separate solar wind modes from [25]. The detailed analysis of the multimode structure of the solar wind in our model is possible based on spectra of type $\upsilon_{\psi ux}=2\upsilon_{\psi u}S_{xu}$ and $\upsilon_{rax}=2\upsilon_{ra}S_{xu}$. From (30) it follows that the velocities $\upsilon_{0\rho}$ and $\upsilon_{h\rho}$ can be interpreted as both the radial and transverse components of the total velocity υ_W. The presence of transverse components $\pm\upsilon_{h\rho}$ of the solar wind near the Sun is confirmed by experimental data collected by the Parker Solar Probe [21 - 24]. The behavior of the transverse component (Fig. 2 from [22]) is stochastic and varies in the range from 50 to –50 km s^{-1}. In [24], such a behavior of the slow solar wind is associated with the presence of equatorial coronal holes in the Sun. A fast solar wind with speeds $2\upsilon_{0\rho}$ occurs near the poles of the Sun.

In our model, it is also possible to describe the multimode structure of the solar and galactic winds at the crossing of the heliopause based on the velocities υ_{eu} from (28), υ_W from (29), υ_{ra}^* from (30) and the corresponding velocity spectra. The experimental data (Fig. 4d from [27], Fig. 2 from [29]) confirm the stochastic behavior and change in the velocity of solar wind particles when the heliopause crosses from 150 km s^{-1} to 100 km s^{-1}. The complex dynamic behavior of the plasma components (Fig. 3, 4 from [29]) with velocities near υ_{eu}, $2\upsilon_{eu}$ inside the heliosphere indicates the presence of a boundary layer near the heliopause.

To estimate the characteristic energies ε_{0A}, E_{0A}, $\varepsilon_{\lambda A}$, effective wavelength λ_A, effective number N_{0n} of particles, we use expressions of the type

$$E_{H0}/\varepsilon_{0A}=E_{0A}/E_G=N_{0A};\quad E_{H0}/E_{0A}=\varepsilon_{0A}/E_G=N_{0n};$$

$$E_{H0}/E_G=N_{HG}=N_{0n}N_{0A};\quad \varepsilon_{\lambda A}^2=\varepsilon_{0A}E_{0A}=E_{H0}E_G;\quad \lambda_A=a_\lambda/\varepsilon_{\lambda A}. \quad (31)$$

Taking into account $N_{0A}=3.557716045\cdot10^5$, $N_{HG}=1.031830522\cdot10^{16}$, a_λ from [6] we find the estimates: $\varepsilon_{0A}=351.4400206$ keV, $E_{0A}=4.311073329$ eV, $\varepsilon_{\lambda A}=1.230887363$ keV, $\lambda_A=1.007114093$ nm, $N_{0n}=2.900261036\cdot10^{10}$.

The presence of a multimode structure of the solar and galactic wind, the Higgs field leads to the replacement $\varepsilon_{\lambda A}$, λ_A by $\varepsilon_{\lambda A}^*$, λ_A^* by the formulas

$$\varepsilon_{\lambda A}^*=\psi_{rc}\varepsilon_{bb};\quad \lambda_A^*=a_\lambda/\varepsilon_{\lambda A}^*=2R_{\lambda A};\quad E_{\lambda A}=R_{\lambda A}/A_G;$$

$$\varepsilon_{bb}=\varepsilon_{0A}(|S_{1u}|+S_{2u});\quad \psi_{rc}=2\Delta_{rc}/E_{0A}=(\varepsilon_{01}-\varepsilon_{02})S_{1u}/\varepsilon_{02}S_{2u}. \quad (32)$$

The values are equal: ε_{bb}= 28.042404 keV , ψ_{rc}= 0.04420725 , Δ_{rc}= 95.290347 meV , $\varepsilon^{*}_{\lambda A}$= 1.239677565 keV , λ^{*}_{A}= 0.999972933 nm , $E_{\lambda A}$= 0.520365996 MeV . The energy $E_{\lambda A}$ (for solar wind particles inside the heliosphere) is associated with the energy $E_{\lambda L}$ (for galactic wind particles behind the heliopause)

$$E_{\lambda A} = (\Omega_{\tau L} + n_g \Omega_{0G}) E_{\lambda L}\,; \quad \Omega_{0G} N_{0A} = 1.5 + \Omega'_{c1} + n'_{F\tau}\,;$$

$$E^2_{rc} = E^2_{0A} - 4\Delta^2_{rc}\,; \quad (E'_{rc})^2 = E^2_{0A} + 4\Delta^2_{rc}\,. \qquad (33)$$

The numerical values are equal: Ω_{0G}= $4.99501253\cdot 10^{-6}$, $E_{\lambda L}$= 213.0772532 MeV , E_{rc}= 4.306858745 eV , E'_{rc}= 4.315283797 eV . The energy estimates ε_{bb} , $E_{\lambda L}$ obtained in our model are consistent with the energies of 28 keV, 213 MeV from [26], and the energy $E_{\lambda A}$ is consistent with the energy of 0.5 MeV from [28].

The magnetic characteristics of solar and galactic wind particles have features of the behavior at the intersection of the heliopause: a jump in the magnetic field from 0.42 to 0.68 nT is observed (Fig. 1a from [27]); components of the magnetic field can have different signs (Fig. 3 from [27]); the presence of a magnetic barrier (Fig. 4a from [27]); a change in the direction of the magnetic field components (Fig. 6b, c from [27]). In our model, to estimate the components of magnetic fields $B_{y\beta x}$, $B^{*}_{y\beta x}$ we use frequency spectra of the type

$$\nu_{y\beta x} = \gamma_n B_{y\beta x} / 2\pi = 2\nu_{y\beta} S'_{0x}\,; \quad \nu^{*}_{y\beta x} = \gamma_n B^{*}_{y\beta x} / 2\pi = 2\nu_{y\beta} S_{ux}\,; \quad y = 0,1,2\,;$$

$$\nu_{y\beta} = \nu_{0y}/N_{ra}\,;\ B'_{2\beta 1} = B^{*}_{2\beta 1} / (1.5 + n'_{zg} + S_{012})\,;\ \nu_{00} = \nu_{H0}\,;\ \nu_{02} = \psi_{02}\nu_{H0}\,. \quad (34)$$

Here we use the well-known nuclear gyromagnetic ratio $\gamma_n / 2\pi$ = 0.6535 MHz/kO for the deuteron (^{2}H) [12], n'_{zg}= 0.114317037 [4]. Based on (34) we find estimates: frequencies $\nu^{*}_{2\beta 1}$= 4.4353480 mHz ; jump of magnetic fields from $B'_{2\beta 1}$= 0.4190147 nT to $B^{*}_{2\beta 1}$ = 0.6787067 nT at the intersection of heliopause. The numerical values of the fields deviations of the type $\delta B = B_{0\beta 1} - B_{0\beta 2}$ = 0.0804015 nT , $\delta B^{*} = B^{*}_{0\beta 1} - B^{*}_{0\beta 2}$ = 0.2019195 nT and the sum of the deviations $\delta B + \delta B^{*}$ = 0.282321 nT are characteristic of the stochastic behavior of the magnetic field on time inside the heliosphere (consistent with data Fig. 6 from [27]).

Conclusions

In fractal quantum systems the model femtoobjects, as active objects with sizes of the order of the classical electron radius, are considered. The main parameters of the model femtoobject, which are coupled with the known parameters from quantum electrodynamics and the Higgs boson, are introduced. To take into

account the stochastic behavior of the parameters, random variables with two possible values and the corresponding probabilities are introduced. It was shown, that the obtained estimates of the proton radius, measurement errors using the example of the hydrogen atom, and estimates of the anomalies in the magnetic moments of leptons are consistent with the experimental data.
The parameters of active nanoobjects and femtoobjects are coupled with cosmological parameters, with new values of the Hubble constants. These active objects can determine the compound, structure and behavior of the solar wind (flows of various particles) near the Sun, Earth and in interstellar space (near the heliopause). The relationships of such active objects with the parameters of the Higgs boson and the Higgs field are determined. Estimates of the main parameters are conformed with the experimental data, obtained by the Planck space observatory (based on Fermi-LAT and Cerenkov telescopes), UTR-2 and URAN-2 radio telescopes, Parker Solar Probe, Voyager 2 and Voyager 1.
The results can be used to find a solution to the problem associated with the Covid-2019 virus (based on active femtoobjects and nanoobjects), in cosmic medicine.

References

1. V.S. Abramov. Gravitational Waves, Relic Photons and Higgs Boson in a Fractal Models of the Universe / C.H. Skiadas and I. Lubashevsky (eds.), 11th Chaotic Modeling and Simulation International Conference, Springer Proceedings in Complexity. Springer Nature Switzerland AG 2019. P. 1-14.
2. V.S. Abramov. Superradiance of Gravitational Waves and Relic Photons from Binary Black Holes and Neutron Stars. Bulletin of the Russian Academy of Sciences: Physics, 83, 3, 364-369, 2019.
3. A.G. Riess, A.V. Filippenko, P. Challis et al. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. Astronomical Journal, 116, 3, 1009-1038, 1998.
4. V.S. Abramov. Active Nanoobjects, Neutrinos and Higgs Fields in Anisotropic Models of Fractal Cosmology. Bulletin of the Russian Academy of Sciences: Physics, 83, 12, 1516-1520, 2019.
5. V.S. Abramov. Supernonradiative states, neutrino and Higgs Boson in fractal quantum systems. Bulletin of the Russian Academy of Sciences: Physics, 84, 3, 284-288, 2020.
6. V.S. Abramov. Active Nanoobjects, Neutrino and Higgs Boson in a Fractal Models of the Universe / C.H. Skiadas and Y. Dimotikalis (eds.), 12th Chaotic Modeling and Simulation International Conference, Springer Proceedings in Complexity. Springer Nature Switzerland AG 2020. P. 1-14.
7. V. Abramov. Super-nonradiative states in fractal quantum systems // XIII International Workshop on Quantum Optics (IWQO-2019). EPJ Web of Conferences, 220, 2 p., 2019.
8. O.P. Abramova, A.V. Abramov. Effect of Ordering of Displacement Fields Operators of Separate Quantum Dots, Elliptical Cylinders on the Deformation Field of Coupled Fractal Structures. / C.H. Skiadas and I. Lubashevsky (eds.), 11th Chaotic Modeling and Simulation International Conference, Springer Proceedings in Complexity. Springer Nature Switzerland AG 2019. P. 15-27,
9. R. Pohl, A. Antognini, F. Nez et al. The size of the proton. Nature, 466, 213-217, 2010.
10. A. Beyer, L. Maisenbacher, A. Matveev et al. The Rydberg constant and proton size

from atomic hydrogen. Science, 358, 79-85, 2017.
11. N. Kolachevsky. 2S–4P spectroscopy in hydrogen atom: the new value for the Rydberg constant and the proton charge radius. IWQO – 2019: Collection of abstracts. Vladimir, September 9 - 14, 2019. Moscow, Trovant, 2019. P. 32.
12. S.V. Vonsovsky. Magnetism of microparticles. Moscow, Nauka, 1973.
13. N. Agafonova, A. Alexandrov, A. Anokhina et al. Final results of the OPERA experiment on ντ appearance in the CNGS neutrino beam. Phys. Rev. Lett. 120, 211801, 1-7, 2018.
14. R.M. Barnett, C.D. Carone, D.E. Groom et al. Review of Particle Physics. Phys. Rev. D54, 1, 1-22, 1996.
15. V.S. Abramov. The Higgs Boson in Fractal Quantum Systems with Active Nanoelements. Bulletin of the Russian Academy of Sciences: Physics, 80, 7,859-865, 2016.
16. V.S. Abramov. Pairs of Vortex-Antivortex and Higgs Boson in a Fractal Quantum System. CMSIM Journal, 1, 69-83, 2017.
17. V.S. Abramov. Cosmological Parameters and Higgs Boson in a Fractal Quantum System. CMSIM Journal, 4, 441-455, 2017.
18. A. Domínguez, R. Wojtak, J. Finke et al. A New Measurement of the Hubble Constant and Matter Content of the Universe Using Extragalactic Background Light γ-Ray Attenuation. The Astrophysical Journal. 885, 2, 137, 2019.
19. N. Schukina. The Sun is the source of life and the cause of catastrophes. Part I. A star named the Sun. Universe, Space, Time, 8 (145), 4-11, 2016.
20. N. Schukina. The Sun is the source of life and the cause of catastrophes. Part II. Light giving life. Universe, Space, Time, 9 (146), 16-23, 2016.
21. D.J. McComas, E.R. Christian, C.M.S. Cohen et al. Probing the energetic particle environment near the Sun, Nature, 576, 223-227, 2019.
22. J.C. Kasper, S.D. Bale, J.W. Belcher et al. Alfvénic velocity spikes and rotational flows in the near-Sun solar wind. Nature, 576, 228-231, 2019.
23. R.A. Howard, A. Vourlidas, V. Bothmer et al. Near-Sun observations of an F-corona decrease and K-corona fine structure, Nature, 576, 232-236, 2019.
24. S.D. Bale, S.T. Badman, J.W. Bonnell et al. Highly structured slow solar wind emerging from an equatorial coronal hole. Nature, 576, 237-242, 2019.
25. N.N. Kalinichenko, M.R. Olyak, A.A. Konovalenko et al. Large-scale structure of solar wind beyond Earth's orbit reconstructed by using data of two-site interplanetary scintillation observations at decameter radiowaves. Kinematics and Physics of Celestial Bodies, 35, 1(205), 27-41, 2019.
26. S.M. Krimigis, R.B. Decker, E.C. Roelof et al. Energetic charged particle measurements from Voyager 2 at the heliopause and beyond. Nature Astronomy, 3, 997-1006, 2019.
27. L.F. Burlaga, N.F. Ness, D.E. Berdichevsky et al. Magnetic field and particle measurements made by Voyager 2 at and near the heliopause. Nature Astronomy, 3, 1007-1012, 2019.
28. E.C. Stone, A.C. Cummings, B.C. Heikkila et al. Cosmic ray measurements from Voyager 2 as it crossed into interstellar space. Nature Astronomy, 3, 1013-1018, 2019.
29. J.D. Richardson, J.W. Belcher, P. Garcia-Galindo et al. Voyager 2 plasma observations of the heliopause and interstellar medium. Nature Astronomy, 3, 1019-1023, 2019.
30. D.A. Gurnett, W.S. Kurth. Plasma densities near and beyond the heliopause from the Voyager 1 and 2 plasma wave instruments. Nature Astronomy, 3, 1024-1028, 2019.

Qubits and fractal structures with elements of the cylindrical type

Olga P. Abramova, Andrii V. Abramov

Donetsk National University, Donetsk, Ukraine
E-mail: oabramova@ua.fm

Abstract: By the method of numerical simulation, the behavior of the deformation field of both separated and related model fractal structures of a cylindrical type was investigated. It is shown, that for the considered structures, the behavior of the deformation field essentially depends on the choice of stochastic processes (realized during iterations) and on the states of the qubit in the perpendicular plane to the axis of the cylinder. It is shown that the structure of the complex deformation field for a circular (elliptical) cylinder essentially depends on the initial basic, superposition states of the qubit. Due to the presence of various qubit states for coupled (using the example of circular and elliptic cylinders) fractal structures, the appearance of random matrices during iterations is characteristic. There is a need to use commutators and anti-commutators, products of separate deformation field operators. At this, the structure of the complex deformation field has own characteristic features of behavior.
Keywords: fractal structure, qubits, random matrices, complex deformation field, ordering of operators, quantum chaos.

1 Introduction

Earlier in [1–3], to describe the total deformation field of coupled fractal structures in an iterative process, the sum of the displacement field operators of separate fractal structures was used. The deformation field of the coupled structure essentially depends on the sequence of separate operators of displacement fields in the iterative process. On the examples of quantum dots [4], elliptic [1, 2] and circular [3, 5] cylinders the influence of the ordering of separate operators of displacement fields on the total deformation field of the coupled structure was shown. The presence of variable semiaxes and variable moduli leads to stochastic behavior of the complex deformation field of such structures. Based on pairs of same fractal structures with opposite orientations of the deformation fields, complex zero operators were introduced [3. 5]. It is shown that changes in the order of the sequence of separate operators in the zero operator for a coupled structure leads to the appearance of a nonzero complex deformation field. At the same time, noise tracks appear on the background of stochastic peaks. The noise track is a stochastic ring, the inside region of which is regular region.
For describe quantum chaos random matrices are used [6]. Elements of random matrices can be formed as a result of an iterative process. In this case, the need arises for the use of commutators and anti-commutators, products of separate

13th CHAOS Conference Proceedings, 9 - 12 June 2020, Florence, Italy

operators, qubit states [7, 8] of the deformation field. Quantum computers [9 - 12] encode information in qubits. The physical systems that realise qubits can be any objects having two quantum states. Different nanostructures and metamaterials [13] can be chosen as active objects. These active objects can be in superposition qubit states and exhibit stochastic properties, quantum entanglement.

The aim of this work is to describe the deformation fields of fractal coupled structures consisting of two separate structures (circular and elliptical cylinders) with different qubit states. In this case, the deformation fields of coupled structures are considered as the sum and product (scalar and matrix) of the deformation fields of separate structures.

2 Description of the deformation field of separate fractal structures in various qubit states

We consider a model fractal structure (circular or elliptical cylinder), located in a bulk discrete lattice $N_1 \times N_2 \times N_3$, whose nodes are given by integers n, m, j. By analogy with [1 – 3, 5] nonlinear equations for the dimensionless displacement function u of the lattice node are

$$u = k_u^2(1 - 2\mathrm{sn}^2(u - u_0, k_u')); \quad (1)$$

$$k_u^2 = (1-\alpha)/Q; \quad k_u' = (1-k_u^2)^{1/2}; \quad p = p_0 + p_1 n + p_2 m + p_3 j; \quad (2)$$

$$Q = p - b_1(n-n_0)^2/n_c^2 - b_2(m-m_0)^2/m_c^2 - b_3(j-j_0)^2/j_c^2. \quad (3)$$

Here u_0 is the constant (critical) displacement; a is the fractal dimension of the deformation field u along the axis Oz (a Î [0,1]); variable modules k_u, k_u' are functions of indices n, m, j nodes of the bulk discrete lattice. The choice of the positive sign of the module k_u' is associated with the choice of the second branch of the displacement function u [14]. Function Q determines the form of the fractal structure, the type of attractors and take into account the interaction of the nodes of both in the main plane of the discrete rectangular lattice N_1´ N_2 as well as interplane interactions. The parameters b_1, b_2, b_3, n_0, n_c, m_0, m_c, j_0, j_c characterize different fractal structures. The choice of function p depends on the choice of parameters p_i, $i = \overline{0,3}$. In this paper, we are limited to consideration of qubit states with $p_1 \neq 0$, $p_2 \neq 0$, $p_3 = 0$ and shift $p_0 \neq 0$.

The iterative procedure on index n for equations (1) - (3) simulates stochastic processes on a rectangular discrete lattice with dimensions $N_1 \times N_2$.

By numerical modelling, it was assumed that $N_1 = 240$, $N_2 = 240$, $\alpha = 0.5$, $u_0 = 29.537$, $p_0 = 1.0423$, $b_1 = b_2 = 1$, $n_0 = 121.1471$, $m_0 = 120.3267$,

j_0= 31.5279 , j_c= 11.8247 , b_3=0 . For a circular cylinder, the semiaxes were equal $n_c = m_c$=57.4327 , and for an elliptical cylinder they were as follows $n_c = 55.2537$, $m_c = 14.9245$.

To describe the deformation field of an separate circular (elliptical) cylinder, the basic states of the qubit and their superposition in the plane *nOm* are introduced. The various states of the qubit in the plane *nOm* are described by nonzero coefficients for linear terms in the functions p , Q .

The initial state of an separate circular (elliptical) cylinder is the state (0,0) with the coefficients $p_1 = 0$, $p_2 = 0$.

The basic states of a qubit are states (1,0), (0,1), (-1,0), (0, -1) with the coefficients $p_1 \neq 0$ or $p_2 \neq 0$. So for state (1,0) are $p_1 = 0.00423$, $p_2 = 0$; for state (0,1) are $p_1 = 0$, $p_2 = 0.00572$; for state (-1,0) are $p_1 = -0.00423$, $p_2 = 0$; for state (0,-1) are $p_1 = 0$, $p_2 = -0.00572$.

For superpositional states of qubits, we have, respectively:

state (1,1) are $p_1 = 0.00423$, $p_2 = 0.00572$;

state (1,-1) are $p_1 = 0.00423$, $p_2 = -0.00572$;

state (-1,1) are $p_1 = -0.00423$, $p_2 = 0.00572$;

state (-1,-1) are $p_1 = -0.00423$, $p_2 = -0.00572$. (4)

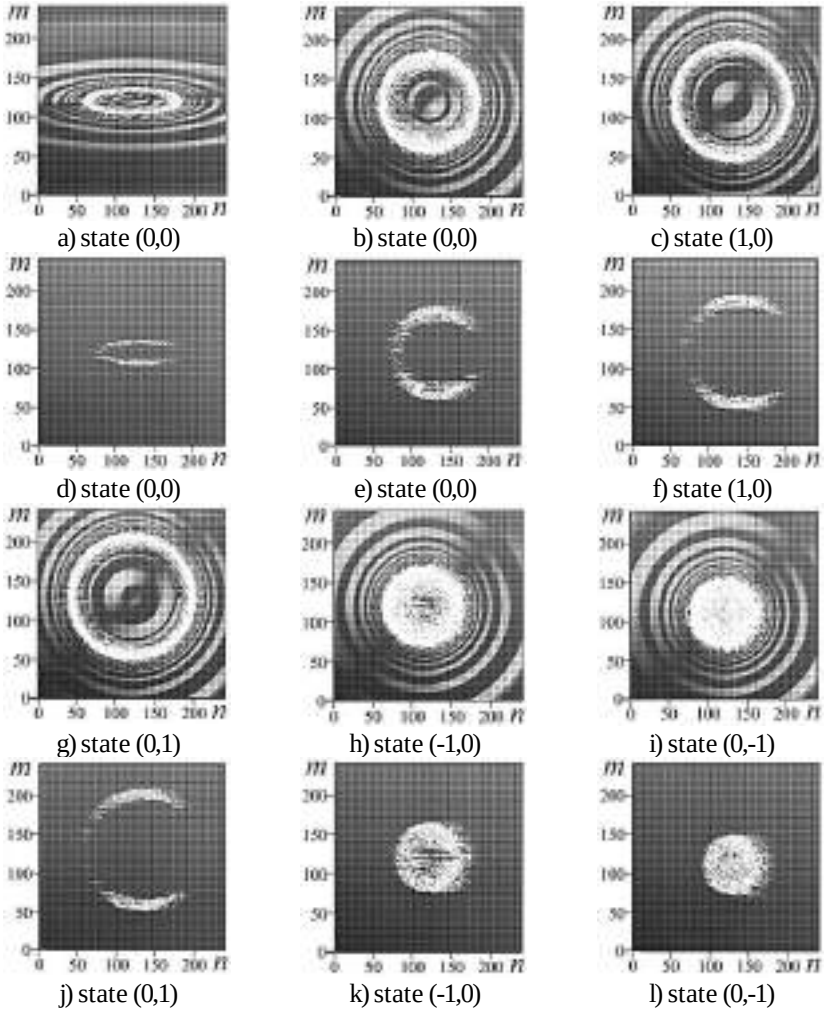
Fig. 1 shows the behavior of the cross sections of the deformation field u for elliptical (Fig. 1a, d) and circular (Fig. 1b, e) cylinders in the initial state (0,0) and basis states of qubits.

Fig. 1c, f, g, h, i, j, k, l shows the behavior of the cross sections of the deformation field u for a circular cylinder in the basic states of qubits: $\mathrm{Re}u$ (Fig. 1c, g, h, i), $\mathrm{Im}u$ (Fig. 1f, j, k, l). In this case, the peak amplitudes and the variation range $\mathrm{Re}u$, $\mathrm{Im}u$ for the elliptical cylinder are smaller than for the circular one.

For the initial (0,0) and basic states (1,0), (0,1) of qubits the regular behavior $\mathrm{Re}u$ in the inner region is characterized (Fig. 1b, c, g).

By changing qubit states, the features of behavior of the deformation field is observed: the concave part of the inner region (Fig. 1b) changes to the convex (Fig. 1c) and then to the convex-concave (Fig. 1g). Such behavior allows the interpretation of the inner region as a membrane with the possible alteration of its states due to a change in the states of qubits.

The regular behavior of the inner region $\mathrm{Re}u$ is limited by the stochastic boundary (stochastic ring). The outer region $\mathrm{Re}u$ is characterized by wave-like behavior, which is explained by the presence of variable modules k_{ui} , k'_{ui} (2) in expressions (1). Localized in the region of boundary rings with discontinuities (Fig. 1e, f, j) the stochastic behavior is characteristic for $\mathrm{Im}u$, while $\mathrm{Im}u = 0$ is characteristic in the inner and outer regions of the rings.



a) state (0,0) b) state (0,0) c) state (1,0)

d) state (0,0) e) state (0,0) f) state (1,0)

g) state (0,1) h) state (-1,0) i) state (0,-1)

j) state (0,1) k) state (-1,0) l) state (0,-1)

Fig. 1. The behavior of the cross sections u (top view) depending on the states of qubits of separate structures: $\mathrm{Re}\,u \in [-1;1]$ - (a, b, c, g h, i); $\mathrm{Im}\,u \in [-1;1]$ - (d, e, f, j, k, l). The initial states of qubits (0,0) for elliptical (a, d) and circular (b, e) cylinders. The basic states of qubits for a circular cylinder (c, f, g - l).

For the other basic states (-1.0), (0, -1) of qubits characteristic stochastic behavior $\mathrm{Re}\,u$ in the inner region and wave-like behavior in the outer region (Fig. 1h, i), that indicates a significant alteration of the structures.

For these states $\mathrm{Im}\,u$ has a stochastic structure, localized in the inner region of the cylinder (Fig. 1k, l), and outside the region $\mathrm{Im}\,u = 0$. The imaginary part

$\mathrm{Im}\,u$ indicates the presence of an effective damping. By changing these basis states, the character of damping changes.

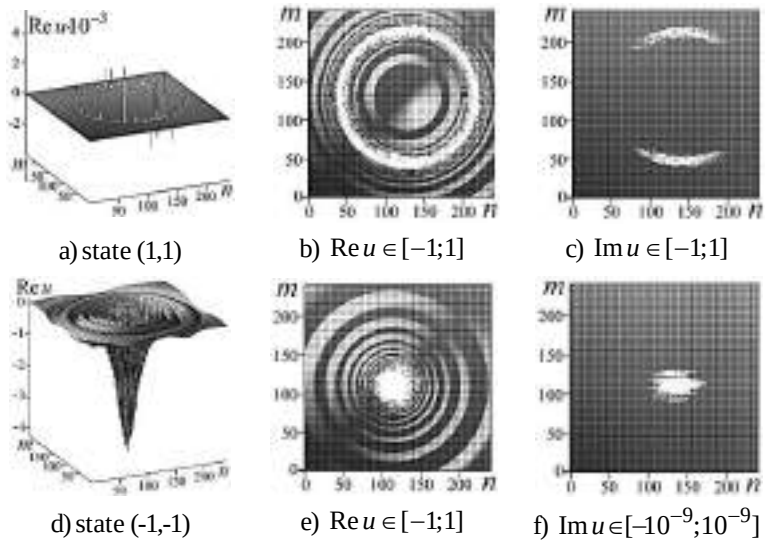


Fig. 2. Superpositional states of qubits of a separate structure (circular cylinder). Behavior $\mathrm{Re}\,u$ (a, d) and cross sections (top view) (b, c, e, f) in the states: (1,1) - (a, b, c); (-1, -1) - (d, e, f).

The presence of superpositional states of qubits in separate structures leads to a change in the behavior of the complex deformation field. As an example, Fig. 2 shows the behavior $\mathrm{Re}\,u$ (Fig. 2a, d) and cross sections (Fig. 2b, c, e, f) of an separate structure (circular cylinder) in superposition states of qubits (1,1) and (-1, -1). The characteristic features of the behavior of the deformation field for state (1.1) (Fig. 2b, c) are close to state (0,0) (Fig. 1b, e). The characteristic features the cross sections behavior of deformation field for the state (-1, -1) (Fig. 2e, f) are close to the states (-1,0) (Fig. 1h, k), (0, -1) (Fig. 1i, l).
However, the behavior $\mathrm{Re}\,u$ for the superposition state (-1, -1) (Fig. 2d) differs significantly from the characteristic behavior $\mathrm{Re}\,u$ of all other superposition states of qubits (1,1) (Fig. 2a), (1, -1), (-1,1). Instead of a structure such as a circular stochastic dislocation (Fig. 2a), a structure like a stochastic funnel (Fig. 2d) arises. In this case, the amplitudes $\mathrm{Re}\,u$ and $\mathrm{Im}\,u$ for the state (-1, -1) are significantly smaller than the amplitudes for other states of qubits.

3 Fractal coupled structures with initial states of qubits of separate structures

Consider the model fractal coupled structures (I,II), (II,I), consisting of two

separate structures (I) and (II) with the same initial qubit states (0,0). By analogy with (1) – (3) nonlinear equations for the dimensionless complex displacement function u of the lattice node are

$$u = \sum_{i=1}^{2} u_{Ri}\ ; \quad u_{Ri} = R_i k_{ui}^2 (1 - 2\mathrm{sn}^2(u - u_{0i}, k'_{ui}))\ ; \tag{5}$$

$$k_{ui}^2 = (1 - \alpha_i)/Q_i\ ; \quad k'_{ui} = (1 - k_{ui}^2)^{1/2}\ ; \quad p_i = p_{0i} + p_{1i}n + p_{2i}m + p_{3i}j\ ; \tag{6}$$

$$Q_i = p_i - b_{1i}(n - n_{0i})^2 / n_{ci}^2 - b_{2i}(m - m_{0i})^2 / m_{ci}^2 - b_{3i}(j - j_{0i})^2 / j_{ci}^2\ . \tag{7}$$

Here, all parameters have the same meaning as for expressions (1) – (3). Parameters R_i (i=1,2) determine the orientation of the deformation fields of separate structures in a coupled system. For separate structures (I) and (II), the deformation fields $u = u_{R1}$ and $u = u_{R2}$ correspond to the matrices $\mathbf{M}_{R1}$ and $\mathbf{M}_{R2}$, whose elements are found independently from each other by the iteration method. In this case, the iterative procedure on index n for equations (5) - (7) simulates two independent stochastic processes on a rectangular discrete lattice with dimensions $N_1 \times N_2$. Earlier in [5], ordered operators of displacement fields of a coupled structure were introduced as the sum of the operators of separate structures. Here, for the sum of the matrices $\mathbf{M}_{R1}$, $\mathbf{M}_{R2}$ the relation is fulfilled

$$\mathbf{M}_{R1} + \mathbf{M}_{R2} = \mathbf{M}_{R2} + \mathbf{M}_{R1}\ . \tag{8}$$

The deformation fields for the coupled structures (I,II), (II,I) correspond to the ordered operators

$$u = u_{(\mathrm{I,II})} = u_{R1} + u_{R2}\ , \quad u = u_{(\mathrm{II,I})} = u_{R2} + u_{R1} \tag{9}$$

and matrices $\mathbf{M}_{(\mathrm{I,II})}$, $\mathbf{M}_{(\mathrm{II,I})}$, whose elements are found by the iteration method. The iterative procedure on index n for equations (5) – (7) simulates two other independent stochastic processes for matrices $\mathbf{M}_{(\mathrm{I,II})}$, $\mathbf{M}_{(\mathrm{II,I})}$. In this case, the relations are fulfilled

$$\mathbf{M}_{(\mathrm{I,II})} = \mathbf{M}_{R1} + \mathbf{M}_{R2} = \mathbf{M}_{R2} + \mathbf{M}_{R1} = \mathbf{M}_{(\mathrm{II,I})}\ ; \quad \mathbf{M}_{(\mathrm{I,II})} - \mathbf{M}_{(\mathrm{II,I})} = 0\ . \tag{10}$$

To describe the deviation of the deformation field of the coupled structures (I,II) and (II,I), we introduce the ordered operator

$$u_\delta = (u_{R1} + f_2(u_{R2})) - (u_{R2} + f_1(u_{R1}))\ , \tag{11}$$

which corresponds to the matrix $\mathbf{M}_\delta$. An iterative procedure on index n simulate stochastic process for a matrix $\mathbf{M}_\delta$, which does not coincide with stochastic processes for matrices $\mathbf{M}_{(\mathrm{I,II})}$, $\mathbf{M}_{(\mathrm{II,I})}$, $\mathbf{M}_{R1}$, $\mathbf{M}_{R2}$. In this case

$$\mathbf{M}_{\delta} \neq \mathbf{M}_{(\mathrm{I,II})} - \mathbf{M}_{(\mathrm{II,I})}; \quad \mathbf{M}_{\delta} \neq 0. \tag{12}$$

From (12) follows, that stochastic processes for matrices $\mathbf{M}_{(\mathrm{I,II})}$, $\mathbf{M}_{(\mathrm{II,I})}$ become dependent. If in (11) assume $f_2(u_{R2}) = u_{R2}$, $f_1(u_{R1}) = u_{R1}$, then $\mathbf{M}_{(\mathrm{I,II})} = \mathbf{M}_{(\mathrm{II,I})}$, what confirms the independence conditions for stochastic processes (10). Attractors of the deformation field of the coupled fractal structure are located on the surface, the core of which is determined from the condition

$$Q_1 \cdot Q_2 = 0. \tag{13}$$

By numerical modeling, it was assumed, that: α_i=0.5, u_{0i}=29.537, p_{0i}=1.0423, b_{1i}=b_{2i}=1, n_{0i} = 121.1471, m_{0i} = 120.3267, n_{c1}= m_{c1}=57.4327, j_{0i}= 31.5279, j_{ci}= 11.8247, p_{1i}=p_{2i}=p_{3i}=0, b_{3i}=0. In this case, in fractal coupled structures (I,II) and (II,I), the structure (I) is a circular cylinder and the structure (II) is an elliptical cylinder with variable semi-axes n_{c2}, m_{c2}. The variable semiaxes were chosen so that the cross-sectional area of the ellipse $S = \pi n_{c2} m_{c2}$ did not change and was equal to the cross-sectional area of the circular cylinder $S = 824.6316\pi$ from [2, 3]. For an elliptical cylinder (II), the semiaxes n_{c2}, m_{c2} were defined as follows:

variant 1 are $n_{c2} = 43.0746$, $m_{c2} = 19.1443$ (the elliptical cylinder is inside the circular cylinder);

variant 2 are $n_{c2} = 55.2537$, $m_{c2} = 14.9245$ (the elliptical cylinder approaches to the circular cylinder along the axis On);

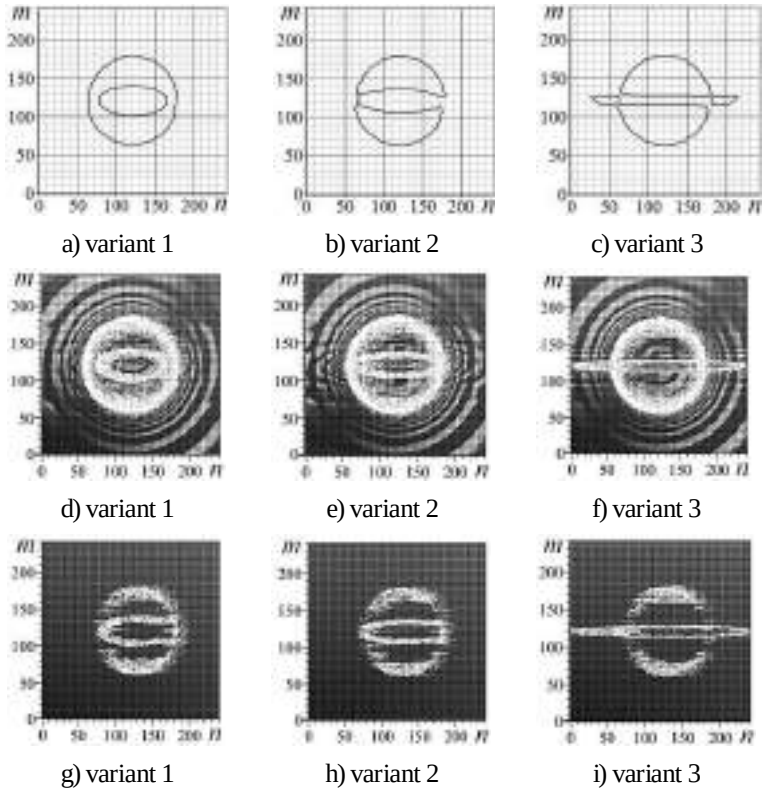
variant 3 are $n_{c2} = 119.9327$, $m_{c2} = 6.8758$ (the elliptical cylinder extends beyond the boundaries of the circular cylinder along the axis On).

Fig. 3 shows the behavior of attractors for all three variants of fractal structures (I,II) and the complex deformation field $u_{(\mathrm{I,II})}$. The different behavior of attractors (Fig. 3a, b, c) and cross sections of the complex deformation field (Fig. 3d - i) confirm the stochastic nature of the deformation field of the structure (I,II) and its dependence on the semiaxes of the elliptic cylinder (II).

The behavior of the deformation field of the structure (II,I) in this paper is not given. However, completed researches performed make it possible to estimate the deviations (12) $\mathrm{Re}\,\mathbf{M}_{\delta} \approx 10^{-9}$, $\mathrm{Im}\,\mathbf{M}_{\delta} \approx 10^{-25}$, which indicates to the dependence of stochastic processes in (11).

As a result of the iterative process, elements of random matrices are formed, which depend on various qubit states of separate structures in a coupled structure. Random matrices are used to describe quantum chaos [6]. In this case, there is a need to use commutators and anti-commutators, products of separate operators of the deformation field. Next, we consider fractal coupled structures

(III) and (IV), the deformation fields of which u_3 and u_4 are described by the product of the deformation fields of separate structures (I) and (II) with the same initial qubit states (0,0). The deformation fields of structures (III) and (IV) correspond to the matrices $\mathbf{M}_3 = \mathbf{M}_{R1} \cdot \mathbf{M}_{R2}$ and $\mathbf{M}_4 = \mathbf{M}_{R2} \cdot \mathbf{M}_{R1}$. Here, the dot symbol describes the operation of ordinary matrix multiplication. Fig. 4 shows the behavior of the complex deformation field for structures (III) and (IV). In this case, structure (II) parameters were chosen corresponding to variant 2. The attractors of structures (III) and (IV) coincide with the attractor from Fig. 3b. Cross sections (Fig. 4b, e), projections onto the plane nOu (Fig. 4a, d) confirm the stochastic and fractal behavior of the deformation field of structure (III), which differs significantly from the behavior of the deformation field of structure (IV) (Fig. 4c, f). This confirms the non-commutativity the operation of ordinary matrix multiplication $\mathbf{M}_3 - \mathbf{M}_4 = \mathbf{M}_{R1} \cdot \mathbf{M}_{R2} - \mathbf{M}_{R2} \cdot \mathbf{M}_{R1} \neq 0$.

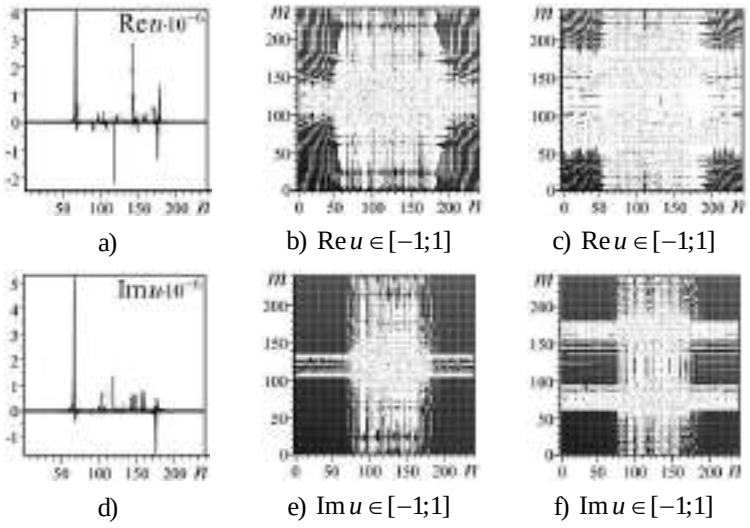


a) variant 1 b) variant 2 c) variant 3

d) variant 1 e) variant 2 f) variant 3

g) variant 1 h) variant 2 i) variant 3

Fig. 3. The behavior of attractors (a, b, c) and the deformation field u of coupled structure (I,II): (d, e, f) – $\mathrm{Re}\, u_{(\mathrm{I,II})} \in [-1;1]$, (g, h, i) – $\mathrm{Im}\, u_{(\mathrm{I,II})} \in [-1;1]$ cross sections (top view).



a) b) $\mathrm{Re}\,u \in [-1;1]$ c) $\mathrm{Re}\,u \in [-1;1]$

d) e) $\mathrm{Im}\,u \in [-1;1]$ f) $\mathrm{Im}\,u \in [-1;1]$

Fig. 4. Deformation fields of structures (III), (IV): $u = u_3$ (a, d) – projections onto the plane nOu, (b, e) – cross sections (top view); $u = u_4$ (c, f) – cross sections (top view).

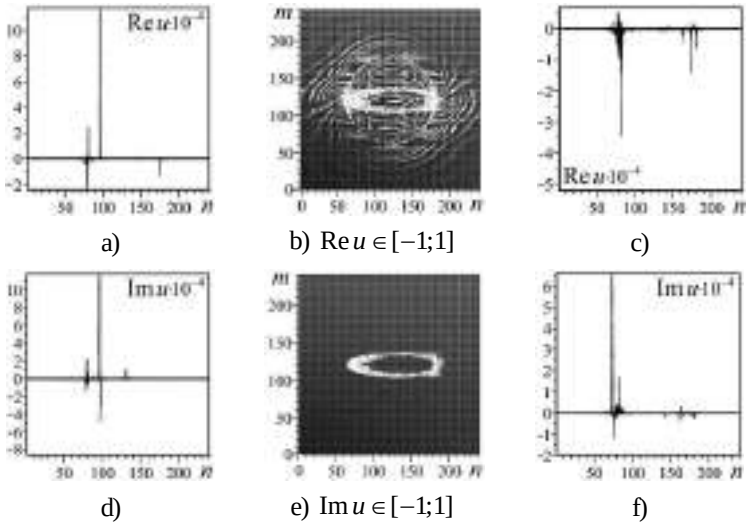
Changing the operation of ordinary matrix multiplication on the scalar multiplication of complex deformation fields leads to the replacement of the coupled structures (III) and (IV) on structures (V) and (VI). In this case, the iterative procedure on index n simulates the coupled (dependent) stochastic processes of the initial independent stochastic processes for separate structures (I) and (II) with the same initial qubit states (0,0).

The deformation fields of structures (V) and (VI) are described by the functions $u_5 = u_{R1} f_5(u_{R2})$ and $u_6 = u_{R2} f_6(u_{R1})$, to which the matrices $\mathbf{M}_5$ and $\mathbf{M}_6$ correspond. If by modeling we use independent iterative processes for structures (I) and (II), then

$$f_5(u_{R2}) = u_{R2}; \quad f_6(u_{R1}) = u_{R1}; \quad u_5 = u_{R1}u_{R2} = u_{R2}u_{R1} = u_6; \quad \mathbf{M}_5 = \mathbf{M}_6. \quad (14)$$

Matrix equality confirms the independence of iterative processes.

Fig. 5 shows the behavior of the complex deformation field for structures (V) and (VI). In this case, structure (II) parameters were chosen corresponding to variant 2. The attractors of structures (V) and (VI) coincide with the attractor from Fig. 3b. Cross sections (Fig. 5b, e), projections onto the plane nOu (Fig. 5a, d) confirm another (compared to Fig. 4) stochastic and fractal deformation field behavior of the structure (V), which also differs significantly from the deformation field behavior of the structure (VI) (Fig. 5c, f). This is due to the dependence of the stochastic processes ($\mathbf{M}_5 \neq \mathbf{M}_6$).



a) b) $\mathrm{Re}\, u \in [-1;1]$ c)

d) e) $\mathrm{Im}\, u \in [-1;1]$ f)

Fig. 5. Deformation fields of structures (V), (VI): $u = u_5$ (a, d) – projections onto the plane nOu, (b, e) – cross sections (top view); $u = u_6$ (c, f) – projections onto the plane nOu.

4 Fractal coupled structures with various superpositional qubits states of separate structures

Next, we consider the superpositional qubits states of fractal coupled structures (V) and (VI). The deformation fields of these structures are described by functions $u_5 = u_{R1} f_5(u_{R2})$ and $u_6 = u_{R2} f_6(u_{R1})$ with the corresponding matrices $\mathbf{M}_5$ and $\mathbf{M}_6$, where the scalar multiplication of complex deformation fields of separate structures (I), (II) is realized. In this case, the iterative procedure on index n simulates coupled (dependent) stochastic processes for the initial independent stochastic processes for structures (I) and (II), the deformation fields of which are described by the functions $u = u_{R1}$ and $u = u_{R2}$. As an example, Fig. 6 shows the behavior of the complex deformation field for structure (V). In this case, the separate structure (I) is a circular cylinder with parameters as for Fig. 1, and the parameters of a separate structure (II) (elliptical cylinder) correspond to variant 2 (the elliptical cylinder approaches the circular cylinder along the axis On). In the coupled structure (V), the separate structures (I), (II) have the same superposition qubit states (1,1) (Fig. 6a, b, d, e) and (-1, -1) (Fig. 6c, f). The behavior of the deformation field of the coupled structure (V) with the same initial qubit states (0,0) is given on Fig. 5a, b, d, e. The presence of same superpositional qubit states (1,1) of separate structures in a coupled system

(Fig. 6a, b, d, e) leads to a change in the complex deformation field compared to Fig. 5a, b, d, e: the decrease amplitudes of peaks, the shift of peaks (Fig. 6a, b), the change of structure (Fig. 6d, e) are observed. An original feature of the deformation field behavior of the coupled structure (V) with the same superpositional states (-1, -1) of separate structures is the absence of the imaginary part of the displacement function in all region ($\mathrm{Im}\, u_5 = 0$), that indicates the absence of effective attenuation. This makes it possible to interpret the coupled structure (V) with the same superpositional states (-1, -1) of the separate structures (I), (II) as a memory cell. For $\mathrm{Re}\, u_5$ the presence of a broadened stochastic peak up is characteristic (Fig. 6c). In this case the cross-sectional structure (Fig. 6f) for state (-1, -1) differs from the cross-sectional structure (Fig. 6d) for state (1,1).

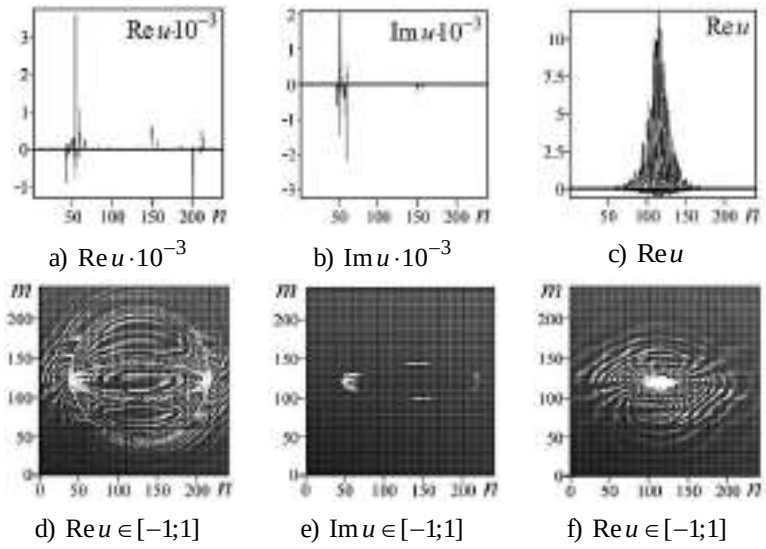
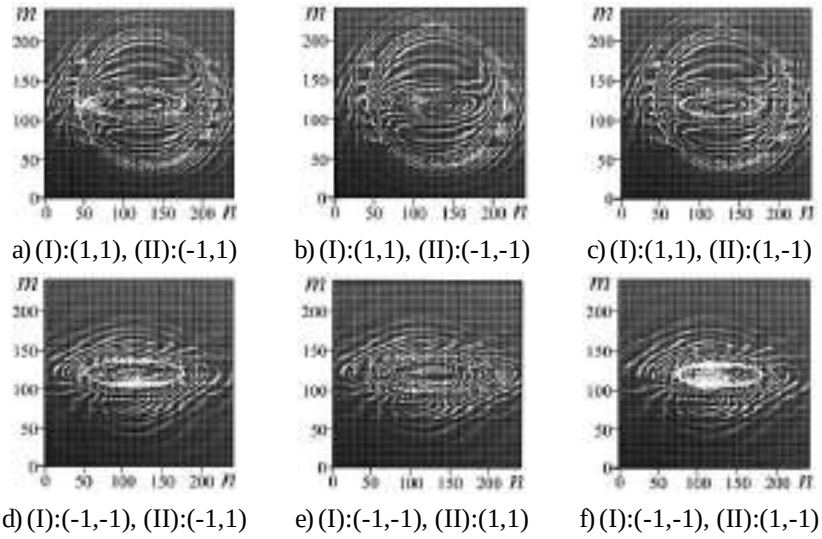


Fig. 6. The behavior of the displacement u of the fractal coupled structure (V): separate structures (I) and (II) have the same superposition qubit states: (1,1) - (a, b, d, e); (-1, -1) - (c, f).

By changing the superposition qubit states of separate structures (I), (II), one can change and control the behavior of the complex deformation field of the coupled structure (V). As an example, Fig. 7 shows the behavior of cross sections $\mathrm{Re}\, u_5$ of the fractal coupled structure (V), when changing superposition qubit states of separate structures (I), (II). If structure (I) is in state (1,1), and the qubit states of structure (II) change (Fig. 7a, b, c), then the complex deformation field of structure (V) changes significantly compared to Fig. 6d, e: for sections $\mathrm{Re}\, u_5$, the effect of mixing of separate trajectories in the

inner region, a change in behavior $\mathrm{Im}\, u_5$ are observed. If structure (I) is in the state (-1, -1), and the qubit states of structure (II) change (Fig. 7d, e, f), then the complex deformation field of structure (V) in comparison with Fig. 6c, f (where $\mathrm{Im}\, u_5 = 0$) arises. In this case, an alteration of the structure of the inner region with the formation of stochastic boundary rings, the effect of mixing of individual trajectories for the cross sections $\mathrm{Re}\, u$ are observed. Using additional (external or internal) action the transitions of separate structures from one qubit state to another can be realized.



a) (I):(1,1), (II):(-1,1) b) (I):(1,1), (II):(-1,-1) c) (I):(1,1), (II):(1,-1)

d) (I):(-1,-1), (II):(-1,1) e) (I):(-1,-1), (II):(1,1) f) (I):(-1,-1), (II):(1,-1)

Fig. 7. The behavior of the cross sections $\mathrm{Re}\, u \in [-1;1]$ (top view) for fractal coupled structure (V). Separate structures (I) and (II) have different superpositional states of qubits.

Similarly, the behavior of the deformation field of the coupled structure (VI), depending on the qubit states of separete structures (II), (I) was studied. In the general case, the deformation field of the coupled structure (VI) is complex. In this case, the conditions

$$u_6 - u_5 = u_{R2} f_6(u_{R1}) - u_{R1} f_5(u_{R2}) \neq 0\,, \quad \mathbf{M}_6 - \mathbf{M}_5 \neq 0\,, \tag{15}$$

are satisfied, that is connected with the dependence of this stochastic processes. This indicates, that the displacement field operators of the separate structures (II), (I) and (I), (II) do not commute in the coupled structures (VI) and (V). As for structure (V), a feature of the deformation field behavior of the coupled structure (VI) with the same superposition states (-1, -1) of separate structures is the absence of effective attenuation in all region ($\mathrm{Im}\, u_6 = 0$). For $\mathrm{Re}\, u_6$ the presence of the

broadened stochastic peak with a structure close to the peak $\mathrm{Re}\,u_5$ (Fig. 6c) is also characteristic, but $\mathrm{Re}\,u_6 - \mathrm{Re}\,u_5 \neq 0$.

Conclusions

By the numerical modelling method the behavior of the deformation field of the coupled fractal structures (circular and elliptical cylinders) in various (initial, basic, superpositional) qubit states was investigated. It is shown, that when the qubit states change, features of the behavior of the complex deformation field of a separate structure are observed. The regular behavior of the inner region $\mathrm{Re}\,u$ is limited by the stochastic boundary (stochastic ring), wherein the concave part of the inner region changes to convex and then to convex-concave. The wave-like behavior for outer region $\mathrm{Re}\,u$ is characteristic. Such behavior allows the interpretation of the inner region as a membrane with the possible alteration of its states due to the change of qubit states. The stochastic behavior for $\mathrm{Im}\,u$, localized in the region of boundary rings with discontinuities is characteristic, wherein in the inner and outer regions of the rings $\mathrm{Im}\,u = 0$.

For fractal coupled structures with initial states of qubits of separate structures, the behavior of attractors and the complex deformation field is considered. It is shown, that the behavior of the deformation field essentially depends on the choice of stochastic processes realized during iterations. As examples, the features of the behavior of the deformation fields resulting from the sum, scalar and matrix products of independent and dependent stochastic processes are investigated.

Fractal coupled structures with various superpositional states of qubits of separate structures are considered. It is shown, that the presence of same superpositional qubit states of separate structures in the coupled system leads to the change in the complex deformation field: there is the decrease in peak amplitudes, peak displacement, and the change in structure. The original feature of the behavior of the deformation field of the coupled structure with the same superpositional states (-1, -1) of separate structures is the absence of effective attenuation ($\mathrm{Im}\,u = 0$), which allows one to interpret the such structure as the memory cell.

By changing the superpositional qubit states of separate structures, one can change and control the behavior of the complex deformation field of the coupled structure. In this case, for the cross sections $\mathrm{Re}\,u$, the alteration of the inner region structure with the formation of stochastic boundary rings, the effect of mixing of separate trajectories is observed. Using additional (external or internal) action transitions of separate structures from one qubit state to another can be realized.

In the general case, the operators of the displacement field of coupled structures depend on the qubit states of separate structures and do not commute.

The results can be used to describe neural networks with variable parameters, in medicine when modeling blood vessels, for quantum information processing.

References

1. O.P. Abramova. Ordering of the Displacement Field Operators of Separate Quantum Dots, Elliptical Cylinders In Coupled Fractal Structures. Bull. of Donetsk National Univer., A, 1, 3-14, 2018.
2. O.P. Abramova, A.V. Abramov. Effect of Ordering of Displacement Fields Operators of Separate Quantum Dots, Elliptical Cylinders on the Deformation Field of Coupled Fractal Structures / C.H. Skiadas and I. Lubashevsky (eds.), 11th Chaotic Modeling and Simulation International Conference, Springer Proceedings in Complexity. Springer Nature Switzerland AG 2019. P. 15-26.
3. O.P. Abramova. Complex Zero Operators in Coupled Fractal Structure with Elements of Cylindrical Type. Bull. of Donetsk National Univer., A, 1, 25-35, 2019.
4. V.S. Abramov. Quantum Dots in a Fractal Multilayer System. Bulletin of the Russian Academy of Sciences. Physics, 81, 5, 625-632, 2017.
5. O.P. Abramova, A.V. Abramov. Coupled Fractal Structures with Elements of Cylindrical Type / C.H. Skiadas and Y. Dimotikalis (eds.), 12th Chaotic Modeling and Simulation International Conference, Springer Proceedings in Complexity. Springer Nature Switzerland AG 2020. P. 15-26.
6. H.-J. Stöckmann. Quantum Chaos. An Introduction. Philipps-Universität Marburg, Germany, 2007.
7. A.N. Omelyanchuk, E.V. Ilyichev, S.N. Shevchenko. Quantum Coherent Phenomena in Josephson Qubits. Kiev, Naukova Dumka, 2013.
8. M.N. Fedorov, I.A. Volkov, Yu.M. Mikhailova. Coutrites and Kukwarts in Spontaneous Parametric Scattering of Light, Correlation and Entanglement of States. JETP, 142, 1(7), 20-43, 2012.
9. M. Nielsen and I. Chuang. Quantum Computation and Quantum Information. Cambridge University Press, New York, 2010.
10. D. Boumeister, A. Eckert, A. Zeilinger. Physics of Quantum Information. Springer, New York, 2001.
11. Y. Ozhigov. Quantum Computers Speed Up Classical with Probability Zero. Chaos Solitons and Fractals, 10, 1707-1714, 1999.
12. D. Castelvecchi. Quantum Computers Ready to Leap out of the Lab. Nature, 541, 9-10, 5 January, 2017.
13. Nonlinearities in Periodic Structures and Metamaterials / ed. Y.S. Kivshar and N.N. Rozanov, Moscow, Fizmatlit, 2014.
14. V.S. Abramov. Alteration of the Stochastic State of the Deformation Field in the Model Multilayer Nanosystem. Bull. of Donetsk National Univer., A, 2, 81-89, 2014.

M-theory as a dynamical system generator

M. Axenides,[1] E. Floratos,[1,2] D. Katsinis,[1,2] G. Linardopoulos[1,2]

[1] Institute of Nuclear and Particle Physics, N.C.S.R. "Demokritos"
153 10, Agia Paraskevi, Greece.

[2] Department of Nuclear and Particle Physics, Faculty of Physics, National and Kapodistrian University of Athens. 157 84 Athens, Greece.

E-mails: axenides@inp.demokritos.gr, mflorato@phys.uoa.gr, dkatsinis@phys.uoa.gr, glinard@inp.demokritos.gr.

Abstract. We review our recent work on ellipsoidal M2-brane solutions in the large-N limit of the BMN matrix model. These bosonic finite-energy membranes live inside $\mathfrak{so}(3) \times \mathfrak{so}(6)$ symmetric plane-wave spacetimes and correspond to local extrema of the energy functional. They are static in $\mathfrak{so}(3)$ and stationary in $\mathfrak{so}(6)$. Chaos appears at the level of radial stability analysis through the explicitly derived spectrum of eigenvalues. The angular perturbation analysis is suggestive of the presence of weak turbulence instabilities that propagate from low to high orders in perturbation theory.

Keywords: Dynamical systems, chaos, M-theory, BMN matrix model, relativistic membranes.

1 Introduction

M-theory By the end of the first superstring revolution (1984-1994), five seemingly different 10-dimensional superstring theories had emerged:

$$\text{Types I, II (IIA, IIB), Heterotic } (\mathfrak{so}(32),\ \mathrm{E}_8 \times \mathrm{E}_8).$$

During the subsequent second superstring revolution (1994-2003), it was found that the 5 superstring theories are connected via a web of dualities (T-duality, S-duality, U-duality, mirror symmetry). What is more, it was realized that the five 10-dimensional superstring theories were just limiting cases of an 11-dimensional theory. This theory was called "M-theory"; it is obtained in the strong-coupling limit ($g_s \to \infty$) of IIA superstring theory. The letter "M" stands for "magic, mystery and matrix" according to one of its founders, E. Witten [1]. Others have associated the letter "M" with "membranes" [2].

Relativistic membranes The idea behind the theory of relativistic membranes is simple: replace 1-dimensional lines (strings) with 2-dimensional surfaces (membranes), much like lines/strings replace 0-dimensional points/particles in the passage from quantum field theory to string theory. Like point particles

13^{th} *CHAOS Conference Proceedings, 9 - 12 June 2020, Florence, Italy*
C. H. Skiadas (Ed)

and strings, membranes are Poincaré invariant objects that can be supersymmetrized. It has been proven that supermembranes can only be defined consistently in 11 spacetime dimensions. Higher-dimensional extended supersymmetric objects (Mp-branes) can be defined in an analogous fashion. Nonetheless, there are reasons to believe that supermembranes (or "M2-branes") are the fundamental objects of the 11-dimensional M-theory, just like strings are the fundamental objects of 10-dimensional string theory.

Matrix models According to the *matrix theory conjecture* of Banks, Fischler, Shenker and Susskind (BFSS) [3], a theory of matrix-discretized supermembranes provides a realization of M-theory in flat spacetime. In the language of matrix models, membranes are fuzzy objects that are represented by $N \times N$ matrices. In the limit of very large matrix dimensions ($N \to \infty$), these matrix models are known to reduce to supermembrane theories.

In 2002, Berenstein, Maldacena and Nastase (BMN) [4] proposed a reformulation of the BFSS matrix model on a particular type of a background that consists of a weakly curved spacetime that is known as a plane-wave, supported by a constant (4-form) field strength:

$$\begin{aligned} ds^2 =& -2dx^+dx^- - \left[\frac{\mu^2}{9}\sum_{i=1}^{3} x_i x_i + \frac{\mu^2}{36}\sum_{j=1}^{6} y_j y_j\right] dx^+dx^+ + \sum_{i=1}^{3} dx_i dx_i + \\ &+ \sum_{j=1}^{6} dy_j dy_j, \qquad F_{123+} = \mu. \end{aligned} \tag{1}$$

Briefly, the BMN matrix model is a deformation of the BFSS matrix model by mass terms and a flux (aka Myers) term. In the large-N limit it is again known [5] that the BMN matrix model reduces to a theory of supermembranes in the 11-dimensional plane-wave background (1). Interestingly, M(atrix) theory has quite recently been applied to the study of chaotic phenomena that take place on the horizons of black holes.

Black holes Black holes (BHs) are regions of spacetime where the force of gravity is so strong that nothing (not even light) can escape. The 2-dimensional surface beyond which it is (classically) impossible for matter or information to escape the gravitational pull of a BH is known as the BH's event horizon. In 1974 Stephen Hawking predicted that it is (quantum-mechanically) possible for BHs to emit thermal radiation and thus slowly evaporate. Because Hawking's radiation is purely thermal, all the information that is stored in BHs seems to get lost.

To resolve the ensuing BH *information paradox* we ultimately need to understand the mechanisms with which information is being stored and processed in BHs. One such mechanism is known as *fast scrambling* or ultra-fast thermalization [6]. More generally, it is widely believed that chaotic phenomena are a dominant feature of BH horizons. Because it is inherently nonlocal, M(atrix) theory turns out to be a valuable tool in the study of information processing by BHs. More precisely, M(atrix) theory can be used to model the dynamics of the microscopic degrees of freedom that are present on BH horizons [7,8].

2 General setup

Our starting point is the Hamiltonian of a bosonic relativistic membrane in the 11-dimensional maximally supersymmetric plane-wave background (1). The Hamiltonian reads, in the so-called light-cone gauge $x^+ = \tau$ [5]:

$$H = \frac{T}{2}\int d^2\sigma \Big[\pi_i^2 + \frac{1}{2}\left\{x_i, x_j\right\}^2 + \frac{1}{2}\left\{y_i, y_j\right\}^2 + \left\{x_i, y_j\right\}^2 + \frac{\mu^2 x^2}{9} + \frac{\mu^2 y^2}{36} - \\ - \frac{\mu}{3}\,\epsilon_{ijk}\left\{x_i, x_j\right\} x_k\Big]. \tag{2}$$

From now on the indices of the coordinates x_i will implicitly be taken to run from 1 to 3, while those of the coordinates y_j will run from 1 to 6.[1] In (2) T stands for the membrane tension and

$$\pi_i^2 \equiv \sum_{i=1}^{3} \dot{x}_i \dot{x}_i + \sum_{j=1}^{6} \dot{y}_j \dot{y}_j, \qquad x^2 \equiv \sum_{i=1}^{3} x_i x_i, \qquad y^2 \equiv \sum_{j=1}^{6} y_j y_j. \tag{3}$$

The definition of the Poisson bracket $\{f, g\}$ that we will be using is

$$\{f\,,\,g\} \equiv \frac{\epsilon_{rs}}{\sqrt{w\left(\boldsymbol{\sigma}\right)}}\,\partial_r f\;\partial_s g = \frac{1}{\sqrt{w\left(\boldsymbol{\sigma}\right)}}\left(\partial_1 f\;\partial_2 g - \partial_2 f\;\partial_1 g\right), \tag{4}$$

where $d^2\sigma = \sqrt{w\left(\boldsymbol{\sigma}\right)}\,d\sigma_1\,d\sigma_2$ is the spatial volume element of the worldvolume and ϵ_{rs} is the 2-dimensional Levi-Civita symbol. In a flat worldvolume it's $w\left(\boldsymbol{\sigma}\right) = 1$ and the usual definition of the Poisson bracket is retrieved.

The Lagrangian equations of motion for the spatial coordinates x and y corresponding to the Hamiltonian (2) are:

$$\ddot{x}_i = \left\{\left\{x_i, x_j\right\}, x_j\right\} + \left\{\left\{x_i, y_j\right\}, y_j\right\} - \frac{\mu^2}{9}\,x_i + \frac{\mu}{2}\epsilon_{ijk}\left\{x_j, x_k\right\} \tag{5}$$

$$\ddot{y}_i = \left\{\left\{y_i, y_j\right\}, y_j\right\} + \left\{\left\{y_i, x_j\right\}, x_j\right\} - \frac{\mu^2}{36}\,y_i. \tag{6}$$

The coordinates x and y can also be shown to obey the Gauss law constraint:

$$\sum_{i=1}^{3}\left\{\dot{x}_i, x_i\right\} + \sum_{j=1}^{6}\left\{\dot{y}_j, y_j\right\} = 0. \tag{7}$$

3 The spherical ansatz

Let us make the following ansatz for the spatial coordinates x and y [9,10]:

$$x_i \equiv x_{1i} = \tilde{x}_{1i}\left(\tau\right) e_1\left(\sigma\right), \qquad i = 1, \ldots, q_1 \tag{8}$$

$$x_{q_1+j} \equiv x_{2j} = \tilde{x}_{2j}\left(\tau\right) e_2\left(\sigma\right), \qquad j = 1, \ldots, q_2 \qquad \& \quad q_1 + q_2 + q_3 = 3 \tag{9}$$

$$x_{q_1+q_2+k} \equiv x_{3k} = \tilde{x}_{3k}\left(\tau\right) e_3\left(\sigma\right), \qquad k = 1, \ldots, q_3 \tag{10}$$

[1] Note also that there's no distinction between upper and lower indices, so that these will be henceforth used interchangeably.

and

$$y_i \equiv y_{1i} = \tilde{y}_{1i}(\tau)\, e_1(\sigma), \qquad i = 1, \ldots, s_1 \tag{11}$$

$$y_{s_1+j} \equiv y_{2j} = \tilde{y}_{2j}(\tau)\, e_2(\sigma), \qquad j = 1, \ldots, s_2 \qquad \& \quad s_1 + s_2 + s_3 = 6 \tag{12}$$

$$y_{s_1+s_2+k} \equiv y_{3k} = \tilde{y}_{3k}(\tau)\, e_3(\sigma), \qquad k = 1, \ldots, s_3. \tag{13}$$

The ansatz (8)–(13) splits the coordinates x and y into three groups

$$x_{ai} = \tilde{x}_{ai}(\tau)\, e_a \qquad \& \qquad y_{bj} = \tilde{y}_{bj}(\tau)\, e_b, \tag{14}$$

where $i = 1, \ldots, q_a$, $j = 1, \ldots, s_b$, $a, b = 1, 2, 3$. Going over to spherical coordinates, $(\sigma_1, \sigma_2) \to (\theta, \phi)$, we define:[2]

$$(e_1, e_2, e_3) = (\cos\phi \sin\theta, \sin\phi \sin\theta, \cos\theta), \qquad \phi \in [0, 2\pi), \quad \theta \in [0, \pi] \tag{16}$$

$$\{e_i, e_j\} = \epsilon_{ijk}\, e_k, \qquad \int e_i\, e_j\, d^2\sigma = \frac{4\pi}{3}\, \delta_{ij}. \tag{17}$$

Note that the Gauss law constraint (7) is automatically satisfied by the ansatz (8)–(13). Now consider the following solutions:

$$\tilde{\mathbf{x}}_1(\tau) = e^{\Omega_{x1}\tau} \cdot \tilde{\mathbf{x}}_{10}, \qquad \tilde{\mathbf{x}}_2(\tau) = e^{\Omega_{x2}\tau} \cdot \tilde{\mathbf{x}}_{20}, \qquad \tilde{\mathbf{x}}_3(\tau) = e^{\Omega_{x3}\tau} \cdot \tilde{\mathbf{x}}_{30} \tag{18}$$

$$\tilde{\mathbf{y}}_1(\tau) = e^{\Omega_{y1}\tau} \cdot \tilde{\mathbf{y}}_{10}, \qquad \tilde{\mathbf{y}}_2(\tau) = e^{\Omega_{y2}\tau} \cdot \tilde{\mathbf{y}}_{20}, \qquad \tilde{\mathbf{y}}_3(\tau) = e^{\Omega_{y3}\tau} \cdot \tilde{\mathbf{y}}_{30}. \tag{19}$$

As in the case of flat space (worked out in [11]) it can be shown that the radii

$$r_{x1}^2 \equiv \tilde{x}_1^2 = \sum_{i=1}^{q_1} \tilde{x}_{10i}\tilde{x}_{10i},\ r_{x2}^2 \equiv \tilde{x}_2^2 = \sum_{j=1}^{q_2} \tilde{x}_{20j}\tilde{x}_{20j},\ r_{x3}^2 \equiv \tilde{x}_3^2 = \sum_{k=1}^{q_3} \tilde{x}_{30k}\tilde{x}_{30k} \tag{20}$$

$$r_{y1}^2 \equiv \tilde{y}_1^2 = \sum_{i=1}^{s_1} \tilde{y}_{10i}\tilde{y}_{10i},\ r_{y2}^2 \equiv \tilde{y}_2^2 = \sum_{j=1}^{s_2} \tilde{y}_{20j}\tilde{y}_{20j},\ r_{y3}^2 \equiv \tilde{y}_3^2 = \sum_{k=1}^{2_3} \tilde{y}_{30k}\tilde{y}_{30k} \tag{21}$$

of the ansatz (18)–(19) can be determined (for all the antisymmetric matrices Ω_{x1}, Ω_{x2}, Ω_{x3}, Ω_{y1}, Ω_{y2}, Ω_{y3}) in terms of the conserved angular momenta

$$(\ell_{x1})_{ij} \equiv \dot{\tilde{x}}_{1i}\tilde{x}_{1j} - \tilde{x}_{1i}\dot{\tilde{x}}_{1j}, \qquad (\ell_{y1})_{ij} \equiv \dot{\tilde{y}}_{1i}\tilde{y}_{1j} - \tilde{y}_{1i}\dot{\tilde{y}}_{1j} \tag{22}$$

$$(\ell_{x2})_{ij} \equiv \dot{\tilde{x}}_{2i}\tilde{x}_{2j} - \tilde{x}_{2i}\dot{\tilde{x}}_{2j}, \qquad (\ell_{y2})_{ij} \equiv \dot{\tilde{y}}_{2i}\tilde{y}_{2j} - \tilde{y}_{2i}\dot{\tilde{y}}_{2j} \tag{23}$$

$$(\ell_{x3})_{ij} \equiv \dot{\tilde{x}}_{3i}\tilde{x}_{3j} - \tilde{x}_{3i}\dot{\tilde{x}}_{3j}, \qquad (\ell_{y3})_{ij} \equiv \dot{\tilde{y}}_{3i}\tilde{y}_{3j} - \tilde{y}_{3i}\dot{\tilde{y}}_{3j}, \tag{24}$$

by minimizing the corresponding effective potential of the membrane. This is completely equivalent to plugging the ansatz (18)–(19) into the equations of motion (5)–(6) and determining the relation between the radii r_{x1}, r_{x2}, r_{x3}, r_{y1}, r_{y2}, r_{y3} and the components of the matrices Ω_{x1}, Ω_{x2}, Ω_{x3}, Ω_{y1}, Ω_{y2}, Ω_{y3} (which in turn always combine to form the conserved angular momenta ℓ_{x1}, ℓ_{x2}, ℓ_{x3}, ℓ_{y1}, ℓ_{y2}, ℓ_{y3}).

[2] We use the volume element in (θ, ϕ) space which implies that $\sqrt{w(\boldsymbol{\sigma})} = \sin\theta$ should be used in the definition (4) of the Poisson bracket. For alternative parametrizations such as

$$(e_1, e_2, e_3) = (cn(\phi|m)\, sn(\theta|n), sn(\phi|m)\, sn(\theta|n), sn(\theta|n)), \tag{15}$$

where $\phi \in [0, 4\mathbb{K}(m))$ and $\theta \in [0, 2\mathbb{K}(n)]$, the corresponding volume element is $\sqrt{w(\boldsymbol{\sigma})} = sn(\theta|n)\, dn(\theta|n)\, dn(\phi|m)$.

4 Effective potentials

The energy of the membrane (2) becomes:

$$E = \frac{2\pi T}{3}\Bigg[\dot{\tilde{x}}_1^2+\dot{\tilde{x}}_2^2+\dot{\tilde{x}}_3^2+\dot{\tilde{y}}_1^2+\dot{\tilde{y}}_2^2+\dot{\tilde{y}}_3^2+\tilde{x}_1^2\tilde{x}_2^2+\tilde{x}_2^2\tilde{x}_3^2+\tilde{x}_3^2\tilde{x}_1^2+\tilde{y}_1^2\tilde{y}_2^2+\tilde{y}_2^2\tilde{y}_3^2+ \\ +\tilde{y}_3^2\tilde{y}_1^2+\tilde{x}_1^2\left(\tilde{y}_2^2+\tilde{y}_3^2\right)+\tilde{x}_2^2\left(\tilde{y}_3^2+\tilde{y}_1^2\right)+\tilde{x}_3^2\left(\tilde{y}_1^2+\tilde{y}_2^2\right)+\frac{\mu^2}{9}\,\tilde{x}^2+ \\ +\frac{\mu^2}{36}\,\tilde{y}^2-2\mu\,\epsilon_{ijk}\,\tilde{x}_{1i}\tilde{x}_{2j}\tilde{x}_{3k}\Bigg]. \tag{25}$$

We now proceed to the following decomposition of the coordinates:

$$\dot{\tilde{x}}_1^2\equiv\dot{\tilde{x}}_{1i}\dot{\tilde{x}}_{1i}=\dot{r}_{x1}^2+\frac{\ell_{x1}^2}{r_{x1}^2},\qquad \dot{\tilde{y}}_1^2\equiv\dot{\tilde{y}}_{1j}\dot{\tilde{y}}_{1j}=\dot{r}_{y1}^2+\frac{\ell_{y1}^2}{r_{y1}^2} \tag{26}$$

$$\dot{\tilde{x}}_2^2\equiv\dot{\tilde{x}}_{2i}\dot{\tilde{x}}_{2i}=\dot{r}_{x2}^2+\frac{\ell_{x2}^2}{r_{x2}^2},\qquad \dot{\tilde{y}}_2^2\equiv\dot{\tilde{y}}_{2j}\dot{\tilde{y}}_{2j}=\dot{r}_{y2}^2+\frac{\ell_{y2}^2}{r_{y2}^2} \tag{27}$$

$$\dot{\tilde{x}}_3^2\equiv\dot{\tilde{x}}_{3i}\dot{\tilde{x}}_{3i}=\dot{r}_{x3}^2+\frac{\ell_{x3}^2}{r_{x3}^2},\qquad \dot{\tilde{y}}_3^2\equiv\dot{\tilde{y}}_{3j}\dot{\tilde{y}}_{3j}=\dot{r}_{y3}^2+\frac{\ell_{y3}^2}{r_{y3}^2}. \tag{28}$$

Plugging (20)–(21) and (26)–(28) into (25), we find that the energy of the membrane becomes

$$E = \frac{2\pi T}{3}\Bigg[\dot{r}_{x1}^2+\dot{r}_{x2}^2+\dot{r}_{x3}^2+\dot{r}_{y1}^2+\dot{r}_{y2}^2+\dot{r}_{y3}^2+\frac{\ell_{x1}^2}{r_{x1}^2}+\frac{\ell_{x2}^2}{r_{x2}^2}+\frac{\ell_{x3}^2}{r_{x3}^2}+\frac{\ell_{y1}^2}{r_{y1}^2}+\frac{\ell_{y2}^2}{r_{y2}^2}+ \\ +\frac{\ell_{y3}^2}{r_{y3}^2}+r_{x1}^2r_{x2}^2+r_{x2}^2r_{x3}^2+r_{x3}^2r_{x1}^2+r_{y1}^2r_{y2}^2+r_{y2}^2r_{y3}^2+r_{y3}^2r_{y1}^2+ \\ +r_{x1}^2\left(r_{y2}^2+r_{y3}^2\right)+r_{x2}^2\left(r_{y3}^2+r_{y1}^2\right)+r_{x3}^2\left(r_{y1}^2+r_{y2}^2\right)+\frac{\mu^2}{9}(r_{x1}^2+ \\ +r_{x2}^2+r_{x3}^2)+\frac{\mu^2}{36}\left(r_{y1}^2+r_{y2}^2+r_{y3}^2\right)-2\mu\,\epsilon_{ijk}\tilde{x}_{1i}\tilde{x}_{2j}\tilde{x}_{3k}\Bigg], \tag{29}$$

so that the corresponding effective potential reads

$$V_{\text{eff}} = \frac{2\pi T}{3}\Bigg[\frac{\ell_{x1}^2}{r_{x1}^2}+\frac{\ell_{x2}^2}{r_{x2}^2}+\frac{\ell_{x3}^2}{r_{x3}^2}+\frac{\ell_{y1}^2}{r_{y1}^2}+\frac{\ell_{y2}^2}{r_{y2}^2}+\frac{\ell_{y3}^2}{r_{y3}^2}+r_{x1}^2r_{x2}^2+r_{x2}^2r_{x3}^2+r_{x3}^2r_{x1}^2+ \\ +r_{y1}^2r_{y2}^2+r_{y2}^2r_{y3}^2+r_{y3}^2r_{y1}^2+r_{x1}^2\left(r_{y2}^2+r_{y3}^2\right)+r_{x2}^2\left(r_{y3}^2+r_{y1}^2\right)+ \\ +r_{x3}^2\left(r_{y1}^2+r_{y2}^2\right)+\frac{\mu^2}{9}\left(r_{x1}^2+r_{x2}^2+r_{x3}^2\right)+\frac{\mu^2}{36}\left(r_{y1}^2+r_{y2}^2+r_{y3}^2\right)- \\ -2\mu\,\epsilon_{ijk}\tilde{x}_{1i}\tilde{x}_{2j}\tilde{x}_{3k}\Bigg]. \tag{30}$$

The above potential (30) contains four different kinds of terms, either repulsive or attractive: (1) kinetic/angular momentum terms (repulsive), (2)

quartic interaction terms (attractive), (3) mass terms (attractive), and (4) a cubic Myers flux term (repulsive). The last two kinds of terms (i.e. the mass terms (3) and the Myers term (4)) are μ-dependent and so they drop out in the $\mu \to 0$ limit (flat space) that was studied in [11]. In both cases (either $\mu = 0$ or $\mu \neq 0$), it is the equilibration of attractive and repulsive forces that determines the extrema of the potential. The two extra repulsive/attractive terms for $\mu \neq 0$ (induced by the plane-wave background) increase the complexity of the resulting dynamical system, as it will become apparent below.

There are three ways to distribute the $\mathfrak{so}(3)$ coordinates x_i $(i = 1, 2, 3)$ into the three groups that are specified by the units e_i in (16), so that we can generally distinguish three main types of membrane configurations. The first two of them (labelled types I and II below) describe rotating membranes (tops) that are point-like (collapsed) in one or two $\mathfrak{so}(3)$ directions and have a vanishing Myers flux term. The third type (III) is probably the most interesting one as it contains all four kinds of repulsive and attractive terms that we described above and extends into the full geometric background of $\mathfrak{so}(3) \times \mathfrak{so}(6)$. Let us now introduce these three types of configurations.

4.1 Type I: $q_1 = 3$, $q_2 = q_3 = 0$

For $q_1 = 3$, $q_2 = q_3 = 0$ we have

$$r_x \equiv r_{x1}, \quad r_{x2} = r_{x3} = 0 \quad \& \quad \ell_x \equiv \ell_{x1}, \quad \ell_{x2} = \ell_{x3} = 0 \tag{31}$$

and the flux term vanishes. The effective potential (30) of the membrane becomes:

$$V_{\text{eff}} = \frac{2\pi T}{3}\left[\frac{\ell_x^2}{r_x^2} + \frac{\ell_{y1}^2}{r_{y1}^2} + \frac{\ell_{y2}^2}{r_{y2}^2} + \frac{\ell_{y3}^2}{r_{y3}^2} + r_{y1}^2 r_{y2}^2 + r_{y2}^2 r_{y3}^2 + r_{y3}^2 r_{y1}^2 + r_x^2\left(r_{y2}^2 + r_{y3}^2\right)\right.$$
$$\left. + \frac{\mu^2 r_x^2}{9} + \frac{\mu^2}{36}\left(r_{y1}^2 + r_{y2}^2 + r_{y3}^2\right)\right]. \tag{32}$$

Apart from the completely symmetric (single-radius) configuration $r = r_x = r_{y1} = r_{y2} = r_{y3}$, $\ell = \ell_x = \ell_{y1} = \ell_{y2} = \ell_{y3}$, the radii and the momenta of the effective potential (32) may be grouped into 5 different axially symmetric (2-radii) configurations and 4 more configurations with 3 different radii. Each of these potentials possesses a local minimum that corresponds to a stationary top solution with time-independent radius and nonzero total angular momentum. There are no static solutions (i.e. having constant radius and zero angular momentum) in this case.

4.2 Type II: $q_1 = 2$, $q_2 = 1$, $q_3 = 0$

For $q_1 = 2$, $q_2 = 1$ and $q_3 = 0$,

$$r_{x3} = 0 \quad \& \quad \ell_{x2} = \ell_{x3} = 0 \tag{33}$$

and the flux term vanishes again. The effective potential (30) becomes:

$$V_{\text{eff}} = \frac{2\pi T}{3}\left[\frac{\ell_{x1}^2}{r_{x1}^2}+\frac{\ell_{y1}^2}{r_{y1}^2}+\frac{\ell_{y2}^2}{r_{y2}^2}+\frac{\ell_{y3}^2}{r_{y3}^2}+r_{x1}^2r_{x2}^2+r_{y1}^2r_{y2}^2+r_{y2}^2r_{y3}^2+r_{y3}^2r_{y1}^2+\right.$$
$$+r_{x1}^2\left(r_{y2}^2+r_{y3}^2\right)+r_{x2}^2\left(r_{y3}^2+r_{y1}^2\right)+\frac{\mu^2}{9}\left(r_{x1}^2+r_{x2}^2\right)+$$
$$\left.+\frac{\mu^2}{36}\left(r_{y1}^2+r_{y2}^2+r_{y3}^2\right)\right]. \tag{34}$$

Although again this case does not lead to any static configuration (with constant radius and zero angular momentum), we may construct one single-radius ($r = r_{x1} = r_{x2} = r_{y1} = r_{y2} = r_{y3}$, $\ell = \ell_{x1} = \ell_{y1} = \ell_{y2} = \ell_{y3}$) solution, 13 axially symmetric (2-radii) tops and 21 tops with 3 different radii.

For example let us consider a type II configuration with all the $\mathfrak{so}(6)$ variables set equal to zero:

$$x_1 = x(\tau)\cdot e_1, \quad x_2 = y(\tau)\cdot e_1, \quad x_3 = z(\tau)\cdot e_2, \quad y_i = 0, \quad i = 1,\dots,6, \tag{35}$$

where the time-dependent part has the form (18). In this case the effective potential (34) becomes:

$$V_{\text{eff}} = \frac{2\pi T}{3}\left[\frac{\ell^2}{x^2+y^2}+\left(x^2+y^2\right)z^2+\frac{\mu^2}{9}\left(x^2+y^2+z^2\right)\right], \tag{36}$$

after setting $\ell_{x1} = \ell$ for simplicity. The corresponding extremisation condition $\nabla V_{\text{eff}} = 0$ implies

$$x\,z^2+\frac{\mu^2 x}{9}-\frac{x\,\ell^2}{\left(x^2+y^2\right)^2} = y\,z^2+\frac{\mu^2 y}{9}-\frac{y\,\ell^2}{\left(x^2+y^2\right)^2} = z\left(x^2+y^2\right)+\frac{\mu^2 z}{9} = 0,$$

which is solved by

$$x^2+y^2 = \frac{3\ell}{\mu} \quad \& \quad z = 0. \tag{37}$$

Complying with (18), we can choose e.g.:

$$x(\tau) = \sqrt{\frac{3\ell}{\mu}}\cos\frac{\mu\,\tau}{3}, \qquad y(\tau) = \sqrt{\frac{3\ell}{\mu}}\sin\frac{\mu\,\tau}{3}, \qquad z(\tau) = 0. \tag{38}$$

Equivalently we could have directly plugged (35) into the equations of motion (5)–(6):

$$\ddot{x}\cdot e_1 = -x\,z^2\cdot e_1-\frac{\mu^2 x}{9}\cdot e_1+\mu\,y\,z\cdot e_3 \tag{39}$$

$$\ddot{y}\cdot e_1 = -y\,z^2\cdot e_1-\frac{\mu^2 y}{9}\cdot e_1+\mu\,x\,z\cdot e_3 \tag{40}$$

$$\ddot{z}\cdot e_2 = -z\left(x^2+y^2\right)\cdot e_2-\frac{\mu^2 z}{9}\cdot e_2. \tag{41}$$

It is easily seen that any solution of the type (18) will again satisfy (37).

