IntechOpen

Complexity in Biological and Physical Systems

Bifurcations, Solitons and Fractals

Edited by Ricardo López-Ruiz

COMPLEXITY IN BIOLOGICAL AND PHYSICAL SYSTEMS - BIFURCATIONS, SOLITONS AND FRACTALS

Edited by **Ricardo López-Ruiz**

Complexity in Biological and Physical Systems - Bifurcations, Solitons and Fractals
http://dx.doi.org/10.5772/intechopen.68687
Edited by Ricardo López-Ruiz

Contributors

Zabidin Salleh, Liyana Abd. Rahim, Raheem Gul, Stefan Bernhard, Carlos Eduardo Maldonado, Vladimir L. Kalashnikov, Evgeni Sorokin, Qiming Liu, Shihua Zhang, Meici Sun, Jon Luke, Georgy Omel'Yanov, Slobodan Zdravkovic, Precious Sibanda, Osman A. I. Noreldin, Sabyasachi Mondal, Eugen Anitas

Notice

First published in London, United Kingdom, 2018 by IntechOpen
eBook (PDF) Published by IntechOpen, 2019
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, The Shard, 25th floor, 32 London Bridge Street
London, SE19SG – United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

Complexity in Biological and Physical Systems - Bifurcations, Solitons and Fractals
Edited by Ricardo López-Ruiz

p. cm.

Print ISBN 978-1-78923-050-5

Online ISBN 978-1-78923-051-2

eBook (PDF) ISBN 978-1-83881-342-0

Meet the editor

Ricardo López-Ruiz, MS, PhD, works as an associate professor in the Department of Computer Science and Systems Engineering, Faculty of Science, University of Zaragoza, Spain. He also serves as an associate researcher in Complex Systems at the School of Mathematics, University of Zaragoza. He also worked as a lecturer in the University of Navarra, the Public University of Navarra, and the UNED of Calatayud. He completed his postdoc with Prof. Yves Pomeau at the École Normale Supérieure of Paris and with Prof. Gabriel Mindlin at the University of Buenos Aires. His areas of interest include statistical complexity and nonlinear models, chaotic maps and applications, multiagent systems, and econophysics.

Contents

Preface

Modeling and simulating biological and physical systems are nowadays active branches of science. The diversity and complexity of behaviors and patterns present in the natural world have their reciprocity in the life systems. Bifurcations, solitons and fractals are some of these ubiquitous structures that can be indistinctively identified in many models with the most diverse applications, from microtubules with an essential role in the maintenance and the shaping of cells, to the nano/microscale structure in disordered systems determined with small-angle scattering techniques. This book collects several works in this direction, giving an overview of some models and theories that are useful for the study and analysis of complex biological and physical systems. It can provide a good guidance for physicists with interest in biology, applied research scientists, and postgraduate students.

The first section of the book presents different biological models with a wide variety of applications. In Chapter 1, Zdravkovic presents three nonlinear mechanical models to explain the dynamics of solitons in microtubules. In Chapter 2, Liu et al. study the behavior of two delayed epidemic spreading models on scale-free networks. In Chapter 3, Salleh and Rahim investigate the existence and stability of equilibria in a nutrient-prey-predator model with intratrophic predation. In Chapter 4, Gul and Bernhard apply global sensitivity analysis to a multicompartment, lumped-parameter model of an arm artery to identify the bifurcation parameters of the arm arteries. The last chapter of this section, Chapter 5 by Maldonado, introduces the idea of biological hypercomputation and analyzes the relationship between matter, energy and information.

The second section of the book presents some physical models showing soliton and fractal behaviors. In Chapter 6, Kalashnikov and Sorokin present the concept of dissipative soliton and its full life cycle as a self-organized object. In Chapter 7, Luke discusses the use of solitons for particle models in the nonlinear Klein-Gordon equation. In Chapter 8, Omel'yanov considers the problem of propagation and interaction of solitons in the generalized KdV equation. In Chapter 9, Noreldin at al. perform a weakly nonlinear stability analysis of the flow of a nanofluid in a porous medium with stress-free boundary conditions. Finally, in the last chapter, Chapter 10, Anitas introduces the concepts of mass and surface fractals in the context of small-angle scattering techniques.

As the editor of this book, I would like to thank all the authors who have contributed to this volume as well as the reviewers for their assessment. Also, I must express my gratitude to the InTechOpen editorial staff for their invitation asking me to be the editor for the second time. With particular help from Ms. Kristina Kardum, the Publishing Process Manager (PPM), we have arrived to convert in this new InTechOpen book. Finally, at this moment where life is a nonsense time flow, I want to dedicate all this effort to the memory of my father, Ricardo López-Barasoain (1935–2015), and to my mother, Amelia Ruiz-Gastón (1935–present), from Villafranca, Navarra, Spain. Of course, the rest of my family and all my friends and advisers are not forgotten in this dedicatory final paragraph.

Ricardo López-Ruiz
University of Zaragoza, Spain

Section 1

Complexity in Biological Systems

Chapter 1

Mechanical Models of Microtubules

Slobodan Zdravković

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71181

Abstract

Microtubules are the major part of the cytoskeleton. They are involved in nuclear and cell division and serve as a network for motor proteins. The first model that describes nonlinear dynamics of microtubules was introduced in 1993. Three nonlinear models are described in this chapter. They are longitudinal U-model, representing an improved version of the first model, radial φ-model and new general model. Also, two mathematical procedures are explained. These are continuum and semi-discrete approximations. Continuum approximation yields to either kink-type or bell-type solitons, while semi-discrete one predicts localized modulated waves moving along microtubules. Some possible improvements and suggestions for future research are discussed.

Keywords: microtubules, partial and ordinary differential equations, kink solitons, breathers

1. Introduction

A cell is defined as eukaryotic if it has a membrane-bound nucleus. Such cells are generally larger and much more sophisticated than prokaryotic ones. Microtubules (MTs) are the basic components of cytoskeleton existing in eukaryotes [1]. They are long structures that spread between a nucleus and a cell membrane. MTs play an essential role in the shaping and the maintenance of cells and are involved in cell division. Also, they represent a network for motor proteins. These proteins move with a velocity of $0.1 - 2\mu m/s$ [2] carrying a certain cargo such as mitochondrion.

All eukaryotic cells produce two kinds of tubulin proteins. These are α and β tubulins, or monomers, and they spontaneously arrange head to tail forming biologically functional subunit that we call a heterodimer, or a dimer for short. When intracellular conditions favor assembly, the dimers assemble into long structures called protofilaments (PFs). Microtubules are usually formed of 13 PFs, as shown in **Figure 1**.

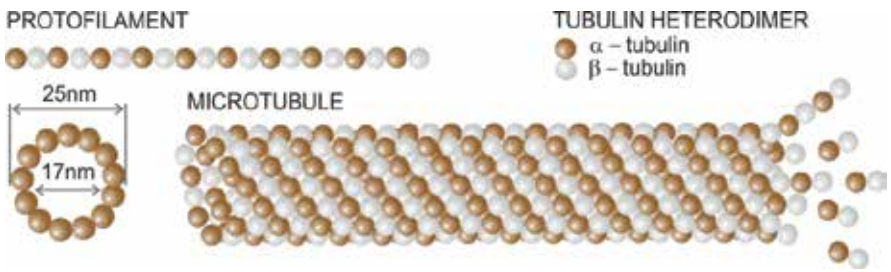

Figure 1. A tubulin dimer, a protofilament and a microtubule [3].

Hence, MTs are long cylindrical polymers whose lengths vary from a few hundred nanometers up to meters in long nerve axons [4]. Each dimer is an electric dipole whose mass and length are $m = 1.8 \times 10^{-22}$ kg and $l = 8$ nm, respectively [5]. The component of its electric dipole moment in the direction of PF is $p = 337\text{Debye} = 1.13 \times 10^{-27}$ Cm [6]. Consequently, MT as a whole appears to be a giant dipole with negatively charged end coinciding with biologically positive end (more active) and vice versa. This is the reason why an intrinsic electric field exists within MT.

MTs in non-neuronal cells are unstable structures. They exhibit dynamic instability behavior existing in phases of elongation (polymerization) or rapid shortening (depolymerization). This size fluctuation has been called as dynamic instability [7, 8]. Notice that the shrinkage rate is bigger than the growth rate (see Ref. [9] and references therein). MTs grow steadily at positive end, corresponding to the β– subunit, and then shrink rapidly by loss of tubulin dimers at the negative end, corresponding to the α–monomer. Many anticancer drugs, *for example,* taxol (paclitaxel), prevent growth and shrinkage of MTs and thus prevent cell proliferation [10].

MTs existing in neuronal cells are stable and, consequently, neurons, once formed, do not divide [4]. This stability is crucial as there are evidences that neuronal MTs are responsible for processing, storage and transduction of biological information in a brain [4, 11].

It was mentioned that MTs represent the traffic road for motor proteins. Some more information can be found in Ref. [9] and in an exhaustive review paper [12]. It suffices now to state that the cellular motors with dimensions of less than 100 nm convert chemical energy into useful work. These small machines have the fundamental role of dissipation in biological systems, which has been confirmed by both the theoretical and the experimental investigations [13]. The molecular motors dissipate continuously and operate as irreversible systems [13].

It is clear that any molecular motor, to start moving, should obtain a certain signal. One of the promising dynamical mechanisms for intracellular signaling is solitary waves, which is explained in this chapter.

2. Mechanical models

MTs, as well as all biological systems, are nonlinear in nature. Strong covalent chemical bonds are usually modeled by linear “springs”, while weak chemical interactions, existing in all biological

systems, are modeled by nonlinear "springs". This means that expressions for energy of biological systems require nonlinear terms, which brings about nonlinear partial differential equations (PDEs) explaining nonlinear dynamics of these systems. This is the topic of the present chapter. We will see that, in case of MTs, the solutions of these nonlinear PDEs are solitary waves.

The word soliton was introduced in 1965 to designate solitary waves describing the propagation of excitations in continuous media with nonlinearity and dispersion [14]. The first qualitative description of solitary waves dates back to 1834 when hydrodynamic engineer John Scott Russell observed them on a surface in a shallow channel [15]. The wave was so stable that the engineer followed it about 1 or 2 miles. From then, there has been tremendous interest for various kinds of solitons in many branches of physics [15–19]. In this chapter, the terms soliton and solitary waves are treated as synonyms, which is commonly accepted in literature.

Solitons are localized waves possessing some interesting properties. The most important is their stability in a sense that they conserve their shape and energy after mutual interaction. In other words, they can pass through one another without annihilation. This was experimentally observed in neurons [20].

To model complex MT dynamics, we should introduce some simplifications. To the best of the author's knowledge, all the models introduced so far have only one degree of freedom per dimer. Hence, for the models explained in this chapter, elementary subunits of PFs are dimers and they perform either longitudinal or angular oscillations and the appropriate models can be called as longitudinal or angular (radial), respectively.

The longitudinal contacts along PFs are much stronger than those between adjacent PFs [21, 22], which allows us to construct a simplified Hamiltonian of MT, which is, practically, Hamiltonian for a single PF only. However, the influence of the neighboring PFs is taken into consideration through the electric field. Namely, each dimer exists in the electric field coming from the dimers belonging to all PFs. Also, the nearest neighbor approximation is assumed.

3. U-model

The first model that describes nonlinear dynamics of MTs is a longitudinal one. It was introduced in 1993 by Satarić et al. [23]. According to the model, the dimers perform angular oscillations but a coordinate u, describing the dimer's displacement, is a projection of the top of the dimer on the direction of PF. Therefore, the displacements are radial but the used coordinate is longitudinal. There is a real longitudinal model assuming longitudinal displacements of the dimers that we call as Z-model [24]. Both U- and Z-models bring about equal crucial differential equations and the latter one will not be studied here.

Somewhat improved and more general version of the first nonlinear model is what we call as U-model [25] and this will be explained in the following paragraphs. Both models are based on the fact mentioned above that the dimers are electric dipoles and that the whole MT can be regarded as ferroelectric [23, 26], which means that the interaction between a single dimer and its surrounding can be modeled by W-potential [23, 27]. This yields to the following Hamiltonian for MT [23, 25, 28]

$$H_u = \sum_n \left[\frac{m}{2} \dot{u}_n{}^2 + \frac{k_u}{2}(u_{n+1} - u_n)^2 - \frac{1}{2} A u_n^2 + \frac{1}{4} B u_n^4 - QE u_n \right] \tag{1}$$

where dot means the first derivative with respect to time while the integer n determines the position of the considered dimer in PF. The first term obviously represents a kinetic energy of the dimer of mass m. The second one is interaction between the neighboring dimers belonging to the same PF in the nearest neighboring approximation and k is an intradimer stiffness parameter. The next two terms represent the W-potential energy mentioned earlier, where the parameters A and B should be determined or, at least, estimated and are assumed to be positive. We should point out that the double-well potential is rather common in physics [27, 29, 30]. The very last term is coming from the fact that the dimer is the electric dipole existing in the field of all other dimers, where $Q > 0$ represents the excess charge within the dipole and $E > 0$ is internal electric field. The last three terms together can be regarded as unsymmetrical W-potential.

Our final goal is the function $u_n(t)$, describing nonlinear dynamics of MT. This function is a solution of so-called dynamical equation of motion, which can be obtained from Eq. (1). To derive it, we introduce generalized coordinates q_n and p_n defined as $q_n = u_n$ and $p_n = m\mathrm{d}u_n/\mathrm{d}t$. Using well-known Hamilton's equations of motion $\mathrm{d}p_n/\mathrm{d}t = -\mathrm{d}H/\mathrm{d}q_\mathrm{n}$ and $\mathrm{d}q_n/\mathrm{d}t = \mathrm{d}H/\mathrm{d}p_\mathrm{n}$, we obtain the following discrete differential equation that should be solved

$$m\ddot{u}_n = k_u(u_{n+1} + u_{n-1} - 2u_n) + Au_n - B{u_n}^3 + QE - \gamma \dot{u}_n \tag{2}$$

The last term is a viscosity force with γ being a viscosity coefficient [23]. Therefore, nonlinear dynamics of MTs has been described by Eq. (2). Obviously, nonlinearity is coming from the fourth degree term in the W-potential.

It was explained earlier that we used some approximations to derive Eq. (2). However, we need one more to solve it. We now explain two mathematical methods for solving this equation. Practically, these two approaches are two approximations. They are continuum and semi-discrete approximations. We will see that the different mathematical procedures yield to different solutions. Therefore, the function $u_n(t)$ depends not only on the physical system but also on the used mathematical method.

Let us explain the continuum approximation first. A question if MTs are discrete or continuum systems was studied in Ref. [31], where it was shown that the continuum approximation is valid. The continuum approximation means a transition $u_n(t) \to u(x,t)$, which allows a series expansion of the terms $u_{n\pm1}$, that is, $u_{n\pm1} \to u \pm \frac{\partial u}{\partial x} l + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} l^2$, where l is the dimer's length explained earlier. In fact, PF can be seen as one-dimensional crystal with l being a period of the lattice. This straightforwardly brings about the following continuum dynamical equation of motion

$$m\frac{\partial^2 u}{\partial t^2} - k_u l^2 \frac{\partial^2 u}{\partial x^2} - QE - Au + Bu^3 + \gamma \frac{\partial u}{\partial t} = 0 \tag{3}$$

This is PDE that cannot be easily solved. Hopefully, this equation can be transformed into an ordinary differential equation (ODE). It is well known that, for a given wave equation, a

traveling wave $u(\xi)$ is a solution which depends upon x and t only through a unified variable ξ as $\xi = \kappa x - \omega t$, where κ and ω are constants. If we substitute the variables x and t by ξ we straightforwardly transform Eq. (3) into the following ODE

$$\alpha_u \Psi'' - \rho_u \Psi' - \Psi + \Psi^3 - \sigma = 0 \tag{4}$$

where $u' \equiv \mathrm{d}u/\mathrm{d}\xi$ and

$$u = \sqrt{\frac{A}{B}}\Psi, \quad \alpha_u = \frac{m\omega^2 - k_u l^2 \kappa^2}{A}, \quad \sigma = \frac{qE}{A\sqrt{A/B}}, \quad \rho_u = \frac{\gamma\omega}{A} \tag{5}$$

Eq. (4) becomes the appropriate one in Ref. [23] for $\alpha_u = -1$. Therefore, the U-model is more general than its predecessor introduced in Ref. [23]. It is crucial that the parameter α_u can be determined together with the function Ψ for known or estimated σ and ρ_u. This is because α_u has very important physical meaning. The first term in Eq. (3) is the inertial term and it is coming from the kinetic energy in Hamiltonian (1), while the second one is the elastic one. Therefore, positive α_u means that the inertial term is bigger than the elastic one and vice versa.

Eq. (4) has already been solved using different mathematical procedures like standard procedure [23, 27, 29, 30] and method of factorization [31, 32]. There exists a group of procedures where the function Ψ is represented as a serious expansion over other known function like $\Psi = \sum_{k=0}^{N} A_k \Phi^k$. The function Φ is usually known and we plug Ψ into Eq. (4) and determine the coefficients A_k. A common example for Φ is a solution of Riccati equation, which is either tangent or tangent hyperbolic. As only the latter function may have physical meaning, we call the method as tangent hyperbolic function method (THFM) [25, 33–35] and extended or modified extended THFM [36]. The function Φ can also be one of Jacobian elliptic functions [37] and, even, unknown [38].

It is very likely that the most general procedure is the simplest equation method (SEM) [39–41] and its simplified version called as modified simplest equation method (MSEM) [42]. According to SEM, the series expansion is [39–41].

$$\Psi = A_0 + \sum_{k=1}^{N} \left(A_k \Phi^k + B_k \left(\frac{\Phi'}{\Phi} \right)^k \right) \tag{6}$$

where A_0, A_k and B_k are coefficients that should be determined and Φ' represents the first derivative. In general, the function $\Phi = \Phi(\xi)$ is known and represents a solution of a certain ODE of lower order than the equation that should be solved. A commonly used example is the Riccati equation [40]

$$\Phi' + \Phi^2 - 2a\Phi - b = 0, \qquad a, b = \mathrm{const} \tag{7}$$

To determine the positive integer N in Eq. (6), we should plug $\Psi = c/\xi^p$, $c = \mathrm{const}$, into Eq. (4) and concentrate our attention on the leading terms [42]. One can easily show that $N = 1$ for Eq. (4) as the leading terms are proportional to $\xi^{-(p+2)}$ and ξ^{-3p}.

The well-known general solution of Eq. (7) is [39, 40]

$$\Phi = a + \sqrt{a^2 + b}\tanh\left[\sqrt{a^2 + b}(\xi - \xi_0)\right] \tag{8}$$

In what follows, we assume $\xi_0 = 0$.

Our next step is determination of the parameters A_0, A_1, B_1, a, b and α_u. According to Eqs. (6) for $N = 1$ and (7), we obtain the expressions for Ψ', Ψ'' and Ψ^3 as required by Eq. (4), which yields to the following expression:

$$K_3\Phi^3 + K_3{}'\Phi^{-3} + K_2\Phi^2 + K_2{}'\Phi^{-2} + K_1\Phi + K_1{}'\Phi^{-1} + K_0 = 0 \tag{9}$$

Obviously, this is satisfied if all the coefficients are simultaneously equal to zero. This brings about a system of seven equations, which can be obtained using Mathematica or similar software [39]. One of them can be written as

$$K_3 \equiv (A_1 - B_1) \times \left[2\alpha_u + (A_1 - B_1)^2\right] = 0 \tag{10}$$

indicating two possible relationships between the parameters A_1 and B_1. Hence, there are a few cases to be studied. They are as follows [39]: **(1)** $B_1 = 0$, $a = 0$; **(2)** $B_1 = 0$, $a \neq 0$; **(3)** $A_1 = B_1$; **(4)** $2\alpha_u = -(A_1 - B_1)^2$, $A_1B_1 \neq 0$; **(5)** $A_1 = 0$, $a \neq 0$ and **(6)** $A_1 = 0$, $a = 0$.