4.3 Type III: $q_1 = q_2 = q_3 = 1$

For $q_1 = q_2 = q_3 = 1$, we write:

$$x_1 = r_{x1} e_1,\ x_2 = r_{x2} e_2,\ x_3 = r_{x3} e_3 \quad \& \quad \ell_{x1} = \ell_{x2} = \ell_{x3} = 0. \tag{42}$$

Note that r_{x1}, r_{x2}, r_{x3} are not radii anymore, but coordinates. The effective potential (30) of the membrane can be written as:

$$\begin{aligned} V_{\text{eff}} = \frac{2\pi T}{3}\Bigg[& \frac{\ell_{y1}^2}{r_{y1}^2} + \frac{\ell_{y2}^2}{r_{y2}^2} + \frac{\ell_{y3}^2}{r_{y3}^2} + r_{x1}^2 r_{x2}^2 + r_{x2}^2 r_{x3}^2 + r_{x3}^2 r_{x1}^2 + r_{y1}^2 r_{y2}^2 + r_{y2}^2 r_{y3}^2 + \\ & + r_{y3}^2 r_{y1}^2 + r_{x1}^2 \left(r_{y2}^2 + r_{y3}^2\right) + r_{x2}^2 \left(r_{y3}^2 + r_{y1}^2\right) + r_{x3}^2 \left(r_{y1}^2 + r_{y2}^2\right) + \\ & + \frac{\mu^2}{9}\left(r_{x1}^2 + r_{x2}^2 + r_{x3}^2\right) + \frac{\mu^2}{36}\left(r_{y1}^2 + r_{y2}^2 + r_{y3}^2\right) - 2\mu r_{x1} r_{x2} r_{x3}\Bigg]. \end{aligned} \tag{43}$$

By combining the various radii (along with the corresponding angular momenta) into groups of one, two or three, we obtain 30 different top configurations, one of which corresponds to a completely symmetric top, 9 to axially symmetric (2-radii) tops and 10 to tops that have 3 different radii.

5 Simple type III solutions

The $\mathfrak{so}(3) \times \mathfrak{so}(3) \times \mathfrak{so}(3) \subset \mathfrak{so}(3) \times \mathfrak{so}(6)$ invariant ansatz

$$x_i = \tilde{u}_i(\tau)\, e_i, \qquad y_j = \tilde{v}_j(\tau)\, e_j, \qquad y_{j+3} = \tilde{w}_j(\tau)\, e_j, \qquad i, j = 1, 2, 3 \tag{44}$$

was studied in [12]. The ansatz (44) is obviously of the form (42) (type III) and it describes rotating and pulsating membranes of spherical topology. The corresponding Hamiltonian

$$H = \frac{2\pi T}{3}\left(\tilde{p}_u^2 + \tilde{p}_v^2 + \tilde{p}_w^2\right) + U, \tag{45}$$

is obtained by integrating out the worldvolume coordinates θ and ϕ. The potential energy U reads

$$\begin{aligned} U = \frac{2\pi T}{3}\Bigg[& \tilde{u}_1^2\tilde{u}_2^2 + \tilde{u}_2^2\tilde{u}_3^2 + \tilde{u}_3^2\tilde{u}_1^2 + \tilde{r}_1^2\tilde{r}_2^2 + \tilde{r}_2^2\tilde{r}_3^2 + \tilde{r}_3^2\tilde{r}_1^2 + \tilde{u}_1^2\left(\tilde{r}_2^2 + \tilde{r}_3^2\right) + \\ & + \tilde{u}_2^2\left(\tilde{r}_3^2 + \tilde{r}_1^2\right) + \tilde{u}_3^2\left(\tilde{r}_1^2 + \tilde{r}_2^2\right) + \frac{\mu^2}{9}\left(\tilde{u}_1^2 + \tilde{u}_2^2 + \tilde{u}_3^2\right) + \\ & + \frac{\mu^2}{36}\left(\tilde{r}_1^2 + \tilde{r}_2^2 + \tilde{r}_3^2\right) - 2\mu\tilde{u}_1\tilde{u}_2\tilde{u}_3\Bigg], \quad \tilde{r}_j^2 \equiv \tilde{v}_j^2 + \tilde{w}_j^2,\ j = 1, 2, 3. \end{aligned} \tag{46}$$

The manifest $\mathfrak{so}(2) \times \mathfrak{so}(2) \times \mathfrak{so}(2)$ symmetry of the Hamiltonian (45)–(46) with respect to the $\mathfrak{so}(6)$ coordinates $\tilde{v}_i$ and $\tilde{w}_i$ implies that any solution of the equations of motion preserves three $\mathfrak{so}(2)$ angular momenta ℓ_i $(i = 1, 2, 3)$.

The kinetic terms of the Hamiltonian (45) can be expressed in terms of the conserved angular momenta ℓ_i as

$$\tilde{p}_v^2 + \tilde{p}_w^2 = \sum_{i=1}^{3} \left(\dot{\tilde{r}}_i^2 + \frac{\ell_i^2}{\tilde{r}_i^2} \right), \tag{47}$$

leading to the effective potential

$$V_{\text{eff}} = U + \frac{2\pi T}{3} \left(\frac{\ell_1^2}{\tilde{r}_1^2} + \frac{\ell_2^2}{\tilde{r}_2^2} + \frac{\ell_3^2}{\tilde{r}_3^2} \right). \tag{48}$$

5.1 The $\mathfrak{so}(3)$ symmetric membrane

Let us now consider the simplest possible subsystem of (44) where the $\mathfrak{so}(6)$ variables $\tilde{v}_i$ and $\tilde{w}_i$ are set to zero [12]:

$$\tilde{v}_i = \tilde{w}_i = 0, \qquad i = 1, 2, 3. \tag{49}$$

Scaling out the mass parameter μ by setting

$$t = \mu\tau, \qquad \tilde{u}_i = \mu u_i \tag{50}$$

leads to the form

$$V_{\text{eff}} = \frac{2\pi T \mu^4}{3} \left[u_1^2 u_2^2 + u_2^2 u_3^2 + u_1^2 u_3^2 + \frac{1}{9} \left(u_1^2 + u_2^2 + u_3^2 \right) - 2 u_1 u_2 u_3 \right] \tag{51}$$

of the membrane effective potential (48) and the Hamilton equations of motion,

$$\dot{u}_1 = p_1, \qquad \dot{p}_1 = -\left[u_1 \left(u_2^2 + u_3^2 \right) + \frac{u_1}{9} - u_2 u_3 \right] \tag{52}$$

$$\dot{u}_2 = p_2, \qquad \dot{p}_2 = -\left[u_2 \left(u_3^2 + u_1^2 \right) + \frac{u_2}{9} - u_3 u_1 \right] \tag{53}$$

$$\dot{u}_3 = p_3, \qquad \dot{p}_3 = -\left[u_3 \left(u_1^2 + u_2^2 \right) + \frac{u_3}{9} - u_1 u_2 \right]. \tag{54}$$

The effective potential (51) is a particular instance of the generalized 3-dimensional Hénon-Heiles potential that was introduced in [13],

$$\begin{aligned} V_{\text{HH}} = & \frac{1}{2} \left(u_1^2 + u_2^2 + u_3^2 \right) + K_3 \, u_1 u_2 u_3 + K_0 \left(u_1^2 + u_2^2 + u_3^2 \right)^2 + \\ & + K_4 \left(u_1^4 + u_2^4 + u_3^4 \right), \end{aligned} \tag{55}$$

with $K_3 = -9$, $K_0 = -K_4 = 9/4$. The critical points of the effective potential (51) are:

$$\mathbf{u}_0 = 0, \qquad \mathbf{u}_{1/6} = \frac{1}{6} \cdot (1, 1, 1), \qquad \mathbf{u}_{1/3} = \frac{1}{3} \cdot (1, 1, 1). \tag{56}$$

6 more critical points can be obtained by flipping the sign of exactly two u_i's. This is consistent with the manifest tetrahedral (T_d) symmetry of the potential (51). The extrema $\mathbf{u}_0$ (point-like membrane) and $\mathbf{u}_{1/3}$ (Myers dielectric sphere) are global degenerate minima of the potential while $\mathbf{u}_{1/6}$ is a saddle point:

$$V_{\text{eff}}(0) = V_{\text{eff}}\left(\frac{1}{3} \right) = 0, \qquad V_{\text{eff}}\left(\frac{1}{6} \right) = \frac{2\pi T \mu^4}{6^4}. \tag{57}$$

Radial spectrum [12] By radially perturbing the 9 critical points u_i^0 in (56) as

$$u_i = u_i^0 + \delta u_i\,(t)\,, \qquad \delta u_i\,(t) = \sum_{k=1}^{6} c_k e^{i\lambda t}\xi_{ik}, \tag{58}$$

we may confirm the above conclusion by examining the corresponding Hessian matrix. It turns out that $\mathbf{u}_0$ and $\mathbf{u}_{1/3}$ are global minima (positive-definite Hessian) and $\mathbf{u}_{1/6}$ is a saddle point (indefinite Hessian). These results are summarized in the following table 1.

critical point	eigenvalues λ^2 (#)	stability
$\mathbf{u}_0$	$\frac{1}{9}$ (3), $\frac{1}{36}$ (6)	center (S)
$\mathbf{u}_{1/6}$	$-\frac{1}{18}$ (1), $\frac{5}{18}$ (2), $\frac{1}{12}$ (6)	saddle point
$\mathbf{u}_{1/3}$	$\frac{1}{9}$ (1), $\frac{4}{9}$ (2), $\frac{1}{4}$ (6)	center (S)

Table 1. Radial spectrum of the $\mathfrak{so}\,(3)$ symmetric membrane.

Angular spectrum [14] We may also perform more general (angular/multipole) perturbations of the following form:

$$x_i\,(t) = x_i^0 + \delta x_i\,(t)\,, \qquad i = 1, 2, 3, \tag{59}$$

where δx_i is expanded in spherical harmonics $Y_{jm}\,(\theta, \phi)$ as

$$x_i\,(t) = \mu u_i\,(t)\,e_i, \quad x_i^0 \equiv \mu u_i^0 e_i, \quad \delta x_i\,(t) = \mu \cdot \sum_{j=1}^{\infty}\sum_{m=-j}^{j} \eta_i^{jm}\,(t)\,Y_{jm}\,(\theta, \phi)\,. \tag{60}$$

For the critical points $\mathbf{u}_0$, $\mathbf{u}_{1/6}$, $\mathbf{u}_{1/3}$ we find the eigenvalues [12]:

$$\mathbf{u}_0: \quad \lambda_P^2 = \lambda_\pm^2 = \frac{1}{9},\ \lambda_\theta^2 = \frac{1}{36} \tag{61}$$

$$\mathbf{u}_{1/6}: \lambda_P^2 = 0,\ \lambda_+^2 = \frac{1}{36}\,(j+1)\,(j+4)\,,\ \lambda_-^2 = \frac{j\,(j-3)}{36},$$
$$\lambda_\theta^2 = \frac{1}{36}\left(j^2 + j + 1\right) \tag{62}$$

$$\mathbf{u}_{1/3}: \lambda_P^2 = 0,\ \lambda_+^2 = \frac{1}{36}\,(j+1)^2\,,\ \lambda_-^2 = \frac{j^2}{9},\ \lambda_\theta^2 = \frac{1}{36}\,(2j+1)^2\,, \tag{63}$$

with multiplicities $d_P = 2j+1$, $d_+ = 2j+3$, $d_- = 2j-1$ and $d_\theta = 6\,(2j+1)$, respectively.

The critical point $\mathbf{u}_0$ (point-like membrane) is obviously stable. $\mathbf{u}_{1/3}$ has a zero mode of degeneracy $2d_P$ while all its other eigenvalues are stable for $j = 1, 2, \ldots$ $\mathbf{u}_{1/6}$ has one $2d_P$-degenerate zero mode for every j and a 10-fold degenerate zero mode for $j = 3$. It is unstable for $j = 1$ (2-fold degenerate) and $j = 2$ (6-fold degenerate). The above results were first obtained by [5] from the matrix model. In the flat-space limit ($\mu \to 0$), we recover the results of [15,16].

5.2 The $\mathfrak{so}(3) \times \mathfrak{so}(3) \times \mathfrak{so}(3)$ symmetric membrane

Similar perturbative analyses can be carried out in the $\mathfrak{so}(3) \times \mathfrak{so}(6)$ sector. A solution of the corresponding equations of motion is given by

$$u_i^0 = u_0, \quad v_j^0(t) = v_0 \cos(\omega t + \varphi_j), \quad w_j^0(t) \equiv v_{j+3}^0(t) = v_0 \sin(\omega t + \varphi_k), \tag{64}$$

where (u_0, v_0) are the critical points of the axially symmetric potential

$$V \equiv \frac{V_{\text{eff}}}{2\pi T \mu^4} = u^4 + 2u^2 v^2 + v^4 + \frac{u^2}{9} + \frac{v^2}{36} - \frac{2u^3}{3} + \frac{\ell^2}{v^2} \tag{65}$$

and $\ell \mu^3 \equiv \ell_1 = \ell_2 = \ell_3$. It can be proven that the critical points (u_0, v_0) always lie within the interval:

$$\frac{1}{6} \leq u_0 \leq \frac{1}{3} \qquad \& \qquad 0 \leq v_0 \leq \frac{1}{12}. \tag{66}$$

Radial spectrum [12] To obtain the radial spectrum we set

$$u_i = u_i^0 + \delta u_i(t), \qquad v_i = v_i^0(t) + \delta v_i'(t), \qquad w_i = w_i^0(t) + \delta w_i'(t), \tag{67}$$

finding six zero eigenvalues and four nonzero ones (quadruply and doubly degenerate):

$$\lambda_{1\pm}^2 = \frac{5u_0}{2} - \frac{1}{9} \pm \sqrt{\frac{1}{9^2} - \frac{u_0}{9} - \frac{5u_0^2}{12} + 4u_0^3} \tag{68}$$

$$\lambda_{2\pm}^2 = \frac{5u_0}{2} - \frac{5}{18} \pm \sqrt{\frac{5^2}{18^2} - \frac{35u_0}{18} + \frac{163u_0^2}{12} - 20u_0^3}. \tag{69}$$

The plots of these eigenvalues can be found in the following figure 1.

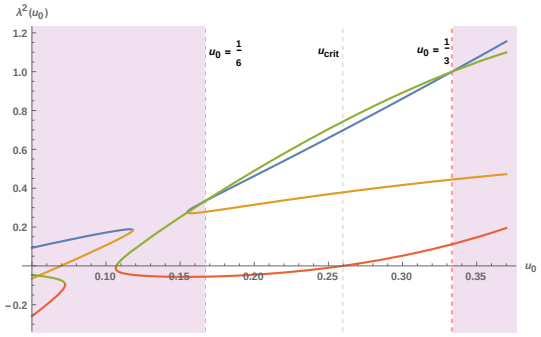


Fig. 1. Radial spectrum of the $\mathfrak{so}(3) \times \mathfrak{so}(6)$ symmetric membrane.

Angular spectrum [14] Going further, we again set out to perform angular/multipole perturbations of the form:

$$x_i = x_i^0 + \delta x_i, \quad i = 1, 2, 3 \qquad \& \qquad y_k = y_k^0 + \delta y_k, \quad k = 1, \ldots, 6, \tag{70}$$

where the $\delta x_i,\ \delta y_k$ are expanded around the classical solution,

$$x_i^0 = \mu u_0 e_i, \quad i = 1,2,3, \qquad y_i^0 = \mu v_i^0(t)\, e_1, \quad i = 1,2 \tag{71}$$

$$y_k^0 = \mu v_k^0(t)\, e_2, \quad k = 3,4 \tag{72}$$

$$y_l^0 = \mu v_l^0(t)\, e_3, \quad l = 5,6, \tag{73}$$

in spherical harmonics $Y_{jm}(\theta,\phi)$:

$$\delta x_i = \mu \cdot \sum_{j,m} \eta_i^{jm}(\tau)\, Y_{jm}(\theta,\phi), \qquad \delta y_k = \mu \cdot \sum_{j,m} \epsilon_k^{jm}(\tau)\, Y_{jm}(\theta,\phi) \tag{74}$$

$$\delta y_l = \mu \cdot \sum_{j,m} \zeta_l^{jm}(\tau)\, Y_{jm}(\theta,\phi), \tag{75}$$

for $i = 1,2,3$, $k = 1,3,5$, $l = 2,4,6$. We find that one of the eigenvalues always vanishes, two others are given by the following analytic expression

$$\lambda_P^2 = \frac{1}{2}\left(j^2+j+2\right)u_0 - \frac{1}{18}\Big(1 + j(j+1) \pm \\ \pm 3\sqrt{144\left(j^2+j-2\right)u_0^3 - 12\left(j^2+j-14\right)u_0^2 - 24u_0 + 1}\Big), \tag{76}$$

while 6 more eigenvalues $\lambda_\pm$ are also known in closed forms but are too complicated to be included here. The corresponding multiplicities of the eigenvalues are $d_P = 2j+1$, $d_+ = 2j+3$, $d_- = 2j-1$. For $j = 1$ four eigenvalues vanish, while two others coincide with the eigenvalues (68)–(69) that were found from radial perturbations:

$$\lambda_P^2 = 4u_0 + \frac{1}{3}, \qquad \lambda_+^2 = \frac{5u_0}{2} - \frac{1}{9} \pm \sqrt{\frac{1}{9^2} - \frac{u_0}{9} - \frac{5u_0^2}{12} + 4u_0^3} \tag{77}$$

$$\lambda_-^2 = \frac{5u_0}{2} - \frac{5}{18} \pm \sqrt{\frac{5^2}{18^2} - \frac{35u_0}{18} + \frac{163u_0^2}{12} - 20u_0^3}. \tag{78}$$

For $j = 2$ there's one zero eigenvalue while $\lambda_P > 0$. We can also plot the $j = 2$ eigenvalues of $\lambda_\pm$ (figure 2):

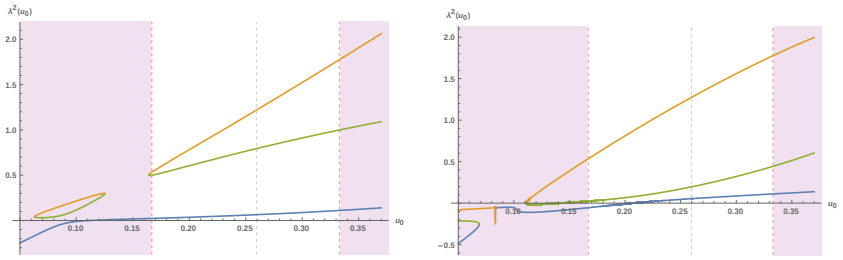


Fig. 2. $\lambda_\pm^2$ for $j = 2$ as a function of u_0.

- The squared nonzero $j = 1$ eigenvalues are all positive/stable in the interval (66), except $\lambda_{-(-)}^2$ which is positive/stable only for $u_{\rm crit} < u_0 < 1/3$, where $u_{\rm crit} \equiv \frac{1}{60}\left(11 + \sqrt{21}\right)$.

- For $j = 2$, the λ_P, λ_+ and one of the λ_- squared eigenvalues are positive/stable in the interval (66). The remaining λ_-^2 eigenvalue is negative/unstable in the interval $\frac{1}{6} \leq u_0 \leq 0.207245 < u_{\text{crit}}$.

- For $j \geq 3$ all the squared eigenvalues are non-negative inside the interval (66) and so the system is stable .

Here's a summary of the angular/multipole spectrum (table 2):

eigenvalues	$j = 1$	$j = 2$	$j \geq 3$	degeneracy
λ_P^2	$0, 0, +$	$0, +, +$	$0, +, +$	$d_P = 2j + 1$
λ_+^2	$0, +, +$	$+, +, +$	$+, +, +$	$d_+ = 2j + 3$
λ_-^2	$0, +, \{0, \pm\}$ $\left(\begin{smallmatrix}\text{positive for}\\ u_0 > u_{\text{crit}}\end{smallmatrix}\right)$	$+, +, \{0, \pm\}$ $\left(\begin{smallmatrix}\text{positive for}\\ u_0 > 0.207245\end{smallmatrix}\right)$	$+, +, +$	$d_- = 2j - 1$

Table 2. Angular spectrum of the $\mathfrak{so}(3) \times \mathfrak{so}(6)$ symmetric membrane.

Higher-order perturbations [17] Beyond linearized perturbation theory (always inside the interval (66)), we anticipate a cascade of instabilities that originates from the $j = 1, 2$ multipoles and propagates towards all higher modes ($j = 3, 4, \ldots$). The perturbative expansion becomes

$$x_i = \sum_{n=0}^{\infty} \varepsilon^n \delta x_i^n = x_i^0 + \sum_{n=1}^{\infty} \varepsilon^n \delta x_i^n, \quad i = 1, 2, 3 \tag{79}$$

$$y_i = \sum_{n=0}^{\infty} \varepsilon^n \delta y_i^n = y_i^0 + \sum_{n=1}^{\infty} \varepsilon^n \delta y_i^n, \quad i = 1, \ldots, 6. \tag{80}$$

It follows that any given mode j at any given order n in perturbation theory couples to all the modes of the previous orders $1, \ldots, n-1$ through an effective forcing term that emerges in the corresponding system of fluctuation equations. The perturbations are expanded in spherical harmonics as

$$\delta x_i^n = \mu \cdot \sum_{j,m} \eta_i^{njm}(\tau) Y_{jm}(\theta, \phi), \qquad \eta_i^{njm}(0) = 0, \qquad i = 1, 2, 3 \tag{81}$$

$$\delta y_i^n = \mu \cdot \sum_{j,m} \theta_i^{njm}(\tau) Y_{jm}(\theta, \phi), \qquad \theta_i^{njm}(0) = 0, \qquad i = 1, \ldots, 6. \tag{82}$$

For example it can be shown that the $(n = 1, j = 1, 2)$ instabilities we found above couple to every mode $(j = 1, 2, \ldots)$ of the second order $(n = 2)$ in perturbation theory.

6 Acknowledgements

The authors would like to thank Christos Skiadas for the invitation to participate to the 13^{th} CHAOS conference. This research is funded in the context of the project "Chaotic dynamics and black holes in BMN theory" E-12386 (MIS 5047794) under the call for proposals "Supporting researchers with an emphasis on young researchers—Cycle B" (EDULLL 103). The project is co-financed by Greece and the European Union (European Social Fund—ESF) by the Operational Programme Human Resources Development, Education and Lifelong Learning 2014-2020.

References

1. E. Witten, *Magic, mystery, and matrix*, *Not. Amer. Math. Soc.* **45** (1998) 1124.
2. M. J. Duff, *Benchmarks on the brane*, hep-th/0407175.
3. T. Banks, W. Fischler, S. H. Shenker, and L. Susskind, *M-theory as a matrix model: A conjecture*, *Phys. Rev.* **D55** (1997) 5112, [hep-th/9610043].
4. D. Berenstein, J. Maldacena, and H. Nastase, *Strings in flat space and pp waves from $\mathcal{N} = 4$ super Yang-Mills*, *JHEP* **04** (2002) 013, [hep-th/0202021].
5. K. Dasgupta, M. M. Sheikh-Jabbari, and M. Van Raamsdonk, *Matrix perturbation theory for M-theory on a pp-wave*, *JHEP* **05** (2002) 56, [hep-th/0205185].
6. Y. Sekino and L. Susskind, *Fast scramblers*, *JHEP* **10** (2008) 065, [arXiv:0808.2096].
7. G. Gur-Ari, M. Hanada, and S. H. Shenker, *Chaos in classical D0-brane mechanics*, *JHEP* **02** (2016) 091, [arXiv:1512.00019].
8. Y. Asano, D. Kawai, and K. Yoshida, *Chaos in the BMN matrix model*, *JHEP* **06** (2015) 191, [arXiv:1503.04594].
9. P. A. Collins and R. W. Tucker, *Classical and quantum mechanics of free relativistic membranes*, *Nucl. Phys.* **B112** (1976) 150.
10. T. Harmark and K. G. Savvidy, *Ramond-Ramond field radiation from rotating ellipsoidal membranes*, *Nucl. Phys.* **B585** (2000) 567, [hep-th/0002157].
11. M. Axenides and E. Floratos, *Euler top dynamics of Nambu-Goto p-branes*, *JHEP* **03** (2007) 093, [hep-th/0608017].
12. M. Axenides, E. Floratos, and G. Linardopoulos, *M2-brane dynamics in the classical limit of the BMN matrix model*, *Phys. Lett.* **B773** (2017) 265, [arXiv:1707.02878].
13. K. Efstathiou and D. Sadovskií, *Perturbations of the 1:1:1 Resonance with tetrahedral symmetry: a three degree of freedom analogue of the two degree of freedom Hénon–Heiles Hamiltonian*, *Nonlinearity* **17** (2004) 415.
14. M. Axenides, E. Floratos, and G. Linardopoulos, *Multipole stability of spinning M2 branes in the classical limit of the BMN matrix model*, *Phys. Rev.* **D97** (2018) 126019, [arXiv:1712.06544].
15. M. Axenides, E. G. Floratos, and L. Perivolaropoulos, *Metastability of spherical membranes in supermembrane and matrix theory*, *JHEP* **11** (2000) 020, [hep-th/0007198].
16. M. Axenides, E. G. Floratos, and L. Perivolaropoulos, *Quadrupole instabilities of relativistic rotating membranes*, *Phys. Rev.* **D64** (2001) 107901, [hep-th/0105292].
17. M. Axenides, E. Floratos, D. Katsinis, and G. Linardopoulos to appear.

Hopf bifurcation analysis for the Fitzhugh-Nagumo model of a spiking neuron

Ilknur Kusbeyzi Aybar

Department of Computer Education and Instructional Technology,
Faculty of Education, Yeditepe University, Istanbul, 34755, Turkey
(E-mail: ikusbeyzi@yeditepe.edu.tr)

Abstract. The Fitzhugh-Nagumo model, which describes a pulse transmission activity in a neuron, is first called the Bonhoeffer-van der Pol model since it is originally transformed from the well-known van der Pol model. The complexity of the neural dynamical models consist of multi-parameter nonlinear systems often allow studying only a particular case for some given values of parameters and prevent obtaining general results. In this study, we present general parameter regions for the existence and the stability of Hopf bifurcation for the Fitzhugh-Nagumo model.
Keywords: Fitzhugh-Nagumo model, Limit cycle, Stability, Periodic solutions.

1 Introduction

Neurons, the smallest members of the brain, transmit information between each other through electrical activities. The electrical activities of a single neuron can be modeled and analyzed by dynamical systems. The communication among neurons observed as firing or spikes occurs as an oscillation formation or loss around a singular point of the dynamical system, i.e., Hopf bifurcation, when a parameter exceeds a threshold value.

The Fitzhugh-Nagumo model, which was proposed by R. Fitzhugh in 1961 and simulated by J. Nagumo et al. in 1962, is governed by the ODE system

$$\begin{aligned} \frac{dx}{dt} &= x(x-a)(1-x) - y = F(x,y), \\ \frac{dy}{dt} &= \epsilon(x - \gamma y) = G(x,y), \end{aligned} \tag{1}$$

where state variables $x(t)$ and $y(t)$ represent the change in the membrane voltage (action potential) and the change in the number of open potassium channels on the membrane of a single neuron, respectively, over time. In this model, the voltage passes the threshold value a; spiking occurs, then it starts to decrease and stabilizes at the neuron membrane's resting potential. The parameter ϵ denotes the speed of the change in the number of open ion channels. The parameter γ states how strongly y depends on x. All system parameters are assumed to be positive, except for the parameter a, which can be negative.

2 Dynamics of the Fitzhugh-Nagumo model

The eigenvalues of the Jacobian matrix can be calculated at the singular points to determine whether the system is stable or unstable at that point. The investigation of the stability for a dynamical system is important because by

this analysis it is then possible to know if all nearby trajectories approach that point of singularity. If the real parts of the eigenvalues of the system's Jacobian matrix are all negative at that point, the system is stable at this point of singularity; otherwise, it is unstable at that point[8]. If the real parts of the eigenvalues of the Jacobian matrix are a pair of pure imaginary values, then the phenomenon called Hopf bifurcation occurs at that point, which implies that at that point, the system is oscillating, that is, this singularity point is an oscillatory solution of the system. The limit cycle can be either stable or unstable, depending on the behaviour of the system. If all nearby trajectories approach the limit cycle inward and outward, it is a stable limit cycle. In order to determine the parameter regions for the existence of a stable limit cycle, we calculate the first Lyapunov coefficient[4]. This study contributes to the comprehension of the electrical activities of neurons by using the dynamical analysis methods since the spikes in the communication of neurons are mathematically stable oscillatory solutions.

Proposition 1. *System* (1) *has three singular points* $E_0 = (0, 0)$,

$$E_- = (\frac{1}{2}(1 + a - \sqrt{(a-1)^2 - \frac{4}{\gamma}}), \frac{1}{2\gamma}(1 + a - \sqrt{(a-1)^2 - \frac{4}{\gamma}})),$$

and

$$E_+ = (\frac{1}{2}(1 + a + \sqrt{(a-1)^2 - \frac{4}{\gamma}}), \frac{1}{2\gamma}(1 + a + \sqrt{(a-1)^2 - \frac{4}{\gamma}})).$$

2.1 System dynamics at E_0

At E_0, the membrane voltage is zero, and the ion channels are all closed. The eigenvalues of the Jacobian matrix at E_0, which is at the resting potential, are

$$\lambda_{1,2} = \frac{1}{2}(-a - \epsilon\gamma \pm \sqrt{(a - \epsilon\gamma)^2 - 4\epsilon}). \tag{2}$$

Proposition 2. *The singular point* E_0 *is a stable singular point when the system parameters* a, ϵ, *and* γ *are all positive or*

$$-\gamma\epsilon < a \leq 0.$$

Theorem 1. *According to the eigenvalues given in* (2), *system* (1) *exhibits Hopf bifurcation when*

$$a < 0, \epsilon > a^2, a = -\epsilon\gamma. \tag{3}$$

Proof. To investigate Hopf bifurcation at E_0, we accept conditions (3) and calculate the first Lyapunov coefficient at this point. We look for a Lyapunov function of the form

$$h(x, y) = \alpha x^2 + \beta xy + \sigma y^2 + h.o.t.$$

that satisfies

$$\frac{\partial h}{\partial x}F(x, y) + \frac{\partial h}{\partial y}G(x, y) = g_1(x^2 + y^2)^2 + ... \tag{4}$$

By solving (4) together with (1), we obtain

$$\alpha = \frac{\beta\epsilon}{2a}$$

and

$$\sigma = -\frac{\beta}{2\epsilon\gamma}.$$

The quadratic form

$$4\alpha\sigma - \beta^2 = -4 + \frac{4}{\epsilon\gamma^2}$$

is positive-definite when we choose $\beta = 2$ because of condition (3) which reduces to

$$\epsilon > 0, 0 < \gamma < \sqrt{\frac{1}{\epsilon}}$$

in this case. We calculate the first Lyapunov coefficient as

$$g_1 = \frac{6 + 2\gamma(\epsilon\gamma(1 - 2\epsilon\gamma) - 2)}{\gamma(3 + \epsilon(2 + \epsilon(3 + 4\gamma^2)))}. \tag{5}$$

Stable oscillations are indicated by a negative first Lyapunov coefficient. Therefore, we study the case $g_1 < 0$ together with the conditions given in (3) and we conclude in the following Theorem.

Theorem 2. *System* (1) *exhibits supercritical Hopf bifurcation when one of the following cases are satisfied:*

$$\begin{aligned}
&(i) 0 < \gamma < 1, \frac{1}{4\gamma} + \frac{1}{4}\sqrt{-3(\frac{(5\gamma - 8)}{\gamma^3})} < \epsilon < \frac{1}{\gamma^2},\\
&(ii) 1 < \gamma \leq \frac{3}{2}, \frac{1}{4\gamma} + \frac{1}{4}\sqrt{-3(\frac{5\gamma - 8}{\gamma^3})} < \epsilon < \frac{1}{\gamma^2},\\
&(iii) \frac{3}{2} < \gamma \leq \frac{8}{5}, 0 < \epsilon < \frac{1}{4\gamma} - \frac{1}{4}\sqrt{-3(\frac{5\gamma - 8}{\gamma^3})},\\
&(iv) \frac{3}{2} < \gamma \leq \frac{8}{5}, \frac{1}{4\gamma} + \frac{1}{4}\sqrt{-3(\frac{5\gamma - 8}{\gamma^3})} < \epsilon < \frac{1}{\gamma^2},\\
&(v) \gamma > \frac{8}{5}, 0 < \epsilon < \frac{1}{\gamma^2}.
\end{aligned} \tag{6}$$

Example 1. To demonstrate supercritical Hopf bifurcation at the origin we choose the parameter set $(a, \epsilon, \gamma) = (-1.75, 0.5, 3.5)$. The eigenvalues of the Jacobian matrix are $\lambda_{1,2} = \pm 0.661438i$. In Fig. 1, we illustrate that system (1) exhibits stable oscillatory regime, i.e. limit cycle. The trajectories outside and inside approach to the limit cycle. In Fig. 1. (a) we have two initial points, one at the outside, and one at the inside of the limit cycle. This is clearly obvious since the first Lyapunov coefficient at this point is $g_1 = -0.0127119$, as we obtain by (5). In Fig. 1. (b), we choose ten different initial points.

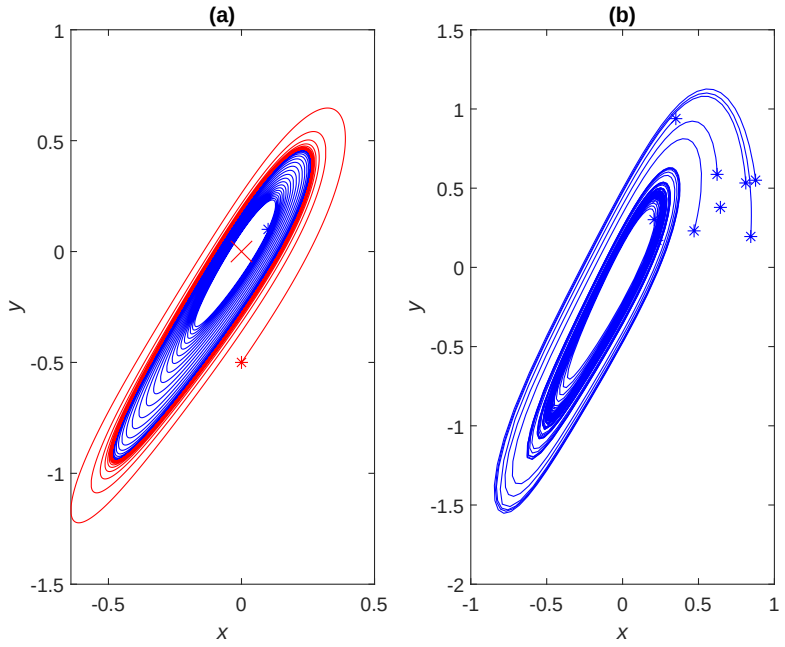


Fig. 1. (a) The supercritical Hopf bifurcation of system (1) when $(a, \epsilon, \gamma) = (-1.75, 0.5, 3.5)$ around the origin. (b) The trajectories moving towards the limit cycle from ten different initial points outside.

2.2 System dynamics at E_- and E_+

In this section, we investigate the local dynamics of system (1) at the singular point E_-. However, same results are obtained for the singular point E_+. First, we make assumption that the state variables and parameter values to be real numbers to represent realistic values, which means that the expressions in the root must be nonnegative. For this reason, we assume the square root terms in the singular points are positive or at least zero. Hence, we begin our investigation for Hopf bifurcation at $E_\pm$ by accepting the condition

$$\Delta = (a-1)^2\gamma - 4 \geq 0. \tag{7}$$

The eigenvalues of the Jacobian matrix at E_- are

$$\lambda_{1,2} = \frac{1}{4\gamma}(2 - \Delta - 2\epsilon\gamma^2 + (1+a)\sqrt{\gamma\Delta} \pm \frac{1}{2}\sqrt{-32\epsilon\gamma^2(\Delta - (1+a)\sqrt{\gamma\Delta}) + 4(2 - \Delta - 2\epsilon\gamma^2 + (1+a)\sqrt{\gamma\Delta})^2}). \tag{8}$$

Proposition 3. *E_- is a stable singular point under one of the following cases:*

$$\begin{aligned} &i.a < 1, \Delta = 0, \epsilon > \frac{2-\Delta}{2\gamma^2}, \\ &ii.a < 1, \Delta > 0, \epsilon > \frac{2-\Delta+\sqrt{\gamma\Delta}(1+a)}{2\gamma^2}, \\ &iii.a > 1, \Delta = 0, \epsilon > \frac{2-\Delta}{2\gamma^2}, \\ &iv.a > 1, \Delta > 0, \epsilon > \frac{2-\Delta+\sqrt{\gamma\Delta}(1+a)}{2\gamma^2}. \end{aligned} \tag{9}$$

Example 2. As an example of Proposition (3), we consider the parameters $(a, \epsilon, \gamma) = (-1, 0.32, 1.31)$. For this parameter set, system (1) is rewritten as

$$\begin{aligned} \frac{dx}{dt} &= x - x^3 - y, \\ \frac{dy}{dt} &= 0.32x - 0.4192y. \end{aligned} \tag{10}$$

The singular points of system (10) are $E_- = (-0.486458, -0.371342)$, $E_0 = (0, 0)$, and $E_+ = (0.486458, 0.371342)$. The eigenvalues of the Jacobian matrix at E_- and E_+ are $\lambda_{1,2} = -0.0645618 \pm 0.440717i$. Since the real parts of the eigenvalues are negative, E_- and E_+ are stable singular points. The origin is a saddle point with a positive, $\lambda_1 = 0.718807$, and a negative, $\lambda_2 = -0.138007$, real eigenvalue. The phase portrait for this parameter set is given in Fig. 2. The solid lines indicate the voltage, $x(t)$, and the dashed lines indicate the number of open channels, $y(t)$. We observe that the state variables end up at E_- and E_+, after little fluctuations.

Hopf bifurcation indicates the birth or death of a periodic solution at a singular point under a small perturbation of a parameter. Hopf bifurcation occurs when a complex conjugate pair of eigenvalues of the Jacobian matrix at a singular point becomes pure imaginary. In this case, a limit cycle, which is a stable (unstable) isolated periodic orbit, exists, and the bifurcation is called supercritical (subcritical) Hopf bifurcation.

Considering assumption (7), we find the conditions for the real part of (8) to be zero and the the expression inside the root to be negative so that the eigenvalues (8) are pure imaginary.

Theorem 3. *System* (1) *undergoes Hopf bifurcation at E_- when $a = -1$ if*

$$1 < \gamma < \frac{3}{2}, \epsilon = \frac{3-2\gamma}{\gamma^2}. \tag{11}$$

Moreover, the Hopf bifurcation at E_- is always supercritical (i.e. there exists a stable limit cycle.).

Proof. To investigate the existence and the stability of Hopf bifurcation, we apply a change of variables to move E_- to the origin by the linear transforma-

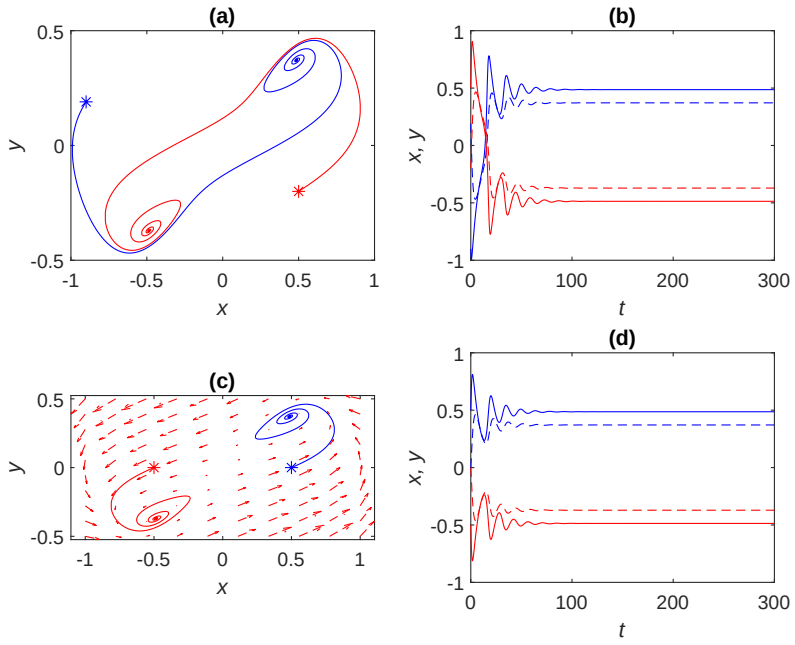


Fig. 2. (a) Stable spiral foci at $E_- = (-0.486458, -0.371342)$, and $E_+ = (0.486458, 0.371342)$ for system (10). (b) Time series plots for the case given in (a). (c) The case given in (a) for different initial points. (d) Time series plots of (c).

tion

$$x = u + \frac{1}{2}(1 + a - \sqrt{\frac{\Delta}{\gamma}}),$$
$$y = v + \frac{1}{2\gamma}(1 + a - \sqrt{\frac{\Delta}{\gamma}}),$$

and obtain the following system:

$$\begin{aligned} \frac{dx}{dt} &= \frac{1}{2\gamma}((2 - \Delta + (1+a)\sqrt{\gamma\Delta})x + (-(1+a)\gamma + 3\sqrt{\gamma\Delta})x^2 - 2\gamma x^3 - 2\gamma y), \\ \frac{dy}{dt} &= \epsilon(x - \gamma y). \end{aligned} \tag{12}$$

System (12) has Hopf bifurcation at the origin when $a = -1$. The eigenvalues of the Jacobian matrix at the singular point at the origin are

$$\lambda_{1,2} = \frac{1}{2\gamma}(3 - \gamma(2 + \epsilon\gamma) + \sqrt{9 + \gamma(-12 + \gamma(4 + \epsilon(2 + \gamma(-4 + \epsilon\gamma))))}. \tag{13}$$

The eigenvalues given in (13) are pure imaginary if

$$1 < \gamma < \frac{3}{2}, \epsilon = \frac{3 - 2\gamma}{\gamma^2}. \tag{14}$$

Hence, system (12) exhibits Hopf bifurcation. Under these conditions, we have positive definite quadratic form.

We observe that when $\Delta = 0$, there is no Hopf bifurcation at the system. Hence, we consider $\Delta > 0$. We also consider $a \neq 1$ for the system to be defined.

Calculating the first Lyapunov coefficient at this point we have

$$\alpha = \frac{\beta\epsilon\gamma}{-2 - \Delta - (1+a)\sqrt{\gamma\Delta}}$$

$$\epsilon = \frac{2 - \Delta + (1+a)\sqrt{\gamma\Delta}}{2\gamma^2}$$

$$\sigma = -\frac{be}{2\epsilon\gamma}$$

We assume the coefficient as $\beta = 2$. Then, we obtain

$$g_1 = -\frac{24(-1+\gamma)\gamma^3}{27 + \gamma(-36 + \gamma(54 + \gamma(-52 + 19\gamma)))} \tag{15}$$

for the first Lyapunov coefficient of system (1) at E_-. When g_1 given in (15) is negative, then there exists a supercritical Hopf bifurcation for system (1). When we solve

$$g_1 < 0$$

and

$$1 < \gamma < \frac{3}{2}, \epsilon = \frac{3 - 2\gamma}{\gamma^2}$$

together, we obtain again condition (14). Therefore, if there is Hopf bifurcation at E_- and E_+, it is a supercritical Hopf bifurcation. Hence, there exists a stable limit cycle.

We explain Theorem (3) with the following numerical example.

Example 3. As an example for Hopf bifurcation, we choose parameter values $(a, \epsilon, \gamma) = (-1, 0.32, 1.25)$, satisfying (3). In this case, system (1) can be written as

$$\begin{aligned} \frac{dx}{dt} &= x - x^3 - y, \\ \frac{dy}{dt} &= 0.32x - 0.4y. \end{aligned} \tag{16}$$

System (16) has the singular points $E_- = (-0.447214, -0.357771)$, $E_0 = (0, 0)$, and $E_+ = (0.447214, 0.357771)$. System (16) possesses Hopf bifurcation at the singular points E_- and E_+, which are located symmetrically on the trajectory plot. This result is due to the eigenvalues of the Jacobian matrix, which are $\lambda_{1,2} = \pm 0.4i$. The stability of Hopf bifurcation is determined by the first Lyapunov coefficient is $g_1 = -1.04639$ as calculated from (15). At the origin, E_0 is a saddle point represented by a positive ($\lambda_1 = 0.712311$) and a negative ($\lambda_2 = -0.112311$) real eigenvalues. The stable limit cycle for this example is illustrated in Fig. 3. In Fig. 3. (a), the blue star is at $(-0.9, 0.19)$

and the red star is at $(0.5, -0.2)$. The trajectories initiating from the blue and the red star approach towards the limit cycle. In Fig. 3. (b), the voltage, $x(t)$, is represented by a solid line, and the number of open ion channels, $y(t)$ is represented by a dashed line. For the two initial points, the blue and the red star, we see the oscillations in $x(t)$ and $y(t)$ for the phase portrait given in Fig. 3. (a). In Fig. 3. (c), two points move along the limit cycle. In Fig. 3. (d), we observe the oscillations of the state variables for the phase portrait given in Fig. 3. (c).

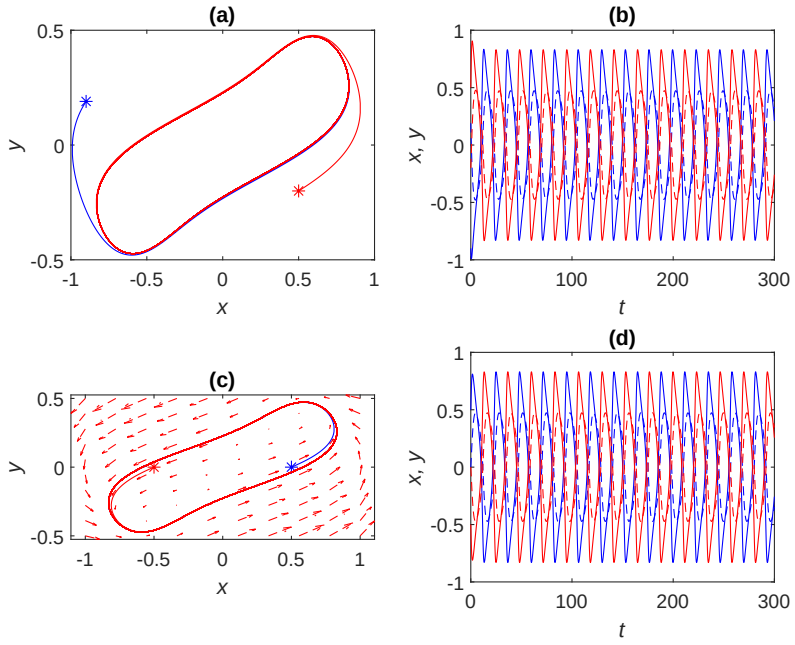


Fig. 3. (a) The stable limit cycle for Example 3 when $(a, \epsilon, \gamma) = (-1, 0.32, 1.25)$ around $E_+ = (0.447214, 0.357771)$ and $E_- = (-0.447214, -0.357771)$. (b) The oscillations in the state variables, i.e. the voltage and the number of open channels given in (a). (c) The trajectories moving towards the limit cycle. (d) The oscillations in the state variables, i.e. the voltage and the number of open channels given in (c).

References

1. I. Kusbeyzi Aybar, O.O. Aybar, B. Fercec, V.G. Romanovski, S. Swarup Samal and A. Weber. Investigation of invariants of a chemical reaction system with algorithms of computer algebra. *MATCH Commun. Math. Comput. Chem.*, 74, 465–480, 2015.
2. L.N.M. Duysens and J. Amesz. Fluorescence spectrophotometry of reduced phosphopyridine nucleotide in intact cells in the near-ultraviolet and visible region. *Biochimica et Biophysica Acta.*, 24, 1, 19–26, 1957.

3. B. Ferčec. *Integrability and local bifurcations in polynomial systems of ordinary differential equations*, Ph.D. Thesis, 2013.
4. Y.A. Kuznetsov. *Elements of Applied Bifurcation Theory*, New York: Springer, 1995.
5. A.M. Liapunov. *Stability of Motion*, with a contribution by V. Pliss. Translated by F. Abramovici and M. Shimshoni, New York, Academic Press, 1996.
6. Mathematica 9.0 Wolfram Research, Inc. https://www.wolfram.com.
7. MATLAB 9.7 MathWorks 2019b https://www.mathworks.com.
8. L. Perko. *Differential Equations and Dynamical Systems*, Texts in Appliesd Mathematics 7, New York: Springer-Verlag, 2001.
9. V.G. Romanovski and D.S. Shafer. *The Center and Cyclicity Problems, A Computational Algebra Approach*, Boston-Basel-Berlin: Birkhauser, 2009.
10. R. Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical J.*, 1, 445–466, 1961.
11. J. Nagumo, S. Arimoto and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proc. IRE.*, 50, 2061–2070, 1962.
12. Y. Xia, M. Grašič, W. Huang and V.G. Romanovski. Limit cycles in a model of olfactory sensory neurons. *Int. J. Bifurcat. Chaos*, 29, 3, 1950038, 2019.
13. L. Zhang and J. Yu. Invariant algebraic surfaces of the FitzHugh-Nagumo system. *J. Math. Anal. Appl.*, 483(2), 123097, 2020.
14. A.G. Korotkov, A.O. Kazakov and T.A. Levanova. Effects of memristor-based coupling in the ensemble of FitzHugh-Nagumo elements. *Eur. Phys. J. Special Topics*, 228, 2325–2337, 2019.

Limit cycles of the Schnakenberg chemical reaction model

I. Kusbeyzi Aybar[1], B. Ferčec[2,3], O. O. Aybar[4], M. Dukarić[3,5]

[1]Department of Computer Education and Instructional Technology, Faculty of Education, Yeditepe University, Istanbul, 34755, Turkey
[2]Faculty of Energy Technology, University of Maribor, Hočevarjev trg 1, 8270 Krško, Slovenia
[3]Center for Applied Mathematics and Theoretical Physics, University of Maribor, Mladinska 3, 2000 Maribor, Slovenia
[4]Piri Reis University, 34940, Tuzla, Istanbul, Turkey
[5]SSI Schaefer, , 2000 Maribor, Slovenia
(E-mail: ikusbeyzi@yeditepe.edu.tr)

Abstract. Dynamical systems contribute to the mathematical modeling of chemical reactions of biological or ecological phenomena producing limit cycle oscillations. In this study, we present a computational approach to examine the bifurcations of limit cycles of the two-dimensional simple chemical reaction model known as the Schnakenberg model. With our approach, we obtain conditions on parameters of the system of the chemical reaction model which gives Hopf bifurcation. Using the Lyapunov function we show the stability of Hopf bifurcation. We illustrate the results with a numerical example.
Keywords: Schnakenberg, Limit cycle, Chemical reaction, Lyapunov function.

1 Introduction

During biochemical reactions, which are the transformations of molecules to other molecules inside the cell, enzymes play roles of biological catalysts and change concentration rates[10]. Biochemical reactions enable cell functions such as digestion and respiration to reproduction and contribute to maintaining the life processes of living organisms. The biochemical reaction, named glycolysis, which is the destruction of glucose to enzymatic acid with enzymes to generate energy, happens in all living organisms. Moreover, in all living organisms, the same enzymes act in this reaction. The fluctuations in the concentrations of the substances lead to glycolytic oscillations, which depend on the concentration rates. The glycolytic oscillations were first observed experimentally by Duysens and Amesz[2]. This phenomenon is known as a fascinating biochemical reaction represented mathematically by a generalized version of the Schnakenberg model[13].

2 The Schnakenberg model

In 1978, J. Schnakenberg introduced the simple chemical reaction system with a limit cycle behaviour, the so called Schnakenberg model, an autocatalytic

chemical reaction model with oscillatory behaviour. This model is characterized by the following three chemical reactions that involve two chemical components and two chemical resources:

$$A \rightleftharpoons X, B \to Y, 2X + Y \to 3X.$$

The dynamical system for this chemical reaction is obtained by considering the state variables x and y as the concentrations at a given time of the chemical substances X and Y which leads to the following system of ordinary differential equations:

$$\begin{aligned} \frac{dx}{dt} &= x^2y - x + b, \\ \frac{dy}{dt} &= -x^2y + a. \end{aligned} \tag{1}$$

Here, the parameters $a > 0$ and $b > 0$ denote the concentration rates of the chemical resources A and B, respectively.

To demonstrate the oscillations in this model, Hwang et al. [4] showed that the dimensionless system of the Schnakenberg model possesses at most one limit cycle in $\mathbb{R}^2_+$. The nonexistence and existence of a positive non-constant steady-state solution to the Schnakenberg system are studied by Li, in 2011 [6]. In this paper, we investigate Hopf bifurcation for the Schnakenberg model by computing the Lyapunov function[12].

System (1) has one unique singular point at $E = (a + b, \frac{a}{(a+b)^2})$, with the eigenvalues of Jacobian matrix given as

$$\lambda_{1,2} = \frac{-1}{2}(1 + (a + b)^2 \pm \sqrt{k(a, b)}), \tag{2}$$

where

$$k(a, b) = (-a + b + (a + b)^3)^2 - 4(a + b)^4. \tag{3}$$

It is obvious that these eigenvalues are pure imaginary if

$$1 + (a + b)^2 = 0, \ k(a, b) < 0, \quad a, b > 0,$$

which simplifies to the condition $0 < a < 1$ and $(a + b)^3 = a - b$. In this case the singularity at E can either be a center (all trajectories are closed), or a focus (all trajectories are spirals)[11]. When the condition

$$(a + b)(1 + (a + b)^2) \neq 0$$

is satisfied, the singular point at E is a focus, if $(a + b)(1 + (a + b)^2) < 0$ all trajectories in a neighborhood of E are moving towards the singularity E (stable focus) and if $(a+b)(1+(a+b)^2) > 0$ all trajectories in a neighborhood of E are moving away the singularity at E. For instance, Figure 1(a) shows stable limit cycle of the system (1) with $a = 0.6$ and $b = 0.17037459017229974$ and eigenvalues are $\lambda_{1,2} = -0.0178966 \pm 0.77016668i$ and Figure 1(b) shows stable limit cycle of the system (1) with $a = 0.6$ and $b = 0.15037459017229974$ and eigenvalues are $\lambda_{1,2} = 0.01806962 \pm 0.75015699i$.[9].

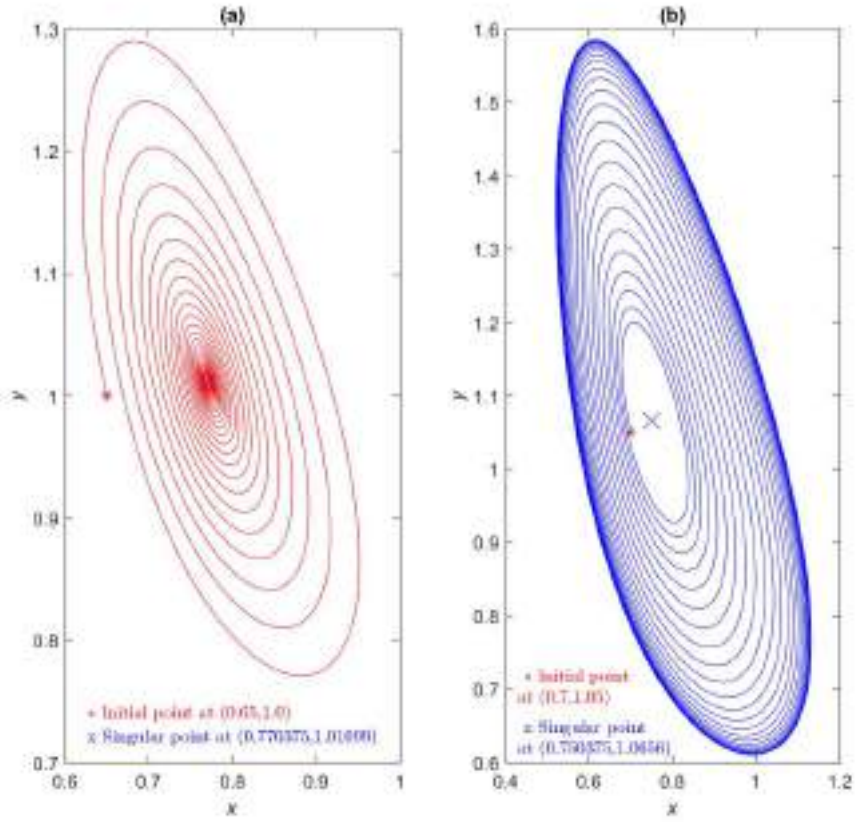


Fig. 1. (a) Stable and (b) Unstable foci of system (1).

3 The existence of Hopf bifurcation

To study Hopf bifurcation for system (1), we first move the singular point E to origin by the linear transformation

$$x \to X + (a+b) \text{ and } y \to Y + \frac{a}{(a+b)^2}, \tag{4}$$

and obtain

$$\begin{aligned} \dot{x} &= \frac{1}{a+b}((a+b)x^2y + ax^2 + +2(a+b)^2xy + (a-b)x + (a+b)^3y) = F(x,y) \\ \dot{y} &= \frac{-1}{a+b}((a+b)x^2y + ax^2 + +2(a+b)^2xy + 2ax + (a+b)^3y) = G(x,y), \end{aligned} \tag{5}$$

where X is rewritten as x, and Y as y. The necessary condition for the existence of Hopf bifurcation at the origin for system (5) is when the trace of the linear

approximation of system (5)

$$\frac{a-b}{a+b} - (a+b)^2,$$

is zero. This condition also satisfies that the real part of the eigenvalues is zero. The second necessary condition is for $k(a,b)$ given in (3) to be negative. These two conditions, with additional two on parameters a and b for the chemical system, $a, b > 0$, form the following system of semi-algebraic equations:

$$\begin{aligned} &(a>0) \ \wedge \ (b>0) \ \wedge ((a+b)^3 - a + b = 0) \ \wedge \\ &(((a+b)^3 - a + b)^2 - 4(a+b)^4 < 0). \end{aligned} \tag{6}$$

System (16) can be solved using `Mathematica` routine `Reduce` [8] and we obtain

$$0 < a < 1 \text{ and } (a+b)^3 = a - b. \tag{7}$$

A commonly used approach for the determination of Hopf bifurcation is the computation of normal forms. However, in this work, we adopt an approach employing the Lyapunov function which we describe now. For a system

$$\begin{aligned} \dot{x} &= -y + P(x,y) = P_1(x,y), \\ \dot{y} &= x + Q(x,y) = Q_1(x,y), \end{aligned} \tag{8}$$

we can always find a function of the form

$$\Psi(x,y) = x^2 + y^2 + \sum_{j+k=3} \psi_{jk} x^j y^k,$$

such that

$$\frac{\partial \Psi}{\partial x} P_1(x,y) + \frac{\partial \Psi}{\partial y} Q_1(x,y) = g_1(x^2+y^2)^2 + g_2(x^2+y^2)^3 + \dots. \tag{9}$$

Based on the Lyapunov Theorem of asymptotic stability [7], we determine the type of the focus stability by using the first nonzero coefficient g_i of the extension of

$$\frac{\partial \Psi}{\partial x} P_1(x,y) + \frac{\partial \Psi}{\partial y} Q_1(x,y).$$

Then, a focus is stable if g_i is negative, and unstable if g_i is positive[1].

If the system is of the form

$$\begin{aligned} \dot{x} &= a_1 x + b_1 y + P(x,y) = P_1(x,y), \\ \dot{y} &= c_1 x - a_1 y + Q(x,y) = Q_1(x,y), \end{aligned} \tag{10}$$

for which the trace of the linear approximation matrix is zero, the resulting expressions involve radicals. To avoid this, we search for a positive-definite Lyapunov function of the form

$$\Psi(x,y) = \alpha x^2 + \beta xy + \gamma y^2 + \sum_{j+k=3} \psi_{jk} x^j y^k$$

which satisfies (9). This is the case if the conditions

$$\alpha = -\frac{c_1\beta}{2a_1},\ \gamma = \frac{b_1\beta}{2a_1} \tag{11}$$

hold, and it is known that the quadratic form

$$\alpha x^2 + \beta xy + \gamma y^2 \tag{12}$$

is positive-definite if $\alpha > 0$ and $4\alpha\gamma - \beta^2 > 0$. Inserting (11) into the expression

$$4\alpha\gamma - \beta^2, \tag{13}$$

we obtain $4\alpha\gamma - \beta^2 = \frac{-\beta^2(a_1^2+b_1c_1)}{a_1^2}$, and when the origin is a center or a focus for system (10), the quadratic form (12) is positive-definite [14].

Theorem 1. *The singular point at the origin (or at E) of system* (5) *(or* (1)*) satisfying conditions* (7) *is a stable focus.*

Proof. When (16) are satisfied, the eigenvalues of the linear approximation matrix of system (1) are

$$\lambda_{1,2} = \pm\frac{\sqrt{-4(a+b)^4 + (-a+b+(a+b)^3)^2}}{2(a+b)}.$$

We look for a Lyapunov function up to degree 8,

$$\Psi_8(x,y) = \sum_{k+s=2}^{8} \psi_{ks}x^k y^s \tag{14}$$

satisfying the equation

$$\frac{\partial\Phi_8}{\partial x}F(x,y) + \frac{\partial\Phi_8}{\partial y}G(x,y) = g_1(x^2+y^2)^2 + g_2(x^2+y^2)^3 + g_3(x^2+y^2)^4. \tag{15}$$

One can see that

$$\alpha = \frac{a\beta}{a-b} \ \text{ and } \ \gamma = \frac{\beta}{2}.$$

This condition can be determined by equating the coefficients of the same monomials on both sides of equation (15).

Let $\beta = 2$, then $\alpha = \frac{2a}{a-b}$, $\gamma = 1$, and $4\alpha\gamma - \beta^2 = \frac{4(a+b)}{a-b} > 0$. For $a > b$, the quadratic form (12) of (14) is positive-definite.

The first nonzero coefficient g_i is

$$g_1 = 4a(a+b)^2(12a^{14} + 168a^{13}b + b^6(1+b^2)(2b^2-5)(1+2b^2)(1+3b^2)+$$
$$+ 2a^{12}(546b^2-1) + 6a^{11}b(728b^2-5) + 7a^9b(5-110b^2+3432b^4)+$$
$$+ 2a^{10}(1-99b^2+6006b^4) + a^7(-57b+300b^3-3564b^5+41184b^7)+$$
$$+ a^5b(-3+167b^2-966b^4-4356b^6+24024b^8) + ab^3(3+5b^2-123b^4-$$
$$+ 325b^6-90b^8+168b^{10}) + a^3b(9-2b^2+13b^4-1860b^6-1430b^8+4368b^{10})+$$
$$+ a^8(-17+33b^2(5-60b^2+1092b^4)) + a^6(6+b^2(-13+924b^4(-5+39b^2)))+$$
$$+ a^4(-3+b^2(-25+3b^2(75-630b^2-990b^4+4004b^6)))+$$
$$+ a^2b^2(-9+b^2(24+b^2(-167+42b^2(-25-11b^2+26b^4)))))/$$
$$((a-b)(2a^6+12a^5b+b^2(2+b^2)(1+2b^2)+a^4(30b^2-3)+a^3(-4b+40b^3)+$$
$$+ a^2(2+6b^2+30b^4)+4ab(-1+3(b^2+b^4)))(3a^{12}+36a^{11}b+b^6(1+b^2)(3+b^2)$$
$$(1+3b^2)+10a^9b(13+66b^2)+a^{10}(13+198b^2)+12a^7b(9+130b^2+198b^4)+$$
$$+ a^8(17+585b^2+1485b^4)+2ab^5(3+42b^2+65b^4+18b^6)+$$
$$+ 4a^3b^3(5+113b^2+390b^4+165b^6)+a^2b^4(5+248b^2+585b^4+198b^6)+$$
$$+ a^6(11+304b^2+2730b^4+2772b^6)+a^5(38b+508b^3+3276b^5+2376b^7)+$$
$$+ 3a^4(16+5b^2(3+38b^2+182b^4+99b^6)))).$$

The semi-algebraic system

$$\begin{aligned}&(g_1 \geq 0) \ \wedge \ (a>0) \ \wedge \ (b>0) \ \wedge ((a+b)^3-a+b=0) \ \wedge \\ &(((a+b)^3-a+b)^2-4(a+b)^4<0).\end{aligned} \tag{16}$$

is an unsolvable system (checked with `Reduce` of `Mathematica`). Since $g_1 < 0$, the derivative with respect to a vector field is negative-definite. Hence the focus is stable[5].

Next theorem summarizes the conditions for the existence of Hopf bifurcation in system (1).

Theorem 2. *For the parameters, a and b, that satisfy the conditions given in* (7)*, Hopf bifurcation can occur at the singular point $E = (a+b, \frac{a}{(a+b)^2})$ of system* (1)*. The bifurcation is always supercritical, i.e., a stable limit cycle is born from E.*

Proof. As demonstrated in the proof of the Theorem 1, the singular point that satisfies conditions (7) is a stable focus. By slightly varying the parameter b, we slightly perturb the system (1), changing the real parts of the eigenvalues (2) to positive. Hence, point E becomes an unstable focus, and the results is a stable limit cycle[3].

3.1 Numerical example

To show the existence of stable limit cycles, we choose parameters a and b as $a = 0.6$ and $b = 0.16037459017229974$. The corresponding system (1) has singular point at $E = (0.760375, 1.03776)$, the eigenvalues given in (2) are $\lambda_{1,2} =$

$\pm 0.760375i$ and the coefficient g_1 of (15) is negative, $g_1 = -0.17187 < 0$. As we can see in Fig. 2. (a), the trajectories move towards the singular point E, i.e. E is stable focus. Now, we perturb b slightly as $b = 0.16037459017229974 - \frac{1}{100} = 0.150375$. Then, eigenvalues become $\lambda_{1,2} = 0.0180689 \pm 0.750157i$ with positive real parts. When we choose initial point $(0.7, 1.05)$, we see that the trajectory moves away from singular point E, on the other hand, the second trajectory plotted from another initial point $(0.6, 0.6)$ moves towards the singular point E. Both trajectories approach to the limit cycle as seen in Fig. 2. (b).

Note that if the eigenvalues are pure imaginary, the local phase portrait in the neighbourhood of the singularity can not be a center, since $g_1 \neq 0$ for positive values of a and b for system (1).

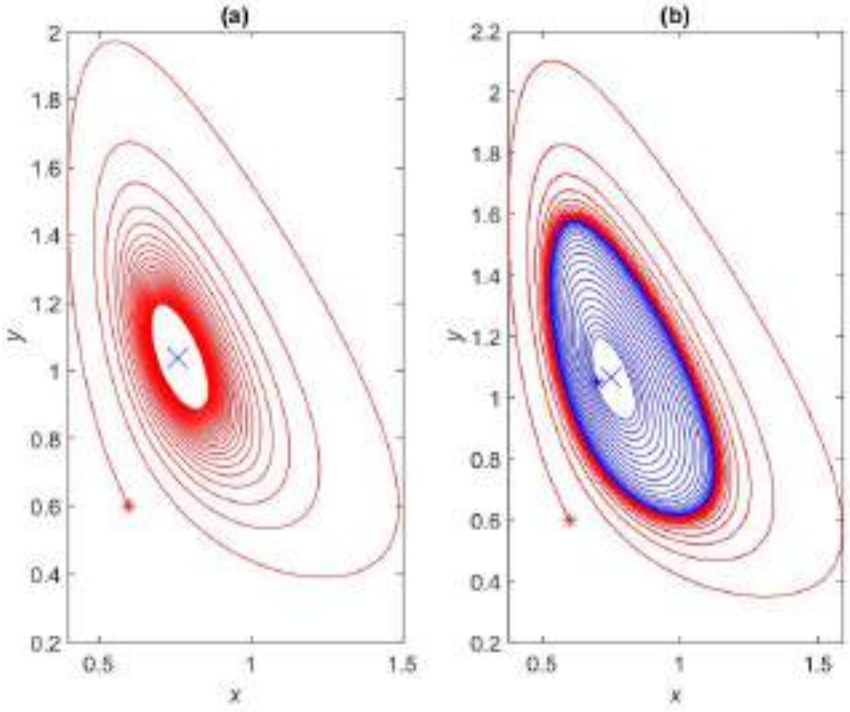


Fig. 2. (a) Stable focus for parameter values $a = 0.6$ and $b = 0.16037459017229974$. (b) A supercritical Hopf bifurcation appearing for system (1) with $a = 0.6$ and $b = 0.150375$.

4 Acknowledgement

This work is supported by the Slovenian Research Agency (research core no. "P1-0306" and the project no. "BI-TR/19-22-003") and the Scientific and Technological Research Council of Turkey (TUBITAK) under project no. 119F017.

References

1. I. Kusbeyzi Aybar, O.O. Aybar, B. Fercec, V.G. Romanovski, S. Swarup Samal and A. Weber. Investigation of invariants of a chemical reaction system with algorithms of computer algebra. *MATCH Commun. Math. Comput. Chem.*, 74, 465–480, 2015.
2. L.N.M. Duysens and J. Amesz. Fluorescence spectrophotometry of reduced phosphopyridine nucleotide in intact cells in the near-ultraviolet and visible region. *Biochimica et Biophysica Acta.*, 24, 1, 19–26, 1957.
3. B. Ferčec. *Integrability and local bifurcations in polynomial systems of ordinary differential equations*, Ph.D. Thesis, 2013.
4. T.-W. Hwang and H.-J. Tsai. Uniqueness of limit cycles in theoretical models of certain oscillating chemical reactions. *J. Phys. A*, 38, 38, 8211–8223, 2005.
5. Y.A. Kuznetsov. *Elements of Applied Bifurcation Theory*, New York: Springer, 1995.
6. Yan Li. Steady-state solution for a general Schnakenberg model *Nonlinear Anal. Real World Appl.*, 12, 1985–1990, 2011.
7. A.M. Liapunov. *Stability of Motion*, with a contribution by V. Pliss. Translated by F. Abramovici and M. Shimshoni, New York, Academic Press, 1996.
8. Mathematica 9.0 Wolfram Research, Inc. https://www.wolfram.com.
9. MATLAB 9.7 MathWorks 2019b https://www.mathworks.com.
10. J.D. Murray. *Mathematical Biology: I. An Introduction*, 3rd edn. Springer, New York, 2002.
11. L. Perko. *Differential Equations and Dynamical Systems*, Texts in Appliesd Mathematics 7, New York: Springer-Verlag, 2001.
12. V.G. Romanovski and D.S. Shafer. *The Center and Cyclicity Problems, A Computational Algebra Approach*, Boston-Basel-Berlin: Birkhauser, 2009.
13. J. Schnakenberg. Simple chemical reaction systems with limit cycle behaviour. *J. theor. Biol.*, 81, 3, 389–400, 1979.
14. Y. Xia, M. Grašič, W. Huang and V.G. Romanovski. Limit cycles in a model of olfactory sensory neurons. *Int. J. Bifurcat. Chaos*, 29, 3, 1950038, 2019.

Fractality and the internal Dirichlet problem

Nilufar A. Azamova.
Mailing address: 100174, Uzbekistan, Tashkent, Vuzgorodok, Universitetskaya str., 4.
E-mail: niluvard@gmail.com
Elena S. Alekseeva.
Mailing address: 100174, Uzbekistan, Tashkent, Vuzgorodok, Universitetskaya str., 4.
E-mail: kometarella@mail.ru.
Alexander A. Potapov.
V. A. Kotelnikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Moscow, Russia
JNU-IREE RAS Joint Lab. of Information Technology and Fractal Processing of Signals, JiNan University, Guangzhou, 510632, China
Mailing address: 125009, Russia, Moscow, Mokhovaya str., 11-7.
E-mail: potapov@cplire.ru
Alexander A. Rassadin.
Mailing address: 603950, Russia, Nizhnij Novgorod, Gagarin ave., 23.
E-mail: brat_ras@list.ru

Headings content

Abstract

In this chapter influence of fractality on solution of the two-dimensional internal Dirichlet problem is analyzed. Two different situations are considered namely the first of them deals with fractal boundary condition on the unit disk. In this case exact solution of the Laplace equation proves to obey to some analog of the de Rham functional equation. Also norm and the Dirichlet integral for this solution has been estimated. In the second situation boundary condition is supposed to be regular but boundary of the domain is fractally perturbed. For clarification of this case both approximate conformal mapping technique and the Potapov concept of physical fractals has been applied.

Abbreviations

IDP – the internal Dirichlet problem.
WF – the Weierstrass function.
TWF – the truncated Weierstrass function.

2.1 The internal Dirichlet problem and fractality of boundary condition

In this section we discuss the internal Dirichlet problem (IDP) on the unit disk Ω_0.

First of all let us consider the two-dimensional Laplace equation:

$$\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\frac{\partial^2 u}{\partial \phi^2}=0, \qquad r<1, \tag{2.1}$$

where r and ϕ are polar coordinates on this disk;

$u(r,\phi)$ is unknown function.

Further equation (2.1) ought to be provided by boundary condition:

$$u(1,\phi)=W(\phi). \tag{2.2}$$

A peculiarity of our point of view on IDP (2.1)-(2.2) is in consideration of fractality of its boundary condition.

As a model of fractal boundary condition we take the well-known Weierstrass function (WF):

$$W(\phi)=\sum_{n=1}^{\infty} a^n\cdot\cos(b^n\cdot\phi), \tag{2.3}$$

where a and b are its parameters.

To provide 2π-periodicity of this function over polar angle ϕ one is obliged to choose function parameter $b=2,\ 3,\ 4,\ \ldots$.

If these parameters obey to inequalities $0<a<1$ and $a\cdot b\geq 1$ then WF (2.3) is a continuous but nowhere differentiable function (see [1] and references there in). Moreover in this case WF possesses by the following fractal dimension:

$$D_F=2+\frac{\ln a}{\ln b}. \tag{2.4}$$

Formula (2.4) demonstrates that one can vary fractal dimension of WF under fixed b continuously from $D_F=1$ to $D_F=2$ by means of changing of its parameter a from $a=1/b$ to $a=1$. That is why we choose WF as the model of fractal boundary condition for IDP (2.1)-(2.2).

General solution of the Laplace equation (2.1) without singulatity in the center of the unit disk is equal to [2]:

$$u(r,\phi)=\sum_{m=0}^{\infty} r^m\cdot(A_m\cdot\cos(m\cdot\phi)+B_m\cdot\sin(m\cdot\phi)). \tag{2.5}$$

Substituting $r=1$ into series (2.5) and comparing the result with WF (2.3) it is easy to find that exact solution of IDP (2.1)-(2.3) is equal to:

$$u(r,\phi)=\sum_{n=1}^{\infty} a^n\cdot r^{b^n}\cdot\cos(b^n\cdot\phi). \tag{2.6}$$

In particular in the center of the disk $u(0,\phi)=0$.

Moreover one can check that function (2.6) obeys to the next functional equation:

$$u(r,\phi) = a \cdot r^b \cdot \cos(b \cdot \phi) + a \cdot u(r^b, b \cdot \phi), \tag{2.7}$$

expressing its self-similarity.

Substituting $r = 1$ into functional equation (2.7) and using formula (2.2) one can easily obtain the de Rham functional equation for WF [1]:

$$W(\phi) = a \cdot \cos(b \cdot \phi) + a \cdot W(b \cdot \phi). \tag{2.8}$$

Further let us calculate on the solution (2.6) the well-known Dirichlet integral:

$$I[u] = \iint_{\Omega_0} (\nabla u)^2 r dr d\phi, \tag{2.9}$$

where ∇ is the operator of two-dimensional gradient.

It is easy to see from functional series (2.6) that vector $\nabla u(1,\phi)$ consists from nowhere differentiable functions therefore it is convenient to calculate functional (2.9) as the following limit:

$$I[u] = \lim_{r \to 1-0} \int_0^{2\pi} u(r,\phi) \cdot \frac{\partial u}{\partial r}(r,\phi) \cdot r \cdot d\phi. \tag{2.10}$$

Under $b = 2,\ 3,\ 4,\ \ldots$ functions $\{\cos(b^n \phi)\}_{n=1}^{\infty}$ are orthogonal on the interval $[0, 2\pi]$:

$$\int_0^{2\pi} \cos(b^n \phi) \cdot \cos(b^m \phi) \cdot d\phi = \pi \cdot \delta_{nm}, \tag{2.11}$$

hence using this relation one can obtain from (2.10) that the Dirichlet integral is equal to:

$$I[u] = \pi \cdot \sum_{n=1}^{\infty} (a^2 \cdot b)^n. \tag{2.12}$$

It is obvious that numerical series (2.12) converges if $a^2 b < 1$ and diverges otherwise. We underline that under growing value of parameter b length of segment of convergence for a: $a \in (1/b, 1/\sqrt{b})$ tends to zero. One can rewrite condition of convergence for sum (2.12) via fractal dimension (2.4) of WF:

$$I[u] = \begin{cases} \dfrac{\pi \cdot a^2 \cdot b}{1 - a^2 \cdot b}, & 1 \le D_F < \dfrac{3}{2} \\ \infty, & \dfrac{3}{2} \le D_F < 2 \end{cases}. \tag{2.13}$$

Formula (2.13) means that if boundary condition (2.2) of IDP in some sense is "weakly" nondifferentiable then series (2.12) is convergent and if WF (2.3) is "strongly" nondifferentiable then series (2.12) is divergent.

Let us now consider norm of function (2.6) in Hilbert space $L^2(\Omega_0)$:

$$\| u \| = \sqrt{\iint_{\Omega_0} u^2(r,\phi) \cdot r dr d\phi}. \tag{2.14}$$

Using formula (2.11) it is not difficult to find that value (2.14) is equal to:

$$\| u \| = \sqrt{\frac{\pi}{2} \cdot \sum_{n=1}^{\infty} \frac{a^{2n}}{b^n + 1}}. \tag{2.15}$$

Under $b \geq 2$ numerical series contained in the right-hand side of the formula (2.15) can be majorized by sum $\sum_{n=1}^{\infty} (a^2 b^{-1})^n$ of convergent geometrical progression. Thus norm (2.14) can be estimated as follows:

$$\| u \| \leq \sqrt{\frac{\pi}{2} \cdot \frac{a^2}{b - a^2}}. \tag{2.16}$$

Upper bound (2.16) is valid under all admissible values of parameter a.

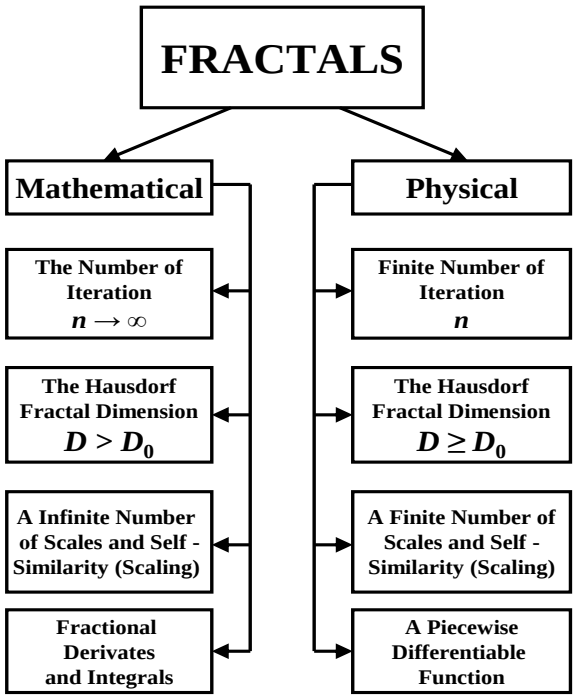


Fig. 2.1. The Potapov scheme of division of fractals on mathematical and physical

In practice to deal with solution (2.6) of IDP (2.1)-(2.2) with boundary condition (2.3) possessing by fractal dimension (2.4) it is convenient to transfer from mathematical fractals to physical ones [1]. The procedure of this transfer is presented on Fig. 2.1. Of course this approach destroys relations (2.7) and (2.8)

but it brings new advantages namely let one take only a limited number N of terms of the series (2.3) then the resulting function:

$$W_N(\phi) = \sum_{n=1}^{N} a^n \cdot \cos(b^n \cdot \phi) \tag{2.17}$$

will be differentiable as many times as required.

Using inequality $|\cos\phi| \le 1$ which is valid for all real ϕ it is not difficult to establish that for any fixed $\delta > 0$ inequality

$$|W(\phi) - W_N(\phi)| < \delta \tag{2.18}$$

is true under

$$N > \left[\frac{\ln \delta(1-a)}{\ln a} \right] + 1. \tag{2.19}$$

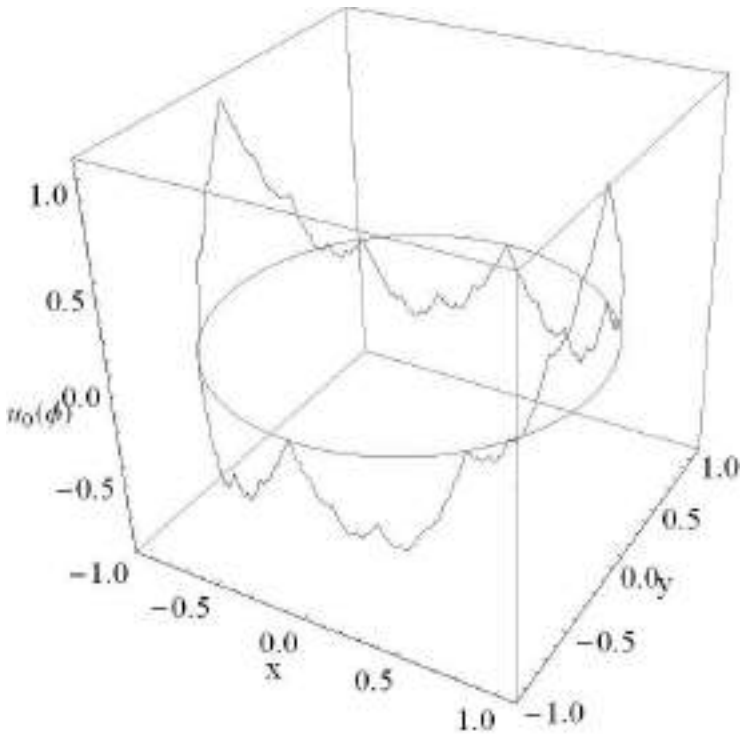


Fig. 2.2. Graph of the boundary condition

In this case we shall call function (2.17) as the truncated Weierstrass function (TWF). In other words if one takes number N for TWF obeying to condition (2.19) then such TWF approximates WF in accordance with inequality (2.18).

Graph of the TWF under $a = 0.75$ and $b = 2$ with $\delta = 0.01$ is presented on Fig. 2.2. In this case $D_F \approx 1.58$ and $N = 21$.

At last let us apply procedure of truncation for series (2.6) with the same number of terms N as in condition (2.19):

$$u_N(r,\phi)=\sum_{n=1}^{N} a^n \cdot r^{b^n} \cdot \cos(b^n \cdot \phi). \tag{2.20}$$

Because of $r<1$ it is easy to estimate that:

$$|u(r,\phi)-u_N(r,\phi)| \le \frac{a^{N+1}}{1-a} \cdot r^{b^{N+1}}. \tag{2.21}$$

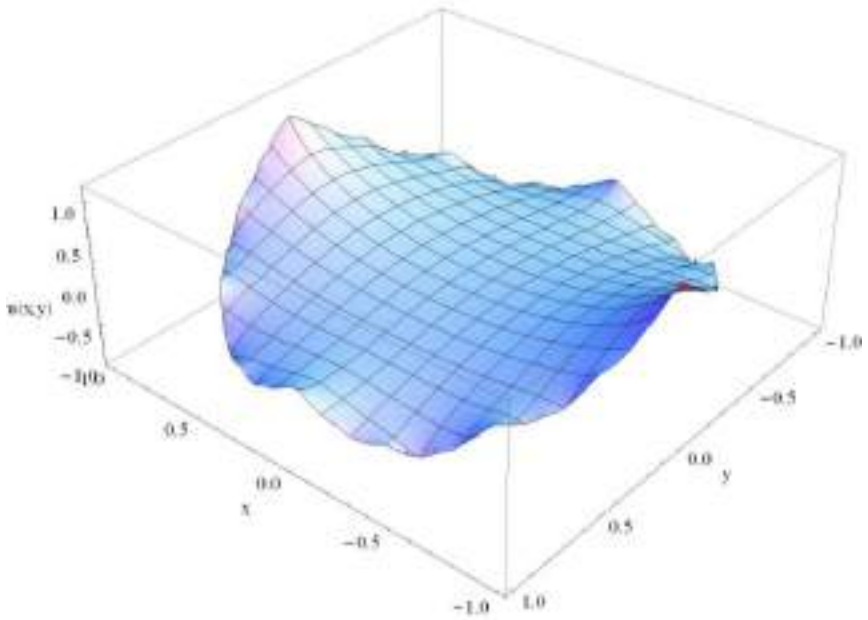


Fig. 2.3. Graph of exact solution

Inequality (2.21) means that in this case function (2.20) approximates function (2.6) even better then TWF approximates WF. From this inequality one can see that in fact fractality of exact solution (2.6) really exists only in narrow ring in the vicinity of the unit circle Γ_0 : $|z|=1$ which is the boundary of the domain Ω_0. Moreover it is clear that under $r<<1$ it is enough only the first term of the series (2.6) to describe behaviour of exact solution of IDP under consideration quite precisely.

Graph of function (2.20) with parameters corresponding to the TWF presented on Fig. 2.2 is shown on Fig. 2.3.

2.2 The internal Dirichlet problem on the nearly circular domain bounded by fractal curve

In this section we consider IDP on the nearly circular domain:

$$\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}=0, \qquad z=x+iy\in\Omega_\varepsilon, \qquad u\big|_{z\in\Gamma_\varepsilon}=u_0(z), \tag{2.22}$$

where Ω_ε is star-shaped domain on complex plane C containing the origin of coordinates;

$\Gamma_\varepsilon=\partial\Omega_\varepsilon$ is boundary of domain Ω_ε;

ε is a small parameter ($0<\varepsilon<<1$) characterizing proximity of domain Ω_ε to the unit disk Ω_0;

$u(z,\varepsilon)$ is unknown function;

$u_0(z)$ is boundary condition.

Let one choose polar equation of Γ_ε in the next form:

$$r(\phi)=1-\varepsilon\cdot\chi(\phi), \qquad \phi\in[0,2\pi], \tag{2.23}$$

where $\chi(\phi)$ is 2π-periodic function over polar angle ϕ and let one suppose that $\chi\in C^2([0,2\pi])$ hence closed curve (2.23) differs slightly from the unit circle Γ_0 both its location and its curvature.

Further the Riemann theorem [3] claims that in this case there is a holomorphic function $w=f(z,\varepsilon)$ realizing conformal mapping of domain Ω_ε on the unit disk Ω_0.

Moreover exact solution of IDP (2.22) can be written via this function $f(z,\varepsilon)$ as follows [4]:

$$u(z,\varepsilon)=\mathrm{Re}\left[\frac{1}{2\cdot\pi\cdot i}\oint_{\Gamma_\varepsilon}u_0(\zeta)\cdot\frac{f(\zeta,\varepsilon)+f(z,\varepsilon)}{f(\zeta,\varepsilon)-f(z,\varepsilon)}\cdot\frac{f'_\zeta(\zeta,\varepsilon)}{f(\zeta,\varepsilon)}\cdot d\zeta\right]. \tag{2.24}$$

Generally speaking one may derive this mapping $f(z,\varepsilon)$ explicitly in the framework of formalism of the harmonic moments of exterior domain $\overline{C}\setminus\Omega_\varepsilon$ [5] but this way is too hard. On the other hand due to representation (2.23) of Γ_ε we can restrict ourselves by construction of approximate conformal mapping of the nearly circular domain on unit disk.

In this case for function $w=f(z,\varepsilon)$ realizing conformal mapping of domain Ω_ε on the unit disk Ω_0 and obeying to conditions $f(0,\varepsilon)=0$ and $f'_z(0,\varepsilon)>0$ the following asymptotic formula is known to be valid [3, 6]:

$$w=f(z,\varepsilon)=z+\varepsilon\cdot f_1(z)+O(\varepsilon^2), \tag{2.25}$$

where

$$f_1(z) = z \cdot \int_0^{2\pi} \chi(\phi) \cdot \frac{\exp(i\phi) + z}{\exp(i\phi) - z} \cdot \frac{d\phi}{2\pi}. \qquad (2.26)$$

A feature of our approach is taking into account a fractality of domain's boundary Γ_ε. On the other hand in accordance with methods developed in [3, 6] for derivation of the desired mapping the boundary Γ_ε must be quite smooth. In order to overcome this obstacle we apply concept of physical fractals (see Fig. 2.1) namely we use as function $\chi(\phi)$ TWF (2.17). Of course this TWF must contain enough terms in correspondence with inequality (2.19) to approximate the input WF with fixed accuracy (2.18). Graph of typical shape of such boundary is shown on Fig. 2.4, dotted line corresponding to the unit circle Γ_0.

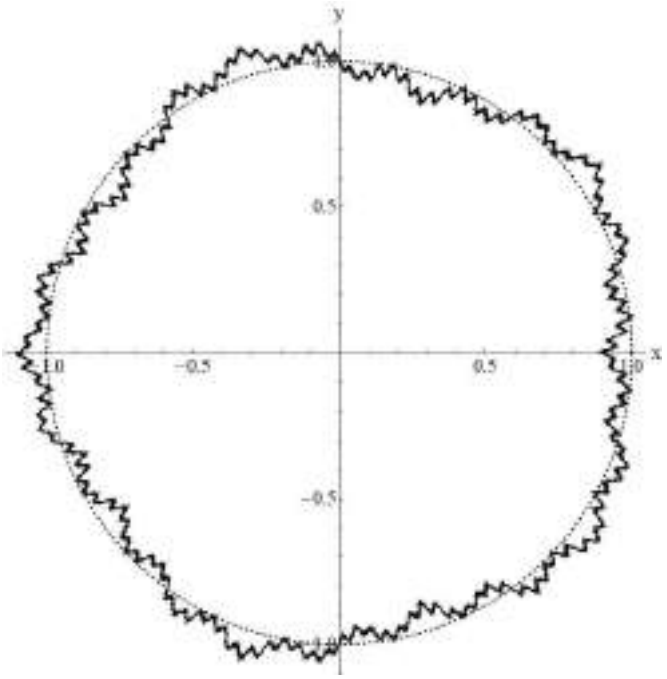


Fig. 2.4. Graph of the domain with fractal boundary

Calculating integral (2.26) with function $W_N(\phi)$ we obtain the next lacunary polynomial of complex variable z:

$$f_1(z) = \sum_{n=1}^{N} a^n \cdot z^{b^n+1}. \qquad (2.27)$$

Substituting expression (2.25) with function (2.27) into formula (2.24) and expanding its integrand on ε one can find approximate solution of IDP (2.22)-(2.23) in the following form:

$$u(z,\varepsilon) = u^{(0)}(z) + \varepsilon \cdot u^{(1)}(z) + O(\varepsilon^2), \quad z \in \Omega_\varepsilon. \tag{2.28}$$

In formula (2.28)

$$u^{(0)}(z) = \mathrm{Re}\left[\frac{1}{2 \cdot \pi \cdot i} \oint_{\Gamma_\varepsilon} u_0(\zeta) \cdot \frac{\zeta + z}{\zeta - z} \cdot \frac{d\zeta}{\zeta} \right] \tag{2.29}$$

and

$$u^{(1)}(z) = u_1^{(1)}(z) + u_2^{(1)}(z) + u_3^{(1)}(z), \tag{2.30}$$

where

$$u_1^{(1)}(z) = \mathrm{Re}\left[\frac{1}{2 \cdot \pi \cdot i} \oint_{\Gamma_\varepsilon} u_0(\zeta) \cdot \left(f_1'(\zeta) - \frac{f_1(\zeta)}{\zeta} \right) \cdot \frac{\zeta + z}{\zeta - z} \cdot \frac{d\zeta}{\zeta} \right], \tag{2.31}$$

$$u_2^{(1)}(z) = -\mathrm{Re}\left[2 \cdot z \cdot \frac{1}{2 \cdot \pi \cdot i} \oint_{\Gamma_\varepsilon} \frac{u_0(\zeta) \cdot f_1(\zeta) \cdot d\zeta}{\zeta \cdot (\zeta - z)^2} \right] \tag{2.32}$$

and

$$u_3^{(1)}(z) = \mathrm{Re}\left[2 \cdot f_1(z) \cdot \frac{1}{2 \cdot \pi \cdot i} \oint_{\Gamma_\varepsilon} \frac{u_0(\zeta) \cdot d\zeta}{(\zeta - z)^2} \right]. \tag{2.33}$$

Zero-order term in asymptotic expansion (2.28) looks like the well-known Poisson integral for the unit disk Ω_0 [3, 4], but we stress that integration in expression (2.29) is performed along the curve Γ_ε with polar equation (2.23) hence to calculate integral in formula (2.29) one ought to substitute into this formula $\zeta(\phi) = (1 - \varepsilon \cdot \chi(\phi)) \cdot \exp(i \cdot \phi)$.

Integrals (2.31), (2.32) and (2.33) representing contribution of boundary roughness into the first order term (2.30) in asymptotic expansion (2.28) must be estimated in the same manner.

This calculation seems to be very awkward but in practice contour integrals in formulae (2.31), (2.32) and (2.33) can be found as sums over residues of its integrands in domain Ω_ε.

To demonstrate this technique let us derive the influence of fractal roughness of boundary of round cylindrical hole in conductor on distribution in it of electrostatic potential and electric field strength. And let us remind that approximate conformal mapping (2.25) of domain with fractal boundary is determined by function (2.27) corresponding to TWF (2.17).

According to general principles of electrostatics let us set:

$$u_0(z) = 1, \quad z \in \Gamma_\varepsilon. \tag{2.34}$$

It is easy to check that in this case calculation in correspondence with formula (2.29) gives us that:

$$u^{(0)}(z) = 1, \quad z \in \Omega_\varepsilon. \tag{2.35}$$

Further let one consider the following function:

$$f_1'(z) - \frac{f_1(z)}{z} = \sum_{n=1}^{N} (a \cdot b)^n \cdot z^{b^n}. \tag{2.36}$$

Lacunary polynomial (2.36) is a holomorphic function of complex variable z on Ω_ε therefore the result of application of the theory of residues to integral in expression (2.31) is equal to:

$$u_1^{(1)}(z) = 2 \cdot \mathrm{Re}\left[f_1'(z) - \frac{f_1(z)}{z} \right]. \tag{2.37}$$

Function $f_1(z)/z$ also has no singularities in domain Ω_ε hence using the Cauchy formula for derivative one can obtain from formula (2.32) that:

$$u_2^{(1)}(z) = -2 \cdot \mathrm{Re}\left[z \cdot \frac{d}{dz} \frac{f_1(z)}{z} \right]. \tag{2.38}$$

At last in accordance with the Cauchy theorem after homotopy of curve Γ_ε into the unit circle Γ_0 in integral (2.33) it is not difficult to see that

$$u_3^{(1)}(z) = 0. \tag{2.39}$$

Thus combining formulae (2.37), (2.38) and (2.39) in correspondence with expression (2.30) we obtain the following unexpected result:

$$u^{(1)}(z) = 0, \quad z \in \Omega_\varepsilon. \tag{2.40}$$

Formula (2.40) means that electrostatic potential of this domain varies only in the second order on ε:

$$u(z, \varepsilon) = 1 + O(\varepsilon^2), \quad z \in \Omega_\varepsilon. \tag{2.41}$$

It is immediately succeed from expression (2.41) that in any point $z \in \Omega_\varepsilon$ electric field strength is equal to:

$$\vec{E}(z, \varepsilon) = -\nabla u(z, \varepsilon) = O(\varepsilon^2). \tag{2.42}$$

Formula (2.42) gives us useful consequence for technique of electric measurements namely under electrostatic screening using round cylindrical hole in conductor one may not take care about precision of boundary processing.

After some generalization on three spatial dimensions the ideology of calculations developed in this section may be useful for description of electrically charged fractal core-shell nanoparticles and elastic tensions around such structures [7].

Acknowledgments

AAP and AER are grateful to the Russian Foundation for Basic Research grant No. 18-08-01356-a. AAP is grateful to the China grant "Leading Talent Program in Guangdong Province" (No. 00201502, 2016-2020) JiNan University (China, Guangzhou).

References

1. Potapov A.A., Gulyaev Yu.V., Nikitov S.A., Pakhomov A.A. and German V.A. The Modern Methods of Image Processing / ed A.A.Potapov – FIZMATLIT, 2008 [in Russian].

2. Morse P.M., Feshbach H. Methods of theoretical physics. Part II. – McGraw-Hill, 1953.

3. Lavrentyev M.A., Shabat B.V. Methods of the Theory of Functions of a Complex Variable – Nauka, 1973 [in Russian].

4. Vladimirov V.S. Equations of Mathematical Physics – Marcel Dekker Inc., 1971.

5. Natanson S. Towards an effectivisation of the Riemann theorem // Ann. Global Anal. Geom. 2005. Vol. 28, no. 3, pp. 233-255.

6. Siryk G.V. On a conformal mapping of near domains // Uspekhi Matem. Nauk. 1956. Vol. 9, no. 5, pp. 57–60 [in Russian].

7. Shishulin A.V., Potapov A.A., Fedoseev V.B. Phase Equilibria in Fractal Core-Shell Nanoparticles of Pb5(VO4)3Cl - Pb5(PO4)3Cl System: The Influence of Size and Shape // In: Advances in Artificial Systems for Medicine and Education II / Ed: Z. Hu, S. Petoukhov, M. He (Part of the Advances in Intelligent Systems and Computing book series – AISC, v. 902). – Cham, Switzerland: Springer Int. Publ., 2019. – P. 405–413.

A New Substitution Box Structure Based on Nose–Hoover Chaotic System

Yaşar Selim Bahçeci[1], Fatih Özkaynak[2]

[1] Department of Software Engineering Fırat University, 23119, Elazig, Turkey
(E-mail: yasarselimbahceci@gmail.com)
[2] Department of Software Engineering Fırat University, 23119, Elazig, Turkey
(E-mail: ozkaynak@firat.edu.tr)

Abstract. A cryptographic protocol should provide two basic requirements for secure communication. These requirements are known as diffusion and confusion. Substitution box structures are needed in order to provide the confusion requirement in block encryption algorithms. These cryptographic blocks must have a nonlinear structure to meet the confusion requirement. Various designs based on chaotic systems have been proposed to ensure the nonlinearity requirement. In this study, a new substitution box structure based on Nose–Hoover Chaotic System is proposed. Successful analysis results showed that the proposed new chaos based substitution box structure could be an alternative to the other three degree chaos based substitution box structures.

Keywords: chaos, cryptography, substitution box, image encryption.

1 Introduction

Our security requirements have changed as everything in our world has changed. The concept of knowledge has become increasingly important in this change [1]. As the concept of knowledge gained importance, the security problem of this information emerged. Researchers have developed many different encryption algorithms to solve this problem. These encryption algorithms must meet various requirements. These requirements are known as diffusion and confusion. Substitution box structures are needed in order to provide the confusion requirement in block encryption algorithms [2]. These cryptographic blocks must have a nonlinear structure to meet the confusion requirement.

Many encryption algorithms use substitution box structures to provide the confusion requirement [3-15]. Although there are many methods for substitution box structures, a design approach that has attracted attention in recent years has been the design approach using chaotic systems. In this study, a substitution box structures approach based on chaotic systems is proposed. The application of this proposed cryptographic structure is shown on an image encryption algorithm.

The rest of the study is organized as follows. In the Section 2, a brief literature summary about chaos based s-box design is given. In the Section 3, the Nose-Hoover chaotic system is introduced. In the Section 4, the proposed new s-box structure is explained and analysis results are given The last section summarizes the study.

2 Related Works

Chaos-based s-box studies have been a remarkable research topic in the last two decades. One of the most important reasons behind this increase in interest in research topic is the developments in cryptanalysis studies. In particular, application attacks allow the attacker to make various inferences about cryptographic protocols using a variety of side channel information. Although designs based on mathematical transformation do not contain weaknesses in terms of performance criteria, their well-defined features allow this side channel information to be easily obtained. Therefore, new searches for alternative s-box structures based on mathematical transformation have accelerated. An important design technique in this aim is chaos based s-box designs.

Chaos based s-box design was first encountered in 2001 [16]. In this study, a s-box design has been realized by using discrete time logistic map and this structure has been used in a block encryption algorithm architecture. Between 2001 and 2010 [17, 18], it has been aimed to improve the performance of s-box structures by using different chaotic maps. However, the performance characteristics of the AES s-box structure could not be approached in these studies. Therefore, performance parameters have been improved with various optimization algorithms [19]. However, the additional transaction costs in these studies have emerged as another problem to be solved .

In 2010, the idea of using continuous time chaotic systems in the design process as an alternative to discrete time chaotic systems has been proposed [20]. The aim of these design studies is to increase the complexity of the chaotic system and improve the nonlinearity properties of s-box designs. Following these studies, several studies aimed at improving performance by using more complex chaotic systems have been proposed. Among these studies, design studies based on hyper chaotic systems [21], time delayed chaotic systems [22] and fractional order chaotic systems [23] draw attention.In chaos based s-box designs, the effect of using only different chaotic systems on performance has not been investigated. In addition to selecting chaotic systems as an entropy source, it is aimed to increase the quality of the entropy source by using various additional procedures [23-30].

In this study, an algorithm based on continuous time chaotic systems has been proposed. The most important aspect of the study is that the system selected as chaotic system does not need any control parameters. This feature of the selected system will have several advantages in the design process of the cryptographic protocol, especially in the process of sharing the secret key of the algorithm (key distribution).

3 Nose–Hoover Chaotic System

Chaos theory is an exciting science. Because it points out that the randomness behind the events have actually mathematical equations. Chaotic behavior first emerged by showing that reason of randomness in weather forecasts modeled by

various differential equations is the internal structure of the system. In the literature, the simplest differential equation models in which chaotic behavior is observed are known as systems like Lorenz, Chua, Chen. The common features of these systems are expressed in third order differential equations.

Nose–Hoover system [31] is a third-order system like the systems mentioned above. The definition using ordinary differential equations is given in Eq. (1). The system in which Eq. (1) is expressed has three initial conditions.

$$\begin{aligned} dx/dt &= y \\ dy/dt &= y * z - x \\ dz/dt &= 1 - y*y \end{aligned} \tag{1}$$

The phase space graphs showing the variation of the state variables of the Nose – Hoover system are shown in Figure 1.

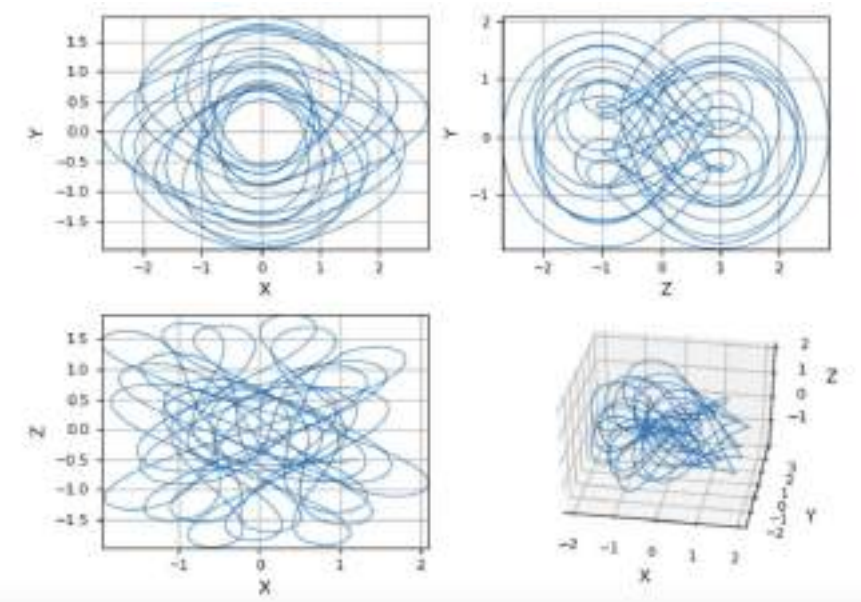


Fig. 1. Phase-space analysis of Nose–Hoover system

4 Proposed S-Box Design Algorithm

The innovative aspect of the study is the chosen chaotic system class. This is the first study in the literature using Nose–Hoover chaotic system in substitution box design. For substitution box design, the recommended method in Ref. [14] is used. The details of the used method can be examined in detail. A program can be produced different substitution box design by running the program repeatedly at different times. The program has a user-friendly design. There is

also an introductory video on how to use the program. There is also an interface for the performance tests of the substitution box structures produced in the program. There are five widely accepted criterion in the literature. These tests are:

- Bijective criterion,
- Nonlinearity criterion,
- Bit independence criterion (BIC),
- Strict avalanche criterion
- Input/output XOR distribution criteria

A sample substitution box structure and performance criteria produced using the proposed chaotic system are given in Table 1. Since the method used for substitution box design automatically provides bijective, this criterion is not included in Table 1.

Table 1. Proposed substitution box structure

s-box																
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**A**	**B**	**C**	**D**	**E**	**F**
0	19	100	82	169	62	29	131	137	16	49	240	105	155	43	152	73
1	36	171	57	18	237	81	247	136	98	9	195	97	228	17	235	165
2	224	218	193	232	177	229	147	227	35	71	46	54	216	58	238	53
3	23	175	139	75	151	33	129	163	252	248	96	61	225	254	68	40
4	21	250	176	78	253	4	200	183	162	66	145	188	243	28	166	64
5	255	22	133	161	39	55	197	191	143	173	104	63	206	83	233	50
6	220	106	154	205	146	181	24	67	25	90	48	111	239	77	101	226
7	74	164	102	204	44	14	87	217	236	91	168	158	120	65	122	119
8	142	10	76	244	189	37	222	207	56	246	174	84	214	60	230	42
9	79	182	221	126	1	2	198	38	245	180	251	116	88	89	134	5
A	231	112	190	69	201	0	72	215	31	167	234	113	209	199	109	186
B	196	95	178	86	52	20	132	128	41	7	13	156	202	3	123	212
C	12	213	160	223	51	93	70	203	242	110	15	125	118	30	80	184
D	187	47	115	208	45	121	210	6	194	108	144	117	138	85	211	148
E	141	192	107	103	124	172	11	241	219	130	159	185	26	170	149	27
F	127	34	99	153	135	140	179	114	59	157	94	92	8	249	32	150

The cryptographic features of the proposed substitution-box structure are given in Table 2.

Table 2. Cryptographic properties of proposed substitution box structure

s-box structure				
Nonlinearity Average	**Strict Avalanche Criterion Average**	**Bit Independence Creation**		**Input / Output XOR Distribution**
		BIC-SAC	**BIC-Nonlinearity**	**Max**
104.25	0.5044	0.502	103.93	10

5 Conclusions

In this study, a substitution box design has been performed which could be an alternative to chaos based substitution box structures in the literature. The innovative aspect of the proposed method is the chaotic system used in the design process. The most important feature that distinguishes this Nose–Hoover system from others is that it does not need any control parameters. The results of the analysis showed that a successful substitution box structure can be obtained.

Acknowledgments

This study is supported by the Firat University Scientific Research Project (TEKF.19.18).

References

1. C. Wu and D. Feng, Boolean Functions and Their Applications in Cryptography. Berlin, Germany: Springer, 2016.
2. M. S. Açikkapi, F. Özkaynak, and A. B. Özer, "Side-channel analysis of chaos-based substitution box structures", IEEE Access, vol. 7, pp. 79030–79043, 2019. doi: 10.1109/ACCESS.2019.2921708.
3. D. Lambić, ''S-box design method based on improved one-dimensional discrete chaotic map,'' J. Inf. Telecommun., vol. 2, no. 2, pp. 181–191, 2018.
4. H. A. Ahmed, M. F. Zolkipli, M. Ahmad, "A novel efficient substitution-box design based on firefly algorithm and discrete chaotic map", Neural Computing and Applications, vol. xx, no. x, pp. xx-yy, May 2018.
5. K. M. Ali and M. Khan, "Application based construction and optimization of substitution boxes over 2D mixed chaotic maps", Int. J. Theor. Phys., pp. 1–27, 2019. doi: 10.1007/s10773-019-04188-3.
6. A. H. Zahid and M. J. Arshad, "An innovative design of substitution-boxes using cubic polynomial mapping", Symmetry, vol. 11, no. 3, p. 437, 2019. doi: 10.3390/sym11030437.
7. M. F. Khan, A. Ahmed, K. Saleem, and T. Shah, "A novel design of cryptographic SP-network based on gold sequences and chaotic logistic tent system", IEEE Access, vol. 7, pp. 84980–84991, 2019. doi: 10.1109/ACCESS.2019.2925081..

8. F. Özkaynak, ‘‘From biometric data to cryptographic primitives: A new method for generation of substitution boxes,’’ in Proc. ACM Int. Conf. Biomed. Eng. Bioinformat., Bangkok, Thailand, Sep. 2017, pp. 27–33. doi: 10.1145/3143344.3143355.
9. L. Liu, Y. Zhang, and X. Wang, “A novel method for constructing the S- box based on spatiotemporal chaotic dynamics”, Appl. Sci., vol. 8, no. 12, p. 2650, 2018. doi: 10.3390/app8122650.
10. F. Özkaynak, “An analysis and generation toolbox for chaotic substitu- tion boxes: A case study based on chaotic labyrinth rene thomas system”, Iranian J. Sci. Technol.-Trans. Elect. Eng., pp. 1–10, 2019. doi: 10.1007/s40998-019-00230-6.
11. M. F. Khan, A. Ahmed, and K. Saleem, “A novel cryptographic substitution box design using Gaussian distribution”, IEEE Access, vol. 7, pp. 15999–16007, 2019. doi: 10.1109/ACCESS.2019.2893176.
12. T. Ye and L. Zhimao, “Chaotic S-box: Six-dimensional fractional Lorenz–Duffing chaotic system and O-shaped path scrambling”, Nonlinear Dyn. vol. 94, no. 3, pp. 2115–2126, 2018. doi: 10.1007/s11071-018- 4478-5.
13. A. H. Zahid, M. J. Arshad, M. Ahmad, (2019). “A Novel Construction of Efficient Substitution-Boxes Using Cubic Fractional Transformation”. Entropy, 21(3), 245.
14. F. Özkaynak, “Construction of robust substitution boxes based on chaotic systems”, Neural Comput. Appl., pp. 1–10, 2017. doi: 10.1007/s00521- 017-3287-y.
15. E. Tanyildizi, F. Özkaynak, A New Chaotic S-Box Generation Method Using Parameter Optimization of One Dimensional Chaotic Maps, IEEE Access, Volume 7, Page(s): 117829 - 117838, DOI: 10.1109/ACCESS.2019.2936447
16. G. Jakimoski, L. Kocarev, ‘‘Chaos and cryptography: Block encryption ciphers based on chaotic maps,’’ IEEE Trans. Circuits Syst. I. Fundam. Theory Appl., vol. 48, no. 2, pp. 163–169, Feb. 2001.
17. F. Özkaynak,, (2020) On the Effect of Chaotic System in Performance Characteristics of Chaos Based S-box Designs, Physica A: Statistical Mechanics and its Applications, Volume 550, 124072, https://doi.org/10.1016/j.physa.2019.124072.
18. G. Tang and X. Liao, ‘‘A method for designing dynamical S-boxes based on discretized chaotic map,’’ Chaos Solitons Fractals, vol. 23, no. 5, pp. 1901–1909, 2005.
19. Z. M. Z. Muhammad, F. Özkaynak, (2020) A Cryptographic Confusion Primitive Based on Lotka–Volterra Chaotic System and Its Practical Applications in Image Encryption, IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET), 25-29 February 2020, Slavske, Ukraine.
20. F. Özkaynak and A. B. Özer, “A method for designing strong S-boxes based on chaotic Lorenz system”, Phys. Lett. A, vol. 374, no. 36, pp. 3733–3738, 2010.
21. E. A. Solami, M. Ahmad, C. Volos, M. N. Doja, and M. M. S. Beg, ‘”Anew hyperchaotic system-based design for efficient bijective substitution- boxes”, Entropy, vol. 20, no. 7, p. 525, 2018. doi: 10.3390/e20070525.
22. F. Artuğer, F. Özkaynak, (2020) A Novel Method for Performance Improvement of Chaos-Based Substitution Boxes, Symmetry 2020, 12, 571, doi:10.3390/sym12040571.

23. A. Belazi, A. A. A. El-Latif, A.-V.Diaconu, R. Rhouma, and S. Belghith, “Chaos-based partial image encryption scheme based on linear fractional and lifting wavelet transforms”, Opt. Lasers Eng., vol. 88, pp. 37–50, Jan. 2017.
24. Ahmad, M. [2018] “Random search based efficient chaotic substitution box design for image encryption,” Int. J. Rough Sets Data Anal. 5(2), 131–147, doi: 10.4018/IJRSDA.2018040107.
25. Ahmed, H.A., Zolkipli, M. F., Ahmad, M. [2018] “A novel efficient substitution-box design based on firefly algorithm and discrete chaotic map”, Neural Comput. Appl. 3, 1-10.
26. Alzaidi, A.A., Ahmad, M., Doja, M.N., Solami, E.A., Beg, M.M.S.[2018] “A New 1D Chaotic Map and beta-Hill Climbing for Generating Substitution-Boxes”, IEEE Access 6, 55405-55418.
27. Alzaidi, A.A., Ahmad, M., Ahmed, H.S., AlSolami, E. [2018] “Sine-Cosine Optimization Based Bijective Substitution-boxes Construction Using Enhanced Dynamics of Chaotic Map”, Complexity, 2018.
28. Farah, M.A.B., Guesmi, R., Kachouri, A. et al. A new design of cryptosystem based on S-box and chaotic permutation. Multimed Tools Appl (2020). https://doi.org/10.1007/s11042-020-08718-8
29. S. S. Jamal, A. Anees, M. Ahmad, M. F. Khan and I. Hussain, "Construction of Cryptographic S-Boxes Based on Mobius Transformation and Chaotic Tent-Sine System," in IEEE Access, vol. 7, pp. 173273-173285, 2019.
30. M. A. Yousaf, H. Alolaiyan, M. Ahmad, M. Dilbar and A. Razaq, "Comparison of Pre and Post-Action of a Finite Abelian Group Over Certain Nonlinear Schemes," in IEEE Access, vol. 8, pp. 39781-39792, 2020.
31. R. Hao, X. Ma (2019) Dynamical Analysis of Nose-Hoover Continuous Chaotic System Based on Gingerbreadman Discrete Chaotic Sequence. In: Jin J., Li P., Fan L. (eds) Green Energy and Networking. GreeNets 2019. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol 282. Springer, Cham

Investigation the spin transport of a DNA chain on the effect of temperature gradient: a multifractal

Sohrab Behnia[1], Fatemeh Nemati[2], and Samira Fathizadeh[3]

[1] Department of Physics, Faculty of Science, Urmia University of Technology, Urmia, Iran
(E-mail: s.behnia@sci.uut.ac.ir)
[2] Department of Physics, Faculty of Science, Urmia University of Technology, Urmia, Iran
(E-mail: fateme.nemti@sci.uut.ac.ir)
[3] Department of Physics, Faculty of Science, Urmia University of Technology, Urmia, Iran
(E-mail: s.fathizadeh@sci.uut.ac.ir)

Abstract. Creating the spin current due to temperature gradient has attended so attention. Researchers are interested in spintronic field since data can be processed and transferred safely. Spin-caloritronic phenomena as, an emerging field, combine the electron's spin and heat in materials.There are severe restrictions through using inorganic materials such as high cost and the lack of materials. Because of those limitations, we choose biological material (especially DNA) with super features such as flexibility, low cost, and adjustable conductance to investigate the spin transfer in an appropriate temperature gradient. The extended PBH model, with spin degree of freedom is chosen for this study. The Nose-Hoover thermostat is used for applying the temperature gradient. We used chaos approach to analyze the system. At first, with simultaneously varying of the biased voltage and magnetic field, we found at E=0.6(mv) and B=0.05, 0.25, 0.4(T), there are maximum spin current. Secondly, we have simultaneously varied the biased voltage and magnetic field and we found the maximum spin current at E=0.6(mv) and T=305, 310, 335 (K). We studied the system using multi-fractal analysis, and Re′nyi dimension. So off and on states were rcognized. This gives a helpful device to recognize the reactions of organic polymer to external factors acting on the system.

Keywords: Spin Switch, Chaos, DNA, Temperature Gradient

1 Introduction

Among the technologies involved in heat transfer, thermoelectric transformation technology (Tes) is an applied technology, that has several advantages, including small size, high reliability, long life [1]. On the other hand, thermoelectric devices can directly convert thermal energy into electrical energy. Semi-conductors, ceramics, and polymers are the most utilized materials in the manufacturing of Tes [2]. Therefore, the transport characteristics of molecular systems can be investigated with thermoelectric devices [3]. The combination of spintronics and thermoelectricity is known as spin-caloritronic phenomena [4, 5]. This phenomenon is utilized to investigate the cooperation of charge, spin, and heat transport. In addition, the use of DNA spin polarization to

transmit information is one of the most interesting issues in information theory[6].

Inorganic materials are widely used in the manufacture of spintronics devices. There are serious restrictions through using inorganic materials such as high cost and the lack of materials. Because of those limitations, we choose biological material (especially DNA) with super features such as flexibility, low cost, and adjustable conductance to investigate the spin transfer in an appearance temperature gradient. The spin polarization property of materials is one of the most important factors in choosing a material in spintronic phenomena. In this regard, a DNA molecule is used as biological material to investigate the DNA nano-switches. By regulating the gate voltage, switching the spin current between on and off states is possible. Spin switch effect was investigated on many inorganic materials such as graphene/MoS_2 [6], MnSi [7] and so on. DNA is a chemical chain consisting of 4 organic bases, and it is stored in the cells of a wide variety of organisms.

DNA is a long polymer consist of nucleotides, and any nucleotide is combined of three basic organic (cytosine, guanine, adenine, thymine), sugar, and phosphate groups [9]. Several theoretical models have illustrated the nonlinear dynamics of DNA, such as Peyrard-Bishop (PB), Peyrard-Bishop-Dauxois (PBD), and PBH (Peyrard-Bishop-Holstein) models [10]. We choose the generalized PBH model with spin degree of freedom because of their advantages such as spin and charge transfer capability. There are several ways to study the effects of the environment on the system, such as the Langevin thermostat, Nos′e-Hoover thermostat, etc. So to investigate the effect of heat on spin transfer, the Nos′e-Hoover thermostat was used as the heat source, Fig.1. Our purpose is creating a spin switch in the DNA chain. In order to find the conditions in which the system acts as a switch, we consider the multi-fractal spectrum. So we studied the Re′nyi dimension and analogous specific heat.

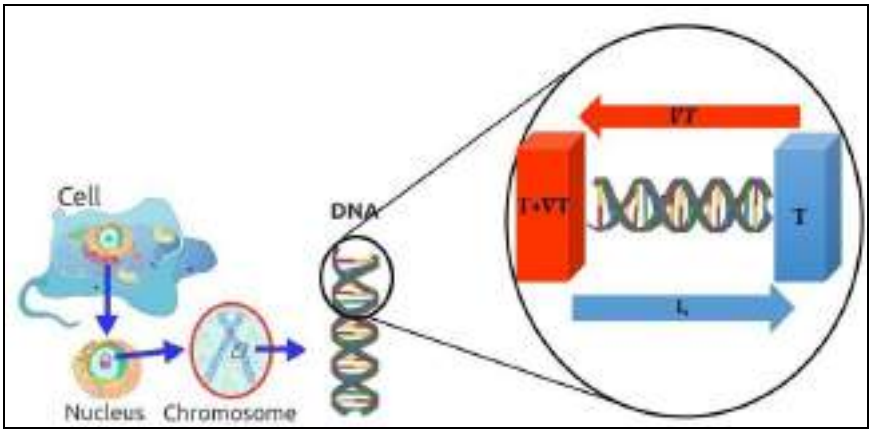


Fig.1. A schematic illustration of the model used to investigate the temperature driven spin transport.

2 Model and Method

We have chosen the N=40 base pairs DNA system (Table1). The Hamiltonian is written as follows [8]:

$$\mathrm{H}_{SYS} = H_{DNA} + H_{car} + H_{\mathrm{int}} + H_{So} + H_{field} \tag{1}$$

Where,

$$\mathrm{H}_{DNA} = \sum_{n} \frac{1}{2} m\dot{y}^2 + V(y_n) + W(y_n, y_{n+1}) \tag{2}$$

H_{DNA} is the DNA lattice Hamiltonian, where m is the mass of base pair, by using a semi-classical approach, $V(y_n) = D_n(e^{a_n y_n} - 1)^2$ and $W(y_n, y_{n+1}) = \frac{k}{2}\left(1 + \rho e^{-\alpha(y_n + y_{n+1})}\right)(y_n + y_{n+1})^2$ are Morse potential and stacking interaction of neighboring base pairs in the chain, respectively. D_n and a_n are depth and the width of the Morse potential. As well as, k is the coupling constant, ρ is the stiffness parameter, and α is the damping coefficient. Also,

$$\mathrm{H}_{car} = \sum_{n,\sigma=\uparrow\downarrow} \left[\varepsilon_n c_n^{\sigma+} c_n^{\sigma} - V_{n,n+1}(c_n^{\sigma+} c_n^{\sigma} + c_{n+1}^{\sigma+} c^{\sigma})\right] \tag{3}$$

$$\mathrm{H}_{\mathrm{int}} = \chi \sum_{n,\sigma=\uparrow\downarrow} y_n c_n^{\sigma+} c_n^{\sigma} \tag{4}$$

Here, $c_n^{\sigma+}$, and c_n^{σ} create or destroy an excitation at the tight-binding site n with $\sigma=\uparrow\downarrow$ and χ is the electron−lattice coupling constant.

Table 1. The DNA sequences [8]

Name	Number of base pairs	sequences
HC1	40	TAAATAAATAAATAAATAAAT AAAATAAATAAAAGCCTTT
CH22	60	AGGGCATCGCTAACGAGGTC GCCGTCCACAGCATCGCTAT CGAGGACACCACACCGTCCA

$$H_{SO} = \sum_n [2it_{SO}\cos(\theta)(c_n^{\uparrow+}c_{n+1}^{\uparrow} - c_n^{\uparrow+}c_{n-1}^{\uparrow} - c_n^{\downarrow+}c_{n+1}^{\downarrow} + c_n^{\downarrow+}c_{n-1}^{\downarrow})] + D_{n,n+1}c_n^{\uparrow+}c_{n+1}^{\downarrow} - D_{n,n+1}^{*}c_n^{\downarrow+}c_{n+1}^{\uparrow} + D_{n,n+1}^{*}c_n^{\downarrow+}c_{n-1}^{\uparrow} - D_{n-1,n}c_n^{\uparrow+}c_{n-1} \quad (5)$$

Where, $D_{n,n+1} = it_{so}\sin\theta\{\sin[n\Delta\varphi] + \sin[(n+1)\Delta\varphi] + i\cos[n\Delta\varphi] + i\cos[(n+1)\Delta\varphi]\}$, t_{so} is spin-orbit interaction constant and θ is the helix angle and $\varphi = n\Delta\varphi$ is the cylindrical coordinate with ϕ the twist angle. The preservation of time reversal symmetry leads to $D_{n,n-1} = D_{n-1,n}^{*}$. The formulation of Nos′e-Hoover thermostat is as follows [8]. The parameters that are used in the PBH model are shown in Table 2.

$$\dot{\xi} = \frac{1}{M}\sum_n m\dot{y}^2 - N\mathrm{K_B}T_0 \quad (6)$$

Here, ξ is the thermodynamics friction coefficient which interacts with the particles. T is the temperature preserved by heat source and M=1000 is the constant of Nos′e –Hoover. In the current study, we investigate the effect of external electrical and magnetic fields on spin transfer in DNA. So, $H_{field} = H_B + H_E$ [8]:

$$H_B = -\mu_B B\sum_n (c_n^{\uparrow+}c_n - c_n^{\downarrow+}c_{n)}$$
$$H_E = -e\sum_{n,\sigma} \mathrm{E}d\cos((n-1)\Delta\varphi)c_n^{\sigma+}c_n^{\sigma} \quad (7)$$

Due to the non-linear equations of the system, the chaos approach can be used to analyze the system. To obtain this approach, the evolution equation of Hamiltonian is extracted and it is necessary to convert the second-order differential equations to first-order equations, and then to extract the corresponding spin-up electric currents to study the information transfer by the spin current. In this regard, tools such as Reni dimension, etc. are used to analyze the behavior of the spin current of the system. So the Heisenberg approach is used to obtain the currents.

Table 2.PBH parameters[8]

Symbol	Unit	Value
m	amu	300
k	eV/A^2	0.04
ρ		0.5
b	A^{-1}	0.35
χ	eV/A	[0.1-0.6]
β_n	A^{-1}	[1-1.7] a_n

$$I^{\uparrow}(t)=\frac{-ie}{\hbar}\sum_{n}\begin{Bmatrix}W_{n,n+1}c_n^{t\uparrow}c_{n+1}^{\uparrow}+W_{n,n-1}^{*}c_{n-1}^{t\uparrow}c_n^{\uparrow}+\\ D_{n,n+1}c_n^{t\downarrow}c_{n+1}^{\downarrow}-D_{n,n-1}c_{n-1}^{t\downarrow}c_n^{\downarrow}\end{Bmatrix} \tag{8}$$

$$I^{\downarrow}(t)=\frac{-ie}{\hbar}\sum_{n}\begin{Bmatrix}W^{*}{}_{n,n+1}c_n^{t\downarrow}c_{n+1}^{\downarrow}+W_{n,n-1}c_{n-1}^{t\downarrow}c_n^{\downarrow}-\\ D^{*}{}_{n,n+1}c_n^{t\uparrow}c_{n+1}^{\uparrow}+D^{*}{}_{n,n-1}c_{n-1}^{t\uparrow}c_n^{\uparrow}\end{Bmatrix} \tag{9}$$

Here, $I^{\uparrow}$ and $I^{\downarrow}$ are spin-up and spin-down currents, respectively. Now, the net spin (I_s) and net charge (I_c) currents are introduced as following:

$$I_S=I^{\uparrow}-I^{\downarrow} \tag{10}$$
$$I_C=I^{\uparrow}+I^{\downarrow}$$

3 Result and Discussion

We have tried to obtain the different electrical responses of the system. Several strategies have been offered to describe strange attractors from a sincerely geometrical point of view. Based on concepts developed in the theory of fractal sets, we have simultaneously varied the biased voltage and magnetic field and studied the response of the system. Various islands have emerged that represent on and off states that are examined and confirmed by multifactorial analysis. The result gives a helpful tool to recognizing the reactions of organic polymer to external factors acting on the system. We have considered the Re′nyi dimension (D_q) [10] to analyze the system:

$$D_q=\frac{1}{q-1}\lim_{l\to 0}\frac{\ln\sum_{i}^{M(l)}p_i^q}{\ln l} \tag{11}$$

Where P_i is the probability that the trajectory on the strange attractor visits box *i*, and *M(l)* is the number of nonempty boxes. The multi-fractal analysis of spin current time series can show the multi-fractal behavior. To achieve a thermodynamic understanding of multi-fractality. We considered $\tau_q = (q-1)D_q$ wherein τ is analogous free energy. So, C_q is defined as follows:

$$C_q = -\frac{\partial^2 T}{\partial q^2} \approx C(q+1) - 2\tau(q) + \tau(q-1) \tag{12}$$

As shown in Fig.2, there are some islands at parameters surface with some different currents. The maximum spin current can be seen at E=0.6 (mV) and B=0.05, 0.25, 0.4 (T). In this way, on and off states are determined.

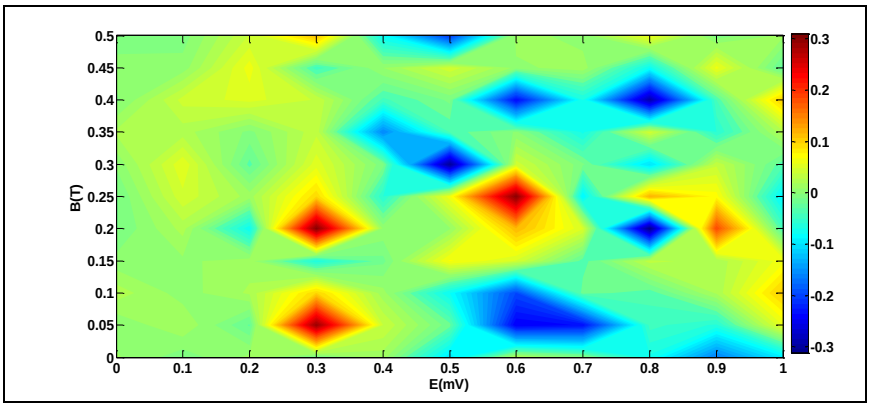


Fig.2. Alteration of spin current due to the concurrent effects of the magnetic field and bias voltage.

Secondly, by simultaneously varying the electrical field and temperature gradient, different behavior of spin current is shown. At E=0.5 (mV), T=305, 310, 335 (K) maximum spin currents are shown, so we can say that on and off states are determined.

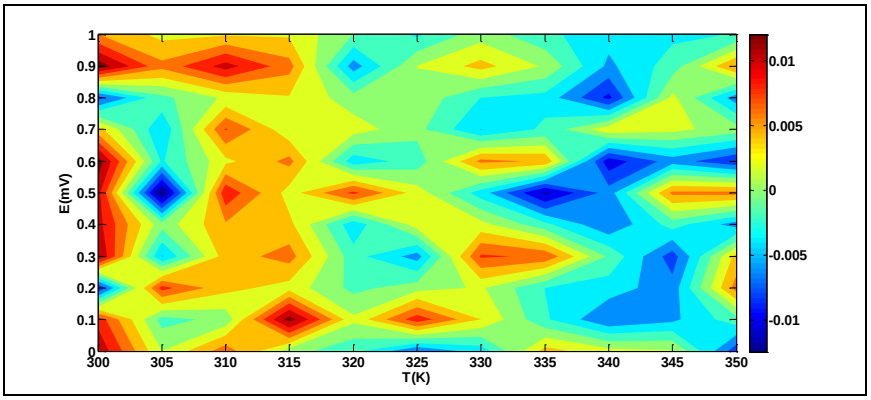


Fig.3. Alteration of spin current due to the concurrent effects of temperature and bias voltage.

The multi-fractal analysis of spin current time series at different magnetic fields shows the multi-fractal behavior. As shown in Fig.4, there is maximum and minimum spin current for B=0.25, 0.05, 0.4 (T) and B=0.15, 0.45 (T), respectively. It means that the region with maximum spin current (ON state) is distinguished from the zero spin current regions (OFF state). Similarly, we studied the Re′nyi dimension spectrum of system for which, temperature gradient is varied. As shown in Fig.5, there is maximum and minimum spin current at T=310, 305, 320 (K) and T=325, 330 (K), respectively. It is clear that the region with the maximum spin current represents the on state, and the region with the minimum spin current represents the off state. So, the spin switch can be designed.

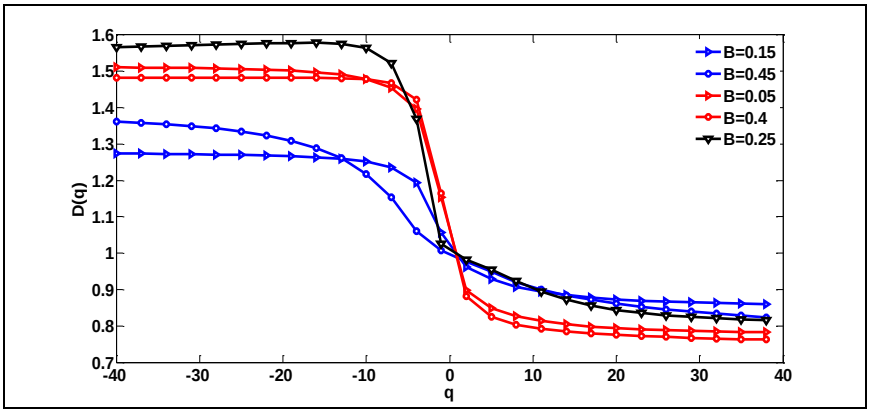


Fig.4. The Re′nyi dimension spectrum at different quantities of the magnetic field at E=0.6 (mV) and T=335(K)

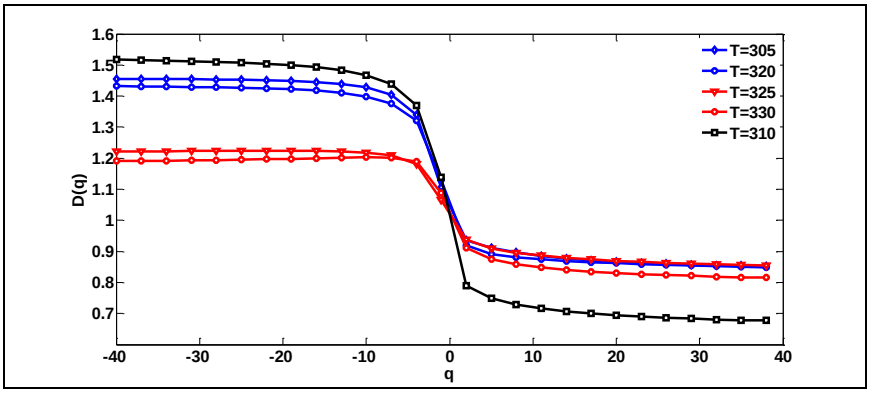


Fig.5. The Re′nyi dimension spectrum at different quantities of the temperature gradient at E=0.5 (mV) and B=0.05(T)

Furthermore, the analogous specific heat C_q is investigated. In this paper, we studied C_q of the magnetic field and temperature. As shown in Figs 6 and 7, there is one single peak for B=0.25 (T) and T=310 (K) wherein, the maximum spin current is observed. The obtained results indicate that higher dimension spectra are relative to the maximum positive currents. This prepares a beneficial gadget to identify the notable responses of the biological polymer.

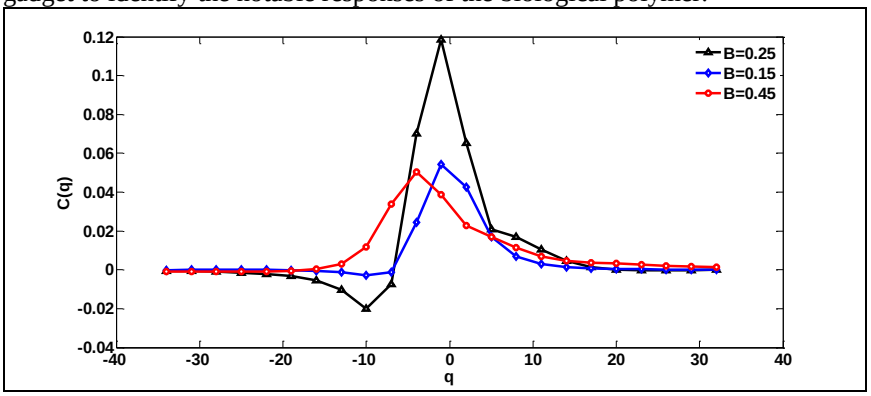


Fig.6. Analogous of specific heat for various values of the magnetic field at E=0.6 (mV) and T=335(K)

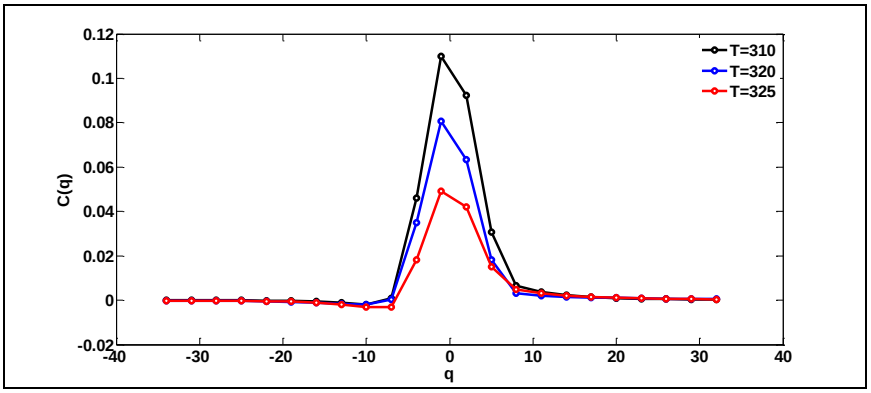


Fig.7. Analogous of specific heat for various values of the temperature gradient at E=0.5 (mV) and B=0.05(T)

Conclusions

We have theoretically studied the various factors to design a spin switch based on the HC1 sequence. Then, we have changed the magnetic field bias voltage and also the temperature and bias voltage at the same time. The on/off behaviors have been shown. We have used the multifractal analysis to understand the switch behavior better. Besides, the on/off behavior of the system can be distinguished via the Re′nyi dimension spectrum. Moreover, The analogous of specific heat can endorsement the results.

References

1. Z. Miao, X. Meng, S. Zhou and M. Zhu. Thermo-mechanical analysis on thermoelectric legs arrangement of thermoelectric modules. Renewable Energy, 147, 2272-2278, 2020.
2. S. M. Pourkiaei, M. H. Ahmadi, M. Sadeghzadeh, S. Moosavi, F. Pourfayaz, L. Chen and R. Kumar. Thermoelectric cooler and thermoelectric generator devices: A review of present and potential applications, modeling and materials. Energy, 2019.
3. S. R. Boona, R. C. Myers, and J. P. Heremans. Spin caloritronics. Energy Environ. Sci. 7, 885–910 2014.
4. R. Korol, M. Kilgour and D. Segal. Thermopower of molecular junctions: Tunneling to hopping crossover in DNA. The Journal of chemical physics, 145(22), 224702, 2016.
5. J. Wang,Y. K. Takahashi and K. I. Uchida. Magneto-optical painting of heat current. Nature Communications, 11(1), 1-7, 2020.
6. Yan, W., Txoperena, O., Llopis, R., Dery, H., Hueso, L. E., & Casanova, F. (2016). A two-dimensional spin field-effect switch. Nature communications, 7, 13372.
7. Neubauer A, Pfleiderer C, Binz B, Rosch A, Ritz R, Niklowitz PG, Bo¨ni P. Topological Hall effect in the a phase of MnSi. Phys Rev Lett. 2009;102(18):186602.

8. S. Behnia, F. Nemati, and S. Fathizadeh. Modulation of spin transport in DNA-based nanodevices by temperature gradient: A spin caloritronics approach. Chaos, Solitons & Fractals, 116, 8-13, 2018.
9. J. B. Okaly, A. Mvogo, R. L. Woulaché AND T. C. Kofané. Nonlinear dynamics of damped DNA systems with long-range interactions. Communications in Nonlinear Science and Numerical Simulation, 55, 183-193, 2018.
10. K. Pawelzik, H. G. Schuster. Erratum: Generalized dimensions and entropies from a measured time series. Physical Review A, 35(1)481, 1987.
11. E. Canessa. Multifractality in time series, Journal of Physics A: Mathematical and Genera, 33.(19), 3637, 2000.

Selective transport of suspending micro-particles in an oscillating fluid through micro-channels.

Philippe Beltrame[1]

UMR1114 EMMAH INRAE - Avignon Université, Avignon, France
(E-mail: philippe.beltrame@univ-avignon.fr)

Abstract. This paper analyzes the possibility to obtain selective transport of microparticles depending their size. The particles are suspended in a fluid confined in modulated channels and a periodic pumping moves back and forth the fluid without net displacement. Using numerical simulation and bifurcation analysis tools, we show the existence of particle drift under the Stokes assumption of the fluid flow. For specific parameter ranges, the particle transport can be selective. The transport solution and the selectivity are related to (de)synchornization transitions in forced non-linear oscillators. We reveal that chaotic transitions are a key factor to drop from a bounded dynamics to a net transport. This transport phenomenon can be relevant for heavy particles in suspended in the air in microgravity environnement.
Keywords: Particle transport, synchronization, Chaotic dynamics, bifurcation analysis, microfluidic.

1 Introduction

Sorting suspended particles in a fluid is an issue in many domains such as the food industry, medical analyses or wastewater treatment. Many processes are based on the microfiltration using a membrane. However, at high permeation rates, this method suffers from the accumulation of non-permeating particles above the membrane surface, thereby blocking the pores [1]. In recent decades, alternative techniques using flow in a periodic and asymmetric structure of micro-channels have been developed. In these systems, the particles are driven mainly by the viscous force. Micro-particles may drift from the streamline mainly due to the lift force. In a confined geometry, the lift effect is strongly dependent on the particle size and induces a selective trajectory. Based on this principle, passive micro-fluidic devices to sort the particles have been developed such as branching channels, pinched flow fractionation, spiral channels or media with a periodic pattern of micro-posts [2–6]. The particle sorting is done using a continuous flow.
Such a method does not apply to the removal of specific particles from a basin. Particle selection using an oscillating flow, i.e. without net displacement of the fluid, is possible as shown in [7]. The particle transport occurs through a periodic structure of triangular columns. For a range of particle sizes, the lift force acts asymmetrically during the back and forth cycle of the fluid flow. This results in a drift that is orthogonal to the oscillating flow for a parameter range. Therefore in all systems the lift force is the key phenomenon of particle

drift and sorting.

In contrast, the present paper focuses on longitudinal transport, i.e. the drift takes place along the axis of fluid oscillations. Indeed few studies are devoted to the onset of transport and to the kind of transition. These questions are useful for the design optimization of these devices.
We consider a micro-device similar to [8] where two basins are connected via modulated channels filled with a liquid. A periodic pumping confers a back and forth fluid motion dragging the particles in suspension. For oscillating Stokes flow, the 1D transport of particles is usually explained by the Stokes drift [9]. The particle follows the flow drift due to the traveling wave. In our context, the fluid is, however, driven by a standing wave which does not lead to fluid drift. The flow therefore needs to be ratchet like. In the early 2000s, the transport of overdamped particles in ratchets in many fields in physics was interpreted as a Brownian motor in which transport results from the action of noise in an asymmetric potential [10,11]. Such a drift ratchet phenomenon may occur in the microfluidic context considered here [12] and the experiment in [8] corroborated this theory. Nevertheless, further experiments revealed that the thermal fluctuations are negligible and the experiment in [8] does not evidence transport due to a Brownian ratchet.
Recently, we highlighted different 1D transport mechanisms in a Stokes flow, called *ratchet flow*, for a simple model of inertial particle [15]. We showed that the spatial variations of the fluid flow induce a 'ratchet effect'. For instance, for moderate damping chaotic dynamics are a key component of the transport. However, the parameter domains of the transport require that the particle radius is not negligible compared to the channel radius [13]. Therefore, the drag coefficient depends on the channel walls and hence on the particle position. Such a variation may induce a friction ratchet [14,10].

The goal of this paper was to determine whether such 1D transport mechanisms as in [15] exist for a particle radius comparable to the channel radius, and then to determine a possible dependence of the transport direction on the particle size. To answer these questions, we computed the friction for the 3D axisymmetric problem and we used bifurcation analysis tools and continuation of periodic orbit to provide a comprehensive overview of the dynamics in the phase space.

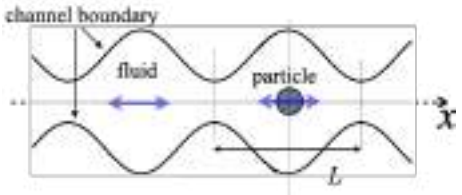


Fig. 1. Sketch of the problem: the particle translates along the x-axis of a periodic modulated channel. It is dragged by the periodic motion of a viscous fluid.

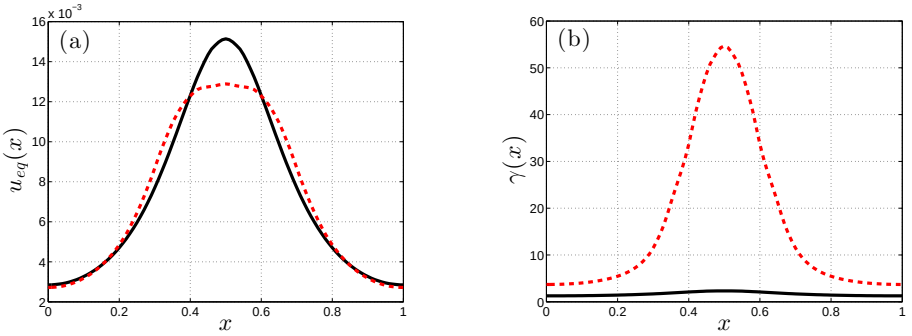


Fig. 2. Profile of (left panel) $\hat{u}_{eq}(x)$ and (right panel) $\hat{\gamma}(x)$ functions for two particle radii: (black line) $r_p = 0.05$ and (red dashed line) $r_p = 0.1$.
fig:profiles

2 Modeling

Let us consider a L-periodically modulated channel infinitely extended along the line (Ox) through which a Newtonian fluid with the viscosity μ is T-periodically pumped. We call 'cell' the channel portion of length L (Fig. 1). The cell is axisymmetric and its radius $r(x)$ varies sinusoidally:

$$R(x) = R_m \left(1 + c_r \cos\left(2\pi x\right)\right) \tag{1}$$

where R_m is the mean radius and c_r is the channel camber. In this study, we fix the channel geometry such as $R_m = 0.14 \cdot L$ and $c_r = 0.56$ which corresponds to the shape shown in Fig. 1. We assume that the flow is a quasi-static Stokes flow. Then, a periodic pumping implies a periodic fluid velocity field in space and time. Moreover, the time dependence of the velocity field is governed by the periodic pumping $\boldsymbol{v}_0(\boldsymbol{r}, t) = \boldsymbol{u}_0(\boldsymbol{r})A(t)$, where $A(t)$ is the pumping amplitude and $\boldsymbol{u}_0(\boldsymbol{r})$ is L-periodic. We note $[p]$ the amplitude of the pressure difference between the cell inlet and outlet. Pressure, length and time are scaled by $[p]$, L and T respectively. We consider spherical particle of mass m with the adimensional radius r_p. If, in addition, we assume that the particle moves only along the axis and the particle does not rotate then the particle position $x(t)$ is governed by the dimensionless ODE:

$$\ddot{x} + \mathcal{P}_\gamma \hat{\gamma}(x)\dot{x} = \mathcal{P}_\gamma \mathcal{P}_v \hat{\gamma}(x)\hat{u}_{eq}(x)A(t). \qquad eq : ode \tag{2}$$

In the latter equation, we have introduced two bifurcation dimensionless parameters:

$$\mathcal{P}_v = \frac{[p]T}{\mu} \tag{3a}$$

$$\mathcal{P}_\gamma = \frac{LT\mu}{m}. \tag{3b}$$

The field $\hat{\gamma}(x) > 0$ is the normalized drag coefficient of the particle. It depends on the channel boundary and on the particle size [17,13]. The field $\hat{u}_{eq}(x)$ is

related to the flow field velocity without particle and also to the particle radius r_p. Fig. 2 displays the fields $\hat{u}_{eq}(x)$ and $\hat{\gamma}(x)$ for a sinusoidal cell profile such as $c_r = 0.56$ $r_{min} = 0.14$. The computation of these coefficients is detailed in [16]. The particle size does not notably affect the veolcity field $\hat{u}_{eq}$ (Fig. 2a) whereas the friction is very sensitive to the particle radius (Fig. 2b). The friction is maximal in the narrow region of the cell and minimal in the larger region. For $r_m = 0.1$, the ratio between the maximum and minimum value is large in contrast to $r_m = 0.05$, for which $\hat{\gamma}(x)$ is almost constant. In the latter case, Eq. 2 is similar to the ratchet flow model studied in [15,23,24] and we expect similar transport dynamics. If $r_p = 0.1$, the large friction contrast may induce, in addition, a friction ratchet [10].

Note that the functions $u_{eq}(x)$ and $\gamma(x)$ are 1-periodic and have the parity-symmetry like the geometry of the problem. If, in addition, the pumping $A(t)$ varies sinusoidally, then the problem is invariant by the parity symmetry $x \to -x$. More precisely, if $x(t)$ is a trajectory given by Eq. 2 then $-x(t+1/2)$ is also the solution for a symmetric initial condition. In order to break of the parity-symmetry, the back and forth phases of the pumping should be different, i.e. this means that $A(t+1/2) \neq -A(t)$. Let us introduce the parameter α such as $0 \leq \alpha < 1$ and define the function $A(t)$:

$$A(t) = \begin{vmatrix} 1-\alpha & , \text{ if } 0 \leq (t \text{ modulo } 1) < \alpha \\ \cos\left(2\pi \frac{t-\alpha}{1-\alpha}\right) - \alpha & , \text{ if } \alpha \leq (t \text{ modulo } 1) < 1 \end{vmatrix} \quad (4)$$

If $\alpha = 0$ is zero, $A(t) = \cos(2\pi t)$ and the problem is symmetric. Otherwise, the pressure difference is constant during the first step in the interval $[t_0, t_0 + \alpha[$ followed by a sinusoidal pumping in the interval $[t_0 + \alpha, t_0 + 1[$. In this case $A(t+1/2) \neq -A(t)$. Note that the mean value of the pumping is still zero if $\alpha \neq 0$.

In this paper, we consider two particle sizes: $r_p = 0.05$ or $r_p = 0.1$. The transport dynamics are explored in the parameter space using the time integration and the path-following method of the periodic solutions. The bifurcation parameters are $\mathcal{P}_\gamma, \mathcal{P}_v$ and α. Branches of T-periodic solution on bifurcation diagrams are represented by the norm $||.||$ such as: $||s|| = [\frac{1}{T}\int_0^T (\dot{x}(t))^2 dt]^{1/2}$.

3 Drift at large drag $\mathcal{P}_\gamma$

In this section we analyze the transport solutions when the drag $\mathcal{P}_\gamma$ is large. According to [15], if $\mathcal{P}_v \ll 1$ we can prove that there are two periodic solutions of the particle motion which are centered at the extrema of the velocity field u_{eq} noted s_0 for the maximum and s_m for the minimum. In the symmetric case ($\alpha = 0$), by increasing $\mathcal{P}_v$, we find out a third solution noted s_a bifurcated from either s_0 or s_m via a spontenaeous symmetry breaking. Fig. 3 shows the time evolution of the three solutions s_0, s_m and s_a. As in [15], one of the three solutions are stable and attracts all of the dynamics if $\alpha = 0$. Therefore, the transport solutions can occur only if $\alpha \neq 0$.

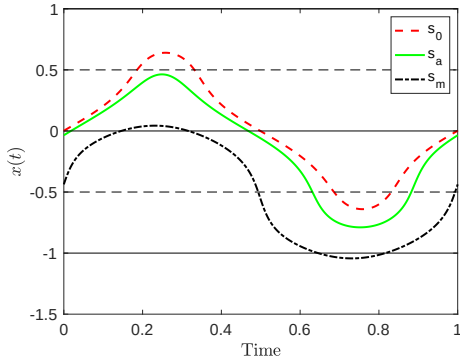


Fig. 3. Time evolution of the periodic solutions s_0, s_m and s_a for $r_p = 0.05$, $\alpha = 0, \mathcal{P}_\gamma = 79$ and $\mathcal{P}_v = 676$.

fig:s3xt

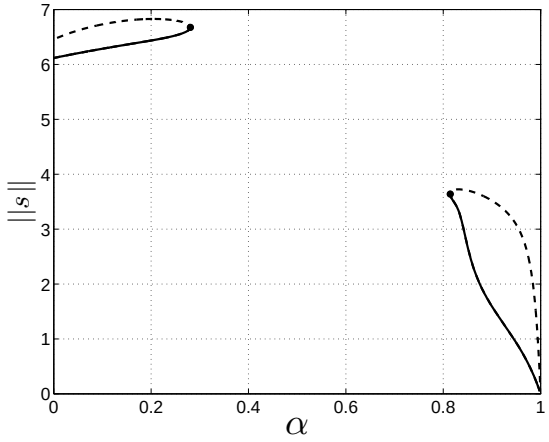


Fig. 4. Continuation of 1-periodic solutions for $r_p = 0.05$, $\mathcal{P}_v = 1350, \mathcal{P}_\gamma = 79$ by varying α. Plain [dashed] line indicates stable [unstable] orbit. Black dots indicate pitchfork bifurcations.

3.1 Transitions to transport solutions

We follow the periodic branches of solutions s_0 and s_m by varying α and we fix $\mathcal{P}_v = 1350$ (fig. 4). The branches s_0 and s_m annihilate in a saddle-node bifurcation for $\alpha \simeq 0.2807$. A similar scenario arises starting from $\alpha = 1$ for which the pumping is zero: a pair of saddle orbits annihilate in a saddle-node for $\alpha \simeq 0.8143$. In the large range $[0.3; 0.8]$, no solution is found. The saddle-node bifurcations correspond to the intermittent bifurcation type-I [19] as explained in [15]. The stroboscopic time evolution of the particle position at every period (fig. 5) displays a regular descending staircase for different values of α. The plateaux correspond to oscillations close to the threshold. The plateaux become longer when α approaches the onset of bifurcations. Such a dynamics is similar to the phase slip of a desynchronisation transition [15]. A well-known consequence is that the drift velocity vanishes as the square root

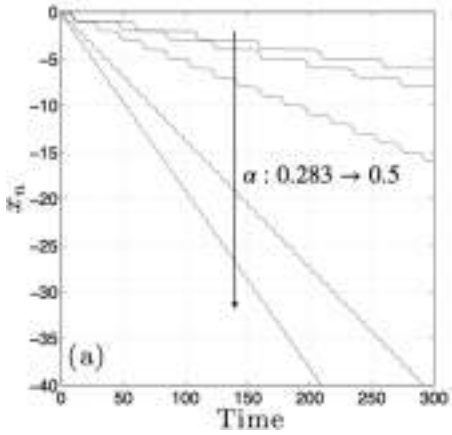


Fig. 5. Discrete particle positions $x_n = x(n)$, $n \in \mathbb{N}$ near the onset. Different values of $\alpha = 0.283; 0.285; 0.3; 0.4; 0.5$, other parameters as in fig. 4.
fig:drift

of the threshold distance: $c \propto \sqrt{\alpha - \alpha_c}$,where α_c is the threshold value [22]. According to the time integration, the drift velocity increases with α till $\alpha \simeq 0.5$ and remains almost constant till 0.6. By further increasing α the velocity decreases to zero when α approaches the critical value of the second saddle node. Consequently, the optimal transport is about $\alpha = 0.5$. In the next section, we fix α at 0.5 and we seek the parameter domains of particle transport. Note that according to the discussion in [15], the particle drift phenomenon is part of a class of dissipative rocking ratchets for which the transport direction is determined by the asymmetry [20,21]. In the current problem, it means that the direction of transport depends only on the sign of the α parameter, i.e. the kind of pumping. The particle drifts to negative values of x if $\alpha > 0$, otherwise the particle drifts to positive values.

3.2 Domain of intermittent drift

We explore the transport domain by varying the parameters $\mathcal{P}_v$ and $\mathcal{P}_\gamma$. We trace the saddle-node loci of the periodic solutions, in the $(\mathcal{P}_\gamma, \mathcal{P}_v)$ plane (Fig. 6) which represents the possible onset of transport. By varying $\mathcal{P}_\gamma$ the two saddle-nodes form a vertical band which ends at a minimal value of $\mathcal{P}_\gamma$ except if $\mathcal{P}_v$ is about 1000 (see Fig. 6). The transport arises in the region outside these bands and when the bands do not overlap (gray region in Fig. 6). Therefore, the transport domain is roughly a sector in the $(\mathcal{P}_v, \mathcal{P}_\gamma)$ plane. Then, the transport occurs if $\mathcal{P}_v$ is large enough, in other words, if the pumping amplitude is large. There is additional tapered vertical spaces for specific values of $\mathcal{P}_v$ for which transport may occur for large $\mathcal{P}_\gamma$ values. The specific $\mathcal{P}_v$ values are only slightly affected by the particle size: the tapered region occurs for $\mathcal{P}_v$ about 1500, 2200 and 2900 regardless of r_p.
In general, the domains of intermittent drift are qualitatively similar by varying r_p. However, the existence domain differs quantitatively, which allows to find

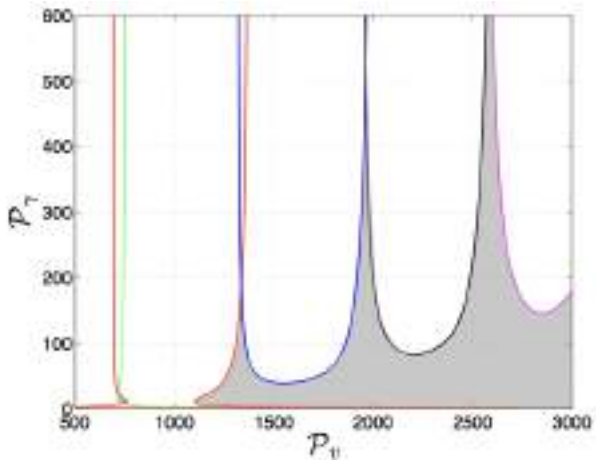


Fig. 6. Saddle-node loci of 1-periodic solutions in the $(\mathcal{P}_\gamma, \mathcal{P}_v)$ plane for $r_p = 0.1$. The gray region displays the domain of intermittent drift.
fig:domain

specific parameters for which the drift arises while for the other particle size the dynamics is still periodic and so bounded. Therefore, the transport only occurs for specific particle sizes.

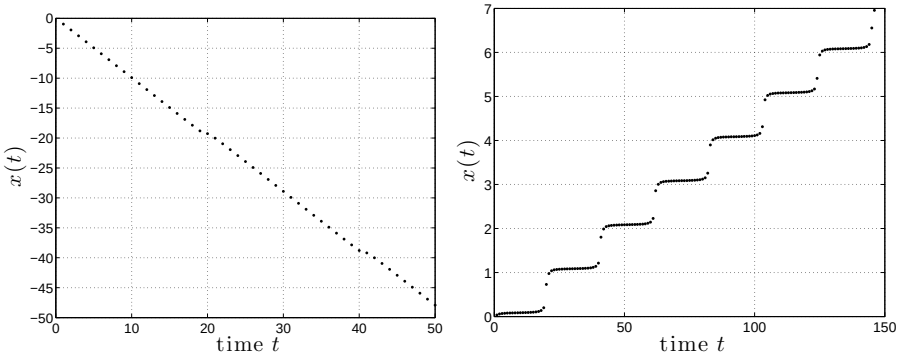


Fig. 7. [left] Discrete particle trajectory x_n near the onset of the synchronized transport $c = -1$. [right] The same discrete particle trajectory represented in the comoving frame $x_n + n$. Parameters are: $r_p = 0.1, \alpha = 0.5, \mathcal{P}_\gamma = 24.2, \mathcal{P}_v = 2250$.
fig:intertosynchro

4 Transport for moderate $\mathcal{P}_\gamma$

In this section, we consider smaller values of $\mathcal{P}_\gamma$. Thus, the particle damping is smaller implying a larger desynchronization between fluid movement and the

particle trajectory. This leads to more complex dynamics especially chaotic dynamics.

4.1 Synchronized transport

sec:sync

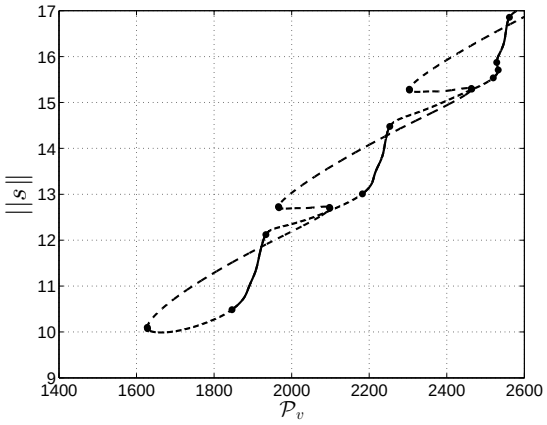


Fig. 8. Continuation of synchronized transport $c = -1$ with $\mathcal{P}_\gamma = 7.46$ and $r_p = 0.1$. Plain [dashed] lines indicate stable [unstable] solutions. Black dots indicate period-doubling bifurcations.

fig:csync

We study the behavior of the intermittent drift when $\mathcal{P}_\gamma$ is decreasing. We observe that the drift velocity c increases and the discrete dynamics presents shorter and shorter plateaus as in Fig. 5. According to Fig. 7, the plateaus may disappear and the transport occurs almost with a constant velocity. The representation of the dynamics in the co-moving frame ($c = -1$) points out an intermittent and a regular behavior. Indeed, the dynamics in the co-moving frame is due to the phase slip at a synchronization transition for weakly non-linear oscillator. It is a similar scenario as described in Section 3 but in the co-moving frame. The threshold is a saddle-node from which a pair of periodic solutions emerge. In the laboratory frame: these periodic solutions correspond to a synchronized transport solution: after a entire number n of time periods the particle moves by an entire number of spatial periods m, then the velocity is a rational $c = m/n$. This result is typical of phase locking (here is the c velocity) of a forced non-linear oscillator [22]. In the example of Fig. 7, we have $c = -1$, but we found also other velocities for other parameters: $c = -1/2$ and $c = -2$.

Another difference with the intermittent drift is that the synchronized transport can be not an attractor especially if $\mathcal{P}_\gamma$ is not large. For instance, for $\mathcal{P}_\gamma = 7.46$, we plotted the bifurcation diagram of the synchronized transport by varying the $\mathcal{P}_v$ parameter (Fig. 8). The transport emerges via a saddle-node bifurcation related to a synchronization phenomenon at about $\mathcal{P}_v = 1628$.

From this saddle, two branches display zigzags with further saddle-nodes and period-doubling bifurcations. Due to the period doubling, the branch changes its stability. According to Fig. 8, there exist $\mathcal{P}_v$ ranges for which the transport $c = -1$ is a stable solution. From the period-doubling bifurcations a cascade of period-doubling bifurcations occurs leading to a chaotic dynamics. This dynamics being bounded in the co-moving frame the transport velocity remains equal to $c = -1$. The scenario is similar to the one described in [15] in this context and it involves universal results of (de)synchronization of periodic oscillators by periodic external action [22]. As a result, in the range delimited by period-doubling, there may exist chaotic transport solutions without locked velocity. The transport velocity is then lower than in the synchronized transport case but the transport does not vanish.

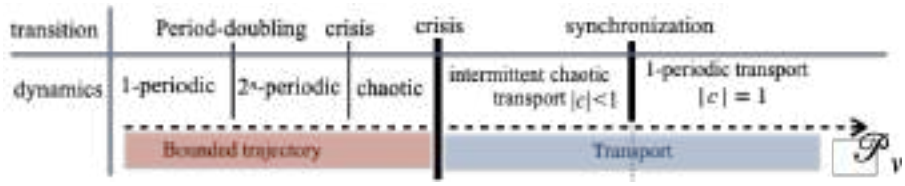


Fig. 9. Route to synchronized transport: dynamics and transitions by increasing $\mathcal{P}_v$ for moderate $\mathcal{P}_\gamma$.

fig:routetosync

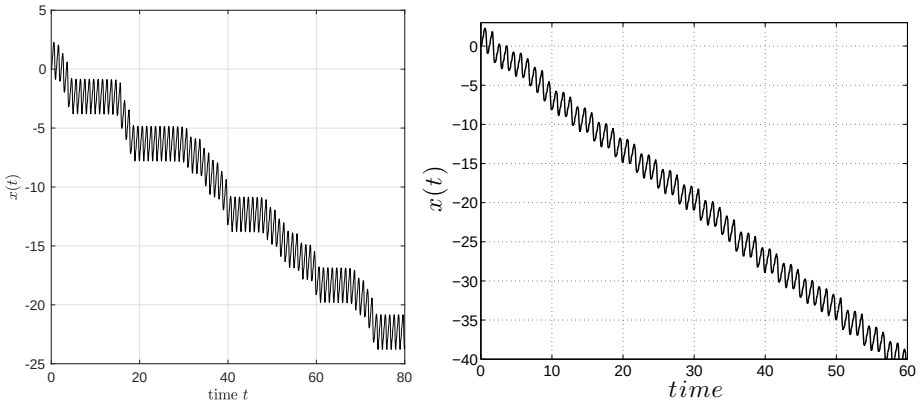


Fig. 10. Chaotic dynamics $r_p = 0.1, \mathcal{P}_\gamma = 7.46, \alpha = 0.5$ for two different values of $\mathcal{P}_v$. (a) near the crisis of the unbounded dynamics ($\mathcal{P}_v = 1458.3$). (b) near the syncrhonization transition ($\mathcal{P}_v = 1467.5$).

fig:transitions

4.2 chaotic transitions

The synchronization transition at the saddle-node of Fig. 8 ($\mathcal{P}_v = 1628, \mathcal{P}_\gamma = 7.46$) differs from the synchronization at $\mathcal{P}_\gamma = 24.1$. The dynamics is no longer quasi-periodic but chaotic. In [15], we found intermittency occuring at chaotic time intervals: this behavior is typical of large forcing of an oscillator [22]. However, the particle dynamics displayed in Fig. 10b does not corroborate the existence of intermittency. We guess that the coexistence of attractors hides the intermittency.
Now, we detail the transitions from the periodic solutions to the synchronized transport by increasing $\mathcal{P}_v$ when $\mathcal{P}_\gamma = 7.46$. Indeed, we retrieve all the bifurcations scenario explained in detail in [15]. The route to synchronized transport is sketched in Fig. 9. The periodic solutions s_0 or s_m (Fig. 2), by increasing $\mathcal{P}_v$, involve a period-doubling cascade leading to a chaotic unbounded dynamics. A merging crisis may appears and because of the spatial periodicity the strange attractor is no longer bounded. If the problem does not have the parity symmetry ($\alpha \neq 0$), we expect a preferential direction. In Fig. 10a, the dynamics of the particle near the onset displays a drift to negative values in a intermittent manner. However, the intermittency is not quasi-periodic but chaotic. By increasing further $\mathcal{P}_v$, the drift velocity increases and the dynamics is still chaotic (Fig. 10b).

4.3 Chaotic drift

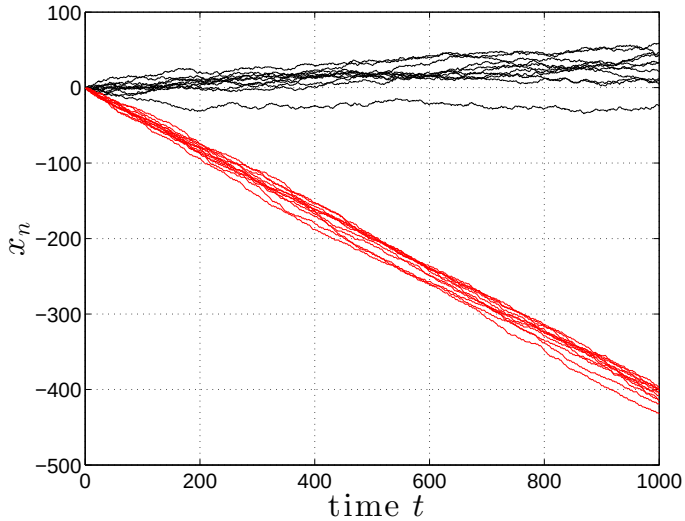


Fig. 11. Stroboscopic particles positions x_n at entire times for radii $r_p = 0.05$ (black lines) and $r_p = 0.1$ (red lines) for eleven initial conditions $(x_0, v_0) = (0, i/10)$ with $i = 0, 1, \ldots, 10$. Other parameters are $\alpha = 0.1, \mathcal{P}_\gamma = 6.7, \mathcal{P}_v = 2000$.
fig:tchaos

The previous transport solutions are obtained for $\alpha = 0.5$ and we find only transport to negative direction. A non-zero value of α is required for the existence of intermittent and quasi-periodic drift. However, if $\mathcal{P}_\gamma$ is not large, unbounded dynamics and also synchronized transport may exist for $\alpha = 0$. The scenario from periodic solution to the unbounded dynamics in Fig. 9 involves, this time, spontaneous symmetry breaking. The chaotic dynamics has diffusion like behavior: no prefential direction is observed. When there is a synchronized transport then the synchronized transport solution in the opposite direction exists too. To have an effective transport, we need to slightly breaks the parity symmetry. As explained in [23], the transport solutions still exist but for slightly different parameter ranges when α is small. Therefore, we expect there is parameter domains for which the transport direction depends on the particle size.
We found for $\alpha = 0.1, \mathcal{P}_\gamma = 6.7$ and $\mathcal{P}_v = 2000$ different drift directions depending on the particle size: a net drift appears for $r_p = 0.1$ and while, for $r_p = 0.05$, a slight drift to positive direction occurs (Fig. 11). The dynamics is chaotic and notably it depends on the initial conditions. For $r_p = 0.05$, the mean value over the initial conditions of the velocity transport is slightly positive ($c \simeq 0.04$). In addition, there is a diffusive-like behavior too: the trajectory deviation increases with the time. Indeed, the chaotic dynamics results from the competition between opposite transport solutions $c = \pm 1$ which are unstable. For $r_p = 0.1$ the diffuse behavior is weak and the mean velocity remains close to $c = -2/5$. A possible explanation is that we are in the vicinity of the onset of the synchronized transport $c = -2/5$. The multiplicity of solutions makes it difficult to find this synchronized transport.