It is obvious that the first case represents nothing but a simpler method called extended tanh-function method. The system mentioned earlier brings about [39]

$$8{A_0}^3 - 2A_0 + \sigma = 0, \qquad \alpha_u = -\frac{{A_1}^2}{2}, \qquad A_1 = -\frac{\rho_u}{3A_0}, \qquad b = \frac{1 - 3{A_0}^2}{{A_1}^2} \tag{11}$$

The final result is [39]

$$\Psi_i = A_{0i} + A_{1i}\Phi_i, \qquad \Phi_i = \sqrt{b_i}\tanh\left(\sqrt{b_i}\,\xi\right) \tag{12}$$

where A_{0i} is the following three real solutions of the first of Eqs. (11)

$$A_{01} = \frac{1}{2\sqrt{3}}\left(\cos F + \sqrt{3}\sin F\right), \quad A_{02} = \frac{1}{2\sqrt{3}}\left(\cos F - \sqrt{3}\sin F\right) \tag{13}$$

$$A_{03} = -\frac{1}{\sqrt{3}}\cos F, \quad F = \frac{1}{3}\arccos\left(\frac{\sigma}{\sigma_0}\right), \quad \sigma_0 = \frac{2}{3\sqrt{3}} \tag{14}$$

Of course, these three solutions exist for $\sigma < \sigma_0$. The case $\sigma > \sigma_0$ was discussed in Ref. [25].

All the three solutions are shown in **Figure 2** for $\sigma \approx 0.9\sigma_0$ and $\rho_u = 1$. Of course, these solutions reproduce previously known results [25]. **Figure 2** shows that the solutions of Eq. (4) are kink and antikink solitons. More detailed analysis of their physical meaning is given in Ref. [25].

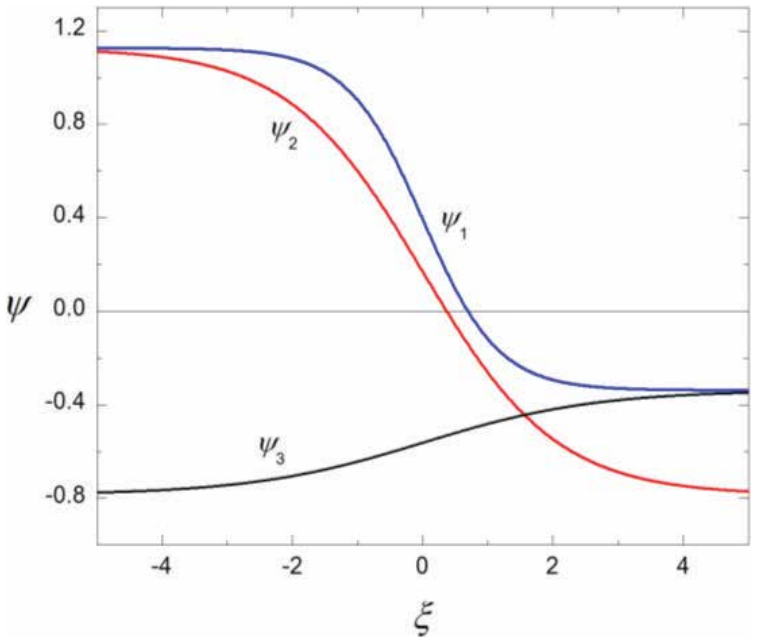

Figure 2. The functions $\Psi(\xi)$ for $\rho_u = 1$ and $\sigma = 0.34$.

It was shown [42] that the second case is equal to the first one indicating that the value of a is irrelevant if $B_1 = 0$. In other words, we could have assumed a simpler version of the Riccati equation neglecting the term $2a\Phi$.

The third case is more interesting. It turns out that, instead of the three lines in **Figure 2**, that is, the three solutions, we obtain infinitely many lines corresponding to each of them [39]. However, they represent three groups of parallel lines, which means that all these solutions are only shifted functions and, consequently, have equal physical meaning. Therefore, this case does not bring about any physically new result.

Case 4 is suggested by Eq. (10). The system of seven equations, mentioned earlier, gives the first and the last term in Eq. (11) as well as

$$a = 0, \qquad A_1 = 2B_1, \qquad \alpha_u = -\frac{{B_1}^2}{2}, \qquad B_1 = -\frac{\rho_u}{3A_0} \tag{15}$$

The final expression for Ψ is

$$\Psi(\xi) = A_0 - \sqrt{1 - 3{A_0}^2}\left(\tanh y + \frac{1}{\sinh y}\right), \qquad y = \frac{3A_0}{2\rho_u}\sqrt{1 - 3{A_0}^2}\,\xi \tag{16}$$

This case yields to a new solution, which was not obtained using less general mathematical methods. However, it may be interesting from mathematical point of view only as Ψ diverges for $\xi = 0$.

Case 5 is a simplified version of SEM, explained in Ref. [42]. The mentioned system brings about $\rho_u = 0$ as well as

$$B_1 = -\frac{\mathrm{A}_0}{a}, \quad \alpha_u = -\frac{{B_1}^2}{2}, \quad {\mathrm{A}_0}^3 - \mathrm{A}_0 - \sigma = 0, \quad b = \frac{a^2\left({\mathrm{A}_0}^2 - 1\right)}{2{\mathrm{A}_0}^2} \tag{17}$$

where a notation A_0 has been introduced to distinguish this parameter from A_0, used in the previous cases. It is interesting to compare the polynomials for A_0 and A_0, existing in Eqs. (11) and (17). We can see that.

$$\mathrm{A}_{0i} = -2A_{0i}, \qquad i = 1,\ 2,\ 3 \tag{18}$$

which means that the values for A_{0i} are given by Eqs. (13), (14) and (18). We can easily show that the final solution for Ψ is [39]

$$\Psi = \mathrm{A}_0 - \frac{\mathrm{A}_0 K^2}{\cosh^2(aK\xi)\ \left[1 + K\tanh(aK\xi)\right]}, \qquad K = \frac{\sqrt{3{\mathrm{A}_0}^2 - 1}}{\mathrm{A}_0\sqrt{2}} \tag{19}$$

Obviously, this function cannot diverge for any value of $a\xi$ but only for $-1 < K < 1$. Also, K should be real and these two requirements eliminate Ψ_2 and Ψ_3 [39], which means that Ψ and A_0 in Eq. (19) are Ψ_1 and A_{01}.

The function $\Psi_1(\xi)$ is shown in **Figure 3** for $a = 0.1$ and for two values of the parameter σ. We notice very interesting result that is a bell-type soliton! This certainly demonstrates the advantage of SEM method over the less general ones.

It is important to study the physical meanings of the parameters a and σ. Eq. (19) indicates that solitonic width is inversely proportional to $|a|$ and that a does not affect maximum of the wave. **Figure 3** shows that the amplitude of Ψ_1 is a decreasing function with respect to σ.

Finally, the last case gives the solution

$$\Psi = \pm\frac{\sqrt{2}}{\sin\left(2\sqrt{-b}\xi\right)}, \qquad b < 0 \tag{20}$$

which is obviously divergent for $\sqrt{-b}\xi = k\pi$, $k = 0,\ \pm 1,\ \pm 2, \ldots$.

Therefore, all the cases are explained and we can see that the continuum approximation yields to both kink solitons and bell-type solitons. The latter may exist only if viscosity is neglected.

It was stated earlier that the coordinate u was the projection of the top of the dimer on the direction of MT. A patient reader may ask how u can be negative when this is the projection. This question is answered in Ref. [28].

It was mentioned earlier that there are two approximations that can be used to solve Eq. (2). Now we get back to Eq. (2) and study semi-discrete one [15, 28, 43]. A mathematical basis for

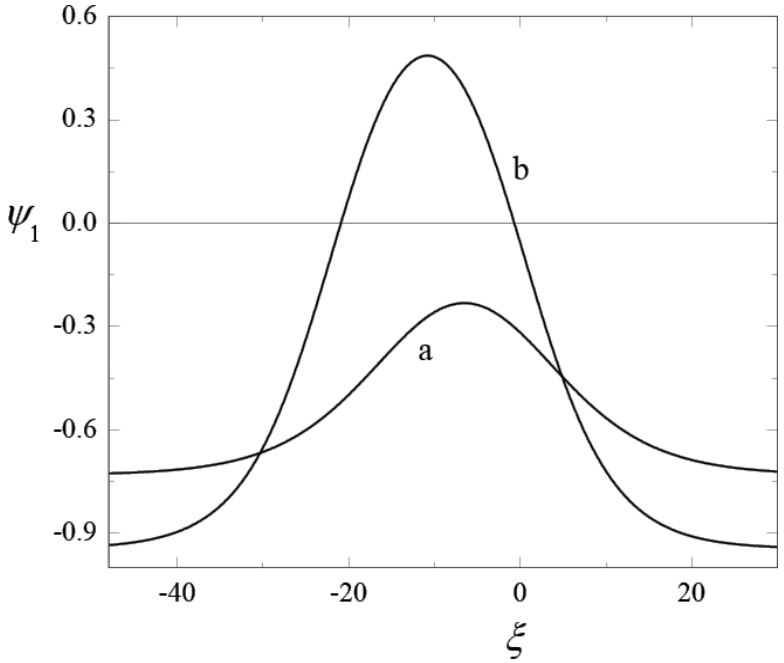

Figure 3. A bell-type soliton for a = 0.1 and σ = 0.34 (a) and σ = 0.1 (b).

the method is a multiple-scale method or a derivative-expansion method [16, 44]. We assume small oscillations

$$u_n(t) \;= \varepsilon\, \Phi_n(t), \qquad \varepsilon << 1 \tag{21}$$

which straightforwardly transforms Eq. (2) into

$$\varepsilon\, m\ddot{\Phi}_n = \varepsilon\, k(\Phi_{n+1} + \Phi_{n-1} - 2\Phi_n) + \varepsilon\, A\Phi_n - \varepsilon^3 B{\Phi_n}^3 + qE + \mathrm{O}\left(\varepsilon^4\right) \tag{22}$$

According to the semi-discrete approximation, we look for wave solution which is a modulated wave, that is [28, 45]

$$\Phi_n(t) = F(\xi)e^{i\theta_n} + \varepsilon\;\; F_0(\xi) + \varepsilon\, F_2(\xi)e^{i2\,\theta_n} + \mathrm{cc} + \mathrm{O}\left(\varepsilon^2\right) \tag{23}$$

$$\xi = (\varepsilon\, nl, \varepsilon\, t), \qquad \theta_n = nql - \omega t \tag{24}$$

where ω is the optical frequency of the linear approximation, $q = 2\pi/\lambda > 0$ is the wave number, cc represents complex conjugate terms and the function F_0 is real. Of course, l is the dimer's length, as mentioned earlier. The function F is continuous and represents an envelope, while $\exp(i\theta_n)$, including discreteness, is a carrier component. Notice that the parameter ε exists in

the function F, but does not in $\exp(i\theta_n)$. This is because the frequency of the carrier wave is much higher than the frequency of the envelope and we need two time scales, t and εt, for those two functions. The same holds for the coordinate scales.

To simplify the problem, a continuum limit $nl \to z$ should be introduced as well as new transformations $Z = \varepsilon z$ and $T = \varepsilon t$. This allows a series expansion of $F(\xi)$, that is

$$F(\varepsilon(n \pm 1)l, \varepsilon t) \to F(Z,T) \pm F_Z(Z,T)\varepsilon l + \frac{1}{2}F_{ZZ}(Z,T)\,\varepsilon^2 l^2 \tag{25}$$

where indexes Z and ZZ denote the first and the second derivative with respect to Z. Hence, the function $\Phi_n(t)$ becomes

$$\Phi_n(t) = F e^{i\theta} + F^* e^{-i\theta} + \varepsilon F_0 + F_2 e^{i2\theta} + {F_2}^* e^{-i2\theta} \tag{26}$$

where * stands for complex conjugate and $F \equiv F(Z,T)$. All this allows us to obtain the expressions existing in Eq. (22), such as $\Phi_{n+1} + \Phi_{n-1} - 2\Phi_n$, Φ_n and ${\Phi_n}^3$, and Eq. (22) becomes [28]

$$\begin{aligned}
&\left(\varepsilon^3 F_{TT} - 2i\varepsilon^2\omega F_T - \varepsilon\omega^2 F\right)e^{i\theta} - \left(4i\varepsilon^3\omega F_{2T} + 4\varepsilon^2\omega^2 F_2\right)e^{i2\theta} + \text{cc} = \\
&= \varepsilon\frac{k}{m}\left\{2F\left[\cos(ql) - 1\right] + 2i\varepsilon l F_Z \sin(ql) + \varepsilon^2 l^2 F_{ZZ}\cos(ql)\right\}e^{i\theta} \\
&+\varepsilon\frac{k}{m}\left\{2\varepsilon F_2\left[\cos(2ql) - 1\right] + 2i\varepsilon^2 l F_{2Z}\sin(2ql)\right\}e^{i2\theta} + \frac{C}{m} \\
&+\varepsilon\frac{A}{m}\left(F e^{i\theta} + \varepsilon F_0 + \varepsilon F_2 e^{i2\theta}\right) - \varepsilon^3\frac{B}{m}\left(3|F|^2 F e^{i\theta} + F^3 e^{i3\theta}\right) + \text{cc} + \mathrm{O}\left(\varepsilon^4\right)
\end{aligned} \tag{27}$$

This crucial expression represents a starting point for a series of important expressions. They can be obtained equating the coefficients for the various harmonics. For example, equating the coefficients for $e^{i\theta}$ and neglecting all the terms with ε^2 and ε^3 one obtains the following expressions for the dispersion relation $\omega = \omega(q)$ and the group velocity $\mathrm{d}\omega/\mathrm{d}q$:

$$\omega^2 = \frac{2k_u}{m}\left[1 - \cos(ql)\right] - \frac{A}{m}, \qquad V_g = \frac{l k_u}{m\omega}\sin(ql) \tag{28}$$

Also, the coefficients for $e^{i0} = 1$ and $e^{i2\theta}$, respectively give [28]

$$\varepsilon^2 F_0 = -\frac{C}{A}, \qquad F_2 = 0 \tag{29}$$

which yields to

$$u_n = \varepsilon F e^{i\theta_n} - \frac{C}{A} + \text{cc} \tag{30}$$

Eqs. (28) and (29) and new coordinates S and τ, defined as $S = Z - V_g T$ and $\tau = \varepsilon T$, allows us to simplify Eq. (27). An explanation for why the parameter ε exists in the time scaling but is

absent in the space scaling is given in Refs. [45, 46]. If we consider the terms for $e^{i\theta}$ again we obtain the well-known nonlinear Schrödinger equation (NLSE) for the function F

$$iF_\tau + PF_{SS} + Q\,|F|^2F = 0 \tag{31}$$

where the dispersion coefficient P and the coefficient of nonlinearity Q are

$$P = \frac{1}{2\omega}\left[\frac{k_u l^2}{m}\cos(ql) - {V_g}^2\right], \qquad Q = -\frac{3B}{2m\omega} \tag{32}$$

Even though Eq. (31) is PDE, its solution exists. This well-known solution, existing for $PQ > 0$, is [15, 47, 48]

$$F(S,\tau) = A_e\,\mathrm{sech}\left(\frac{S - u_e\tau}{L_e}\right)\exp\frac{iu_e(S - u_c\tau)}{2P}, \qquad u_e > 2u_c \tag{33}$$

where parameters u_e and u_c represent envelope and carrier component velocities, while the amplitude A_e and the soliton width L_e have the forms

$$A_e = \sqrt{\frac{{u_e}^2 - 2u_eu_c}{2PQ}}, \qquad L_e = \frac{2P}{\sqrt{{u_e}^2 - 2\,u_e\,u_c}} \tag{34}$$

It is very difficult to deal with the parameters u_e and u_c as $u_e > 2u_c$ is completely unprecise statement. However, $u_c/u_e < 0.5$ seems to be more practical. Hence, new parameters U_e and η have been introduced as $U_e = \varepsilon\,u_e$, $\eta = u_c/u_e$ and $0 \le \eta < 0.5$ [45]. Finally, we can easily obtain the expression for the longitudinal displacement of the dimer at the position n

$$u_n(t) = A_0\mathrm{sech}\left(\frac{nl - V_e t}{L}\right)\cos(\Theta nl - \Omega t) - \frac{C}{A} \equiv U_n(t) - \frac{C}{A} \tag{35}$$

where

$$A_0 \equiv 2\varepsilon A_e = 2|U_e|\sqrt{\frac{1-2\eta}{2PQ}}, \qquad L \equiv \frac{L_e}{\varepsilon} = \frac{2P}{|U_e|\sqrt{1-2\eta}} \tag{36}$$

$$V_e = V_g + U_e, \qquad \Theta = q + \frac{U_e}{2P}, \qquad \Omega = \omega + \frac{\left(\mathrm{V}_g + \eta U_e\right)U_e}{2P} \tag{37}$$

One more parameter can be eliminated using the idea of coherent mode [49]. This mode means that the envelope and the carrier wave velocities are equal. It follows from Eq. (35) that $V_e = \Omega/\Theta$, which yields to the function $U_e(\eta)$. This means that the wave $u_n(t)$ is the one phase function, preserving its shape in time.

To plot the function $u_n(t)$ or, equivalently, $U_n(t)$ we should know or estimate the values of a couple of the parameters. Of course, if 2D plot is chosen, $U_n(t)$ can be presented as either a

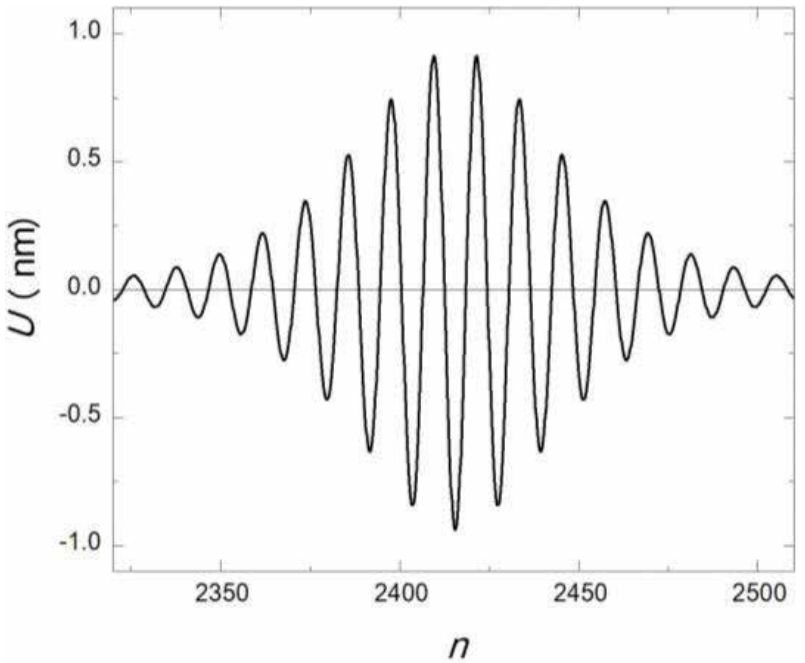

Figure 4. Function $U(n)$ for $t = 50$ns, $A = 2.9 \times 10^{-3}$N/m, $B = 1.7 \times 10^{14}$N/m^3, $k_u = 150A$, $N = 15$ and $\eta = 0.43$.

function of t at a certain position n or as a function of n for chosen t. Very detailed analysis of the parameter selection was done in Ref. [28]. One example for $q = 2\pi/Nl$ is shown in **Figure 4**. Obviously, this is a localized modulated wave usually called as breather. We can see that its width is about 200 nm, which means that it covers about 25 dimers.

As a conclusion, we can state that the two mathematical procedures bring about even three results, that is, three different solitons. These are kinks, bell-type solitons and breathers. They may be signals for the motor proteins to start moving, as explained in Introduction.

Obviously, viscosity has been neglected. This will be explained in the following section, within the φ-model.

4. φ-model

A weak point of the U-model is the last term in Eq. (1). A scalar product $-\vec{p} \cdot \vec{E} = -QdE\cos\varphi_n$ would be better choice for the potential energy, where d is the distance between the centers of positive and negative charges within the dipole. This potential indicates the angle as a coordinate instead of the projection u and the Hamiltonian for the radial model, which we call as φ-model, is [50, 51]

$$H_\varphi = \sum_n \left[\frac{I}{2} \dot{\varphi}_n{}^2 + \frac{k_\varphi}{2} \left(\varphi_{n+1} - \varphi_n \right)^2 - pE \cos \varphi_n \right] \tag{38}$$

where I is a moment of inertia of the dimer at the position n. Notice that the W-potential does not exist in Eq. (38) even though the terms including $\varphi_n{}^2$ and $\varphi_n{}^4$ appear as a result of a series expansion of the cosine function. Instead of viscosity force introduced in the previous section, we introduce viscosity momentum $M_v = -\Gamma \dot{\varphi}$, where Γ is the viscosity coefficient [51–53]. Following the procedure explained earlier, we obtain

$$\alpha_\varphi \Psi'' - \rho_\varphi \Psi' + \Psi - \Psi^3 = 0 \tag{39}$$

where

$$\varphi = \Psi \sqrt{6}, \qquad \alpha_\varphi = \frac{I\omega^2 - k_\varphi l^2 \kappa^2}{pE}, \qquad \rho_\varphi = \frac{\omega \Gamma}{pE} \tag{40}$$

Like above, the solutions are kink solitons [50].

It is very interesting to compare the expressions for α_u and α_φ, given by Eqs. (5) and (40). They can be written as

$$\alpha_u = \frac{m\kappa^2}{A} \left(v^2 - c_u{}^2 \right), \qquad c_u{}^2 = \frac{k_u l^2}{m} \tag{41}$$

and

$$\alpha_\varphi = \frac{I\kappa^2}{pE} \left(v^2 - c_\varphi{}^2 \right), \qquad c_\varphi{}^2 = \frac{k_\varphi l^2}{I} \tag{42}$$

where $v = \omega/\kappa$ is the soliton velocity, while c_u and c_φ are corresponding sound velocities. According to Eq. (11), we can see that the U-model predicts $c_u > v$ as A is positive. On the other hand, $\alpha_\varphi = 2\rho_\varphi{}^2/9 > 0$ [50] means that, according to the φ-model, the kink belongs to the class of supersonic solitons. We will return to this issue in the next section.

Now, we switch to the semi-discrete approximation within the φ-model to solve the dynamical equation of motion, which is [51]

$$\varepsilon I \ddot{\Phi}_n = \varepsilon \ k_\varphi (\Phi_{n+1} + \Phi_{n-1} - 2\Phi_n) - \varepsilon pE\Phi_n + \varepsilon^3 pE\Phi_n{}^3 + \mathrm{O}\left(\varepsilon^4\right) \tag{43}$$

where $\varphi = \varepsilon \Phi$ has been used. Of course, Eq. (43) is analog to Eq. (22). Following the procedure explained in the previous section, we straightforwardly obtain $F_0 = 0$ and $F_2(\xi) = 0$, as well as

$$\omega^2 = \omega_0^2 + \frac{4k}{I} \sin^2 (ql/2), \qquad \omega_0 = \sqrt{pE/I}, \qquad V_g = \frac{lk}{I\omega} \sin (ql) \tag{44}$$

where ω_0 is the lowest frequency of the oscillations [59]. Also, we easily obtain NLSE (31), where

$$P = \frac{1}{2\omega}\left[\frac{l^2 k}{I}\cos(ql) - {V_g}^2\right], \qquad Q = \frac{3pE}{2I\omega}. \tag{45}$$

The final solution $\varphi_n(t)$ is the same as $U_n(t)$ except that P and Q are different. Therefore, both the U- and the φ-models predict the breather waves moving through MT.

Finally, viscosity should be introduced in the semi-discrete approximation [51]. Due to viscosity momentum $M_v = -\Gamma\dot{\Phi}_n$, the final result $\varphi_n(t)$ includes the expected exponential term $e^{-\beta t}$, where $\beta = \Gamma/2I$ [51].

5. General model of MTs

It was mentioned earlier that the weak point of the U-model is the last term in Eq. (1). Also, it is better to use the radial coordinate φ than the longitudinal one as we assume angular oscillations of the dimers. The scalar product $-\vec{p}\cdot\vec{E} = -QdE\cos\varphi_n$, existing in the φ-model, solved these problems but the W-potential has been missing. In fact, a series expansion of $\cos\varphi_n$ gives ${\varphi_n}^2$ and ${\varphi_n}^4$ terms but with opposite signs from those in the U-model. These two terms are, practically, a potential that looks like W in a mirror having only one minimum surrounded by two maxima and, due to its shape, can be called as M-potential [54]. This potential brings about $\alpha_\varphi > 0$, which is disputable result.

Therefore, we want to solve the mentioned problem regarding the U-model but to keep the W-potential, the coordinate φ and, probably, $I\omega^2 < kl^2\kappa^2$, that is, $\alpha < 0$. This suggests the following Hamiltonian

$$H = \sum_n \left[\frac{I}{2}{\dot{\varphi}_n}^2 + \frac{k}{2}\left(\varphi_{n+1} - \varphi_n\right)^2 - \frac{A}{2}{\varphi_n}^2 + \frac{B}{4}{\varphi_n}^4 - pE\cos\varphi_n\right] \tag{46}$$

where $A > 0$, $B > 0$ and φ_n has the same meaning as in the φ-model. Let us call the model as general one (GM). The procedure mentioned earlier brings about

$$\left(I\omega^2 - kl^2\kappa^2\right)\varphi'' - \Gamma\omega\varphi' - (A - pE)\varphi + \left(B - \frac{pE}{6}\right)\varphi^3 = 0 \tag{47}$$

where, of course, $\varphi \equiv \varphi(\xi)$. If we consider Eqs. (3) and (47), we can see that the last two terms in Eq. (47) may be the first derivatives of either W- or M-potential, depending on the sign of the terms in the brackets. However, these brackets may have different signs or can be zero. Therefore, the possible cases are:

Case 1: $(A - pE)\left(B - \frac{pE}{6}\right) > 0$, **Case 2:** $(A - pE)\left(B - \frac{pE}{6}\right) < 0$,

Case 3: $A = pE$, $B \neq \frac{pE}{6}$, **Case 4:** $A \neq pE$, $B = pE/6$.

All of them are studied in Ref. [54] and they will be explained here briefly.

Case 1 straightforwardly yields to

$$\alpha_1\Psi'' - \rho_1\Psi' + \Psi - \Psi^3 = 0 \tag{48}$$

where

$$\varphi = \pm\sqrt{\frac{A - pE}{B - pE/6}}\Psi \equiv K\Psi, \qquad \alpha_1 = \frac{I\omega^2 - kl^2\kappa^2}{pE - A}, \qquad \rho_1 = \frac{\Gamma\omega}{pE - A} \tag{49}$$

and the final solution is [54]

$$\varphi(\xi) = \frac{K}{2}\left[1 + \tanh\left(\frac{3}{4\rho_1}\xi\right)\right], \qquad \alpha_1 > 0 \tag{50}$$

Eq. (50) holds for both positive and negative ρ_1. Therefore, $\varphi(\xi)$ represents kink soliton if $\rho_1 > 0$ and antikink one for negative ρ_1, which is shown in **Figure 5**.

One of the advantages of the GM over the φ-model is the value of amplitude. Namely, the amplitude of the kink soliton, according to the φ-model, is $\sqrt{6}$, coming from Eq. (40). This is unrealistic, too big value. Instead of $\sqrt{6}$, the appropriate factor, existing in the GM, is K, given by Eq. (49).

If viscosity is neglected, the GM brings about

$$\varphi_0 = K\tanh(\xi/a), \qquad a^2 = 2\alpha_1 \tag{51}$$

Case 2 straightforwardly yields to

$$\alpha_2\Psi'' - \rho_2\Psi' + \Psi + \Psi^3 = 0, \quad \varphi = \pm\sqrt{\frac{pE - A}{B - pE/6}}\Psi \equiv K'\Psi \tag{52}$$

and to the final results

$$\varphi_2 = i\frac{K'}{2}\left[1 + \tanh\left(\frac{3}{4\rho_2}\xi\right)\right], \quad \rho_2 \neq 0 \tag{53}$$

and

$$\varphi_{20} = K'\tan(\xi/a), \quad a^2 = -2\alpha, \quad \rho_2 = 0 \tag{54}$$

It is obvious that these results do not have physical meaning as φ_2 is complex, while φ_{20} may diverge.

Case 3 brings about

$$\alpha_3 \Psi'' - \rho_3 \Psi' + \Psi^3 = 0, \qquad \alpha_3 = \frac{I\omega^2 - kl^2\kappa^2}{B - pE/6}, \qquad \rho_3 = \frac{\Gamma\omega}{B - pE/6} \tag{55}$$

as well as $a_0 = a = \alpha_3 = \rho_3 = 0$, which certainly means that Eq. (55) does not have any solution having physical sense.

The remaining Case 4 linearizes Eq. (47) and will not be studied here.

Therefore, the GM yields to the kink solitons as the previous two models do. However, this is the radial model and the problems with both the last term in Eq. (1) and the huge amplitude in the case of the φ-model have been solved. We should study one more issue. It was mentioned earlier that the U-model predicts the subsonic kink soliton, while the φ-model predicts the supersonic wave. How about the GM? It was shown that Case 1 yields to the solutions having physical sense and that $\alpha_1 > 0$. According to Eq. (49), we easily reach the final conclusion:

a. If $A > pE$ and $B > pE/6$, then $\rho_1 < 0$ and the function $\varphi(x,t)$ is subsonic soliton, kink for the positive K in Eq. (50) and antikink otherwise.

b. If $A < pE$ and $B < pE/6$, then $\rho_1 > 0$ and the function $\varphi(x,t)$ is supersonic soliton, antikink for the positive K in Eq. (50) and kink otherwise.

All this certainly suggests the advantages of the GM with respect to the previous two.

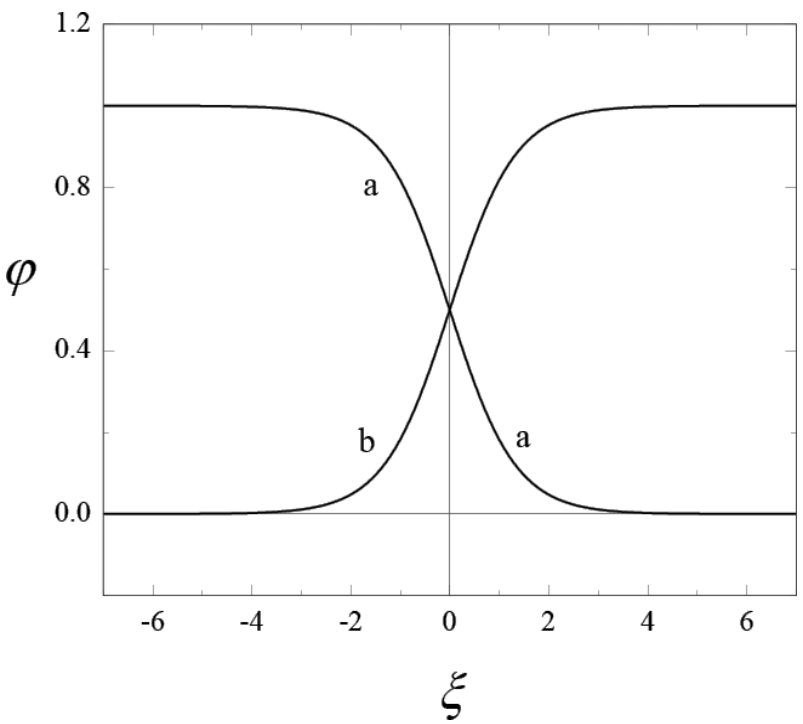

Figure 5. Kink soliton for $\rho_1 = -1$ (a) and antikink soliton for $\rho_1 = 1$ (b) for $K = 1$.

6. Conclusion and future research

In this chapter, the three models describing nonlinear dynamics of MTs are shown. The first one, the U-model, is the improved version of the first nonlinear model and it predicts subsonic kink solitons moving along MT. The second one, the radial φ-model, predicts the supersonic kinks. Finally, the GM is explained. This is the radial model which yields both possibilities regarding the kink's speed. If we assume that the kink soliton is subsonic wave then we know the minimum value of the parameter A, that is, $A > pE$, as explained earlier.

Two mathematical procedures are explained, continuum and semi-discrete approximations. It is very interesting that the final result depends not only on the physical system but on the mathematical methods as well. These solutions are the kink soliton and the breather. The question is which one, if any, really moves along MTs. This is not known in the moment and cannot be without experimental results.

It was demonstrated that the GM is better than the previous two models. However, this does not mean that it should not be improved. For example, there has been an attempt to improve the model introducing Morse potential instead of the harmonic one [55]. The harmonic potential energy assumes that attractive and repulsive forces are equal. Morse potential is not symmetric and is good for both strong and weak interactions.

In this chapter, the dimers are considered as elementary units. However, their structure is more complicated and they include tubulin tales (TTs). Consequently, nonlinear dynamics of TTs should also be studied and some results already exist [56, 57].

The W-potential has two minima which means that it assumes existence of the two angles between the dimer and the direction of PF around which the dimer oscillates. One of the future tasks should be measuring these angles. First of all, such experiment would check if the W-potential is correct or not. If it is, then our knowledge of their values would improve the theory a lot.

One of the future research goals should be two-component model. This may mean that we should construct the model assuming two degrees of freedom. However, one of these degrees can be an internal one, which means that oscillations if monomers within the dimer should be taken into consideration. Notice that the two-component model may be the one studying electro-acoustic wave excitations [58].

Finally, we should bear in mind the cytological and medical applications of the research explained in this chapter [59–61].

Acknowledgements

This work was supported by funds from Serbian Ministry of Education, Sciences and Technological Development (grant No. III45010).

Author details

Slobodan Zdravković

Address all correspondence to: szdjidji@vin.bg.ac.rs

Vinča Institute of Nuclear Sciences, University of Belgrade, Atomic Physics Laboratory, Serbia

References

[1] Dustin P. Microtubules. Berlin: Springer; 1984

[2] Cifra M, Pokorny J, Havelka D, Kučera O. Electric field generated by axial longitudinal vibration modes of microtubule. Biosystems. 2010;**100**:122

[3] Taken from Internet on the 13th of May, 2016 (https://www.google.rs/search?q=microtubule&biw=1680&bih=913&tbm=isch&tbo=u&source=univ&sa=X&sqi=2&ved=0ahUKEwi_4duGgdfMAhXlDsAKHWVACEAQsAQIJw)

[4] Hameroff S, Penrose R. Consciousness in the universe: A review of the 'Orch OR' theory. Physics of Life Reviews. 2014;**11**:39

[5] Sahu S, Ghosh S, Hirata K, Fujita D, Bandyopadhyay A. Multi-level memory-switching properties of a single brain microtubule. Applied Physics Letters. 2013;**102**:123701

[6] Havelka D, Cifra M, Kučera O, Pokorný J, Vrba J. High-frequency electric field and radiation characteristics of cellular microtubule network. Journal of Theoretical Biology. 2011;**286**:31

[7] Desai A, Mitchison TJ. Microtubule polymerization dynamics. Annual Review of Cell and Developmental Biology. 1997;**13**:83

[8] Maurer SP, Fourniol FJ, Bohner G, Moores CA, Surrey T. EBs recognize a nucleotide-dependent structural cap at growing microtubule ends. Cell. 2012;**149**:371

[9] Zdravković S. Microtubules: A network for solitary waves. Journal of the Serbian Chemical Society. 2017;**82**(5):469

[10] Chattoraj S, Bhattacharyya K. Biological oscillations: Fluorescence monitoring by confocal microscopy. Chemical Physics Letters. 2016;**660**:1

[11] Hameroff SR, Watt RC. Information processing in microtubules. Journal of Theoretical Biology. 1982;**98**:549

[12] Chowdhury D. Stochastic mechano-chemical kinetics of molecular motors: A multidisciplinary enterprise from a physicist's perspective. Physics Reports. 2013;**529**:1

[13] Lucia U. Molecular machine as chemical-thermodynamic devices. Chemical Physics Letters. 2013;**556**:242

[14] Zabusky NJ, Kruskal MD. Interaction of solitons in a collisionless plasma and the recurrence of initial states. Physical Review Letters. 1965;**15**:240

[15] Dauxois T, Peyrard M. Physics of Solitons. Cambridge, UK: Cambridge University Press; 2006

[16] Dodd RK, Eilbeck JC, Gibbon JD, Morris HC. Solitons and Nonlinear Wave Equations. London: Academic Press, Inc.; 1982

[17] Remoissenet M. Waves Called Solitons. Berlin, Heidelberg: Springer-Verlag; 1989

[18] Lakshmanan M, Rajasekar S. Nonlinear Dynamics. Berlin, Heidelberg: Springer-Verlag; 2003

[19] Scott A. Nonlinear Science Emergence and Dynamics of Coherent Structures. Moscow: Fizmatlit; 2007 (In Russian)

[20] Gonzalez-Perez A, Mosgaard LD, Budvytyte R, Villagran-Vargas E, Jackson AD, Heimburg T. Solitary electromechanical pulses in lobster neurons. Biophysical Chemistry. 2016;**216**:51

[21] Drabik P, Gusarov S, Kovalenko A. Microtubule stability studied by three-dimensional molecular theory of solvation. Biophysical Journal. 2007;**92**:394

[22] Nogales E, Whittaker M, Milligan RA, Downing KH. High-resolution model of the microtubule. Cell. 1999;**96**:79

[23] Satarić MV, Tuszyński JA, Žakula RB. Kinklike excitations as an energy-transfer mechanism in microtubules. Physical Review E. 1993;**48**:589

[24] Zdravković S, Satarić MV, Zeković S. Nonlinear dynamics of microtubules – A longitudinal model. Europhysics Letters. 2013;**102**:38002

[25] Zdravković S, Kavitha L, Satarić MV, Zeković S, Petrović J. Modified extended tanh-function method and nonlinear dynamics of microtubules. Chaos Solitons Fract. 2012;**45**:1378

[26] Satarić MV, Tuszyński JA. Relationship between the nonlinear ferroelectric and liquid crystal models for microtubules. Physical Review E. 2003;**67**:011901

[27] Collins MA, Blumen A, Currie JF, Ross J. Dynamics of domain walls in ferrodistortive materials. I. Theory, Physical Review B. 1979;**19**(7):3630

[28] Zdravković S, Zeković S, Bugay AN, Satarić MV. Localized modulated waves and longitudinal model of microtubules. Applied Mathematics and Computation. 2016;**285**:248

[29] Gordon A. Nonlinear mechanism for proton transfer in hydrogen-bonded solids. Physica B. 1987;**146**:373

[30] Gordon A. Kink dynamics in hydrogen-bounded solids. Physica B. 1988;**151**:453

[31] Zdravković S, Maluckov A, Đekić M, Kuzmanović S, Satarić MV. Are microtubules discrete or continuum systems? Applied Mathematics and Computation. 2014;**242**:353

[32] Đekić M. Employment of the method of factorization for solving problems in nonlinear dynamics of microtubules. Kragujevac Journal of Science. 2014;**36**:59

[33] El-Wakil SA, Abdou MA. New exact travelling wave solutions using modified extended tanh-function method. Chaos Solitons and Fractals. 2007;**31**:840

[34] Ali AHA. The modified extended tanh-function method for solving coupled MKdV and coupled Hirota–Satsuma coupled KdV equations. Physics Letters A. 2007;**363**:420