5 Concluding remarks

We have shown using time-integration and bifurcation analysis that a selective transport of micro-particles is possible depending on the particle size.
For large drag $\mathcal{P}_\gamma$, the possible slow drift is quasi-periodic and determined by the pumping asymmetry, i.e. the parameter α. Then, the drift requires a non zero value of α typically a value about 0.5. By decreasing the drag $\mathcal{P}_\gamma$ we found out synchronized transport solutions. The velocity is then locked to a rational value. If α is about 0.5, the transport velocity is negative. Therefore, for α about 0.5, we found either bounded periodic solutions or transport in the negative direction. However, if the selectivity is possible, we do not find a change of transport direction depending on the particle size.
Such a property was found for moderate value of $\mathcal{P}_\gamma < 10$ and α about 0.1. Indeed, for this parameters values, the asymmetry of the problem is no longer required. We retrieved the scenarios of the ratchet problem of a point-like particle. In particular, the dynamics can be chaotic. We found parameters such as the direction of the parameter drift depends on the particle size.
This theoretical study may have application of heavy particles in the air in micro-gravity environment. Indeed, the small value of $\mathcal{P}_\gamma$ requires a low density fluid and the gravity could break the phenomena transport. Because of the

chaotic dynamics could be strongly influenced by the noise [24], a further work would be to study the influence of noise on the transport selectivity.

References

1. Kulrattanarak T., van der Sman R., Schroënn C. and Boom R. Advances in Colloid and Interface Science142200853 .
2. Pamme N. Lab Chip2007716441659.
3. Huang R., Barber T. A., Schmidt M. A., Tompkinsand R. G., Toner M., Bianchiand D. W., Kapur R. and Flejter W. L. PRENATAL DIAGNOSIS282008892.
4. Hasni A. E., Göbbels K., Thiebes A., Bräunig P., Mokwa W. and Schnakenberg U. Procedia Engineering2520111197 eurosensorsXXV.
5. Autebert J., Coudert B., Bidard F.-C., Pierga J.-Y., Descroix S., Malaquin L. and Viovy J.-L. Methods572012297 flow Cytometry and Cell Sorting: the Next Generation.
6. Lee G.-H., Kim S.-H., Ahn K., Lee S.-H. and Park J. Y. Journal of Micromechanics and Microengineering262016013003.
7. Loutherback K., Puchalla J., Austin R. H. and Sturm J. C. Phys. Rev. Lett.1022009045301.
8. Matthias S. and Müller F. Nature424200353.
9. Santamaria F., Boffetta G., Afonso M. M., Mazzino A., Onorato M. and Pugliese D. EPL (Europhysics Letters)102201314003.
10. Reimann P. Physics Reports361200257.
11. Hänggi P. and Marchesoni F. Rev. Mod. Phys.812009387.
12. Kettner C., Reimann P., Hänggi P. and Müller F. Phys. Rev. E612000312.
13. Makhoul M., Beltrame P. and Joelson M. International Journal of Mechanics92015260.
14. Luchsinger R. H. Phys. Rev. E622000272.
15. Beltrame P., Makhoul M. and Joelson M. Phys. Rev. E932016012208.
16. Makhoul M., Beltrame P. and Joelson M. Particle drag force in a periodic channel: wall effects in proc. of Topical Problems of Fluid Mechanics : Proceedings, Prague 2015 pp. 141–148.
17. Brenner H. Journal of Fluid Mechanics181964pp 144.
18. Makhoul M. Modélisation du transport de particule dans un écoulement de stokes à effet cliquet Ph.D. thesis Université d'Avignon (juillet 2016).
19. Pomeau Y. and Manneville P. Commun. Math. Phys.741980189.
20. Wickenbrock A., Cubero D., Wahab N. A. A., Phoonthong P. and Renzoni F. Phys. Rev. E842011021127.
21. Cuesta J. A., Quintero N. R. and Alvarez-Nodarse R. Phys. Rev. X32013041014.
22. Pikovsky A., Rosenblum M. and Kurths J. International Journal of Bifurcation and Chaos1020002291.
23. Beltrame P. The Foundations of Chaos Revisited: From Poincaré to Recent Advancements Understanding Complex Systems (Springer) 2016 Ch. Absolute Negative Mobility in a Ratchet Flow pp. 249–261.
24. Beltrame P. Journal of Chaotic Modeling and Simulation12018101.
25. Bhagat A. S., Kuntaegowdanahalli S. S. and Papautsky I. Microfluidics and Nanofluidics72009217.
26. Dijkshoorn J., Schutyser M., Wagterveld R., Schron C. and Boom R. Separation and Purification Technology173201786 .

27. G. D'Amico, J. Janssen and R. Manca. Homogeneous semi-Markov reliability models for credit risk management. *Decisions in Economics and Finance*, 28, 2, 79–93, 2005.
28. J. Janssen and R. Manca. *Semi-Markov risk models for finance, insurance and reliability*, Springer, New York, 2007.
29. J. Janssen and C.H. Skiadas. Dynamic modelling of Life-Table data, *Applied Stochastic Models and Data Analysis*, 11, **1**, 35–49, 1995.
30. G. Saporta. *Probabilités, Analyse des Données et Statistique*, Editions Technip, Paris, 1990.
31. C. H. Skiadas and C. Skiadas. *Chaotic Modeling and Simulation: Analysis of Chaotic Models, Attractors and Forms*, Taylor and Francis/CRC, London, 2009.

Chaotic Modeling and Simulation (CMSIM) : –, 2020

Application of New 4-D Chaotic Map for Secure IP-Communications

Belqassim Bouteghrine[1], Camel Tanougast[2], and Said Sadoudi[3]

[1] LCOMS, Université de Lorraine, 57070 Metz, France
(E-mail: belqassim.bouteghrine@univ-lorraine.fr)
[2] LCOMS, Université de Lorraine, 57070 Metz, France
(E-mail: Camel.Tanougast@univ-lorraine.fr)
[3] Ecole Militaire Polytechnique, 16000 Algiers, Algeria
(E-mail: sadoudi.said@gmail.com)

Abstract. Chaos systems have been studied for decades due to their applications in several domains such as: economy, communications, cryptography...etc. In recent years, designing and proposing new and higher dimensional chaotic systems become an increased tendency in particular for chaotic systems applied in network security domain.
In this paper, we propose a new 4-dimension chaotic map with four (04) control parameters and five (05) non-linear terms. Then, we investigate the chaotic behaviors of the proposed system by considering the bifurcation and the Lyapunov exponents (LE) theories. The proposed map is applied for generating cipher keys to perform data encryption and secure an Internet Protocol (IP) communication.
Keywords: Chaotic, Dimension, Bifurcation, Lyapunov exponents, Network security, Encryption, IP-communication.

1 Introduction

Chaos systems have been studied for decades due to their applications in several domains such as: electronic circuits [1],[2], network security [3],[4], encryption domain [5],[6] and in power control [7]. In one hand, chaotic systems' properties have been studied and investigated using the bifurcation and the Lyapunov exponents theories [8],[9]. In the other hand, among the existing chaotic systems, researchers have investigated for new and more complex systems. By combining two coexisting attractors, in [10] the evolution of a new 4-dimension chaotic system is presented and analyzed by using bifurcation diagrams and Lyapunov exponents' spectrum. In [11] authors proved that using only sine or cosine functions and modifying two (02) variables in the function, n-scroll attractors can be generated. Moreover looking for higher dimension, a new 4-D hyperchaotic continuous time system is introduced in [15] and its main specifications are analyzed by means of equilibrium points, stabilities and power

Received: / Accepted:

ISSN 2241-0503

spectrum. In [16] authors showed that, using 4-D Lorenz system for secure TCP communicationonl, consume a huge amount of ressources and computations.
Consequently, we present through this paper an optimized new 4-dimensional discrete time system for encryption purpose. Then, the chaotic behavior of the proposed map is investegated based on the bifurcation and Lyapunov exponent theories. The rest of this paper is structured as follows. The 4-D map is introduced and the chaotic behavior is investigated in section II. In Section III, software implementation of secure communication is illustrated using the proposed map. Finally, a conclusion is given in section IV.

2 The Proposed 4-D Map

2.1 System Description

The proposed 4-D map with (04) bifurcation parameters and (05) non-linear terms is given as follows:

$$\begin{aligned} X(n+1) &= 1 - a * X(n)^2 + (Y(n) * Z(n) * P(n)) \\ Y(n+1) &= 1 - b * Y(n)^2 + (X(n) * Z(n) * P(n)) \\ Z(n+1) &= c * (X(n) * Y(n) * P(n)) \\ P(n+1) &= d * X(n) \end{aligned} \tag{1}$$

Where X, Y, Z and P are the state variables and a, b, c, d are the control parameters or the bifurcation parameters.

2.2 Chaos Behavior Investigation

Following the lines given in [12] and [13], the chaotic behavior of the proposed system (1) is investigated by considering mainly the bifurcation and the Lyapunov exponents (LE) theories.

Bifurcation process
Bifurcation theory is concerned with changes in the solutions' behavior of the proposed system (1) as the parameters a, b,c and d are varying.
Figure 1 shows that chaotic behavior appears in several intervals of the parameters $a \in$[0.22, 1.00], $b \in$[0.90, 2.00], $c \in$[-0.95, 0.95] and for the values of the parameter $d \in$[-1.05, 2.5].

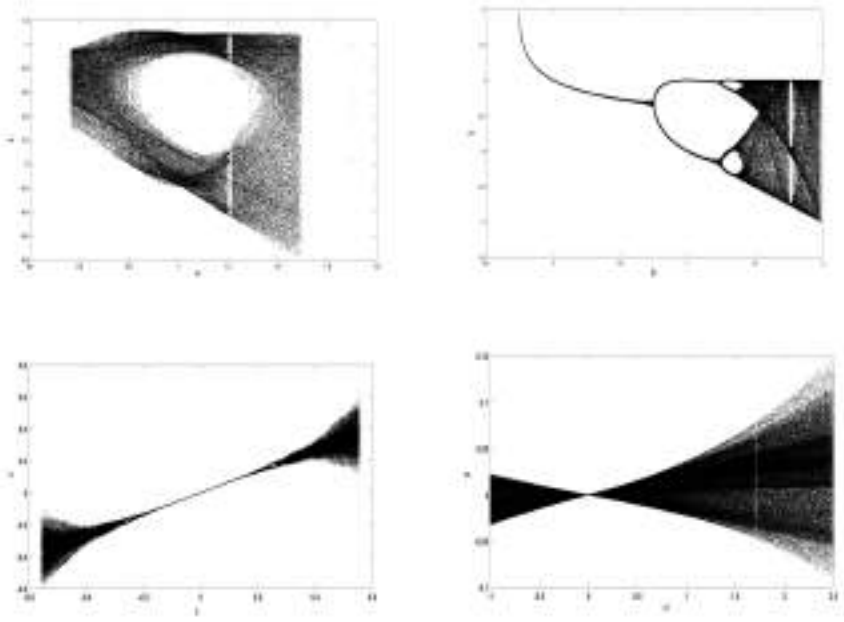


Fig. 1. The bifurcation graphs of the proposed 4-D map.

LE process

Computing the LE values gives a possibility not only to detect all resonances in the response function, but also to detect the presence of chaos [14]. To search for strange attractors or for chaos in the proposed model (1), we proceed as follows:

First, we select the values of the parameters a, b,c and d calculated in the previous step (bifurcation process). Then we iterate equations (1) repeatedly until the Lyapunov exponent becomes small or negative, in which case the solution is probably a fixed point or limit cycle. In either event, we choose a different combination of a, b,c, d and start over the process.

If, after a few thousand iterations, the solution is bounded (not enormous) and the Lyapunov exponent is positive, then it is likely that you have found a strange attractor that corresponds to the chosen values of a, b,c and d.

As a final result, the chaotic behavior is obtained and illustrated by the trajectory graphs and the signal graphs in figures 3 and 4, respectively.

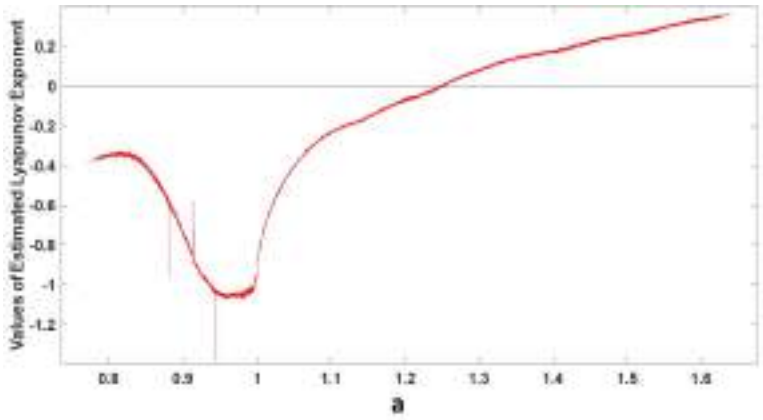


Fig. 2. The LE graph of the proposed 4-D map.

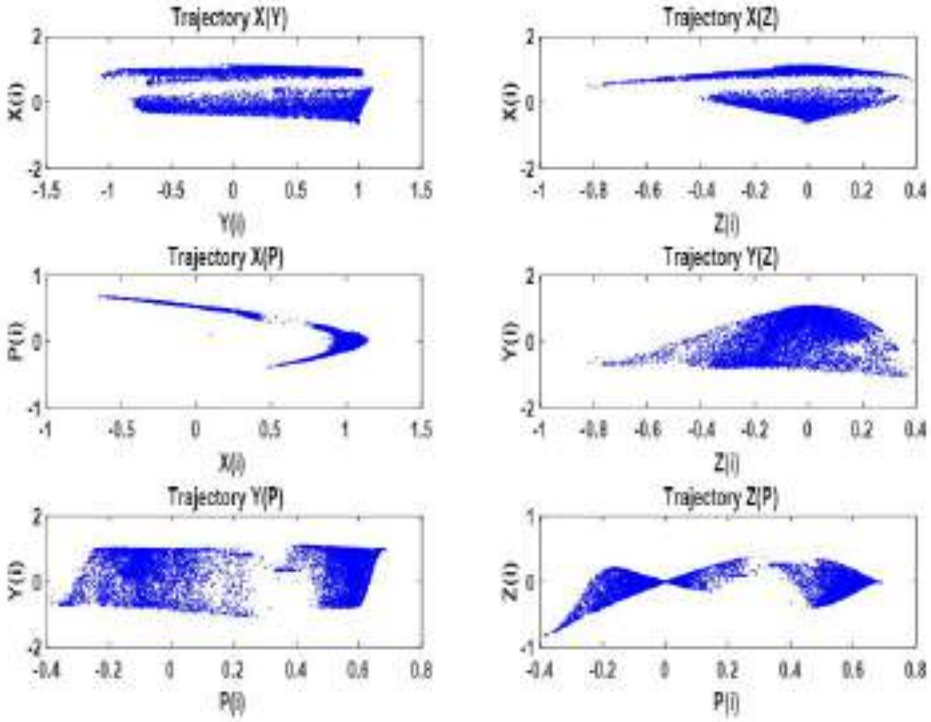


Fig. 3. The trajectory graphs of the proposed 4-D map.

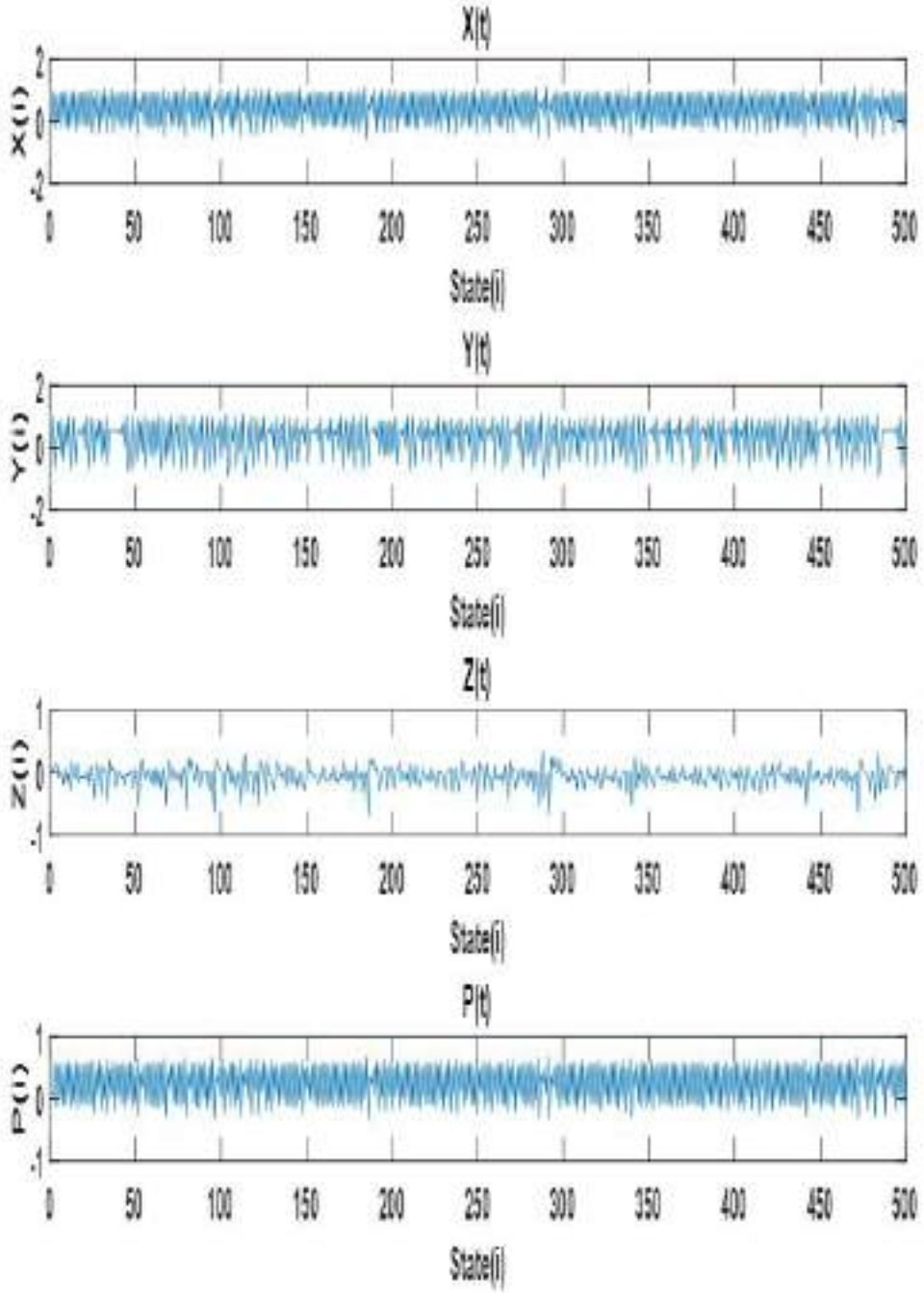


Fig. 4. The signal graphs of the proposed 4-D map.

3 software implementation of secure Communication

Chaotic systems have been introduced in the network communication security domain since that they are characterized by their sensitivity, unpredictability and their widespread spectrum. Hence; we propose to introduce the 4-D chaos map for securing Client-Server Communication. In the first order, we prepare the Client-Server platform composed of two (02) client stations connected through an Ethernet network to one (01) server machine (Figure 5).

Fig. 5. The platform test bench.

The proposed platform works considering the following steps :
- The server listens for client connections;
- The client initiates the connection;
- The connection is established;
- The key-Generator load the first random sequence and the next sample;
- The client uses the first key to encrypt and send data to the server;
- The server use the same key to decrypt and send confirmation of receiving data;

Finally, the client terminates the communication by closing the channel, and the key-Generator module saves samples for next use with the server.

In the second order, we develop a C# software application for exchanging securely messages between connected computers (figure 7).The developed application includes a chaos-based cryptosystem using the proposed 4-D map which is described by the system (2); with a=1.55; b= 1.7; c= 1.6; d= 0.40. Considering that we are targeting to generate a 32-bit key using the chosen map, then we have the key space of the cryptosystem $2^{32*(4+4)}=2^{256}$ which is considered very good value in for encryption since that the required value

should be more than 2^{100} to satisfy the encryption requirements for key space [16].

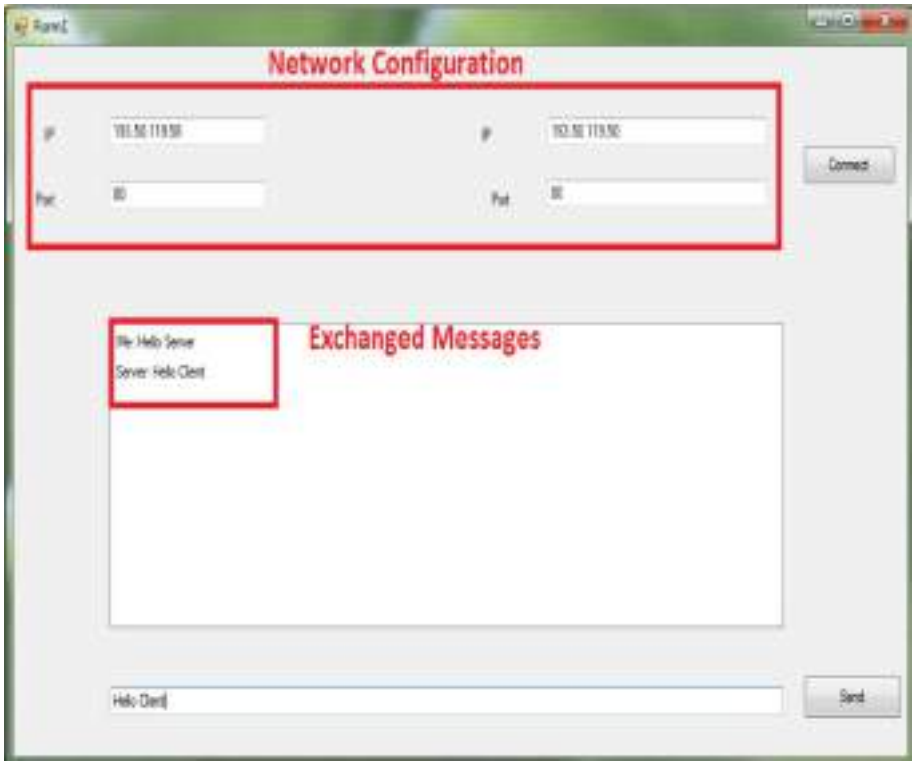


Fig. 6. The developped application's interface.

4 Conclusion

In recent years; researching and developping new and higher dimensional chaotic systems become a rising trend in perticular for secure communication purposes. In this paper, we propose new 4-Dimensional chaotic map in first order. Secondly, the chaotic behavior of the proposed system is investigated using an algorithm based on the bifurcation and the Lyapunov exponents properties. Finally, we apply the proposed system to generate random keys in order to perform a secure IP-communication.
As future work, statistical and security tests of the proposed system will be considered as well as the hardware implementation of the proposed scheme.

References

1.Y. Zhou, Z. Hua, C. M. Pun and C. L. Philip Chen. Cascade Chaotic System With Applications. *IEEE Trans. Cybern.*, vol. 45, no 9, pp. 2001-2012, Sept. 2015.

2.G. Mykolaitis. Temperature-stable hyperchaotic oscillator.*IEEE Trans. Circuits Syst. Fundam. Theory Appl.*, vol. 45, no 11, pp. 1198-1200, Nov. 1998.

3.A. Chonka, J. Singh and W. Zhou. Chaos theory based detection against network mimicking DDoS attacks.*IEEE Commun. Lett.*, vol. 13, no 9, pp. 717-719, Sept. 2009.

4.Y. Chen, X. Ma and X. Wu. CDDoS detection algorithm based on preprocessing network traffic predicted method and chaos theory.*IEEE Commun. Lett.*, vol. 17, no 5, pp. 1052-1054, May. 2013.

5.X. Zhang, L. Wang, Z. Zhou, and Y. Niu. A Chaos-Based Image Encryption Technique Utilizing Hilbert Curves and H-Fractals.*IEEE Access*, vol. 7, pp. 74734-74746, 2019.

6.U. Cavusoglu, A. Akgul, S. Kacar, I. Pehlivan, and A. Zengin. A novel chaos-based encryption algorithm over TCP data packet for secure communication.*Secur. Commun. Netw.*, vol. 9, no 11, pp. 1285-1296, 2016.

7.A. Al Talabani, A. Nallanathan and H. X. Nguyen. A Novel Chaos Based Cost Function for Power Control of Cognitive Radio Networks. *IEEE Commun. Lett.*, vol. 19, no 4, pp. 657-660, Apr. 2015.

8.J. C.Sprott. *Chaos and Time-Series Analysis.* Oxford: Oxford University Press, vol. 69.2003.

9.Z. Sandor, B. Erdi, A. Szell, and B. Funk. The Relative Lyapunov Indicator: An Efficient Method of Chaos Detection.*Celest. Mech. Dyn. Astron.*, vol. 90, no 1, pp. 127-138, Jul. 2004.

10.Q. Lai, C. Chen, X. W. Zhao, J. Kengne, and C. Volos. Constructing Chaotic System With Multiple Coexisting Attractors.*IEEE Access*, vol. 7, pp. 24051-24056, 2019.

11.W. K. S. Tang, G. Q. Zhong, G. Chen, and K. F. Man. Generation of n-scroll attractors via sine function.*IEEE Trans. Circuits Syst. Fundam. Theory Appl.*, vol. 48, no 11, pp. 1369-1372, Nov. 2001.

12.C. Shen, S. Yu, J. Lu, and G. Chen. A Systematic Methodology for Constructing Hyperchaotic Systems With Multiple Positive Lyapunov Exponents and Circuit Implementation.*IEEE Trans. Circuits Syst. Regul. Pap.*, vol. 61, no 3, pp. 854-864, mar 2014.

13.C. J. Feng, X. Z. Li, and L. Y. Liu. Bifurcation and global dynamical behavior of the f(T) theory.*Mod. Phys. Lett. A*, vol. 29, no 07, pp. 1450033, Mar. 2014.

14.J. K. Hunter. *Introduction to dynamical systems.* U.C.Davis Mathematics MAT A, 207. 2011.

15.M. A.ABABNEH. New four-dimensional chaotic attractor. *Ain Shams Engineering Journal*, vol. 9, no 4, pp. 1849-1854, 2018.

16.B. Bouteghrine, M. Rabiai, C. Tanougast and S. Sadoudi. Hardware Implementation of Secured Socket Communication based on Chaotic Cryptosystem. *2019 International Conference on Cyber Security and Protection of Digital Services (Cyber Security)*, Oxford, United Kingdom, 2019, pp. 1-4, doi: 10.1109/CyberSecPODS.2019.8884879.

Liutex-based analysis of vortex in two-phase flow past 2D cylinder using LBM on GPU

Pengxin Cheng[1], Nan Gui[1], Xingtuan Yang[1], Jiyuan Tu[1,2], Shengyao Jiang[1], and Haijun Jia[1]

[1] Institute of Nuclear and New Energy Technology, Collaborative Innovation Center of Advanced Nuclear Energy Technology, Key Laboratory of Advanced Reactor Engineering and Safety of Ministry of Education, Tsinghua University, Beijing, 100084, People's Republic of China.
(E-mail: guinan@mail.tsinghua.edu.cn)
[2] School of Engineering, RMIT University, Melbourne, VIC 3083, Australia

Abstract. Two-phase flow past obstacle is widely applied in industries and engineering, where the interaction of different phases coupled with influence of solid is rather complex. In this context, the flow characteristics and vortex field have been investigated to explore the mechanism of two-phase drag and vortex variation. The Lattice Boltzmann Method is utilized to study the multi-component multiphase flow. The computation implemented on GPU is remarkably accelerated owing to the natural parallelism of LBM. The process of two-phase flow past a 2D cylinder is thoroughly examined. The drag forces including the total force and the components caused by dispersed phase and continuous phase respectively, are illustrated, and the mechanisms for the variations have also been explained. Meanwhile, the vortex-identification approaches based on the Liutex as well as traditional methods are compared. The relationship between the breakup as well as coalescence process and extremums of different vortex identification variables is analyzed.
Keywords: Two-phase flow, flow past cylinder, LBM, vortex identification, Liutex.

1 Introduction

Flow past obstacles is a ubiquitous phenomenon encountered in various industries including chemical process, food processing and power engineering. For decades, numerous experimental and numerical efforts have been made focusing on drag and lift force, vortex separation, flow-induced vibration, etc. The two-phase flow past obstacles plays a significant role in practical application, such as the packed bed reactor in petroleum chemical process and heat exchangers in power industry.

In view of the impact of bubbles on the solid, the influence of bubble movement on vortex shedding and the involvement of bubbles into the vortex, the interaction in two-phase cross flow between three phases, namely solid, liquid and gas, is coupled and intricate. Under this circumstance, the traditional computational fluid dynamics (CFD) methods like Finite Volume Method (FVM) might encounter difficulties and fail to acquire the details of flow field from macroscopic scale.

Endowed with the advantage of flexible geometry characteristics, inherent parallelism and simplicity of implementation, the Lattice Boltzmann method (LBM) has

witnessed rapid development in the past decades, particularly in the field of multiphase flow and flow in complex geometry. Based on the molecular kinetic theory, the LBM provides a novel and potent approach with solid background of physics and efficient algorithm in multi-scale analysis from the mesoscopic perspective. With the computation largely restricted to local nodes, the LBM is ideal for implementation with parallelized hardware such as GPU, which shows an apparent advantage over CPU in High Performance Computing. The implementation of LBM on GPU contributes to boosting computation speed and achieving reliable acceleration performance.

As a fundamental research target, the vortex is of vital significance for an intuitive understanding of turbulence. After decades of intensive research, there are still plenty of ambiguous issues left to be solved including a widely-acknowledged definition of vortex. The vortex identification method has evolved from the first generation intuitive vorticity-based method that fails to discriminate rotational motion from a shear layer in viscous turbulent flow, to the second generation that rely on other eigenvalue-based parameters including Δ, λ_2, λ_{ci}, etc., which are unfortunately dependent on case-related uncertain threshold and unable to capture the vortical strength to some extent.

Recently Gao and Liu put forward the third generation Liutex method [1,2], which provides a more precise perspective of defining vortex. According to the theory, the vorticity could be decomposed into the non-rotational part, mainly shear and the rotational part called Liutex, namely $\vec{\omega} = \vec{S} + \vec{R}$. The direction of Liutex vector, representing the local rotation axis, is defined via the real eigenvector of the velocity gradient tensor. The magnitude of Liutex, used to quantify the local rotational strength, could be simply determined by a explicit formula proposed by Wang [3], which follows the idea of getting rid of the non-rotational part from the vorticity. With the aid of Liutex, continuous research has been conducted on various applications including the hairpin vortices [4] and vortex in swirling jets [5,6].

This paper will present a comprehensive investigation on the flow and vortex field and drag force of multiphase flow past cylinder. The variation of multiphase drag force with the movement of fluid will be analyzed. The detailed evolution of the primary variables during the bubble deformation will be discussed for a deep understanding of its relationship with vortex generation and development. A wide range of identification parameters including the Liutex will be evaluated. The relationship between extremums of different vortex identification variables and bubble deformation process, mainly the breakup as well as coalescence, will be analyzed.

2 Models and methods

2.1 Basics of Lattice Boltzmann Method

Lattice Boltzmann Method could be derived from the Boltzmann equation based on the kinetic molecular dynamics.

$$\frac{\partial f}{\partial t} + \vec{\xi} \cdot \nabla_{\vec{x}} f + \vec{a} \cdot \nabla_{\vec{\xi}} f = \Omega(f), \tag{1}$$

where $f(\vec{x},\vec{\xi},t)$ denotes the particle distribution function that represents the density of particles with velocity $\vec{\xi}$ at position $\vec{x}$ and time t. The collision operator Ω_f characterizes microscopic interaction between molecules and could be simplified by single-relaxation-time approximation proposed by Bhatnagar–Gross–Krook (BGK) [7].

$$\Omega(f) = -\frac{f - f_{eq}}{\tau} \tag{2}$$

Through discretizing in the physical space, velocity space and time, the discrete Lattice Boltzmann Equation (LBE) could be derived,

$$f_i(\vec{x} + \vec{c}_i\delta t, t + \delta t) - f_i(\vec{x},t) = (\Omega(\vec{x},t) + F_i(\vec{x},t))\delta t \tag{3}$$

where $\vec{c}_i$ represents the discrete velocity direction, and $f_i(\vec{x},t)$ is the particle distribution function along the direction of $\vec{c}_i$. The widely adopted discretization model is usually expressed as DdQm, where d and m denotes number of dimensions and velocities. For 2D simulation D2Q9 is adopted.

The macroscopic parameters such as density, velocity and internal energy can be calculated statistically from moments of the discrete distribution function.

$$\rho = \sum_i f_i,\ \rho\vec{u} = \sum_i c_i f_i,\ \rho e = \frac{1}{2}\sum_i (\vec{c}_i - \vec{u})^2 f_i. \tag{4}$$

By means of the multi-scale expansion techniques such as Chapman-Enskog analysis, the LBGK model could recover continuity equation and Navier-Stokes equation at the macroscopic scale under the limit of small Mach number.

2.2 Multiphase Lattice Boltzmann methods

In multiphase flow simulation, the Shan-Chen model [8] incorporates a repulsive or attractive Shan-Chen force, which does not need any explicit interface tracking. Compared with other models, the Shan-Chen model has satisfactory performance in both efficiency and accuracy and is adopted in the multiphase LBM simulation.

For single-component multiphase flow, the interaction force density acting on fluid at $\vec{x}$ can be computed via integral over all possible interaction sites $\vec{x}'$.

$$F_{\mathrm{SC}}(\vec{x}) = -\int (\vec{x}' - \vec{x}) G(\vec{x},\vec{x}')\psi(\vec{x})\psi(\vec{x}')\mathrm{d}^3\vec{x}' \tag{5}$$

where $\psi(\rho)$ also called pseudo-potential denotes the effective density. $G(\vec{x},\vec{x}')$ is a Green function to determine the strength and range of interaction, with the simplest form concerning nearest lattice neighbors.

$$G(\vec{x},\vec{x}') = \begin{cases} w_i G & \vec{x}' = \vec{x} + \vec{c}_i\delta t \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

In a multiphase fluid system constituted by n components, the LBE of Shan-Chen model can be expressed as

$$f_i^\sigma(\vec{x}+\vec{c}_i\delta t, t+\delta t) - f_i^\sigma(\vec{x},t) = \frac{f_i^{eq(\sigma)} - f_i^\sigma}{\tau^\sigma}\delta t + (1-\frac{\delta t}{2\tau^\sigma})F_i^\sigma(\vec{x},t) \tag{7}$$

where $f_i^\sigma, f_i^{eq(\sigma)}, \tau^\sigma, F_i^\sigma$ is the distribution function, equilibrium distribution function, relaxation time and forcing term of the σ component.

The total interaction force acting on the σ component can be computed by adding the force imposed by all the components.

$$F_{\mathrm{SC}}^\sigma(\vec{x}) = -\psi^\sigma(\vec{x})\sum_{\sigma'} G_{\sigma\sigma'}\sum_i w_i \psi^{\sigma'}(\vec{x}+\vec{c}_i\delta t)\vec{c}_i\delta t \tag{8}$$

In the present work we focus on the two-component multiphase system without phase change, namely $n = 2$ and $G_{\sigma\sigma} = 0$. $G_{\sigma\sigma'}$ is set as positive for repulsive force between different components. The overall density, weighed average velocity and pressure of the fluid are calculated as follows.

$$\rho = \sum_\sigma \rho_\sigma,\ \vec{u} = \frac{\sum_\sigma(\rho_\sigma\vec{u}_\sigma/\tau_\sigma)}{\sum_\sigma(\rho_\sigma/\tau_\sigma)},\ p = c_s^2(\rho_\sigma + \rho_{\sigma'}) + c_s^2 G\rho_\sigma\rho_{\sigma'} \tag{9}$$

2.3 Momentum Exchange Algorithm of the Bounce-Back method

The widely adopted method to implement no-slip wall boundary condition in LBM is the Bounce-Back method, with the advantage of simplicity of implementation and guaranteed mass conservation. However, the traditional Bounce-Back method in the regular lattice could only approximate arbitrary boundaries with stair-case shapes, which might introduce large errors at cells nearby boundaries. While by increasing the mesh resolution the error could be reduced and the results approaches precise solution.

In the context of Bounce-Back scheme, the Momentum Exchange Algorithm (MEA) is adopted to calculate the force in this paper. The basic idea of MEA is to identify the populations across the boundary and calculate the net momentum transfer consisting of two parts, namely the momentum carried by the fluid to the wall f_i^{in} and the momentum transported from the wall to the fluid $f_{\bar{i}}^{out}$. The procedure includes the following steps [9].

- Identify the momentum links between solid and boundary nodes x_i^w.
- Evaluate the incoming populations f_i^{in} and bouncing populations $f_{\bar{i}}^{out}$ of each momentum link.
- Calculate the transferred momentum during the time steps and acting force on the boundary.

The drag and lift forces are calculated using MEA in the following section. The drag force F_D is to quantify the resistance in the opposite direction of flow. The lift force F_L is in the orientation perpendicular to the flow direction. Thus they can be evaluated by the components of total force in X and Y direction. Since the force is calculated via the populations f, for the multi-component case in present paper, the force of the σ component could be obtained based on the populations of each component f^σ.

2.4 Parallel Implementation of LBM on GPU

For practical convenience, the main process of calculation of LBE can be decomposed into two parts, namely collision and streaming. With the features of 'non-linearity is local' in collision step and 'non-locality is linear' in streaming step, LBM is naturally suitable for computation on parallel architecture including GPU. GPU is capable of executing a multitude of threads in parallel simultaneously. which could take advantage of thousands of cores in GPU and remarkably boost the computation. The implementation of LBM on GPU has aroused wide research interest and proved successful in delivering reliable performance. In this paper, the open-source software Sailfish is adopted for the simulation. The computation was implemented on NVIDIA Tesla K20 GPU.

2.5 Vortex identification method and Liutex

Several well-acknowledged identification parameters are selected including the vorticity ω [10], Q [11], R [12], Ω [13], Ω_R [14] and Liutex R [15], which are defined by:

$$\begin{aligned}
\omega &= \|\vec{\omega}\| = \|\nabla \times \vec{u}\| \\
Q &= \frac{1}{2}(\|\boldsymbol{B}\|_F^2 - \|\boldsymbol{A}\|_F^2) \\
\Omega &= \frac{\|\boldsymbol{B}\|_F^2}{\|\boldsymbol{A}\|_F^2 + \|\boldsymbol{B}\|_F^2} = \frac{b}{a+b} \\
\Omega_R &= \frac{\beta^2}{\alpha^2 + \beta^2 + \epsilon} \\
\vec{R} &= R\vec{r} = \left(\langle\vec{\omega}, \vec{r}\rangle - \sqrt{\langle\vec{\omega}, \vec{r}\rangle^2 - 4\lambda_{ci}^2}\right)\vec{r}
\end{aligned} \tag{10}$$

where $\boldsymbol{A} = \frac{1}{2}[\nabla\vec{u} + (\nabla\vec{u})^T]$, $\boldsymbol{B} = \frac{1}{2}[\nabla\vec{u} - (\nabla\vec{u})^T]$ are the symmetric and antisymmetric parts of the velocity gradient tensor, respectively. $a = \|\boldsymbol{A}\|_F^2$, $b = \|\boldsymbol{B}\|_F^2$, where $\|\cdot\|_F^2$ represents the Frobenius norm. $\alpha = \frac{1}{2}\sqrt{\left(\frac{\partial V}{\partial Y} - \frac{\partial U}{\partial X}\right)^2 + \left(\frac{\partial V}{\partial X} + \frac{\partial U}{\partial Y}\right)^2}$ and $\beta = \frac{1}{2}\left(\frac{\partial V}{\partial X} - \frac{\partial U}{\partial Y}\right)$, where while U, V and W represent the velocity components in the XYZ coordinate satisfying $\nabla\vec{V} = \boldsymbol{Q}\nabla\vec{u}\boldsymbol{Q}^T$ by a rotation matrix $\boldsymbol{Q}$ and $[\nabla\vec{V}]_{1,3} = [\nabla\vec{V}]_{2,3} = 0$. ϵ is a small positive parameter. λ_{ci} and $\vec{r}$ are the imaginary part of the complex eigenvalue and the real eigenvector of $\nabla\vec{u}$, respectively. In this work, these criteria will be used for comparative analysis of vortex motion.

2.6 Validation

In this section, the LBM model is firstly validated against 2D single-phase flow past a cylinder to verify the accuracy of force evaluation. Then the Laplace test is conducted to examine the multiphase model.

For two-dimensional case, with the characteristic length corresponding to the obstacle's diameter, the drag and lift coefficient and the Strouhal number are calculated as follows.

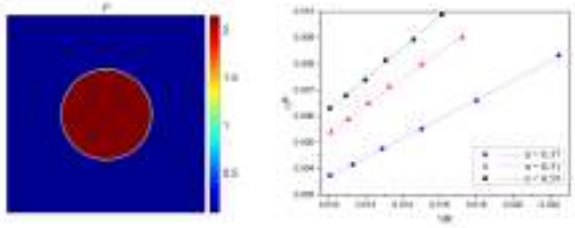
$$C_D = \frac{F_D}{\frac{1}{2}\rho D\bar{u}^2},\ C_L = \frac{F_L}{\frac{1}{2}\rho D\bar{u}^2},\ St = \frac{fD}{\bar{u}} \tag{11}$$

The comparison of simulation results with validation reference is displayed in Table 1. On the whole, the simulation results show in good conformity with benchmark in Schaefer's paper[16].

As for the multi-component Laplace test, a spherical droplet of one phase with radius R is initialized surrounded by the another liquid phase. The interface is determined where the density reaches the average of the bulk densities inside and outside the droplet. The result of Laplace test is displayed in Fig. 1, where the relationship between measured pressure difference Δp and $1/R$ is linear, with coefficient of determination of the fitted lines above 0.999, in desirable agreement with the Laplace Law, namely $\Delta p = \sigma/R$.

Concerning the two-phase model validation, it has been completed in our earlier work [17], by the validating the bubble departure diameter and release period. In addition, the GPU performance of LBM has also been tested and validated via the cavity driven flow case in another earlier work [18].

On the whole, the LBM mode, two-phase scheme, GPU performance have already been well validated.



(a) Density contour **(b)** Data fitting

Fig.1. The validation of Laplace test for multi-component case

3 Results and discussion

3.1 Numerical setup

The flow domain is discretized by 256×1024 Cartesian mesh grids. The two-dimensional cylinder obstacle with the diameter of 100 nodes is placed right in the center of channel and referred to as *cylinder* in the following sections. Initially a circular bubble of the

Table 1. Validation of flow past a 2D cylinder

Re	20		100		
	C_D	C_L	C_{Dmax}	C_{Lmax}	St
Present result	5.5499	0.0118	3.2877	0.9842	0.3062
Benchmark lower bound	5.57	0.0104	3.22	0.99	0.295
Benchmark upper bound	5.59	0.0110	3.24	1.01	0.305
Benchmark bounds refer to the results in Schaefer's paper[16]					

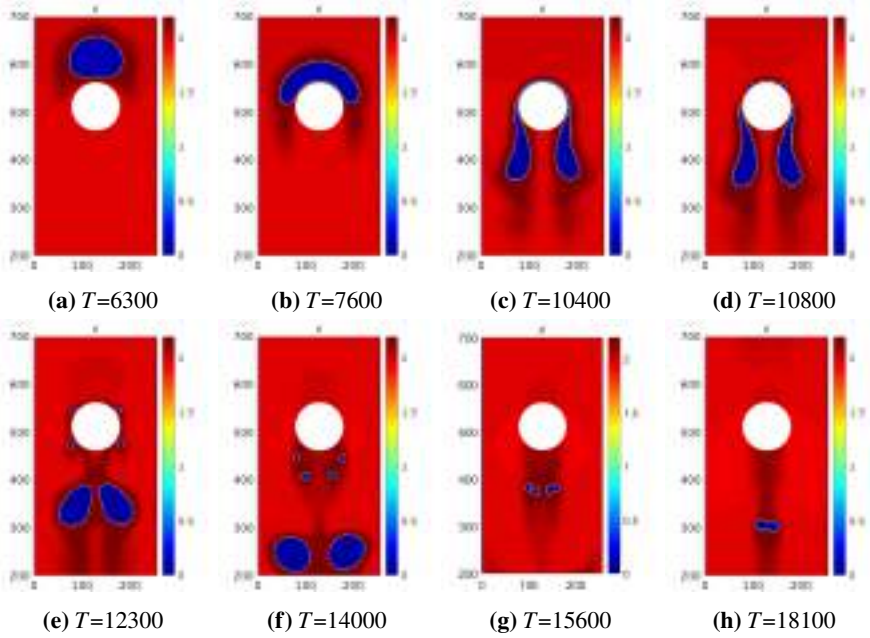
dispersed phase is set up at the upside of channel with the diameter equal to 100 nodes, which will be called *bubble* later. The rest of field is filled with the continuous phase. Note that the density contour denotes the density of the continuous phase. Referring to previous multi-component multiphase research [19], the viscosity of the two phases are set as the same and the parameter G is determined via $G\rho = 1.8$. The initial velocity of the bubble and the force is in the downward direction.

3.2 The evolution of two-phase flow

The evolution of density contour is shown in Fig. 2, where the typical image during the progress is presented. With initial velocity and imposed force, the bubble moves towards the cylinder. As the nearly round bubble approaches the obstacle, the hindrance from the obstacle flattens the bubble at the bottom (at T=6300). Subsequently it is separated by the solid (at T=7600) and stretches to a great extent (at T=10400), till the filament between the bubble becomes unstable and breaks up into several satellite bubbles (at T=10800). The four small bubbles move along the periphery of cylinder and successively detach from the solid (at T=12300). Influenced by the vortex field behind the obstruction, bubbles move towards center-line and become closer (at T=14000). Involved in the vortex behind the obstacle, the distance between two bubbles reduced and they coalesce for the first time (at T=15600) and finally merge into one bubble (at T=18100).

3.3 The evolution of drag force

The evolution of drag force is shown in Fig. 3. As the bubble approaches the cylinder, the drag force of continuous phase F_D^c increment steadily while that of the disperse phase F_D^d declines inversely, leading to the increase of total drag force F_D^{total}. This variation results from the compression of the fluid between the bubble and solid due to the movement of the bubble. This trend continues until F_D^c attains its maximum at T=7625 and F_D^d reaches the minimum at T=7689, when the bubble surrounds nearly the upper half part of cylinder. Then the bubble covers the majority of cylinder periphery and the tendency of the curves converts into the opposite, namely F_D^c decreases till it reaches the minimum at T=10751 and conversely F_D^d increases till its maximum at T=10766. F_D^{total} plunges and attained its minimum at T=10741. With the portion of



(a) T=6300 **(b)** T=7600 **(c)** T=10400 **(d)** T=10800

(e) T=12300 **(f)** T=14000 **(g)** T=15600 **(h)** T=18100

Fig.2. The contour of density at different time steps

bubble coverage of cylinder periphery increasing, F_D^d consequently enhances and F_D^c reduces, until the bubble is over stretched and break up into pieces. Therewith the continuous phase regains the contact with cylinder and F_D^c boosts, arriving at a local maximum at T=11633. In contrast F_D^d drops towards a local minimum at T=11617. Afterwards the curves gradually flatten out.

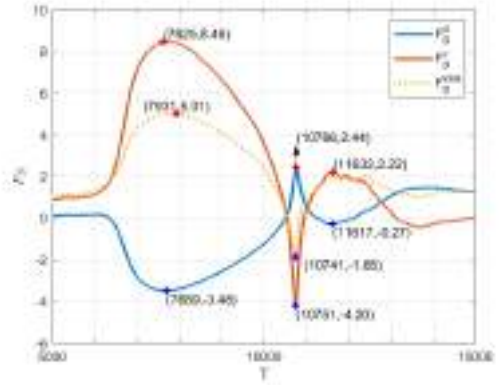


Fig.3. The evolution of drag force

3.4 The evolution of R, S, ω and comparison of vortex identification methods

Based on the relation of $\vec{R} + \vec{S} = \vec{\omega}$, the evolution of R, S and vorticity ω at different time steps are shown in Fig. 4 and Fig. 5.

For the evolution of breakup process in Fig. 4, before bubble brings into contact with solid, some zones with negative Liutex appear inside the bubble in the bottom and lateral at T=6300. As the deformation affects the distribution of Liutex, there exist negative zone in the center and positive zone on the sides at T=7600. At T=10400 there exist zones with positive Liutex between and ahead of the stretched bubbles, as well as zones with negative Liutex laterally adjacent to them. When the break-up of bubble occurs at T=10800, magnitude of Liutex near the position of fracture is relatively high in positive or negative value. As the surface tension transform the bubble into a nearly round shape at T=11600 and T=12300, zones with high magnitude of negative Liutex inside the bubbles are observed.

For the evolution of coalescence process in Fig. 5, when the small bubbles merge together, the distortion results in zones with high level of magnitude of Liutex both inside and around the bubble. Taking the coalescence of two bubbles in the right as an example, at T=15500 there appear negative Liutex zone near the line of bubble contact. At T=15600, there are positive Liutex zone inside the bubble and alternatively positive and negative Liutex outside the bubble. At T=15700, the negative Liutex zones have disappeared and the positive zones also shrink in area and decrease in magnitude.

On the whole, S and Ω show similarity with each other in distributions and magnitudes. Mainly, the bubble's motion and shape dominate the variations of shear and vorticity, whose distributions and magnitudes are so close that the variations of R plays merely a fairly secondary role. Because of the relatively low velocity, shear deformation plays the dominating role on the vortex evolution characteristics. However, indicated by R, the pure rotational motion also shows clear features dominated by the formation of bubbles deformation, breakup and coalescence.

Moreover, considering the immediate moment when the bubble breaks up for example, Fig. 6 shows the comparison of the varied vortex identification approaches defined by Eq. (10), where R, Q, ω, S, Ω and Ω_R can all show the basic characteristics of vortex structure partly yet some differences still exist.

At the moment of bubble breakup, rotational vortices are formed around the four small satellite bubbles near the cylinder. In the meantime, in the downside of the square obstacle, some small pure rotational vortices are also formed. The rotational vortexes are formed around the locations with sufficiently large or locally maximum curvatures of interface, such as around the ends of the stretched bubbles. The reason lies in that the vortex can only be generated when and where the shear is too large for the fluid to resist. Thus vortices may be generated during the bubble breakup near the region with sufficiently large curvatures of interface.

In addition, R and Q show comparatively similar distributions in pairs, which evaluate the vortex by absolutely rotational strength. Whereas Ω and Ω_R show comparatively similar distributions in pairs, which determine the vortex by relatively levels of strengths of rotation to shear.

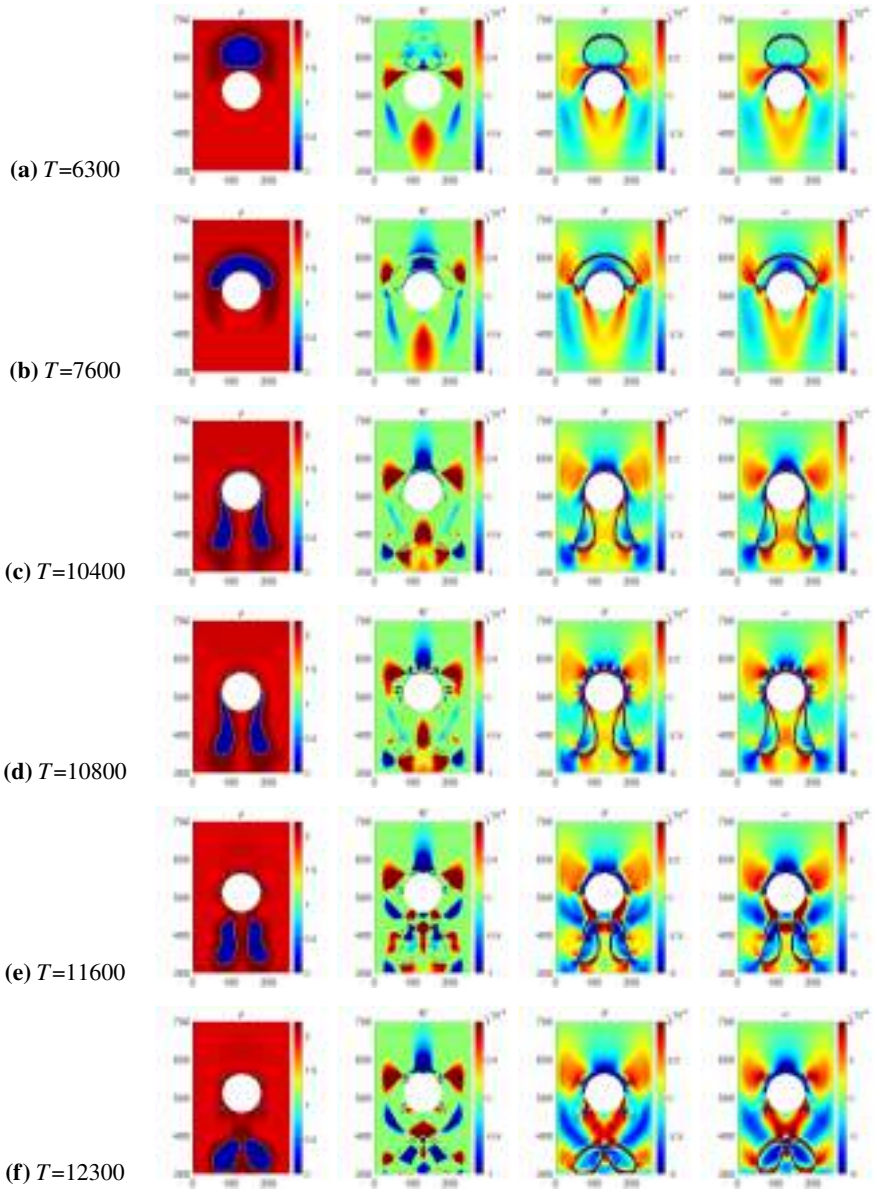


Fig.4. The evolution of R, S and ω at T=6300, 7600, 10400, 10800, 11600 and 12300 respectively

(a) T=15500

(b) T=15600

(c) T=15700

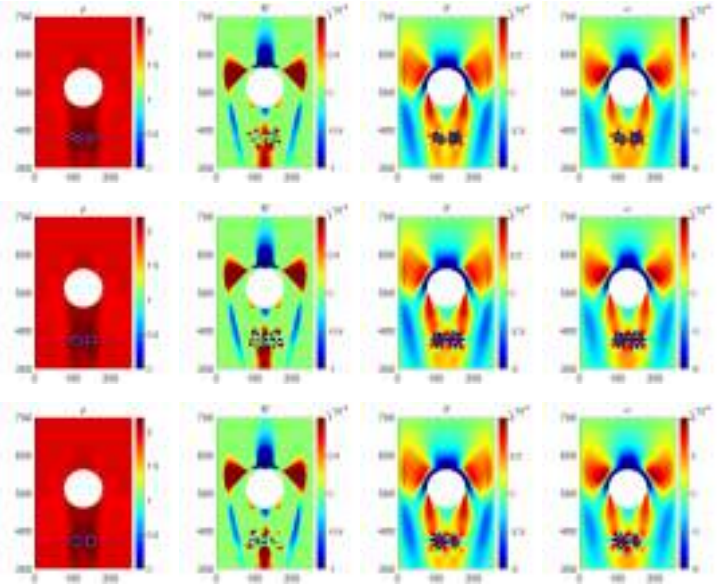


Fig.5. The evolution of R, S and ω at T=15500, 15600 and 15700 respectively

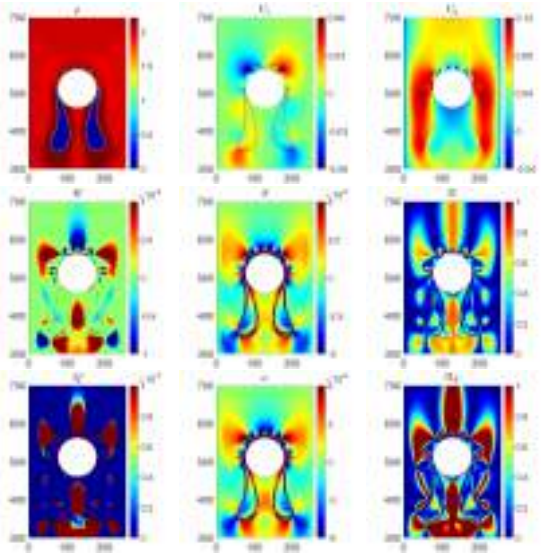


Fig.6. The contour of primary variables at T=10800

3.5 The evolution of R^{max} and S^{max}

The evolution of statistics of primary variables of the dispersed and continuous phases will be discussed in this and following sections. The local maximums and minimums are represented by means of red and blue marks in the figures. In order to relieve the burden of computation, the data is possessed every ten time steps, resulting in the exhibited tenfold extremes.

Firstly, the maximum value of R (namely R^{max}) is shown in Fig. 7. In Fig. (7a), it is noteworthy that the data of R^{max} is only selected from the interior of the bubble (dispersed phase). There exist three distinct peaks at T=10830, T=15480 and T=18050. Thus both the maximum and minimum of Liutex experience extremes at T=10830. Whereas in Fig. (7b) the data of R^{max} is only selected from the exterior of the bubble (continuous phase). The remarkable peaks locate at T=10770, T=15460 and T=18030.

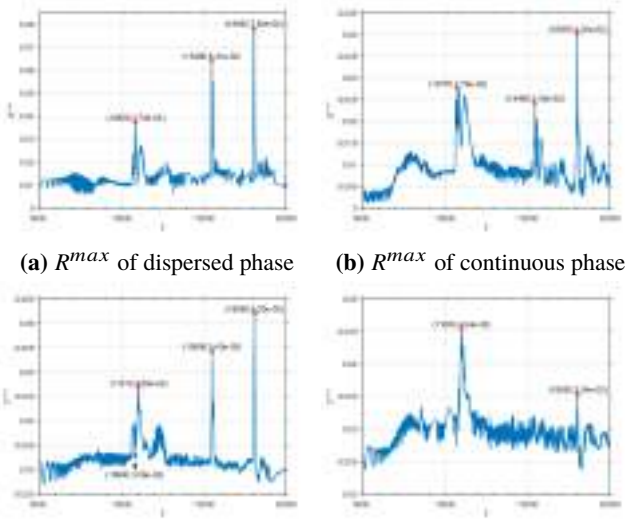
By Fig. (7b), the suddenly sharp increase of R^{max} at about $T = 10770$ is caused by the very moment the bubble begins to break up into small satellite bubbles, whereas the peak of R^{max} at $T = 15460$ corresponds to the moment when small bubbles coalesce into large ones for the first time. Similarly, the peak of R^{max} at $T = 18030$ corresponds to the moment when bubbles merge into one eventually. In Fig. (7a) for the dispersed phase, the sharp peaks of R^{max} correspondingly coincide with that of the continuous phase in Fig. (7b). The small differences of peak times of R^{max} between Fig. (7a) and Fig. (7b) are caused by the tiny differences of rotational motion of vortices of the two phases.

Moreover, detailed inspection also indicates that the peaks of R^{max} of the dispersed phase are later than that of the continuous phase. It may possibly be deduced that the rotational vortex or the vortical motion is transferred from the exterior to the interior of the bubble. In other words, the vortex inside the bubble is induced by the external shear or rotational motions.

On the other hand, with regard to the evolution of the maximum of shear, S^{max}, in Fig. (7c) for the dispersed phase, the three remarkable peaks locate at T=11010, T=15500 and T=18080, with the location approximate to that in Fig. (7a). In Fig. (7d) for the continuous phase, the curve manifests a distinct peak at T=11000 and a minor peak at T=18050.

Comparing Figs. (7c) and (7d) to Figs. (7a) and (7b), it is also clear that the peaks of R^{max} and S^{max} almost take place simultaneously. In traditional applications, the vortex is always identified by vorticity, which in current situation is mostly dominated by shear rather than pure rotation. However, like in Fig. 7, as the maximum of real vortex indicated by R^{max} can also be identified by the maximum shear S^{max}, using vorticity identification methods is somewhat usable if the most accurate description of vortex is not needed.

Nevertheless, the first peak of R^{max} is closer to the moment of bubble break-up than that of S^{max}. Besides there is no peak near the first coalescence of bubbles in Figs. (7d). Therefore, using R to quantify the pure rotational vortex is mostly encouraged in two-phase flows, either for vortices in the dispersed phase or in the continuous phase.



(a) R^{max} of dispersed phase **(b)** R^{max} of continuous phase

(c) S^{max} of dispersed phase **(d)** S^{max} of continuous phase

Fig.7. The evolution of maximum of R and S

3.6 The evolution of Q^{max}

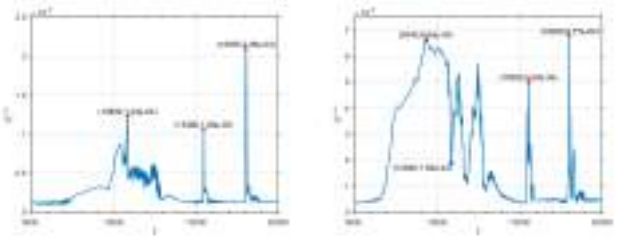
The evolution of Q^{max} is shown in Fig. 8. Respecting the maximum of Q^{max}, in Fig. (8a) for the dispersed phase, there exist three remarkable peaks located at T=10830 and the other two at T=15480 and T=18050. Referring to previous figures, these peaks are on the location near the bubble break-up and coalescence. In Fig. (8b) for the continuous phase, two peaks lie at T=15550 and T=18030, while the curve reaches minimum at T=10880 and experiences drastic fluctuations.

To sum up, it is found that Q^{max} could also indicate the two-phase vortical motion. In other words, it is also an applicable indicator for shape deformation of bubbles and rotational motion of fluids to some extent, yet not as good as R^{max}.

3.7 The evolution of Ω^{ave} and Ω_R^{ave}

Finally, the evolution of Ω^{ave} and Ω_R^{ave} are shown in Fig. 9. In this section, the averaged values of Ω and Ω_R over the dispersed or continuous phases are illustrated.

Concerning Ω^{ave}, Figs. (9a) and (9b) show the mean values of Ω^{ave} of the dispersed phase and continuous phase respectively. The Ω^{ave} of dispersed phase first climbs to a maximum at T=6180. Subsequently it falls back yet rises again and exhibits drastic fluctuations afterwards. This kind of variation of Ω^{ave} of the dispersed phase clearly indicates the variation of vortical motion inside the bubbles, which is clearly more smoothly and varied more slowly than the variation of Ω^{ave} of the continuous



(a) Q^{max} of dispersed phase **(b)** Q^{max} of continuous phase

Fig.8. The evolution of maximum of Q

phase (Fig. (9b)). As a whole, neither Ω^{ave} of the dispersed phase nor that of the continuous phase indicate the instantaneous variations of the vortex motion of the two-phases as good as using R^{max}.

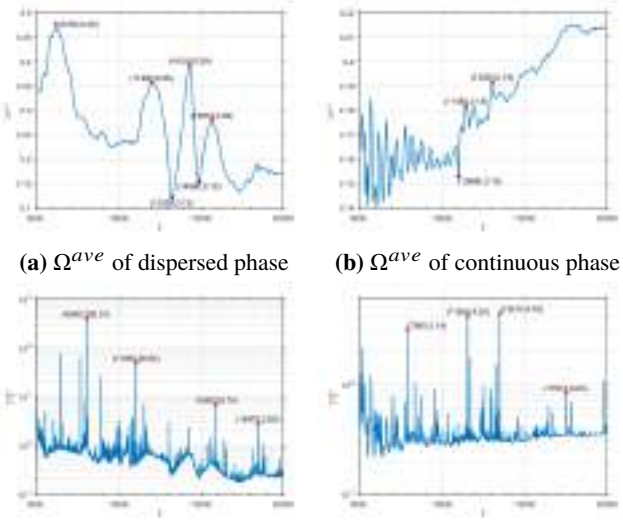
Figs. (9c) and (9d) show the variations of Ω_R^{ave} of the dispersed and continuous phases. In this condition, the variation of Ω_R^{ave} shows much clearer features than Ω^{ave}. To say specifically, Ω_R^{ave} is always fluctuating more evidently and intermittently than Ω^{ave}. Moreover, the peaks of Ω_R^{ave} also take place around the points of suddenly sharp change of former vortex identification methods, e.g. R^{max} or Q^{max}, such as the peaks at T=11050, T=15920 and T=18470.

Referring to Eq. (10), the mechanism may lie in the definition of Ω_R^{ave}. Generally speaking, α and β are better than $\boldsymbol{A}$ and $\boldsymbol{B}$ for defining the pure rotational motion of fluids. α and β are used to define R or the pure rotational vortical motion. Conversely, $\boldsymbol{A}$ and $\boldsymbol{B}$ with possible shear deformation, are not indicating the pure rotational motion features. Hence, it is still very clear to evaluate even the total level or strength of the rotational motion of fluids by using α and β.

4 Conclusion

In this work, we presented the characteristics of vortex and drag force of two-phase flow past a cylinder using Liutex-based analysis through validated multiphase LBM model on GPU computation. The main findings can be summarized as follow:

- Regarding the evolution feature, the bubble may be deformed and stretched greatly until its final breakup during the process of flow past the cylinder. After that, the small satellite bubbles may gradually coalesce to make larger ones successively in the downstream.
- Dominated by the bubble behavior, the drag force components caused by the dispersed bubble phase and continuous fluid phase show the opposite trend. The mechanism of variation of force components results from the specific variations of velocity and density of the two-phases.
- Influenced by the two-phase interaction behavior, the vortex identification method R successfully describes the region of pure rotation of fluids, though its magnitudes are small compared to the shear of fluids.



(a) Ω^{ave} of dispersed phase **(b)** Ω^{ave} of continuous phase

(c) Ω_R^{ave} of dispersed phase **(d)** Ω_R^{ave} of continuous phase

Fig.9. The evolution of average of Ω and Ω_R

- Via averaging over the respective phase fields, Ω_R shows evident advantages compared to Ω, since Ω_R can not be weakened by the averaging process, whereas the Ω^{ave} fails to show such kind of features.

Acknowledgments

The authors are grateful for the support of this research by the National Science and Technology Major Project (Grant No. 2011ZX06901-003), the National Natural Science Foundations of China (Grant No. 51576211), the National High Technology Research and Development Program of China (863) (2014AA052701).

References

1. C. Liu, Y. Gao, S. Tian, and X. Dong, "Rortex a new vortex vector definition and vorticity tensor and vector decompositions [J]," *Physics of Fluids*, vol. 30, no. 3, 2018.
2. Y. Gao and C. Liu, "Rortex and comparison with eigenvalue-based vortex identification criteria [J]," *Physics of Fluids*, vol. 30, no. 8, pp. 085 107–, 2018.
3. Y.-q. Wang, Y.-s. Gao, J.-m. Liu, and C. Liu, "Explicit formula for the liutex vector and physical meaning of vorticity based on the liutex-shear decomposition [J]," *Journal of Hydrodynamics*, vol. 31, no. 3, pp. 464–474, 2019.
4. Y.-s. Gao, J.-m. Liu, Y.-f. Yu, and C. Liu, "A liutex based definition and identification of vortex core center lines [J]," *Journal of Hydrodynamics*, vol. 31, no. 3, pp. 445–454, 2019.
5. N. Gui, H.-b. Qi, L. Ge, P.-x. Cheng, H. Wu, X.-t. Yang, J.-y. Tu, and S.-y. Jiang, "Analysis and correlation of fluid acceleration with vorticity and liutex (rortex) in swirling jets [J]," *Journal of Hydrodynamics*, vol. 31, no. 5, pp. 864–872, 2019.
6. N. Gui, L. Ge, P.-x. Cheng, X.-t. Yang, J.-y. Tu, and S.-y. Jiang, "Comparative assessment and analysis of rortex vortex in swirling jets [J]," *Journal of Hydrodynamics*, vol. 31, no. 3, pp. 495–503, 2019.
7. P. L. Bhatnagar, E. P. Gross, and M. Krook, "A model for collision processes in gases. i. small amplitude processes in charged and neutral one-component systems [J]," *Physical Review*, vol. 94, no. 3, pp. 511–525, 1954.
8. X. Shan and H. Chen, "Lattice boltzmann model for simulating flows with multiple phases and components [J]," *Physical Review E*, vol. 47, no. 3, pp. 1815–1819, 1993.
9. K. Timm, H. Kusumaatmaja, A. Kuzmin, O. Shardt, G. Silva, and E. Viggen, "The lattice boltzmann method: principles and practice [M]," Berlin, Germany, Springer, 2016.
10. J.-m. Liu, Y.-q. Wang, Y.-s. Gao, and C. Liu, "Galilean invariance of omega vortex identification method [J]," *Journal of Hydrodynamics*, vol. 31, no. 2, pp. 249–255, 2019.
11. Y. Wang, Y. Gao, H. Xu, X. Dong, J. Liu, W. Xu, M. Chen, and C. Liu, "Liutex theoretical system and six core elements of vortex identification [J]," *Journal of Hydrodynamics*, 2020.
12. Y. Gao, J. Liu, Y. Yu, and C. Liu, "A liutex based definition and identification of vortex core center lines [J]," *Journal of Hydrodynamics*, vol. 31, p. 445–454, 2019.
13. X. Dong, Y. Wang, X. Chen, Y. Dong, Y. Zhang, and C. Liu, "Determination of epsilon for omega vortex identification method [J]," *Journal of Hydrodynamics*, vol. 30, p. 541–548, 2018.
14. C. Liu, Y. Gao, X. Dong, Y. Wang, J. Liu, Y. Zhang, X. Cai, and G. N, "Third generation of vortex identification methods: Omega and liutex/rortex based systems [J]," *Journal of Hydrodynamics*, vol. 31, p. 205–223, 2019.
15. J. Liu, Y. Gao, Y. Wang, and C. Liu, "Objective omega vortex identification method [J]," *Journal of Hydrodynamics*, vol. 31, p. 455–463, 2019.
16. M. Schaefer and S. Turek, "Benchmark computations of laminar flow around a cylinder [M]," vol. 48, pp. 547–566, 1996.
17. T. Sun, N. Gui, X. Yang, J. Tu, and S. Jiang, "Numerical study of patterns and influencing factors on flow boiling in vertical tubes by thermal lbm simulation [J]," *International Communications in Heat & Mass Transfer*, vol. 86, p. 32–41, 2017.
18. P. Cheng, N. Gui, X. Yang, J. Tu, and S. Jiang, "Application of lattice boltzmann methods for the multiphase fluid pipe flow on graphical processing unit [J]," *Journal of Computational Multiphase Flows*, vol. 10, no. 3, p. 109–118, 2018.
19. H. Huang, D. T. Thorne, M. G. Schaap, and M. C. Sukop, "Proposed approximation for contact angles in shan-and-chen-type multicomponent multiphase lattice boltzmann models [J]," *Physical Review E*, vol. 76, no. 6, p. 066701, 2007.

Marco Desogus[1]

Systemic effects of the credit crunch. A dynamic mathematical model

Abstract

This paper began with the hypothesis that there is an intrinsic mutualistic dependence between the bio-economic performance of banks and that of enterprises. Indeed this supposition is sustained by the comprehensive presentation herein of the correlations calculated on the data of Italian system. At the very least, this law seems plausible in relations among financial intermediaries and smaller enterprises.
Therefore, this research concentrated on developments within the bank-enterprise system (and, as a repercussion, in households), since clearly, the positive effects, including the impact on macroeconomics, generated by an efficient banking sector in the supply of funding to the productive infrastructure, so that it can remain healthy and efficient, should not be overlooked (nor, on the contrary, should the negative effects produced by the disappearance such a virtuous cycle). Hence, through the use of dynamic models, this paper posits a mathematical argument, which evaluates the structural trends of the populations of banks and companies with respect to the more or less expansive strategies of the former, in granting credit.
Empirical observations of the negotiation of loans to enterprises can also be represented as a "critical stress utilization" of the capacity of (micro)enterprises to generate positive economic variations using ever less financial leverage. This has led to the assessment of stable and unstable points of equilibrium, of the bifurcations generated and of their irreversibility (hysteresis), resulting in, on the banking side, stopping profits and an increase in non-performing loans.
Finally, it was investigated whether there is an optimal minimum level of credit leverage to the system, also in light of the current destabilization caused by the SARS-COVID19 emergency, and therefore of the better "sealing actions" that must be conferred to the system itself.

KEYWORDS: *credit crunch; dynamic modelling; simulation; credit big data; economic systems; financial stability; non-linear analysis.*

JEL CODES: C61, C62, E32, E44, G21

[1] PhD in Economics; Economist and Credit Financial Advisor – Practitioner. Contract lecturer of Mathematics for Economics and Social Science, Dept. of Economics and Business Science – University of Cagliari, Italy. E-mail: marco.desogus79@unica.it – marcodesogus@gmail.com. Address: viale Sant'Ignazio, 84 – 09123 Cagliari ITALY.
With the supervision of prof. Beatrice Venturi and thanks to the collaboration of Elisa Casu, BSc.

1. Introduction

Banks, especially commercial banks – that is to say, institutions, which cannot, due to regulatory restrictions (or limitations), engage in business other than financial intermediation – are acutely dependent on the performance of their customers. It has been equally well established that banks, in performing their lending functions, implement their strategies based on the criteria of commerce and profit: loans are represented by risk portfolios whose overall weighting is derived from the aggregate of segmented units, which in turn are based on precise evaluations of each individual counterparty. This paper endeavours to study the effects of these two objectives being systematised, in order to define whether there is an ideal level – a quantitative *optimum* – of total loans to be disbursed by banks to companies. Starting from the reasoning that there is an average portfolio of clients, where all the n positions that it comprises[2] have the same PD – Probability of Default, the same LGD – Loss Given Default and the same EAD – Exposure at Default, the total expected losses TEL_n – which represent a cost and which will be accounted for in the bank's income statement – will be [3]:

$$TEL_n = n * PD * EAD * LGD \qquad [1]$$

Then add the unexpected losses, which represent a capital constraint for the banks and which determine the level of caution that they will increasingly adopt in the disbursement of loans. By applying the "Vasicek model", with the observance of a time t = 1 year:

$$TL_n = \sum_{i=1}^{n} U_i LGD_i EAD_i \qquad [2]$$

Where:
TL_n is the total loss on the portfolio over the time t;
LGD_i and EAD_i where there will be the LGD and the EAD of the i-th company respectively;
U_i is the Boolean indicator, which takes on a value of 1 if the i-th enterprise has reached default within time t, or the value 0 if the i-th enterprise has continued to be *performing* in the range considered. This is a function of the amount x, which represents a Gaussian variable of the macroeconomic context and which indicates an additional scenario in which the enterprises in the portfolio operate:

$$x_i = \upsilon K + \tau e_i \qquad [3]$$

where K is the random variable strictly interpreted from the macroeconomic context, while the e_i element indicates the random idiosyncratic factor of the individual counterparty. Then υ and τ are the respective multipliers, one of which is often set as known a priori. It should be pointed out that a certain component of e_i is however dependent on K. Therefore:

$$TL_n = \sum_{i=1}^{n} U_i(K, e_i) LGD_i EAD_i \qquad [4]$$

$U_i(K, e_i)$ replaces the PD in the formula and indicates the likelihood that the default of enterprise i will occur, based on its own idiosyncratic factor e_i and on the conditions of its macroeconomic context[4]. Since, LGD_i and EAD_i are assumed to be known *a priori*, the need to maximise e_i and K [5] appears to be clear, limiting ourselves to the bank-enterprise relationship, by both of these entities in cohabitation within the economic system: with the banks needing to minimize losses and reduce capital absorption, and the enterprises needing to obtain credit and do *business* sustainably and profitably.

---

[2] Cf. Conti (2016).
[3] Losses in the portfolio measured *ex post* with minimal deviations.
[4] For a more thorough development of this reasoning, see Desogus & Venturi (2019).
[5] A targeted analysis of the K component, based on the same initial postulate, was conducted by Desogus & Casu (2020).

Clearly, there are even more macroscopic components inside of *K*, including monetary policy, international economic influences, credibility and trust of the country, etc. This is why we are focusing our attention on the dynamics of the bank-enterprise system (and, in turn, households), since we cannot overlook the positive influence, including macroeconomic effects, generated by an efficient banking sector in the transmission of liquidity to the productive complex so that it can remain healthy and performing (and, on the contrasting condition, where there are the negative effects produced by the disappearance of this virtuous cycle)[6]. We shall call the population of the banks Z_1 and the population of the companies (in portfolio) Z_2 and that $z_1 = z_1(t)$ and $z_2 = z_2(t)$ are the manifestations of loans *performing* (or *in bonis*)[7] over the time *t* of the two populations Z_1 and Z_2 respectively:

$$\begin{cases} \frac{dz_1}{dt} = z_1 f_1(z_1, z_2) \\ \frac{dz_2}{dt} = z_2 f_2(z_1, z_2) \end{cases} \qquad [5]$$

This system of differential equations indicates the dependence of the number of *performing* of each population on that of both[8]: the link between the two populations is in fact intrinsically mutual

$$\frac{\partial f_1}{\partial z_2} > 0 \quad \text{and} \quad \frac{\partial f_2}{\partial z_1} > 0 \qquad [6]$$

If the above can be viewed as empirically acceptable – a healthy productive sector resonates with a healthy banking sector, which continues to be capable of strengthening the good condition of companies – it is necessary to delve into the dynamics of the bank-enterprise system, taking the limits set by the *K* and *e* components into account. These limits, which are at first felt at each company within the Z_2 population, provide z_2 a logistical trend with a maximum value that is always less than the totality of Z_2 and the procedural and strategic constraints on the banking business at the time of deliberation and approval.

2. Methodological notes and Italian data-set

Hence, the first step will be to analyse the complexity and non-linearity of the interactions between the population of banks and the population of companies, and to examine the equilibrium points in the dynamic system arising from these interactions. Specifically, there will be the defence of the idea that the contraction of credit, especially to micro and small enterprises, in their correlation with the increase in the mortality rate of those same businesses and the deterioration of existing credit, produces "bio-economic" effects, which are very similar to the observations of exploitation (*more precisely*, over-exploitation) of resources in the same system – inhabited by both populations – by the *dominant species* or *player*. It can be empirically deduced how excessive exploitation by a generic agent may cause difficulties in the regeneration of resources, which could therefore lead to negative consequences for the entire system.

As has been acknowledged, each company can generate positive flows towards the banking system that can be broken down into different forms:

- the creation of income, part of which becomes bank deposits through savings and thus a form of funding for the banks;

[6] Cf. Iyer *et al*. (2014), and Petrosky-Nadeau (2013).

[7] The rationale for default levels could be developed in contrasting terms: the final result would be the same.

[8] Cf. Düllmann & Kick (2014).

- the payment of intermediation charges (broadly understood), which constitute income for the bank.

For the purposes of this paper, it will also be necessary to semantically adapt the concept of the exploitation of the "enterprise" resource. If, in fact, the role of banks' intervention does change in its depth and adequacy through the reduction of credit leverage below a certain threshold, we will see a phenomenon that is only deceptively positive for banks, and only in the short-term, given the reduction in supervisory provisions (and greater available liquidity) as well as more contained risk indicators in the portfolio. However, an excessive restriction of credit, which is gradually tightened over time, means jeopardising the favourable environment for the resilience of the system of production and business in general, as can be seen from the data in the tables below, extracted from the Italian case, taken as an example[9].

Figure 1 – Loans to enterprises in Italy, in million Euros

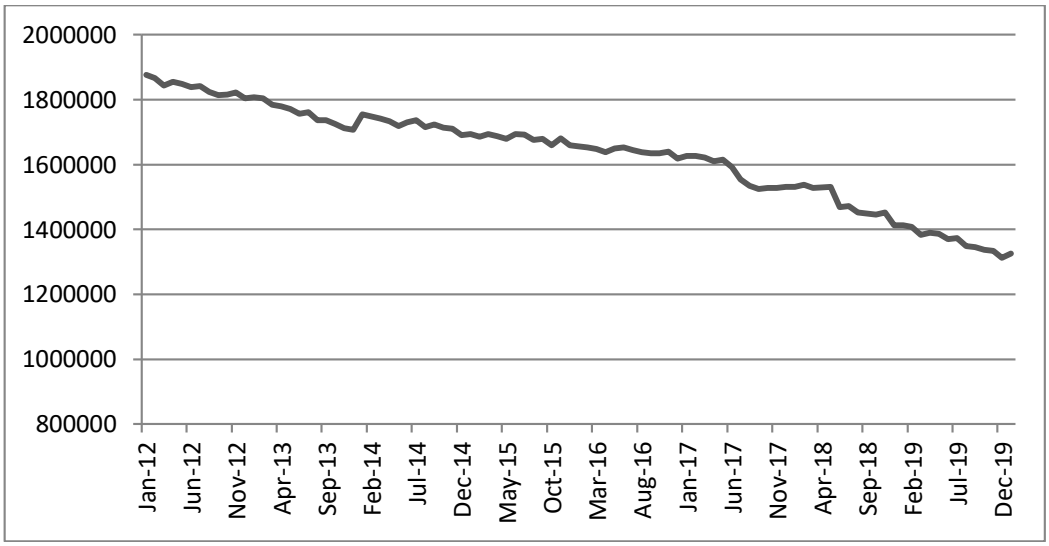
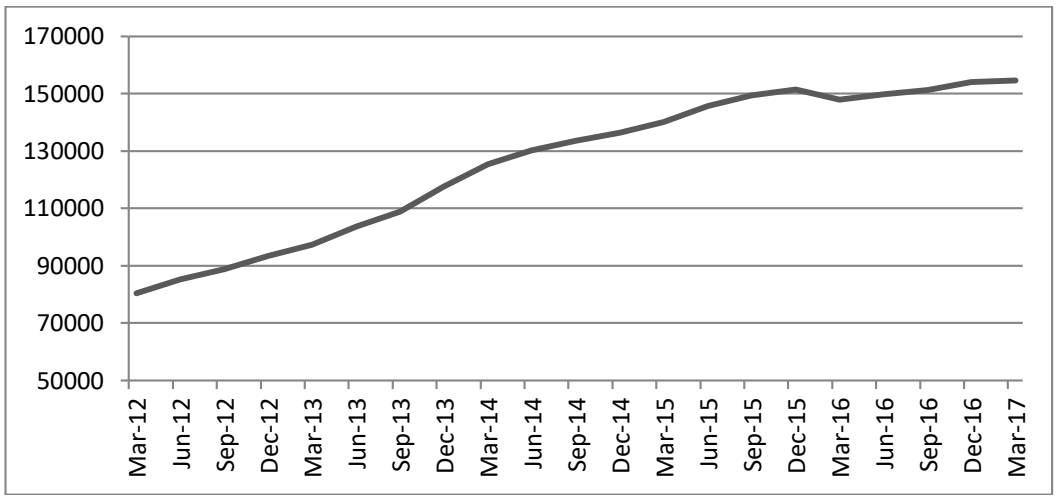


Figure 2 – Net non performing loans in Italy, in million Euros (March 2012 – March 2017)



[9] Tables in the appendix. Processing based on data from the Bank of Italy, ISTAT and Chambers of Commerce.