[35] Kavitha L, Akila N, Prabhu A, Kuzmanovska-Barandovska O, Gopi D. Exact solitary solutions of an inhomogeneous modified nonlinear Schrödinger equation with competing nonlinearities. Mathematical and Computer Modelling. 2011;**53**:1095

[36] Fan E. Extended tanh-function method and its applications to nonlinear equations. Physics Letters A. 2000;**277**:212

[37] Zeković S, Muniyappan A, Zdravković S, Kavitha L. Employment of Jacobian elliptic functions for solving problems in nonlinear dynamics of microtubules. Chinese Physics B. 2014;**23**:020504

[38] Zdravković S, Zeković S. Nonlinear dynamics of microtubules and series expansion unknown functions method. Chinese Journal of Physics. 2017;**55**:2400

[39] Zdravković S, Gligorić G. Kinks and bell-type solitons in microtubules. Chaos. 2016;**26**: 063101

[40] Kudryashov NA. Exact solitary waves of the fisher equation. Physics Letters A. 2005; **342**:99

[41] Kudryashov NA. Simplest equation method to look for exact solutions of nonlinear differential equations. Chaos Soliton and Fractals. 2005;**24**:1217

[42] Kudryashov NA, Loguinova NB. Extended simplest equation method for nonlinear differential equations. Applied Mathematics and Computation. 2008;**205**:396

[43] Remoissenet M. Low-amplitude breather and envelope solitons in quasi-one-dimensional physical models. Physical Review B. 1986;**33**:2386

[44] Kawahara T. The derivative-expansion method and nonlinear dispersive waves. Journal of the Physical Society of Japan. 1973;**35**:1537

[45] Zdravković S. Helicoidal Peyrard-bishop model of DNA dynamics. Journal of Nonlinear Mathematical Physics. 2011;**18**(Suppl. 2):463

[46] Remoissenet M, Peyrard M. Soliton dynamics in new models with parameterized periodic double-well and asymmetric substrate potentials. Physical Review B. 1984;**29**:3153

[47] Zakharov VE, Shabat AB. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. Soviet Physics JETP. 1972;**34**(1):62 http://www.jetp.ac.ru/cgibin/dn/e_034_01_0062.pdf

[48] Scott AC, Chu FYF, McLaughlin DW. The Soliton: A new concept in applied science. Proceedings of the IEEE. 1973;**61**:1443

[49] Zdravković S, Satarić MV. Single molecule unzippering experiments on DNA and Peyrard-Bishop-Dauxois model. Physical Review E. 2006;**73**:021905

[50] Zdravković S, Satarić MV, Maluckov A, Balaž A. A nonlinear model of the dynamics of radial dislocations in microtubules. Applied Mathematics and Computation. 2014; **237**:227

[51] Zdravković S, Bugay AN, Aru GF, Maluckov A. Localized modulated waves in microtubules. Chaos. 2014;**24** 023139

[52] Das T, Chakraborty S. A generalized Langevin formalism of complete DNA melting transition. Europhysics Letters. 2008;**83**:48003

[53] Tabi CB, Mohamadou A, Kofané TC. Modulated wave packets in DNA and impact of viscosity. Chinese Physics Letters. 2009;**26**:068703

[54] Zdravković S, Satarić MV, Sivčević V. General model of microtubules. To be published in Nonlinear Dynamics

[55] Zdravković S, Bugay AN, Parkhomenko AY. Application of Morse potential in nonlinear dynamics of microtubules. Nonlinear Dynamics. 2017;**90**:2841

[56] Sekulic DL, Satarić BM, Zdravković S, Bugay AN, Satarić MV. Nonlinear dynamics of C-terminal tails in cellular microtubules. Chaos. 2016;**26**:073119

[57] Sataric MV, Sekulic DL, Zdravkovic S, Ralevic NM. A biophysical model of how α–tubulin carboxy–terminal tails tune kinesin–1 processivity along microtubule. Journal of Theoretical Biology. 2017;**420**:152

[58] Bugay AN. Nonlinear waves as signals in microtubules. Nonlinear Phenomena in Complex Systems. 2015;**18**:236

[59] Rahnama M, Tuszynski JA, Bókkon I, Cifra M, Sardar P, Salari V. Emission of mitochondrial biophotons and their effect on electrical activity of membrane via microtubules. Journal of Integrative Neuroscience. 2011;**10**:65

[60] Pokorný J. Physical aspects of biological activity and cancer. AIP Advances. 2012;**2**: 011207

[61] Sataric MV, Sekulic DL, Sataric BM. Actin filaments as the fast pathways for calcium ions involved in auditory processes. Journal of Biosciences. 2015;**40**:549

Chapter 2

The Dynamics Analysis of Two Delayed Epidemic Spreading Models with Latent Period on Heterogeneous Network

Qiming Liu, Meici Sun and Shihua Zhang

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71087

Abstract

Two novel delayed epidemic spreading models with latent period on scale-free network are presented. The formula of the basic reproductive number and the analysis of dynamical behaviors for the models are presented. Meanwhile, numerical simulations are given to verify the main results.

Keywords: epidemic spreading, scale-free network, basic reproductive number, time delay, stability

1. Introduction

Following the seminal work on scale-free network, in which the probability of $p(k)$ for any node with k links to other nodes is distributed according to the power law $p(k) = Ck^{-\gamma}$ $(2 < \gamma \leq 3)$, suggested by Barabási and Albert [1], the researches of complex network have attracted more and more interests. It was found that many relevant networks, for instance, the internet, the World Wide Web (WWW), the patterns of human sexual contacts, biology network, transportation infrastructure, etc., exhibit power-law or "scale-free" degree distributions.

The dynamical behaviors of epidemic diseases have been studied for a long time. The epidemic spreading process on network is primarily dominated by two factors: one is the macroscopic topology of the underlying network, and the other is the microscopic infection scheme, which includes properties of disease, infection pattern, individual differences, infectivity of individuals, etc. The traditional epidemic dynamics is based on homogeneous network, and the infectivity rate is equally likely over all links [2]. However, the real disease transmission network exhibits scale-free properties, and the spreading of epidemic disease (e.g., computer

virus spreading, epidemic disease between human beings) on heterogeneous network, i.e., scale-free network, has been studied by many researchers [3–27].

A handful of existing works address the complex behavior of epidemic spreading using compartmental differential equations [2, 15]. Comparing with the ordinary differential equation models, more realistic models should be retarded functional differential equation models which can include some of the past states of these systems. Time delay plays an important role in the process of the epidemic spreading, for instance, the latent period of the infectious diseases or computer virus, the infection period of infective members, and the immunity period of the recovered individuals can be represented by time delays [2]. Recently, some researchers discussed the epidemic spreading model with time delays [7, 15–17]. Susceptible-Infected-Removed (*SIR*) model is a basic and important epidemic model, Zou and Wu discussed a delayed SIR model without birth and death [15], and Wang and Wang et al. discussed a delayed SIR model with birth rate and death rate [16]. However, it is suitable to divide the nodes being considered into disjoint classes of susceptible, exposed, infective, and recovered nodes in modeling disease transmission [2, 10], i.e., a susceptive node first through an incubation period (and it is said to become exposed) after infection and before becoming infectious. For example, the latent period of epidemic cholera is about 1–3 days, hepatitis B virus 100 days, measles 10–11 days, chincough 7–10 days, diphtheria 2–4 days, scarlatina 2–5 days, poliomyelitis 7–14 days, and so on [25]; the resulting model is Susceptible-Exposed-Infected-Removed (*SEIR*) model. In addition, some diseases confer temporary immunity, and the recovered nodes cycle back into the susceptive class after an immune period; the resulting model is Susceptible-Exposed-Infected-Removed-susceptible (*SEIRS*) model.

In this paper, we will present a suitable *SEIR* model with time delay and a suitable *SEIRS* model with time delay on heterogeneous network by using functional differential equation to investigate the dynamical behaviors of epidemic spreading.

The rest of this paper is organized as follows: In Section 2, the *SEIRS* model with time delay on scale-free network is discussed. The *SEIRS* model with time delay on scale-free network is discussed in Section 3. Finally, the main conclusions of this work are summarized in Section 4.

2. Analysis of the *SEIR* model with time delay

2.1. The *SEIR* model

Suppose that the size of the network is a constant N and the degree of each node is time invariant during the period of epidemic spreading, $p(k)$ denotes the degree distribution of the network. We classify all the nodes in the network into n groups such that the nodes in the same group have the same degree. That is, each node in the kth group has the same connectivity ($k = m, m+1, \cdots, n$). Let $S_k(t), E_k(t), I_k(t)$ and $R_k(t)$ be the relative density of susceptible nodes, exposed nodes, infected nodes, and recovered nodes of connectivity k at time t, respectively, where $k = m, m+1, \cdots, n$ (m and n are the minimum and maximum degree in network topology) and n is related to the network age, measured as the number of nodes N [3]:

$$n = mN^{1/(\gamma-1)}. \tag{1}$$

Let τ be the latent period of the disease, i.e., each exposed node becomes an infected node after τ. The relative density $S_k(t)$, $E_k(t)$, $I_k(t)$ and $R_k(t)$, at the mean-field level, satisfy the following set of coupled different equations when $t>0$ [15, 16]:

$$\begin{cases} \dot{S}_k(t) &= -\lambda(k)S_k(t)\Theta(t), \\ \dot{E}_k(t) &= \lambda(k)S_k(t)\Theta(t) - \lambda(k)S_k(t-\tau)\Theta(t-\tau), \\ \dot{I}_k(t) &= \lambda(k)S_k(t-\tau)\Theta(t-\tau) - \mu I_k(t), \\ \dot{R}_k(t) &= \mu I_k(t) \end{cases} \tag{2}$$

with the normalization condition

$$S_k(t) + E_k(t) + I_k(t) + R_k(t) = 1 \tag{3}$$

holds due to the fact that the number of total nodes with degree k is a constant $p(k)N$ during the period of epidemic spreading. Where $\lambda(k)$ is the degree-dependent infection rate such as $\lambda(k)=\lambda k$ [3] and $\lambda C(k)$ [4], μ is the recovery rate of the infected nodes; $\Theta(t)$ represents the probability that any given link points to an infected node. Assuming that the network has no degree correlations [7, 16, 22], we have

$$\Theta(t) = \frac{1}{\langle k\rangle}\sum_{k=m}^{n}\phi(k)p(k)I_k(t) \tag{4}$$

in which $\langle k\rangle = \sum_k p(k)k$ stands for the average node degree and $\phi(k)$ means the occupied edges which can transmit the disease (i.e., represents the infectivity of infected nodes) [22]; they have many different forms, such as $\varphi(k)=A$ in [5], $\varphi(k)=ak^{\alpha}/(1+bk^{\alpha})$, $0<\alpha<1$ in [7], and so on. Here, we point out that the delay τ in the model (2) in this paper is different from one in the model (2)–(4) in [15]. The incubation period τ in the model in [15] is another kind of time period, during which the infectious agents develop in the vector and the infected vector becomes infectious after that time.

Note that we obtain from the third equation of system (2) that

$$E_k(t) = \lambda(k)\int_{t-\tau}^{t} S_k(s)\Theta(s)ds \tag{5}$$

and the normalization condition becomes the following mathematical form

$$S_k(t) + \lambda(k)\int_{t-\tau}^{t} S_k(s)\Theta(s)ds + I_k(t) + R_k(t) = 1. \tag{6}$$

The initial conditions of system (2) are

$$S_k(\theta) = \varphi_k(\theta), I_k(\theta) = \Psi_k(\theta), R_k(t) = \zeta_k(\theta), \theta\in[-\tau, 0] \tag{7}$$

and satisfy $S_k(0) + \lambda(k)\int_{-\tau}^{0} S_k(s)\Theta(s)ds + I_k(0) + R_k(0) = 1$ which guarantees the normalization condition holds. And $\Phi_k = (\phi_k(\theta), \Psi_k(\theta), \zeta(\theta), k = m, m+1, \cdots, n-m+1) \in C$ are nonnegative continuous on $[-\tau, 0]$, $\phi_k(0) > 0, \Psi_k(0) > 0$, and $\zeta(\theta) = 0$ for $\theta = 0$. C denotes the Banach space $C([-\tau, 0], R^{3(n-m+1)})$ with the norm, where $|f(\theta)|_\tau = \sup_{-\tau \le \theta \le 0} |f(\theta)|$. $\|\omega\| = \left(\sum_{i=m}^{n} \left(|\Psi_i(\theta)|_\tau^2 + |\phi_i(\theta)|_\tau^2 + |\zeta_i(\theta)|_\tau^2\right)\right)^{1/2}$.

2.2. The main results for the model

In this section, we first discuss the final size relation of solutions for system (2).

It is easy to know that system (2) only has a disease-free equilibrium set

$$M_0 = \left\{\left(\widehat{S}, \widehat{E}, \widehat{I}, \widehat{R}\right) | E_k = I_k = 0, S_k + R_k = 1, R_k, S_k \ge 0, k = m, m+1, \cdots, n\right\} \tag{8}$$

in which $\widehat{S} = (S_m, S_{m+1}, \cdots, S_n)$, $\widehat{E} = (E_m, E_{m+1}, \cdots, E_n)$, $\widehat{I} = (I_m, I_{m+1}, \cdots, I_n)$, $\widehat{R} = (R_m, R_{m+1}, \cdots, R_n)$.

Supposing $f(t)$ is an arbitrary nonnegative continuous function $f(t)$, we adopt the following convention:

$$f(+\infty) = \lim_{t \to +\infty} f(t) \tag{9}$$

and we obtain from the last equation of system (2) that

$$R_k(+\infty) - R_k(0) = \mu \int_0^{+\infty} I_k(s)ds. \tag{10}$$

According to the last equation of system (2), $R_k(t)$ is increasing and bounded above by 1, and it has a limit as $t \to +\infty$. Thus, the left-hand side of (10) is finite due to boundedness of $R_k(+\infty)$, and $R_k(0)$ exits, i.e., $0 < \int_0^{+\infty} I_k(u)du < +\infty$. Since $I_k(t)$ is smooth nonnegative function, we know $I_k(+\infty) = 0$, i.e., $\lim_{t \to +\infty} I_k(t) = 0$.

Furthermore, we have from (5) that

$$0 \le E_k(t) = \lambda(k)\int_{t-\tau}^{t} S_k(s)\Theta(s)ds \le \lambda(k)\int_{t-\tau}^{t} \Theta(s)ds, \tag{11}$$

In addition, by using mean value theorem for integrals, we have

$$\int_{t-\tau}^{t} \Theta(s)ds = \Theta(\xi)\tau, (t - \tau \le \xi \le t). \tag{12}$$

We obtain from $I_k(+\infty) = 0$ that $\lim_{\xi \to +\infty} \Theta(\xi)\tau = 0$ and then $\lim_{t \to +\infty} E_k(t) = 0$. Hence, M_0 is globally attractive [27].

In addition, it follows from (2) that

$$\dot{S}_k(t) + \dot{E}_k(t) + \dot{I}_k(t) = -\mu I_k(t). \tag{13}$$

Integrating (13) from 0 to +∞, we obtain that

$$S_k(0) - S_k(+\infty) + E_k(0) - E_k(+\infty) + I_k(0) - I_k(+\infty) = \mu \int_0^{+\infty} I_k(s)ds. \tag{14}$$

Noting that $I_k(+\infty)=0$, $E_k(0)=E_k(+\infty)=0$, and $S_k(+\infty)$ exists due to existence of $\int_0^{+\infty} I_k(s)ds$, we have from (14) that

$$\int_0^{+\infty} I_k(s)ds = \frac{1}{\mu}(S_k(0) + I_k(0) - S_k(+\infty)). \tag{15}$$

Additionally, integrating the first equation of system (2) from 0 to $+\infty$, we have

$$\ln \frac{S_k(0)}{S_k(\infty)} = \frac{\lambda(k)}{\langle k\rangle} \sum_k \phi(k)p(k) \int_0^{+\infty} I_k(s)ds. \tag{16}$$

Substituting (15) into (16), we obtain that

$$\ln \frac{S_k(0)}{S_k(+\infty)} = \frac{\lambda(k)}{\langle k\rangle} \sum_k \phi(k)p(k) \frac{1}{\mu}(S_k(0) + I_k(0) - S_k(+\infty)). \tag{17}$$

Because there are only several infective nodes at the beginning of disease spreading, we take $S_k(0)\approx 1$ and obtain from (17) that

$$\ln S_k(+\infty) = \frac{\lambda(k)}{\mu\langle k\rangle} \sum_k \phi(k)p(k)(S_k(+\infty) - 1). \tag{18}$$

Consequently,

$$R_k(\infty) = 1 - S_k(+\infty). \tag{19}$$

Hence, we have the following result.

Theorem 2.1. *The equilibrium set* $M_0 = \left\{ \left(\widehat{S}, \widehat{E}, \widehat{I}, \widehat{R}\right) | E_k = I_k = 0, S_k + R_k = 1, k = 1, 2, \cdots, n \right\}$ *of system* (2) *is globally attractive, i.e.,* $\lim_{t\to+\infty} I_k(t)=0$, $\lim_{t\to\infty} E_k(t)=0$. *And* $R_k(+\infty)$, $S_k(+\infty)$ *are given by formulas* (18) *and* (19).

Note that it is impossible for every susceptible to be infected. Supposing $S_k(+\infty)=0$, we know from (16) that

$$+\infty = \frac{\lambda(k)}{\mu\langle k\rangle} \sum_k \phi(k)p(k) \tag{20}$$

Obviously, Eq. (20) does not hold, i.e., $S_k(+\infty)=0$. Similar results were obtained in the early literature [19].

Secondly, we discuss the basic reproductive number of model (2). The basic reproductive number is an important conception; it represents the average number of secondary infections

infected by an individual of infective during the whole course of disease in the case that all the members of the population are susceptible [2].

Theorem 2.2. *For system* (2)*,*

$$R_0 = \frac{\langle \lambda(k)\phi(k) \rangle}{\mu \langle k \rangle} \tag{21}$$

is the basic reproductive number for system (2).

Proof. Note that $\sum_{k=m}^{n} \varphi(k)p(k)I_k(t)$ may be considered as the force of infection [15] and $\Theta(t)$ may be considered as the average force of infection. Letting $\Theta(t)$ be an auxiliary function and computing its time derivative along the solution of (2), we get

$$\begin{aligned} \frac{d\Theta(t)}{dt} &= \frac{1}{\langle k \rangle} \sum_k \phi(k)p(k)\dot{I}_k(t) \\ &= \frac{1}{\langle k \rangle} \sum_k \phi(k)p(k)\big(\lambda(k)S_k(t-\tau)\Theta(t-\tau) - \mu I_k(t)\big) \\ &= \Theta(t-\tau)\frac{1}{\langle k \rangle} \sum_k \lambda(k)\phi(k)p(k)S_k(t-\tau) - \mu\Theta(t). \end{aligned} \tag{22}$$

We have

$$\left.\frac{d\Theta(t)}{dt}\right|_{t=0} = \Theta(-\tau)\frac{1}{\langle k \rangle} \sum_k \lambda(k)\phi(k)p(k)S_k(-\tau) - \mu\Theta(0)). \tag{23}$$

Since each exposed node becomes infected node after τ, $I_k(-\tau) = I_k(0)$. It follows that $\Theta(-\tau) = \Theta(0)$. Meanwhile, $S_k(-\tau) \approx 1$. Hence, we have from (23) that

$$\left.\frac{d\Theta(t)}{dt}\right|_{t=0} = \mu\left(\frac{1}{\mu\langle k \rangle} \sum_k \lambda(k)\phi(k)p(k) - 1\right)\Theta(0) = \mu(R_0 - 1)\Theta(0). \tag{24}$$

If $R_0 > 1$, $\left.\frac{d\Theta(t)}{dt}\right|_{t=0} > 0$, which means that $\Theta(t)$ increases at the beginning of the epidemic and there exists at least one outbreak.