Figure 3 – Total number of Italian enterprises

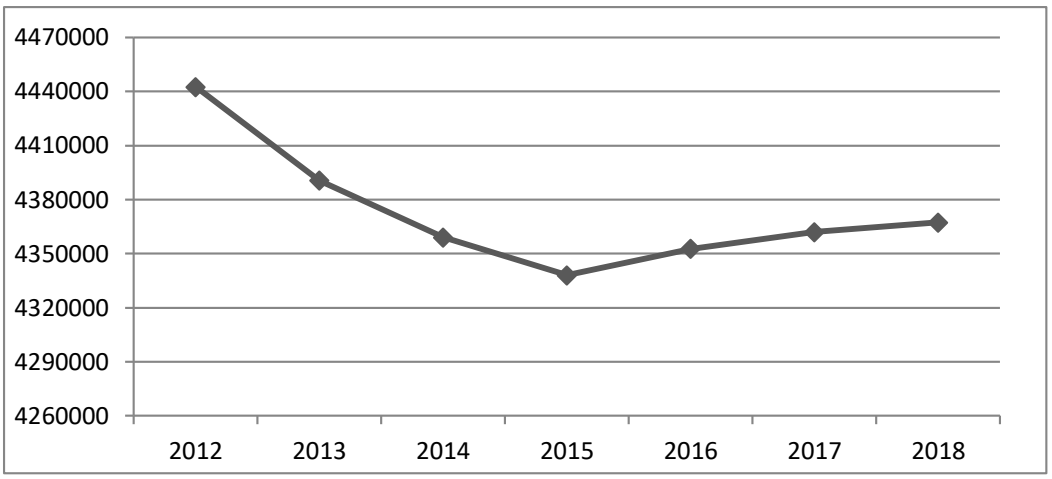
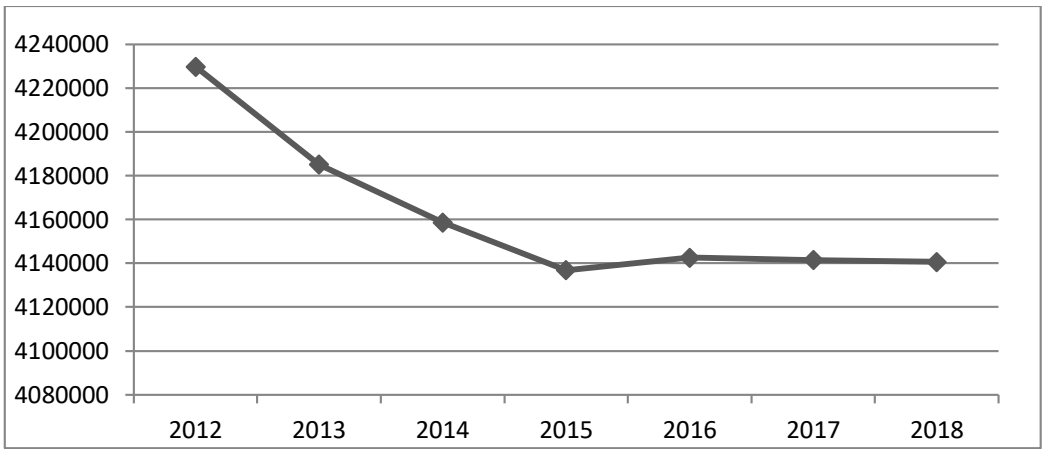


Figure 4 – Number of Italian micro enterprises

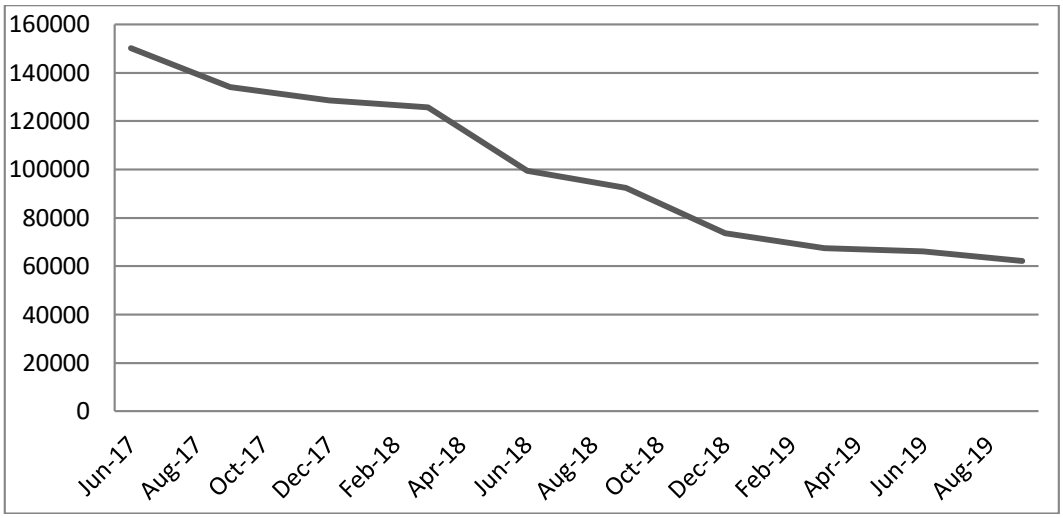


The scenario, presented here among the figures shown above, begins from the progressive reduction in the disbursement of loans – which we assume to be an independent variable – to which a general increase in net non-performing loans until the first quarter of 2017 can be reasonably correlated (mean Bravais-Pearson correlation coefficient at March 2017 = –0.57). Nevertheless, the repercussion of this manifestation has different results according to the size of the business. On the one hand, micro enterprises leave the market more quickly than *new-co* companies of a similar size, on the other, macro enterprises seem not to be affected by this phenomenon. However, the incidence of the macro-enterprise component corrects the total figure only slightly.

In fact, from the *data-set* adopted, the correlation between the reduction in the volume of credit granted to the production system and the number of SMEs operating on the market came to +0.73.

For the sake of methodological completeness, another observation should be made on the trend in non-performing loans from March 2017 to December 2019, still net of the NPLs sold.

Figure 5 – Net non performing loans in Italy, in million euros (June 2017 – September 2019)



As can be seen in Figure (5), the data transmitted by the banks seem to show a reduction in impaired loans. Instead, this phenomenon is due to recent ECB provisions on the management of non-performing loans: specifically, the March 2017 guidelines and subsequent addenda, which strengthened monitoring procedures and at the same time, provided incentives for the sale of those loans. These actions – evidently – led to provisions and losses, respectively, in the banks' balance sheets. ABI (Italian Banking Association), in its 2018 and 2019 reports, posted disposals for 50-70 billion euro per period. Moreover, even before the SARS-COVID-19 emergency, it estimated an outlook for a new growth in non-performing loans, even net NPLs, for the 2020-2021 two-year period.

Therefore, this point of view will interpret banking business as aimed at exploiting the system, continuing to attract enterprise savings and reducing risks and risk capital. This, as widely cited in the literature[10], causes shocks in enterprise resources and reduces those same resources, even leading to extremely critical scenarios. Evidently, this also penalizes the very stability of the system and will end up affecting the banks' balance sheets in terms of lower stocks of funding as well as higher costs due to deteriorated assets and net losses[11]. Expanding the frame of reference to the whole economic system, since income from salaried employment is also transferred from companies, the harmful effects will have a procyclical affect on the entire nation's macroeconomics[12].

3. The evolution of the population of enterprises in relation to available leverage. A first dynamic model

Let us proceed with some dynamic analyses, [13] and indicate with X the periodic stock of Z_2

$$X(t+1) = F\big(X(t)\big) = X(t) + GX(t) - C(t) \qquad [7]$$

The formula [7] represents, over time, the *metaphorical* quantity of the *renewable* resource made up of enterprises, with the connotation given above. t identifies the time, represented (for now) by discrete intervals, which, in line with the characteristics of the system

[10] *Ex multis*, cf. Bernanke *et al*. (1996).
[11] Cf. Wehinger (2014).
[12] Cf. Buera *et al.* (2015).
[13] The following mathematical analyses is partly inspired by and conducted based on – with the appropriate adaptations to the contents here treated –the contribution of Bischi *et al*. (2005).

considered, can be quarters, half-years or years. $X(t)$ is the number of the enterprise resources; G is its growth rate in the unit of time and for each population unit; $C(t)$ gives an account of the quantity of the resources that are lost ("exploited") and therefore the enterprises that have left the market or that have shown defaults as a result of the lesser banking support.

The temporal evolution of the model is described by the function that provides the law, which determines the evolution of the quantity of the enterprise resource in '$t+1$', it being known at the time t and hence increased by its natural growth capacity or decreased by the quota lost.

The first deduction we can make using this equation concerns the existence of a condition of equilibrium when $X(t+1)$ is equal to $X(t)$: that is to say, when the population remained constant over the range and that therefore $GX(t)$ is equal to $C(t)$. Scrutinizing this last relation, it becomes equally clear that the proposed law takes both the bio-ecological characteristics (G factor), and the action of the agent banks (C factor) into account. If the latter, in t, is less than the G function, then there will be an increase in the population of enterprises over the period '$t+1$' and vice versa there will be a reduction in the number of enterprises. The Italian data highlight this situation, among other things, as a performance trend in the n intervals examined. This model could be represented by means of a continuous variable and therefore with instantaneous evolution through the use of the derivative dX/dt. Nevertheless, it seems that the phenomenon being discussed in this section would be more correctly represented through the periodic growth model because of its own aforementioned intrinsic characteristics.

In the absence of a credit crunch, we can approximate $C(t)=0$ and the population of enterprises will describe a quantitative evolution based solely on the micro- and macro-economic context in which it is operating. In this scenario, it is assumed that in each period γX new individuals enter the market and lX leave due to "natural causes". We can rewrite the equation [7] as follows:

$$X(t+1) = X(t) + \gamma X(t) - lX(t) = (1+\gamma-l)\,X(t) \qquad [8]$$

Where $G=\gamma-l$ is the population growth rate and the $X(t+1)=\gamma X(t)$ model is linear and depicts a geometric progression due to g.

$$X(1) = gX(0);\;\; X(2) = gX(1) = g^2X(0);\;\; \ldots;\;\; X(t) = g^{\mathrm{t}}X(0);\;\; \ldots \qquad [9]$$

With $g<l$ (and therefore $g<1$), the progression converges to zero exponentially. If $g=l$ the population number remains the same and will therefore have a constant value. If instead $g>l$, and therefore $g>1$, there will be a population growth trend moving toward infinity with an exponential function[14].

The mortality rate of the enterprises l, is determined not only by the disappearance of leverage, but is also subject to other micro- and macro-economic causes. For example, there is the context of the country in which they are inserted or the level of competition that exists among them. As concerns the latter, we can write $l=\lambda X$ where the parameter λ precisely indicates the greater difficulty – and, conversely, the easiest exit from the market – in the case of an increase in product competitors. G therefore, subsequent to this additional step, is equal to $\gamma-\lambda X$. This function, represented by a straight line, is zero in $K=\gamma/\lambda$ and this is an equilibrium point with zero growth, called 'carrying capacity'.

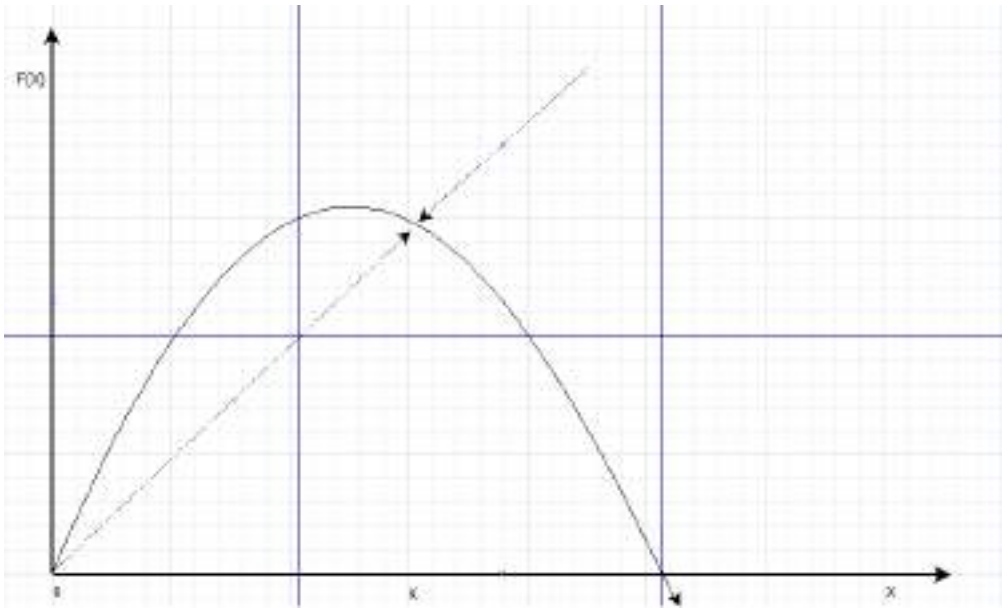
$$X(t+1) = F\big(X(t)\big) = (1+\gamma)\,X(t) - \lambda X(t)^2 \qquad [10]$$

We are here facing a model that has become non-linear, where the $F(X)$ function is a parabola with the intersection of the abscissae in $X=0$ and $X=(1+\gamma)/\lambda$. The system's equilibrium

[14] Clearly, the exponential trend moving toward infinity is precisely an abstract situation, since the "finite" environment in which companies exist and operate would in any case provoke an initial change in the function's concavity until its arrest.

points correspond with the intersections between the parabola and the bisector $F(X) = X$, that is to say $X(t+1) = X(t)$, which are $X = 0$ (extinction equilibrium point) and $X = K$ (carrying capacity).

Figure 6 - Logistic growth function

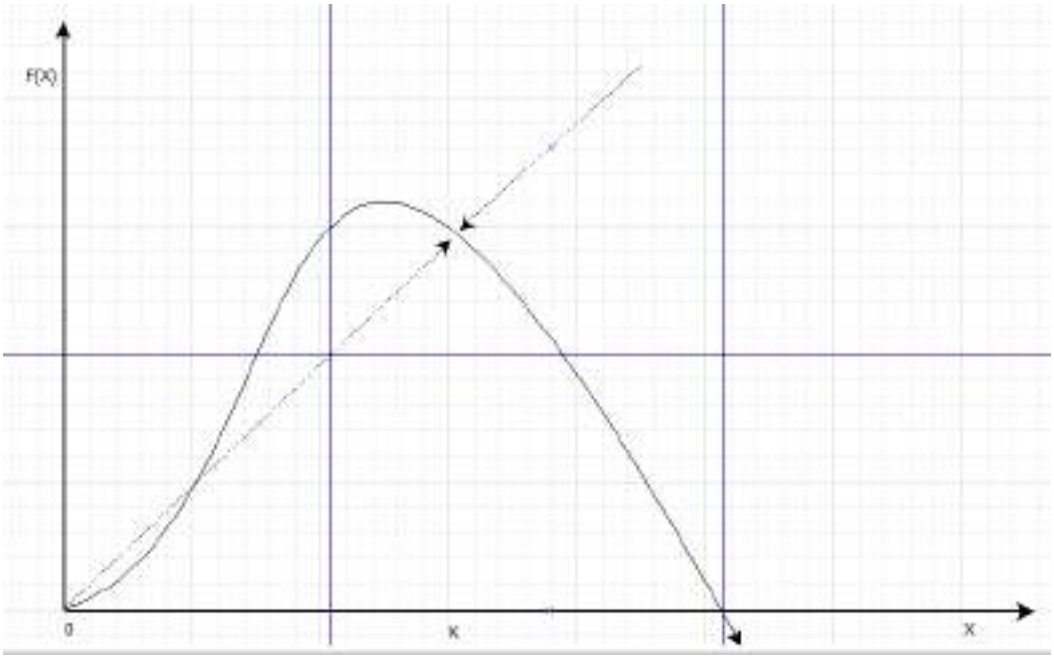


Since the latter is a "natural equilibrium point", the forces acting on it would be more easily compensated by endogenous automatic processes – for small deviations – which can be traced back to $X = K$. This is because, in the logistic growth function, $F(X) > X$ and $F(X) < X$ are to the left and to the right of the equilibrium point respectively. Vice versa in the case of $X = 0$, since any increase of the resources, whether spontaneous or stimulated, will develop an accelerated departure from this equilibrium point in the various successive time intervals, and will still tend towards carrying capacity.

These general considerations should be further broken down by typing the growth functions. From a positive starting value, the intrinsic growth rate $r = G(0)$, we reach another intersection with the abscissae in the value $K > 0$ (i.e. $G(K) = 0$), that is to say the oft-mentioned *natural* equilibrium point. If the function thus described is increasing and continuous – though not necessarily linear – the previous observations remain valid: $F(X) = X\big(1 + G(X)\big)$ is in fact concave and unimodal and not particularly different from the parabola used for the general case.

Nevertheless, in dealing with the population of enterprises within an economic system, we must have recourse to the specific growth function $G(X)$, since we have already had the opportunity to mention how the increase in the number of these enterprises is conditioned by the total number of enterprises insisting on the system, and above all by those that undertake the same commercial objective. For this reason, the population of enterprises will be described by a growth function with a maximum point in an intermediate value of that same population: this situation is called growth with *depensation*. (Allee effect)

Figure 7 - Growth with depensation function



The attitude of the curve is always characterized by one single maximum – in compliance with the unimodality – however, it will go through a first phase of convexity for small values of X and a point of inflection that will make it concave as X grows. Even here we notice two equilibrium points: the fixed extinction point in $X = 0$ and the carrying capacity point in $X = K > 0$. This characteristic change of concavity assumes particular relevance in the effects of the credit crunch, as will be seen below.

Beginning again from the equation [7], we set $C(t) > 0$, with $C(t)$ being the number of companies that exit the market or go into default due to the reduction in leverage they are granted. The condition of equilibrium is described by $F\big(X(t)\big) = C(t)$.

Let us now introduce the parameter S into this discussion. This parameter highlights the banking system strategy that implements a reduction of credit to the system, in order to minimize its capital allocations and maximize its profits, even from non-commercial business (e.g. the purchase of sovereign bonds). In addition, we will also consider the coefficient a, which takes the degree of "aggressiveness" in the implementation of this credit crunch-strategy into account. Thus, we will have $C(t) = aSX(t)$. So, we will rewrite the equation [10] as follows:

$$X(t+1) = F\big(X(t)\big) = X(t)\big(1 + \gamma - aS - \lambda X(t)\big) \qquad [11]$$

The $F(X)$ function is now a parabola passing through the origin with the vertical line $X = (1 + \gamma - aS)/(2\lambda)$ as its axis of symmetry. Setting $F(X) = X$, the equilibrium points will be $X_0 = 0$ and $K_S = (\gamma - aS)/\lambda$. From a natural equilibrium point with carrying capacity obtained by $S = 0$ (absence of credit crunch) in $K = K_0 = \gamma/\lambda$, we can obtain different scenarios with the increase of S:

- With $aS < \gamma$, the K_S equilibrium point remains positive and stable, whilst at the $X_0 = 0$ point, instability is noted.
- Upon the additional increase of the parameter S, the value of the positive equilibrium point is reduced because K_S is a decreasing function of S.

- If the level of S, that is, the level of the a coefficient increases up to $aS = \gamma$, an overlap of the two equilibrium points will be created.
- An additional increase of S will determine a bifurcation in which the K_S equilibrium point will become negative and unstable; at the same time this will cause an attraction in the point X_0.

In other words, when aS exceeds the value of the parameter γ – that is, the enterprises' 'demographic rate' – the system seems to inexorably lead towards the (theoretical) extinction of the enterprise resource.

Hence, let us assume that the banks have only the two extreme strategies available:

a) $aS > \gamma$
b) $aS < \gamma$

The utility functions associated with each of these are in figure (8).

Figure 8

players		Strategy **a**	Strategy **b**
	Banks	$u(\eta)$	$u(\theta)$
	Companies	$u(\nu)$	$u(\xi)$

As set out above, the banks' pay-off will be $\theta > \eta$ (at least over the medium- to long-term); the companies' pay-off will be $\xi > \nu$ (always)[15]. What emerges is a game with a non-cooperative, dominant player (Z_1 – the banks – who will decide the strategy), with only one win-win solution (figure 8) in choice b, which is reinforced by the application of the von Neumann-Morgenstern rationality axioms[16].

Figure 9

Strategy	↓ Z_1 Z_2→	**a**	**b**
	a	$u(\eta), u(\nu)$	
	b		$u(\theta), u(\xi)$

To formalize the banks' strategy for financing enterprises, particularly concerning the reduction of liquidity, we write $Q = aSK_S$, where Q indicates the net performance of banks in terms of profit and provisions relative to their credit activity:

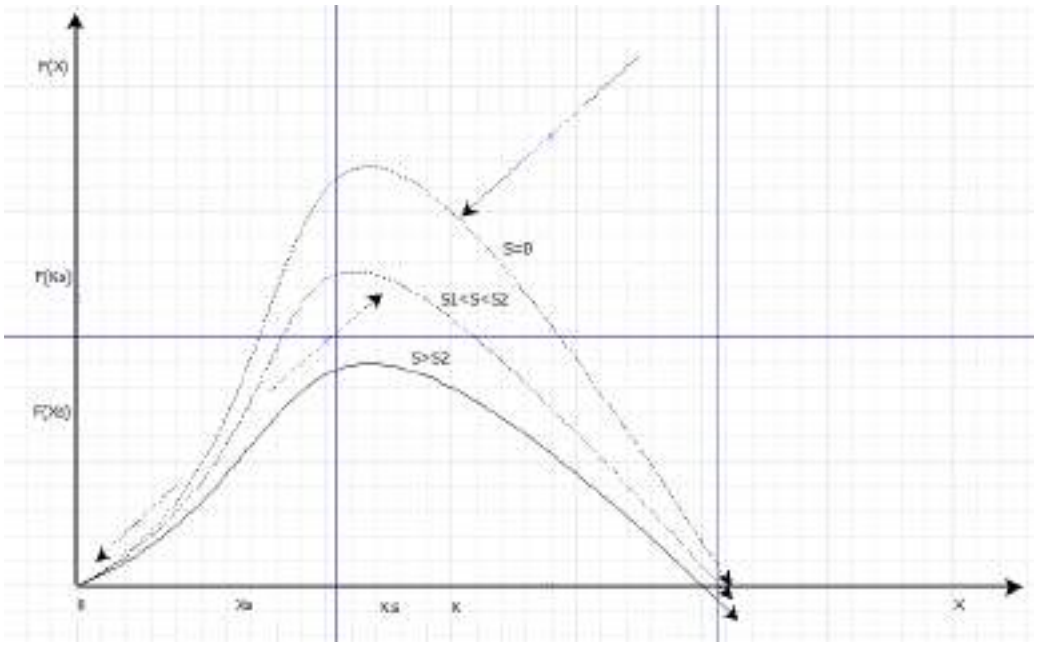
- if $S < \gamma/a$, then $Q = [aS(\gamma - aS)]/\lambda$;
- if $S > \gamma/a$, then $Q = 0$.

The best and most *remunerative* level of Q, with full preservation of the bank-enterprise system's stability, is obtained by $S = \gamma/(2a)$.

[15] Cf. Orton *et al.* (2015).
[16] Cf. Desogus & Casu (2020).

If, therefore, a greater incidence of the restrictive policy represented by aS offers greater advantages in the short-term, how the bank-enterprise system, having reached equilibrium, suffers a progressive depletion of the enterprise population and an increase in losses for the banks over the long-term (with $S > \gamma/(2a)$) begins to take shape, until it determines unfavourable outcomes for $S = \gamma/a$. If, consequently S_{MSQ} was the maximum level of sustainable credit crunch, the situation $S > S_{MSQ}$ would lead to an *over-contraction*, which is a concept very similar to *overexploitation* borrowed from the field of Bio-economics.
As mentioned above, the growth function of the enterprise population is characterized by depensation. This phenomenon is described by the change of concavity of the function F. This creates complications in the presence of credit reduction activity that will bring about a new equilibrium, which we will call X_S and which can be traceable between the equilibrium of the extinction point and the carrying capacity point K_S (that is to say, $0 < X_S < K_S$). X_S is also unstable and represents a junction between stock values: when $X < X_S$ its evolution will tend towards extinction; with $X > X_S$, the system is directed towards the equilibrium point K_S. X_S is therefore *the survival threshold*: the value of S in correspondence with the threshold establishes the bifurcation value, which will be indicated by S_1.

Figure 10



If $S > S_1$, then, equilibrium will be established in $X = 0$, with a survival threshold of $X_S > 0$. Upon the further increase of S, one can observe a progressive approach of the two points of equilibrium. This has the effect of reducing the value of the stock at a stable equilibrium, which at the same time, creates greater vulnerability of the system, caused by the rise of the survival threshold, below which the population of enterprises will be heading towards extinction. If the credit contraction strategy S becomes even more significant, we will be able to observe X_S still growing, and K_S shrinking, until X_S and K_S overlap and then cancel themselves out. This places us before another bifurcation, S_2, which will lead to only one possible solution: the theoretical *exhaustion* of the enterprise resource.
The measures adopted by the central banks, their monetary policy in general and the public injection of liquidity into the productive system were made necessary to avoid this "trap", which undermines the survival of the entrepreneurial fabric, the stability of the banking system and of the entire economic system. Yet these measures should always necessarily be

implemented before the critical value of S_2 has been reached: that is to say, before the stock $X(t)$, which has fallen below the survival threshold, slips into the range of attraction of the equilibrium point of extinction.

Actually, the phenomena analysed produce such *extreme* effects over a very long period during which there are no containment measures whatsoever. Moreover, the equilibrium point of extinction is asymptotic: or rather, there is an asymptote located at a distance from the extinction point equal to a radius p. This distance is primarily due to the presence in the population Z_2 of enterprises – especially the largest – able to make up for the lack of financial leverage with their own and endogenous means and resources. Furthermore, the factor $x_i = \alpha K + \beta e_i$ – formula [3] illustrated in the introduction – impedes the population of enterprises from *plummeting* to zero. This can be demonstrated with the relation $\Delta x_i \Delta p = m$, with m parametric: in other words, if the radius between the number X of Z_2 and the extinction point is significantly shortened, the contribution of the K (macro-economic) and e_i (idiosyncratic) factors will proportionally counterbalance the result of the product leaving the value of m unchanged. This is the resilience of the super-system. Nevertheless, the point m is still at a critical height and the entire system is labile.

The instruments known as long term refinancing operations (LTRO) and the main refinancing operations (MRO) on a weekly or bimonthly basis (so-called '*overnight*' loans), as well as 'quantitative easing' – with a *bridge*, precisely, for banks – were intended to bring back and/or maintain the value of the stock X above the survival threshold, reducing S until $X(t)$ will (again) be higher than the threshold value X_S. This process will be successful when $S < S_1$ is achieved. The irreversibility described above, which induces regulators to act according to the precautionary principle, is called "hysteresis".

Beginning again from the equation [7], we can now see how central (banking) policy can intervene "upstream" to avoid the situation of the non-linear system discussed previously and referable to the "catastrophe theory". This, therefore, applies even before the implementation of recovery measures, and where there is still $S_2 < S < S_1$.

Control over the credit crunch can be modelled by keeping under observation the levels of default or exit from the market by companies, caused by the reduction of their financial leverage, and making $C(t)$ equal to a constant c:

$$X(t+1) = F\big(X(t)\big) = X(t)\big(1 + \gamma - \lambda X(t)\big) - c \qquad [12]$$

The parabola of the growth function will be shifted down by h units. Solving the equilibrium equation $F(X) = X$, or rather $\lambda X^2 - \gamma X + c = 0$, we will have two solutions:

$$X_c = \frac{r - \sqrt{\gamma^2 - 4c\lambda}}{2\lambda} \quad \text{and} \quad K_c = \frac{r + \sqrt{\gamma^2 - 4c\lambda}}{2\lambda} \qquad [13]$$

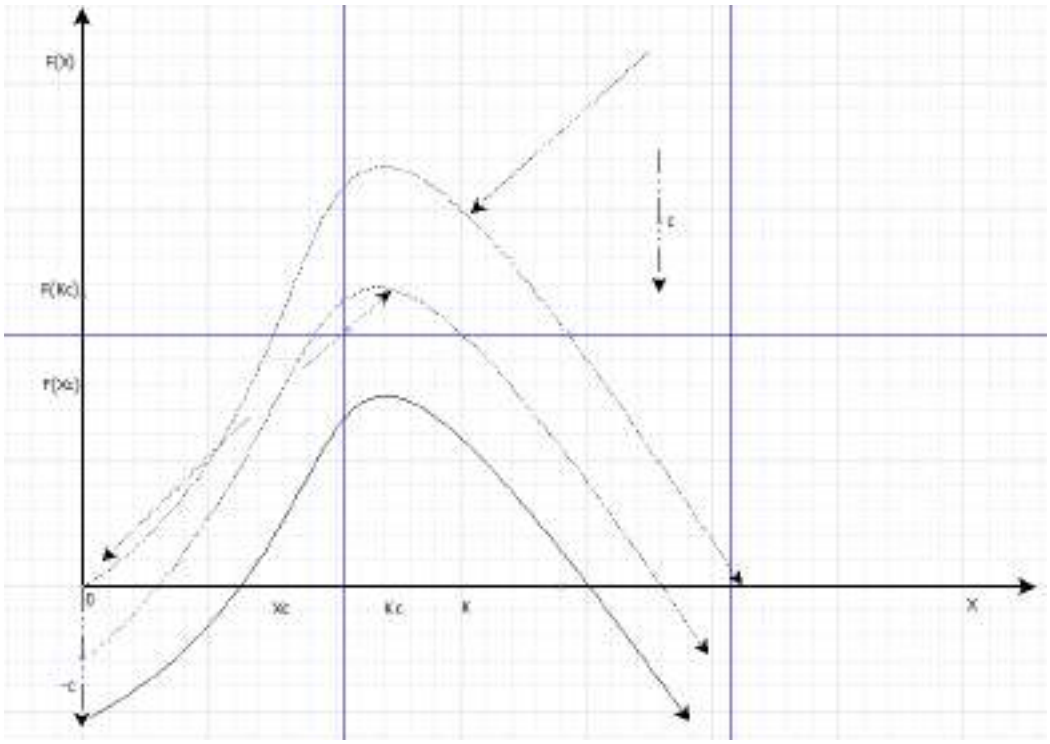
Both solutions are real and positive for $c < \gamma^2/(4\lambda)$. At one not excessive level of enterprise population reduction c, there will therefore be two points of equilibrium: X_c and K_c: the first is unstable and the second is stable.

The instability of X_c is motivated by the fact that, when the value of the stock $X(t)$ over a certain period t, is located in a right hand interval of X_c – or rather, $X(t)$ will be slightly greater than X_c – then in the next interval it will tend to move further away: being $F(X) > X$ to the right of X_c, so then $X(t+1) > X(t)$.

If, vice versa, $X(t)$ were to be slightly less than X_c, that is to say if it was in a left hand interval, since $F(X) < X$ is to the left of X_c, then $X(t+1) < X(t)$ will be the result, and here too, a further departure will be observed. X_c is therefore also a *survival* threshold value, since if the value of the stock $X(t)$ were to, at a specific instant, become less at point X_c – due to an external shock, that is to say an immediate and idiopathic restriction of the credit granted – the system would begin to move towards negative values and once again towards the (theoretical) extinction of the population of the enterprises within a certain '$t + n$'. On the other hand, if an unfavourable contraction, even though reducing the stock $X(t)$, were to not

cause it to fall below the point X_c, and hence $X(t) > X_c$, the system would be spontaneously attracted towards the stable equilibrium point K_c.
Furthermore, since K_c is in any case lower than the carrying capacity of the population of enterprises K (equal to γ/λ), it can be noted that a credit crunch, even if controlled, would take the system to a stable equilibrium value of the population of enterprises, less than that which would have been the case in the total absence of funding retention strategies. In this scenario, the two equilibrium points obviously depend on the height c: whose rise will show an increase in the value of X_c and the decrease in the value of K_c. The system becomes more fragile, since as the survival threshold increases the natural equilibrium value decreases. In $c = \gamma^2/(4\lambda)$ there is the overlap of the two points X_c and K_c, and in these the parabola becomes tangent to the bisector. Once the value of c has been surpassed, which is an additional point of bifurcation, we are again in a situation where the only final evolution of the model is extinction.

Figure 11



By activating similar forms of control, based on the identification of a loss share-threshold for individuals-enterprises, following a restrictive credit strategy by the banks – in the absence of structural or exogenous crises – a stable equilibrium of the system is achieved, albeit with a density lower than the carrying capacity of the population observed in the absence of contractions of financial leverage. This is on the condition that the stock X does not fall below a certain threshold X_c, that is, that the height c does not pass the bifurcation point.

4. Is there therefore an optimal minimum level of disbursed financial leverage? A second dynamic model

Wanting to take another step forward – to offer another point of view – we will now reconsider the introduction – expanding it slightly – for which, at adequate levels of credit

leverage granted, the population of companies will grow by $\frac{dZ_2}{dt} = GZ_2$ with $G > 0$. Vice versa – taking the abstraction to an extreme – if the population of companies were to disappear as a result of an unsustainable financial squeeze, the banks would go into default. So if $Z_2 = 0$, $\frac{dZ_1}{dt} = -HZ_1$ with $H > 0$.

The number of (positive) transactions between banks and businesses is clearly proportional to the product of their populations that are *performing*. We have said that a credit retention strategy produces, in an immediate future[17], a positive effect on banks' balance sheets – namely $\alpha Z_2 Z_1$ – whilst, instead, this burdens enterprises with an outcome on the number, which is expressed by the relation $-\beta Z_2 Z_1$. Even α and β are positive constants and measure the immediate result of the interaction – in terms of credit disbursed/received – for the two populations.

As a result of these assumptions, the previous equations will undergo a formal development:

$$\frac{dZ_2}{dt} = GZ_2 - \beta Z_2 Z_1 \quad \text{and} \quad \frac{dZ_1}{dt} = -HZ_1 + \alpha Z_2 Z_1 \qquad [14]$$

The objective is now – once the positive starting values have been arbitrarily set (at the time t) for the two populations Z_1 and Z_2 – to identify and define the qualitative behaviour of the trajectories of the system. [18]

Solving

$$\begin{cases} Z_2(G - \beta Z_1) = 0 \\ Z_1(-H + \alpha Z_2) = 0 \end{cases} \qquad [15]$$

We will find the critical points $A(0,0)$ and $B\left(\frac{H}{\alpha}, \frac{G}{\beta}\right)$

Near point A the corresponding linear system will be:

$$\frac{d}{dt}\begin{pmatrix} Z_2 \\ Z_1 \end{pmatrix} = \begin{pmatrix} G & 0 \\ 0 & -H \end{pmatrix}\begin{pmatrix} Z_2 \\ Z_1 \end{pmatrix} \qquad [16]$$

With the eigenvalues and eigenvectors:

$$\rho_1 = G, \qquad \kappa_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\rho_2 = -H, \qquad \kappa_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

And solution:

$$\begin{pmatrix} Z_2 \\ Z_1 \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{Gt} + C_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-Ht} \qquad [17]$$

From which it may be seen how the point $A(0,0)$ – the origin – is a saddle point, and therefore unstable: the entry is an ordinate and all the different trajectories move away from the critical point.

In the same way, we shall proceed with an observation of the surroundings of the point B, setting

$$Z_2 = \left(\frac{H}{\alpha}\right) + w \quad \text{and} \quad Z_1 = \left(\frac{G}{\beta}\right) + v \qquad [18]$$

Therefore

$$\frac{d}{dt}\begin{pmatrix} w \\ v \end{pmatrix} = \begin{pmatrix} 0 & -\beta H/\alpha \\ \alpha G/\beta & 0 \end{pmatrix}\begin{pmatrix} w \\ v \end{pmatrix} \qquad [19]$$

[17] At least as long as the client enterprises are able to remain in the market and properly repay their quotas of the borrowed capital and the related finance charges.

[18] An (extensive) application of the Lotka-Volterra equations is proposed. *Ex multis,* cf. Hernàndez-Bermejo & Fairén (1997), Costanzini (2009).

With eigenvalues (imaginary) $\rho_{1/2} = \pm i\sqrt{GH}$, which demonstrate the existence of a stable centre in $B\left(\frac{H}{\alpha}, \frac{G}{\beta}\right)$. The system's trajectories will be given by:

$$\frac{d}{dt} = \frac{dv/dt}{dw/dt} = -\frac{\left(\frac{\alpha G}{\beta}\right)w}{\left(\frac{\beta H}{\alpha}\right)v} \qquad [20]$$

or even:

$$\alpha^2 G w\, dw + \beta^2 H v\, dv = 0 \qquad [21]$$

From which:

$$\alpha^2 G w^2 + \beta^2 H v^2 = k \qquad [22]$$

With k constant of integration (not negative).
The solutions can be formalized in:

$$w = \frac{H}{\alpha} J \cos\left(\sqrt{(GH)}t + \psi\right) \quad \text{and} \quad v = \frac{G}{\beta}\sqrt{\frac{H}{G}} J \sin\left(\sqrt{(GH)}t + \psi\right) \qquad [23]$$

where the constants J and ψ are determined by the initial conditions. Replacing in [18], we get:

$$Z_2 = \frac{H}{\alpha} + \frac{H}{\alpha} J \cos\left(\sqrt{(GH)}t + \psi\right) \quad and \quad Z_1 = \frac{G}{\beta} + \frac{G}{\beta}\sqrt{\frac{H}{G}} J \sin\left(\sqrt{(GH)}t + \psi\right) \qquad [24]$$

The trajectories of the corresponding linear system are therefore almost elliptical and move anticlockwise. The equations [24] offer a good approximation of the representation.
Let us return to the non-linear system, reducing it to the equation:

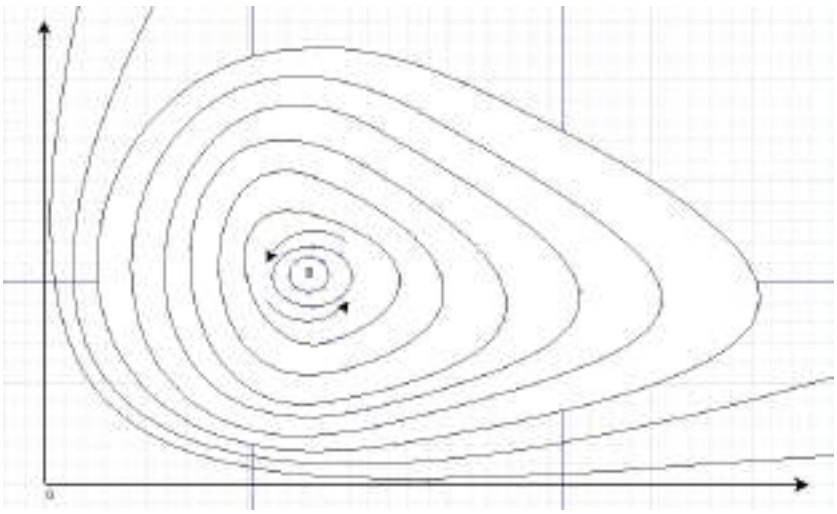
$$\frac{dZ_1}{dZ_2} = \frac{dZ_1/dt}{dZ_2/dt} = -\frac{Z_1(-H + \alpha Z_2)}{Z_2(G - \beta Z_1)} \qquad [25]$$

Which, by separating it, leads to the solution:

$$G \ln(Z_{1)} - \beta Z_1 + H \ln(Z_2) - \alpha Z_2 = q \qquad [26]$$

Setting q – which is also a constant of integration – the graph of the equation is still a curve surrounding the critical point $B\left(\frac{H}{\alpha}, \frac{G}{\beta}\right)$, which is therefore a centre for the non-linear system as well.

Figure 12 - Phase representation of the nonlinear system



Depending on the initial conditions, the trajectories on the stable critical point may show slight variations in the abscissa (Z_2) and in the ordinate (Z_1), or they may be more pronounced and have a greater difference compared to an ellipse. Finally, since Z_2 and Z_1 are periodic functions of t, a non-overlapping oscillatory movement will follow: the curves will intersect at levels of leverage sufficient to support the financial, productive and economic health of enterprises such that their systemic confidence and *performing* presence in the market can be increased. At these points, banks will have the convenience of restricting credit volumes to increase their short-term performance, producing new mortality in the entrepreneurial fabric and an increase in their non-performing loans *in the system*, which will lead them again towards a greater expansion of lending and the reactivation of the cycles.
We reiterate that this model has greater validity for banks that operate mainly in the credit market, since the simultaneous diversification of assets by banks would mitigate the cause-effect mechanism of the proposed model.

Position conclusions

Therefore, the preservation of the bank-enterprise relationship seems necessary, for the maintenance of the macrosystem. In particular, as far as we are concerned, the stabilization of the levels of leverage in the productive and entrepreneurial sectors should be consistent with the dynamic models outlined above. At the very least, the recursive phases highlighted above should also be guaranteed through regulatory standardization. This becomes more significant in fragile and/or partially compromised economic scenarios or when unfavourable economic conditions prevail.
Getting back to the Italian case, specifically to the measures enacted to support the economy made necessary by the current SARS-COVID-19 emergency, we will be analysing the lines of intervention dictated by the Italian government and we will briefly see how these are activated with the main objective of keeping the bank-enterprise system in working order.
It should first of all be pointed out that – in all evidence – the 2020 economic crisis has arisen from exogenous elements. Therefore, it was not caused by internal *pathologies* within the same economy (as was true, for example, of the 2009-2011 crisis). This consideration will assume particular importance during the management of the resolution of the current crisis. This must be undertaken first to ensure temporary compensation for the sudden cessation of productive activities, and therefore allaying concerns about restoring economic and financial features that have been damaged. The liquidity so far allocated coming to Euro 350 billion (leverage on nominal € 25 billion) plus Euro 400[19] billion (again, as a result of leverage), was not directly injected into the production system (or households), with the "helicopter money" rationale or through other forms however direct they may have been. Instead, the intervention of the Italian government was aimed at strengthening the public guarantee funds, leaving the role of financial intermediation by the banks unchanged and at the same time encouraging the latter to undertake quantitative expansion of credit disbursed.
Indeed, both recent Decree-Laws, No. 18 and No. 23 of 2020, expressly addressed all resources to direct public guarantees, immediately enforceable and with zero weighting on provisions, as well as to reinsurance for the world of credit guarantee consortia, with free access and coverage up to 100% (for small loans), and in any case between 80% and 90% for most financing operations.

[19] Our intention here is not to discuss the merits of macroeconomic arguments for stability, public deficit or European Community loan instruments – bonds or no bonds – so that the State can raise these resources.

These measures - assessed here only for the aspects pertinent to this paper - should be able to impede a further credit crunch, which would certainly have manifested itself in the face of the generalised worsening of creditworthiness brought about by the lockdown, while avoiding critical developments in the tendency towards default by the populations of enterprises and banks in the nation.

References

ABI-CERVED (2018 and 2019). Outlook on non-performing loans. In 'Mercati, Analisi, Scenario'. https://www.abi.it/Pagine/Mercati/Analisi/Scenario-e-previsioni/Outlook-ABI-Cerved-su-sofferenze-imprese.aspx.

BERNANKE, B.S., GERTLER, M. and GILCHRIST, S. (1996). The Financial Accelerator and the Flight to Quality. *The Review of Economics and Statistics*, 78/1, 1-15.

BISCHI, G. I., LAMANTIA, F. and SBRAGIA L. (2005). Modelli matematici e politiche per la gestione della pesca. In ANTONELLI, G., BISCHI, G. I. and VIGANÒ, E. *La sostenibilità nel settore della pesca. Modelli, politiche, esperienze in un'area del litorale romagnolo-marchigiano* (33-63). Milano: FrancoAngeli.

BUERA, F.J., FATTAL-JAEF, R.N. and SHIN, Y. (2015). Anatomy of a credit crunch: From capital to labor markets. *Review of Economic Dynamics*, 18/1, 101-117.

CONTI, G. (2016). Matematica e rischio di credito. http://www.mathisintheair.org/wp.

COSTANZINI, S. (2009). *Analisi dei sistemi dinamici e della loro stabilità nelle applicazioni economiche*. [Unpublished manuscript]. Department of Mathematics, University of Bologna. Retrieved from https://www.dm.unibo.it/~fioresi/tesi/TesiCostanzini.pdf

CROSS, R. B. (1993). On the foundations of hysteresis in economic systems. *Economics and Philosophy*, 9, 53–74.

DESOGUS, M. and CASU, E. (2020). A contribution on relationship banking. Economic, anthropological and mathematical reasoning, empirical evidence from Italy. *International Research Journal of Finance and Economics*, 178, 25-49.

DESOGUS, M. and VENTURI, B. (2019). Bank Crashes and Micro Enterprise Loans. *International Journal of Business and Social Science*, 12, 35-53.

DÜLLMANN, K. and KICK, T. (2014). Stress testing German banks against a global credit crunch. *Financial Markets and Portfolio Management*, 28, 337-361.

ELSTER, J. (1976). A note on hysteresis in the social sciences. *Synthèse*, 33, 371–391.

ECB - European Central Bank (2017). Guidance to banks on non-performing loans. https://www.bankingsupervision.europa.eu/ecb/pub/pdf/guidance_on_npl.en.pdf

FONG, H.G. and VASICEK, O.A. (1984). A Risk Minimizing Strategy for Portfolio Immunization. *The Journal of Finance*, 39(5), 1541-6.

GOODWIN, R. M. (1951). The nonlinear accelerator and the persistence of business cycles. *Econometrica*, 19, 1–17.

GUASTELLO, S. J. (1995). Chaos, catastrophe, and human affairs: Applications of nonlinear dynamics to work, organizations, and social evolution. Mahwah, NJ: Lawrence Erlbaum Associates.

HERNÀNDEZ-BERMEJO, B. and FAIRÉN, V. (1997). Lotka-Volterra representation of general nonlinear systems. *Mathematical Biosciences*, *140*, 1–32.

HOLLING, C. S. (1973). Resilience and stability of ecological systems. *Annual Review of Ecology and Systematics*, 4, 1–24.

IYER, R., PEYDRÓ, J.L., DA ROCHA LOPEZ, S. and SCHOAR, A. (2014). Interbank Liquidity Crunch and the Firm Credit Crunch: Evidence from the 2007–2009 Crisis. *The Review of Financial Studies*, 27/1, 347-372.

KUZNETSOV, YU., MURATORI, S. and RINALDI, S. Bifurcations and chaos in a periodic predator prey model. *International Journal of Bifurcation and Chaos,* 2, 117–128 (1992)

ORTON, P., ANSELL, J. and ANDREEVA, G. (2014). Exploring the performance of small- and medium-sized enterprises through the credit crunch. *Journal of the Operational Research Society*, 66/4, 657-663.

PERKO L. (2001). Differential Equations and Dynamical Systems. New York: Springer.

PETROSKY-NADEAU N. (2013). TFP during a credit crunch. *Journal of Economic Theory*, 148/3, 1150-1178.

PUU, T. (2000). Attractors, bifurcations, and chaos. Heidelberg: Springer-Verlag.

RINALDI, S., MURATORI, S. and KUZNETSOV, YU. (1993). Multiple attractors, catastrophes and chaos in seasonally perturbed predator-prey communities. *Bulletin* of *Mathematical Biology,* 55, 15–35.

ROSSER, J.B. Jr. (1991). From catastrophe to chaos: A general theory of economic discontinuities. Boston: Kluwer Academic Publishers (2nd edition, vol. I, 2000).

SAEZ, E. and ZUCMAN, G. (2020). Keeping business alive: the government as buyer of last resort, Mimeo.

VENTURI, B., CASULA, G., DESOGUS, M. and PILI, A. (2020). Mathematical tools in economic and financial models. Rome: Aracne.

WEHINGER, G. (2014). SMEs and the credit crunch. *OECD Journal: Financial Market Trends*, 115-148.

Appendix

***Dataset 1**: Loans to enterprises in Italy, in million euros*

Feb-12	1,867,069	**Feb-14**	1,748,482	**Feb-16**	1,653,726	**Feb-18**	1,538,125
Mar-12	1,843,099	**Mar-14**	1,741,682	**Mar-16**	1,648,836	**Mar-18**	1,528,030
Apr-12	1,855,380	**Apr-14**	1,734,242	**Apr-16**	1,638,352	**Apr-18**	1,528,992
May-12	1,848,450	**May-14**	1,719,438	**May-16**	1,650,041	**May-18**	1,531,149
Jun-12	1,839,294	**Jun-14**	1,730,636	**Jun-16**	1,652,764	**Jun-18**	1,468,328
Jul-12	1,842,289	**Jul-14**	1,737,299	**Jul-16**	1,644,173	**Jul-18**	1,471,657
Aug-12	1,824,729	**Aug-14**	1,714,682	**Aug-16**	1,637,576	**Aug-18**	1,453,087
Sep-12	1,814,339	**Sep-14**	1,723,485	**Sep-16**	1,635,688	**Sep-18**	1,449,988
Oct-12	1,816,006	**Oct-14**	1,714,305	**Oct-16**	1,634,588	**Oct-18**	1,445,756
Nov-12	1,822,061	**Nov-14**	1,709,791	**Nov-16**	1,639,756	**Nov-18**	1,452,218
Dec-12	1,803,779	**Dec-14**	1,690,082	**Dec-16**	1,618,884	**Dec-18**	1,412,341
Jan-13	1,807,279	**Jan-15**	1,694,068	**Jan-17**	1,627,127	**Jan-19**	1,412,799
Feb-13	1,804,572	**Feb-15**	1,686,081	**Feb-17**	1,626,021	**Feb-19**	1,407,968
Mar-13	1,784,326	**Mar-15**	1,694,929	**Mar-17**	1,622,248	**Mar-19**	1,383,853
Apr-13	1,779,044	**Apr-15**	1,688,135	**Apr-17**	1,610,564	**Apr-19**	1,389,698
May-13	1,771,135	**May-15**	1,679,991	**May-17**	1,614,943	**May-19**	1,385,884
Jun-13	1,757,391	**Jun-15**	1,694,540	**Jun-17**	1,591,435	**Jun-19**	1,370,106
Jul-13	1,761,553	**Jul-15**	1,693,195	**Jul-17**	1,554,869	**Jul-19**	1,374,303
Aug-13	1,737,363	**Aug-15**	1,675,366	**Aug-17**	1,534,324	**Aug-19**	1,349,538
Sep-13	1,737,088	**Sep-15**	1,679,762	**Sep-17**	1,524,810	**Sep-19**	1,346,075
Oct-13	1,725,214	**Oct-15**	1,659,496	**Oct-17**	1,527,558	**Oct-19**	1,337,474
Nov-13	1,712,062	**Nov-15**	1,680,464	**Nov-17**	1,528,311	**Nov-19**	1,334,656
Dec-13	1,706,800	**Dec-15**	1,659,189	**Dec-17**	1,531,909	**Dec-19**	1,312,601
Jan-14	1,755,443	**Jan-16**	1,656,962	**Jan-18**	1,531,414	**Jan-20**	1,325,769

***Dataset 2**: non performing loans in Italy, in million euros*

Mar-12	80,372	**Mar-14**	125,350	**Mar-16**	147,868	**Mar-18**	125,784
Jun-12	85,174	**Jun-14**	130,275	**Jun-16**	149,683	**Jun-18**	99,454
Sep-12	88,628	**Sep-14**	133,524	**Sep-16**	151,241	**Sep-18**	92,281
Dec-12	93,420	**Dec-14**	136,323	**Dec-16**	154,034	**Dec-18**	73,545
Mar-13	97,330	**Mar-15**	140,097	**Mar-17**	150,485	**Mar-19**	67,456
Jun-13	103,642	**Jun-15**	145,662	**Jun-17**	150,251	**Jun-19**	66,077
Sep-13	108,895	**Sep-15**	149,286	**Sep-17**	133,973	**Sep-19**	62,205
Dec-13	117,511	**Dec-15**	151,423	**Dec-17**	128,586		

Dataset 3: *total number of enterprises in Italy*

Year	2012	2013	2014	2015	2016	2017	2018
B: extraction of minerals from quarries and mines	2,451	2,336	2,257	2,186	2,250	2,317	2,332
C: manufacturing	417,306	407,344	396,422	389,317	399,458	404,529	406,508
D: supply of electricity, gas, steam and air conditioning	8,926	10,169	10,459	10,775	10,015	10,041	10,056
E: supply of water, sewerage, waste management and environmental remediation services	8,967	9,121	9,146	9,231	9,060	9,230	9,301
F: construction	572,412	549,846	529,103	511,405	534,824	537,348	537,853
G: wholesale and retail trade, repair of motor vehicles and motorcycles	1,163,413	1,153,640	1,123,134	1,105,227	1,128,117	1,129,702	1,130,596
H: transport and storage	131,755	129,865	125,688	123,625	127,651	127,817	128,172
I: accommodation and food service businesses	307,878	313,207	312,013	315,464	312,000	311,478	312,009
J: information and communications services	97,280	95,989	96,997	98,381	96,933	96,915	97,079
K: financial and insurance service businesses	91,434	93,031	95,209	96,173	93,199	93,341	93,394
L: real estate businesses	235,434	243,564	239,134	238,273	237,137	237,094	237,067
M: professional, scientific and technical businesses	710,017	691,700	705,895	714,934	700,468	700,308	700,406
N: rental and travel agencies, business support services	143,770	139,362	139,898	139,595	139,959	140,415	140,724
P: education	26,890	27,677	29,088	29,566	28,360	28,256	28,304
Q: healthcare and social services	259,400	261,056	277,295	285,231	269,170	269,050	269,191
R: arts, sports, entertainment and amusement businesses	63,054	62,704	64,169	65,022	63,165	63,350	63,404
S: other service businesses	202,065	199,902	203,180	203,680	200,831	200,794	200,857
TOTAL	**4,442,452**	**4,390,513**	**4,359,087**	**4,338,085**	**4,352,597**	**4,361,985**	**4,367,253**

Dataset 4: *number of micro enterprises in Italy*

Year	2012	2013	2014	2015	2016	2017	2018
B: extraction of minerals from quarries and mines	1,907	1,850	1,775	1,712	1,796	1,795	1,795
C: manufacturing	345,293	338,015	328,486	321,837	330,613	330,526	330,459
D: supply of electricity, gas, steam and air conditioning	8,380	9,610	9,916	10,205	9,448	9,445	9,443
E: supply of water, sewerage, waste management and environmental remediation services	6,485	6,688	6,748	6,816	6,628	6,626	6,625
F: construction	548,709	528,592	509,648	492,388	515,477	515,341	515,237
G: wholesale and retail trade, repair of motor vehicles and motorcycles	1,124,546	1,116,087	1,086,631	1,068,659	1,089,768	1,089,481	1,089,262
H: transport and storage	119,126	117,430	113,241	110,756	114,173	114,143	114,120
I: accommodation and food service businesses	288,119	294,007	292,996	295,706	290,253	290,177	290,119
J: information and communications services	91,274	89,895	91,020	92,279	90,353	90,329	90,311
K: financial and insurance service businesses	88,998	90,637	92,831	93,799	90,799	90,775	90,757
L: real estate businesses	234,738	242,874	238,492	237,637	236,437	236,374	236,327
M: professional, scientific and technical businesses	702,053	683,778	698,154	707,020	691,902	691,720	691,581
N: rental and travel agencies, business support services	132,452	128,082	128,721	128,394	128,327	128,294	128,268
P: education	25,239	25,957	27,351	27,781	26,359	26,352	26,347
Q: healthcare and social services	253,160	254,655	270,894	278,646	262,123	262,054	262,001
R: arts, sports, entertainment and amusement businesses	60,658	60,382	62,001	63,011	60,997	60,981	60,969
S: other service businesses	198,593	196,542	199,755	200,185	197,103	197,051	197,011
TOTAL	**4,229,730**	**4,185,081**	**4,158,660**	**4,136,831**	**4,142,556**	**4,141,465**	**4,140,633**

Generalised univariable fractal interpolation functions

Vasileios Drakopoulos[a]*, **DuYong Pak**[b], **SongIl Ri**[b]

[a] Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia, Greece

[b] Department of Mathematics, University of Science, Pyongyang, D.P.R. of Korea

Abstract

We show how to construct a generalised iterated function system whose graph is the attractor, a fractal set, of some continuous function which interpolates a given set of data. Moreover, Rakotch contractions and vertical scaling factors as (continuous) 'contraction functions' are used in order to obtain generalised fractal interpolation functions with extensive practical applications, including data fitting and approximation of functions. A special generalised fractal interpolation function is introduced as an explicit illustrative example to show the effectiveness of the proposed method as compared to other existing methods. In particular, fractal interpolation functions which are widely presented in the literature can be obtained as particular cases of our construction.

Mathematical Subject Classification: 37C45; 28A80; 37L30

Keywords: Iterated function system (IFS); Fractal interpolation function (FIF); Rakotch contraction; function vertical scaling factors.

1 Introduction

The concept of fractal interpolation functions (FIFs) was introduced by Barnsley [1] on the basis of the theory of iterated function systems (IFSs). Barnsley defined a fractal interpolation function (FIF) in 1986 and presented a construction of fractal functions by fractal interpolation. In the developments of theory of FIFs, many researchers have generalized the notion of FIFs in different ways. The fractal interpolation functions have been discussed in detail in the literature (see [1]-[10]). Fractal interpolation functions have become a powerful tool for modeling many natural objects and have wide applications in mathematics

*Corresponding author at: Department of Computer Science and Biomedical Informatics, University of Thessaly, Lamia, Greece E-mail address: vdrakop@uth.gr (V. Drakopoulos).

and several other areas of applied sciences. For example, the fractal interpolation functions have been widely used in approximation theory, image compression, computer graphics and modeling of natural functions (or surfaces) such as rocks, metals, terrains and so on. As we know, a fractal interpolation function is generated by an IFS that consists of a finite set of some continuous functions on a complete metric space. Vertical scaling factors in the continuous functions have a decisive influence on the shape of the corresponding FIF because the vertical scaling factors uniquely determine the corresponding FIF provided that the interpolation points are prescribed in advance (see [12]).

How to construct fractal functions (rough functions) and analyse their complexity has become one of the most important topics in fractals (see [10]). The graph of a fractal interpolation function is an attractor of some iterated function system (see [1]). The concept of iterated function system was introduced as a natural generalization of the well-known Banach contraction principle (see [1], see [12], cf. [10]). Iterated function systems have become powerful tools for construction and analysis of new fractal interpolation functions. In order to ensure more flexibility in modeling natural shapes and phenomena or in image processing, researchers proposed many types of fractal interpolation functions by using iterated function systems (see [1], [12], [2], [10]).

The connectivity of attractors of iterated function systems is very important in the construction of fractal interpolation functions. The graphs of linear one variable fractal interpolation functions are always continuous functions. In usual approaches, the existence of linear FIFs follows from Banach's fixed point theorem (see [1]). Furthermore, in almost all the papers, the various types of FIFs are limited within the cases of constant vertical scaling factors and Banach's fixed point theorem. In 2011, to get the FIFs with more flexibility and diversity in a more general sense, Wang and Fan introduced a natural generalization of Barnsley's affine fractal interpolation function by using special function vertical scaling factors and Banach's fixed point theorem (see [**?**]). In order to construct new iterated function systems and fractal interpolation functions, one can use the well-known fixed point results obtained in the fixed point theory (see [**?**], [5], [9], [10]). As far as we know, the first significant generalization of Banach's principle was obtained by Rakotch in 1962 (see [4], p.124).

In 2017, Ri presented a method to generate generalized FIFs by using the Rakotch fixed point theorem ([7]) instead of the Banach fixed point theorem in certain concrete case (see [10]). However, results of [10] do not directly apply to the general case which often occurs in practical applications. In fact, in general case, the methods of proof of results in [10] fail because generalized transformations can involve special function vertical scaling factors that are not constant vertical scaling factors. The results of [**?**] and [10] inspire us to find possible vertical scaling factors (not necessarily constant vertical scaling factors) and contractions (not necessarily Banach contractions) for the existence of generalized FIFs. In this paper, in order to obtain generalized fractal interpolation functions, we use Rakotch contractions and special function vertical scaling factors. Dealing with generalized fractal interpolation functions is better than the one provided by [10]. In particular, we give an explicit illustrative

example to demonstrate the effectiveness of obtained results.

This paper is organized as follows. In Section 2 we recall some results needed in constructing general FIFs. In Section 3 we introduce a new type of IFSs that will be used in our discussion for a special class of FIFs with function vertical scaling factors. In Section 4 we give a generalized FIF with special function vertical scaling factors as the fixed point of certain Read-Bajraktarević operator. In Section 5 we ensure that a generalized IFS with special function vertical scaling factors has a unique invariant set. In Section 6 we give an explicit illustrative example to demonstrate the effectiveness of the preceding theory. Finally, in Section 7 we draw our conclusions.

2 Preparatory facts

In this section, we describe some basic notions and theorems on fixed point theory. The following results will be the key in the proof of our main results.

Definition 2.1. *(see [?], p.100, see [7], see [4], p.144) (1) If for some function* $\varphi : (0, +\infty) \to (0, +\infty)$ *and a self-map* f *of a metric space* (X, d)*, we have*

$$\forall_{x,y\in X} \quad d(f(x), f(y)) \leq \varphi(d(x, y)),$$

then we say that f *is a* φ*-contraction. (2) If* f *is a* φ*-contraction for some function* $\varphi : (0, +\infty) \to (0, +\infty)$ *such that for any* $t > 0$*,* $\alpha(t) := \frac{\varphi(t)}{t} < 1$ *and the function* $(0, +\infty) \ni t \to \frac{\varphi(t)}{t}$ *is non-increasing, then we call such a function a Rakotch contraction.*

Remark 2.2. *(see [4], p.144, diagram) Each Banach contraction is a Rakotch contraction, since a map* $f : X \to X$ *is a Banach contraction iff it is a* φ*-contraction for a function* $\varphi(t) = \alpha t$*, for some* $0 \leq \alpha < 1$*.*

Theorem 2.3. *(see [7], cf. [?], cf. [5], cf. [4]) (1) Let* X *be a complete metric space and* $f : X \to X$ *be a Rakotch contraction. Then there is a unique fixed point* $k \in X$ *of* f*, and for each* $x \in X$*,*

$$\lim_{n\to+\infty} f^n(x) = k.$$

(2) Let X *be a complete metric space and* $\{X; f_1, \cdots, f_N\}$ *be an iterated function system consisting of Rakotch contractions. Then there is a unique non-empty compact set* $K \subset X$ *such that*

$$K = \bigcup_{i=1}^{N} f_i(K).$$

Now we describe some basic results on fractal interpolation theory. Let N be a positive integer greater than one and $I := [x_0, x_N] \subset \mathbb{R}$. Let a set of data points $\{(x_i, y_i) \in I \times \mathbb{R} : i = 0, 1, 2, \cdots, N\}$ be given, where $\{x_0, x_1, \cdots, x_N\}$ is a partition of I (i.e., $x_0 < x_1 < x_2 < \cdots <$

x_N) and $y_0, y_1, \cdots, y_N$ are given real numbers. Set $I_i := [x_{i-1}, x_i] \subset I$ and let $l_i \colon I \to I_i$ for $i = 1, 2, \cdots, N$ be contractive homeomorphisms such that

$$l_i(x_0) = x_{i-1}, l_i(x_N) = x_i,$$
$$|l_i(x') - l_i(x'')| \leq \lambda |x' - x''| \ \ whenever \ \ x', x'' \in I$$

for some $0 \leq \lambda < 1$. Furthermore, let mappings $F_i : I \times \mathbb{R} \to \mathbb{R}$ be continuous with, for some $k \geq 0$ and $0 \leq \alpha < 1$,

$$F_i(x_0, y_0) = y_{i-1}, F_i(x_N, y_N) = y_i.$$
$$|F_i(x', y') - F_i(x'', y'')| \leq k|x' - x''| + \alpha |y' - y''|$$

for all $x', x' \in I$, $y', y'' \in \mathbb{R}$, and $i = 1, 2, \cdots, N$.

Now define functions $w_i \colon I \times \mathbb{R} \to I \times \mathbb{R}$ for $i = 1, 2, \cdots, N$ by

$$w_i := (l_i(x), F_i(x, y)).$$

Barnsley presented the following famous result.

Theorem 2.4. *(cf. [1], p.306) The IFS$\{I \times \mathbb{R}, w_i : i = 1, 2, \cdots, N\}$ defined above has a unique nonempty compact set $G \subset \mathbb{R}^2$ such that*

$$G = \bigcup_{i=1}^{N} w_i(G).$$

Then G is the graph of a continuous function $f : I \to \mathbb{R}$ which obeys

$$f(x_i) = y_i \quad for \quad i = 0, 1, \cdots, N.$$

The function f whose graph is the attractor of an IFS is called a fractal interpolation function (FIF) corresponding to the data $\{(x_i, y_i) : i = 0, 1, \cdots, N\}$ (cf. [1], p.306).

Remark 2.5. *In accordance with the idea of Barnsley, researchers proposed many types of FIFs. In [1], [12] and [10],*

$$l_i(x) := \frac{x_i - x_{i-1}}{x_N - x_0} x + \frac{x_N x_{i-1} - x_0 x_i}{x_N - x_0}.$$

(1) In [1], [12] and [2], the maps $w_i(x, y)$ are chosen so that functions $F_i(x, y)$ are Banach contractions with respect to the second variable.
In the affine fractal interpolation function (cf. [1], p.308),

$$F_i(x, y) := c_i x + d_i y + f_i,$$

where $|d_i| < 1$, and in the fractal interpolation function with function vertical scaling factors (see [?], see [12], p.3-4, cf. [2]),

$$F_i(x, y) := d_i(x)y + q_i(x),$$

where $\sup_{x\in I} |d_i(x)| < 1$. (2) In [10], one type of fractal interpolation functions is considered, where the maps $w_i(x, y)$ are chosen so that functions $F_i(x, y)$ are Rakotch contractions with respect to the second variable. In the nonlinear fractal interpolation function (see [10]),

$$F_i(x, y) := c_i x + s_i(y) + f_i,$$

where s_i is a Rakotch contraction.

3 A certain generalised IFS

In this section, we introduce a new type of IFSs that will be used in our discussion for a special class of FIFs with function vertical scaling factors. Barnsley's functional condition for existence of a fractal interpolation function can be replaced by another functional conditions (see [10]). In order to obtain a new generalized fractal interpolation function, we use Rakotch contractions and special function vertical scaling factors in the construction of a generalized IFS (cf. [10], cf. [12]).

Let N be a positive integer greater than one and $I := [x_0, x_N] \subset \mathbb{R}$. We will work in the complete metric space $I \times \mathbb{R}$, with the Euclidean metric d_0. Let a set of data points $\{(x_i, y_i) \in I \times \mathbb{R} : i = 0, 1, 2, \cdots, N\}$ be given, where $x_0 < x_1 < x_2 < \cdots < x_N$ and $y_0, y_1, y_2, \cdots, y_N \in \mathbb{R}$. Set $I_i := [x_{i-1}, x_i] \subset I$ and define contractive homeomorphisms $l_i : I \to I_i$ by

$$l_i(x) := a_i x + e_i,$$

where for all $i = 1, 2, \cdots, N$, the real numbers a_i, b_i are chosen to ensure that $l_i(I) = I_i$. Let $\varphi : (0, +\infty) \to (0, +\infty)$ be a non-decreasing continuous function such that for any $t > 0$, $\alpha(t) := \frac{\varphi(t)}{t} < 1$ and the function $(0, +\infty) \ni t \to \frac{\varphi(t)}{t}$ is non-increasing. Let $d_i \colon I \to \mathbb{R}$ be a continuously differentiable function such that

$$\max_{x\in I} |d_i(x)| \leq 1.$$

Then by the differential mean value theorem and the existence theorem of maximum value and minimum value of continuous function, we can see that for some $L_{d_i} > 0$,

$$|d_i(x') - d_i(x'')| \leq L_{d_i}|x' - x''|,$$

where $x', x'' \in I$. Hence d_i is Lipschitz function defined on I satisfying $\max_{x\in I} |d_i(x)| \leq 1$. Consider an IFS of the form $\{I \times \mathbb{R}; w_i, i = 1, 2, \cdots, N\}$ in which the maps are generalized transformations of the special structure

$$w_i \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} l_i(x) \\ F_i(x, y) \end{pmatrix} = \begin{pmatrix} a_i x + e_i \\ c_i x + d_i(x)s_i(y) + f_i \end{pmatrix},$$

where the transformations are constrained by the data according to

$$w_i \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix}, w_i \begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$$

for $i = 1, 2, \cdots, N$, and s_i are some Rakotch contractions (with the same function φ). Then for all $(x, y'), (x, y'') \in I \times \mathbb{R}$,

$$\begin{aligned} |F_i(x, y') - F_i(x, y'')| &= |d_i(x)||s_i(y') - s_i(y'')| \\ &\leq |s_i(y') - s_i(y'')| \leq \varphi(|y' - y''|). \end{aligned}$$

That is, each $w_i(x, y)$ is chosen so that function $F_i(x, y)$ is Rakotch contraction with respect to the second variable. Also, analytically, we obtain (compare with a_i, e_i, c_i, f_i of [10]).

$$\begin{aligned} a_i &= \frac{x_i - x_{i-1}}{x_N - x_0}, \\ e_i &= \frac{x_N x_{i-1} - x_0 x_i}{x_N - x_0} \\ c_i &= \frac{y_i - y_{i-1}}{x_N - x_0} - \frac{d_i(x_N)s_i(y_N) - d_i(x_0)s_i(y_0)}{x_N - x_0}, \\ f_i &= \frac{x_N y_{i-1} - x_0 y_i}{x_N - x_0} - \frac{x_N d_i(x_0)s_i(y_0) - x_0 d_i(x_N)s_i(y_N)}{x_N - x_0}. \end{aligned}$$

Remark 3.1. *1) Our bivariate function $d_i(x)s_i(y)$ is a generalization of bivariate function $d_i(x)y$ in the fractal interpolation function with function vertical scaling factors (see [**?**], see [12], p.3-4, cf. [2]). In fact, in the case when $0 < \max_{x \in I} |d_i(x)| < 1$ (see [12], p.3), obviously,*

$$d_i(x)y = \frac{d_i(x)}{\max_{x \in I} |d_i(x)|} \max_{x \in I} |d_i(x)|y.$$

Let $s_i(y) := \max_{x \in I} |d_i(x)|y$ and $d_i^*(x) := \frac{d_i(x)}{\max_{x \in I} |d_i(x)|}$. Then $d_i(x)y = d_i^*(x)s_i(y)$, $\max_{x \in I} |d_i^*(x)| = 1$ and s_i is a Banach (or Rakotch) contraction.

2) Our functional condition $\max_{x \in I} |d_i(x)| \leq 1$ is the essential condition to show the difference between Banach contractibility of $F_i(\cdot, y)$ and Rakotch contractibility of $F_i(\cdot, y)$ (compare with [12]). In fact, since $\varphi(t) < t$ for any $t > 0$,

$$\begin{aligned} |F_i(x, y') - F_i(x, y'')| &= |d_i(x)||s_i(y') - s_i(y'')| \\ &\leq \max_{x \in I} |d_i(x)||s_i(y') - s_i(y'')| \\ &\leq \max_{x \in I} |d_i(x)|\varphi(|y' - y''|) \\ &\leq \max_{x \in I} |d_i(x)||y' - y''|, \end{aligned}$$

where $(x, y'), (x, y'') \in \mathbb{R}^2$.
Hence, if $\max_{x\in I} |d_i(x)| < 1$, *as can be seen, notwithstanding each* s_i *is a Rakotch contraction that is not Banach contraction, each* F_i *is Banach contraction with respect to the second variable because*

$$|F_i(x, y') - F_i(x, y'')| \leq \max_{x\in I} |d_i(x)||y' - y''|.$$

On the other hand, if $\max_{x\in I} |d_i(x)| = 1$, *then we can conclude that each* F_i *is Rakotch contraction (that is not Banach contraction) with respect to the second variable whenever each* s_i *is a Rakotch contraction (that is not Banach contraction) because*

$$|F_i(x, y') - F_i(x, y'')| \leq \max_{x\in I} |d_i(x)|\varphi(|y' - y''|).$$

4 Fixed point of a certain operator

In this section, we introduce a generalized FIF with special function vertical scaling factors as the fixed point of certain Read-Bajraktarević operator (see [6]). By using Rakotch fixed point theorem, we show that the graph of a generalized FIF with special function vertical scaling factors is the invariant set of some generalized IFS.

Denote by $C(I)$ the set of continuous functions $f : I = [x_0, x_N] \to \mathbb{R}$. Let $C^*(I) \subset C(I)$ denote the set of continuous functions $f\colon I \to \mathbb{R}$ such that $f(x_0) = y_0$ and $f(x_N) = y_N$, that is,

$$C^*(I) := \{f \in C(I) : f(x_0) = y_0, f(x_N) = y_N\}.$$

Let $C^{**}(I) \subset C^*(I) \subset C(I)$ be the set of continuous functions that pass through the given data points $\{(x_i, y_i) \in I \times \mathbb{R} : i = 0, 1, 2, \ldots, N\}$, that is,

$$C^{**}(I) := \{f \in C^*(I) : f(x_i) = y_i, i = 0, 1, \cdots, N\}.$$

Define a metric $d_{C(I)}$ on $C(I)$ by

$$d_{C(I)}(g, h) := \max_{x\in[x_0,x_N]} |g(x) - h(x)|$$

for all $g, h \in C(I)$. Then $(C(I), d_{C(I)})$, $(C^*(I), d_{C(I)})$ and $(C^{**}(I), d_{C(I)})$ are complete metric spaces. For all $f \in C^*(I)$, define a mapping $T : C^*(I) \to C(I)$ by

$$\begin{aligned} Tf(x) :&= F_i(l_i^{-1}(x), f(l_i^{-1}(x))) \\ &= c_i l_i^{-1}(x) + d_i(l_i^{-1}(x))s_i(f(l_i^{-1}(x))) + f_i \end{aligned}$$

for $x \in [x_{i-1}, x_i]$ and $i = 1, 2, \cdots, N$. Obviously, T is a form of Read-Bajraktarević operator as defined in [6].

Lemma 4.1. $Tf \in C^{**}(I)$ *for all* $f \in C^*(I)$. *That is,* $T : C^*(I) \to C^{**}(I)$ *and* $T^n : C^{**}(I) \to C^{**}(I)$ *for all* $n \geq 2$.

Proof. Since

$$w_i \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} x_{i-1} \\ y_{i-1} \end{pmatrix}, w_i \begin{pmatrix} x_N \\ y_N \end{pmatrix} = \begin{pmatrix} x_i \\ y_i \end{pmatrix}$$

for $i = 1, 2, \cdots, N$, we obtain $l_i(x_0) = x_{i-1}$, $l_i(x_N) = x_i$, $l_i^{-1}(x_{i-1}) = x_0$, $l_i^{-1}(x_i) = x_N$, $F_i(x_0, y_0) = y_{i-1}$ and $F_i(x_N, y_N) = y_i$ for $i = 1, 2, \cdots, N$. Hence if $x_i \in [x_{i-1}, x_i]$ for $i = 1, 2, \cdots, N$, then since $f \in C^*(I)$, we obtain

$$\begin{aligned} Tf(x_i) &= F_i(l_i^{-1}(x_i), f(l_i^{-1}(x_i))) \\ &= c_i l_i^{-1}(x_i) + d_i(l_i^{-1}(x_i)) s_i(f(l_i^{-1}(x_i))) + f_i \\ &= c_i x_N + d_i(x_N) s_i(f(x_N)) + f_i \\ &= F_i(x_N, f(x_N)) = F_i(x_N, y_N) = y_i \end{aligned}$$

and if $x_i \in [x_i, x_{i+1}]$ for $i = 0, 1, 2, \cdots, N-1$, then since $f \in C^*(I)$, we obtain

$$\begin{aligned} Tf(x_i) &= F_{i+1}(l_{i+1}^{-1}(x_i), f(l_{i+1}^{-1}(x_i))) \\ &= c_{i+1} l_{i+1}^{-1}(x_i) + d_{i+1}(l_{i+1}^{-1}(x_i)) s_{i+1}(f(l_{i+1}^{-1}(x_i))) + f_{i+1} \\ &= c_{i+1} x_0 + d_{i+1}(x_0) s_{i+1}(f(x_0)) + f_{i+1} \\ &= F_{i+1}(x_0, f(x_0)) = F_{i+1}(x_0, y_0) = y_i. \end{aligned}$$

So $f(x_i) = y_i$ for all $i = 0, 1, 2, \cdots, N$ and $Tf(x)$ is continuous at each of the points $x_1, x_2, \cdots, x_{N-1}$. By definition of the mapping T, $Tf(x)$ is continuous on the interval $[x_{i-1}, x_i]$ for all $i = 1, 2, \cdots, N$. Hence $Tf \in C^{**}(I)$ and $T^n : C^{**}(I) \to C^{**}(I)$ for all $n \geq 2$. □

Using Lemma 4.1 and the technique introduced in [10], we can obtain the following Theorem that will be used in our discussion for a special class of FIFs with function vertical scaling factors.

Theorem 4.2. *Let N be a positive integer greater than one. Let $\{I \times \mathbb{R}; w_i, i = 1, 2, \cdots, N\}$ denote the IFS defined above, associated with the set of data*

$$\{(x_i, y_i) : i = 0, 1, \cdots, N\}.$$

Then the operator T is a Rakotch contraction (considered as a map $T\colon C^*(I) \to C^*(I)$). Hence there is a unique continuous function $f\colon I \to \mathbb{R}$ which is a fixed point of T. In particular, $f(x_i) = y_i$ for $i = 0, 1, \cdots, N$. Moreover, the graph G of f is invariant with respect to $\{I \times \mathbb{R}; w_1, \cdots, w_N\}$, i.e.,

$$G = \bigcup_{i=1}^{N} w_i(G).$$

Proof. Since $\max_{x\in I}|d_i(x)|\leq 1$, we obtain that for all $g,h\in C^*(I)\subset C(I)$,

$$\begin{aligned}
d_{C(I)}(Tg,Th)&=\max_{x\in[x_0,x_N]}|Tg(x)-Th(x)|\\
&=\max_{i=1,2,\cdots,N}\max_{x\in[x_{i-1},x_i]}|Tg(x)-Th(x)|\\
&=\max_{i=1,2,\cdots,N}\max_{x\in[x_{i-1},x_i]}|c_il_i^{-1}(x)+d_i(l_i^{-1}(x))s_i(g(l_i^{-1}(x)))+f_i\\
&\quad -c_il_i^{-1}(x)+d_i(l_i^{-1}(x))s_i(h(l_i^{-1}(x)))+f_i|\\
&=\max_{i=1,2,\cdots,N}\max_{x\in[x_{i-1},x_i]}|d_i(l_i^{-1}(x))s_i(g(l_i^{-1}(x)))-d_i(l_i^{-1}(x))s_i(h(l_i^{-1}(x)))|\\
&\leq\max_{i=1,2,\cdots,N}\max_{x\in[x_{i-1},x_i]}|s_i(g(l_i^{-1}(x)))-s_i(h(l_i^{-1}(x)))|\\
&\leq\max_{i=1,2,\cdots,N}\sup_{x\in[x_{i-1},x_i]}\varphi(|g(l_i^{-1}(x))-h(l_i^{-1}(x))|),
\end{aligned}$$

where $\varphi\colon(0,+\infty)\to(0,+\infty)$ is some non-decreasing function such that $\varphi(t)<t$ for $t>0$ and $t\to\frac{\varphi(t)}{t}$ is non-increasing. Since $\varphi\colon(0,+\infty)\to(0,+\infty)$ is non-decreasing continuous function and $l_i^{-1}:[x_{i-1},x_i]\to[x_0,x_N]$ for all $i=1,2,\cdots,N$, we obtain that for $i_0\in\{1,2,\cdots,N\}$ and $x_0\in[x_{i_0-1},x_{i_0}]$,

$$\begin{aligned}
\varphi(|g(l_{i_0}^{-1}(x_0))-h(l_{i_0}^{-1}(x_0))|)\leq&\varphi(\max_{x\in[x_{i_0-1},x_{i_0}]}|g(l_{i_0}^{-1}(x))-h(l_{i_0}^{-1}(x))|)\\
\leq&\varphi(\max_{x\in[x_0,x_N]}|g(x)-h(x)|)\\
=&\varphi(d_{C(I)}(g,h)).
\end{aligned}$$

Since x_0 was arbitrary,

$$\sup_{x\in[x_{i_0-1},x_{i_0}]}\varphi(|g(l_{i_0}^{-1}(x)-h(l_{i_0}^{-1}(x))|)\leq\varphi(d_{C(I)}(g,h))$$

and since i_0 was arbitrary,

$$\max_{i=1,2,\cdots,N}\sup_{x\in[x_{i-1},x_i]}\varphi(|g(l_i^{-1}(x)-h(l_i^{-1}(x))|)\leq\varphi(d_{C(I)}(g,h)).$$

Hence we obtain

$$\begin{aligned}
d_{C(I)}(Tg,Th)\leq&\max_{i=1,2,\cdots,N}\sup_{x\in[x_{i-1},x_i]}\varphi(|g(l_i^{-1}(x)-h(l_i^{-1}(x))|)\\
\leq&\varphi(d_{C(I)}(g,h))=\varphi(d_{C(I)}(g,h)).
\end{aligned}$$

So we conclude that $T\colon C^*(I)\to C^{**}(I)\subset C^*(I)$ is a Rakotch contraction (with the same function φ) on the complete metric space $(C^*(I),d_{C(I)})$. Theorem 2.3 (1) implies that T possesses a unique fixed point in $C^*(I)$. That is, there exists a continuous function $f\in C^*(I)$ such that for all $x\in[x_0,x_N]$,

$$Tf(x)=f(x).$$

Since $T : C^*(I) \to C^{**}(I)$ (by Lemma 4.1), we have $f = Tf \in C^{**}(I)$. That is, there is a continuous function f that passes through the given data points $\{(x_i, y_i) \in I \times \mathbb{R} : i = 0, 1, 2, \ldots, N\}$. Let G denote the graph of $f \in C^{**}(I)$, that is, $G := \{(x, f(x)) : x \in [x_0, x_N]\}$. Since f is a fixed point of the operator T and if $x \in [x_{i-1}, x_i]$, then

$$Tf(x) = F_i(l_i^{-1}(x), f(l_i^{-1}(x))),$$

we obtain that for all $x \in [x_0, x_N]$,

$$\begin{aligned} f(l_i(x)) =& Tf(l_i(x)) \\ =& F_i(l_i^{-1}(l_i(x)), f(l_i^{-1}(l_i(x)))) \\ =& F_i(x, f(x)). \end{aligned}$$

Since $w_i(x, y) = (l_i(x), F_i(x, y))$ for all for $i = 1, 2, \cdots, N$, we obtain that

$$\begin{aligned} w_i(G) =& w_i(\{(x, f(x)) : x \in [x_0, x_N]\}) \\ =& \{w_i(x, f(x)) : x \in [x_0, x_N]\} \\ =& \{(l_i(x), F_i(x, f(x))) : x \in [x_0, x_N]\} \\ =& \{(l_i(x), f(l_i(x))) : x \in [x_0, x_N]\} \\ =& \{(x, f(x)) : x \in [x_{i-1}, x_i]\}. \end{aligned}$$

Hence

$$\begin{aligned} G =& \{(x, f(x)) : x \in [x_0, x_N]\} \\ =& \bigcup_{i=1}^{N} \{(x, f(x)) : x \in [x_{i-1}, x_i]\} \\ =& \bigcup_{i=1}^{N} w_i(G). \end{aligned}$$

This completes the proof. □

Remark 4.3. *In the case where the vertical scaling factor parameters are constants, Barnsley investigated the existence of affine FIFs by using the Banach fixed point theorem (see [1]), and Wang and Fan introduced a natural generalization of Barnsley's affine FIFs by using special function vertical scaling factors and Banach's fixed point theorem (see [?]). Here, we study the existence of generalized FIFs with function vertical scaling factors by using the Rakotch fixed point theorem, and the techniques used in Theorem 4.2 is completely different from those used in [1]-[2]. But Theorem 4.2 does not ensure that the IFS$\{I \times \mathbb{R}; w_i, i = 1, 2, \cdots, N\}$ has a unique invariant set. The uniqueness of invariant set is determined explicitly in Theorem 5.1.*

5 Attractor of a certain IFS

In this section, we ensure that a generalized IFS$\{I \times \mathbb{R}; w_i, i = 1, 2, \cdots, N\}$ with special function vertical scaling factors has a unique invariant set (attractor). Theorem 5.1 that is our main theorem in this paper improves upon a result proved by [10], and the proof of Theorem 5.1 is based on arguments first applied in [10]. If we combine the both of Theorem 4.2 and Theorem 5.1 into a party, we can easily know that the graph of a generalized FIF with special function vertical scaling factors is a unique attractor of a certain IFS.

Theorem 5.1. *Let N be a positive integer greater than one. Let each s_i be a bounded Rakotch contraction. Let $\{I \times \mathbb{R}; w_i, i = 1, 2, \cdots, N\}$ denote the IFS defined above, associated with the set of data*

$$\{(x_i, y_i) : i = 0, 1, \cdots, N\}.$$

Then there is a metric d_θ on $I \times \mathbb{R}$, equivalent to the Euclidean metric d_0, such that for all $i = 1, \cdots, N$, w_i are Rakotch contractions with respect to d_θ. In particular, there exists a unique nonempty compact set $G \subset I \times \mathbb{R}$ such that

$$G = \bigcup_{i=1}^{N} w_i(G).$$

Proof. We define a metric d_θ on $I \times \mathbb{R}$ by

$$d_\theta((x', y'), (x'', y'')) := |x' - x''| + \theta|y' - y''|,$$

where θ is a positive real number which is specified below. Since $|d_i(x') - d_i(x'')| \leq L_{d_i}|x' - x''|$ and $F_i(x, y) := c_i x + d_i(x)s_i(y) + f_i$,

$$\begin{aligned}
|F_i(x', y') - F_i(x'', y'')| = & \\
&= |c_i x' + d_i(x')s_i(y') + f_i - (c_i x'' + d_i(x'')s_i(y'') + f_i)| \\
&\leq |c_i||x' - x''| + |d_i(x')s_i(y') - d_i(x'')s_i(y'')| \\
&\leq |c_i||x' - x''| + |d_i(x')||s_i(y') - s_i(y'')| + |s_i(y'')||d_i(x') - d_i(x'')| \\
&\leq |c_i||x' - x''| + |s_i(y') - s_i(y'')| + \sup_{y'' \in D(s_i)} |s_i(y'')||d_i(x') - d_i(x'')| \\
&\leq (|c_i| + \sup_{y'' \in D(s_i)} |s_i(y'')|L_{d_i})|x' - x''| + |s_i(y') - s_i(y'')|,
\end{aligned}$$

where $D(s_i) \subset \mathbb{R}$ is the domain of definition of s_i. Let

$$k := \max_{i=1,2,\cdots,N} (|c_i| + \sup_{y'' \in D(s_i)} |s_i(y'')|L_{d_i}),$$

Then for all $(x', y'), (x'', y'') \in I \times \mathbb{R}$,

$$|F_i(x', y') - F_i(x'', y'')| \leq k|x' - x''| + \varphi(|y' - y''|),$$

where $\varphi : (0,+\infty) \to (0,+\infty)$ is some non-decreasing function such that $\varphi(t) < t$ for $t > 0$ and $t \to \frac{\varphi(t)}{t}$ is non-increasing. That is, each $F_{i,j}$ is a Rakotch contraction (with the same function φ) with respect to the second variable, and Lipschitz with respect to the first variable. Hence we obtain for all $(x', y'), (x'', y'') \in I \times \mathbb{R}$,

$$\begin{aligned}
d_\theta(w_i(x', y'), w_i(x'', y'')) =& d_\theta((l_i(x''), F_i(x', y')), (l_i(x''), F_i(x'', y'))) \\
=& |l_i(x') - l_i(x'')| + \theta |F_i(x', y')) - F_i(x'', y'')| \\
\leq& |a_i||x' - x''| + \theta(k|x' - x''| + \varphi(|y' - y''|)) \\
=& |a_i||x' - x''| + \theta k|x' - x''| + \theta\varphi(|y' - y''|) \\
\leq& (|a_i| + \theta k)|x' - x''| + \theta\varphi(|y' - y''|).
\end{aligned}$$

Let $(x', y'), (x'', y'') \in I \times \mathbb{R}$ and $(x', y') \neq (x'', y'')$. Since $\varphi : (0,+\infty) \to (0,+\infty)$ is non-decreasing continuous function and $\varphi(t) < t$ for all $t > 0$, we obtain that

$$\begin{aligned}
d_\theta(w_i(x', y'), w_i(x'', y'')) \leq& (|a_i| + \theta k)|x' - x''| + \theta\varphi(|y' - y''|) \\
=& (|a_i| + \theta k)|x' - x''| + \theta \frac{\varphi(|y' - y''|)}{|x' - x''| + |y' - y''|}(|x' - x''| + |y' - y''|) \\
=& (|a_i| + \theta k + \theta \frac{\varphi(|y' - y''|)}{|x' - x''| + |y' - y''|})|x' - x''| \\
&+ \theta \frac{\varphi(|y' - y''|)}{|x' - x''| + |y' - y''|}(|y' - y''|) \\
\leq& (|a_i| + \theta k + \theta \frac{\varphi(|x' - x''| + |y' - y''|)}{|x' - x''| + |y' - y''|})|x' - x''| \\
&+ \theta \frac{\varphi(|x' - x''| + |y' - y''|)}{|x' - x''| + |y' - y''|}(|y' - y''|) \\
\leq& (|a_i| + \theta k + \theta)|x' - x''| + \theta \frac{\varphi(|x' - x''| + |y' - y''|)}{|x' - x''| + |y' - y''|}|y' - y''|) \\
\leq& \max\{|a_i| + \theta k + \theta, \frac{\varphi(|x' - x''| + |y' - y''|)}{|x' - x''| + |y' - y''|}\}(|x' - x''| + \theta|y' - y''|) \\
=& \max\{|a_i| + \theta k + \theta, \frac{\varphi(|x' - x''| + |y' - y''|)}{|x' - x''| + |y' - y''|}\} d_\theta((x', y'), (x'', y'')) \\
\leq& \max\{\max_{i=1,2,\cdots,N} |a_i| + \theta k + \theta, \frac{\varphi(|x' - x''| + |y' - y''|)}{|x' - x''| + |y' - y''|}\} d_\theta((x', y'), (x'', y'')).
\end{aligned}$$

Since $N > 1$, we obtain $0 < a_i := \frac{x_i - x_{i-1}}{x_N - x_0} < 1$ for all $i = 1, 2, \cdots, N$.
Let

$$\theta := \frac{1 - \max_{i=1,2,\cdots,N} |a_i|}{2(k+1)}.$$

Then $0 < \max_{i=1,2,\cdots,N} |a_i| + \theta k + \theta < 1$ and since $k \geq 0$, we obtain $0 < \theta < 1$.
Let for all $t > 0$,

$$\beta(t) := \max\{\max_{i=1,2,\cdots,N} |a_i| + \theta p + \theta, \frac{\varphi(t)}{t}\}.$$

Then because $\alpha(t) := \frac{\varphi(t)}{t}$ and $\alpha : (0, +\infty) \to [0, 1)$ is a non-increasing, we can see that $\beta : (0, +\infty) \to [0, 1)$ is a non-increasing and for each $(x', y'), (x'', y'') \in I \times \mathbb{R}, (x', y') \neq (x'', y'')$,

$$d_\theta(w_i(x', y'), w_i(x'', y'')) \leq \beta(d((x', y'), (x'', y')))d_\theta((x', y'), (x'', y'')),$$

where $d((x', y'), (x'', y'')) := |x' - x''| + |y' - y''|$. Since $0 < \theta < 1$, for all $(x', y'), (x'', y'') \in I \times \mathbb{R}, (x', y') \neq (x'', y'')$,

$$|x' - x''| + \theta|y' - y''| \leq |x' - x''| + |y' - y''|.$$

That is,

$$d_\theta((x', y'), (x'', y'')) \leq d((x', y'), (x'', y'')).$$

Since $\beta : (0, +\infty) \to [0, 1)$ is a non-increasing, we can see that

$$\begin{aligned} d_\theta(w_i(x', y'), w_i(x'', y'')) \leq & \beta(d((x', y'), (x'', y')))d_\theta((x', y'), (x'', y'')) \\ \leq & \beta(d_\theta((x', y'), (x'', y')))d_\theta((x', y'), (x'', y'')). \end{aligned}$$

Hence w_i are Rakotch contractions in $(I \times \mathbb{R}, d_\theta)$. On the other hand, metric d_θ is equivalent to the Euclidean metric d_0 on $I \times \mathbb{R}$ (see [10]). So $(I \times \mathbb{R}, d_\theta)$ is a complete metric space. Hence $w_i : I \times \mathbb{R} \to I \times \mathbb{R}$ is a Rakotch contraction in $(I \times \mathbb{R}, d_\theta)$ and by Theorem 2.3 (1), there exists an unique fixed point in $I \times \mathbb{R}$. By Theorem 2.3 (2), for the complete metric space $(I \times \mathbb{R}, d_\theta)$, there is a unique nonempty compact set $G \subset I \times \mathbb{R}$ such that

$$G = \bigcup_{i=1}^{N} w_i(G).$$

By the definition of Hausdorff metric, equivalence of two metrics implies the equivalence of Hausdorff metrics generated by them (see [9], p.91, Lemma 3.6). Hence for $(I \times \mathbb{R}, d_0)$, there is a unique nonempty compact set $G \subset I \times \mathbb{R}$ such that

$$G = \bigcup_{i=1}^{N} w_i(G).$$

This completes the proof. □

Remark 5.2. *The boundedness of s_i is the essential condition to establish a unique invariant set of an iterated function system.*
In the fractal interpolation function with function vertical scaling factors, $0 < \max_{x \in I} |d_i(x)| <$

1 *(see [?], see [12], p.3-4, cf. [2]). Let* $M := \max_{x\in I} |c_i x + f_i|$ *and* $h \geq \frac{M}{1-\max_{x\in I}|d_i(x)|}$. *Then for all* $y \in [-h, h]$,

$$\begin{aligned} |F_i(x,y)| &= |c_i x + d_i(x)y + f_i| \\ &\leq M + \max_{x\in I} |d_i(x)||y| \\ &\leq M + \max_{x\in I} |d_i(x)|h \leq h. \end{aligned}$$

So for all $(x, y) \in I \times [-h, h]$, *we can see that* $F_i(x, y) \in [-h, h]$. *That is, an iterated function system,*

$$\{I \times [-h, h]; w_i : i = 1, 2, \cdots, N\},$$

has been constructed (cf. [3], p.1897). Thus $D(s_i) = [-h, h]$ *and* $s_i(y) := \max_{x\in I} |d_i(x)|y$ *is bounded in* $D(s_i)$ *(see Remark 3.1). Hence the boundedness of* s_i *in* $D(s_i)$ *is the essential condition to establish a unique invariant set of an iterated function system (see the proof of Theorem 5.1, cf. [3], p.1897, the proof of Theorem 2.1).*

Remark 5.3. *Our result is a substantial generalization of [?, ?, ?]. The function whose graph is the attractor of an IFS as described in Theorem 4.2 and Theorem 5.1 generalizes the affine fractal interpolation function (see [1]), the fractal interpolation function with function vertical scaling factors (see [?], see [12]) and the nonlinear fractal interpolation function (see. [10]). (1) In the affine fractal interpolation function (cf. [1], p.308, Example 1), for all* $t > 0$,

$$\varphi(t) := \max_{i=1,2,\cdots,N} |d_i|t,$$

where $|d_i| < 1$ *for all* $i = 1, 2, \cdots, N$.
(2) In the fractal interpolation function with function vertical scaling factors (cf. [?], cf. [12], p.3), for all $t > 0$,

$$\varphi(t) := \max_{i=1,2,\cdots,N} \max_{x\in I} |d_i(x)|t,$$

where $d_i(x)$ *is Lipschitz function defined on* I *satisfying* $\sup_{x\in I} |d_i(x)| < 1$ *for all* $i = 1, 2, \cdots, N$. *(3) In the nonlinear fractal interpolation function (cf. [10]),* $d_i(x) \equiv 1$ *and* $\varphi : (0, +\infty) \to (0, +\infty)$ *is a non-decreasing continuous function such that for any* $t > 0$, $\alpha(t) := \frac{\varphi(t)}{t} < 1$ *and the function* $(0, +\infty) \ni t \to \frac{\varphi(t)}{t}$ *is non-increasing. Thus, we improve upon results proved by [1], [12] and [10].*

6 A certain generalized fractal interpolation function

In this section we focus on generalized FIFs with a special structure by means of results obtained in the previous sections. For this purpose, we give an extremely explicit simple example to demonstrate the effectiveness of the preceding theory. We may assume, without loss of generality, that $[x_0, x_N] = [0, 1]$. This special case can always be achieved by means of an affine transformation (which does not change the existence of FIF) (see [2]).

Let $\varphi(t) := \frac{t}{1+t}$ for $t \in (0, +\infty)$. Then $\varphi : (0, +\infty) \to (0, +\infty)$ is a non-decreasing continuous function and $t \to \frac{\varphi(t)}{t}$ is non-increasing continuous function.
Let a set of data $\{(x_i, y_i) : i = 0, 1, \cdots, N\}$ be given, where $x_i, y_i \in [0, 1]$ for all $i = 0, 1, \cdots, N$. Let for all $i = 1, 2, \cdots, N$,

$$d_i(x) := 2^{2i} x^i (1 - x)^i.$$

Then

$$\max_{x \in [x_0, x_N]} |d_i(x)| = 1$$

and by differential mean value theorem, for all $x', x'' \in [0, 1]$, there is $L_{d_i} > 0$ such that

$$|d_i(x') - d_i(x'')| \leq L_{d_i} |x' - x''|.$$

Let for $y \in [0, +\infty)$ and $i = 1, 2, \cdots, N$,

$$s_i(y) := \frac{y}{1 + iy}.$$

Then, for $y', y'' \in [0, +\infty)$,

$$\begin{aligned} |s_i(y') - s_i(y'')| &= |\frac{y'}{1 + iy'} - \frac{y''}{1 + iy''}| \leq \frac{|y' - y''|}{1 + i|y' - y''|} \\ &\leq \frac{|y' - y''|}{1 + |y' - y''|} = \varphi(|y' - y''|). \end{aligned}$$

That is, each s_i is Rakotch contraction (with the same function φ) that is not Banach contraction on $[0, +\infty)$ (see [10], cf. [11], p.848, cf. [8], p.262).
Let for all $i = 1, 2, \cdots, N$,

$$w_i(x, y) := (a_i x + e_i, c_i x + d_i(x) s_i(y) + f_i),$$

where

$$\begin{aligned} a_i &= x_i - x_{i-1}, \quad e_i = x_{i-1}, \\ c_i &= y_i - y_{i-1}, \quad f_i = y_{i-1}. \end{aligned}$$

Then, by Theorem 4.2 and Theorem 5.1, there exists a continuous function $f : [0, 1] \to \mathbb{R}$ that interpolates the given data $\{(x_i, y_i) : i = 0, 1, \cdots, N\}$. Moreover, the graph G of f is invariant with respect to $\{[0, 1] \times \mathbb{R}; w_1, w_2, \cdots, w_N\}$, i.e.,

$$G = \bigcup_{i=1}^{N} w_i(G).$$

This clearly shows that our method is much more effective than the method due to [1]-[10].

Remark 6.1. *We refer to f as a generalized fractal interpolation function with function vertical scaling factors. The reason for this name is that the functions F_i take the form*

$$F_i(x, y) = c_i x + d_i(x)s_i(y) + f_i,$$

where $\max_{x\in I} |d_i(x)| \leq 1$ and each s_i is Rakotch contraction.
That is, each F_i, in general, is generalized with respect to the second variable y (cf. [10]). In fact, in [1]-[12], since $0 < |d_i(x)| \equiv |d_i| < 1$ or $0 < \max_{x\in I} |d_i(x)| < 1$ and

$$d_i(x)y = \frac{d_i(x)}{\max_{x\in I} |d_i(x)|} \max_{x\in I} |d_i(x)|y,$$

we can see that

$$\begin{aligned} F_i(x, y) &= c_i x + d_i(x)y + f_i \\ &= c_i x + d_i^*(x)s_i(y) + f_i, \end{aligned}$$

*where $d_i^*(x) := \frac{d_i(x)}{\max_{x\in I} |d_i(x)|}$ and $s_i(y) := \max_{x\in I} |d_i(x)|y$, and thus each s_i is a special Banach contraction and linear. That is, each $F_i(x, y)$ is a special Banach contraction and linear with respect to the second variable y. Then the corresponding FIF is an affine FIF introduced by Barnsley (see [1]) or a FIF with function vertical scaling factors (see [**?**], see [12], cf. [2]). Obviously, we can say that the generalized FIFs with function vertical scaling factors may have more flexibility and applicability.*

In below, we give the graph of a linear FIF, the graph of a nonlinear FIF of [10] that is not a linear FIF, and the graph of a generalized FIF that is not a FIF of [10] (see Fig. 1, Fig. 2, Fig. 3). Here we omit their details to avoid the repetition.

7 Conclusion

The FIFs have been widely used in approximation theory, image compression, computer graphics and modeling of natural surfaces such as rocks, metals, terrains and so on. In order to get more flexibility and diversity in modeling natural shapes and phenomena or in image processing, we introduced new generalized FIFs which generalize widely used linear FIFs. In order to obtain new generalized FIFs, we use Rakotch contractions and special function vertical scaling factors, and we have presented the principle and the method of generalized fractal interpolation in detail. Dealing with generalized fractal interpolation functions is better than the one provided by [10]. Theorem 4.2 and Theorem 5.1 ensure that an attractor of constructed generalized IFS is a graph of some continuous function which interpolates the given data. In particular, an explicit illustrative example shows that our result remains still true under essentially weaker conditions on the maps of IFS. Comparing linear FIFs with function vertical scaling factors and generalized FIFs with function vertical scaling factors, we can know that the FIFs considered in this paper have more flexibility and diversity and are more suitable to the fitting and approximation of many complicated functions.

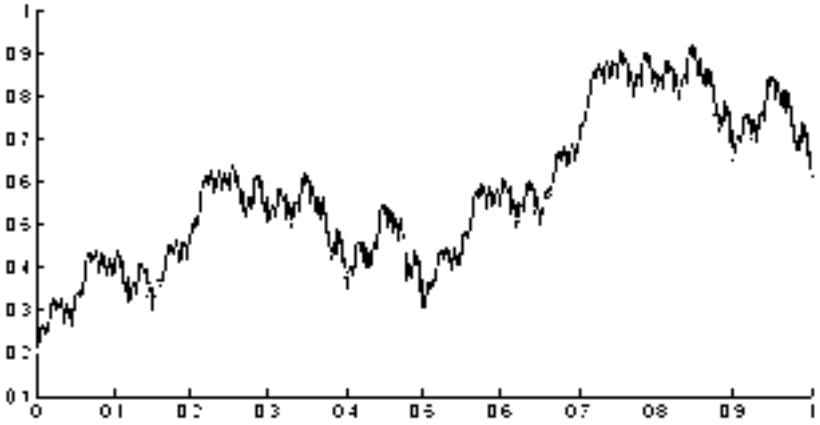


Figure 1: A linear FIF.

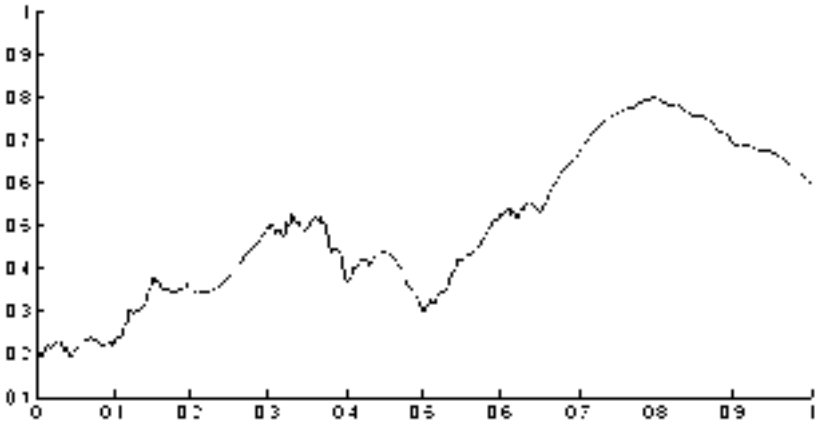


Figure 2: A nonlinear FIF of [10] that is not a linear FIF.

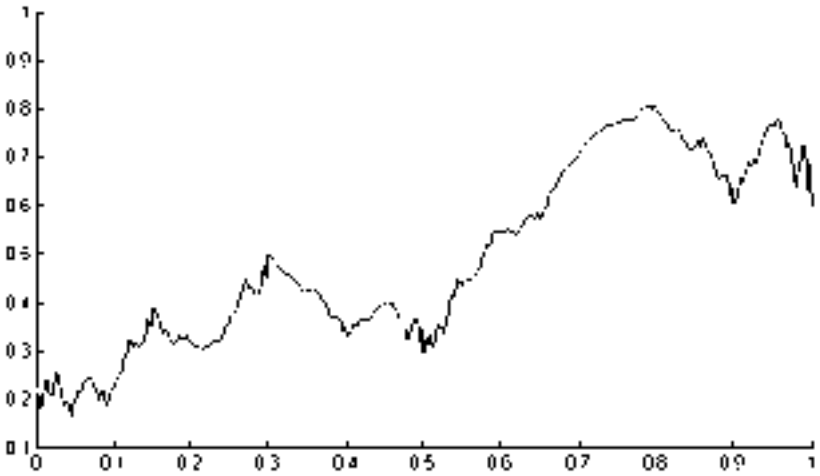


Figure 3: A generalized FIF that is not a FIF of [10].

References

[1] M. F. Barnsley. Fractal functions and interpolation. *Constr. Approx.*, 2:303–329, 1986.

[2] M. F. Barnsley and P. R. Massopust. Bilinear fractal interpolation and box dimension. *J. Approx. Theory*, 192:362–378, 2015.

[3] Zhigang Feng, Yizhuo Feng, and Zhenyou Yuan. Fractal interpolation surfaces with function vertical scaling factors. *Appl. Math. Lett.*, 25(11):1896–1900, 2012.

[4] J. Jachymski and I. Jóźwik. Nonlinear contractive conditions: a comparison and related problems. *Banach Center Publ.*, 77:123–146, 2007.

[5] G. G. Łukawska and J. Jachymski. The Hutchinson-Barnsley theory for infinite iterated function systems. *Bull. Aust. Math. Soc.*, 72:441–454, 2005.

[6] P. R. Massopust. *Fractal functions, fractal surfaces and wavelets.* Academic Press, San Diego, 1994.

[7] E. Rakotch. A note on contractive mappings. *Proc. Amer. Math. Soc.*, 13:459–465, 1962.

[8] B. E. Rhoades. A comparison of various definitions of contractive mappings. *Trans. Am. Math. Soc.*, 226:257–290, 1977.

[9] Song-il Ri. A new fixed point theorem in the fractal space. *Indagationes Mathematicae*, 27(1):85–93, 2016.

[10] Songil Ri. A new nonlinear fractal interpolation function. *Fractals*, 25(6):1750063 (12 pages), 2017.

[11] N.-A. Secelean. Generalized iterated function systems on the space $l^{\infty}(X)$. *J. Math. Anal. Appl.*, 410(2):847–858, 2014.

[12] Hong-Yong Wang and Jia-Shan Yu. Fractal interpolation functions with variable parameters and their analytical properties. *J. Approx. Theory.*, 175:1–18, 2013.

Hidden Bifurcation kind to Multiscroll Chaotic Attractors via Saturated Function Series

Zaamoune Faiza[1] and Tidjani Menacer[2]

[1] Department of Mathematics, University Mohamed Khider, Biskra, Algeria
(E-mail: zaafaiza25@gmail.com)
[2] Department of Mathematics, University Mohamed Khider, Biskra, Algeria
(E-mail: tidjanimenacer@yahoo.fr)

Abstract. In this paper, design Hidden Bifurcation kind to Multiscroll Chaotic Attractors via Saturated Function Series are reconnoitred. The idea was taken from the work Zaamoune, et al(2019) and the method introduced by Menacer, et al. (2016) for Chua multiscroll attractors. These idea (hidden bifurcation kind) depends on how you appear the scrolls odd or even were the number of scrolls is even. We have studied many examples to prove this idea, mentioned that the number n of scrolls satisfies $n = p + q + 2$ in LÜ , Chen et al (2004).
Keywords: Hidden bifurcation, multiscroll chaotic attractor, saturated function series..

1 Introduction

It is well known that the method hidden attractor generating one-directional(1-D) -scroll has been studied in the last few years [1]-[3]-[4]. Since the method was presented by Leonov, et al [4] in the Chua attractor, they have proposed efficient technique for the numerical localization of the hidden attractors in one-directional dynamical systems. Hidden attractor has many applications in a real word like mechanics, electronics, chemistry, biology but the most important in electronic circuits (hysteresis circuit, and saturated circuit). In 2016, the auteur's Menacer, et al changed the type of discrete parameters by presented a generating multi-spirals, and this new method they called a "hidden bifurcation", the cause of this name it's has a change in the number of spirals. In this paper, we study design hidden bifurcation kind method in the attractor for generating one-directional (1-D) -scroll by saturated function series. In 2004 Chen, et al design and analysis of multi scroll chaotic attractors from saturated function series [8], but we present a new idea in hidden bifurcation, where we know that the role of method hidden bifurcation it's a parametres control in the number of spirals, her in this work p_1 and q_1 in function 12 it's a bossed. In this article, we change the values of system Chen parameters 11 a, b ,c and d_1 we found a new attractor different about Chen attractor in [8], we change every time the parameter ϵ in 0.55 to 1 and the parameters p_1 and q_1 and we noted that it's not only the parameter ϵcontrol in the appearance of numbers scrolls, the parameters p_1 and q_1 also control in appearance of numbers scrolls it's a new idea. This paper is disposed of as follows : In Section 2, the analytical-numerical method for hidden attractor proposed by Leonov in [6],[7]. the model of $1-D$ scroll chaotic attractors generated by saturated function series proposed

.In Section 3, the model of $1-D$ scroll chaotic attractors generated by saturated function series proposed. In section 4 , the localization technique presented in [2] for hidden bifurcation in $1-D$ scroll chaotic attractors, we introduced the results for a new idea. Finally, in Sec. 6, a terse conclusion is pictured.

2 Analytical-numerical procedure for attractors localization

Leonov et al. [6],[9], [10] found a procedure to discover numerically hidden attractor for Chua attractor. The technique developed in [2], discovering hidden bifurcations in the multispiral Chua attractor. *To improv, this numerical method, consider a system with one-directional (1-D) -scroll*

$$\frac{dx}{dt} = Px + \beta\Psi(\kappa^T x), \quad x \in \mathbb{R}^3. \tag{1}$$

were P is a constant $(n \times n)$ matrix,β, κ are constant $n-$dimensional vectors, T is a transposition operation, $\Psi(\varsigma)$ is a continuous piecewise-differentiable vector-function, and $\Psi(0) = 0$. Consider a coefficient of Harmonic linearization k at like the matrix P_0 as :

$$P_0 = P + k\beta\kappa^T \tag{2}$$

wich $\pm i\omega_0$ $(\omega_0 > 0)$ eigenvalues the matrix P_0 and the rest have negative real parts. Suppose that such k occurs. So, rewrite system 1 as.

$$\frac{dx}{dt} = P_0 x + \beta\varphi(\kappa^T x), \tag{3}$$

were $\varphi(\varsigma) = \Psi(\varsigma) - k\varsigma$. We display a fixed sequence of functions $\varphi^0(\varsigma), \varphi^1(\varsigma), \cdots, \varphi^n(\varsigma)$, that the function $\varphi^0(\varsigma)$ is small, and $\varphi^m(\sigma) = \varphi(\varsigma)$. In this state the smallness of function $\varphi^0(\varsigma)$, permit one to practise the procedure of harmonic linearization for the system

$$\frac{dx}{dt} = P_0 x + \beta\varphi^0(\kappa^T x) \tag{4}$$

and conclude a stable nontrival periodic solution $x^0(t)$. So, the localization of the attractor of a system(19), design numerically the transformation of this periodic solution. So, we obtain the primary condition $x^0(0)$ of the periodic solution, system (4) can be changed by S $(X = SY)$ to the form :

$$\begin{cases} \dot{y_1} = -\omega_0 y_2 + v_1\varphi^0(y_1 + c_3^t Y_3) \\ \dot{y_2} = \omega_0 y_1 + v_2\varphi^0(y_1 + c_3^t Y_3) \\ \dot{Y_3} = A_3 Y_3 + V_3\varphi^0(y_1 + c_3^t Y_3) \end{cases} \tag{5}$$

So V_3 and c_3 is an $(n-2)-$dimensional vector, y_1, y_2 are scalar values; Y_3 is an $(n-2)-$dimensional vector, v_1 and bv_2 are real numbers; A_3 is an $(n-2)\times(n-2)$ matrix, where all of its eigenvalues have negative real parts. Supposed that for the matrix A_3 there exists a $d_3 > 0$ such that

$$Y_3^t(A_3 + A_3^t)Y_3 \leq -2d_3 |Y_3|^2, \ \forall Y_3 \in \mathbb{R}^{n-2}. \tag{6}$$

In one-directional(1-D) -scroll, case, present the describing function Φ of a real variable ς as follows:

$$\Phi(\varsigma) = \int_0^{2\pi/\omega_0} cos(\omega_0 t)\varsigma) \cos((\omega_0 t)\, dt \tag{7}$$

Theorem 1. *If a positive ς_0 satisfies that*

$$\Phi(\varsigma_0) = 0,\ b_1 \frac{d\Phi(\varsigma)}{d\varsigma}\bigg|_{\varsigma=\varsigma_0} < 0 \tag{8}$$

So, for the initial condition of the periodic solution $X^0(0) = S(y_1(0),\ y_2(0),\ Y_3(0))^T$ at the first step of algorithm, one has

$$y_1(0) = \varsigma_0 + O(\epsilon),\ y_2(0) = 0,\ Y_3(0) = O_{n-2}(\epsilon) \tag{9}$$

In Application, to find k and ω_0, can uses the transfer function of system (4),

$$W(\rho) = \kappa^T (M - \rho I)^{-1} \beta \tag{10}$$

where ρ is a complex variable. The number ω_0 is calculated from the equation $ImW(i\omega_0) = 0$ and k is then determined using the formula $k_0 = -ReW(i\omega_0)$.

3 $1 - D$ Scroll Chaotic Attractor via a saturated function series

Here, to generate $1 - D$ n scroll chaotic attractor we presented a system from saturated function series follows :

$$\begin{cases} \dot{x} = y \\ \dot{y} = z \\ \dot{z} = -ax - by - cz + d_1 f(x; k_1; h_1; p_1; q_1) \end{cases} \tag{11}$$

where

$$f(x; k_1; h_1; p_1; q_1) = \begin{cases} y_1 & if\ x > q_1 h_1 + 1 \\ y_2 & if\ |x - ih_1 \leq 1, -p_1 \leq_1 i \leq q_1| \\ y_3 & if\ ih_1 + 1 < x < (i+1)\, h_1 - 1 \\ \text{and} & -p_1 < i < q_1 - 1 \\ y_4 & if\ x < -q_1 h_1 - 1 \end{cases} \tag{12}$$

$y_1 = (2q_1 + 1)\, k_1$, $y_2 = k_1\,(x - ih_1) + 2ik_1$, $y_3 = (2i+1)\, k_1$, and $y_4 = -\,(2q_1 + 1)\, k_1$here, a, b, c, d_1 are real numbers and the parameters p_1, q_1, h and k are integers. The formula to calculated the number n of scrolls, it's explained in Chen, et al [8] and in Zaamoune, et al[11] For k = 11,h = 22,$p_1 = 3$, $q_1 = 3$, $a = d_1 = 0.8$, $c = 0.72$ and $b = 0.6$ a 8-scroll attractor is generated as the verged attractor of the system (11-12), see Fig 1

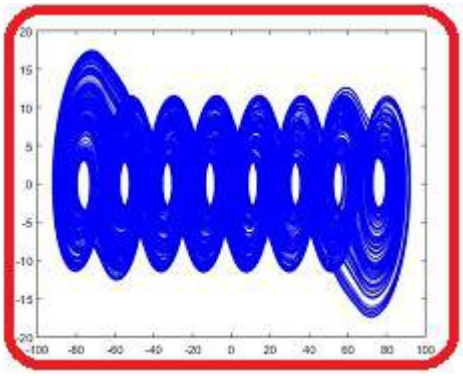


Fig. 1. One-directional 8-scroll chaotic attractors.

4 Generalized hidden bifurcation with scalar nonlinearity.

Now, we apply the above procedure; rewrite a system 11 in the form

$$\frac{dx}{dt} = Px + \beta\Psi(\kappa^t x), \qquad x \in \mathbb{R}^3 \tag{13}$$

Here

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a & -b & -c \end{pmatrix}, \quad X = \begin{pmatrix} x \\ y \\ z \end{pmatrix},$$

$$\kappa = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \beta \begin{pmatrix} 0 \\ 0 \\ d_1 \end{pmatrix} \quad and \quad \psi(\varsigma) = f(\varsigma).$$

Define the coefficient k and a small parameter ϵ, system (13) can be rewrite in the forme :

$$\frac{dx}{dt} = P_0 x + \beta\epsilon\varphi(\kappa^T x), \tag{14}$$

where

$$P_0 = P + k\beta\kappa^T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ k_0 d_1 - a & -b & -c \end{pmatrix},$$

$$\lambda_{1,2}^{P_0} = \pm i\omega_0, \quad \lambda_3^{P_0} = -d$$

The transfer function $W_P(\lambda)$ of system (14) can be given by

$$W_{P_0}(\lambda) = \kappa^T (P - \lambda I)^{-1}\beta \tag{15}$$

where λ is a complex variable. By the transformation $X = SY$, system (14) is changed to the form

$$\frac{dY}{dt} = HY + v\varepsilon\varphi(c^T Y) \tag{16}$$

where

$$H = \begin{pmatrix} 0 & -\omega_0 & 0 \\ \omega_0 & 0 & 0 \\ 0 & 0 & -d \end{pmatrix}, \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ Y_3 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ 1 \end{pmatrix}$$

and $c = \begin{pmatrix} 1 \\ 0 \\ -h \end{pmatrix}$. The transfer function of system (16) can be written as

$$\begin{aligned} W_H(\lambda) = & \quad c^T(H - I\lambda)^{-1}v \\ = & \frac{h}{d+\lambda} - \lambda\frac{v_1}{\lambda^2+\omega_0^2} + \omega_0\frac{v_2}{\lambda^2+\omega_0^2} \end{aligned}$$

So, we could obtain the implies k, d, h, v_1, v_2 by using the equality of transfer functions of systems (14) and (16):

$$W_H(\lambda) = \kappa^T(M_0 - \lambda I)^{-1}\beta \tag{17}$$

This implies the following relations:

$$\begin{aligned} & k_0 = \tfrac{a-\omega_0^2 d}{d_1} \\ & d = c \\ & h = \tfrac{-d_1}{\omega_0^2+d^2} \\ & v_1 = \tfrac{-d_1}{\omega_0^2+d^2} \\ & v_2 = \tfrac{-cd_1}{\omega_0(\omega_0^2+d^2)} \end{aligned} \tag{18}$$

We utilized the transformation $X = SY$the following relationships can be gotten :

$$H = S^{-1}P_0S,\ b = S^{-1}\beta,\ c^T = \kappa^t S \tag{19}$$

So, by 19 we found this matrix :

$$S = \begin{pmatrix} S_{11} = 1 & S_{12} = 0 & S_{13} = -h \\ S_{21} = 0 & S_{22} = -\omega_0 & S_{23} = dh \\ S_{31} = -\omega_0^3 & S_{32} = 0 & S_{33} = d^2h \end{pmatrix}$$

So, thr first step in above procedure is determine the initial data, as

$$X(0) = SY(0) = S\begin{pmatrix} \varsigma_0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \varsigma_0 S_{11} \\ \varsigma_0 S_{21} \\ \varsigma_0 S_{31} \end{pmatrix} \tag{20}$$

For system the Chen, the initial condition is :

$$X^0(0) = \left(x^0(0) = \varsigma_0,\ y^0(0) = 0,\ z^0(0) = -\varsigma_0\omega_0^3 \right) \tag{21}$$

5 Numerical Results of Hidden Bifurcation kind

In this work, we introduced a new idea in hidden bifurcation behavior in the attractor by we changed the values of parameters of the system generated via saturated function serie. So, we presented the system (11-12) with parameter values

$$a = d_1 = 0.8, c = 0.72, b = 0.6 p_1 = q_1 = 3, k = 11 and h = 22$$

By folowing the above method we are started first calculation the frequency ω_0 and a coefficient of harmonic linearization k as well :

$$\omega_0 = 0.7745 \quad and \quad k = 0.46$$

Then, we presented fours cases numbers of scrolls $4, 6, 8, 10$, by numerous sequentially ε from the value $\varepsilon = 0.55$ to $\varepsilon = 1$. while in the case of $p = 8$ and $q = 0$,the initial conditions are : $(x_0(0) = -3.3256,\ y_0(0) = 0,\ z_0(0) = 1.5450.)$

When, $p = q = 3$ the hidden bifurcation kind is even as $2, 4, 6, 8$ and for $p = 0, q = 6$ the hidden bifurcation kind is odd as $1, 2, 3, 4, 5, 6, 7, 8$ that means the number of scrolls appearance is one by one, see figures (4, 5).

The case 4 scrolls, $p = q = 1$ the hidden bifurcation kind is even as $2, 4$and for $p = 0, q = 2$ the hidden bifurcation kind is odd as $1, 2, 3, 4$, see figures (2, 3).

The case 6 scrolls, $p = q = 2$ the hidden bifurcation kind is even as $2, 4, 6$ and for $p = 0, q = 4$ the hidden bifurcation kind is odd as $1, 2, 3, 4, 5, 6$.

The case 10 scrolls, $p = q = 4$ the hidden bifurcation kind is even as $2, 4, 6, 8, 10$ and for $p = 0, q = 8$ the hidden bifurcation kind is odd as $1, 2, 3, 4, 5, 6, 7, 8, 9, 10$.

The role idea of the hidden bifurcation kind was 'even or odd' if the number of the scrolls was even. This idea based on the principle idea of [11]. View that when the parameters p or q equalizes zero the hidden bifurcation behavior it was odd, as for, $p = q$ the hidden bifurcation behavior was even and it's explicated in tables below (1) and figures (2, 3, 4, 5)

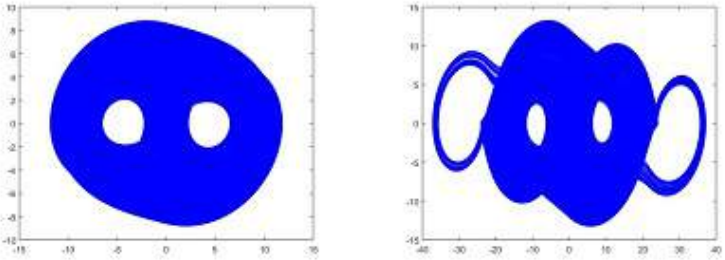
$n = p + q + 2$	The number of scrolls is even
p or q value Zero	The scrolls appearance is odd
$p = q$	The scrolls appearance is even

Table 1. The behavior of hidden bifurcation kind

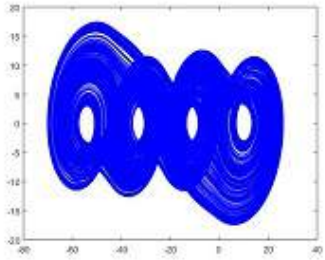
References

1. LÜ, J. & CHEN. G. (2006). Generating multiscroll chaotic attractors: Theories, methods and applications. International Journal of Bifurcation and Chaos,16(4), 775–858.

2. MENACER, T., LOZI, R. & CHUA, L.O. (2016). Hidden bifurcations in the multispiral Chua attractor. International Journal of Bifurcation and Chaos,16(4), 1630039–1630065.
3. KUZNETSOV, N. V., LEONOV, G. A. & PRASAD, A. (2016). Hidden attractors in dynamical systems. Physics Reports, 637, 1-50.
4. DENG, Q. & WANG, C. (2019). Multi-scroll hidden attractors with two stable equilibrium points. Chaos 29, 093112. https://doi.org/10.1063/1.5116732
5. ZHANG, X., & WANG C. (2019), Multiscroll Hyperchaotic System with Hidden Attractors and Its Circuit Implementation. International Journal of Bifurcation and Chaos, 29 (09), 1950117.
6. LEONOV, G. A., & KUZNETSOV, N. V. (2011). Localization of hidden Chua's attractors. Phys. Lett. A, 375, 2230–2233.
7. LEONOV, G. A. (2010). Effective methods for periodic oscillations search in dynamical systems. Appl. Math. Mech, 74, 37–73.
8. LÜ, J., CHEN, G. YU, X & LEUNG, H. (2004). Design and analysis of multiscroll chaotic attractors from saturated function series. IEEE Trans. Circuits Syst. I, 51(12), 2476–2490.
9. LEONOV, G. A. VAGAITAEV, V. I. & KUZNETSOV, N. V. (2010). Algorithm for localizing Chua attractors based on the harmonic linearization method. Dokl. Math, D, 663–666.
10. LEONOV, G. A., KUZNETSOV, N. V. (2013). Hidden attractors in dynamical systems. From hidden oscillations in Hilbert–Kolmogrov, Aizerman, and Kalman problems to hidden chaotic attractor in Chua circuits International Journal of Bifurcation and Chaos, 23, 1330002-1–1330002-69.
11. Zaamoune, F. T. Menacer, R. Lozi. & Chen, G. (2019). Symmetries in Hidden Bifurcation Routes to Multiscroll Chaotic Attractors Generated by Saturated Function Series. Journal of Advanced Engineering and Computation, 3,4, 511–522.

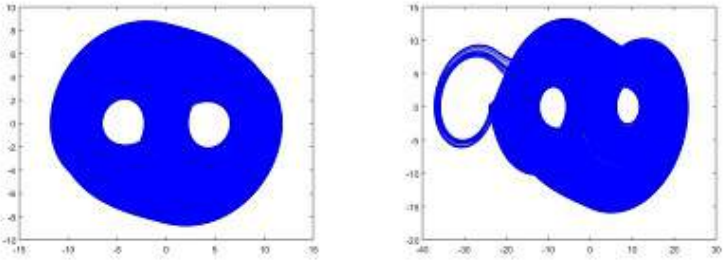


(a) (b)

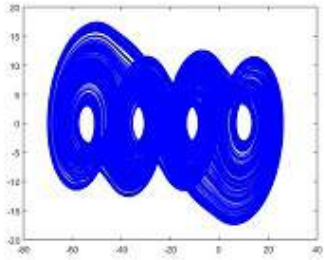


(c)

Fig. 2. The increasing number of spirals of system (14) according to increasing ε values, when $p = 1$ and $q = 1$, k=11 and h=22. (a) : The first scroll for ε=0.55, (b) : The second scroll for ε=0.92, (c) : The third scroll for ε=1.



(a) (b)



(c)

Fig. 3. The increasing number of spirals of system (14) according to increasing ε values, when $p = 2$ and $q = 0$, k=11 and h=22. (a) : The first scroll for ε=0.55, (b) : The second scroll for ε=0.92, (c) : The third scroll for ε=1.

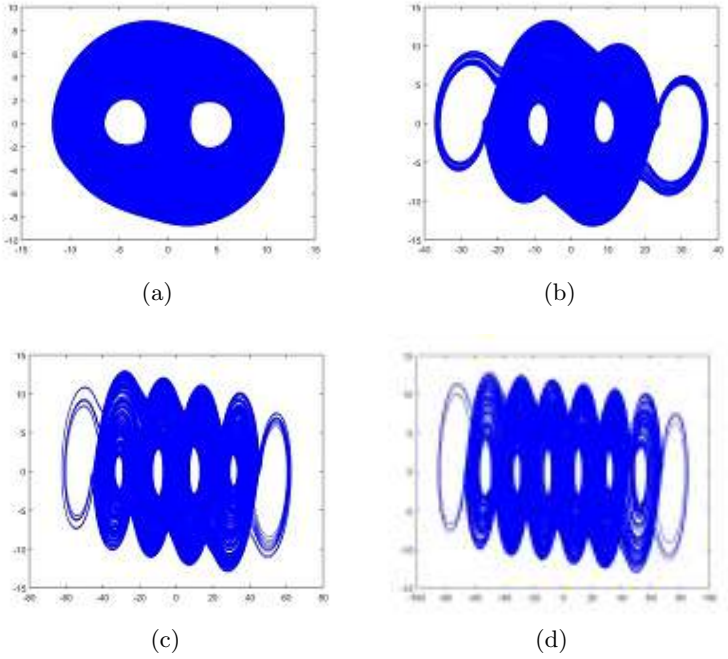


Fig. 4. The increasing number of spirals of system (14) according to increasing ε values, when $p = 3$ and $q = 3$, k=11 and h=22. (a) : The first scroll for ε=0.55, (b) : The second scroll for ε=0.92, (c) : The third scroll for ε=0.975, (d) : The last scroll for ε=1

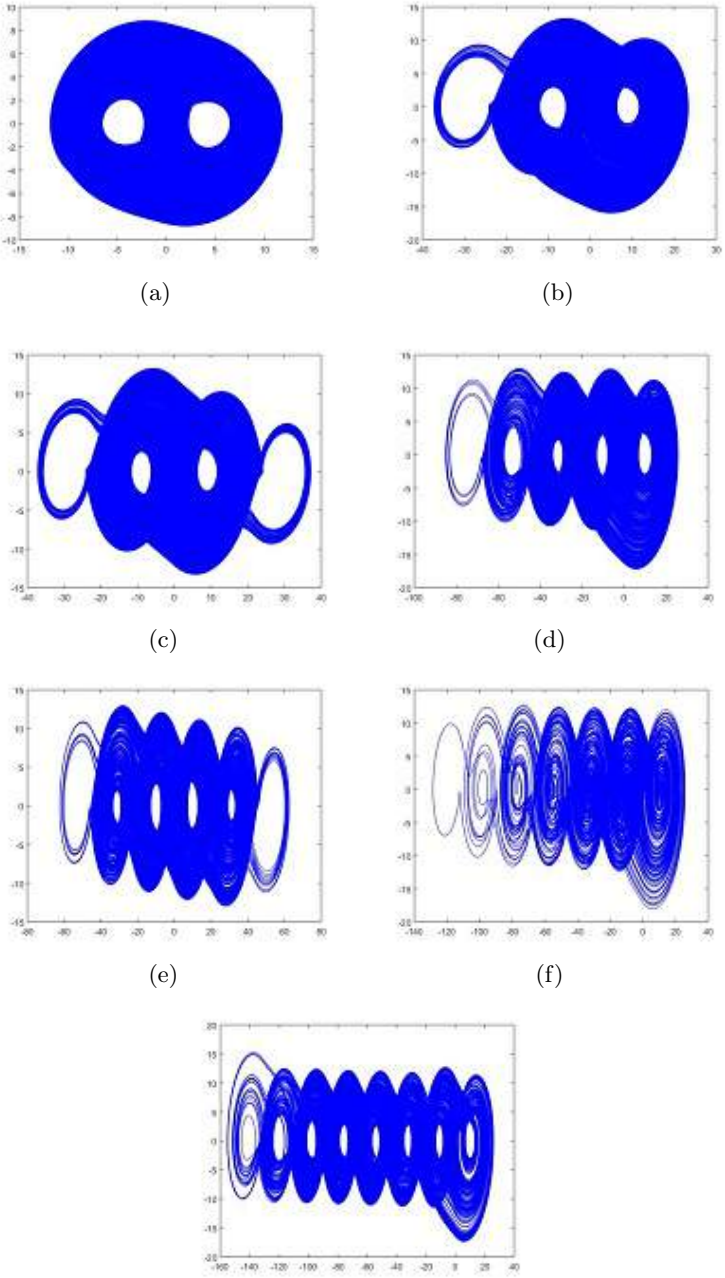


Fig. 5. The increasing number of spirals of system (14) according to increasing ε values, when $p = 6$ and $q = 0$, k=11 and h=22. (a) : The first scroll for ε=0.55, (b) : The second scroll for ε=0.92, (c) : The third scroll for ε=0.98, (d) : The fourth scroll for ε=0.985, (e) : The fifthly scroll for ε=0.99, (f) : The sixthly scroll for ε=0.992, (g) : The last scroll for ε=1

Analysis of thermal and quantum escape times of Josephson junctions for signal detection

G. Filatrella[1], C. Barone[2], C. Guarcello[3], A. S. Piedjou Komnang[3], V. Pierro[4], A. Rettaroli[5], and S. Pagano[2]

[1] INFN Gruppo Collegato Salerno and Dept. of Sciences and Technologies, Univ. of Sannio, Via F. De Sanctis, I-82100 Benevento, Italy
(E-mail: filatrella@unisannio.it)
[2] INFN, Gruppo Collegato Salerno and Dept. of Physics, Univ. of Salerno, Via Giovanni Paolo II, I-84084 Fisciano (SA), Italy
[3] Dept. of Physics, Univ. of Salerno, Via Giovanni Paolo II, I-84084 Fisciano (SA), Italy
[4] INFN Gruppo Collegato Salerno and Dept. of Engineering, University of Sannio, Corso Garibaldi 107, I-82100 Benevento, Italy
[5] INFN, Laboratori Nazionali di Frascati, Frascati (Roma) Italy and Dept. of Mathematics and Physics, University of Roma Tre, I-00100 Roma, Italy

Abstract. In this work we investigate the limits to the possibility to reveal the existence of weak microwave signals through Josephson junctions. Even if the Josephson element is capable to reveal the electromagnetic field, thermal noise is to be quantified by means of signal theory, as a confounding factor that limits the detection. We show how the decision problem can be embedded in the frame of signal detection. As a consequence, the optimization of the detection probability and the minimization of the false alarm probability give a guide to select the Josephson junction parameters that best suit the purpose.
Keywords: Josephson junctions, Escape time, Signal detection, Particle detection.

1 Introduction

When Josephson junctions (JJ) are employed as detectors of microwave signals, they can reach a very high response (of the order of kV/W [31]), close to the quantum sensitive limit [14,15,26]. Moreover, as superconducting elements, the devices can work no matter how the temperature is lowered, thus allowing to minimize thermal noise, at least until the contribution to the escapes through quantum tunneling processes becomes statistically dominant. These premises are, in a nutshell, the basis for the intensity of the efforts devoted to the development of highly-sensitive detectors based on JJ [3], as well as Josephson calorimeters [35,16]. To reveal even very weak electromagnetic signals, the energy to induce the transition between two states of the JJ should be close to the single photon energy. As the interaction between the JJ and the signal is generally mediated by a resonant cavity, also the latter excitation should be on the same energy scale [20]. Thus, for a more detailed description of the device potentialities and limits, it is necessary to embed the problem in the context of microwave photon manipulation through superconducting electronics [13,1]. In doing so, the energy landscape of the device, as well as the antenna interaction

with the microwave photon field, are essential for a device description [4]. It is in fact the JJ potential that dictates the disturbing switches due to thermal escapes and quantum tunneling [4,32]. A model for the detector consists of a current drive that perturbs the JJ dynamics and favors the switches to a finite voltage state; the statistics of the escape times are analyzed to highlight the presence of the perturbation.

Several remarkable achievements have been reached so far. The minimum photon content that can be revealed through JJ has been estimated in the order of 10^2, a limit that can be possibly lowered [36]. It has also been demonstrated that it is possible to resolve the number of photons in a propagating mode [11]. Quite naturally, to decide about the presence of a weak signal, it is necessary to achieve a silent enough state, that is a device capable to stay quite if no signal is applied. Some schemes that exploit the phase diffusion regime have been recently proposed with such purpose [30]. However, for the purpose to exploit the high sensitivity of JJ in the search for photons resulting from elusive particles as axions [7–9,28,10,22,37,3], it is necessary to set up a different detection scheme to ascertain the existence of extra photons, above and beside the background [5]. The purpose of this work is to analyze a scheme for the detection of such excess photons through the analysis of the JJ *switching currents* [2,29], embedded in the frame of signal detection [12]. In fact, as the Josephson phase is of quantum nature it is not directly accessible, detection is possible only if the photons cause a *switch*, a passage from the superconducting state to a finite voltage state (mathematically, the problem amounts to determine the first passage time across the separatrix of a potential well [17]). As the passage also occurs because of thermal fluctuations, a careful analysis is necessary to prove that the passages are a consequence of some external field. The framework of signal detection allows to make simple estimates of the temperature constraints and of the experimental set-up. As a consequence, the optimization of the detection probability (and the minimization of the false alarm probability) gives a guide to select the JJ parameters that best suit to reveal weak microwave signals. In brief, detection amounts to the following question: how is it possible to infer the presence of microwave extra photons, apart the thermal ones, from the analysis of the switches to the finite voltage of a JJ? A tentative answer to this question is the subject of the present paper, organized as follows: in Sect. 2 the problem will be extensively formulated, alongside with the description of an electrical model for the dynamics of a JJ and of a perturbation. In Sect. 3 the signal analysis indications for the performances of the detection will be recounted. Last Sect. 4 concludes.

2 The problem

We shall consider a small tunnel JJ coupled to the environment through the bias current:

$$C\frac{\hbar}{2e}\frac{d^2\varphi}{dt'^2}+\frac{1}{R}\frac{\hbar}{2e}\frac{d\varphi}{dt'}+I_0\sin\varphi=I_b+I_N(t')+I_S(t') \tag{1}$$

The right hand side of this equation summarizes the current supplied to the junction. In this formulation, therefore, one assumes that the impedance of the Josephson element is much lower than the impedance of any source, and consequently the external world is seen as a current source (although other configurations are possible, in which a voltage bias is more appropriated [24]). Eq. (1) includes inertia (determined by the capacitance C), dissipation (as governed by the dissipative element R) and fluctuations (the random current $I_N(t')$ supplied by the resistance), the nonlinear periodic term (the oscillating Josephson current of amplitude I_0), a constant bias current I_b, and the signal current $I_S(t)$ (that summarizes the effects of the photon field). Fluctuations are assumed to be Gaussian with:

$$\langle I_N(t') \rangle = 0, \tag{2}$$

$$\langle I_N(t') I_N(t'-s') \rangle = \frac{2k_B T}{R} \delta(t'-s'). \tag{3}$$

where k_B is the Boltzmann constant, T is the absolute temperature, δ the Dirac function, $\langle \cdot \rangle$ is the expectation operator.

The usual normalized units [6,33] are as follows (here, as usual, $\hbar$ is the reduced Planck constant, and e is the elementary charge):

- The current is normalized to the critical current I_0:

$$\gamma = \frac{I_b}{I_0}. \tag{4}$$

- Time is normalized to ω_J^{-1}, where $\omega_J = \sqrt{\hbar C / 2eI_0}$ is the frequency of the linear oscillator. :

$$t = t' \omega_J. \tag{5}$$

Introducing the normalized temperature $D = k_B T \omega_J / R I_0^2$ and the normalized conductance $1/\beta_c = (1/R)\sqrt{\hbar/(2eI_0C)}$, leads to the normalized versions of Eqs.(1,2,3):

$$\frac{d^2\varphi}{dt^2} + \frac{1}{\beta_c}\frac{d\varphi}{dt} + \sin\varphi = \gamma + \gamma_N(t) + \gamma_S(t), \tag{6}$$

$$\langle \gamma_N(t) \rangle = 0, \tag{7}$$

$$\langle \gamma_N(t) \gamma_N(t-s) \rangle = 2D\delta(t-s), \tag{8}$$

where γ's indicate the normalized current terms. The dynamics of the JJ is characterized by the bias-dependent small oscillation frequency:

$$\omega_0(\gamma) = \left(1-\gamma^2\right)^{1/4} \tag{9}$$

and the energy barrier that cages the dynamics is:

$$\Delta U(\gamma) = 2\left[\sqrt{1-\gamma^2} - \gamma \cos^{-1}(\gamma)\right]. \tag{10}$$

Finally, the photon field γ_S is supposed to be modeled as a succession of normalized impulses of amplitude A and duration $\delta\tau$ that arrive regularly with a period T. The response of a JJ to such a field, and how the response can be exploited to infer the existence of the field, is the problem dealt with in this paper. Some indications are presented in the next Section.

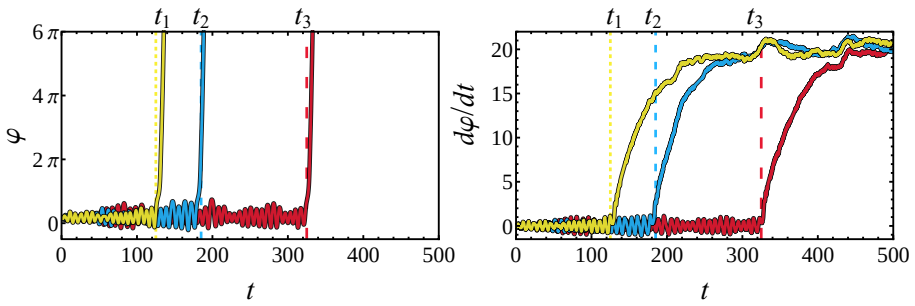


Fig. 1. Time dependent dynamics of the phase φ and the voltage $d\varphi/dt$. The dashed lines denote the time at which a passage to the finite voltage is detected.

3 Results

A JJ described by Eqs.(6,7,8) can undergo switches between the zero voltage and the finite voltage, as shown in Fig.1. When the system switches, the phase difference increases and, according to the a.c. Josephson relation $V = (h/2e)d\varphi/dt$ [18,19], a voltage drop across the junction can be measured. The switches can occur either because of thermal current (7,8), or because of the current pulse train. The starting point is therefore to collect the sequence of times at which a passage has occurred in the absence of the photon field drive:

$$t_1^0,\, t_2^0, \ldots,\, t_N^0. \tag{11}$$

For this case, one expects that, on average, the escapes (the passages across the separatrix between the localized oscillations and unbound runaways) occur with a rate given by the Kramers approximation [17]:

$$r_0(\gamma, D) = \frac{\omega_0}{2\pi} e^{-\frac{\Delta U(\gamma)}{D}} \tag{12}$$

(for the Gaussian noise; other kinds of noise give pretty different results [25]).

In the presence of a signal, which mimic the absorption of some photons that is capable to induce a current into the JJ, the sequence of switches is presumably altered, and the photons can be possibly revealed through deviations from the purely thermal sequence (11); let us call the sequence of the escapes in the presence of the extra-photons:

$$t_1^1,\, t_2^1, \ldots,\, t_N^1. \tag{13}$$

This sequence is expected to have a larger number of events in the same observation time respect to the unperturbed case (11). An example of two histograms for $\{t^0\}$ (no signal) and $\{t^1\}$ (with signal) are shown in Fig. 2. We underline that, at variance with the cases in which one is interested in counting the number of photons (e.g., [20]), to reveal a source of photons (e.g., the presence of an axion field [3]) it suffices to statistically determine if an observed collection of switches is more likely drawn from sequence (11) or from sequence (13).

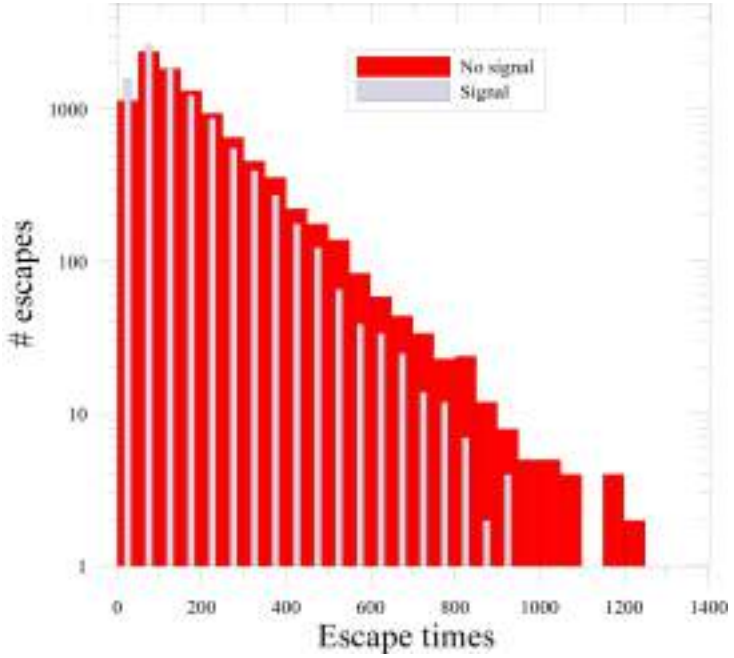


Fig. 2. Histograms of $N = 10^4$ switching times. Parameters of the simulations are: $\gamma = 0.8$, $D = 0.1$, $1/\beta_c = 0.025$. The impulses that mimic the photon fields are of amplitude $A = 0.5$, duration $\delta\tau = 10$, and arrive regularly with a period $T = 100$.

In the simplest approximation, one expects that the rate r_1 associated to the escapes (13) in the presence of photons is just the sum of the unperturbed rate r_0 and an additional rate r_A due to the photons:

$$r_1(\gamma, D, \gamma_S) = r_0(\gamma, D) + r_A. \tag{14}$$

In our settings, where the photons are modeled as current pulses with a period T, we have $r_A = 1/T$ for a most efficient detector.

Apart the analytical approximations embodied in Eq.(14), for our line of reasoning it is important to underline that the rates r_0 and r_1 are statistical averages, for Eqs.(12,14) predict what happens on average. Thus, in a given measurement time P one expects $1/\langle t_i^0 \rangle = n_0 = Pr_0$ escapes in the purely thermal case, and $1/\langle t_i^1 \rangle = n_1 = Pr_1$ escapes in the presence of the extra photons, with the obvious inequality $n_1 > n_0$. In an actual measurement, one observes a certain number of switches, say n, and a decision is to be made: in which sense the measured number favors an hypothesis (the switches are just due to thermal activation) or the other (there are extra switches due to the photon field)? Naturally, the more the two rates are apart, the more likely is that the measurement is a clear cut decision. Also, no matter how close they are, with a sufficiently long measurement time P it is always possible to discriminate the two conditions. To make these ideas quantitative is the objective of the next Section.

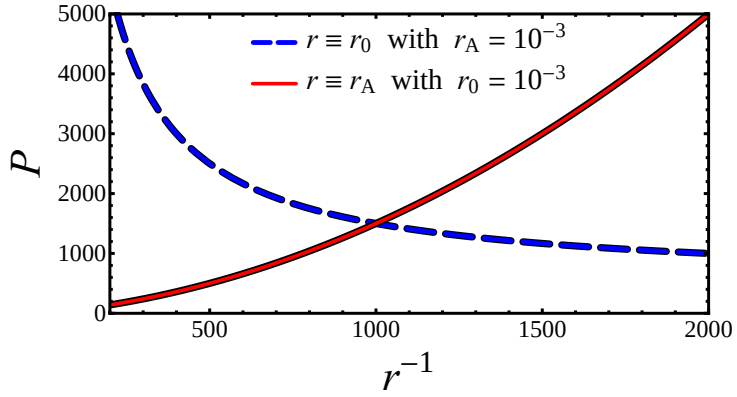


Fig. 3. Length of the observation period P, according to estimate (16), as a function of r_A at a fixed $r_0 = 0.001$ (solid red curve) and as a function of r_0 at a fixed $r_A = 0.001$ (dashed blue curve).

3.1 Statistical analysis of the switching times

To quantify the efficiency of the detection of a photon field one can introduce the Kumar-Caroll (KC) index d_{KC} [21], in analogy with the detection of continuous sinusoidal signals [12,2]:

$$d_{KC} = \frac{|\langle t^1 \rangle - \langle t^0 \rangle|}{\sqrt{\frac{1}{2}\left[\sigma^2(t^1) + \sigma^2(t^0)\right]}}. \tag{15}$$

where $\langle t^{0,1} \rangle$ is the average switching time in the absence (presence) of the signal, and $\sigma^2(t^{0,1})$ the corresponding variances. This index is a proxy for the Signal-to-Noise-Ratio (SNR) [21], and as such will be used in this paper.

Assuming a large number of events in the measurement time P, and that the escapes due to the photons are an additional rate independent of the thermal rate, and that a reliable detection requires at least $d_{KC} = 1$, one obtains the relation [34]:

$$P\, r_A^2 - \frac{1}{2} r_A - r_0 = 0, \tag{16}$$

between the rate of the photons r_A, the thermal escape rate r_0, and the observation time P. Inspection of Eq.(16) reveals that the measurement time length P is positively related to the thermal rate r_0, see Fig. 3.

Some further elaborations of the estimate (16) are relevant. To begin with, one can define a ratio between the photon arrivals and the thermal spontaneous escapes:

$$r_0 = x\, r_A$$

that allows to explicitly observe that:

$$d_{KC} = \sqrt{\frac{2 P r_A}{2x+1}}. \tag{17}$$

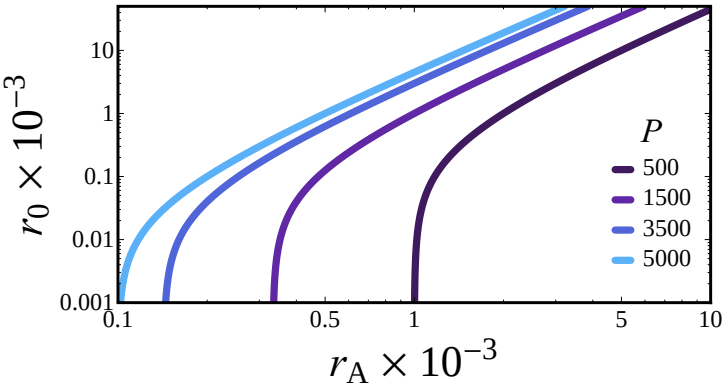


Fig. 4. Relation between the thermal escape rate as a function of the photon arrival time, according to estimate (16), for different values of P.

This equation confirms the intuitive scaling between the observation time P, the photon arrival rate, and the ratio between the thermal and photon rates. In particular, one can insert the matching condition $P = 1/r_A$ in Eq. (17) to obtain the relation between the observation time, the temperature and bias point through Eqs. (9)–(12).

In Fig. 3 it is displayed the behavior of the observation time P, calculated according to Eq. (16), as a function of the inverse photon rate, r_A^{-1} (see the red solid curve). If the photon rate is low, e.g. around 10^{-3} Hz, the observation time for a $d_{KC} = 1$ reads ~ 1500 s. It is also noticeable that the dependence is quadratic. Should the photon rate be much smaller than the assumed mHz, the detection could prove unfeasible. Conversely, if the photon arrival rate is relatively high, one can considerably increase also the thermal rate, and hence higher temperatures are allowed.

In Fig. 4 it is displayed the behavior of the thermal escape rate, r_0, as a function of the photon arrival time, r_A, for different values of the observation time P. It is evident that the thermal escape rate (that is, the temperature of the system) can be increased as the photon rate increases. However, for any value of P, there is a threshold value of r_A (e.g., $r_A \simeq 10^3$ for $P = 500$) below which thermal escapes to achieve a $d_{KC} = 1$ become vanishingly small. The latter condition entails extremely low temperatures.

In brief, if some SNR is to be reached, the trade-off between the parameters of the experiment can be evaluated. A more detailed analysis of the problem requires to retrieve the index (15) from numerical simulations of the model equations (6,7,8), as we shall do below.

3.2 An example of the Kumar-Carrol index usage

To illustrate an application of the signal-to-noise index (15), let us suppose that the index has been computed for several values of the bias current γ to optimize the bias level. To fix the ideas, let us suppose that the highest value of d_{KC} corresponds to $\gamma = 0.8$; the escapes retrieved for such choice are shown

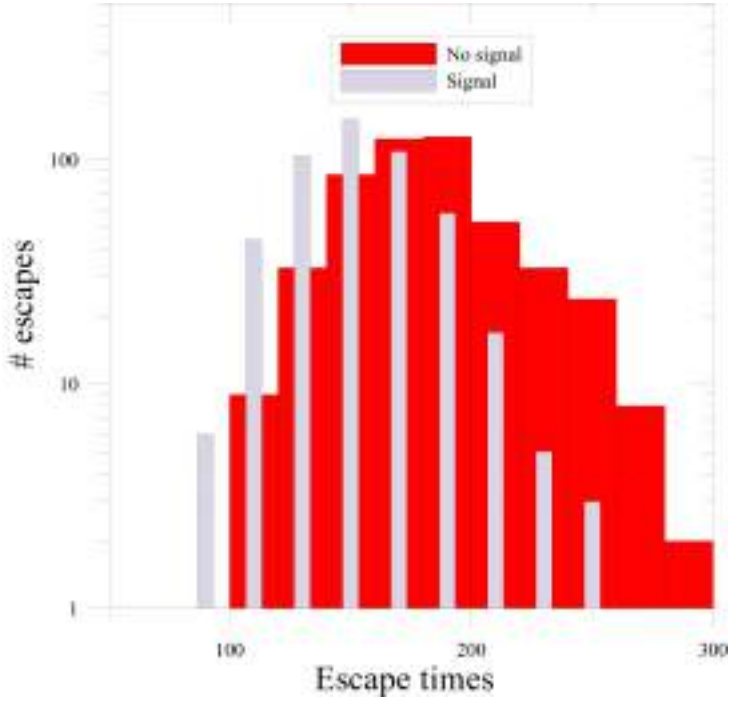


Fig. 5. Histograms of the average over 20 switching times, for the same data of Fig. 2, and thus consisting of $N' = 500$ data. The other parameters are therefore the same as in Fig. 2.

in Fig.2. The index reads $d_{KC} = 20.8$, that indicates a very good SNR – would it be possible to collect 10000 escapes, the average of the exit times could give a clear cut indication of the presence, or not, of the excess photons. In fact the statistics of the escapes can be summarized in Tab. 1.

Signal	N	$\langle t_i \rangle$	$\sigma(t_i)$	$\sigma(\langle t_i \rangle_N)$
absent	10^4	182.7	151	1.24
present	10^4	154.0	124	1.51

Table 1. Statistics of the escape times of Fig.2. Here N is the number of switches, $\langle t_i \rangle$ the average escape time, $\sigma(t_i)$ the standard deviation, $\sigma(\langle t_i \rangle_N)$ the standard deviation of the average . The resulting KC-index reads $d_{KC} = 20.8$.

It is evident that for the case under examination the detection performs extremely well. With a KC-index around 20 the SNR is extremely high, and also intuitively there is no doubt that, if the average escape time decreases from ~ 180 to ~ 150, something has happened and a signal is present. This is quite reasonable, for the switches occur on average with the same rate as in the incoming pulses period $T = 100$; therefore the efficiency is very high

(almost each pulse causes a switch) and the number of data is conspicuous ($N = 10^4$). Under these circumstances, the statistical analysis is just a confirm of the intuition. However, the approach proves useful for the design of an experiment if it is not possible to collect as many as 10^4 switches. Let us suppose, *ceteris paribus*, that the rate of arrival of the pulses is extremely low, say $1/r_A \sim 1h$, and therefore the number of events that can be collected in a day of measurements is around $N = 20$. A principal question would be: how many days of data collection should be planned to decide, with good confidence, about the existence of the pulses?

Let us thus suppose that the actual sample of data to be analyzed consists of M escape times, $M \ll N$, say $M = 20$ to fix ideas. A collection of 500 average escape times (over $M = 20$ events) retrieved binning the 10^4 data of Fig. 2, is shown in Fig. 5. In a single measurement run in which 20 switches are collected, a single average escape time would be obtained, and not always the same value for the statistical fluctuations. Let us summarize the data in Tab. 2.

Signal	N'	$\langle t_i \rangle$	$\sigma(t_i)$	$\sigma(\langle t_i \rangle_{N'})$
absent	500	182.7	33.8	1.24
present	500	154.0	27.7	1.51

Table 2. Statistics of the escape times of Fig.5, that is the escape times averaged over $M = 20$ events. Here $N' = N/M = 500$ is the number of means, $\langle t_i \rangle$ the average escape time over $N' \times M$ escapes, $\sigma(t_i)$ the standard deviation of the N' means, $\sigma(\langle t_i \rangle_{N'})$ the standard deviation of the overall average. Naturally, $\langle t_i \rangle$ and $\sigma(\langle t_i \rangle_{N'})$ coincide with Table 1, as they are computed on the same set of data. The resulting KC-index for a single measurement (that is, an average over 20 data) reads $d_{KC} = 0.93$.

In the first place, let us remind of a subtle difference between Fig. 2 and Fig. 5, as perhaps better explained in the corresponding Table 1 and Table 2. Figure 2 is a collection of all data, any single switch that has occurred. Figure 5 is a collection of hypothetical repetition of the average over 20 switches. The statistical test to decide about the presence of the signal is to be performed on the average over all data in the case of Fig. 2, and hence the very high SNR, compared to the test on a single event of Fig. 5, to which pertains a much smaller SNR.

For the central limit theorem one can assume that the distribution of the averages is approximately Gaussian, centered on the population average (that is estimated through the average of the $N = 10^4$ data) and with a standard deviation which is smaller than the population standard deviation.

Applying the signal analysis means to determine the error of the first type α and of the second type β when a decision on the existence of the photon field is to be made, on the basis of the escape time average. Let us do so with the help of the Gaussian approximation, that is to assume that the histogram of the average escape time of Fig. 5 is Gaussian distributed, as schematically

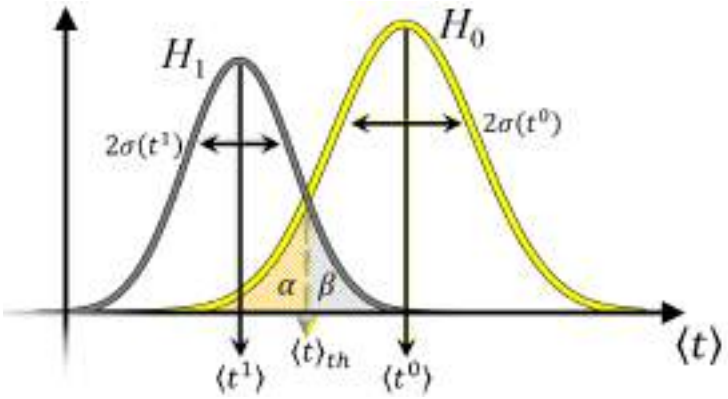


Fig. 6. Application of the detection scheme to the averaged data assumed to be Gaussian distributed.

illustrated by Fig. 6. To decide between the hypothesis one usually places a threshold for the measured average escape time, $\langle t \rangle_{th}$, see Fig. 6. The decision will favor the hypothesis "1" (the photon field is there) if the actual measured time is below the threshold, and obviously the complementary hypothesis (there is no photon field) will be selected if the actual measured time is above. The choice of the threshold reflects on the decision errors: either the probability of false alarm (type I error α), or the probability of missing a signal (type II error β). These features are combined in the so-called receiver operator characteristic of the test statistic, that is, of the combination of the errors α and β for each particular choice of the threshold. It is natural, if there are not particular reasons to do otherwise, to choose the case $\alpha = \beta$ that unequivocally individuates a threshold and hence the errors. The features of the detection can thus be obtained by a straightforward application of the central limit theorem with the data of histogram of Fig. 2: it suffices to notice that the standard deviation $\sigma(t_i)$ that appears in the definition of d_{KC} (15) is smaller of a factor $\sim \sqrt{M} \simeq 4.5$, as it is confirmed by Tables 1,2. In this manner one can connect the SNR estimated through d_{KC} to the size of type I and II errors, as illustrated in Fig. 7. In this figure, we show the d_{KC} index as a function of the sample size M (open circles and solid line). From each estimated d_{KC} index, according to the condition $\alpha = \beta$, we can uniquely determine the size of the errors of the test (triangles), see Fig. 6. In the same figure, the solid line displays the d_{KC} behavior estimated assuming that the standard deviations in Eq. (15) decrease as the square root of the number of data in a sample, $\sqrt{M}$. It is clear that such behavior scales nicely for $M \ll N$. Therefore, if a sufficiently long simulation is available, it is possible to carefully design the experiment to achieve the desired error bound.

Let us summarize how it is possible to put to a good use the definition of the SNR through the KC-index (15). First, it is useful to select the most appropriated parameter values to achieve the best SNR; in the present case we have supposed that the optimization of the SNR has given the best bias point $\gamma = 0.8$. For such parameter, a consistent number of events has been collected,

$N = 10^4$. From the data so collected, it has been possible to determine the number of experiments that could suffice to achieve the desired level of error (type I and II, chosen to be identical).

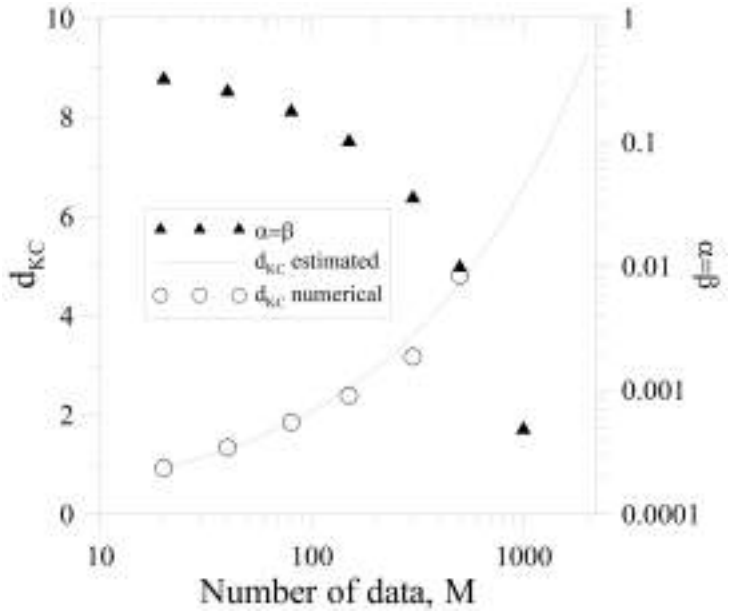


Fig. 7. Application of the detection scheme to the averaged data assumed to be Gaussian distributed. The circles (left axis) indicate the SNR ratio estimated by the index d_{KC}. The triangles (right axis) indicate the $\alpha = \beta$ level of the errors in the detection through the sample average as a function of the number of switches M of the average. The solid line is the estimate obtained assuming that the standard deviations scale as the square root of the number of points N. The other parameters are the same as in Fig. 2.

4 Conclusions

We have demonstrated that the application of signal analysis to switching event of a Josephson junction subject to a periodic train of current pulses can be used to carefully plan experiments devised to decide about the existence of the perturbation. If the pulse train is to be interpreted as a photon field that irradiates the junction, this scheme can be applied to decide about the existence of elusive particles [3]. In particular the scheme can be useful if the arrival frequency of the pulses is very low, and therefore it is particularly cumbersome to collect a large number of events to distinguish the signal from the external field from spontaneously, thermally activated, events. Under these circumstances, it is necessary to resort to statistical test. When this is the case, the SNR to be reached for a significant test is the guideline for an accurate

experiment design. As the estimate of the SNR calls for extensive simulations, both to optimize the parameters and to collect a statistically relevant number of events, it is probably necessary to resort to parallel simulations, possibly with CUDA architecture [27].

Let us add a word of caution. The analysis here presented is based on sample mean, detection can be improved with maximum likelihood estimators [2], that exploit the full information content of the escape distribution. However, the analytical estimate of the distribution of the escapes is a relatively complicated problem [23] to give a reliable solution for the estimate evaluation.

Acknowledgments

The authors wish to acknowledge financial support from Italian National Institute for Nuclear Physics INFN through the Project SIMP and from University of Salerno through projects FARB17PAGAN, FARB19PAGAN.

References

1. B. Abdo, O. Jinka, N. T. Bronn, S. Olivadese and M. Brink. On-chip single-pump interferometric Josephson isolator for quantum measurements. *arXiv:2006.01918v1*, 2020.
2. P. Addesso, G. Filatrella and V. Pierro. Characterization of escape times of Josephson junctions for signal detection. *Physical Review E*, 85, 016708-01–10, 2012.
3. D. Alesini, D. Babusci, C. Barone et al. Status of the SIMP Project: Toward the Single Microwave Photon Detection. *Journal of Low Temperature Physics*, 199, 348–354, 2020; Development of a Josephson junction based single photon microwave detector for axion detection experiments. *Journal of Physics: Conference Series*, 1559, 012020-1–7, 2020.
4. D. V. Anghel, K. Kulikov, Y. M. Galperin and L. S. Kuzmin, Electromagnetic radiation detectors based on Josephson junctions: Effective Hamiltonian. *Physical Review B*, 101, 024511-1–10, 2020.
5. E. Aprile it et al. Observation of Excess Electronic Recoil Events in XENON1T. *arXiv:2006.09721 [hep-ex]* , 2020.
6. A. Barone and G. Paternó *Physics and applications of the Josephson effect*
7. C. Beck, Testing axion physics in a Josephson junction environment. *Modern Physics Letters A* 26, 2841-2852, 2011.
8. C. Beck, Possible resonance effect of axionic dark matter in Josephson junctions. *Phys. Rev. Lett.* 111, 231801, 2013.
9. C. Beck, Axion mass estimates from resonant Josephson junctions. *Physics of the Dark Universe* 7–8, 6-11, 2015.
10. C. Beck, Cosmological flux noise and measured noise power spectra in SQUIDs. *Scientific Reports* 6, 28275, 2016.
11. R. Dassonneville, R. Assouly, T. Peronnin, P. Rouchon and B. Huard, Number-resolved photocounter for propagating microwave mode. *arXiv:2004.05114 [quant-ph]*, 2020.
12. G. Filatrella and V. Pierro. Detection of noise-corrupted sinusoidal signals with Josephson junctions. *Physical Review E*, 82, 046712-1–9, 2010;
13. Xiu Gu, A. F. Kockum, A. Miranowicz, Yu. Liu and F. Nori, Microwave photonics with superconducting quantum circuits. *Physics Reports*, 718–719, 1–102, 2017.

14. C. Guarcello, D. Valenti, B. Spagnolo, V. Pierro and G. Filatrella. Anomalous transport effects on switching currents of graphene-based Josephson junctions. *Nanotechnology*, 28, 134001, 2017.
15. C. Guarcello, D. Valenti, B. Spagnolo, V. Pierro and G. Filatrella. Josephson-based Threshold Detector for Lévy-Distributed Current Fluctuations. *Physical Review Applied*, 11, 044078-01-09, 2019.
16. C. Guarcello, A. Braggio, P. Solinas, G. P. Pepe, and F. Giazotto. *Phys. Rev. Applied*, 11, 054074, 2019.
17. P. Hänggi, P. Talkner and M. Borkovec. Reaction-rate theory: fifty years after Kramers. *Review of Modern Physics*. 62, 251-341, 1990.
18. B. Josephson. Possible new effects in superconductive tunnelling. *Physics Letters*, 1, 251, 1962.
19. B. D. Josephson. The discovery of tunnelling supercurrents. *Rev. Mod. Phys.*, 46, 251, 1974.
20. B. Kubala, J. Ankerhold and A.D. Armour. Electronic and photonic counting statistics as probes of nonequilibrium quantum dynamics. *New Journal of Physics*, 22, 023010-1–10, 2020.
21. B.V.K.V Kumar and C.W. Carrol. Performances of Wigner distribution function based detection methods. *Optical Engineer*, 23, 732–737, 1984.
22. L. S. Kuzmin, A. S. Sobolev, C. Gatti, D. Di Gioacchino, N. Crescini, A. Gordeeva and E. Il'ichev. Single Photon Counter Based on a Josephson Junction at 14 GHz for Searching Galactic Axions. *IEEE Transactions on Applied Superconductivity*, 28, 1–5, 2018
23. B. Lindner. Moments of the First Passage Time under External Driving. *Journal of Statistical Physics*, 117, 703–737, 2004.
24. M. Nashaat, A. E. Botha, and Yu. M. Shukrinov. "Devil's staircases in the IV characteristics of superconductor/ferromagnet/superconductor Josephson junctions", Phys. Rev. B 97, 224514 (2018).
25. V. V Palyulin, G. Blackburn, M. A. Lomholt, N. W. Watkins, R. Metzler, R. Klages and A. V. Chechkin. First passage and first hitting times of Lvy flights and Lévy walks. *New Journal of Physics*, 21, 103028, 2019.
26. F. Paolucci, N. Ligato, V. Buccheri, G. Germanese, P. Virtanen, and F. Giazotto. Hypersensitive tunable Josephson escape sensor for gigahertz astronomy. *arXiv:2003.05966v2*, 2020.
27. V. Pierro, L. Troiano, E. Mejuto and G. Filatrella. Stochastic first passage time accelerated with CUDA. *Journal of Computational Physics*, 361, 136–149, 2018.
28. V.A. Popov, Resonance detection of dark matter axions using a dc SQUID. *Journal of Experimental and Theoretical Physics* 122, 236242, 2016.
29. O. V. Pountougnigni, R. Yamapi , C. Tchawoua, V. Pierro and G. Filatrella. Detection of signals in presence of noise through Josephson junction switching currents. *Physical Review E* 101, 052205-1–10, 2020.
30. L. S. Revin, Andrey L. Pankratov, A. V. Gordeeva, A. A. Yablokov, I. V. Rakut, V. O. Zbrozhek and L. S. Kuzmin. Microwave photon detection by an Al Josephson junction. *Beilstein Journal of Nanotechnology*, 11, 960965, 2020.
31. S. J. Rezvani, D. Di Gioacchino, C. Gatti, C. Ligi, M. Cestelli Guidi, S. Cibella, M. Fretto, N. Poccia, S. Lupi and A. Marcelli. Proximity Array Device: A Novel Photon Detector Working in Long Wavelengths. *Condensed Matter*, 5, 33-1–9, 2020.
32. A. M. Sokolov and F. K. Wilhelm, A superconducting detector that counts microwave photons up to two, *arXiv:2003.04625v1*, 2020.
33. F. Tafuri (Ed.) *Fundamentals and Frontiers of the Josephson Effect*, Springer Series in Materials Science, Volume 286, 2019.

34. SIMP collaboration. Analysis of Josephson junctions switching current distributions for the optimal detection of microwave single photons. Presented at the Erice meeting *New trends in nonequilibrium stochastic multistable systems and memristors*. Ocotober 18-21, 2019, Italy.
35. E. D. Walsh, D. K. Efetov, G. Lee, M. Heuck, J. Crossno, T. A. Ohki, P. Kim, D. Englund and K. Chung. Graphene-Based Josephson-Junction Single-Photon Detector, *Physical Review Applied*, 8, 024022-1–11, 2017.
36. A .A . Yablokov , V.M. Mylnikov, A.L. Pankratov, E.V. Pankratova and A.V. Gordeeva. Suppression of switching errors in weakly damped Josephson junctions, *Chaos, Solitons and Fractals*, 136, 109817-1–6, 2020.
37. J. Yan, C. Beck. Nonlinear dynamics of coupled axion–Josephson junction systems. *Physica D: Nonlinear Phenomena* 403, 132294, 2020.

Control for Set-Valued Movements of Dynamical Systems under Uncertainty with Applications

Tatiana F. Filippova

Krasovskii Institute of Mathematics and Mechanics, Russian Academy of Sciences, Ural Federal University, Ekaterinburg, Russian Federation
(E-mail: ftf@imm.uran.ru)

Abstract. The guaranteed control problems for nonlinear dynamical systems with uncertainty in initial states and parameters are studied. The case is investigate when only the bounding sets for initial system states and for system parameters are given without any additional statistical or probabilistic information on these values. Applying the previously developed approaches and new results developed here to evaluating trajectory tubes and reachable sets, we study the properties of optimal control that solves the problem of control for the trajectory tube of a dynamic system with uncertainty and nonlinearity of a quadratic type.
Keywords: Nonlinear dynamics, Control, Estimation, Uncertainty, Ellipsoidal calculus, Funnel equations.

1 Introduction

The paper investigates the problems associated with the study of reachable sets of a nonlinear control dynamical system (and of a corresponding differential inclusion) with incomplete information on the initial states of the system or on other system parameters, limited by specifying only some special sets containing the unknown elements (Kurzhanski[14], Kurzhanski and Varaiya[16], Allgöwer and Zheng[1], Milanese *et al.*[18], Scweppe[22], Walter and Pronzato[23]). As indicated in many studies, the geometry of the reachable sets of nonlinear dynamical systems may be very complicated. In these cases, the approximation of reachable sets by domains of a certain canonical form is of interest. As such canonical figures, the most natural are ellipsoids, parallelepipeds, polyhedra and some other canonical figures. A number of important approaches are relevant for assessing the unknown states of control systems and corresponding trajectory tubes of differential inclusions through approximation by canonical sets and tubes of motions with an accurate description of their parameters and dynamic characteristics (Kurzhanski and Valyi[15], Chernousko[5], Kostousova[12], Polyak *et al.*[21]).