Meanwhile, if $R_0 \leq 1$, we obtain from (24) that $\left.\frac{d\Theta(t)}{dt}\right|_{t=0} \leq 0$. Let $t^* = \sup\{T \geq 0 : \Theta(t)$ decreases on $[0, T]\}$. Then, it follows from the above discussion that $T \geq 0$. We will prove that $T = +\infty$. Note that we obtain from the first equation of system (2) that

$$S_k(t) = S_k(0)\mathrm{e}^{-\lambda(k)\Psi(t)} \tag{25}$$

in which $\Psi(t) = \frac{1}{\langle k \rangle}\sum_k \int_0^t \varphi(k)p(k)I_k(u)du$. Hence, it follows from Eqs. (22) and (25) that

$$\frac{d\Theta(t)}{dt} = \Theta(t-\tau)\frac{1}{\langle k \rangle} \sum_k \lambda(k)\phi(k)p(k)S_k(0)\mathrm{e}^{-\lambda(k)\Psi(t-\tau)} - \mu\Theta(t). \tag{26}$$

By way of contradiction, supposing that $T<+\infty$, then we have $\frac{d}{dt}\Theta(t^*)=0$, and there exists a $t_1\in(t^*, t^*+\tau]$ such that $\frac{d}{dt}\Theta(t_1)>0$. It follows that there is a $t_2\in[t^*, t_1)$ such that $\frac{d}{dt}\Theta(t_2)=0$ and $\Theta(t_2)<\Theta(t_1)$. Note that $\Theta(t_2-\tau)\geq\Theta(t_1-\tau)$. It follows from (25) that

$$0<\frac{d}{dt}\Theta(t_1)\leq\frac{d}{dt}\Theta(t_2)=0, \tag{27}$$

which is a contradiction. Hence, $\Theta(t)$ decreases on $[0,+\infty)$, and there is no one outbreak when $R_0\leq 1$. Hence, R_0 is the basic reproductive number for system (2).

It follows from Theorems 2.1 and 2.2 that R_0 is the basic reproductive number for system (2), which is irrelative to τ. There exists at least one outbreak for the spreading of epidemic if $R_0>1$, and there is no outbreak if $R_0\leq 1$. Whether or not there exists one outbreak for the spreading of epidemic, $\lim_{t\to+\infty} I_k(t)=0$ due to global attractivity of M_0.

Besides, if we let $\tau=0$, $\varphi(k)=k$, $\lambda(k)=\lambda k$, $\mu=1$, the model (2) reduces to the model in [9]. Furthermore, the basic reproductive number for system (2) is $R_0=(\lambda\langle k^2\rangle)/(\langle k\rangle)$, which is identical with the results that the epidemic threshold $\lambda_c=(\lambda\langle k\rangle)/(\langle k^2\rangle)$ in [9]. And, R_0 is always more than unity when N is large enough [3, 7], and it means the lack of any basic reproductive number. This result is consistent with the results in epidemic dynamics on heterogeneous network [3, 10].

2.3. Numerical simulation for the model

Now, we present numerical simulations to support the results obtained in previous sections and analyze the effect of time delay on behaviors of disease spreading.

The degree distribution of scale-free network is $p(k)=Ck^{-\gamma}$, and C satisfies $\sum_{k=1}^{n}p(k)=1$. Here, we set the maximum degree $n=100$ and the minimum degree $m=1$. Consider system (2), let $\varphi(k)=ak^{\alpha}/(1+b^{\alpha})$ in which $a=0.5, \alpha=0.75, b=0.02$ and $\lambda(k)=\lambda k$, and let $\gamma=2.5$. **Figures 1–4** show the dynamic behaviors of system (2) with the initial functions satisfying condition (7).

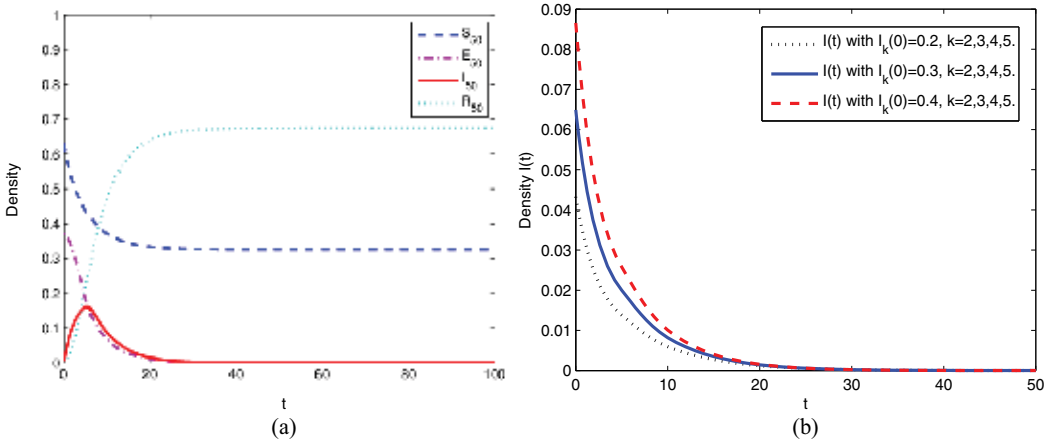

Figure 1. (a) The time evolutions for system (2) with $\lambda=0.5, \mu=0.4, \tau=5$ and $I_k(0)=0.2, k=2, 3, 4, 5, I_k(0)=0$ for the other k and $R_0=0.4006$. (b) The time evolutions for system (2) with different initial values, $\lambda=0.4, \mu=0.4, \tau=5$ and $R_0=0.4006$.

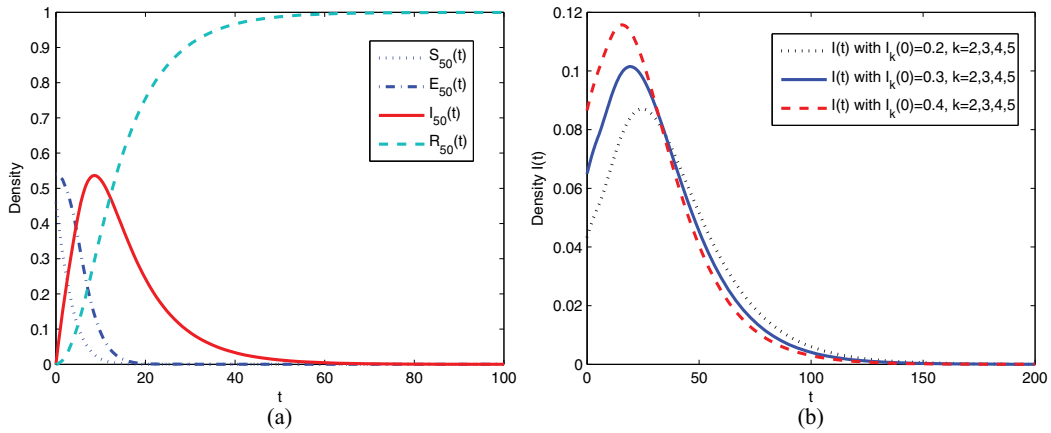

Figure 2. (a) The time evolutions for system (2) with $\lambda = 1$, $\mu = 0.1$, $\tau = 5$ and $I_k(0) = 0.2$, $k = 2, 3, 4, 5$, $I_k(0) = 0$ for the other k and $R_0 = 3.2051$. (b) The time evolutions for system (2) with different initial values, $\lambda = 1$, $\mu = 0.1$, $\tau = 5$ and $R_0 = 3.2051$.

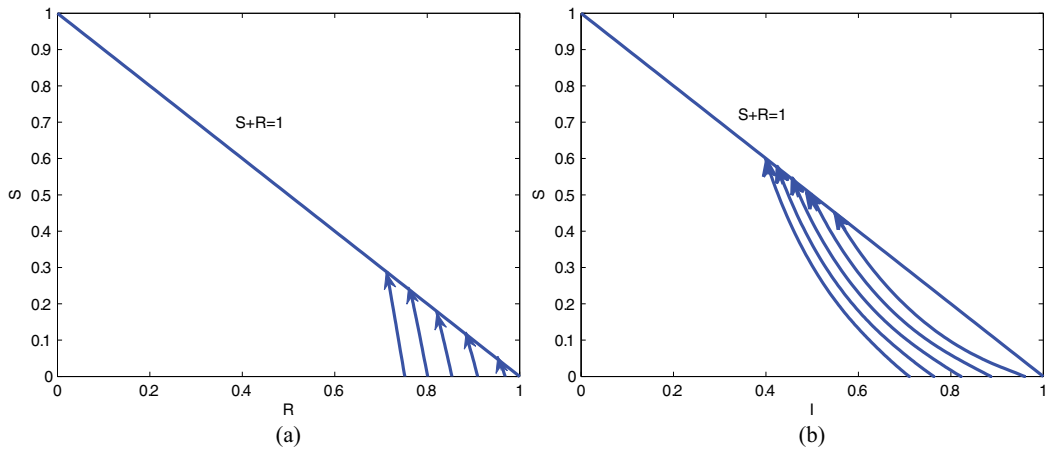

Figure 3. (a) Phase trajectories on SR-plane of system (2) with different initial values, $\lambda = 0.5$, $\mu = 0.4$, $\tau = 5$ and $R_0 = 0.4006$. (b) Phase trajectories on SR-plane of system (2) with different initial values, $\lambda = 1$, $\mu = 0.1$, $\tau = 5$ and $R_0 = 3.2051$.

Denote that

$$S(t) = \sum_{k=m}^{n} p(k)S_k(t), I(t) = \sum_{k=m}^{n} p(k)I_k(t), R(t) = \sum_{k=m}^{n} p(k)R_k(t). \tag{28}$$

They are the relative average density of susceptible nodes, exposed nodes, infected nodes, and recovered nodes at time t, respectively.

First, **Figures 1** and **2** show that the infection eventually disappears, whatever $R_0 < 1$ or not, and the outbreak of disease spreading appears when $R_0 > 1$ and the outbreak of disease spreading does not appear when $R_0 \leq 1$. Meanwhile, **Figure 3** shows that phase trajectories on SR-plane of system (2) with different initial values tend to be $S(t) + R(t) = 1$, i.e., $\sum_k p(k)(S_k(t) + R_k(t)) = 1$, which is consistent with the fact that the equilibrium M_0 is globally attractive. The numerical simulation results are identical with Theorems 2.1–2.2.

http://dx.doi.org/10.5772/intechopen.71087

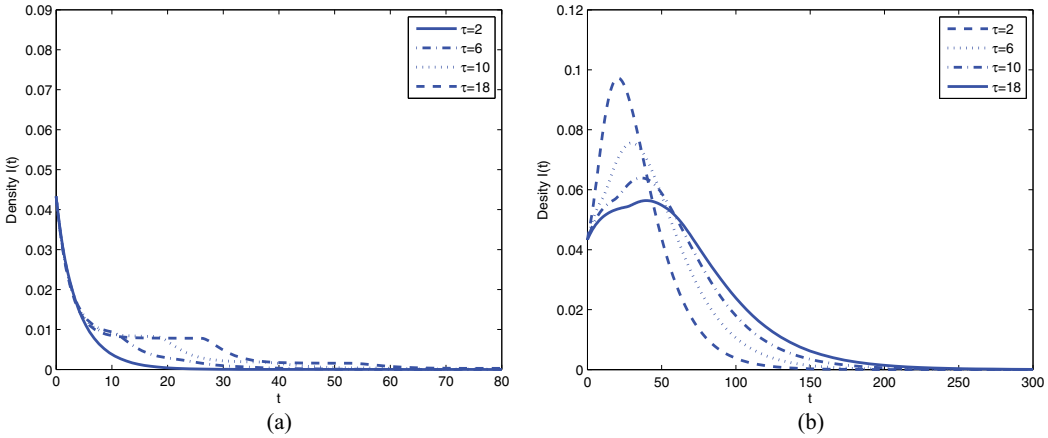

Figure 4. (a) The time evolutions of the average relative density $I(t)$ for system (2) with different τ as well as $I_k(0) = 0.2, k = 2, 3, 4, 5, I_k(0) = 0$ for the other k, $\lambda = 0.5, \mu = 0.4$ and $R_0 = 0.4006$. (b) The time evolutions of the average relative $I(t)$ for system (2) with different τ as well as $I_k(0) = 0.2, k = 2, 3, 4, 5, I_k(0) = 0$ for the other k, $\lambda = 1, \mu = 0.1$ and $R_0 = 3.2051$.

Second, time delay τ has no effects on the basic reproductive number R_0 according to (21), but it has much impact on the of process of the disease; the slower the relative density of infected nodes converges to zero, the larger τ gets, i.e., time delay may slow down the speed of disappearing the disease spreading on network. Meanwhile, time delay may effectively reduce the peak value of the relative density of infected nodes when $R_0 > 1$. Thus, the delay cannot be ignored.

At last, we know from **Figure 5** that time evolutions of the average force of infection for system (2) is consistent with time evolutions of the average relative density $I(t)$. However, there is only one outbreak, which is different from the phenomenon in [15].

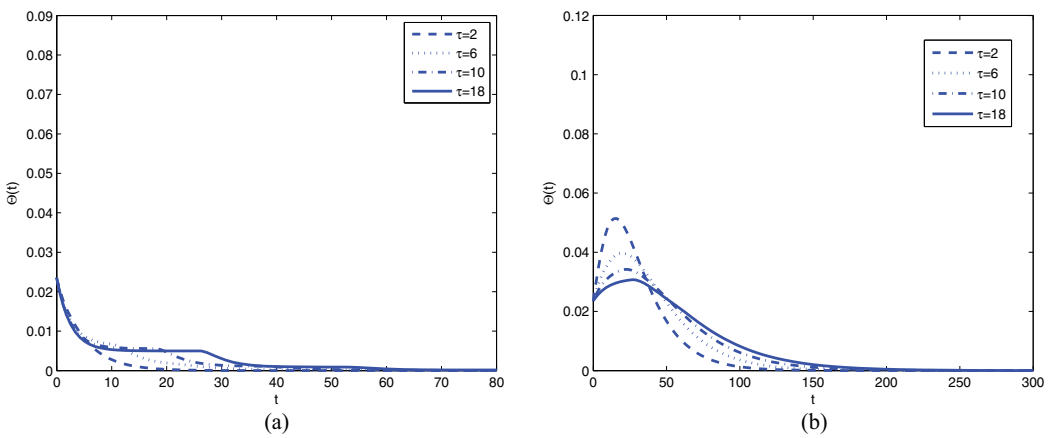

Figure 5. (a) The time evolutions of the average force of infection $\Theta(t)$ for system (2) with different τ as well as with $I_k(0) = 0.2, k = 2, 3, 4, 5, I_k(0) = 0$ for the other k, $\lambda = 0.5, \mu = 0.4$ and $R_0 = 0.4006$. (b) The time evolutions of the average force of infection $\Theta(t)$ for system (2) with different τ as well as $I_k(0) = 0.2, k = 2, 3, 4, 5, I_k(0) = 0$ for the other k, $\lambda = 1, \mu = 0.1$ and $R_0 = 3.2051$.

3. Analysis of the *SEIRS* model with time delays

3.1. The *SEIRS* model

Since some diseases confer temporary immunity, the recovered nodes cycle back into the susceptive class after an immune period. Let ω be furthermore the immune period of the recovered node, and the recovered node cycles back into the susceptive class after an immune period ω. Denote that $\sigma = \max\{\tau, \omega\}$. Based on the model (2), the relative densities $S_k(t)$, $E_k(t)$, $I_k(t)$ and $R_k(t)$, at the mean-field level, satisfy the following set of coupled different equations when $t > 0$:

$$\begin{cases} \dot{S}_k(t) = -\lambda(k)S_k(t)\Theta(t) + \mu I_k(t-\omega), \\ \dot{E}_k(t) = \lambda(k)S_k(t)\Theta(t) - \lambda(k)S_k(t-\tau)\Theta(t-\tau), \\ \dot{I}_k(t) = \lambda(k)S_k(t-\tau)\Theta(t-\tau) - \mu I_k(t), \\ \dot{R}_k(t) = \mu I_k(t) - \mu I_k(t-\omega) \end{cases} \tag{29}$$

with the normalization condition (3).

Furthermore, we obtain from the third equation and the fourth equation of system (29) that

$$E_k(t) = \lambda(k)\int_{t-\tau}^{t} S_k(s)\Theta(s)ds, R_k(t) = \mu\int_{t-\omega}^{t} I_k(s)ds. \tag{30}$$

Hence, the normalization condition becomes the following mathematical form:

$$S_k(t) + \lambda(k)\int_{t-\tau}^{t} S_k(s)\Theta(s)ds + I_k(t) + \mu\int_{t-\omega}^{t} I_k(s)ds = 1. \tag{31}$$

Obviously, if we discuss the dynamical behaviors of system (29), we just need to discuss the following system:

$$\begin{cases} \dot{S}_k(t) = -\lambda(k)S_k(t)\Theta(t) + \mu I_k(t-\omega), \\ \dot{I}_k(t) = \lambda(k)S_k(t-\tau)\Theta(t-\tau) - \mu I_k(t) \end{cases} \tag{32}$$

with the normalization condition (31).

The initial conditions of system (32) are

$$S_k(\theta) = \varphi_k(\theta), I_k(\theta) = \Psi_k(\theta), \theta \in [-\sigma, 0], \tag{33}$$

which satisfy $S_k(0) + \lambda(k)\int_{-\tau}^{0} S_k(s)\Theta(s)ds + I_k(0) + \mu\int_{-\omega}^{0} I_k(s)ds = 1$. Hence, the normalization condition (31) holds. And, $\Phi_k = (\phi_k(\theta), \Psi_k(\theta), k = m, m+1, \cdots, n-m+1) \in C$ are nonnegative continuous on $[-\sigma, 0]$, $\phi_k(0) > 0$, $\Psi_k(0) > 0$, and $\zeta(\theta) = 0$ for $\theta = 0$. C denotes the Banach space $C([-\sigma, 0], R^{2(n-m+1)})$ with the norm $\|\omega\| = \left(\sum_{i=m}^{n}\left(|\Psi_i(\theta)|_\sigma^2 + |\varphi_i(\theta)|_\sigma^2\right)\right)^{1/2}$, where $|f(\theta)|_\sigma = \sup_{-\tau\le\theta\le 0}|f(\theta)|$.

3.2. The main results for the model

Denote that

$$R_0 = \frac{1}{\mu}\frac{\langle \lambda(k)\varphi(k)\rangle}{\langle k\rangle}, \tag{34}$$

where $\langle f(k)\rangle = \sum_{k=m}^{n} f(k)p(k)$ in which $f(k)$ is a function.

Theorem 3.1. *System* (32) *always has a disease-free equilibrium* E_0 $(1, 1, \cdots, 1, 0, 0, \cdots, 0)$*, and it has a unique endemic equilibrium* $E_* = \left(S_m^*, S_{m+1}^*, \cdots, S_n^*, I_m^*, I_{m+1}^*, \cdots, I_n^*\right)$ *when* $R_0 > 1$.

Proof. Obviously, the disease-free equilibrium E_0 of system (32) always exists. Now, we discuss the existence of the endemic equilibrium of system (32). Note that the equilibrium E_* should satisfy:

$$\begin{aligned} &-\lambda k S_k^* \Theta^* + \mu I_k^* = 0, \\ &S_k^* + \lambda(k)\tau S_k^* \Theta^* + I_k^* + \omega\mu I_k^* = 1, \end{aligned} \tag{35}$$

where

$$\Theta^* = \frac{1}{\langle k\rangle}\sum_{k'} \phi(k')p(k')I_k^*. \tag{36}$$

We obtain from (35) that

$$I_k^* = \frac{\lambda k \Theta^*}{\mu + \lambda k\left(1 + \mu\tau + \omega\mu\right)\Theta^*}. \tag{37}$$

Substituting it into Eq. (4), we obtain the self-consistency equality:

$$\Theta^* = \frac{\lambda}{\langle k\rangle}\sum_{k=m}^{n} \phi(k)p(k)\frac{\lambda k\Theta^*}{\mu + \lambda k\left(1 + \mu\tau + \omega\mu\right)\Theta^*} = f(\Theta^*). \tag{38}$$

Note that

$$f'(\Theta^*)\big|_{\Theta^*=0} = \frac{\lambda}{\langle k\rangle}\sum_{k=m}^{n} \phi(k)p(k)\frac{\lambda k\mu}{\left(\mu + \lambda k\left(1 + \mu\tau + \omega\mu\right)\Theta^*\right)^2}\bigg|_{\Theta^*=0} = \frac{\lambda\langle k\phi(k)\rangle}{\mu\langle k\rangle} = R_0 \tag{39}$$

and

$$f''(\Theta^*) = \frac{-2\lambda}{\langle k\rangle}\sum_{k=m}^{n} \phi(k)p(k)\frac{\lambda^2 k^2\mu\left(1 + \mu\tau + \omega\mu\right)\Theta^*}{\left(\mu + \lambda k\left(1 + \mu\tau + \omega\mu\right)\Theta^*\right)^3} < 0. \tag{40}$$

Hence, if $R_0 > 1$, Eq. (38) has a unique positive solution. Consequently, system (32) has a unique endemic equilibrium $E_*\left(S_1^*, S_2^*, \cdots, S_n^*, I_1^*, I_2^*, \cdots, I_n^*\right)$ since (35) and (37) hold.

Theorem 3.2. *If $R_0 \le 1$, the disease-free equilibrium E_0 of system* (32) *is globally attractive.*

Proof. We define a Lyapunov function $V(t)$ as

$$V(t) = \frac{1}{2}\Theta^2(t) + \gamma \int_{t-\tau}^{t} \Theta^2(\mu)d\mu, \tag{41}$$

where γ is a constant to be determined. Let $G = \{\phi : \dot{V}(\phi) = 0\}$, and M is the largest set in G which is invariant with respect to system (32). Clearly, M is not empty since $E_0 \in M$. Calculating the derivative of $V(t)$ along the solution of (32), we get

$$\begin{aligned}
\dot{V}(t)\big|_{(3.3)} &= \Theta(t)\left[\frac{1}{\lambda(k)}\rangle \sum_k \varphi(k)p(k)\big(-\lambda k S_k(t-\tau)\Theta(t-\tau) - \mu\Theta(t)\big)\right] + \gamma\Theta^2(t) - \gamma\Theta^2(t-\tau) \\
&\le \Theta(t)\left[\frac{1}{\langle k\rangle}\big\langle \lambda(k)\phi(k)\big\rangle\Theta(t-\tau) - \mu\Theta(t)\right] + \gamma\Theta^2(t) - \gamma\Theta^2(t-\tau) \\
&\le \frac{1}{2\langle k\rangle}\big\langle \lambda(k)\phi(k)\big\rangle\Theta^2(t) + \frac{1}{2\langle k\rangle}\big\langle \lambda(k)\phi(k)\big\rangle\Theta^2(t-\tau) - \mu\Theta^2(t) + \gamma\Theta^2(t) - \gamma\Theta^2(t-\tau) \\
&= \left(\frac{1}{2\langle k\rangle}\big\langle \lambda(k)\phi(k)\big\rangle - \mu + \gamma\right)\Theta^2(t) + \left(\frac{1}{2\langle k\rangle}\big\langle \lambda(k)\phi(k)\big\rangle - \gamma\right)\Theta^2(t-\tau).
\end{aligned} \tag{42}$$

Note that $R_0 \le 1$ implies $\frac{1}{\langle k\rangle}\langle \lambda(k)\varphi(k)\rangle \le \mu < 0$; if we let $\gamma = \frac{1}{2\langle k\rangle}\langle \lambda(k)\varphi(k)\rangle$, we have from (42) that

$$\dot{V}(t)\big|_{(3.3)} \le \left(\frac{1}{\langle k\rangle}\big\langle \lambda(k)\phi(k)\big\rangle - \mu\right)\Theta^2(t) \le 0. \tag{43}$$

It follows from $S_k(t) + E_k(t) + I_k(t) + R_k(t) = 1$ that $M = E_0$. Therefore, by the LaSalle invariance principle [24], the disease-free equilibrium E_0 is globally attractive.

Lemma 3.1. [28] *Consider the following equation:*

$$\dot{x}(t) = a_1 x(t-\tau) - a_2 x(t), \tag{44}$$

where $a_1, a_2, \tau > 0$; $x(t) > 0$ for $-\tau \le t \le 0$. We have

i. if $a_1 < a_2$, then $\lim_{t \to +\infty} x(t) = 0$,

ii. if $a_1 > a_2$, then $\lim_{t \to +\infty} x(t) = +\infty$.

Lemma 3.2. ([29], p 273–280) *Let X be a complete metric space, $X = X^0 \cup \partial X^0$, where ∂X^0, assumed to be nonempty, is the boundary of X^0. Assume the C^0 – semigroup $T(t)$ on X satisfies $T(x) : X^0 \to X^0$, $T(x) : \partial X^0 \to \partial X^0$ and*

i. there is a t_0 such that $T(t)$ is compact for $t > t_0$.

ii. $T(t)$ is point dissipative in X.

iii. $\tilde{A}_\partial$ is isolated and has an acyclic covering M.

Then, $T(t)$ is uniformly persistent if and only if, for each $M_i \in M$,

$$W^s(M_i) \cap X^0 = \varnothing, \tag{45}$$

where $\tilde{A}_\partial = \bigcup_{x \in A_\partial} \omega(x)$, $\omega(x)$ is the omega limit set of $T(x)$ through x, and A_∂ is global attractor of $T_\partial(t)$ in ∂X^0 in which $T_\partial(t) = T(t)|_{\partial X^0}$.

Theorem 3.3. *For system* (32)*, if* $R_0 > 1$*, the disease is uniformly persistent, i.e., there exists a positive constant* ε *such that* $\lim_{t \to +\infty} \inf I(t) > \varepsilon$*, where* $I(t) = \sum_{k=m}^{n} \phi(k) p(k) I_k(t)$.

Proof. Denote that

$$X = \left\{ \left(\overline{S}, \overline{\Psi}\right) : \Psi_k(\theta) \geq 0, \textit{for all } \theta \in [-\zeta, 0], k = m, m+1, \cdots, n \right\}, \tag{46}$$

$$X^0 = \left\{ \left(\overline{S}, \overline{\Psi}\right) : \Psi_k(\theta) > 0, \textit{for some } \theta \in [-\zeta, 0], k = m, m+1, \cdots, n \right\}, \tag{47}$$

and, consequently,

$$\partial X^0 = X / X^0 = \left\{ \left(\overline{S}, \overline{\Psi}\right) : \Psi_i(\theta) = 0, \textit{for all } \theta \in [-\sigma, 0], i \in \{m, m+1, \cdots, n\} \right\}, \tag{48}$$

where $\left(\overline{S}, \overline{\Psi}\right) = (S_m, S_{m+1}, \cdots, S_n, \Psi_m, \Psi_{m+1}, \cdots, \Psi_n)$.

Let $(S_m(t), I_m(t), \cdots, S_n, I_n(t)) = (S_m(t, \omega), I_m(t, \omega), \cdots, S_n(t, \omega), I_n(t, \omega))$ be the solution of (32) with initial function $\omega = (\zeta_m(\theta), \Psi_m(\theta), \cdots, \Psi_n(\theta), \phi_n(\theta))$ and

$T(t)(\omega)(\theta) = (S_m(t+\theta, \omega), I_m(t+\theta, \omega), \cdots, S_n(t+\theta, \omega), I_n(t+\theta, \omega))$, $\theta \in [-\sigma, 0]$. Obviously, X and X^0 are positively invariant sets for $T(t)$. $T(t)$ is completely continuous for $t > 0$. Also, it follows from $0 < S_k(t), I_k(t) \leq 1$ for $t > 0$ that $T(t)$ is point dissipative. E_0 is the unique equilibrium of system (32) on ∂X^0, and it is globally stable on ∂X^0, $\tilde{A}_\partial = \{E_0\}$, while E_0 is isolated and acyclic.

Finally, the proof will be done if we prove $W^s(E_0) \cap X^0 = \varnothing$, where $W^s(E_0)$ is the stable manifold of E_0. Suppose it is not true, then there exists a solution $\left(\overline{S}, \overline{I}\right)$ in X^0 such that

$$\lim_{t \to +\infty} \inf S_k(t) = 1, \ \lim_{t \to +\infty} \inf I_k(t) = 0, k = 1, 2, \cdots, n. \tag{49}$$

Since $R_0 > 1$, we may choose $0 < \eta < 1$ such that $\alpha = \eta(\lambda(k)\tau\langle\varphi(k)\rangle + 1 + \mu\omega)$ satisfies $(1 - \alpha)R_0 > 1$. At the same time, there exists a $t_1 > \tau$ such that $I_k(t) < \eta$ for $t > t_1$ due to $\lim_{t \to +\infty} \inf I_k = 0$.

When $t > t_1$, we obtain from (32) that

$$S_k(t) = 1 - \left(I_k(t) + \lambda(k) \int_{t-\tau}^{t} S_k(s) \Theta(s) ds + \mu \int_{t-\omega}^{t} I_k(s) ds \right) \geq 1 - \left(\eta + \lambda(k)\tau\eta\langle\phi(k)\rangle + \mu\omega\eta \right) = 1 - \alpha. \tag{50}$$

On the other hand, for $t > t_1$ we have from (4) and (50) that

$$\begin{aligned}\dot{\Theta}(t) &= \frac{1}{\langle k\rangle}\sum_{k=m}^{n}\phi(k)p(k)\dot{I}_k(t)\\ &= \frac{1}{\langle k\rangle}\sum_{k=m}^{n}\phi(k)p(k)\left[\lambda(k)S_k(t-\tau)\Theta(t-\tau)-\mu I_k(t)\right]\\ &\geq (1-\alpha)\frac{\langle\lambda(k)\phi(k)\rangle}{\langle k\rangle}\Theta\big(t-\tau)\big)-\mu\Theta(t)\end{aligned} \tag{51}$$

Note that $(1-\alpha)R_0>1$, and it follows with $(1-\alpha)\frac{\langle\lambda(k)\varphi(k)\rangle}{\langle k\rangle}>\mu$. Hence, we obtain from (51) that $\lim_{t\to+\infty}\Theta(t)=+\infty$ according to Lemma 3.1 contradicts $\lim_{t\to+\infty}\Theta(t)=0$ due to $\lim_{t\to+\infty}I_k(t)=0$. Then, $W^s(E_0)\cap X^0=\varnothing$.

Hence, the infection is uniformly persistent according to Lemma 3.2, i.e., there exists a positive constant ε such that $\lim_{t\to+\infty}\inf I_k>\varepsilon$ and, consequently, $\lim_{t\to+\infty}\inf I(t)>\sum_{k=m}^{n}p(k)\varepsilon=\varepsilon$. This completes the proof.

In addition, Liu and Zhang discussed a simple *SEIRS* model without delay in [25], and the basic productive number for the model in [25] is $\lambda A/\gamma$, which is consistent with R_0 for the model (32) in which $\phi(k)=A$ in this paper.

3.3. Numerical simulations for the model

Now, we present the results of numerical simulations. The degree distribution of the scale-free network is $p(k)=Ck^{-\gamma}$, and C satisfies $\sum_{k=1}^{n}p(k)=1$. Here, we set still the maximum degree $n=100$ and the minimum degree $m=1$.

Consider system (32). Let $\varphi(k=ak^{\alpha}/(1+b^{\alpha})$ in which $a=0.5$, $\alpha=0.75$, $b=0.02$ and $\lambda(k)=\lambda k$, and let $\gamma=2.5$ and $\mu=0.06$. **Figures 1–4** show that the dynamic behaviors of system (32) with the initial functions satisfy condition (33) in which $I_k(s)=0.45, k=2,3,4,5$ for $s\in[-\sigma,0]$ and $I_k=0$, $k\neq2,3,4,5$.

Although R_0 is irrelative to τ and ω. **Figures 6** and **7** show that both the delay τ and ω have certain influence on the relative density of the infected nodes when $R_0<1$, for example, the faster the relative density of infected nodes converges to zero, the larger ω gets or the smaller τ gets. In addition, **Figures 6** and **7** show that the average relative density of the infected nodes $I(t)$ monotonically decreases to zero, whereas the relative density of infected nodes of connectivity k always breaks out first and then decreases to zero; the reason of the phenomenon appears that the spreading network is a scale-free one. Note that **Figures 8** and **9** show that the delay τ and ω have much impact on the steady state of density of the infected nodes when $R_0>1$, the density of infected decreases as the delay τ and ω increase, which is consistent with the formula (37). We also know from (37) that $I(t)\to0$ as $\omega\to+\infty$ or $\tau\to+\infty$.

Especially, system (29) reduces to the following *SIRS* model [26]:

$$\begin{cases} \dot{S}_k(t) = -\lambda(k)S_k(t)\Theta(t) + \mu I_k(t-\omega), \\ \dot{I}_k(t) = \lambda(k)S_k(t)\Theta(t) - \mu I_k(t), \\ \dot{R}_k(t) = \mu I_k(t) - \mu I_k(t-\omega). \end{cases} \tag{52}$$

with the normalization condition

$$S_k(t) + I_k(t) + R_k(t) = 1. \tag{53}$$

Figure 10 shows that when $R_0 > 1$, the quarantine delay ω can impact the density of infected nodes at the stationary state, and raising the quarantine period will suppress the viruses when ω is not large enough, which coincides with formula (37). Moreover, there exists periodic oscillation near the endemic equilibrium when ω is large enough. This is an interesting phenomenon which means that a bifurcation may appear.

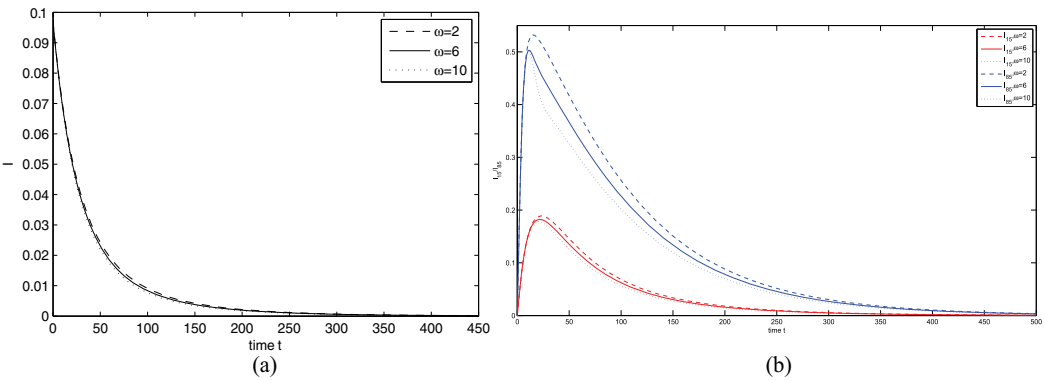

Figure 6. (a) Evolutions of $I(t)$ for system (32) with $\tau = 3$, $\lambda(k) = 0.03k$, and $R_0 = 0.8037$. (b) Evolutions of I_{15} and I_{85} for system (32) with $\tau = 3$, $\lambda(k) = 0.03k$, and $R_0 = 0.8037$.

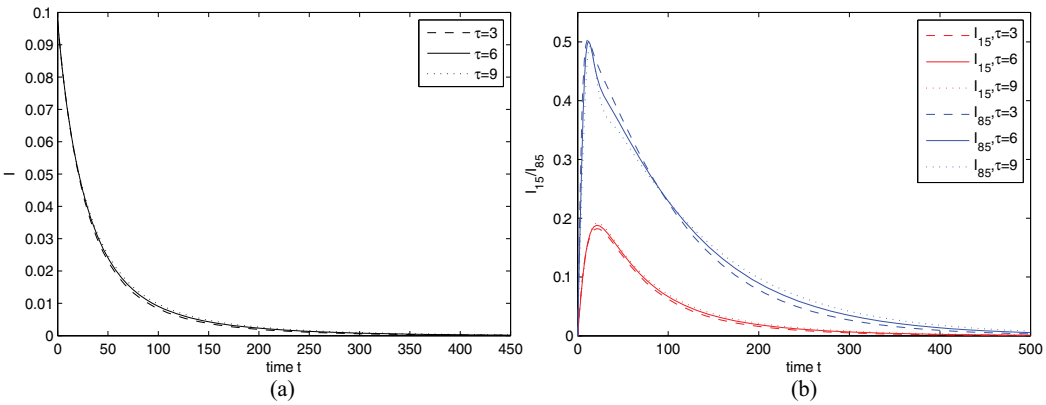

Figure 7. (a) Evolutions of $I(t)$ for system (32) with $\omega = 6$, $\lambda(k) = 0.03k$, and $R_0 = 0.8037$. (b) Evolutions of I_{15} and I_{85} for system (32) with $\omega = 6$, $\lambda(k) = 0.03k$, and $R_0 = 0.8037$.

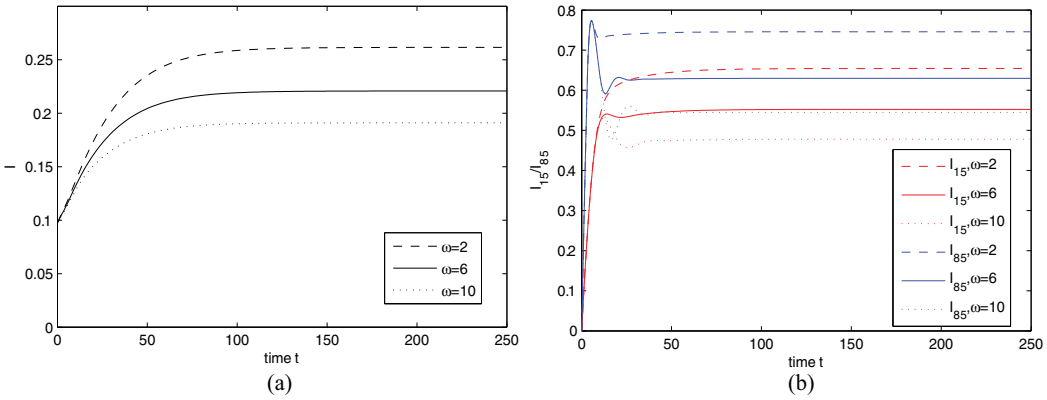

Figure 8. (a) Evolutions of $I(t)$ for system (32) with $\tau = 3$, $\lambda(k) = 0.14k$, and $R_0 = 3.7505$. (b) Evolutions of I_{15} and I_{85} for system (32) with $\tau = 3$, $\lambda(k) = 0.14k$, and $R_0 = 3.7505$.

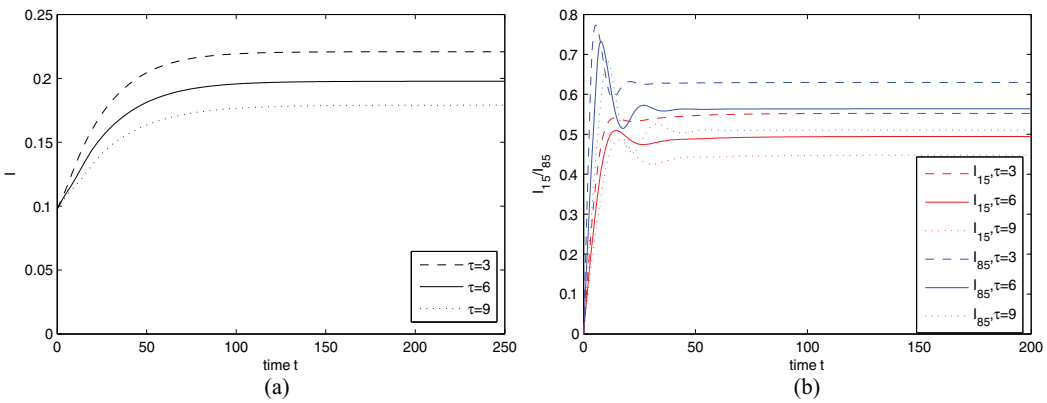

Figure 9. (a) Evolutions of $I(t)$ for system (32) with $\omega = 6$, $\lambda(k) = 0.14k$, and $R_0 = 3.7505$. (b) Evolutions of I_{15} and I_{85} for system (32) with $\omega = 6$, $\lambda(k) = 0.14k$, and $R_0 = 3.7505$.