Currently the principal facts and results of the theory of linear differential systems with uncertain parameters are well developed, a number of important

13^{th} *CHAOS Conference Proceedings, 9 - 12 June 2020, Florence, Italy*
C. H. Skiadas (Ed)

and computationally useful algorithms have been constructed for finding the external and internal (with respect to the inclusion of sets) approximations of the set-valued states of dynamical systems in the case of a linear system dynamics. However the presence of nonlinear terms in the state velocities of the control systems causes a loss of the convexity of the reachable sets and, therefore, raises many theoretical questions and therefore requires the development of related mathematical tools and algorithms that are adequate to the indicated problems of nonlinear analysis. Some ideas and approaches to the study of set-valued motions (trajectory tubes) for a number of differential systems with nonlinearity and uncertainty in dynamics were presented earlier in Filippova[7], Filippova and Lisin[8], Filippova and Matviychuk[9] (see also references in the indicated publications).

In this paper we assume that in a dynamic system there are two types of nonlinearity, namely, we have a combination of bilinear and quadratic functions in the state velocities. Earlier, we examined the problems of evaluating the reachable sets of systems under study taking into account all possible controls at once. Knowing the areas of reachability with respect to all parameters of the system under study (for all possible initial states, disturbances, controls) is very useful, since it helps to evaluate the capabilities of the system. However, it seems important to have a description of the trajectory tube generated by a specific choice of a control function, it will allow solving optimization problems for set-valued movements of the considered systems under uncertainty. Note that in this paper we consider a special class of control systems with nonlinearity and uncertainty under other informational assumptions than was done in a recent paper Filippova and Matviychuk[10]. Thus, this research continues and complements developments in the field of mathematical control theory related to the study of the dynamics of multivalued states of nonlinear control systems. The approaches and algorithms presented here may be applied in the study of models with nonlinearity and uncertainty in real systems in robotics, economics, biology and other fields (considered e.g. in Allgöwer and Zheng[1], Bayen and Rapoport[2], Cecarelli *et al.*[4], Keller *et al.*[11]).

2 Problem formulation

2.1 Basic notations

The main notations used in the paper are basic; however, we define here some additional, most frequently used and important constructions.

We denote by $\mathbb{R}^n$ the n-dimensional vector space and by $\operatorname{comp}\mathbb{R}^n$ the set of all compact subsets of $\mathbb{R}^n$. Also $\mathbb{R}^{n\times m}$ denotes the set of all $n\times m$-matrices.

The usual inner product of $x, y \in \mathbb{R}^n$ is $x'y = (x, y) = \sum_{i=1}^n x_i y_i$ with prime as a transpose and also the

$$\|x\| = \|x\|_2 = (x'x)^{1/2}, \quad \|x\|_\infty = \max_{1\le i\le n} |x_i|$$

are corresponding norms for $x \in \mathbb{R}^n$.

For the identity matrix we use the symbol $I \in \mathbb{R}^{n\times n}$. Denote by $\mathrm{Tr}(A)$ a trace (a sum of diagonal elements) of $n \times n$-matrix A. Let $B(a,r) = \{x \in \mathbb{R}^n : \|x-a\| \le r\}$ be a ball in $\mathbb{R}^n$ with a center $a \in \mathbb{R}^n$ and with a radius $r > 0$.

We use here also the notation

$$E(a,Q) = \{x \in \mathbb{R}^n : \ (Q^{-1}(x-a),(x-a)) \le 1\}$$

for the ellipsoid in $\mathbb{R}^n$, where $a \in \mathbb{R}^n$ is its center and a $n \times n$-matrix Q is symmetric and positive definite.

2.2 Main problem

We study here the nonlinear control system

$$\begin{gathered} \dot{x} = A(t)x + f(x)d + u(t), \\ x_0 \in \mathcal{X}_0, \quad t_0 \le t \le T, \end{gathered} \tag{1}$$

here $x, d \in \mathbb{R}^n$, $\|x\| \le K$ $(K > 0)$, the function $f(x)$ is quadratic in x, that is $f(x) = x'Bx$, with a positive definite and symmetric $n \times n$-matrix B.

Functions $u(t)$ ("controls") in (1) are assumed to be Lebesgue measurable on $[t_0, T]$ and

$$u(t) \in \mathcal{U}, \quad for \ \ a.e.\ t \in [t_0, T].$$

We assume that the constraint set $\mathcal{U}$ is given and $\mathcal{U} \in \mathrm{comp}\,\mathbb{R}^n$. The $n \times n$-matrix function $A(t)$ in (1) has the form

$$A(t) = A^0 + A^1(t), \tag{2}$$

where the $n \times n$-matrix A^0 is given and the measurable $n \times n$-matrix $A^1(t)$ is unknown but bounded, $A^1(t) \in \mathcal{A}^1$ for $t \in [t_0, T]$, namely we have

$$A(t) \in \mathcal{A} = A^0 + \mathcal{A}^1, \tag{3}$$

$$\mathcal{A}^1 = \big\{A{=}\{a_{ij}\}{\in}\mathbb{R}^{n\times n} : |a_{ij}|{\le}c_{ij},\ i,j{=}1,\dots n\big\},$$

where $c_{ij} \ge 0$ $(i,j = 1,\dots n)$ are given numbers. The latter relations mean that all elements of the matrix $A(t)$ are known only up to certain errors, the values of which are given (this does not exclude the case when some elements of the matrix can be known exactly, this corresponds to the situation when some $c_{ij} = 0$).

Assume that we have the ellipsoid as an initial set $\mathcal{X}_0$ in (1), that is

$$\mathcal{X}_0 = E(a_0, Q_0),$$

with a symmetric and positive definite matrix $Q_0 \in \mathbb{R}^{n\times n}$ and with a center a_0.

If it will be necessary we will use also a notation $x(t; u(\cdot)) = x\big(t; u(\cdot), A(\cdot), x_0\big)$ with indication of additional parameters $A(\cdot), x_0$ for an absolutely continuous function $x(t)$ which is the solution to (1)–(3) with initial state $x_0 \in \mathcal{X}_0$, with admissible control $u(\cdot)$ and with a matrix $A(\cdot)$ satisfying (2)–(3).

Due to the fact that some quantities are unknown but bounded, we are forced to consider all possible versions of motions compatible with additional data as a generalized solution to the control system, that is, we need to replace a single-valued trajectory by a bundle or tube of motions of the following form $\mathcal{X}(t;u(\cdot))$.

Definition 1. For each admissible control $u(\cdot)$ the *generalized solution tube* $\mathcal{X}(t;u(\cdot))$ (with $t\in[t_0,T]$) of system (1)–(3) is defined as follows,

$$\mathcal{X}(t;u(\cdot))=\{x\in\mathbb{R}^n:\exists\, x_0\in\mathcal{X}_0,\ \exists\, A(\cdot)\in\mathcal{A},$$
$$x=x(t)=x\big(t;u(\cdot),A(\cdot),x_0\big)\}.$$

Let us consider the following main problems.

Problem 1. For each feasible control $u(\cdot)\in\mathcal{U}$, find the optimal external ellipsoidal estimate $E(\hat{a},\hat{Q};T,u(\cdot))$ of the reachable set $X(T;u(\cdot))$ of the system (1)-(4), such that

$$\mathcal{X}(T;u(\cdot))\subset E(\hat{a},\hat{Q};T,u(\cdot)).$$

Remark 1. Here we understand the optimality of the desired ellipsoidal estimate, bearing in mind the closest operation with respect to inclusion of related sets.

Problem 2. Given a vector $x^*\in R^n$ find the feasible control $u^*(\cdot)\in\mathcal{U}$ such that the related ellipsoidal estimate is optimal, that is we have

$$d(x^*,E(\hat{a^*},\hat{Q}^*;T,u^*(\cdot)))=\inf_{u(\cdot)\in\mathcal{U}}d(x^*,E(\hat{a^*},\hat{Q}^*;T,u(\cdot)))=\epsilon^*.$$

3 Main results

First, we define an auxiliary parameter k, which is required to formulate the main result (see also Filippova[7]). To do this, consider the matrix $B^{1/2}Q_0B^{1/2}$ and denote its maximal eigenvalue as k^2, that is we have

$$E(a_0,Q_0)\subseteq E(a_0,(k_0^+)^2B^{-1}),\tag{4}$$

and k_0^+ is the smallest positive number for which this estimate (4) is true.

Theorem 1. *The upper ellipsoidal estimate is true*

$$\mathcal{X}(t_0+\sigma;u(\cdot))\subseteq E(a^*(t_0+\sigma),Q^*(t_0+\sigma)\mid u(\cdot))+o(\sigma)B(0,1)\tag{5}$$

with $\sigma^{-1}o(\sigma)\to 0$ *for* $\sigma\to+0$ *and*

$$a^*(t_0+\sigma)=\tilde{a}(t_0+\sigma)+\sigma(\hat{a}+a_0'Ba_0\cdot d+k^2d)+\sigma u(t_0),\tag{6}$$

and with functions $\tilde{a}(t)$, $Q^*(t)$ *satisfying the following equations*

$$\dot{\tilde{a}}=\tilde{A}^0\tilde{a},\quad t_0\le t\le T,\ \tilde{a}(t_0)=a_0,\tag{7}$$

$$\dot{Q}^* = \tilde{A}^0 Q^* + Q^*(\tilde{A}^0)' + qQ^* + q^{-1}G,\ Q^*(t_0) = Q_0,\quad t_0 \le t \le T, \tag{8}$$

where

$$\tilde{A}^0 = A^0 + 2d \cdot a_0' B,\quad q = \big(n^{-1}\,\mathrm{Tr}\,((Q^*)^{-1}G)\big)^{1/2}, \tag{9}$$

$$G = \mathrm{diag}\Big\{(n-v)\big[\sum_{i=1}^{n} c_{ji}|\tilde{a}_i| + \big(\max_{\sigma=\{\sigma_{ij}\}} \sum_{p,q=1}^{n} Q^*_{pq} c_{jp} c_{jq} \sigma_{jp}\sigma_{jq}\big)^{1/2}\big]^2\Big\}, \tag{10}$$

with a maximum in (10) *calculated over numbers* $\sigma_{ij} = \pm 1$, $i, j = 1, \ldots, n$, *such that we have* $c_{ij} \neq 0$ *and* v *is a number of such indices* i *for which* $c_{ij} = 0$ *for all* $j = 1, \ldots, n$.

Proof. The relation (5) is established along the main lines and ideas presented in Filippova[7]. Indeed, from the funnel equation Panasyuk[20] we have

$$X(t_0 + \sigma; u(\cdot)) \subseteq \bigcup_{\tilde{x} \in E(0, k_0^{+2} B^{-1})} (a_0 + \tilde{x} + \sigma(A_0 + \mathcal{A}_1)(a_0 +$$

$$\tilde{x}) + \sigma(a_0 + \tilde{x})' B(a_0 + \tilde{x})) + \sigma u(t_0) + o(\sigma) B(0,1). \tag{11}$$

We remind that we may use here the property that at the boundary points $\tilde{x}$ of the ellipsoid $E(0, (k_0^+)^2 B^{-1})$ we have the equality $\tilde{x}' B \tilde{x} = (k_0^+)^2$ (for a more simple case detailed explanations of the last property may be found also in Filippova[7]). With this property and rearranging the terms in (11), we come to the formulas (5)-(10). ■

Remark 2. We see here that the ellipsoidal estimates of the tube $X(t; u(\cdot))$ for each fixed control $u(\cdot)$ are under investigation here and therefore the parameters of the estimation procedures depend on $u(\cdot)$. We can complicate the problem by additionally assuming the presence of state constraints or by considering a slightly more general class of uncertainty, e.g. in the coefficients of the matrix of linear terms of the state velocities.

Remark 3. It follows from Theorem 1 that we can construct a discrete tube $E(\hat{a}, \hat{Q}; T, u(\cdot))$ with ellipsoidal cross-sections that solves Problem 1 and for which we have the inclusion

$$X(T; u(\cdot)) \subseteq E(\hat{a}^+(T), \hat{Q}^+(T); u(\cdot)) + o(\epsilon) B(0,1). \tag{12}$$

We emphasize that this discrete construction may be used as a basis for related computational schemes and algorithms allowing to find the trajectory tubes numerically.

Using the results Filippova and Matviychuk[9], we may derive the following result.

Theorem 2. *Let* ϵ^*, $u^*(\cdot)$ *be the optimal values of the Problem 2. Then we have the relations*

$$\epsilon^* = \min_{u(\cdot) \in \mathcal{U}} \max_{||l||=1} \{r^+(T; u(\cdot))(l' B^{-1} l)^{1/2} +$$

$$l'(a^+(T; u(\cdot)) - x^*)\} = \max_{||l||=1} \{r^+(T; u^*(\cdot))(l' B^{-1} l)^{1/2} + \tag{13}$$

$$l'(a^+(T; u^*(\cdot)) - x^*)\}.$$

Proof. First, we find the minimal positive number ϵ such that the following inclusion is true

$$E(a^+(T), Q^+(T); T, u(\cdot)) \subseteq B(x^*, \epsilon),$$

or equivalently

$$\rho(l|E(a^+(T), Q^+(T); T, u(\cdot)) \leq \rho(l|B(x^*, \epsilon)), \quad \forall l \in \mathbb{R}^n.$$

Appling the result of Theorem 1, we get the relation

$$l'a^+(T) + (l'Q^+(T)l)^{1/2} \leq l'x^* + \epsilon||l||,$$

and from the above relations we conclude that

$$\epsilon^* = \min_{u(\cdot)} \max_{||l||=1} ((l'Q^+(T)l)^{1/2} + l'(a^+(T) - x^*)).$$

Taking into account the equality $Q^+(T) = r^+(T)B^{-1}$ we get the equations (13).■

The proposed results may be used as the basis for the development of computational algorithms for solving applied problems of controlling and estimating the movements of real systems operating in conditions of uncertainty and nonlinearity, in particular, in the fields of robotics, economics and finance, biology and other fields. Related algorithms with computational examples (for lower dimensional systems) that illustrate the approach may be found e.g. in Filippova and Matviychuk[9]. In the next section a more complicated example of a dynamical system in the space $\mathbb{R}^3$ is given and discussed.

4 Numerical simulations

Example. Consider the following control system

$$\begin{cases} \dot{x}_1 &= -x_1 + x_1^2 + x_2^2 + 2x_3^2 + u_1(t), \\ \dot{x}_2 &= x_2 + u_2(t), \\ \dot{x}_3 &= x_3 + u_3(t), \end{cases} \tag{14}$$

Assume that $\mathcal{U} = B(0,1)$, $x_0 \in X_0 = B(0,1)$ and $t \in [0,T]$ with $T = 0.4$. The projections of reachable sets $X(t)$ together with related estimating ellipsoids $E^+(t) = E(a^+(t), Q^+(t))$ onto the planes of state coordinates (related planes are (x_1, x_2), (x_1, x_3) and (x_2, x_3), respectively) are shown in Fig. 1-3 for time grid $t = 0.1; 0.15; 0.2; 0.25; 0.3; 0.35; 0.4$ (we need to specify here that for simplicity we put $u(t) = 0$ here, in other cases calculations and pictures are similar).

The last Fig. 4 shows the upper estimating ellipsoid $E^+(t) = E(a^+(t), Q^+(t))$ and the reachable set $X(t)$ as they are in the related space $\mathbb{R}^3$ of state variables $\{x_1, x_2, x_3\}$ for $t = 0.4$.

Note that the evaluating ellipsoid touches the reachable set (that is, the external estimate is tight), which implies that without changing the structure of parameters (for example, without changing the main matrix of coefficients), it cannot be reduced to a smaller ellipsoid.

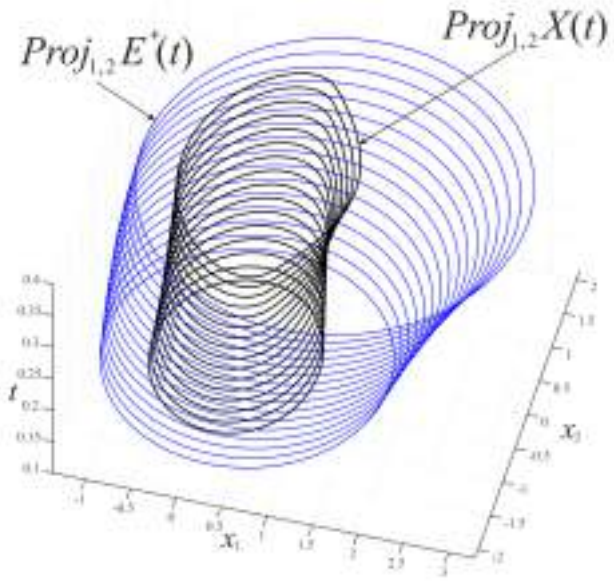


Fig. 1. Projections $Proj_{1,2}E^{+}(t)$ of ellipsoids $E^{+}(t) = E(a^{+}(t), Q^{+}(t))$ (blue color) and projections $Proj_{1,2}X(t)$ of reachable sets (black color) $X(t)$ at the plane of $\{x_1, x_2, t\}$-coordinates.

5 Further theoretical directions and possible applications

Theoretical schemes and related numerical algorithms for evaluating trajectory tubes and methods for solving control problems for set-valued motions based on Theorems 1-2 can be developed further in many directions, among them we note the following areas:

- studies of optimization and robust stabilization problems for uncertain nonlinear systems with impulsive control functions,
- problems of viability and control for dynamical systems described by nonlinear differential equations and differential inclusions,
- improvement and development of new numerical methods for estimating set-valued motions of nonlinear dynamical systems (ensembles of trajectories) based on the proposed ideas for high-dimensional systems,
- research of new, more complex classes of nonlinearity in the dynamics of controlled systems with uncertain factors,
- development of theoretical approaches to the estimation of set-valued motions using approximations for set-valued motions based on the use of discrete schemes of the theory of differential inclusions with a large order of accuracy.

The applications of the problems discussed here are in the nonlinear control and estimation theory and related nonlinear models with unknown but bounded errors. Numerous application models can be noted here, in particular, real models in robotics, in transportation systems, in biology, medicine and

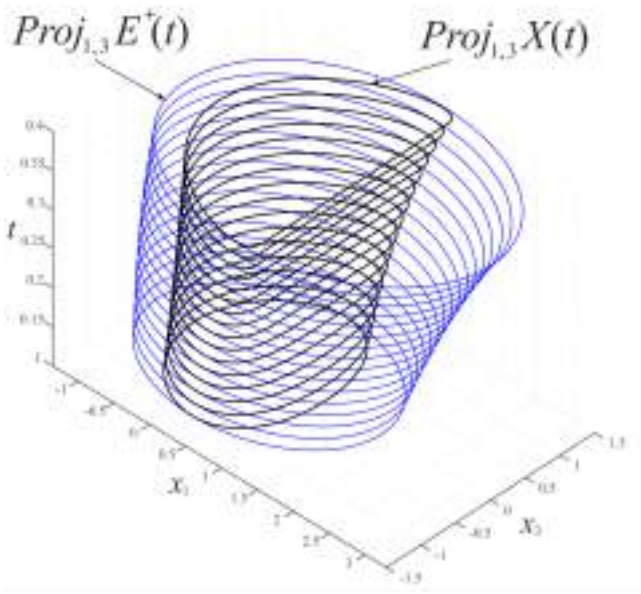


Fig. 2. Projection $Proj_{1,3}E^{+}(t)$ of ellipsoids $E^{+}(t) = E(a^{+}(t), Q^{+}(t))$ (blue lines) and projections $Proj_{1,3}X(t)$ of reachable sets (black lines) $X(t)$ at the plane of $\{x_1, x_3, t\}$-coordinates.

economics. In these aspects, we would like to highlight, in particular, the studies and results obtained earlier by Bayen and Rapoport[2], Cecarelli *et al.*[4], Koller *et al.*[11]), Filippova and Matviychuk[9], Kuntsevich and Volosov[13], Malyshev and Tychinskii[17], Ovsyannikov[19].

6 Conclusion

The paper deals with the state estimation problems for uncertain dynamical control systems for which we assume that the initial state is unknown but bounded with given constraints. We consider here a special case of uncertainty and nonlinearity when the matrix parameters in state velocities are unknown but bounded.

The system nonlinearity under study is generated also by the presence of bilinear terms and quadratic forms in related differential equations. The problem is reformulated as the control problem for the motion of related set-valued states.

Using the ideas developed earlier for some classes of uncertain systems we solve here the control problem with a new class of uncertainty and with a special structure of nonlinearity. So we construct the external ellipsoidal estimates of reachable sets for the system under study and find the solution of the related optimization problem.

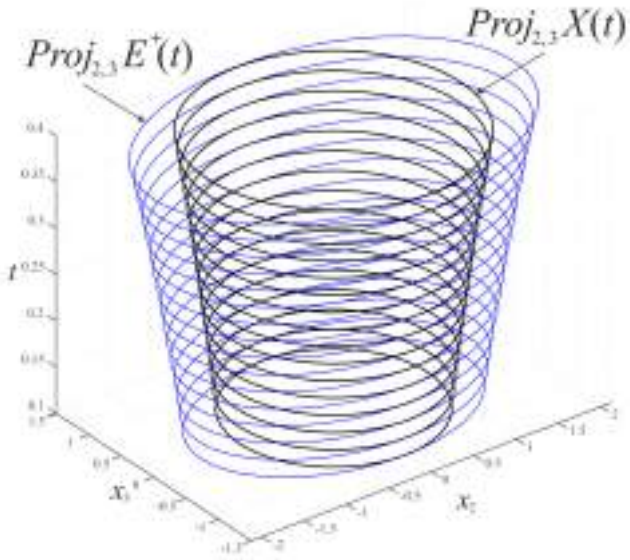


Fig. 3. The projections $Proj_{2,3}E^{+}(t)$ of estimating ellipsoids $E^{+}(t) = E(a^{+}(t), Q^{+}(t))$ (indicated in blue lines) and projections $Proj_{2,3}X(t)$ of reachable sets (indicated in black lines) $X(t)$ at the plane of $\{x_2, x_3, t\}$-coordinates.

Acknowledgements

The study was partially supported by the Russian Foundation for Basic Researches (RFBR Project No.18-01-00544).

References

1. F. Allgöwer and A. Zheng. (Eds.) *Nonlinear Model Predictive Control*, Birkhauser, Basel, 2000.
2. T. Bayen and A. Rapaport. Minimal time crisis versus minimum time to reach a viability kernel: a case study in the prey-predator model. *Optimal Control Applications and Methods*, Wiley, 40(2), 330–350, 2019.
3. R. W. Brockett. On the reachable set for bilinear systems. *Lecture Notes in Economics and Mathematical Systems*, 111, 54–63, 1975.
4. N. Ceccarelli, M. Di Marco, A. Garulli and A. Giannitrapani. A set theoretic approach to path planning for mobile robots. *Proc. 43rd IEEE Conference on Decision and Control, Atlantis, Bahamas, Dec. 2004*, 147-152, 2004.
5. F. L. Chernousko. *State Estimation for Dynamic Systems*, CRC Press, Boca Raton, 1994.
6. F. L. Chernousko and D. Ya. Rokityanskii. Ellipsoidal bounds on reachable sets of dynamical systems with matrices subjected to uncertain perturbations. *Journal of Optimization Theory and Applications*, 104, **1**, 1-19, 2000.
7. T. F. Filippova. Differential equations of ellipsoidal estimates for reachable sets of a nonlinear dynamical control system. *Proceedings of the Steklov Institute of Mathematics (Supplementary issues)*, 271, suppl.1, S75-S84, 2010.

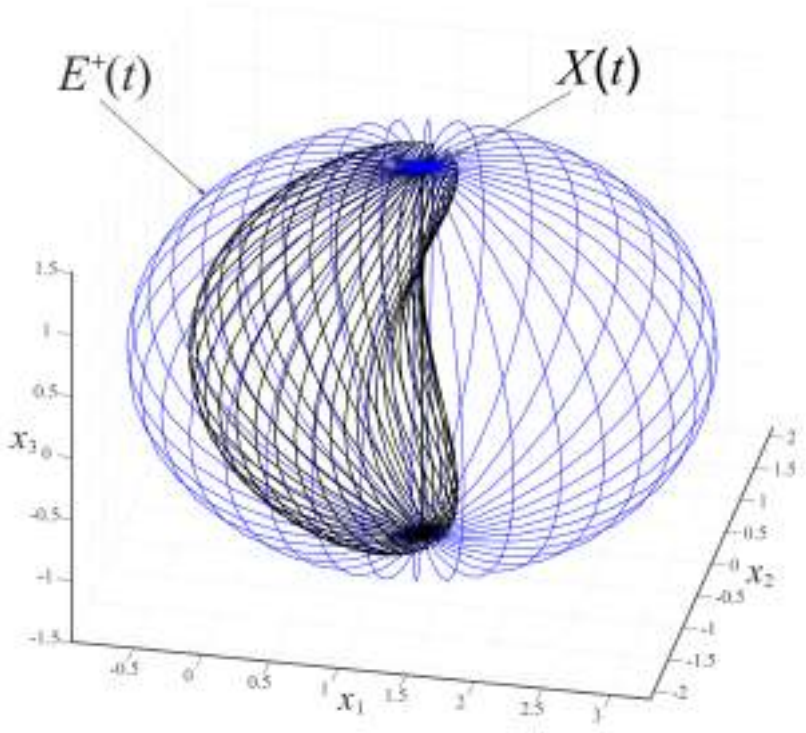


Fig. 4. Reachable set $X(t)$ and its upper ellipsoidal estimate $E^+(t) = E(a^+(t), Q^+(t))$ for $t = 0.4$ (3d-picture in the plane of state variables $\{x_1, x_2, x_3\}$).

8. T. F. Filippova and D. V. Lisin. On the estimation of trajectory tubes of differential inclusions. *Proc. Steklov Inst. Math.: Problems Control Dynam. Systems*. Suppl. Issue 2, S28–S37, 2000.
9. T. F. Filippova and O. G. Matviychuk. Approaches to estimating the dynamics of interacting populations with impulse effects and uncertainty. In: C. Skiadas and Y. Dimotikalis (Eds.), *12th Chaotic Modeling and Simulation International Conference, CHAOS 2019*, Springer Proceedings in Complexity, Springer, Cham, 85–99, 2020.
10. T. F. Filippova and O. G. Matviychuk. Control problems for set-valued motions of systems with uncertainty and nonlinearity. In: A. Tarasyev, V. Maksimov and T. Filippova (Eds.), *International Conference Stability, Control, Differential Games, SCDG2019, September 16 20, 2019, Yekaterinburg, Russia*, Lecture Notes in Control and Information Sciences - Proceedings, Springer, Cham, 379–389, 2020.
11. T. Koller, F. Berkenkamp, M. Turchetta and A. Krause. Learning-based model predictive control for safe exploration and reinforcement learning, arXiv:1803.08287v3 [cs.SY], https://arxiv.org/pdf/1803.08287.pdf , 1–9, 2018.
12. E. K. Kostousova. On polyhedral estimates of reachable sets of discrete-time systems with uncertain matrices and integral bounds on additive terms. In: Y. Sergeyev and D. Kvasov (Eds.), *Numerical Computations: Theory and Algorithms. NUMTA 2019*, Lecture Notes in Computer Science, Springer, Cham,

11974, 124–138, 2000.
13. V. M. Kuntsevich and V. V. Volosov. Ellipsoidal and interval estimation of state vectors for families of linear and nonlinear discrete-time dynamic systems. *Cybernetics and Systems Analysis*, 51(1), 64–73, 2015.
14. A. B. Kurzhanski. *Control and Observation under Conditions of Uncertainty*, Nauka, Moscow, 1977.
15. A. B. Kurzhanski and I. Valyi. *Ellipsoidal Calculus for Estimation and Control*, Birkhauser, Boston, 1997.
16. A. B. Kurzhanski and P. Varaiya. *Dynamics and Control of Trajectory Tubes. Theory and Computation*, Springer-Verlag, New York, 2014.
17. V. V. Malyshev and Yu. D. Tychinskii. Construction of attainability sets and optimization of maneuvers of an artificial earth satellite with thrusters in a strong gravitational field. *Proc. of RAS, Theory and Control Systems*, 4, 124-132, 2005.
18. M. Milanese, J. P. Norton, H. Piet-Lahanier and E. Walter (Eds.). *Bounding Approaches to System Identification*, Plenum Press, New York, 1996.
19. D. A. Ovsyannikov. Mathematical modeling and optimization of beam dynamics. *IFAC Proceedings Volumes*, 34(6), 13–22, 2001.
20. A. I. Panasyuk. On the equation of an integral funnel and its applications. *Differ. Equations*, 24(11), 1263-1271, 1988.
21. B. T. Polyak, S. A. Nazin, C. Durieu and E. Walter. Ellipsoidal parameter or state estimation under model uncertainty. *Automatica*, 40, 1171-1179, 2004.
22. F. Schweppe. *Uncertain Dynamic Systems*, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.
23. E. Walter and L. Pronzato. *Identification of Parametric Models from Experimental Data*, Springer-Verlag, Heidelberg, 1997.

Battle of Salamis: Greeks were destined to win

Konstantina Founta[1], Loukas Zachilas[2]

[1] Department of Economics, University of Thessaly, Volos, Greece
(E-mail: kfounta@uth.gr)
[2] Department of Economics, University of Thessaly, Volos, Greece
(E-mail: zachilas@uth.gr)

Abstract. In this paper we present a new non–linear, discrete, dynamical system trying to model the historic battle of Salamis (480 BC) between Greeks and Persians. The model describes the most effective strategic behavior between two participants during a battle or in a war. Moreover, we compare the results of the dynamical analysis to Game Theory, considering this conflict as a dynamic game.
Keywords: Battle of Salamis, Discrete Systems, Modeling Strategic Behavior, Game Theory.

1 Introduction

The model approaches short–term conflicts between two participants (players), where one is weaker than the other opponent. Also, the parameters (that we use in Eq. 1, see below) are the most crucial factors in order to highlight the optimal way to achieve a decisive victory.
The solution of the equations of the dynamical system (called equilibrium point) shows what kind of behavior each player should adopt, i.e. an aggressive or defensive one. In this way, according to Game Theory, the optimum strategic behavior is called Nash Equilibrium [9].
One of the most representative games of Evolutionary Game Theory is the so-called game "Hawk – Dove", which was originally developed by Smith and Price [10] to describe animal conflicts and is quite similar to our attempt. We will present below the game and its results.
The game "Hawk – Dove" has many applications in everyday life. There are two animals (or two players) fighting for the same resource. Each of them can behave either as a hawk (i.e. fight for the resource) or as a dove (i.e. abandon the resource before the conflict escalates into a fight). Individuals have a benefit B if they win and a cost C if lose.
If a Hawk meets a Hawk, they will fight and one of them will win the resource; average payoff is *(B-C)/2*. If a Hawk meets a Dove, the Dove immediately withdraws, so the payoff of the Dove is *zero*, while the payoff of the Hawk is *B*. If a Dove meets a Dove, the one who first gets hold of the resource keeps it, while the other does not fight for it; average payoff *B/2*. The strategic form of the game is given by the payoff matrix:

$$Payoff_{H,D} = \begin{pmatrix} (B-C)/2 & B \\ 0 & B/2 \end{pmatrix}$$

2 Solution of the game "Hawk – Dove"

We set the benefit $B = 2$ if a player wins, and the cost $C = 1$ if a player loses. Using the Gambit[1] software (16.0.1), we find Nash equilibriums and the dominant strategy.

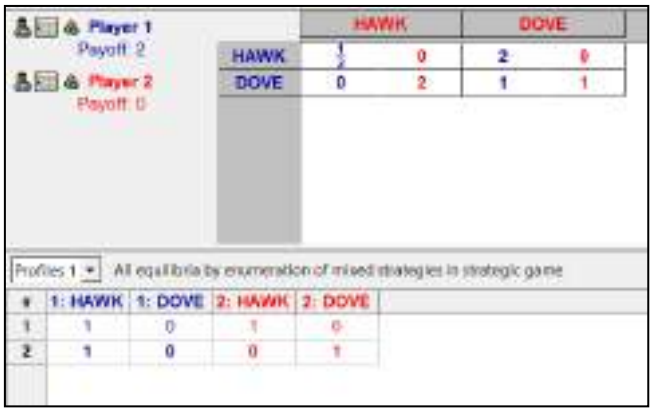


Fig. 1: The results of "Hawk - Dove" game.

Fig. 1 shows us the payoff matrix and the two Nash equilibriums. If both players behave as a Hawk, the one who first injures the other wins. We set the player 1 starts and injures the player 2, thus player 1 wins. If someone behaves as a Hawk and the other behaves as a Dove, then the player with the aggressive behavior (Hawk) wins and takes all the resource. If both players behave as a Dove, then they share the resource.

Regarding Nash equilibriums, there are two pure strategies. On the one hand, both players behave as Hawks and on the other hand, player 1 behaves as a Hawk and player 2 as a Dove. Moreover, we can observe that player 1 behaves as a Hawk in both cases and player 2 behaves either as a Hawk or as a Dove, but in each case player 1 wins.

We should note that if player 2 injures first player 1, the Nash equilibriums would be different.

[1] McKelvey, Richard D., McLennan, Andrew M., and Turocy, T. L., 2014.

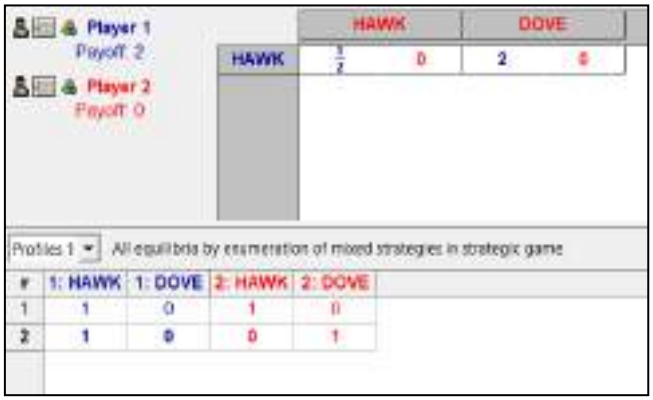


Fig. 2: Dominant Strategy.

Fig. 2 shows the dominant strategy of the game, where player 1 behaves as a Hawk independently of the player's 2 behavior (i.e. Hawk or Dove). Therefore, the first dominant strategy may not be effective, because both players behave as Hawks and player 1 wins the half of the resource and does not maximize his profit. Although, if the player behaves as a Hawk, knowing that the other player behaves as a Dove, then he takes all the resource (maximum profit). Thus, we believe that the second Nash equilibrium is more effective and optimum strategy.

3 The Dynamical Model

It is widely acknowledged that the military strategy is the combination of ends, ways and means [7]. In our attempt to study the strategic behavior of two warring parties, we developed a new non-linear discrete system of two equations based on the above phrase. The main objective of the model is to simulate the way by which the two opponents behave strategically, where the one is weaker than the other.
At the same time, in Game Theory, the war is considered as a dynamic game where the strategies of the players are studied by calculating their optimal strategy (Nash equilibrium). In this research, we tried to compare the results of the Game Theory with those from the analysis of the discrete dynamical system. At the end of the analysis, the optimum and effective strategy for both participants (players) will be suggested.
The model, which is applied in short-term conflicts and describes the strategic behavior of each participant, is given by Eq. 1:

$$\begin{cases} x_{t+1} = P_x + TN_x - G \cdot \left(D_y + E_x\right) \cdot 4y_t \cdot (1 - y_t) \\ y_{t+1} = P_y + TN_y - (1 - G) \cdot \left(D_x + E_y\right) \cdot 4x_t \cdot (1 - x_t) \end{cases} \quad [1]$$

where:
x_t: The strategic behavior of any form of social organization x (state, country, etc.) at the time t.

y_t: The strategic behavior of any form of social organization *y* (state, country, etc.) at the time *t*.
x_{t+1}: The optimal strategic behavior of any form of social organization *x* at the (next moment of) time *t* + 1.
y_{t+1}: The optimal strategic behavior of any form of social organization *y* at the (next moment of) time *t* + 1.
We consider $x_t, y_t, x_{t+1}, y_{t+1} \in [0,1]$, because the logistic equation is defined in [0,1], which is derived from the study of biological populations reproduced in discrete time [6]. It's the evolution of the population model of Malthus [8] and shows that the exponential growth cannot tend to infinity, but there is a critical point, i.e. a saturation. In other words, it is not possible for someone to win and the other to lose continuously. Also, each optimal strategic behavior, at the time *t*, affects the next move – strategic behavior, at the time *t* + 1, of the opponent.
In addition, we can interpret the values of variables (and parameters, as shown below) as percentages or probabilities, which help us to explain the results; these are also explained through the Game Theory.
Moreover, if the value of x_{t+1} (or y_{t+1}, respectively) equals to 0, it indicates the fully defensive strategic behavior of participant *x* (or *y* respectively), while if it equals to 1, then it indicates the fully aggressive behavior of participant *x* (or *y* respectively).
The parameters of Eq. 1 are the main and most important factors that could affect the strategic behavior of *x* (or *y*, respectively). In particular:
The parameter $\boldsymbol{P_x}$ represents the strength (economic, military, population, territorial) of *x* and $\boldsymbol{P_y}$ is the strength of *y*, respectively. These two parameters indicate the substance of each form of social organization compared to the other.
$\boldsymbol{TN_x}$ and $\boldsymbol{TN_y}$ represents the Technological Naval capability and evolution of *x* and *y*, respectively. These two parameters are also defined in comparison with the technological capability and evolution of the other participant and describe the *means* mentioned by [7].
The parameter $\boldsymbol{G}$ represents the geographical location (geophysical terrain) of the area where the battle or the war is taking place. We believe that this is another part of the military strategy, namely the *ways* [7]. Trying to emphasize the importance of this parameter and how it can be an advantage or disadvantage for each participant, we set in the first equation as $\boldsymbol{G}$ and in the second equation as $\boldsymbol{1 - G}$. The closer to the 1 the value of the parameter, the easier the geophysical terrain of the area is.
The parameter $\boldsymbol{D_x}$ represents the damages caused by *x* to *y* and respectively, $\boldsymbol{D_y}$ represents the damages that *y* brings to *x*. The damages which we refer to may be economic, territorial, military, etc. or even deception and damaging of the psychological part of the opponent. Moreover, these two parameters complete the last part of the military strategy, namely the *ends* [7].
The parameter $\boldsymbol{E_x}$ represents the expenses of participant *x* and $\boldsymbol{E_y}$ the expenses of participant *y*, respectively. In other words, these denote the preparation costs of each participant for a battle (or war), compared to each other.

All the parameters that have been presented above should belong to [0,1]. Namely, $P_x, P_y, TN_x, TN_y, G, D_x, D_y, E_x, E_y \in [0,1]$.

In the next section, we present the dynamic analysis and the results from the application of Eq. 1 in naval battle of Salamis.

4 The case of (naval) Battle of Salamis

The naval battle of Salamis was an important battle of the second Persian invasion in Greece and has been estimated to being held on September 28th, 480 BC in the Salamis straits (in the Saronic Gulf near Athens). The two warring parties were the Greeks (Hellenic alliance) and the Persian Empire [2].

After the fall of Thermopylae, the Persians proceeded to Athens. The Greeks had been advised by the Oracle of Delphi, that only the "wooden walls" would save them and they considered that this referred to a fight in the sea [5].

A few days before the battle, the meeting of the Greek admirals had to decide the geographic location of the battle. On the one hand, the Spartan General Evriviades proposed to fight in the Isthmus of Corinth, under the main argument that in case of failure it would be possible for them to continue to fight into the center of the Peloponnese. On the other hand, the Athenian General Themistocles insisted to fight in Salamis straits. He believed that if he forced the Persians to attack there, the numerous Persian ships couldn't extent highlighting their dominance. Ultimately, the council considered that Themistocles' argument was better and decided to support it [1].

The Greek fleet was estimated by Herodotus in 380 triremes and Aeschylus gave a round 300 triremes, but we can't be certain for the exact number. On contrary, the Persian fleet was estimated in 500-600 triremes[2]. Herodotus describes the Persian ships as "*better sailing*", when compared to the Greek fleet. This may be attributable to a combination of factors such as lightness of materials and structure of the ship, better seamanship and more extensive naval experience. The triremes of Hellenic alliance were heavier and more durable. However, Herodotus reports that these ships were equipped with an embolism, with which they sank the enemy ships. They used two attacking maneuvers: *diekplous*, (i.e. attack from the rear or sides with a sharp turn) and *periplous*, (flanking or enveloping move, which generally gave an extra benefit against superior numbers in open water). The purpose of both was to ram the enemy in the side. In this way, they achieved serious damages or even the complete destruction of the Persians ships. On the contrary, the Persian tactic was "ramming and boarding" [11].

[2] Aeschylus, writing decades earlier, also gives 1,207 triremes, but Herodotus writes, shortly before battle took place, that the Persian fleet wasn't much bigger than Greek. Because of a weather phenomenon (storms) 600 ships sank (400 at the coast of Magnesia, north of Artemisium and 200 in Euboea).

Fig. 3: The battle of Salamis.
Source: Burn, A. R. (1962). Persia and the Greeks, New York: Minerva Press

At dawn (if the date of the battle was indeed 28th Sept.), the two fleets were ready for the naval conflict. Xerxes, sure of his victory, sat on a throne on mountain Aigaleo, to enjoy the war spectacle. The narrowness of the space and the limited extent of the sea did not allow the Persians to use the major of their force in the front line. Thus, the number of ships was approximately equal. In this naval battle, the bravery and dexterity of the Greek fleet played an important role. They fought aggressively to defend their moral values and their freedom [1].

Herodotus reports that "*the Greeks fought with discipline and held their formation, but the Persians did not seem to be following any plan, so things were bound to turn out for them as they did*". Also, Aeschylus mentions that Themistocles must be given the credit for their battle and the winning tactics. The turning point of the battle came as the Persians "*suffered their greatest losses when the ships in their front line were put to fight and those following, pressing forward to impress the King* (i.e. Xerxes) *with their deeds, became entangled with them as they tried to escape*", as Herodotus comments [11].

The naval battle evolved rapidly and by the noon it was visible that the Greeks would win. The Persian fleet had crushed, while the Greek fleet continued to haunt them, killing the helpless, non–swimming soldiers. This brought the battle to an end, leaving the Greek force in full control of the straits [1].

When the battle was over, a Roman source mentions that Greeks lost more than 40 triremes and Persians more than 200 ones [11]. The victory of the Greek force was of major importance, since they managed to cause the collapse of the Persian morale, which is evidenced by the abandonment of the battle. In addition, the right

decision of Themistocles for the geographic location of the naval battle was one of the most intelligent movements to bring the Greek victory.

5 Applying the model in naval battle of Salamis – Approaching the reality

Starting the dynamical analysis of the naval battle of Salamis, we set the initial conditions in Eq. 1, which represent as much as possible the historical events of the battle. Specifically:

(a) We set Greeks as the weak participant – player (x) and Persians as the powerful participant – player (y).
(b) The strength of Hellenic alliance, $P_x = 0.25$ and the strength of Persian empire, $P_y = 0.8$.
(c) The technological naval capability of Greeks, $TN_x = 0.7$ and the technological naval capability of Persians, $TN_y = 0.35$.
(d) The geographic location of the naval battle, G = 0.4, i.e. the Salamis straits, which are an advantage point for the Greek fleet.
(e) The damage caused to Persian side was huge, so we set $D_x = 0.8$ and $D_y = 0.2$.
(f) The preparation costs of this battle for each participant: $E_x = 0.3, E_y = 0.7$, respectively. According to Kyriazis and Zouboulakis [5], 100 new Athenian triremes were built under the Athenian Naval Law of Themistocles. Each one cost one talent (6000 ancient drachmae), so the total cost was 100 talents (or 600.000 ancient drachmae). In 480 BC, the Athenian fleet was comprised of 200 triremes, equivalent to the two thirds of the total Greek strength. However, the Persian ships were similar in shape, so we assume that the cost of each ship was similar. Thus, it is obvious that the Persians spent more money to support their expedition to the Greek territories than the Greeks.

With these initial conditions, we solve the system (Eq.1), by using the mathematical software Maxima[3] (5.39.0), calculating the equilibrium points. Then, we study more extensively the behavior of the model and we present bifurcation diagrams and timeseries diagrams using the software E&F Chaos[4].
Solving the system (Eq.1), there are two equilibrium points: E_1 (x^* = 0.75, y^* = 0.475) and E_2 (x^{**} = 0.96, y^{**} = 1.012). According to Game Theory, these two fixed points are considered as Nash Equilibriums [9]. Below, the stability of the fixed points will be examined.
The Jacobian matrix is:

$$J = \begin{pmatrix} 0 & 0.8y - 0.8(1-y) \\ 3.6x - 3.6(1-x) & 0 \end{pmatrix}$$

[3] https://sourceforge.net/projects/maxima/files/Maxima-Windows/5.39.0-Windows/
[4] E & F Chaos: written by Diks, C., Hommes, C., Panchenko, V., van der Weide, R., (2008).

We calculate the Jacobian matrix at the equilibrium point E_1:

$$J^* = \begin{pmatrix} 0 & 0.038 \\ 1.803 & 0 \end{pmatrix}$$

The determinant of J^* is det (J^*) = 0.069 > 0.
The trace of J^* is trace(J^*) = 0.
The eigenvalues of J^* is (0.264i, -0.264 i); two complex roots.
The discriminant $\Delta = trace(J^*)^2 - 4 \cdot det(J^*) = -0.2788 < 0$.
Therefore, the equilibrium point E_1 is a <u>stable – center</u>.
Studying the second fixed point E_2, the Jacobian matrix at the equilibrium point is:

$$J^{**} = \begin{pmatrix} 0 & 0.82 \\ 3.314 & 0 \end{pmatrix}$$

The determinant of J^{**} is det (J^{**}) = $-2.718 < 0$.
The trace of J^{**} is trace(J^{**}) = 0.
The eigenvalues of J^{**} is (1.648, -1.6487); two real roots.
The discriminant is $\Delta = trace(J^{**})^2 - 4 \cdot det(J^{**})$ 10.874 > 0.
Therefore, the equilibrium point E_2 is a <u>saddle point</u>.
Consequently, we accept the fixed point E_1 (x^* = 0.75, y^* = 0.475) and reject E_2 (x^{**} = 0.96, y^{**} = 1.012), because the value of y^{**} is greater than 1.
Thus, we continue the analysis for the fixed point E_1. Interpreting this equilibrium point, we confirm the aggressive (strategic) behavior of Greeks; since the value of x^* is close to 1 and the mild (strategic) behavior of Persians; since they thought it would be an "*easy win*".
Indeed (historically), the courage of the Greeks, their technological naval skills, and the advantageous geographical location contributed in this aggressive behavior. As far as the Persians are concerned, their mild (strategic) behavior is due to the fact that they underestimated their enemy, since they regarded that the Greeks are an easy target, and they would achieve a decisive victory.
Connecting the game "Hawk – Dove" to the naval battle of Salamis, player 1 (red) is "Persians" and player 2 (blue) is "Greeks" (Fig. 4). The Hellenic alliance had an aggressive behavior (Hawk) and the Persians behaved as a Dove. According the Nash equilibriums that have been mentioned above (See 2), the Greeks (player 2) should behave as a Hawk (i.e. aggressive), regardless of Persian's behavior, so as to win this battle.

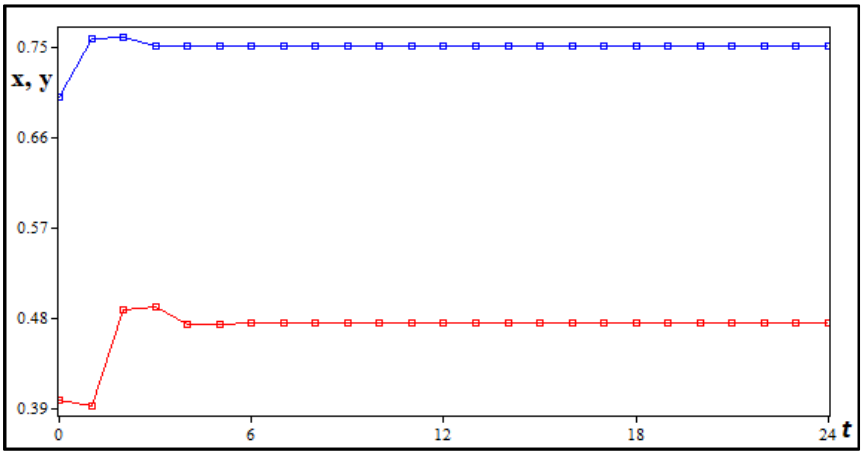


Fig. 4: Time series diagram - x (Greeks; blue) and y (Persians; red).

Fig. 4 shows us how the two warring parties behave (strategically). In particular, it represents the optimal strategic behavior of Greeks and Persians in Salamis straits for a time interval of 24 hours. We can observe an oscillation, at the beginning, until t = 6 h. (both lines) and then it is normalized and balanced. That means that the duration of the main battle was approximately 6 hours. Indeed, according to historical documents, the battle started at dawn (approximately at 06:00 am) and the Greek victory was visible at noon.

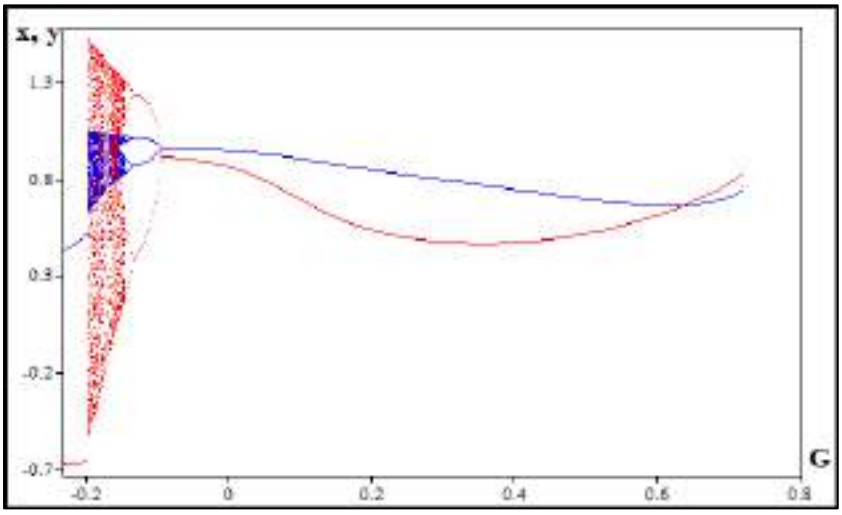


Fig. 5: Bifurcation Diagram for different values of *G*.
x (Greeks; blue) and *y* (Persians; red).

Fig. 5 presents the strategic behavior of Greeks (blue) and Persians (red) as the parameter G changes. We observe for the positive values of G, the blue line is

above the red until $G = 0.64$ (critical value) and for $G > 0.64$ the red line is above the blue. The increase of the value of parameter signifies the change of the geographical location (a more open sea), which becomes more difficult for Greeks and in contrary easier for Persian. Thus, we approve that if the location of the naval battle was in an open sea, the Persians would have a crucial advantage, which would possibly lead to win this conflict.

Although we did not study the negative values of parameter G, we believe that there are some unpredictable geophysical factors (e.g. meteorological phenomena to influence the outcome of the conflict), which are surprisingly interesting. Specifically, we refer to weather conditions, such as air, ripple, etc., which can affect the geophysical terrain of the area. Due to these weather phenomena, period doubling bifurcations and chaos appear and we cannot predict what could happen in the battle for these values of G.

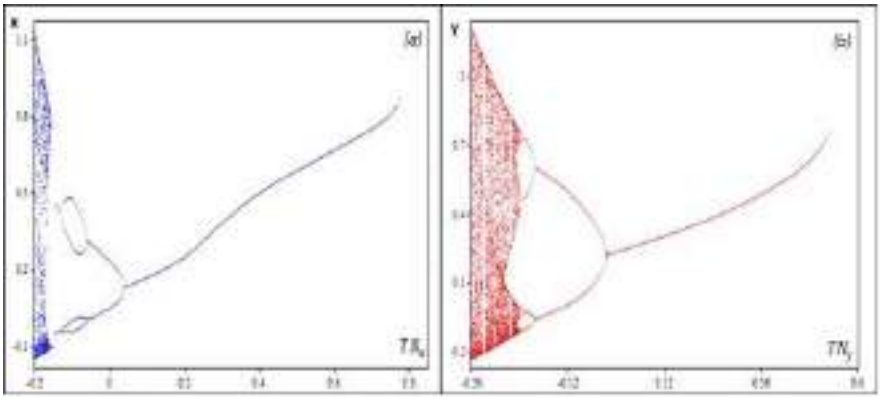


Fig. 6: Bifurcation diagrams for different values of parameters TN_x and TN_y.

Fig. 6a: x (Greeks; blue) and Fig. 6b: y (Persians; red).

Fig. 6 depicts the technological evolution and capability of x (Greeks; blue) and y (Persians; red), respectively. In the left diagram (Fig. 6a), for the negative values of parameter TN_x, we can distinguish a pair of bubble bifurcations, while afterwards we have the well-known period-doubling scenario to chaos. A possible interpretation of this chaotic scenario is the uncertainty of Greeks in technological capability – first attempts to construct ships. The first ships, as Krasanakis [4] mentions, were floating planks and carved tree trucks only with oars. Since the ships were primitive, the situation was unstable (there is chaos in this range of values) because they were not capable to fight in naval battles. Later, the sails were invented, which gave high speed to ships, and they were consisted no more than wood but iron. For this reason, we have bubble bifurcations, which indicate the technological alternatives that existed for the construction of the ships. In the interval of positive values of TN_x, there is stability with two fixed points. Here, it's the beginning of better shipbuilding ability and new expertise ship construction. Finally, there is one equilibrium point which shows the better version of ships, of that period, namely Triremes. Triremes were wooden

warships which move either with sails or oars. Moreover, in the positive values of the parameter TN_x, the increasing of the slope of the curve is visible, which, on the one hand, it means that in 480 BC the triremes were an innovation in shipbuilding and on the other hand, it shows the excellent naval capability of the Greeks.
Persians, through the years, developed technological equipment because of their expansive mania to conquer Greece. Comparing the Figures 6*a* and 6*b*, it seems that Persians had a lower technological development than Greeks, since they focused more on land army than on warships. Their ships were mainly used as troopships rather than battleships [11].

Conclusions

In this article, a new non – linear discrete model has been presented, which simulates the optimum strategic behavior of two warring parties for short–term battles. In addition, we try to compare this model with the game "Hawk – Dove", applying this attempt in the naval battle of Salamis. Based on the results we have extracted, we (mathematically) proved the historical events of this conflict. Specifically, the Greek's strategic behavior fits with the aggressive behavior of the Hawk and as well as the Persian's strategic behavior fits with the more defensive behavior of the Dove in the game. Moreover, the estimated duration of the battle was proved as well as the dominance of the Greek fleet in the Salamis straits. Finally, the technological naval capability of the Greek alliance was able to cause serious damages to the opponent and led them to a crashing defeat.

References

1. Burn, A. R. Persia and the Greeks, New York: Minerva Press, 1962.
2. Goodwin, W. The battle of Salamis. Harvard Studies in Classical Philology, Vol. 17, pp. 74-101, 1906.
3. Green. P. The Greco – Persian Wars, University of California Press, 1998.
4. Krasanakis, A. Naval History of the Greek Nation. Athens Press, 2008.
ISBN: 9789604841592.
5. Kyriazis, N., Zouboulakis, M. Democracy, Sea Power and Institutional Change: An Economic Analysis of the Athenian Naval Law. European Journal of Law and Economics, 17: 117–132, Kluwer Academic Publishers, 2004.
6. Lorenz, E. "The problem of deducing the climate from the governing equations", Tellus 16, 1964.
7. Lykke, A. F. Jr. Defining Military Strategy, *Military Review* 69, no. 5, 1989.
8. Malthus, T. An Essay on the Principle of Population. Publisher: J. Johnson, London,1798.
9. Osborne, M. J. and Rubinstein, A. A Course in Game Theory. Cambridge, MA: MIT, 1994. ISBN: 9780262150415.
10. Maynard Smith, J., and Price, G. R. The logic of animal conflict. Nature 246(5427): 15–18,1973.
11. Shepherd, W. Salamis 480 BC: The naval campaign that saved Greece. Osprey Publishing, 2010. ISBN: 9781846036842.

Software
Diks, C., Hommes, C., Panchenko, V., van der Weide, R., (2008), "E&F Chaos: A user-friendly software package for nonlinear economic dynamics", Computational Economics, vol. 32, pp. 221- 244.
McKelvey, Richard D., McLennan, Andrew M., and Turocy, Theodore L. (2014). Gambit: Software Tools for Game Theory, Version 16.0.1. http://www.gambit-project.org.
https://sourceforge.net/projects/maxima/files/Maxima-Windows/5.39.0-Windows/

Life evolves in experimentally confirmed ‘half-chaos’ of not fully random networks, but not ‘on the edge of chaos’

Andrzej Gecow[1]

[1] No affiliation, Warsaw Poland, (E-mail: andrzejgecow@gmail.com , Home-page: https://sites.google.com/site/andrzejgecow/home)

Abstract. An experiment cannot refute mathematical theory, but decides whether a given theory is suitable for describing the indicated phenomena. I do not question the theory of chaos in the field of infinite and continuous spaces, but I present a hitherto unknown class (half-chaotic systems) of finite discrete networks in which this theory gives erroneous results and I explain why. It is easy to repeat the computer simulation experiments described here, also much wider descriptions are available. The widely known Kauffman hypothesis 'life on the edge of chaos' is based on the recognition made in random autonomous Kauffman networks. This recognition, supported by mathematical theory of chaos, gave an image in which systems can be either ordered or chaotic with a fairly fast phase transition between them. Only for parameters in the immediate vicinity of this phase transition, the changes have properties suitable for describing stability of adaptive evolution and typically modeled objects. This limitations for the variables are strong. However, modeled adapted systems are not fully random, they are usually stable, but the estimated parameters are usually “chaotic”-they place the fully random networks in the chaotic regime, far from the narrow phase transition. The half-chaotic network has such “chaotic parameters”, it simultaneously exhibits in similar share both small (ordered) and large (chaotic) reactions for small disturbances. The discovery of half-chaos frees modeling of adapted systems from sharp restrictions; it allows to use “chaotic parameters” and get a nearly stable system more similar to modeled one. It gives a base for identity criterion of an evolving object, simplifies the definition of basic Darwinian mechanism and changes “life on the edge of chaos” to “life evolves in the half-chaos of not fully random systems”.

Keywords: Kauffman networks, complex networks, chaos, life on the edge of chaos, phase transition to chaos, damage spreading.

1 Introduction

This is empirical work using computer simulation. It concerns dynamics in complex autonomous Kauffman networks that are finite and discrete, shows that current theory of deterministic chaos used for them, based on Lyapunov coefficient in infinite, continuous space, implies false expectations. Such a method is an approximation. It loses a few important phenomena present in such the networks, but absent in the infinite and continuous space. Current view is based on the assumption that networks are fully random, however, interesting phenomena concerning life occur in not fully random networks due to natural selection. Also limitation to Boolean networks in statistical investigation makes incorrect picture. Due to such the reasons, expectations of the theory that life is on the edge of chaos can be and are inadequate.

The description of the investigation and the arguments for introducing the half-chaos given in this article is necessarily shortened and simplified. A much more extensive investigation and description is available in supplement **(**Gecow [9]**)** to this article. Wider versions of the article are available in preprints **(**Gecow [8]**)**. The data (programs and its sources, results of simulations) analyzed here and additional explanations are available from the author on any request.

2 Kauffman network and chaos in it

The considerations concern the statistical stability of the deterministic Kauffman networks (Kauffman [11, 12, 13]). A node in such a network receives signals at the K inputs, converts them uniquely to the output signal called the state of the node, then sends it to other nodes by k outputs. Up to now, 2 (logical) signal states have been used. In the simplest case it was assumed the same probability and full randomness of connections and functions. K (called "connectivity", see **(**Turnbull *et al.* [24]**)**) was the basic variable for Kauffman.
The conflict (Aldana *et al.* [1, 2], Turnbull *et al.* [24]) of a size of K in the Kauffman model and K estimated from nature is a problem solved here. Kauffman postulates that the natural property of the random ordered systems ('order for free' Kauffman [14]) is the source of stability, but then K should be extremely small ($K \leq 2$) (Derrida and Pomeau [4]). The attempts to prove that the real genetic network is ordered (Serra *et al.* [19, 20], Shmulevich *et al.* [22]) assume such a source of stability. Different circumstances allowing for greater K in the ordered phase were indicated (p.48 in Aldana *et al.* [1]), such as a significant difference in probabilities of logical states (Derrida and Pomeau [4]), or deviation from the randomness of the function (canalizing Kauffman *et al.* [15]), but these and other suggestions are not satisfactory for many reasons (Gecow [7]).
Synchronous computing is used, i.e. the states of nodes from the discrete time t are input signals and arguments of function of other nodes, and the results of these functions are node states at the next moment (t+1). Considerations have been limited to autonomous systems – they do not take the signal from the environment. Determining the states and functions of all nodes and the connections between nodes uniquely determines the trajectory - consecutive states of the whole network (sets of states of all nodes).
Same K for all N nodes of the network are taken. The size of a change in a network function at time t after a small disturbance is measured by the number A (from Avalanche Serra *et al.* [20]) of the nodes, which have a different state from the pattern network - identical but without disturbance. The value d=A/N is called damage, its distribution characterizes stability and is the most important result (Fig.1a,b).
The main characteristic of the chaotic behavior of dynamical systems is high sensitivity to initial conditions, leading to maximally different effects for very similar initial conditions. I use the term 'chaos' in such the meaning, similarly as Kauffman [13] does. For chaotic Kauffman networks a small initiation of damage typically causes a large avalanche of damage which spreads onto a big

part of the discrete and finite system and ends at a Derrida equilibrium level (Derrida and Pomeau [4], Gecow [7], Derrida and Weisbuch [5]), which is a maximal lose of information about previous system. The existence of this limitation is the main difference between this 'chaos' and the more commonly taken definition (Schuster [21]) used for continuous variables on infinite space. The term 'chaos' is not reserved for one of those separate areas. The distribution of damage size is the experimental base to classify particular system of Kauffman network as chaotic or ordered using levels of damage equilibrium calculated from Derrida's annealed approximation.

According to my previous (Gecow [7]) suggestions, here I also study a larger number of s (>2, usually 4) of equally probable signal states, which in random networks for every sensible K (≥2) always gives chaos. Attempts to introduce more signal states already exist (Luque, F.J. Ballesteros [17], Sole *et al.* [23]), but they assume the possibility of an ordered phase for the random network, that these states cannot be equally probable.

Several types of networks are considered. They differ in the rules of their creation and k distributions: sf (scale-free Barabási,*et al.* [3]), er (classical Erdős and Rényi [6] "random"), and ss (single-scale). In the figures, the second letter of these shortcuts indicates the network type. It is also looked after the node functions are correctly random, but this assumption cannot always be fully met, so the impact of the derogations is checked. The vector type,s,K is the basic variable here. Further, parameters s,K, which in the case of a random system give chaos, we will refer to in short as "chaotic".

3 Main results in brief, half-chaos

The research has shown that among the systems with chaotic parameters, the strongly increased stability desirable in modeling many interesting objects is in not fully random systems with short attractors, here called half-chaotic. Functioning of such the system after a small disturbance is quite different (chaotic) or very similar (ordered), comparing to a undisturbed system. Medium damage is practically nonexistent. Both of these options occur in the same system similarly often, which is surprising, since it was previously thought that the system may be either chaotic or ordered. As addition to the short attractor, this ratio is influenced (Gecow [8]) among other by negative feedback, modularity, and 'in-ice-modularity' detected in these studies.

Kauffman is trying to describe living systems and similar using several easy to show and the main parameters of his model, he simplifies the rest of them assuming their randomness, but natural selection works on all possible parameters destroying their randomness. Indeed, it is difficult to imagine the possibility of the existence of half-chaotic systems. In fact, after the system is drawn, it is either chaotic or ordered, and the set of random systems contains all the possible ones. Only near the phase transition the changes of function can statistically be small, and such are necessary for the evolution of modeled objects, as Kauffman stated in the well-known hypothesis that "life is on the edge of chaos and order." In the interpretation of the results of this approach, it has not been seen that the statistical absence of intermediate systems does not

imply a small number of such systems. There are a lot of half-chaotic systems, but their share is negligible, because chaotic systems with given parameters (e.g., K) are radically more - for larger N not imaginable many. In light of the results, it must be stated that "life evolve in the half-chaos of not fully random systems".

4 A point attractor system is half-chaotic

Demonstration of the existence of half-chaotic systems can be made simply by indicating such systems, e.g. in the form of an algorithm for their construction. A good and simple example is a point-attractor system. You can get this system from the random system slightly change the function of nodes: for current input signals, insert as the function value the current state of this node. This is a very slight deviation from randomness. By investigating the various static statistic characteristics of such a network, it is difficult to detect that this is a unique network. This uniqueness is not only because a point attractor, but it is a half-chaotic network.

The distribution of the damage size in the half-chaotic network has two peaks - the left of very small changes of network function and the right - of the large changes in the function. Between them is a big gap - intermediate changes practically do not occur (Fig.1a). This allows us to naturally determine the "level of order" q as the content of the left peak. This q is the most important characteristic of the damage distribution (Fig.1c, 2a,f), which clearly distinguishes the half-chaotic systems (J in Fig.2f, Fig.1c without X, Fig.3a,b) from the chaotic ones (X in Fig.1c, 2f, 3c). The right peak was predicted by Derrida (Derrida and Pomeau [4], Derrida and Weisbuch [5], Gecow [7, 8]) in the annealed approximation model - it is a chaotic reaction, the maximal loss of information about earlier function. It can be said that the system ceases to be itself from before of the disturbation, completely loses its adaptation (if it was adapted), or more generally - purposeful action.

Disturbation initiating of network function change can be very different. In my main study I used permanent change of node function in one node only for one (initial) state of inputs. For s = 4 we can change the value of the function to 3 other values. In another investigation I use addition and removal of node.

The natural question arises: Will leaving such disturbing changes, modeling the evolution of the object, results in the loss of the half-chaos? It turns out that leaving only the changes that triggered a minor change in the range of left peak allows for a long evolution without losing half-chaos. This feature, termed the ‘evolutionary stability of the half-chaos’, is included into definition of the half-chaos. The system is changing slowly - it remains "self", although it evolves and may lose the point attractor. Changes in the network functioning are small but not negligible. However, the attractor typically remain small.

Accepting only minor changes in the naturally separated left peak is the real basis of the ‘natural identity criterion’. Acceptance of one initiating change giving a chaotic change of function (from the right peak) causes (experiments X in Fig.2f, 1c) that the feature of half-chaos disappears, and the system becomes

typically chaotic. It is a good model of death, which is necessary for elimination in the Darwinian basic mechanism. Simultaneously, elimination is the opposition of identity continuity defined by the left peak. It simplifies and clarifies the definition of the Darwinian mechanism and allowing it to be used to define life.

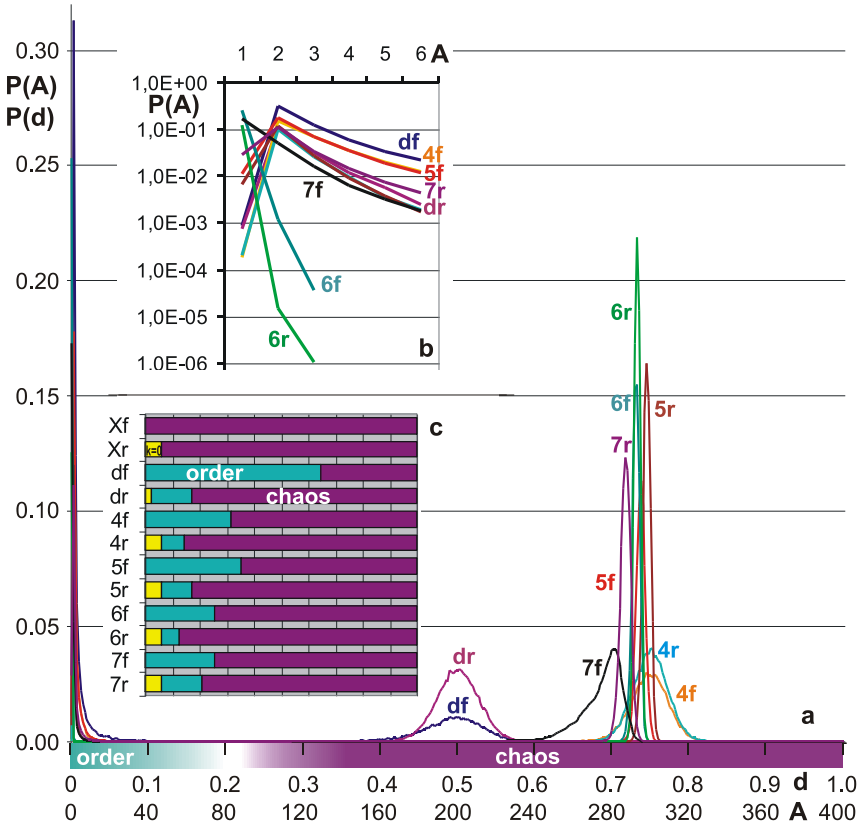


Fig.1. The main result. a, b – distribution of damage size, a – full, b – left peak . A – Avalanche, number of different node states at maximal t (tmx) in disturbed and undisturbed networks. A is averaged over the last 50 timesteps. In the networks there are N = 400 nodes, and damage, d = A/N. The first character of the curve description ('d', '4', '5', '6', '7') indicates the experiment, while the second shows the network type ('f' – scale-free, 'r' – Erdős-Rényi 'random'). Only the networks in experiment 'd' are Boolean (number of equally probable signal variants, s = 2) but connectivity (number of input links), K = 4. In all remaining experiments, s = 4 and K = 3. Fully random networks with such values of parameters s and K are chaotic (only the right peak exists), but here, the left peak exists and its share is not negligible (see q in **c**). Such a picture, with two peaks, is for each particular network. They are neither ordered nor chaotic; thus, I call them 'half-chaotic'. The results from a few hundred networks for each experiment are summarized here. Errors are not calculated due to the presence of many types of rare

causes, which make such calculation inadequate—the smoothness of the curves is enough.

c – Half-chaos – fractions of ordered events (q) and chaotic. In the range of q an order resulting from the absence of output in some nodes (k = 0) in the network er is isolated as yellow. For ‘d’ and X (from ‘7’) there are no evolution, the results concern the network immediately after generation of half-chaos, but for X also after acceptance of one chaotic change, which gives a typical chaos (see also Fig.2f). In the remaining methods (‘5’, ‘6’ and ‘7’) result is a sum of the results of 4 stable set M, as in **a,b**, (see Fig.2). s,K = 4,3 except ‘d’ where s,K = 2,4.

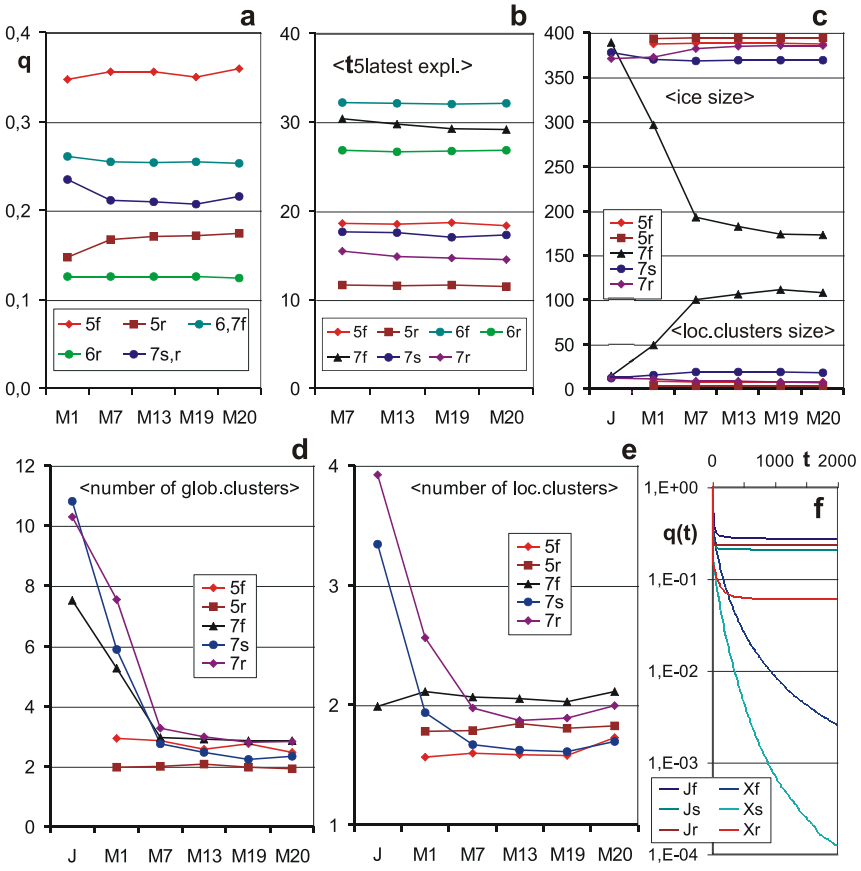


Fig.2. The stabilization of basic parameters during evolution is the main argument for evolutionary stability of half-chaos.
The similarity of half-chaos based on in-ice-modularity, despite the differences in the way (‘5’,’7’) of it obtaining. For ‘6’ it is no in-ice-modularity, then lack ‘6’ in (c,d,e). Sets J, M1, M7, M13, M19 and M20 of initiation are full - no blocking of reverse changes (‘5-7’, N=400). Initial set J and M1, see (c,d,e) are not yet stabilized, therefore they are not summarized in Fig.1.

a - Stability of parameter q (degree of order of the system, the contents of the left peak in Fig.1) shows lack of moving towards the chaos (smaller q) during the evolution.
b - The average time of five latest explosion to the chaos does not grow. In the chaotic networks such explosions (Fig.3c) happen almost until the not yet exploded processes exist.
c - The average size of the ice and of the local clusters. In '6' there is no ice. 7f has a specific derogation, but it also stabilizes.
d - The average number of global clusters.
e - The average number local clusters.
f - Average q(t) in sets of J and X, starting after construction of in-ice-modular system ('7'). In this experiment (part of '7') N=800, tmx=2000, no evolution and M sets. J gives a typical picture for half-chaotic systems - q quickly stabilizes and is high. X starts after accumulation one disturbing permanent change, which has given large damage avalanche. Here q(t) drops up to tmx and probably further, but not for er due to k=0.

5 Short attractor and in-ice-modularity

A more detailed analysis of the mechanisms for maintaining the increased stability of systems with chaotic parameters has shown that the basic condition of the statistically significant presence of the left peak in the damage distribution is a sufficiently short attractor. It allows to limit the number of different circumstances in which secondary initiations (disturbations) appear in effect of encountering an input state by a node for which the value of the function was changed. For one expiration of the damage is relatively probable, but the fade out of each of the many independent disturbations is practically impossible.
It was checked by the method '6' that such the condition is sufficient to obtain a half-chaos, but the shape of the left peak (Fig.1b) was significantly different from obtained starting from point attractor (method '5'). Practically it contains only lack of functional change (A = 0) and therefore such system is not suitable for modeling the evolution of objects like administrative units, technological processes, technical constructions and living organisms.
The differences in the half-chaos mechanisms obtained starting from the point attractor ('5') and from the small attractor ('6') were investigated more in details. It turned out that the mechanism in the first case is very similar to the indicated by Kauffman in the vicinity of the phase transition – "small lakes of activity in the ice" (originally: "unfrozen islands") (Kauffman [12]). Kauffman calls ice nodes that do not change their state. Starting from the small attractors ('6') practically does not create ice. Point attractor system is a completely frozen, disturbation unfreezes small subset of nodes (Fig.2c), and inside it attractor is usually small. After accumulating such the change, the next disturbation unfreezes another independent piece (clusters in the figures), and an image similar to modularity is created. I called it in-ice-modularity. In-ice-modules are also the classic modules, but this is only one, supporting, but less important factor. The main property of the in-ice-modules is the activity - changes of the states of forming them nodes. The ice surrounds them and isolates from the other in-ice-modules. In-ice-modules are the result of the

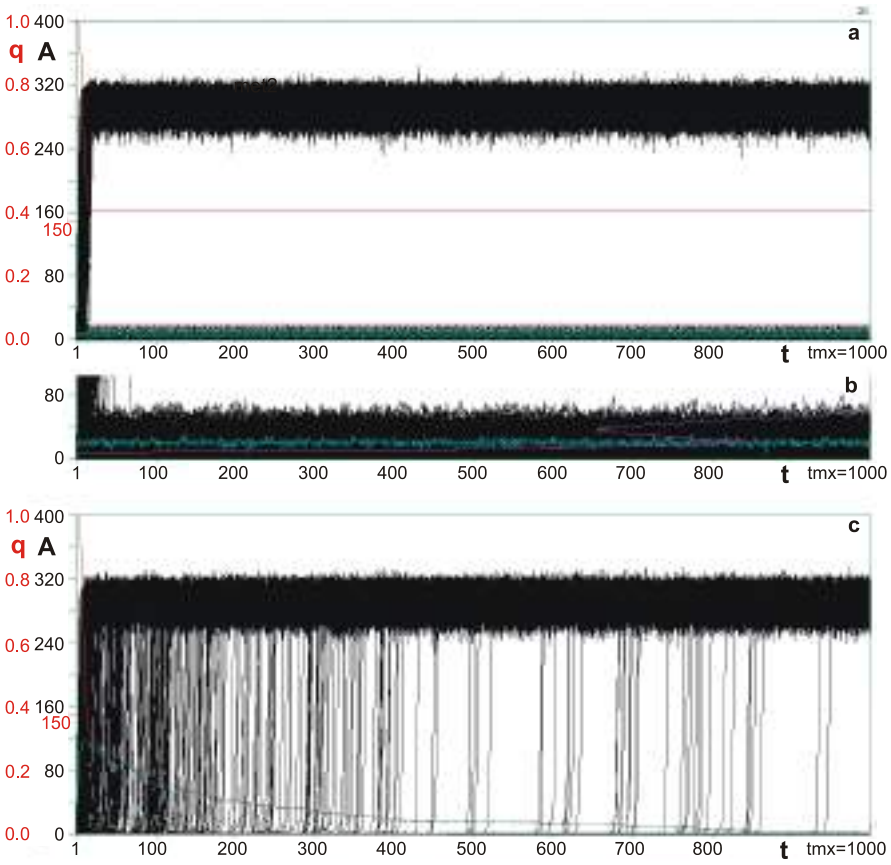


Fig.3. Half-chaos and chaos in the presentation of A(t) from simulation of a full set of small permanent disturbations.
A - number of the nodes states different than in the not disturbed system. t is a number of steps of network calculation from damage initiation. The details should be watched in enough magnification - they were presented on the screen pixels. Showed examples are typical for all experiments or network types. Here s =4, it indicates level of Derrida equilibrium near A=290 (see Fig.1a). Typical N = 400, tmx = 1000 is used, so the rectangle has a dimension of 400*1000 pixels. After each of 1200 initiation A(t) is drawn with a continuous black line. As can be seen, the transition to chaos in the vicinity Derrida balance is rapid in several to over a dozen steps, where A increases drastically, so I call it "explosion." Threshold is here chosen at A=150 which is marked in red. It is used for q calculation. After the end of initiation set, the red curve q(t) - the share of processes that in the time t did not cross the threshold, was added to the figure. q=1 for A=N. Red description of the left has been added for readability. Blue line describe the share of processes that currently have A = 0, i.e. damage fade out.
a – Half-chaos. The red curve q(t) quickly stabilizes at a high level near q=0.4. (See Fig.2f.) In the lower part of the graph many trajectories are drawn (there are 476 of 1200). The last explosion occurs at t=18. These, and lot of other data are presented on the

screen and saved, they are described in details in (Gecow [8, 9]). Blue line is much lower than red. It means, that most of stable processes are not a case of damage fade out (A=0).
b – A premise of in-ice-modules. The lower part of the similar to (a) image, but a case with larger fluctuation. The q level is here high, near q = 0.55. A distribution of damage size lower than threshold 150 was studied on section from t = 600 to tmx (purple curve on the right frame). It is one of ways to see the in-ice-modules existence. The seen significant peak is probably an effect of stimulating of one hypothetical large in-ice-module.
c – Chaos. Typical form for sf network type, for er and ss explosions typically end earlier. Here one process has not yet exploded at tmx. Experiments X (see Fig.2f) have also such the picture. Here, q(t) is steadily decreased until not 'exploded' processes exist. For the chaotic system processes which fade out (A=0, blue line) are the main part of not yet exploded which build q, the secondary initiations lead typically to their explosion.

functioning defined by the functions and states of nodes in a given structure. This state really gives the half-chaos based on small attractors, but they are local attractors in in-ice-modules, and their assembly can give quite a large global attractor for the whole system. Local clusters (present in one process after one permanent change) are defined as sets of nodes with the same period of state changes. There are usually several local clusters (Fig.2e) at the same time, and ice occupies most of the network (Fig.2c). In the method '6' there is usually one cluster covering almost the entire network, there is lack of in-ice-modules but the classic modules are present like in method '5'. In one process of evolution very close clusters are frequent, they freeze for many initial changes and reappear, often changing their period. A set of such similar clusters in the process of evolution of a given network is called global cluster (Fig.2d), but its definition is fuzzy and abounds in many different phenomena difficult to resolve. Among them are assembling and disintegration of clusters.
An algorithm is developed to create a in-ice-modular system based on such a description, in the networks with random structure, without starting from the point attractor. It effectively gives half-chaos ('7') with similar properties observed when starting from the point attractor ('5'). It is unnecessary to force the small attractors in the in-ice-modules.

6 The supports for stability

It is generally believed that the stability of the various systems results from homeostasis based on regulation by negative feedback. Kauffman [14] pointed instead to the property of ordered phase (order for free) as the most important reason, but for it extremely small K should be expected. Searching for mechanisms for increasing the stability of systems with chaotic parameters I began with an attempt to show that increasing the share of negative feedback is sufficient (Gecow [8]). However, this claim turn out to be wrong, a short attractor is necessary, but negative feedback indeed (Gecow [8, 9]) causes very strong support for stability. Narrowing of functions and always present modularity also are insufficient, but similarly supportive. However, research on modularity is rather superficial and the subject needs to be deepened.

7 The basic details of simulations

Simulation studies and their analysis include many important details that are unfeasible to include in this article. They are described in more than 200 pages of reports (Gecow [9]), only selected of them in article (Gecow [8]). Only basic ones are listed here.
The typical size of the studied network is N=400 nodes, but experiments with N=800 and even 4000 also exist. The basic parameters are: s - the number of equally probable states of the signal, uniquely s=2 (the Boolean network denoted by ‘d’ in Figs.1) or usually s=4; and K - number of inputs to the node, constant for the whole network, K=4 for s=2 and K=3 for s=4. A(t) - the number of node states different from the undisturbed pattern at a given moment t is the principal measured value. This value is shown in Fig.3 and is also used (Figs.2f, 3) to determine q(t). The particular system is calculated at tmx discrete time steps t after disturbation, and then at t=tmx, more adequate value for the final results (Fig.1) is recorded as averaged A over the last 50 counting steps t. Typically, tmx is 1000, but longer stretches are also used, even up to 20000. The tmx value was arbitrarily determined on the basis of preliminary simulations, but it is checked whether the increase would not change the results.
Network types sf, ss and er are used in experiments described here. The second letter indicates the network type in the figures. The number of networks with parameters type,s,K in a particular final simulation is usually about 600, it happens 100, but often experiments were repeated in a similar way, giving a much greater certainty. Due to the strong influence of various factors, often sporadic, formal errors in the obtained results are not calculated, judging such a calculation as clearly inadequate and misleading. This problem is limited to the similarity of results from the similar simulations and the visual evaluation of fluctuations. The evolution of the experiments ‘5’, ‘6’, ‘7’ has many additional constraints to increase credibility. They are similar in these three cases for the basic results, but moreover, a series of simulations of other rules are performed. After initial set of initiation (J) there are 20 sets M, in most of them the reversal of the accepted changes are blocked. This results in the exclusion of a large number of initiations from the measurements and leads to a significant slowdown of evolution. After several such set, the reversal is allowed (M1, M7, M13, M19, M20), assuming that the change has already another circumstances. It also allows to correct measure of various phenomena that illustrate evolution (Fig.2). Since the attractor is decreasing spontaneously, making it difficult to move away from the point attractor, it is also forbidden to reduce the global attractor to less than 7, and in the M20 (at ‘5’, ’6’, ’7’) to reduce the attractor. Also, the point of initiation of damage is shifted by t=50 after each accumulated change, because the magnitude of this shift turn out to be a significant factor.

8 Conclusion

Various complex objects adapted by natural selection or humans were attempted to model by dynamic complex networks. For that, Kauffman's Boolean networks are perfect. It was estimated that mean connectivity K should be clearly larger than 2. However, such networks, according to current views, should be chaotic, and this did not match the features of modeled objects. The discovery of the half-chaos in which K and s can take larger values than the extreme 2, and yet in the model remains the order, removed from this modeling sharp constraints, which until now are the typical basis of many considerations (Nghe *et al.* [18], Serra *et al.* [19, 20], Kauffman [12, 14], Aldana *et al.* [1, 2], Luque and Ballesteros [17], Sole *et al.* [23], Kauffman *et al.* [15], Shmulevich *et al.* [22], Iguchi *et al.* [10], Villani *et al,* [25], Kinoshita, [16]). This opens the door to adequate models with complex networks.

Kauffman's well-known hypothesis (Kauffman [12, 13]) "lives on the edge of chaos and order" as a result of deepening the model and finding the half-chaos is here modified to "life evolves in the half-chaos of not fully random systems". This change entails deeper interpretive conclusions, indicates the natural basis of the identity of the evolving object and the model of its death necessary for Darwinian elimination.

Of course, such a brief summary is necessarily a great simplification. Basing the conclusions on the finite network simulations in complex nonlinear phenomena is well-justified, but is in conflict with the habit of defining chaos through Lyapunov exponent, whose analog for networks (coefficient of damage propagation (Gecow [7]) and eq.2.3 in (Aldana *et al.* [1])) in the case of half-chaos turns out to be misleading. The presented results can be treated as a guide to developing more advanced methods, but in the case of finite, dynamic complex networks, it will not necessarily be a more comfortable and more relevant description.

References

1. M. Aldana, S. Coppersmith, L. P. Kadanoff, “Boolean Dynamics with Random Couplings” in Perspectives and Problems in Nonlinear Science, Applied Mathematical Sciences Series, ed. E. Kaplan, J. E. Marsden, K. R. Sreenivasan, Springer-Verleg, Berlin 2003.
2. M. Aldana, Dynamics of Boolean Networks with Scale Free Topology. Physica D 185, 45-66, 2003.
3. A.-L. Barabási, R. Albert, H. Jeong, Mean-field theory for scale-free random networks. Physica A 272, 173–187, 1999.
4. B. Derrida, Y. Pomeau, Random Networks of Automata: A Simple Annealed Approximation. Europhys. Lett., 1(2), 45–49, 1986.
5. B. Derrida, G. Weisbuch, Evolution of Overlaps Between Configurations in Random Boolean Networks. Journal De Physique 47, 1297–1303, 1986.
6. P. Erdős, A. Rényi, Random graphs. Publication of the Mathematical Institute of the Hungarian Academy of Science, 5, 17–61, 1960.

7. A. Gecow, Emergence of Matured Chaos During Network Growth, Place for Adaptive Evolution and More of Equally Probable Signal Variants as an Alternative to Bias p. in: Chaotic Systems, E.Tlelo-Cuautle (ed.), ISBN: 978-953-307-564-8, 280-310 http://www.intechopen.com, 2011.
8. A. Gecow, Life is not on the edge of chaos but in a half-chaos of not fully random systems. Definition and simulations of the half-chaos in complex networks, http://arxiv.org/abs/1712.09609 v3, 2019.
9. A. Gecow, Report of simulation investigations, a base of statement that life evolves in the half-chaos http://vixra.org/abs/1603.0220 , 2016, with its second part: Report of simulation investigations, part II, a growth of half-chaotic autonomous networks http://vixra.org/abs/1711.0467, 2017.
10. K. Iguchi, S. Kinoshita, H. Yamada, Boolean dynamics of Kauffman models with a scale-free network. J Theor Biol 247; 138–151, 2007.
11. S. A. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets. J. Theor. Biol. 22, 437-467, 1969.
12. S. A. Kauffman, Requirements for Evolvability in Complex Systems - Orderly Dynamics and Frozen Components. Physica D 42, 135–152, 1990.
13. S. A. Kauffman, The Origins of Order: Self-Organization and Selection in Evolution, Oxford University Press New York 1993.
14. S.A. Kauffman, At Home in the Universe, Oxford University Press USA 1996.
15. S.A. Kauffman, C. Peterson, B. Samuelsson, C. Troein, Genetic networks with canalyzing Boolean rules are always stable. PNAS vol. 101 no.49, 17102-17107, 2004.
16 S. Kinoshita, H. Yamada Role of Self-Loop in Cell-Cycle Network of Budding Yeast. Open Journal of Biophysics, 9, 10-20, 2019 https://doi.org/10.4236/ojbiphy.2019.91002.
17. B. Luque, F.J. Ballesteros, Random walk networks. Physica A 342 207–213, 2004.
18. P. Nghe, W. Hordijk, S. A. Kauffman, S. I. Walker, F. J. Schmidt, H. Kemble, J. A. M. Yeates, N. Lehman, Prebiotic network evolution: Six key parameters. Molecular BioSystems, 11(12), 3206-3217. DOI: 10.1039/c5mb00593k , 2015.
19. R. Serra, M. Villani, A. Semeria, Genetic network models and statistical properties of gene expression data in knock-out experiments. J. Theor. Biol. 227, 149-157, 2004.
20. R. Serra, M. Villani, A. Graudenzi, S. A. Kauffman, Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data. J.Theor.Biol, 246 449-460, 2007.
21. H. Schuster, Deterministic Chaos: An Introduction. Physik-Verlag, 1984.
22. I. Shmulevich, S. A. Kauffman, M. Aldana, Eukaryotic cells are dynamically ordered or critical but not chaotic. PNAS 102 (38), 13439–13444, 2005.
23. R.V. Sole, B. Luque, S. Kauffman, "Phase transitions in random networks with multiple states", Technical Report 00-02-011, Santa Fe Institute 2000.
24. L. Turnbull, M.T. Hütt, A.A. Ioannides, S. Kininmonth, R. Poeppl, K. Tockner, L.J. Bracken, S. Keesstra, L. Liu, R. Masselink, A.J. Parsons, Connectivity and complex systems: learning from a multi-disciplinary perspective, Applied Network Science; 3:11, 2018. https://doi.org/10.1007/s41109-018-0067-2
25. M. Villani, L. La Rocca, S.A. Kauffman, R. Serra, Dynamical Criticality in Gene Regulatory Networks. Complexity Hindawi, 2018. https://doi.org/10.1155/2018/5980636

Albert Einstein and the doubling of the deflection of light

Jean-Marc Ginoux[1]

One of the three consequences of Einstein's theory of general relativity was the curvature of light passing near a massive body. In 1911, he published a first value of the angle of deflection of light, then a second value in 1915, equal twice the first. In the early 1920s, when he received the Nobel Prize in Physics, a violent controversy broke out over this result. It was then disclosed that the first value he had obtained in 1911 had been calculated more than a century before by a German astronomer named Johann von Soldner. The aim of this article is therefore to compare the methods used by Soldner and then by Einstein leading to this first value and to explain the importance of the doubling of this value in the framework of Einstein's theory of gravitation.

A. The Genesis of General Relativity and the Curvature of Light

Two years after the publication of his article on special relativity, Albert Einstein considered generalizing his theory. Thus, in 1907, he wrote an article entitled "Relativitätsprinzip und die aus demselben gezogenen Folgerungen" ("On the Principle of Relativity and the Conclusions Drawn from it"), at the request of Johannes Stark, editor of the *Jahrbuch der Radioaktivität*, in which he presented for the first time one of the consequences of his theory. In paragraph V. entitled "The Principle of Relativity and Gravitation", he wrote:

[1] Laboratoire des Sciences de l'Information et des Systèmes, LIS, CNRS, UMR 7020 – Archives Henri Poincaré, AHP, CNRS, UMR 7117.

> "From this it follows that those light rays that do not propagate along the ξ-axis are bent by the gravitational field[2]…"

At that time Einstein thought that the effect of gravitational field on rays of light was too weak to be detected as evidenced by his conclusion:

> "Unfortunately, the effect of the terrestrial gravitational field is so small according to our theory (because of the smallness of $\frac{\gamma x}{c^2}$) that there is no prospect of a comparison of the results of the theory with experience[3]."

During the autumn 1911, Einstein became full professor of theoretical physics at the German Charles-Ferdinand University in Prague. Einstein lived more than a year at number 7 Leniscka Street with his first wife Mileva Maric and their two children Hans Albert and Eduard. Here he wrote a second article, considered as the starting point of his theory of general relativity: "Einfluss der Schwerkraft auf die Ausbreitung des Lichtes" ("On the influence of Gravitation on the Propagation of Light") in which he gave a first value of the deflection of light passing near a massive body. In paragraph 4 entitled Curvature of light rays in the gravitational field, he wrote:

> "By equation (4) a light-ray passing by a heavenly body suffers a deflection to the side of the diminishing gravitational potential, that is, to the side directed toward the heavenly body, of the magnitude

[2] Albert Einstein, "Relativitätsprinzip und die aus demselben gezogenen Folgerungen" ("On the Principle of Relativity and the Conclusions Drawn from it") *Jahrbuch der Radioaktivität*, 4, 411–462, 1907.
[3] *Ibid*.

$$\alpha = \frac{1}{c^2} \int_{v=-\frac{\pi}{2}}^{v=+\frac{\pi}{2}} \frac{kM}{r^2} \cos(v)\, ds = \frac{2kM}{c^2 \Delta},$$

where *k* denotes the constant of gravitation, *M* the mass of the heavenly body, Δ the distance of the ray from the center of the body (and r and υ are as shown in Fig. 3). *A light-ray going past the Sun would accordingly undergo deflection by the amount of* $4.10^{-6} = 0.83$ *seconds of arc*. The angular distance of the star from the center of the Sun appears to be increased by this amount. As the fixed stars in the parts of the sky near the Sun are visible during total eclipses of the Sun, this consequence of the theory may be compared with experimental evidence[4]."

Then, he added:

"It would be urgently wished that astronomers take up the question here raised, even though the considerations presented above may seem insufficiently established or even bizarre. For, apart from any theory, there is the question whether it is possible with the equipment at present available to detect an influence of gravitational fields on the propagation of light[5]."

In 1913 Einstein was appointed, on the recommendation of Marie Curie and Henri Poincaré, to a chair of mathematical physics at the Swiss Federal Polytechnic in Zürich (later the Eidgenössische Technische Hochschule, ETH) where he had studied a few years earlier. The following year, he joined Berlin to become director of the Kaiser-Wilhelm Institute of Physics,

[4] Albert Einstein, "Einfluss der Schwerkraft auf die Ausbreitung des Lichtes" ("On the influence of Gravitation on the Propagation of Light"), *Annalen der Physik*, 4(35), 898–908, 1911.
[5] *Ibid*.

and university professor. On November 1915, he submitted four papers to the journal of the Prussian Academy of Sciences[6]. Then, he published in 1916 in *Annalen der Physik* a 40-page article entitled "Die Grundlage der algemeinen Relativitätstheorie" ("The Foundations of the General Theory of Relativity") considered as the final and complete version of his theory of general relativity. In the last section he presented the three observable physical consequences of his theory and in particular the curvature of the light rays and explained that this latter consequence could be verified by exact astronomical experiments:

> "From this it follows that the spectral lines of light reaching us from the surface of large stars must appear displaced towards the red end of the spectrum*.
>
> *According to E. Freundlich, spectroscopical observations on fixed stars of certain types indicate the existence of an effect of this kind, but a crucial test of this consequence has not yet been made[7]."

Then, he provided a second value of the deflection of light passing near a massive body:

> "We examine the curvature undergone by a ray of light passing by a masse M at the distance Δ. If we choose the system of co-ordinates in agreement with the

[6] Albert Einstein, "Grundgedanken der allgemeinen Relativitätstheorie und Anwendung dieser Theorie in der Astronomie" ("Fundamental Ideas of the General Theory of Relativity and the Application of this Theory in Astronomy"), *Preussische Akademie der Wissenschaften*, *Sitzungsberichte*, 315, 1915, "Zur allgemeinen Relativitätstheorie" ("On the General Theory of Relativity"), *Preussische Akademie der Wissenschaften*, *Sitzungsberichte*, 778-786 & 799-801, November 4, 1915, "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie" ("Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity"), *Preussische Akademie der Wissenschaften*, *Sitzungsberichte*, 831-830, November 18, 1915, "Feldgleichungen der Gravitation" ("The Field Equations of Gravitation"), *Preussische Akademie der Wissenschaften*, *Sitzungsberichte*, 844-847, November 25, 1915.

[7] Albert Einstein, "Die Grundlage der allgemeinen Relativitätstheorie" ("The Foundations of the General Theory of Relativity"), *Annalen der Physik*, 4(49), 769–822, 1916.

accompanying diagram, the total bending of the ray (calculated positively if concave towards the origin) is given in sufficient approximation by[8]

$$B = \frac{\kappa M}{2\pi\Delta}$$

According to this, a ray of light going past the sun undergoes a deflexion of 1.7''…"

It is very surprising to notice that this second value of the angle of deflection of light provided by Einstein is equal twice the first. It will be learned some years later that the first value provided by Einstein in 1911 was in fact identical to that published a century earlier by the German physicist and astronomer Johann Georg von Soldner[9]. According to Abraham Pais :

"An Argentinian eclipse expedition which had gone to Brazil in 1912 and which had the deflection of light on its experimental program was rained out. In the summer of 1914, a German expedition led by Erwin Freundlich and financed by Gustav Krupp, in a less familiar role of benefactor of humanity, headed for the Crimea to observe the eclipse of August 21. (Russian soldiers and peasants were told by their government not to fear evil omens: the forthcoming eclipse was a natural phenomenon.). When the war broke out, the party was warned in time to

[8] The value $\kappa = 8\pi K/c^2$ where K represents the Cavendish's constant.

[9] This fact has been reported since by many historians of science such as Thomas Glick, *The Comparative Reception of Relativity,* Boston Studies in the Philosophy and History of Science, Vol. 103, Springer Netherlands, Dordrecht, D. Reidel, 1987; Jean Eisenstaedt, « De l'influence de la gravitation sur la propagation de la lumière en théorie newtonienne. L'archéologie des trous noirs », *Archive for History of Exact Sciences*, Vol. 42 (4), (September 1991), p. 315-386. Jürgen Renn, The Genesis of Relativity, Vol. 1-4, Boston Studies in the Philosophy and History of Science, Vol. 250, Springer Netherlands, Dordrecht, D. Reidel, 2007; Milena Wazeck, *On Einstein's opponents: the public controversy about the theory of relativity in the 1920s*, Cambridge; New York: Cambridge University Press, 2014. Nevertheless, neither of these references contains any mathematical analysis nor comparison of Soldner's and Einstein's results.

return and some did so. Those who hesitated were arrested, eventually returned home safely but of course without results. Frustration continued also after November 18, 1915, the day on which Einstein announced the right bending of 1”74. […]

An opportunity to observe an eclipse in Venezuela in 1916 had to be passed up because of the war. Early attempts to seek deflection in photographs taken during past eclipses led nowhere. An American effort to measure the effect during the eclipse of June 1918 never gave conclusive results. It was not until May 1919 that two British expeditions obtained the first useful photographs and not until November 1919 that their results were formally announced[10].”

When the armistice was signed on November 11, 1918, two expeditions were mounted, one to Sobral in Brazil, led by Andrew Crommelin from the Greenwich Observatory, and one to Principe Island off the coast of Spanish Guinea, led by Eddington. After the return of the expeditions, data analysis began. Einstein could not hide his enthusiasm in the expectation of the results. On September 22, 1919, Hendrik Lorentz sent a telegram to Einstein announcing:

“Eddington found stellar shift at solar limb,
tentative value between nine-tenths of a second and twice that.”

On the afternoon of November 6, 1919, at Burlington House in Piccadilly, the Astronomer Royal, Sir Frank Dyson, had the honour of presenting the results of the two expeditions. He described in detail the equipment, the photographs, and the complexities of the computations. His conclusion, however, was simple.

[10] Abraham Pais, *Subtle Is the Lord: The Science and the Life of Albert Einstein*, Oxford University Press, 1982, p. 304.

> "The results of the expeditions to Sobral and Principe leave little doubt that a deflection of light takes place in the neighbourhood of the sun and that it is of the amount demanded by Einstein's generalized theory of relativity."

Thus, it was considered that the astronomical observations had ''demonstrated'' the curvature of space. The day after, Einstein's name became legendary. The *London Times* published an article entitled "Revolution in Science, New Theory of the Universe, Newtonian ideas overthrown". On November 9, 1919, the *New York Times* published the following article:

"Diversion of light Rays Accepted as Affecting Newton's Principles[11]."

Thus, in 1919 the expedition led by Crommelin and Eddington "confirmed" Einstein's second prediction, that is to say, that the value of the deflection of the light passing near the Sun was equal to 1.7 seconds of arc[12]. Moreover, the question, which was the subject of intense controversy, was whether or not Einstein was aware of Soldner's work when he published his article in 1911.

B. Controversy around the curvature of light

In the early Twenties, Einstein and his theory of relativity were subject to many attacks of various natures[13]. The first, the most odious, had an anti-Semitic character. Then, in 1921, scientists such as Charles Lane Poor, a professor of astronomy, a specialist in Celestial

[11] See Jean-Marc Ginoux, *Albert Einstein: a biography through the Time(s)*, Hermann, Paris, 2016.
[12] Eddington's results for the solar eclipse observation of the apparent displacement of stars of 1919 have been widely disputed by many historians of science. See for example John Earman and Clark Glymour, "Relativity and Eclipses: The British Eclipse Expeditions of 1919 and Their Predecessors," *Historical Studies in the Physical Sciences*, Vol. 11, No. 1 (1980), p. 49-85.
[13] See Milena Wazeck, *On Einstein's opponents: the public controversy about the theory of relativity in the 1920s*, Cambridge; New York: Cambridge University Press, 2014.

Mechanics at Columbia University, tried to prove that Einstein's theories of relativity were false. To this aim he published an article entitled « Is Einstein wrong? – A debate[14] ». When Einstein was invited to the Collège de France in 1922 by Paul Langevin, his theories were challenged by Édouard Guillaume who had worked with him as a patent examiner at the Swiss patent office in Bern and had come purposely from Swiss in order to "destroy relativity". Indeed, he had published a few weeks beforehand an article entitled: "Y a-t-il une erreur dans le premier mémoire d'Einstein ?" which let no doubt concerning his intentions[15].

Unfortunately, this anti-relativistic attitude was shared by a part of the French scientific community and more particularly by many Academicians of Sciences. It has been initially convened that Einstein presents his work at the Academy but some of the members of the Academy had decided as a protest against his presence to rise and leave the hall as soon as he entered, Einstein had to renounce[16]. In August 1922, Einstein, who had received many death threats, decided to leave temporarily Europe. On October 8, 1922, Albert Einstein and his second wife Elsa came aboard the S.S. Kitano Maru in Marseille (south France) and reached then their final destination at Fukuoka in Japan where Einstein gave a lecture at the Daihaku Theater. On November 15, during a stopover in an hotel of Shanghai, Einstein received a telegram from Sweden announcing him that he has been awarded the Nobel Prize of Physics for "his contribution to theoretical physics and more particularly for his discovery of the law of photoelectric effect" and not for his relativity theories.

This news almost immediately triggered new reactions. Einstein's most virulent opponent was the German physicist Philipp Lenard, Nobel Prize in 1905 (the year of the relativity theory)

[14] Charles Lane Poor, "Is Einstein wrong? – A debate" in *The Forum*, June 1924, p. 705-715.
[15] Édouard Guillaume, "Y a-t-il une erreur dans le premier mémoire d'Einstein ?" *Revue Générale des Sciences Pures et Appliquées*, vol. 33, 15 janvier 1922, p. 5-10.
[16] See Michel Biezunski, *Einstein à Paris*, Paris, Presses universitaires de Vincennes, Saint Denis, 1991, p. 26.

and active proponent of the Nazi ideology. Then, in April 1923 Professors Ernst Gehrcke[17] of Berlin, P. Lenard of Heidelberg, O. E. Westin[18] of Stockholm charged Einstein with downright plagiarism, saying:

> "From these facts the conclusion seems inevitable that Einstein cannot be regarded as a scientist of real note. He is not an honest investigator." Thus Westin protested to the Directorate of the Nobel Foundation against the reward of Einstein[19]."

What were "theses facts" invoked by the three professors against Einstein? *The New York Times* of April 13 provides the answer:

> "… in 1801 Dr. J. von Soldner, a German physicist of eminence in his day, actually derived the formula recently used by Einstein. This was 122 years ago. Einstein never once mentions Soldner in his writings. This is bad enough, but the worst is yet to come.
>
> It has been shown by Professor Dr. E. Gehrcke, Director of the Imperial Physical and Technical Institute of Berlin, a position first filled by Helmohltz and by Professor P. Lenard of Heidelberg, winner of the Nobel Prize in Physics, that Soldner omitted a certain factor in his formula of l801, which error Einstein also copied when he appropriated the Einstein-Soldner formula in the Einstein paper of 1911. In a subsequent paper to the Berlin Academy of Science, 1915, Einstein

[17] E. Gehrcke (1878-1960) was a German experimental physicist, Director of the Imperial Institute of Berlin. He was a *Privatdozent* at the *Friedrich-Wilhelms-Universität* from 1904 to 1921 and an *außerordentlicher Professor* (*extraordinarius professor*) from 1921 to 1946.

[18] Oscar Edward Westin (1848-1930) was a Swedish engineer, professor of mechanical engineering at the *Royal Institute of Technology* in Stockholm.

[19] See *The New York Times*, April 13, 1923. This event is also mentioned in Philipp Frank, *Einstein: His Life and Times*, New York, A. A. Knopf, 1947, p. 202 and next.

> camouflaged this fraud as best he could, yet could not prevent its discovery and exposure by Professors Lenard of Heidelberg, Gehrcke of Berlin and Westin or Stockholm."

To these three scientists was added the professor Arvid Reuterdahl, Dean of the Engineering Department of the University of St. Thomas, St. Paul, Minnesota. He sent a letter to the Editor of *The New York Times* published on June 3, 1923 and in which he claimed:

> "There are two episodes in the Einsteinian development of the bending of light. In Einstein's 1911 paper the value of the deflection is given as 0.83 of a second. Soldner's value was 0.84 of a second. The two formulae are identical except in the matter of the convenient substitution by Einstein of different letters than those used by Soldner. Compared, letter for letter, the meanings are, however identical[20]. In his 1916 paper Einstein modified his 1911 value to read 1.7 of a second. Einstein has never taken the world into his confidence concerning the reason of this change. He has never admitted that either one or the other of these values must be erroneous[21]."

Indeed, in 1916, in his famous article entitled "The Foundations of the General Theory of Relativity[22]", Einstein realized that his earlier result on the bending of light, he had presented in his previous article of 1911, was too small by a factor of 2. This factor will be proved to be decisive because it enables to reveal a strict separation between Newtonian and Einsteinian theory of gravitation. This was due to the fact that in his 1911 paper Einstein had not taken

[20] See § C.3.
[21] We will see in § C.4 that this statement is partially inaccurate.
[22] Albert Einstein, "Die Grundlage der allgemeinen Relativitätstheorie" ("The Foundations of the General Theory of Relativity"), *Annalen der Physik*, 4(49), 769–822, 1916.

into account in his computations the curvature of space but, only included the effect of Newtonian gravitational interaction on the four-dimensional space-time that he will express later in a Minkowskian metric. According to Abraham Pais :

> "Let us briefly recapitulate Einstein's progress in understanding the bending of light.
>
> 1907. The clerk at the patent office in Bern discovers the equivalence principle, realizes that this principle by itself implies some bending of light, but believes that the effect is too small to ever be observed.
>
> 1911. The professor at Prague finds that the effect can be detected for starlight grazing the sun during a total eclipse and finds that the amount of bending in that case is 0''87. He does not yet know that space is curved and that, therefore, his answer is incorrect. He is still too close to Newton, who believed that space is flat and who could have himself computed the 0''87 (now called the Newton value) from his law of gravitation and his corpuscular theory of light.
>
> 1912. The professor at Zurich discovers that space is curved. Several years pass before he understands that the curvature of space modifies the bending of light.
>
> 1915. The member of the Prussian Academy discovers that general relativity implies a bending of light by the sun equal to 1''74, the Einstein value, twice the Newton value. This factor of 2 sets the stage for a confrontation between Newton and Einstein[23]."

[23] Abraham Pais, *op. cit.*, p. 303.

C. Comparison of the work of Einstein and Soldner

Such a comparison has been subject to many studies and several historians of science have analyzed both works of Soldner and Einstein from a mathematical point of view during these last decades. In 1975, Hans Fuchs published an article entitled "On the history of ideas about the effect of gravity on the light[24]" in which he presented the proofs leading to the value of the deflection of light passing near a massive body obtained by Soldner in 1801 and then by Einstein in 1911 and finally in 1915. After comparing Einstein's paper of 1911 with that of Soldner of 1801, Fuchs wrote: "Man erhält also trotz Zuhilfenahme des Aquivalenzprinzips wieder den alten falschen klassischen Wert! Wie ist das möglich?[25]". Then, he explained: "Obwohl wir von der Erde aus eine Verlangsamung der Vorgänge feststellen, werden wir doch mit gleichen Uhren auf der Sonne wie auf der Erde die gleichen Frequenzen und die gleiche Lichtgeschwindigkeit messen, weil die Uhren in gleichem Masse wie die Naturvorgänge verlangsamt werden!

Diese logisch einwandfreie Sicht der Dinge erlangte Einstein erst durch das tiefere Verständnis, das durch die allgemeinen Relativitätstheorie gebracht wurde (nach 1915)[26]." By using the Fields Equations of Gravitation and the Schwarzchild metric (including the so-called Ricci tensors) Fuchs showed that Einstein was then able to give the "der korrekte relativistische Wert" ("the correct relativistic value"). He also compared the relativistic and classical calculations and stated that the two results differ by a factor of two. However, his proof is different from those presented in this paper.

---

[24] Hans Fuchs, "Zur Geschichte der Ideen über die Wirkung der Schwerkraft auf das Licht," *Orion*, Vol. 33 (151), (December 1975), p. 183-193.

[25] "Thus, despite the aid of the principle of equivalence, one gets the old false classic value again! How is that possible?"

[26] "Although we observe a slowing down of the earth from the earth, we will measure the same frequencies and the same speed of light with the same clocks on the sun as on the earth, because the clocks are slowed down to the same extent as the natural processes!
Einstein achieved this logically correct view of things only through the deeper understanding that came with the general theory of relativity (after 1915)."

In 1978, Stanley Jaki published an article in which he recalled the historical context of the bending of light[27]. Starting from the seminal works of Newton and Laplace, he presented the controversy triggered out by Lenard, Gehrcke and Westin (see § B. above). Then, he proposed an English translation of Soldner's article allowing historians of science to study and compare his works with those of Einstein[28]. He gives many details and references concerning Soldner's life but didn't provide any mathematical analysis of his article.

In 1980, John Earman and Clark Glymour published a long article in which they compared Einstein's results of 1911 and 1915. As Fuchs[29], they recalled: "Einstein had not by 1911 yet absorbed the four-dimensional geometrical way of viewing space-time urged by Minkowski. In certain respects his thinking about space-time was still classical[30]." Then, they explained: "Einstein gave two arguments for the deflection of light passing near a massive body such as the sun; one argument, given in 1911 before the general theory was in hand, relied on his "principle of equivalence," while the other, given in 1916, used Einstein's own approximate solution to his gravitational field equations together with Huygens' principle from classical optics. The former derivation gave a value for the deflection at the limb of the sun of 0.83" of arc, the latter 1.7" of arc." They concluded that: "By 1916 Einstein had obtained two different expressions for the angular deflection of a light ray by a massive gravitational source, both giving the angle as a hyperbolic function of distance of closest approach to the massive body. The two expressions, one from the principle of equivalence and the other from the general theory, differ only by a factor of two." Nevertheless, they didn't compare Einstein's results with Soldner's.

---

[27] See also Jean Eisenstaedt, « De l'influence de la gravitation sur la propagation de la lumière en théorie newtonienne. L'archéologie des trous noirs », *Archive for History of Exact Sciences*, Vol. 42 (4), (September 1991), p. 315-386.

[28] Stanley L. Jaki, "Johann Georg von Soldner and the gravitational bending of light, with an English translation of his essay on it published in 1801," *Foundations of Physics*, December 1978, Vol. 8, Issue 11-12, p. 927-950.

[29] Hans Fuchs, "Zur Geschichte der Ideen über die Wirkung der Schwerkraft auf das Licht," *Orion*, Vol. 33 (151), (December 1975), p. 183-193.

[30] John Earman and Clark Glymour, "Relativity and Eclipses: The British Eclipse Expeditions of 1919 and Their Predecessors," *Historical Studies in the Physical Sciences*, Vol. 11, No. 1 (1980), p. 49-85.

In 1981, Hans-Jürgen Treder and Gerhard Jackisch published an article in which they considered that "A factor 2, which had been the occasion for misinterpretation, has to be attributed to the terminology used by German physicists and astronomers of that time[31]." They concluded that: "Soldner did nowhere draw false inferences but fell a victim to the printer's devil, and it is indisputable that Soldner obtained the Newtonian value of the deflection of light, which with respect to the constants of his times amounts to 0"84, and not to Einstein's value." Nevertheless, it is obvious that they compared Soldner's result with that provided by Einstein in 1915 and not with that of 1911. Moreover, their conclusion has been challenged by historians of science such as Ledo Stefanini[32] who wrote: "Some scholars attribute the numerical errors appearing in Soldner's memoir (correctly pointed out by Lenard in 1921 republication) to typographical errors[33], but this does not suffice to clear the issue." So, the aim of this work is to mathematically compare the formula obtained by Soldner with the one stated by Einstein in his paper of 1911 and to verify if they are identical or not. By using a simple first-order series expansion, it will thus be proved (to our knowledge for the first time) that both Soldner's and Einstein's formula are perfectly identical. Then, a mathematical analysis of the second formula concerning the bending of light established by Einstein in 1915 will enable to explain the importance of the doubling of this value in the framework of Einstein's theory of gravitation.

[31] Hans-Jürgen Treder and Gerhard Jackisch, "On Soldner's value of Newtonian deflection of light," *Astronomische Nachrichten*, Vol. 302 (6), (May 1981), p. 275-277.
[32] Ledo Stefanini, "A misunderstanding in Soldner's interpretation of the gravitational deflection of light," *Lettera Matematica*, Vol. 4 (3-4), (March 2017), p. 167-172.
[33] Stefanini quotes Hans-Jürgen Treder and Gerhard Jackisch.

1. Value of deflection of light by Soldner in 1801

Johann Georg von Soldner (1776-1833) was a German physicist, mathematician and astronomer. He first worked in the Berlin Observatory (*Berliner Sternwarte*) and later in 1808 in Munich where he became a member of the Academy of Sciences and the director of the observatory in Bogenhausen. In a paper written in March 1801 and published in 1804, he calculated the amount of deflection of a light ray by a star based on Newton‘s corpuscular theory of light and wrote:

> “It is, of course, true that already through observations and otherwise one was aware of considerable deviations from an assumed law; such as was the case with the aberration of light. There can, however, be deviations which are so small that it is difficult to decide whether they are true deviations or errors of observation. There can also be deviations which are considerable but, being combined with magnitudes one has not yet succeeded in clearly identifying, escape the observer. Of the latter kind may be the deviation of a light ray from straight line when it passes close by a celestial body and is considerably exposed to its attraction[34].”

Soldner then presented the following diagram (see Fig. 1).

[34] See J. von Soldner, “Ueber die Ablenkung eines Lichtstrals von seiner geradlinigen Bewegung,” (“On The Deviation Of A Light Ray From Its Motion Along A Straight Line Through The Attraction Of A Celestial Body Which It Passes Close By”), *Berliner Astronomisches Jahrbuch*, 1804, p. 161-172. See also Stanley L. Jaki, “Johann Georg von Soldner and the gravitational bending of light, with an English translation of his essay on it published in 1801,” *Foundations of Physics*, December 1978, Vol. 8, Issue 11-12, p. 927-950.

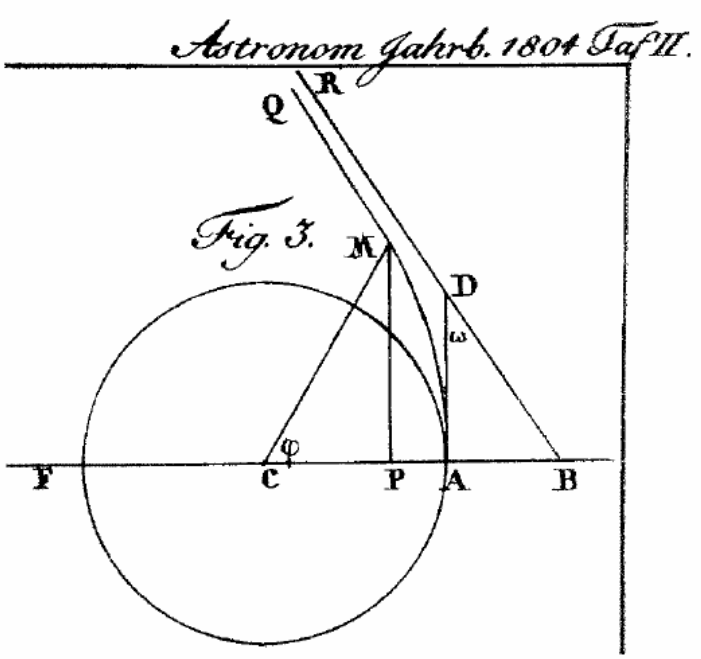


Fig. 1. Soldner's diagram for deflexion of a light ray.

Then, he indicated that "C (Fig. 1) is the center of the attracting body, A is the location at its surface. From A, a light ray goes into the direction AD or in the horizontal direction, by a velocity with which it traverses the way *v* in a second. Yet the light ray, instead of travelling at the straight line AD, will forced by the celestial body to describe a curved line AMQ, whose nature we will investigate (…) *g* be the gravitational acceleration at the surface of the body. Furthermore CP = *x*, MP = *y* and the angle $MCP = \varphi$." Then, he explained:

> "The force with which the light ray at M will be pulled by the body in the direction *MC*, will be $2gr^{-2}$. This force can be decomposed into two others, $\frac{2g}{r^2}cos(\varphi)$ and $\frac{2g}{r^2}sin(\varphi)$ according to the directions *x* and *y*; and therefore one obtains the following two equations (s. *Traité de mécanique céleste* par Laplace, Tome I, p. 21)

$$\frac{d^2x}{dt^2} = -\frac{2g}{r^2}\cos(\varphi)$$

$$\frac{d^2y}{dt^2} = -\frac{2g}{r^2}\sin(\varphi)\ldots"$$

These two equations correspond to the projection along the *x* and *y* directions of Newton's second law according to which the product of the mass by the acceleration (left hand side) is equal to Newton's gravitational force (right hand side).

Soldner's simplification of the "mass of a ray of light" on either side of this equality is consistent with Newton's theory of light then considered as made up of small discrete particles called "corpuscles". Indeed, it is only in 1803, three years after the writing of Soldner's article, that Thomas Young (1773-1829) performed his famous double-slit experiment from which he proposed a wave theory of the light. Thus, starting from the two preceding equations and after a demonstration which does not present any great difficulties, Soldner draws the following conclusion:

> "The light ray, however, comes in the direction DA to the eyes of the observer; thus ADB will be the angle of perturbation. If one calls this angle ω then one has, since the triangle ABD at A is a right triangle

$$\tan\omega = \frac{AB}{AD}$$

> If one puts these values for AB and AD in the expression for tan ω, then one has

$$\tan \omega = \frac{2g}{v\sqrt{v^2 - 4g}}$$

If one substitutes in the formula for tan ω the acceleration of gravity on the surface of the sun, and one takes the radius of that body for unity, then one finds ω = 0''.84. If one could observe the fixed stars very close to the sun, then one would have to take this very much into account. But since this is not known to happen, the perturbation caused by the sun can also be neglected[35]."

At the time, such observations were impossible; Soldner therefore concluded that these effects were minute. He ended his article as follows:

"Hopefully, no one would find it objectionable that I treat a light ray as a heavy body. That light rays have all the absolute [basic] properties of matter one can see from the phenomenon of aberration which is possible only because light rays are truly material. And furthermore, one cannot think of a thing which exists and works on our senses that would not have the property of matter[36]."

Thus, it appears that Soldner based his computations on the Newton's emission theory, according to which light is made up of particles. As far as Einstein is concerned, he made use of the Huygens principle, that is to say, the variation of the direction of the wavefront as a function of the luminous frequency, as will be seen in the next section.

[35] *Ibid.*
[36] *Ibid.*

2. Value of deflection of light by Einstein in 1911

In his 1911 paper, Einstein first established that "the velocity of light in the gravitational field is a function of the location[37]." Then, by using Huygens's principle, he stated that "light-rays propagated across a gravitational field undergo deflection[38]."

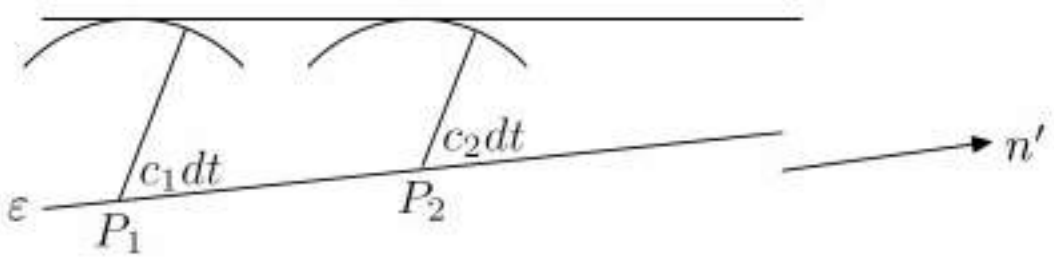


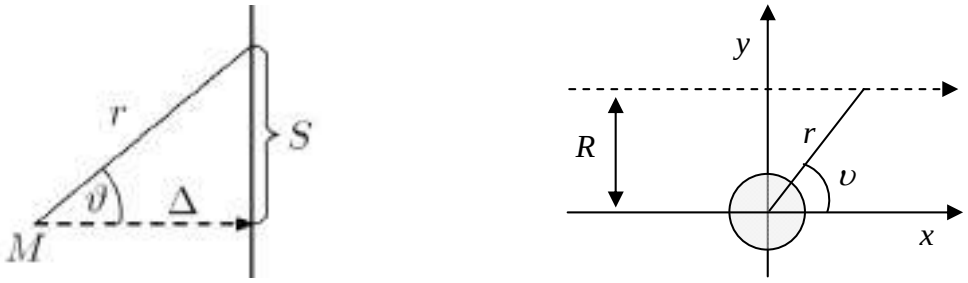
Fig. 2. Einstein's diagram for deflexion of a light ray.

From this figure (see Fig. 2), Einstein stated that the direction of the wave front changes by an amount equal to $-\partial c/\partial n'$ per unit of distance along the direction of the wave (where c is the velocity of light) and the "angle of deflection per unit of path of the light-ray is $-\frac{1}{c}\frac{\partial c}{\partial n'}$". Finally, he obtained "for the deflection α, which a light-ray experiences toward the side n' on any path (s) the expression

$$\alpha = -\frac{1}{c^2}\int \frac{\partial \Phi}{\partial n'} ds \text{ "}$$

[37] Albert Einstein, "Einfluss der Schwerkraft auf die Ausbreitung des Lichtes" ("On the influence of Gravitation on the Propagation of Light"), *Annalen der Physik*, 4(35), 898–908, 1911.
[38] *Ibid*.

where the integral goes from $-\infty$ to $+\infty$ and $\Phi = kM/r$ is the gravitation potential. Then, Einstein changed variables to polar coordinates as highlighted on the following figure (see Fig. 3a). In order to simplify the understanding of his approach, let's pose in what follows: $n' = y$, $s = x$ et $S = R$ (see Fig. 3b)[39].



(a) (b)

Fig. 3. Deflection of a light ray in polar coordinates.

According to Fig. 3b, we have $r = \sqrt{x^2 + y^2}$. Thus, the gravitation potential reads:

$$\boxed{\Phi = \frac{kM}{\sqrt{x^2 + y^2}}}$$

It follows that $\frac{\partial \Phi}{\partial n'} = \frac{\partial \Phi}{\partial y} = \frac{\partial}{\partial y}\left(\frac{kM}{\sqrt{x^2 + y^2}}\right) = -\frac{kM}{r^3} y$. Einstein then considers that nearly all of the deflection occurs within some reasonable proximity of the gravitating body. So, we can simply set $y = R$ in the integral which reads:

[39] See also Kevin Brown, *Reflections on Relativity*, lulu.com, Mars 2017.

$$\alpha = \frac{1}{c^2}\int_{-\infty}^{+\infty}\frac{kM}{r^3}\,yds = \frac{kMR}{c^2}\int_{-\infty}^{+\infty}\frac{1}{r^3}\,dx\,.$$

The following variable changes $x = R\,tan(\upsilon)$, $y = R$ leads Einstein to:

$$\alpha = \frac{kMR}{c^2}\int_{-\infty}^{+\infty}\frac{1}{\left(x^2+y^2\right)^{3/2}}\,dx = \frac{kMR}{c^2}\int_{-\pi/2}^{+\pi/2}\frac{Rd\upsilon/cos^2(\upsilon)}{\left(R^2tan^2(\upsilon)+R^2\right)^{3/2}} = \frac{kM}{c^2R}\int_{-\pi/2}^{+\pi/2}cos(\upsilon)\,d\upsilon = \frac{2kM}{c^2R}$$

Then, Einstein ended his 1911 paper by this sentence:

"*A light-ray going past the Sun would accordingly undergo deflection by the amount of* 4.10^{-6} = 0,83 *seconds of arc*[40]."

Let's notice that although the reasonings and the computations are different, the result of Einstein is exactly the same as that of Soldner (as shown in the next section). In other words, a light ray passing near the sun will undergo a deflection of nearly 0.83''.

3. Comparison of Einstein and Soldner formulas

This section aims to prove that both formulas established by Soldner in 1801 and Einstein in 1911 are identical except in the choice of letters.

Soldner's formula of 1801 reads: $\tan\omega = \dfrac{2g}{v\sqrt{v^2-4g}}$

- g is the gravitational acceleration at the surface of the body

- v is its velocity.

[40] *Ibid.*

In his article, Soldner wrote:

> "On the presupposition that light needs 564''.8 decimal seconds of time to come from the sun to the earth, one finds that it traverses in one-tenth of a second 15.562085 earth radii. Thus v = 15.562085. If one takes among the geographical latitudes that whose square of the sine is 1/3 (corresponding to a latitude of 35°16'), the earth's radius as 6,369,514 meters, and the acceleration of gravity there as 3.66394 meters (see *Traité de mécanique céleste* by Laplace, Vol. I, p. 118), then expressed in earth radii g = 0.000000575231."

We deduce that the dimension of the acceleration of gravity on the surface of the body g is s^{-2} and that of the velocity is s^{-1}. Indeed, Soldner made use of a kind of "normalisation" of these two physical variables which can be written as $g = g_0/\Delta$ and $v = c/\Delta$ where Δ is the radius of the considered body (in this case the Sun). Soldner's formula can be rewritten as

$$\tan\omega = \frac{2g}{v\sqrt{v^2-4g}} = \frac{2g}{v^2}\frac{1}{\sqrt{1-\frac{4g}{v^2}}} = \frac{2g}{v^2}\left[1-\frac{4g}{v^2}\right]^{-\frac{1}{2}}$$

But since, according to Soldner $g << v$, a first-order[41] series expansion can be made and reads:

$$\tan\omega = \frac{2g}{v^2}\left[1-\frac{4g}{v^2}\right]^{-\frac{1}{2}} = \frac{2g}{v^2}\left[1+\frac{4g}{v^2}+\cdots\right] \approx \frac{2g}{v^2}+O\left(v^{-4}\right)$$

[41] In his 1911 and 1915 articles Einstein also made use of first-order approximations.

The angle ω is considered as infinitely small, so we deduce that $\omega \approx \frac{2g}{v^2}$.

Then, by posing $g = g_0/\Delta$ and $v = c/\Delta$, we have:

$$\boxed{\omega \approx \frac{2g_0\Delta}{c^2}}$$

In 1911, Einstein proposed the following formula: $\alpha = \frac{2kM}{c^2\Delta}$ where

- k is the constante of gravitation (Cavendish's constant),
- M is the mass of the heavenly body,
- Δ is the distance of the ray from the center of the body, i.e., its radius.

In the case of the Sun, we have: $g_0 = \frac{kM}{\Delta^2}$. So, it gives

$$\alpha = \frac{2kM}{c^2\Delta} = \frac{2\Delta}{c^2}\left(\frac{kM}{\Delta^2}\right) = \frac{2g_0\Delta}{c^2}$$

Thus, both Soldner and Einstein's formulas are identical.

$$\boxed{\omega = \alpha = \frac{2g_0\Delta}{c^2}}$$

4. Value of deflection of light by Einstein in 1915

During the autumn 1915, Einstein completed his general theory of relativity. He thus modified his gravitation potential Φ while taking into account as previously the Newton's gravitational interaction (see § C.2) but also the curvature of space near a massive body[42]. Then, he expressed the gravitation potential Φ as

$$\boxed{\Phi = \frac{kM}{\sqrt{x^2 + y^2}} + \frac{mx^2}{\left(x^2 + y^2\right)^{3/2}}}$$

In this case $\frac{\partial \Phi}{\partial n'} = \frac{\partial \Phi}{\partial y} = \frac{\partial}{\partial y}\left(\frac{kM}{\sqrt{x^2 + y^2}} + \frac{kMx^2}{\left(x^2 + y^2\right)^{3/2}}\right) = -\frac{4x^2 + y^2}{r^5} kMy$

The integral reads then $\alpha = \frac{1}{c^2}\int_{-\infty}^{+\infty} \frac{4x^2 + y^2}{r^5} kMyds = \frac{kMR}{c^2}\int_{-\infty}^{+\infty} \frac{4x^2 + R^2}{r^5} dx$. By using the same variable changes, $x = R\,tan(\upsilon)$, $y = R$, Einstein obtained:

$$\alpha = \frac{kMR}{c^2}\int_{-\infty}^{+\infty} \frac{4x^2 + R^2}{r^5} dx = \frac{kMR^3}{c^2}\int_{-\pi/2}^{+\pi/2} \frac{4tan^2(\upsilon) + 1}{\left(R/cos(\upsilon)\right)^5} \frac{Rd\upsilon}{cos^2(\upsilon)} = \frac{4kM}{c^2 R}$$

[42] Cf. Kevin Brown, *Op. Cit.*

In his publication of 1916 Einstein[43] provided for the deflection of a light ray the value $B = \frac{\kappa M}{2\pi\Delta}$ which results of the previous integration. He defined at the page 818 of this same article the constant $\kappa = \frac{8\pi k}{c^2}$ (see Eq. (69)). By replacing into the value of *B*, we have:

$$\boxed{B = \frac{4kM}{c^2 R} = 2\alpha}$$

Einstein wrote in his conclusion:

"According to this, a ray of light going past the sun undergoes a deflexion of 1.7''…"

Thus, it appears that Einstein's computation of the value of deflection of a light ray performed in 1915 led him to twice the amount derived in his 1911 paper.

Where does this doubling come from? How did Einstein justify it?

In fact, contrary to what Arvid Reuterdahl (see § B) claimed, Einstein has really "taken the world into his confidence concerning the reason of this change". Indeed, as early as 1915, Einstein wrote:

> "By use of the Huygens principle, one finds through a simple calculation, that a light ray from the Sun at distance Δ undergoes an angular deflection of magnitude $2\alpha/\Delta$, while the earlier calculation had given the value α/Δ. A

[43] Albert Einstein, "Die Grundlage der allgemeinen Relativitätstheorie" ("The Foundations of the General Theory of Relativity"), *Annalen der Physik*, 4(49), 769–822, 1916.

corresponding light ray from the surface rim of the Sun should give a deviation of 1.7'' (instead of 0.85'')[44]."

In 1920, in the appendix 3 of the third edition of his book written in 1916, he wrote:

"As a result of this theory, we should expect that a ray of light which is passing close to a heavenly body would be deviated towards the latter. For a ray of light which passes the sun at a distance of Δ sun-radii from its centre, the angle of deflection (α) should amount to

$$\alpha = \frac{1.7 \text{ seconds of arc}}{\Delta}$$

It may be added that, according to the theory, half of this deflection is produced by the Newtonian field of attraction of the sun, and the other half by the geometrical modification ("curvature") of space caused by the sun[45]."

D. Conclusion

The author of one of Einstein's most famous biographies, Ronald Clark had written that he was "the man who had bent the light". One would be led to believe that the concept of "curvature of light" which was conceived by Soldner in the early nineteenth century was rediscovered a century later by Einstein. In fact, many authors have shown that this idea was

[44] Albert Einstein, "Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie" ("Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity"), *Preussische Akademie der Wissenschaften*, *Sitzungsberichte*, 831–839, 1915.
[45] Albert Einstein, *Relativity*: *The Special and General Theory*, London, Methuen & Co Ltd, 1920.

already present in the works of Isaac Newton. Indeed, in the first of the famous Queries of his work entitled *Opticks*, Newton wrote in 1704:

« Do not Bodies act upon Light at a distance, and by their action bend its Rays, and is not this action *(caeteris paribus)* strongest at the least distance? »

It thus appears that neither Soldner nor Einstein is the inventor of the concept of "curvature of light". One can then ask whether Einstein had any knowledge of Soldner's work. It is naturally very difficult to answer this question. The elements we have today allow us only to affirm that the values of the deflection of light rays passing near a massive body obtained by Soldner in 1801 and by Einstein in 1911 are perfectly identical, although the computations of Soldner were based on Newton's corpuscular theory, while those of Einstein were based on the Huygens's principle. First, it is important to note that the impossibility of measuring the deflection of light during the eclipses of 1912 and 1914 was an extraordinary opportunity for Einstein. Indeed, without this providential rain and without the declaration of war the observations of the astronomers would have absolutely not confirmed the first value that it had provided in 1911 and they would certainly have invalidated his theory.

Note then that the method of computing the deflection of light that Einstein used in 1915 is exactly the same one he used in 1911. The only difference is the expression of gravitation potential which took into account the curvature of the space in the vicinity of a massive body. Thus, in his article in 1915, Einstein provided a value (1.7'') which was well contained within the range of values observed by the expeditions led by Crommelin and Eddington in 1918 (0.9'' to 1.8''). It is clear from this analysis that the plagiarism accusations against Einstein which are part of an anti-relativist and anti-Semitic movement are absolutely baseless. Even if it has been established that the values of the deflection of a light ray provided by Soldner in

1801 and Einstein in 1911 are identical, they are both wrong because they don't take into account the curvature of space and so, are not consistent with those "obtained" by Crommelin and Eddington and confirmed after by more accurate astronomical observations.

E. References

Michel Biezunski, *Einstein à Paris*, Paris, Presses universitaires de Vincennes, Saint Denis, 1991.

Kevin Brown, *Reflections on Relativity*, lulu.com, Mars 2017.

Ronald W. Clark, *Einstein the life and times*, London, Hodder and Stoughton, 1973, p. 494.

John Earman and Clark Glymour, "Relativity and Eclipses: The British Eclipse Expeditions of 1919 and Their Predecessors," *Historical Studies in the Physical Sciences*, Vol. 11, No. 1 (1980), p. 49-85.

Jean Eisenstaedt, « De l'influence de la gravitation sur la propagation de la lumière en théorie newtonienne. L'archéologie des trous noirs », *Archive for History of Exact Sciences*, Vol. 42 (4), (September 1991), p. 315-386.

Albert Einstein, "Relativitätsprinzip und die aus demselben gezogenen Folgerungen" ("On the Principle of Relativity and the Conclusions Drawn from it") *Jahrbuch der Radioaktivität*, 4, 411–462, 1907.

Albert Einstein, “Einfluss der Schwerkraft auf die Ausbreitung des Lichtes” (“On the influence of Gravitation on the Propagation of Light”), *Annalen der Physik*, 4(35), 898–908, 1911.

Albert Einstein, “Erklärung der Perihelbewegung des Merkur aus der allgemeinen Relativitätstheorie” (“Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity”), *Preussische Akademie der Wissenschaften*, *Sitzungsberichte*, 831–839, 1915.

Albert Einstein, “Die Grundlage der allgemeinen Relativitätstheorie” (“The Foundations of the General Theory of Relativity”), *Annalen der Physik*, 4(49), 769–822, 1916.

Albert Einstein, *Relativity*: *The Special and General Theory*, London, Methuen & Co Ltd, 1920.

Philipp Frank, *Einstein: His Life and Times*, New York, A. A. Knopf, 1947

Hans Fuchs, “Zur Geschichte der Ideen über die Wirkung der Schwerkraft auf das Licht,” Orion, Vol. 33 (151), (December 1975), p. 183-193.

Thomas Glick, *The Comparative Reception of Relativity,* Boston Studies in the Philosophy and History of Science, Vol. 103, Springer Netherlands, Dordrecht, D. Reidel, 1987;

Jaume Giné, “On the origin of the deflection of light,” *Chaos, Solitons & Fractals*, 2008, vol. 35(1), p. 1-6.

Jean-Marc Ginoux, *Albert Einstein : une biographie à travers le temps*, Hermann, Paris, 2016.

Jean-Marc Ginoux, *Albert Einstein: a biography through the time(s)*, Hermann, Paris, 2016.

Édouard Guillaume, « Y a-t-il une erreur dans le premier mémoire d'Einstein ? » *Revue Générale des Sciences Pures et Appliquées*, vol. 33, 15 janvier 1922, p. 5-10.

Stanley L. Jaki, "Johann Georg von Soldner and the gravitational bending of light, with an English translation of his essay on it published in 1801", *Foundations of Physics*, December 1978, Vol. 8, Issue 11-12, p 927-950.

Charles Lane Poor, « Is Einstein wrong? – A debate » in *The Forum*, June 1924, p. 705-715.

Abraham Pais, *Subtle Is the Lord: The Science and the Life of Albert Einstein*, Oxford University Press, 1982.

Jürgen Renn, The Genesis of Relativity, Vol. 1-4, Boston Studies in the Philosophy and History of Science, Vol. 250, Springer Netherlands, Dordrecht, D. Reidel, 2007

Johann von Soldner, "Ueber die Ablenkung eines Lichtstrals von seiner geradlinigen Bewegung," ("On The Deviation Of A Light Ray From Its Motion Along A Straight Line Through The Attraction Of A Celestial Body Which It Passes Close By"), *Berliner Astronomisches Jahrbuch*, 1804, p. 161-172.

Ledo Stefanini, “A misunderstanding in Soldner’s interpretation of the gravitational deflection of light,” *Lettera Matematica*, Vol. 4 (3-4), (March 2017), p. 167-172.

Hans-Jürgen Treder and Gerhard Jackisch, “On Soldner’s value of Newtonian deflection of light,” *Astronomische Nachrichten*, Vol. 302 (6), (May 1981), p. 275-277.

Milena Wazeck, *On Einstein’s opponents: the public controversy about the theory of relativity in the 1920s*, Cambridge; New York: Cambridge University Press, 2014.

A Statistical Ensemble Based Approach for Entropy in Cryptocurrencies Markets

Luca Grilli[1] and Domenico Santoro[2]

[1] Università degli Studi di Foggia, Dipartimento di Economia, Management e Territorio, Via Da Zara, 11, I-71121 Foggia, Italy
(E-mail: luca.grilli@unifg.it)
[2] Università degli Studi di Foggia, Dipartimento di Economia, Management e Territorio, Via Da Zara, 11, I-71121 Foggia, Italy
(E-mail: domenico_santoro.553606@unifg.it)

Abstract. In this paper, we point out an "affinity" between the system of agents trading in cryptocurrencies and statistical mechanics. In particular, we try to extend the concept of entropy in the sense of Boltzmann such a definition to a model in which the particles are replaced by N economic subjects (agents), that are completely described by their ability to buy and to sell a certain quantity of cryptocurrencies. In addition, by applying this model to the closing prices of some of this we show that entropy can be used as an indicator to forecast the price trend of cryptocurrencies.
Keywords: Cryptocurrency, Entropy, Prices Forecast, Boltzmann, Blockchain.

1 Introduction

The concept of entropy was first introduced by Clausius[23], whose definition was applied to a thermodynamic system that performs a transformation. Since the mid-19th century, entropy has been a key element linking mechanics to thermodynamics; however, this entropy suffered from a conceptual problem which, as demonstrated by Gibbs[28], was revealed in the case of identical gases (*Gibbs Paradox*). He solved this problem by changing the count of states. On the other hand, Boltzmann[16] presented his statistical interpretation of thermodynamic entropy, managing to link the macroscopic properties of a system with the microscopic ones. Based on Gibbs, in 1949 Shannon[8] developed a theory capable of evaluating the amount of information that is lost in receiving a message from a source to a recipient. This form of entropy was generalized by Rényi[3], Tsallis[7], Adler et al.[15] (in topology), redefined by Pincus[19] (*approximate entropy*) and - more recently - by Chen et al.[27] as a time series regularity measure.
The application of entropy in sectors such as economics or finance is linked to the work of Brissaud[4] that assimilated **entropy** to disorder, so as to make this tool that has always been applied the physics part of the economy. The forms of entropy most used in this case are Shannon entropy and the generalizations by Rényi and Tsallis, who contributed to creating a new line of application for the management of financial portfolios. For example, these new types of entropy has been used by Philippatos and Wilson[6], Usta and Kantar[10], Jana et al.[22], Gulko[17], and Dionisio et al.[2].
What we want to demonstrate in this paper is that it is possible to assimilate

a cryptocurrency system to a thermodynamic system. In this way, we are able to determine entropy in the sense of Boltzmann so that we can make price predictions related to the possibility that they move in a more or less wide range; unlike all the recent applications concerning theories based on Shannon entropy and its derivations. Innovation is linked to the reinterpretation of the monetary system of cryptocurrencies. In this sense, we can apply physical theories to a social science. Once the system has been described, our goal is to verify that entropy calculated in the physical sense also occurs in the economic context to allow us to make assumptions on how the process could move in the next future. This type of conjecture has been presented by Sergeev[20], Zakiras et al.[26] and Smith and Foley[9]. In particular McCauley[13], based on this previous theory, maintains that the illiquidity of the markets does not allow for the application of the concepts of statistical mechanics.
The paper structure is the following: in Section 2 we analyze cryptocurrencies and their key characteristics, focusing on the fact that they have a supply limit; in Section 3 we describe the evolution of a system of a particle in statistical thermodynamics and how to determine its entropy, subsequently applying these notions to our monetary system; in Section 4 we define the theoretical assumptions we can link to the system created previously to study the price evolution in these currency markets and we analytically describe the calculation of entropy using real data; finally in section 5 some conclusions are drawn.

2 Cryptocurrency

Cryptocurrencies represent a digital currency system with no guarantee institution and no transaction control. The main cryptocurrencies, by media coverage or by the possibility that some financial intermediaries offer to use them as a payment instrument, are: Bitcoin, Ethereum and Ripple. Unlike traditional financial assets, their value is not based on tangible assets such as the economy of a country or a company, but it is based on the security of an algorithm that tracks transactions. Their definition is controversial since by some entities [11] they are considered intangible assets (IFRS) while according to the German financial supervisory authority (BaFin) they are officially financial instruments [5]. All the cryptocurrencies have been based on the *Bitcoin*, a currency created by Nakamoto[24] who in 2009 released a software capable of implementing transactions. The currency itself is a unique alphanumeric string that represents a certain transaction, a transaction which will then be entered in a public register called *blockchain*.
The blockchain is the fulcrum of these systems and is essentially a register in which the data of the owners of the currency are entered, transactions occur in an encrypted manner. The blockchain is a data structure consisting of a list of transaction blocks linked together so that each refers to the previous one in the chain. A block is a data structure that aggregates transactions to include them in the public register. The block is made of a header, containing metadata, followed by a long list of transactions. A complete block, with all transactions, is, thus, 1000 times larger than the block header [1]. The integrity of the blockchain network is guaranteed through consensus algorithms

such as *Proof-of-Work* (PoW) and *Proof-of-Stake* (PoS), that solve the Byzantine Generals Problem[12] (problem of consent in the presence of errors). A consensus algorithm is a mechanism used by the network to reach consensus, i.e. ensuring that the protocol rules are followed and that transactions occur correctly so that coins can only be spent once.
The cryptocurrency generation process is called *mining*, which adds money to the supply. Cryptocurrencies are "minted" during the creation of each block at a fixed and decreasing rate [1]: each block generated on average every 10 minutes contains new currency. For example, if we consider Bitcoin, every 210000 blocks the currency issue rate decreases by 50% (the availability of new coins grows as a geometric series every 4 years). It is estimated that around the year 2140, the production of the last block will be reached (6930000) and the number of coins produced will tend to its upper limit of 21 million (precisely 20999999.97690000), value introduced by Nakamoto himself and contained in the variable "MAX_MONEY" as can be read in the source code present on GitHub. This value represents a sanity check, especially used to avoid bugs in which it is possible to generate currency from nothing and therefore moving towards a situation in which the blockchain diverges into different potential paths (called *fork*).

3 Methodology

The main assumption in this paper is that the prices of cryptocurrencies behave like a thermodynamic system, so it is possible to determine entropy by using the Boltzmann formula. In order to present the theoretical framework and the methodology, we need to briefly introduce the main physical results. In Statistical Mechanics a macroscopic system is made up of N molecules ($N \sim 10^{24}$ is the Avogadro's constant) whose mechanics provide the evolution of $6N$ dynamic variables describing completely the microscopic states of this system. Motion in the phase space can be studied using the $3N$ position components and the $3N$ momenta components, indicated with $\{\mathbf{q}_i\}$ and $\{\mathbf{p}_i\}$ whose evolution is driven by Hamilton's equations. Mechanics, therefore, provides a very detailed description of the system contrary to thermodynamics which studies the collective variations; for this reason, the mechanical point of view can be defined *microscopic* and the thermodynamic one *macroscopic*. The study of the system from a microscopic point of view concerns experimental observation on one or a few molecules.
Everything that happens from the microscopic side can be expressed in macroscopic terms through thermodynamics, defined in this case as a large amount of microscopic variables. We consider an isolated system of N particles described by the $3N$ coordinates and the $3N$ momenta in a $6N$-dimensional space at a certain time t. Particles are subject to the laws of classical mechanics and therefore $\mathbf{X}(t)$ evolves according to Hamilton's equations. Since the Hamiltonian $H(p, q)$ does not depend on time, the energy E is a conserved quantity during motion and develops on a fixed hypersurface. We want, for example,

Source: https://github.com/bitcoin/bitcoin/blob/master/src/amount.h

to measure an observable $A(\mathbf{X})$ (a function defined in the phase space) of the system in thermodynamic equilibrium, but since the scale of macroscopic times is much larger than the microscopic one, we can consider a datum as the result of a system that has gone through a large series of microscopic states; this implies that the observable must be compared with an average performed along with the evolution of the system calculated over very long times $\bar{A}$. The calculation of $\bar{A}$ would require knowledge of both the microscopic state at a certain moment and the determination of the corresponding trajectory in the phase space, which corresponds to a practically inexhaustible request. To determine the observable, the *ergodic* theory intervenes, according to which each energy surface is completely accessible to any motion with the given energy and the average residence time in a certain region is proportional to its volume. If these conditions are satisfied, the average $\bar{A}$ can be calculated as the average of $A(\mathbf{X})$ in which the states with the fixed energy contribute with equal weight. In applications it is convenient to consider on average all states with energy within a fixed range $[E, E + \Delta E]$; furthermore, we are only interested in some macroscopic properties such as particle number N and the volume V. There is an infinite number of systems that satisfy these conditions: these form the *Gibb's ensemble* which is represented by a set of points in the phase space characterized by a density function $\rho(p, q, t)$ defined so that $\rho(p, q, t)\, d^{3N}p\, d^{3N}q$ corresponds to the number of representative points of the system during the instant t contained in the infinitesimal volume of the phase space $d^{3N}p\, d^{3N}q$. Furthermore, since energy, volume and number of particles are constants of motion, the total number of systems in an ensemble is conservative.
We can thus introduce the *postulate of equal a priori probability* who claims that when a macroscopic system is in thermodynamic equilibrium its state can be with equal probability each of those which satisfies the macroscopic conditions of the system. This postulate implies that the system under consideration belongs to an ensemble called *microcanonic* with density function

$$\rho(p, q) = \begin{cases} \rho^* & if\ E < H(p, q) < E + \Delta \\ 0 & otherwise \end{cases} \tag{1}$$

where ρ^* is constant and all members of the ensemble have the same number of particles and equal volume.
We can define $\Gamma(E)$ the volume occupied by the microcanonical ensemble in the phase space as:

$$\Gamma(E) \equiv \int_{E<H(p,q)<E+\Delta E} d^{3N}p\, d^{3N}q \tag{2}$$

and $\Sigma(E)$ the volume bounded by the energy surface E:

$$\Sigma(E) \equiv \int_{H(p,q)<E} d^{3N}p\, d^{3N}q \tag{3}$$

so that

$$\Gamma(E) = \Sigma(E + \Delta E) - \Sigma(E). \tag{4}$$

Entropy, then, can be defined as:

$$\begin{aligned} S_\Gamma &= \int_{E\leq H\leq E+\Delta E} d^{3N}p\, d^{3N}q\, \rho(-\kappa_B \ln \rho) \\ &= \int_{E\leq H\leq E+\Delta E} d^{3N}p\, d^{3N}q\, \frac{1}{\Gamma}\left(-\kappa_B \ln \frac{1}{\Gamma}\right) \\ &= \frac{1}{\Gamma}\kappa_B \ln \Gamma \int_{E\leq H\leq E+\Delta E} d^{3N}p\, d^{3N}q \\ &= \frac{1}{\Gamma}\kappa_B \ln \Gamma\cdot\Gamma = \kappa_B \ln \Gamma(E) \end{aligned} \tag{5}$$

where $\kappa_B \sim 1.3806 * 10^{-23}$ is the Boltzmann constant. To analytically calculate $\Gamma(E)$, which represents the number of states accessible to the system at temperature T, we must consider that a microcanonical ensemble is made up of J identical copies of the closed system, each of which is located in a microstate $(\mathbf{p}_i, \mathbf{q}_i)$ of the phase space. Being all on the same hypersurface E, we can divide it into cells of equal size, where in each there are j_i systems such that $J = \sum_i j_i$. To define the system it is necessary to find the most probable distribution of the j_i microstates, that is, to count the total number of ways in which we can obtain a certain macrostate. In the Boltzmann paradigm with an ideal gas consisting of identical particles under the same conditions, we can say that

$$\Gamma(E) = \frac{J!}{\prod_i j_i!} \tag{6}$$

The idea that entropy is connected to volumes in the phase space finds its origin in the *Helmholtz Theorem*, whose goal is to exactly bring thermodynamics down from mechanics.

Let us now try to translate this physical theory into a financial dress. Viaggiu et al.[25] have developed a representation of an economic model relating to money from a thermodynamic point of view. In their description the ensemble is made up of the N interacting economic subjects, entirely described by two variables $\{x_i, y_i\}$ which represent money and credit/debt capacity and which are not conjugated in the sense of mechanics Hamiltonian. The key characteristic is to consider a representative function of the total currency as a conservative law, to be able to exploit the ergodic hypothesis.

Our idea is to go back to their hypothesis by applying it to the case of cryptocurrencies. We consider a model in which the particles are replaced by N economic subjects (agents) who intend to trade in cryptocurrencies (compared only to a reference currency, such as the USD). These agents are completely described by 2 variables, which we can, however, identify as $\{x_i, y_i\}$, where x_i and y_i indicate, respectively, the ability to buy and to sell a certain quantity of cryptocurrencies (both expressed in monetary terms). The latter hypothesis is possible according to the fact that the market to which we refer is influenced only by the supply and demand leverage. As for [25], even if the complete Hamiltonian formalism is not respected, we can consider as a conserved quantity the total number of cryptocurrencies in circulation which by their definition is constant over a suitable time interval through the function $M(x_i, y_i)$ (as in

the particular case of Bitcoins for which the supply limit is fixed at 21 million). However, since the supply limit has not yet been reached by any cryptocurrency we consider this quantity constant concerning the currency in circulation in a precise time t, therefore:

$$M = \sum_{i=1}^{N} x_i + y_i. \tag{7}$$

In this sense, the sum of the ability to sell and buy of the N agents fully describes the cryptocurrencies in circulation. The ergodic hypothesis allows us, given a certain function $f(x_i, y_i)$, to express its average with respect to the time in terms of an average over the ensemble at fixed M:

$$\bar{f} = \int_{M=const} f(x, y)\rho(x, y)\, dx\, dy \tag{8}$$

where $\rho(x, y)$ denotes the probability distribution of the ensemble. Through these assumptions we can verify the economic transformations through thermodynamics; in particular, as in statistical mechanics, we can calculate the *volume in the phase space* [25] . If we integrate over all the available volume of the configuration space spanned by $\{x, y\}$ with $\bar{M} = m$ (where $\bar{M}$ denotes the average over the whole configuration space) we have $\int_{\bar{M}=m} d^N x\, d^N y = 0$. So introducing a thick shell Δ where $\Delta \ll m$ we can define:

$$\Gamma(m) = \int_{m<M<m+\Delta M} \frac{d^N x\, d^N y}{k^{2N}} \tag{9}$$

where $d^N x\, d^N y$ is understood as the phase space and k is a normalization factor such that Γ is dimensionless. This functional represents the number of microscopic realizations of the system under examination and allows us to calculate the entropy S as described in the equation (5).

4 The model

In this section, however, we try to define, through a new type of approach, how it is possible to calculate entropy considering essentially the prices obtainable from the currency markets (FOREX).
First, we know that cryptocurrencies are used by an approximate number of economic entities equal to 44 million (based on the number of blockchain portfolios[21]) for which $N \gg 1$. We also know that every subject in our system is fully described by its ability to buy and sell ($\{x_i, y_i\}$). Let us consider that these two variables are summarized in the *last prices* of the cryptocurrency on the currency markets, a type of price used to keep track of changes in the value of an asset throughout a session. In this sense, the latest prices allow us to understand whether, compared to the previous session, the ability to buy or sell prevailed. We can summarize this price capability in the sentence *"prices describe the strength with which agents position themselves in the phase space"*.

The key point is that we can use the function M (described above) because in a certain time t the quantity of cryptocurrencies is constant and quantifiable, in this way we can go back to the previous economic model and determine Γ as described in the equation (9). Analytically, we do not consider the number of economic subjects present in the market but indirectly deduce their "position" in the phase space from the difference between the closing prices. In particular, first we cluster the closing price series based on a certain reference interval (5 days); as for each cluster there is a maximum and a minimum price, we calculate the difference in terms of necessary steps to pass from one to the other obtaining a certain value of **gap G** (this assumption is based on the idea that the distance between maximum and minimum is a measure of the dispersion of agents in our phase space); to calculate the "volume" occupied by the disposition of the agents we use combinatorial analysis, therefore:

$$\Gamma = G^5 \tag{10}$$

Once the value of Γ is determined, entropy can be calculated by using the Boltzmann formula:

$$S = \kappa_B \ln \Gamma. \tag{11}$$

Finally, precisely because Boltzmann's constant is of the order of Avogadro's number, we can *"rationalize"* this entropy value obtained by multiplying it by 10^{23}. Our data analysis shows that in situations where entropy is drastically reduced, in the following phase it must grow in an "almost obligatory" way; this in terms of cryptocurrency prices indicates that in situations in which the gap between the maximum and the minimum is drastically reduced in the transition from one cluster to another "almost compulsorily" follows a situation in which it is certainly wider than the previous one. This type of price-based entropy defines how agents move in the phase space, so it allows us to understand if there is more movement towards one area rather than another.

4.1 Dataset

The empirical analysis has been applied to the closing prices of three cryptocurrencies, all related to the US dollar (USD), that are:

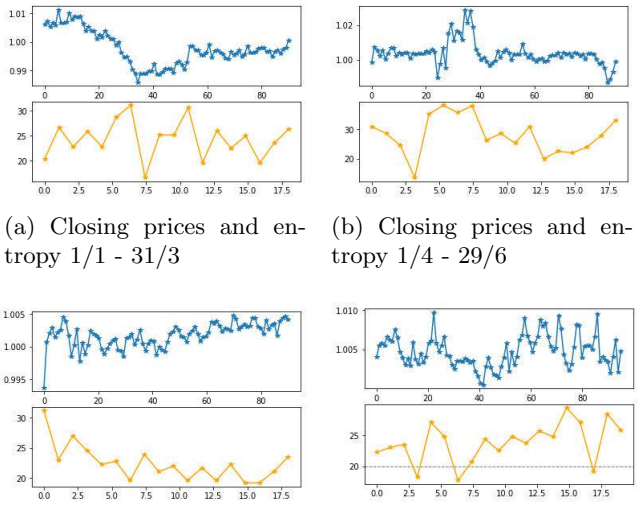
- Tether, whose price with 4 decimal places requires a step equal to 0.0001;
- Bitcoin Cash, whose price with 2 decimal places requires a step equal to 0.01;
- Litecoin, whose price with 3 decimal places requires a step equal to 0.001.

Prices are considered with a daily time frame over 1 year, from 1/1/2019 to 31/12/2019 and they are clustered in 5 days. To make the figures more clear, the 1-year interval has been divided into 4 trimesters. Furthermore, to better test the idea, the same test was carried out also on daily prices at 1 minute of 1/4/2020 recorded from 10:56 to 11:52, instead of clustered in 5 minutes. The difference from the daily case is that these prices were collected, always from the same source, but observed on different currency markets.

Source: Investing.com

4.2 Numerical examples

We can start the analysis from the annual case. The first cryptocurrency analyzed is Tether (USDT/USD), whose price moves in a neightborhood of 1 and consists of 4 decimal places; distinguish the trend of entropy compared to prices in the 4 ranges previously defined.



(a) Closing prices and entropy 1/1 - 31/3

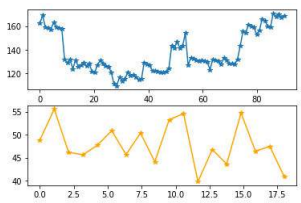
(b) Closing prices and entropy 1/4 - 29/6

(c) Closing prices and entropy 30/6 - 27/9

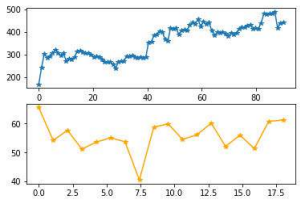
(d) Closing prices and entropy 28/9 - 31/12

Fig. 1: Prices (blue) and entropy (orange) Tether in the period 1/1 - 31/12

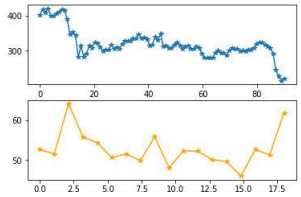
As can be seen graphically, when entropy reaches a point of relative minimum falling below a certain threshold (it therefore undergoes a sharp reduction) it is forced in the next cluster to grow, almost as if to rebalance itself. In terms of prices, this implies that in the cluster in which the entropy descent occurred there was a very small gap and, in the subsequent cluster, since entropy increases the gap also increases. In this case, the range of variation of prices is very "narrow" and every movement is important. It is possible, however, to notice for example looking at the figure 1 (d) what is the gap value and therefore the entropy threshold that, if "under"-passed, will cause an immediate growth in the next future.
The next cryptocurrency analyzed is Bitcoin Cash (BCH/USD):



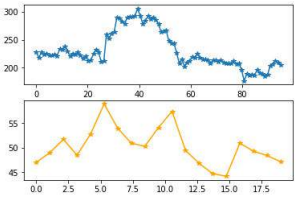
(a) Closing prices and entropy 1/1 - 31/3



(b) Closing prices and entropy 1/4 - 29/6



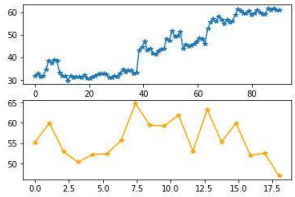
(c) Closing prices and entropy 30/6 - 27/9



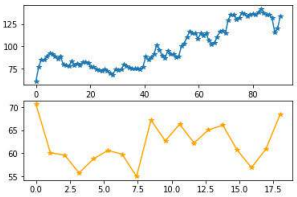
(d) Closing prices and entropy 28/9 - 31/12

Fig. 2: Prices (blue) and entropy (orange) Bitcoin Cash in the period 1/1 - 31/12

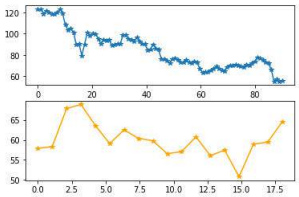
In this case the figure 2 (d) shows how the gap threshold below which a sharp drop in entropy occurs can also be quite high (especially in currencies where high volatility allows it to to move many points from one price to another). The last cryptocurrency we have considered is Litecoin (LTC/USD):



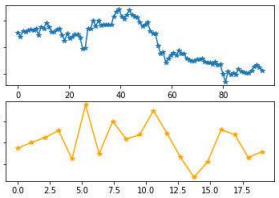
(a) Closing prices and entropy 1/1 - 31/3



(b) Closing prices and entropy 1/4 - 29/6



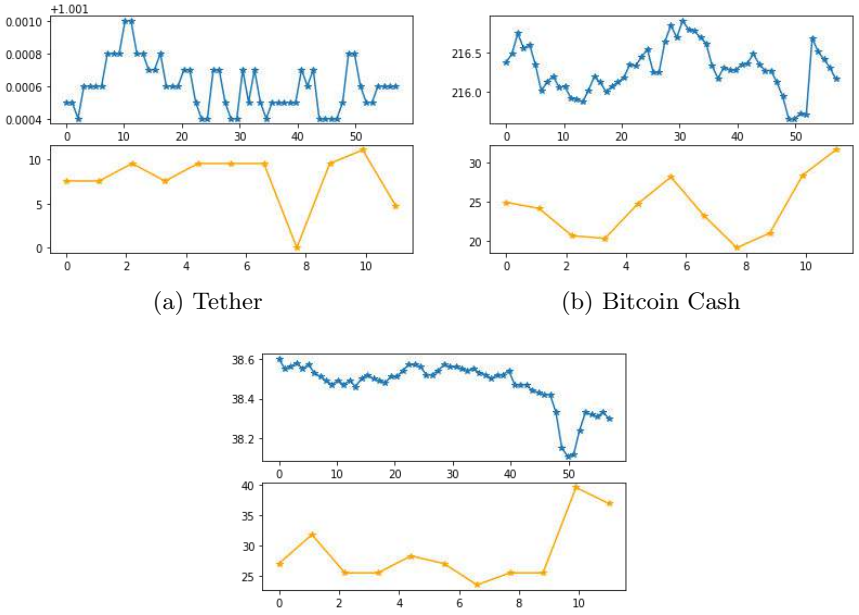
(c) Closing prices and entropy 30/6 - 27/9



(d) Closing prices and entropy 28/9 - 31/12

Fig. 3: Prices (blue) and entropy (orange) Litecoin in the period 1/1 - 31/12

Also in this cryptocurrency all the situations defined above occur, in particular from the figure 3(d) it can be seen how, following the fact that the first 4 clusters are growing despite the gap value being quite low, the gap threshold to define the drastic descent of entropy is quite low. As for the case of 1-minute prices, we can summarize the trend of the different cryptocurrencies together as shown in figure 4 which shows how all the assumptions made in the previous case are also respected for prices of this type



(a) Tether

(b) Bitcoin Cash

(c) Litecoin

Fig. 4: Prices (blue) and entropy (orange) of cryptocurrencies based on 1 minute

In particular, the hypotheses made previously are very evident in the case of the Tether (figure 4 (a)).

4.3 Possible effects on prediction

Thanks to previous results we can use entropy as an indicator to make predictions on the price trend of cryptocurrencies in the currency markets. For example, we can suppose that we are in a certain cluster X where entropy has declined sharply. As previously defined, we expect entropy to grow in the next cluster and this leads to an increase in the price gap. The hypothesis we can make is that the value of the gap in the cluster $X+1$ is at least one unit higher than the value in the cluster X: we can use this information to understand what the future price range will be. In this case, knowing the value of the gap in the cluster X, we can create a bifurcation that represents the possible evolution of the price in the event of a bullish or bearish trend. Assuming, moreover, that the first cryptocurrency price close enough to the last price of the previous cluster what we can expect is such a situation: if the second closing price of the cluster $X+1$ is **higher** than the previous price in the same cluster and assuming an upward trend we can assume that the series of prices continues in an area that we have defined as Gap^{-}; while if the second closing price of the cluster $X+1$ is **lower** than the previous price in the same cluster and assuming a bearish trend we can assume that the price series continues in an area that we have defined as Gap^{+}. Such information can be fundamental

for example for an investor who intends to choose the ideal moment to enter (or exit) the market or balance any price limits.

5 Conclusions

In this paper, we have defined a similarity between a thermodynamic system and a currency system. Thanks to this assumption, we have shown how it's possible to apply Boltzmann's entropy to cryptocurrencies. This system is characterized by the presence of N subjects interested in buying (or selling) this type of currency. Assuming that the quantity of money at a certain moment t is fixed and determinable, it is possible to hypothesize that the position of each economic entity is summarized by the last price of the cryptocurrency itself in the currency markets, as an indicator characterized by the ability to buy and sell. With this hypothesis, it was possible to determine the entropy using the Boltzmann formula, dividing the time interval into clusters and calculating the gap between the different prices. This analysis has shown that when entropy falls sharply then it must necessarily grow shortly; which in terms of price corresponds to a situation in which the gap between maximum and minimum is wider than the previous one.

References

1. Antonopoulos A. A. *Mastering Bitcoin: Unlocking Digital Cryptocurrencies*. O'Reilly Media, 2014.
2. Dionisio A., Menezes R, and Menezes D. A. An econophysics approach to analyse uncertainty in financial markets: an application to the portuguese stock market. *The European Physical Journal*, 2006.
3. Rényi A. Probability theory. *North-Holland Series in Applied Mathematics and Mechanics*, 1970.
4. Brissaud J. B. The meanings of entropy. *Entropy*, 2005.
5. BaFin. Merkblatt: Hinweise zum tatbenstand des kryptoverwahrgeschäfts. 2020.
6. Philippatos G. C. and Wilson C. J. Entropy, market risk and the selection of efficient portfolios. *Applied Economics*, 1972.
7. Tsallis C. Possible generalization of boltzmann-gibbs statistics. *Springer-Verlag*, 1988.
8. Shannon C. E. and Weaver W. The mathematical theory of communication. *University of Illinois Press*, 1949.
9. Smith E. and Foley D. K. Is utility theory so different from thermodynamics? *SFI Preprint*, 2002.
10. Usta I. and Kantar Y. M. Mean-variance-skewness-entropy measures: a multi objective approach for portfolio selection. *Entropy*, 2011.
11. IFRS. Holdings of cryptocurrencies - agenda paper 12. *IFRS Staff Paper*, 2019.
12. Chen J., Duan K., Zhang R., Zeng L., and Wang W. An ai based super nodes selection algorithm in blockchain networks. *arXiv:1808.00216*, 2018.
13. McCauley J. Thermodynamic analogies in economics and finance: instability of markets. *MPRA*, 2004.
14. Huang K. *Statistical Mechanics*. Wiley, 2008.
15. Adler R. L., Konheim A. G., and McAndrew M. H. Topological entropy. *American Mathematical Society*, 1965.

16. Boltzmann L. Vorlesungen *ü*ber gastheorie. *Leipzig: J. A. Barth.*, 1896.
17. Gulko L. The entropy theory of stock option pricing. *Int. J. Theoretical Appl. Finance*, 1999.
18. Falcioni M. and Vulpiani A. *Meccanica Statistica Elementare*. Springer-Verlag, 2015.
19. Pincus S. M. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Science*, 1991.
20. Sergeev V. M. Rationality, property rights, and thermodynamic approach to market equilibrium. *Journal of Mathematical Sciences*, 2006.
21. Szmigiera M. Number of blockchain wallet users globally 2016-2019. *Statista*, 2020.
22. Jana P., Roy T. K., and Mazumder S. K. Multi-objective possibilistic model for portfolio selection with transaction cost. *J. Comput. Appl. Math.*, 2009.
23. Clausius R. Ueber verschiedene f*ü*r die anwendung bequeme formen der hauptgleichungen der mechanischen w*ä*rmetheorie. *Ann. Phys. Chem.*, 1865.
24. Nakamoto S. Bitcoin: A per-to-peer electronic cash system. *Bitcoin.org*, 2008.
25. Viaggiu S., Lionetto A., Bargigli L., and Longo M. Statistical ensembles for money and debt. *arXiv:1109.0891*, 2018.
26. Zakiras V., Christopoulos A. G., and Rendoumis V. L. A thermodynamic description of the time evolution of a stock market index. *European Journal of Economics, Finance and Administrative Sciences*, 2009.
27. Chen W., Wang Z., Xie H., and Yu W. Characterization of surface emg signal based on fuzzy entropy. *IEEE Transaction*, 2007.
28. Gibbs J. W. On the equilibrium of heterogeneous substances. *Am. J. Sci.*, 1878.