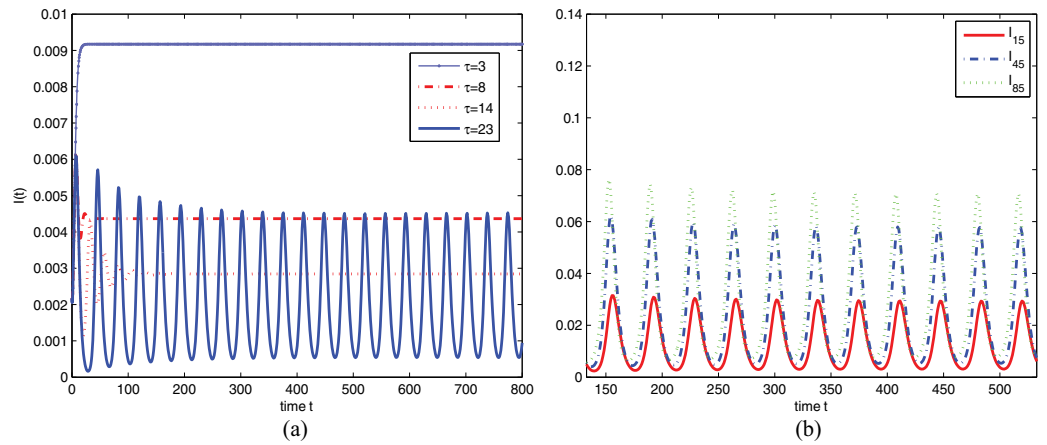

Figure 10. (a) Evolutions of $I(t)$, for system (52) with $\omega = 3, 8, 14, 23$, respectively, when $R_0 = 1.3881 > 1$. (b) Evolutions of I_{15}, I_{45}, I_{85}, for system (1) with $\omega = 23$, respectively, when $R_0 = 1.3881 > 1$.

4. Conclusion and discussion

An *SEIR* model with time delay on the scale-free network, which formulated a disease or computer virus transmission with constant latent period, is presented. For *SEIR* model, the basic reproduction number is

$$R_0 = \frac{1}{\mu}\frac{\langle \lambda(k)\phi(k)\rangle}{\langle k\rangle}. \tag{54}$$

When $R_0 \le 1$, there is no outbreak of the disease spreading, and the infection eventually disappears. When $R_0 > 1$, there exists at least one outbreak for the spreading of epidemic, and then $\lim_{t\to+\infty} I_k(t) = 0$ due to global attractivity of M_0. If the recovered nodes cycle back into the susceptive class after an immune period, we obtain a *SEIRS* model with two time delays on the scale-free network, which formulated a disease transmission with constant latent and immune periods. For *SEIRS* model, the basic reproduction number is still R_0 shown in (54). If $R_0 \le 1$, although the equilibrium E_0 is globally stable and the infection eventually disappears, the equilibrium E_0 may lose its stability when $R_0 > 1$ and the infection will always exists.

Although R_0 is irrelevant to time delays, they influence the dynamical behaviors of the model such as slowing down the speed disappear of disease spreading on network, depressing the density of infected nodes at the stationary state.

In addition, for *SEIRS* model, numerical simulations show that the endemic equilibrium E_* may be globally asymptotically stable under some conditions when $R_0 > 1$ (shown in **Figures 8** and **9**). We would like to mention here that it is interesting but challenging to discuss the stability of equilibrium E_* when $R_0 > 1$.

Furthermore, more and more researchers realize the fundamental role of the stochastic nature of diseases on their dynamics. In order to gain analytical insight into the behavior of the epidemic spreading, we also may extend the models (2) and (29) to the ones with random perturbations, i.e., stochastic differential equation models.

Acknowledgements

This research was supported by the Hebei Provincial Natural Science Foundation of China under Grant No. A2016506002. Qiming Liu would like to thank Professor Zhen Jin of Shanxi University in China for his helpful suggestions that led to truly significant improvements of this chapter.

Author details

Qiming Liu*, Meici Sun and Shihua Zhang

*Address all correspondence to: lqmmath@163.com

Shijiazhuang Branch, Arm Engineering University, Shijiazhuang, China

References

[1] Barabási AL, Albert R. Mint: Emergence of scaling in random networks. Science. 1999;**286**:509-512. DOI: 10.1126/science.286.5439.509

[2] Ma Z, Li J. Dynamical Modelling and Analysis of Epidemics. Singapore: World Scientific Publishing Company; 2009

[3] Pastor-Satorras R, Vespignani A. Mint: Epidemic dynamics in finite size scale-free networks. Physical Review E. 2002;**65**:035108. DOI: 10.1103/PhysRevE.65.035108

[4] Olinky R, Stone L. Mint: Unexpected epidemic thresholds in heterogeneous networks: The role of disease transmission. Physical Review E. 2004;**70**:030902. DOI: 10.1103/PhysRevE.70.030902

[5] Yang R, Ren J, et al. Mint: Epidemic spreading on heterogeneous networks with identical infectivity. Physics Letters A. 2007;**364**:189-193. DOI: 10.1016/j.physleta.2006.12.021

[6] Cheng X, Liu X, Chen Z, Yuan Z. Mint: Spreading behavior of SIS model with non-uniform transmission on scale-free networks. The Journal of Universities of Posts and Telecommunications. 2009;**16**:27-31. DOI: 10.1016.16/S1005-8885(08)60173-9

[7] Zhang H, Fu X. Mint: Spreading of epidemics on scale-free networks with nonlinear infectivity. Nonlinear Analysis. 2009;**70**:3273-3278. DOI: 10.1016/j.na.2008.04.031

[8] Fu X, Michael S, David M, Zhang H. Mint: Epidemic dynamics on scale-free networks with piecewise linear infectivity and immunization. Physical Review E. 2008;**77**:036113. DOI: 10.1103/PhysRevE.77.036113

[9] Moreno Y, Pastor-Satorras R, Vespignani A. Mint: Epidemic outbreaks in complex heterogeneous networks. The European Physical Journal B. 2002;**26**:521-529. DOI: 10.1140/epjb/e20020122

[10] Zhang J, Jin Z. Mint: The analysis of epidemic network model with infectious force in latent and infected period. Discrete Dynamics in Nature and Society. 2010;**2010**:604329. DOI: 10.1155/2010/604329

[11] Zhu G, Fu X, Chen G. Mint: Global attractivity of a network-based epidemics SIS model with nonlinear infectivity. Communications in Nonlinear Science and Numerical Simulation. 2013;**17**:2588-2594. DOI: 10.1016/j.cnsns.2011.08.039

[12] Gong Y, Song Y, Jiang G. Mint: Epidemic spreading in scale-free networks including the effect of individual vigilance. Chinese Physics B. 2012;**21**:010205. DOI: 10.1088/1674-1056/21/1/010205

[13] Xia C, Li W, Sun S, Wang J. Mint: An SIR model with infection delay and propagation vector in complex networks. Nonlinear Dynamics. 2012;**69**:927-934. DOI: 10.1007/s11071-011-0313-y

[14] Xu X, Chen G. The SIS model with time delay on complex networks. International Journal of Bifurcation and Chaos. 2009;**19**:623-626. DOI: 10.1142/S021812740902324X

[15] Zou S, Wu J, Chen Y. Multiple epidemic waves in delayed susceptible-infected-recovered modeled on complex networks. Physical Review E. 2011;**83**:056121. DOI: 10.1103/PhysRevE.83.056121

[16] Liu Q, Deng C, Sun M. Mint: The analysis of an epidemic model with time delay on scale-free networks. Physica A: Statistical Mechanics and its Applications. 2014;**410**:79-87. DOI: 10.1016/j.physa.2014.05.010

[17] Wang J, Wang J, Liu M, Liu Y. Mint: Global stability analysis of an SIR epidemic model with demographics and time delay on networks. Physica A: Statistical Mechanics and its Applications. 2014;**410**:268-275. DOI: 10.1016/j.physa.2014.05.011

[18] Kang H, Fu X. Mint: Epidemic spreading and global stability of an SIS model with an infective vector on complex networks. Communications in Nonlinear Science and Numerical Simulation. 2015;**27**:30-39. DOI: 10.1016/j.cnsns.2015.02.018

[19] Kermack WO, McKendrick AG. Mint: A contributions to the mathematical theory of epidemics. Proceedings of the Royal Society. A: Mathematical, Physical and Engineering Sciences. 1927;**115**:700-721. DOI: 10.1098/rspa.1927.0118

[20] Li K, Small M, Zhang H, Fu X. Mint: Epidemic outbreaks on networks with effective contacts. Nonlinear Analysis: RWA. 2010;**11**:1017-1025. DOI: 10.1016/j.nonrwa.2009.01.046

[21] Zhang J, Jin Z. The analysis of an epidemic model on networks. Applied Mathematics and Computation. 2011;**217**:7053-7064. DOI: 10.1016/j.amc.2010.09.063

[22] Wang Y, Jin Z, Yang Z, et al. Global analysis of an SIS model with an infective vector on complex networks. Nonlinear Analysis: RWA. 2012;**13**:543-557. DOI: 10.1016/j.nonrwa.2011.07.033

[23] Li C, Tsai C, Yang S. Mint: Analysis of epidemic spreading of an SIRS model in complex heterogeneous networks. Communications in Nonlinear Science and Numerical Simulation. 2014;**19**:1042-1054. DOI: 10.1016/j.cnsns.2013.08.033

[24] Li T, Wang Y, Guan Z. Mint: Spreading dynamics of a SIQRS epidemic model on scale-free networks. Communications in Nonlinear Science and Numerical Simulation. 2014;**19**:686-692. DOI: 10.1016/j.cnsns.2013.07.010

[25] Liu J, Zhang T. Mint: Epidemic spreading of an SEIRS model in scale-free networks. Communications in Nonlinear Science and Numerical Simulation. 2011;**16**:3375-3384. DOI: 10.1016/j.cnsns.2010.11.019

[26] Liu Q, Deng C. An computer virus spreading model with delayed quarantine in Internet. In: Proceedings of the 6th International Conference on Communication and Network Security, Singapore (ICNNS2016); 26-29 November, 2016; Singapore. New York: ACM; 2016. pp. 1-5

[27] Hale J. Theory of Functional Differential Equations. New York: Springer-Verlag; 1977

[28] Gao S, Chen L, Teng Z. Mint: Pulse vaccination of an SEIR epidemic model with time delay. Nonlinear Analysis: RWA. 2008;**9**:599-607. DOI: 10.1016/j.nonrwa.2006.12.004

[29] Kuang Y. Delay Differential Equations with Applications in Population Dynamics. Boston: Academic Press; 1993

Chapter 3

Stability and Hopf Bifurcation Analysis of a Simple Nutrient-Prey-Predator Model with Intratrophic Predation in Chemostat

Zabidin Salleh and Liyana Abd Rahim

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71624

Abstract

A 3-dimensional nutrient-prey-predator model with intratrophic predation is proposed and studied. Some elementary properties such as invariance of nonnegativity, boundedness and dissipativity of the system are presented. The purpose of this chapter is to study the existence and stability of equilibria along with the effects of intratrophic predation towards the positions and stability of those equilibria of the system. We also investigate the occurrence of Hopf bifurcation. In the case when there is no presence of predator organisms, intratrophic predation may not give impact on the stability of equilibria of the system. We also analysed global stability of the equilibrium point. A suitable Lyapunov function is defined for global stability analysis and some results of persistence analysis are presented for the existence of positive interior equilibrium point. Besides that, Hopf bifurcation analysis of the system are demonstrated.

Keywords: chemostat, intratrophic predation, local and global stability, Hopf bifurcation

1. Introduction

Chemostat, a piece of laboratory apparatus is frequently used in mathematical ecology. This device carries an important role for ecological studies because the mathematics are tractable, and the relevant parameters are readily measurable. Chemostat also can be used for modelling microorganism systems such as lake and wastewater treatment. For detailed information regarding chemostat device, see [1].

There are abundant of research journals and papers for analysing chemostat models. One of them is Li and Kuang [2], considered a simple food chain with distinct removal rate, which in this case,

conservation law has failed. Therefore, to overcome the problem, they constructed a Lyapunov function in global stability analysis of predator-free steady state. Local and global stability of other steady states were shown in the paper along with persistence analysis of the system.

Another paper regarding chemostat model is by El-Sheikh and Mahrouf [3] that presented a 4-dimensional food chain in a chemostat with removal rates. They studied local and global stability of equilibria along with elementary properties including boundedness of solutions, invariance of nonnegativity, dissipativity and persistence analysis. Hopf bifurcation theory was applied.

Recently, several analysis on chemostat models are carried out, for example, by Hamra and Yadi [4] and Yang et al. [5]. In the work of [4], they studied a chemostat model with constant recycle sludge concentration. Number of parameters are reduced by considering a dimensionless model. Next, they successfully proved the existence of a positive uniform attractor for the model with different removal rates by using dissipative theory. Hence, they used methods of singular perturbation theory in order to investigate the asymptotic behaviour of the chemostat model under small pertubations. Thus, it is shown that in the case of two species in competition, the positive unique equilibrium point is globally asymptotically stable.

In the research of Yang et al. [5], piecewise chemostat models which involve control strategy with threshold window are proposed and analysed. They investigated the qualitative analysis such as existence and stability of equilibrium points of the system and it is proved that the regular equilibria and pseudo-equilibrium cannot coexist. The global stability analysis of both the regular equilibria and pseudo-equilibrium have been studied using qualitative analysis techniques of non-smooth Filippov dynamic systems. Furthermore, the bifurcation analysis of the system is investigated with theoretical and numerical techniques.

Moreover, one of the interesting topic is the research involving intratrophic predation. Intratrophic predation is a situation where members of a trophic group consume other members of the same trophic group (for the purpose of mathematical modelling). Several previous studies based on intratrophic predation have been discovered. One of them is by Pitchford and Brindley [6]. They studied a predator-prey model with intratrophic predation and successfully shown that the model had desirable and plausible features. Furthermore, to investigate the effect of intratrophic predation towards the position and stability of equilibrium points of a model, they had developed a simple asymptotic method.

For Hopf bifurcation analysis, a study by Mada Sanjaya et al. [7] has been carried out. They introduced an ecological model of three species food chain with Holling Type-III functional response. Two equilibrium points are obtained. They proved that the system has periodic solutions around those equilibrium points. Not only that, they also investigated the dynamical behaviours of the system and found that it was sensitive when the parameter values varied.

Tee and Salleh [8] investigated Hopf bifurcation of a nonlinear modified Lorenz system using normal form theory that was the same technique used in Hassard et al. [9]. Then, the dynamics on centre manifold of the system was presented as it will be applied in the technique of normal form theory. Another research based on Hopf bifurcation analysis is carried out by [10]. They considered a three-species food chain models with group defence. It is proved that the model without delay undergone Hopf bifurcation by using the carrying capacity of the environment

as the bifurcation parameter. In the analysis of Hopf bifurcation in a delay model, they used a computer code BIFDD to determine the stability of bifurcation solutions.

Furthermore, a simple 3-dimensional food chain model in chemostat with variable yield for prey population and constant yield for predator population is proposed by Rana et al. [11]. In this model, the prey consumes the nutrient and the predator consumes the prey but the predator does not consume the nutrient. The functional response functions are assumed in Michaelis-Menton type. The stability of equilibrium points, the existence of limit cycles, the Hopf bifurcation and the positive invariant set for the system are discussed by qualitative analysis of differential equations. Finally, numerical simulations are carried out in support of the theoretical results.

Our work is a modification of the models in Li and Kuang [2] and El-Sheikh and Mahrouf [3]. It should be emphasised that this work is different from [2, 3]. The modified model contains parameter b, known as the measure of intensity of intratrophic predation which is not in their models. The parameter b and term $\left(1+\frac{D_1x}{1+x+D_1by}\right)$ are added to differential equations x' and y' in the interest of studying the behaviour of the modified model. By this motivation, we analysed the stability and Hopf bifurcation of the model with intratrophic predation, as intratrophic predation analysis is rarely considered in mathematical models of populations biologically [6].

The purpose of this chapter is to study the existence and stability of equilibria along with the effects of intratrophic predation towards the positions and stability of those equilibria of the system. We shall also investigate the occurrence of Hopf bifurcation towards the system. Hence, we introduced a simple nutrient-prey-predator model in chemostat with intratrophic predation. Some notations regarding the model will be explained. Not only that, several elementary properties such as nonnegativity, boundedness and dissipativity along with definitions and several results will be presented. The local and global stability together with existence of the equilibrium points are shown. For global stability analysis, a suitable Lyapunov function is defined. Lastly, we applied Hopf bifurcation theorems (see [9]) in the analysis of Hopf bifurcation.

2. The model

In this work, we consider a nutrient-prey-predator model with one prey organism and one predator organism in chemostat with intratrophic predation. Biologically, the predator organisms will feed upon the prey organisms, while the prey organisms will consume the nutrient in the chemostat. Precisely, the modified model of [2, 3] is

$$\begin{cases} s'(t) = (s^0 - s)D_0 - \dfrac{1}{\beta_1} f_1(s)x, \\ x'(t) = \left(f_1(s) - D_1\right)x - \dfrac{1}{\beta_2} f_2(x)\left(y + \dfrac{D_1xy}{D_2 + x + by}\right), \\ y'(t) = f_2(x)\left(y + \dfrac{D_1xy}{D_2 + x + by}\right) - y, \end{cases} \tag{1}$$

where $s(0) = s_0 > 0$, $x(0) = x_0 > 0$, $y(0) = y_0 > 0$, $D_i = 1$; $i = 0, 1, 2$, $\beta_j \leq 1$; $j = 1, 2$ and $0 \leq b \leq 1$. This leads to the following assumptions on functional response f_n, $n = 1, 2$:

$$\begin{aligned} &(i)\ f_n : \mathbb{R}^+ \to \mathbb{R}^+ \text{ and } f_n, n = 1, 2 \text{ are continuously differential equations,} \\ &(ii)\ f_n(0) = 0, \\ &(iii)\ f_1'(s) > 0, f_2'(x) > 0 \text{ for all } s, x > 0. \end{aligned} \tag{2}$$

Generally, functional response is a common component in predator-prey system. The term 'functional response' was first stated by Solomon [12] which defined the relationship between the rate of predation, i.e., the number of prey organisms consumed per predator organism in time, t with the density of prey organisms.

By referring to system (1), s represents the concentration of nutrient at time t while s^0 is the input nutrient concentration. The variables x and y represent the concentration of prey and predator at time t, respectively. Parameters β_1 and β_2 denote the yield constants, D_0 is the washout rate of the chemostat while D_1 and D_2 are the removal rates of x and y, respectively. Removal rate is the sum of washout rate and death rate θ_i, i.e., $D_i = D_0 + \theta_i$, $i = 1, 2$. $f_1(s)$ and $f_2(x)$ denote the specific growth rate of prey x and predator y, respectively, while b is the measure of intensity of intratrophic predation in predator organisms y. These observations are based on numerical simulations. We can rescale system (1) by reducing the number of parameters using standard change of variables such as.

$$\bar{s} = \frac{s}{s^0}, \qquad \bar{x} = \frac{x}{D_0\beta_1 s^0}, \qquad \bar{y} = \frac{y}{D_0 D_1 \beta_1 \beta_2 s^0},$$

$$f_1(s) = \frac{\tilde{f_1}(s^0\bar{s})}{D_0}, \qquad f_2(x) = \frac{\tilde{f_2}(D_0\beta_1 s^0 \bar{x})}{D_0}, \qquad b = \frac{\bar{b}}{\beta_2},$$

$$\bar{t} = D_i t, i = 0, 1, \qquad D_2 = D_0\beta_1 s^0.$$

To make our works more convenient, we just drop the bars and tildes. So after rescaling, the system (1) becomes

$$\begin{cases} s' = 1 - s - f_1(s)x, \\ x' = \left(\dfrac{f_1(s)}{D_1} - 1\right)x - f_2(x)y\left(1 + \dfrac{D_1 x}{1 + x + D_1 by}\right), \\ y' = f_2(x)y\left(1 + \dfrac{D_1 x}{1 + x + D_1 by}\right) - y, \end{cases} \tag{3}$$

where

$$s(0) > 0, x(0) > 0, y(0) > 0 \text{ and } D_1 > 0. \tag{4}$$

Now, the variables are non-dimensional and the discussion is in $\mathbb{R}^3_+ = \{(s, x, y) : s > 0, x > 0, y > 0\}$.

2.1. Elementary properties of system (3)

In this section, we present nonnegativity, boundedness and dissipativity of the system (3) with respect to a region in $\mathbb{R}^3_+$. Firstly, we consider dissipativity of the system (3).

Definition 1 [13]. A system with differential equations $x' = f(x)$ is defined to be dissipative if there exists a bounded subset Γ of $\mathbb{R}^3$, such that there is a time t_0 for any $x^0 \in \mathbb{R}^3$ which depends on x^0 and Γ so that the solution of the system $\phi(t, x^0) \in \Gamma$ for $t \geq t_0$.

Theorem 1. *Let H be the region*

$$H = \left\{ (s,x,y) \in \mathbb{R}^3_+ : \frac{1}{P_{max}} - q \leq s + x + y \leq \frac{1}{P_{min}} + q \right\},$$

where $P_{max} = \max\left\{1, f_1(s) + \frac{f_1(s)}{D_1} - 1\right\}$, $P_{min} = \min\left\{1, f_1(s) + \frac{f_1(s)}{D_1} - 1\right\}$ *and q is a positive constant.*

Then,

i. *H is positively invariant.*

ii. *All nonnegative solutions of (3) with initial values in* $\mathbb{R}^3_+$ *are uniformly bounded and they eventually attracted into region H.*

iii. *The system is dissipative.*

Remark 1. We must show that the solutions of system (3) are nonnegative and bounded, so that the system becomes biologically meaningful.

Proof of Theorem 1.

(i): First, we must show that $\frac{1}{P_{max}} - q \leq s + x + y \leq \frac{1}{P_{min}} + q$. Let us define

$$P_{max} = \max\left\{1, f_1(s) + \frac{f_1(s)}{D_1} - 1\right\} \text{ and } P_{min} = \min\left\{1, f_1(s) + \frac{f_1(s)}{D_1} - 1\right\}.$$

By adding those differential equations s', x' and y' in (3), we shall get

$$s' + x' + y' = 1 - \left[s + \left(f_1(s) + \left(\frac{f_1(s)}{D_1} - 1 \right) \right) x + y \right].$$

Hence, this leads to

$$1 - P_{max}(s + x + y) \leq 1 - \left[s + \left(f_1(s) + \left(\frac{f_1(s)}{D_1} - 1 \right) \right) x + y \right] \leq 1 - P_{min}(s + x + y).$$

Thus, solving the above inequality gives

$$\frac{1}{P_{max}} - q \leq s + x + y \leq \frac{1}{P_{min}} + q,$$

as q is the positive constant

(ii) and (iii): Let $0 < s_0 < 1$ and consider.

$$s' = 1 - s - f_1(s)x < 1 - s.$$

Then $s(t) < 1 - \frac{e^{-t}}{1-s_0}$ for $0 < s_0 < 1$, and hence $\lim_{t\to\infty} \sup s(t) < 1$ for $0 < s_0 < 1$.

Now, consider the x' equation;

$$x' = \left(\frac{f_1(s)}{D_1} - 1\right)x - f_2(x)y\left(1 + \frac{D_1 x}{1 + x + D_1 by}\right) < \left(f_1(s) - 1\right)x < \alpha_1 x,$$

where $\alpha_1 = \max_{s\in H}\left(f_1(s) - 1\right)$. We assume that $\alpha_1 = \left(\max_{s\in H}\left(f_1(s) - 1\right)\right) < 0$. Hence, $x(t) < x_0 e^{\alpha_1 t}$, $\alpha_1 < 0$, and thus, $\lim_{t\to\infty} \sup x(t) < x_0$, for $x_0 > 0$.

Similarly, consider the y' equation,

$$y' = f_2(x)y\left(1 + \frac{D_1 x}{1 + x + D_1 by}\right) - y < \alpha_2 y,$$

where $\alpha_2 = \max_{x\in H}\left(f_2(x) + \frac{f_2(x)D_1 x}{1+x+D_1 by}\right)$. We shall assume that $\max_{x\in H}\left(f_2(x) + \frac{f_2(x)D_1 x}{1+x+D_1 by}\right) < 0$. Then, $y(t) < y_0 e^{\alpha_2 t}$ where $\alpha_2 < 0$. Hence, $\lim_{t\to\infty} \sup\, y(t) < y_0$ for $y_0 > 0$. Thus, system (3) is proved to be uniformly bounded and dissipative, following Definition 1. ∎

2.2. Existence of equilibrium points

According to Robinson [14], a point is termed 'equilibrium point' because of the forces are in equilibrium and the mass did not move. Therefore, in order to find equilibrium points of the system (3); E_1, E_2 and E_3, we equate the differential equations s', x' and y' to zero and solve the resulting equations simultaneously. The possible equilibrium points are as follows;

i. $E_1(1, 0, 0)$ where no predator organism y and prey x exist. By equating system (3) to zero, we will get $s = 1$ from the equation $1 - s - f_1(s)x = 0$. Then, from equation $x' = 0$, we get $x = 0$ when $y = 0$.

ii. $E_2\left(\zeta_s, \frac{1-\zeta_s}{D_1}, 0\right)$. $s = \zeta_s$ is the unique solution of $\frac{f_1(s)}{D_1} - 1 = 0$. Let $y = 0$, then Eq. (3) give $1 - s - f_1(s)x = 0$ and $\left(\frac{f_1(s)}{D_1} - 1\right)x = 0$. When $x \neq 0$, $\frac{f_1(s)}{D_1} - 1 = 0$. Therefore, $f_1(s) = D_1$. Next, we find x. From equation $1 - s - f_1(s)x = 0$, $x = \frac{1-s}{f_1(s)}$. By substituting $f_1(s) = D_1$ and $s = \zeta_s$, we get $x = \frac{1-\zeta_s}{D_1}$. Thus, E_2 is the one possible equilibrium point that consists only prey organisms and not predator organisms.

iii. $E_3(s^*, x^*, y^*)$ denotes the positive interior equilibrium point with $s^*, x^*, y^* > 0$. s^* is a unique solution of $1 - s - f_1(s)x = 0$. From this, we have $x^* = \frac{1-s^*}{f_1(s^*)}$. Next, let the

equation $y' = 0$. Since $y \neq 0$, $f_2(x)\left(1 + \frac{D_1 x}{1+x+D_1 by}\right) = 1$. We solve for y, and get $y^* = \frac{(2x^*+1)f_2(x^*)-x^*-1}{bD_1\left(1-f_2(x^*)\right)}$. Thus, the equilibrium point $E_3(s^*, x^*, y^*)$ is $\left(s^*, \frac{1-s^*}{f_1(s^*)}, \frac{(2x^*+1)f_2(x^*)-x^*-1}{bD_1\left(1-f_2(x^*)\right)}\right)$, where $x^* = \frac{1-s^*}{f_1(s^*)}$.

To show the existence of E_2, we let the system (3) be restricted to $\mathbb{R}^+_{sx}$ as

$$\begin{cases} s' = 1 - s - f_1(s)x, \\ x' = \left(\frac{f_1(s)}{D_1} - 1\right)x, \end{cases} \tag{5}$$

where $s(0) > 0$ and $x(0) > 0$. Thus, Lemma 1 below shows the existence of non-trivial equilibrium point E_2.

Lemma 1. Suppose that a point $(\zeta_s, \widehat{x})$ exists in $\mathbb{R}^+_{sx}$ such that $\zeta_s + D_1\widehat{x} - 1 = 0$ as time, t approaches ∞. Then, the non-trivial equilibrium point $E_2\left(\zeta_s, \frac{1-\zeta_s}{D_1}, 0\right)$ exists.

Proof. We will get two surfaces by equating system (5) to zero;

$$1 - s - f_1(s)x = 0,$$

$$\left(\frac{f_1(s)}{D_1} - 1\right)x = 0.$$

Then,

$$\frac{1-s}{x} = f_1(s) \text{ and } D_1 = f_1(s).$$

Thus, $1 - s = D_1 x$, i.e., $x = \frac{1-s}{D_1}$. By taking $s = \zeta_s$, we will have $x = \widehat{x} = \frac{1-\zeta_s}{D_1}$ and satisfying $\zeta_s + D_1\widehat{x} - 1 = 0$. Hence, the equilibrium point $E_2\left(\zeta_s, \frac{1-\zeta_s}{D_1}, 0\right)$ exists. ∎

2.3. Stability analysis of equilibrium points E_1, E_2 and E_3

In this section, we analyse the stability of equilibrium points as it plays an important part in ordinary differential equations with their applications. As we cannot easily identify the positions of equilibrium point in applications of dynamical system, but only approximately, so the equilibrium point must be in stable state to be biologically meaningful [15]. Several definitions and theorem from Ref. [14] are stated to make us understand clearly.

Definition 2. [14] A fixed point x^* is Lyapunov stable or L-stable, provided that any solution $\phi(t; x_0)$ stays near x^* for all future time $t \geq 0$ if the initial condition x_0 starts near enough to x^*. Specifically, a fixed point x^* is L-stable, provided that for any $\epsilon > 0$, there is $\delta > 0$, such that if $\|x_0 - x^*\| < \delta$, then $\|\phi(t; x_0) - x^*\| < \epsilon$ for all $t \geq 0$.

Another form of stability is asymptotically stable. This is stated in the following Definition 3. below. Before going further to understand this concept, we need the definition of ω-limit set as follows.

Definition 3. [14] A point k is an ω-limit point of the trajectory of x_0, if $\phi(t;x_0)$ keeps coming near k as $t \to \infty$ (i.e., there is a sequence of times t_j, with $t_j \to \infty$ as $j \to \infty$ such that $\phi(t_j;x_0)$ converges to k). Certainly, if $\left\|\phi(t;x_0) - x^*\right\| \to 0$ as $t \to \infty$, then x^* is the only ω-limit point of x_0. There can be more than one point that is an ω-limit point of x_0. The set of all ω-limit points of x_0 is denoted by $\omega(x_0)$ and is called the omega limit set of x_0.

Definition 4. [14] A fixed point x^* is weakly asymptotically stable, if there exists $\delta_1 > 0$ such that $\omega(x_0) = \{x^*\}$ for all $\|x_0 - x^*\| < \delta_1$ (i.e., $\left\|\phi(t;x_0) - x^*\right\| \to 0$ as time $t \to \infty$ for all $\|x_0 - x^*\| < \delta_1$). Therefore, a fixed point is weakly asymptotically stable, if the stable manifold contains all points in a neighbourhood of the fixed point (i.e., all points are sufficiently close). A fixed point x^* is asymptotically stable if it is both L-stable and weakly asymptotically stable. An asymptotically stable fixed point is also called sink.

Moreover, Theorem 2 below clearly shows that if a fixed point x^* is hyperbolic (the real parts of all eigenvalues of the Jacobian matrix at x^*, i.e., $DF_{(x^*)}$ are nonzero), then the stability type of the fixed point for the nonlinear system is the same as for the linearized system at that fixed point.

Theorem 2. [14] *Let $\dot{x} = F(x)$ be a differential equation in n variables, with a hyperbolic fixed point x^*. Suppose thal F, $\frac{\partial F_i}{\partial x_j}(x)$ and $\frac{\partial^2 F_i}{\partial x_j \partial x_k}(x)$ are all continuous. Then, the stability type of the fixed point for the nonlinear system is the same as for the linearized system at that fixed point.*

a. *If the real parts of all the eigenvalues of $DF_{(x^*)}$ are negative, then the fixed point is asymptotically stable for the nonlinear equation (i.e., if the origin of the linearized system is asymptotically stable, then x^* is asymptotically stable for the nonlinear equation). In this case, the basin of attraction $W^S(x^*)$ is an open set that contains some solid ball about the fixed point.*

b. *If there is at least one eigenvalue of $DF_{(x^*)}$ has positive real part, then the fixed point x^* is unstable. (For the linearized system, the fixed point can be a saddle, unstable node, unstable focus, etc.)*

c. *If one of the eigenvalues of $DF_{(x^*)}$ has zero real part, then the situation is more complicated. In particular, for $n = 2$, if the fixed point is an elliptic center (eigenvalues $\pm\beta i$) or one eigenvalue is zero of multiplicity one, then the linearized system does not determine the stability type of the fixed point.*

Now by utilising the Theorem 2, we are going to analyse the stability of the fixed points E_1, E_2 and E_3.

(i) Stability analysis of E_1.

Now, we will discuss the stability type of the equilibrium point $E_1(1,0,0)$. The Jacobian matrix of the system (3) is

$$J=\begin{bmatrix} -f_1'(s)x-1 & -f_1(s) & 0 \\ \frac{f_1'(s)x}{D_1} & \frac{f_1(s)}{D_1}-f_2(x)y\left[\left(\frac{D_1}{1+x+D_1by}\right)-\left(\frac{D_1x}{(1+x+D_1by)^2}\right)\right]-f_2{}'(x)y\left(\left(\frac{D_1x}{1+x+D_1by}\right)+1\right)-1 & \left(\frac{f_2(x)D_1^2bxy}{(1+x+D_1by)^2}\right)-f_2(x)\left(1+\frac{D_1x}{1+x+D_1by}\right) \\ 0 & f_2(x)y\left[\left(\frac{D_1}{1+x+D_1by}\right)-\left(\frac{D_1x}{(1+x+D_1by)^2}\right)\right]+f_2'(x)y\left(1+\frac{D_1x}{1+x+D_1by}\right) & f_2(x)\left(1+\frac{D_1x}{1+x+D_1by}\right)-\left(\frac{f_2(x)D_1^2bxy}{(1+x+D_1by)^2}\right)-1 \end{bmatrix}.$$

Then,

$$J_{E_1}=\begin{pmatrix} -1 & -f_1(1) & 0 \\ 0 & \frac{f_1(1)}{D_1}-1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

As Jacobian matrix J_{E_1} above is an upper triangular, therefore the diagonal is the value of all of its eigenvalues; $\lambda_1=-1$, $\lambda_2=-1$ and $\lambda_3=\frac{f_1(1)}{D_1}-1$. If all the eigenvalues of J_{E1} are negatives, then this leads to the following result.

Theorem 3. *If the eigenvalue λ_3 such that*

$$\lambda_3=\frac{f_1(1)}{D_1}-1<0,$$

then the trivial equilibrium point $E_1(1,0,0)$ is locally asymptotically stable.

(ii) Stability analysis of E_2.

Next, we analyse the local stability of the system (3) that restricted to the neighbourhood of the equilibrium point $E_2\left(\zeta_s,\frac{(1-\zeta_s)}{D_1},0\right)$. The Jacobian matrix at $E_2\left(\zeta_s,\frac{(1-\zeta_s)}{D_1},0\right)$ is given as

$$J_{E_2}=\begin{pmatrix} -\frac{f_1'(\zeta_s)(1-\zeta_s)}{D_1}-1 & -f_1(\zeta_s) & 0 \\ \frac{f_1'(\zeta_s)(1-\zeta_s)}{D_1^2} & \frac{f_1(\zeta_s)}{D_1}-1 & -f_2\left(\frac{1-\zeta_s}{D_1}\right)\left(1+\left(\frac{1-\zeta_s}{1+\left(\frac{1-\zeta_S}{D_1}\right)}\right)\right) \\ 0 & 0 & f_2\left(\frac{1-\zeta_s}{D_1}\right)\left(1+\left(\frac{1-\zeta_s}{1+\left(\frac{1-\zeta_S}{D_1}\right)}\right)\right)-1 \end{pmatrix}.$$

We can see from both Jacobian matrices J_{E1} and J_{E2} that there are no parameter b involved when the predator organisms y is zero. This means that intratrophic predation does not affect the local stability and existence of E_1 and E_2 when no predator organisms involved. The characteristic equation is $\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3 = 0$ where

$$\begin{aligned}c_1 = \frac{1}{D_1(D_1-\zeta_s+1)}\Bigg(&f_1'(\zeta_s)+3D_1-f_1(\zeta_s)+f_1'(\zeta_s)D_1-D_1f_2\left(\frac{1-\zeta_s}{D_1}\right)\\&-2\zeta_sf_1'(\zeta_s)-3D_1\zeta_s-2D_1^2f_2\left(\frac{1-\zeta_s}{D_1}\right)+\zeta_s^2f_1'(\zeta_s)-f_1(\zeta_s)D_1+\zeta_sf_1(\zeta_s)+3D_1^2-D_1\zeta_sf_1'(\zeta_s)\\&+D_1\zeta_sf_2\left(\frac{1-\zeta_s}{D_1}\right)+D_1^2\zeta_sf_2\left(\frac{1-\zeta_s}{D_1}\right)\Bigg),\end{aligned}$$

$$\begin{aligned}c_2 = \frac{1}{D_1(D_1-\zeta_S+1)}\Bigg(&2f_1(\zeta_s)-3D_1-2f_1'(\zeta_s)-2D_1f_1'(\zeta_s)+f_2\left(\frac{1-\zeta_s}{D_1}\right)f_1'(\zeta_s)\\&+2D_1f_2\left(\frac{1-\zeta_s}{D_1}\right)+4\zeta_sf_1'(\zeta_s)+3D_1\zeta_s+4D_1^2f_2\left(\frac{1-\zeta_s}{D_1}\right)-2\zeta_s^2f_1'(\zeta_s)\\&+2D_1f_1(\zeta_s)-f_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)-2\zeta_sf_1(\zeta_s)-3D_1^2-2D_1f_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)\\&+\zeta_sf_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+2D_1f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+2D_1\zeta_sf_1'(\zeta_s)\\&-2\zeta_sf_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)-2D_1\zeta_sf_2\left(\frac{1-\zeta_s}{D_1}\right)+\zeta_s^2f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)\\&-2D_1^2\zeta_sf_2\left(\frac{1-\zeta_s}{D_1}\right)-3D_1\zeta_sf_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+D_1\zeta_s^2f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+D_1\zeta_sf_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)\Bigg),\end{aligned}$$

$$\begin{aligned}c_3 = \frac{1}{D_1(D_1-\zeta_s+1)}\Bigg(&f_1(\zeta_s)-D_1-f_1'(\zeta_s)-D_1f_1'(\zeta_s)+f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+D_1f_2\left(\frac{1-\zeta_s}{D_1}\right)\\&+2\zeta_sf_1'(\zeta_s)+D_1\zeta_s-\zeta_s^2f_1'(\zeta_s)+f_1(\zeta_s)D_1-f_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)-\zeta_sf_1(\zeta_s)\\&-D_1^2-2D_1f_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+\zeta_sf_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+2D_1f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)\\&+D_1\zeta_sf_1'(\zeta_s)-2\zeta_sf_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)-D_1\zeta_sf_2\left(\frac{1-\zeta_s}{D_1}\right)\\&+\zeta_s^2f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)-D_1^2\zeta_sf_2\left(\frac{1-\zeta_s}{D_1}\right)-3D_1\zeta_sf_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)\\&+D_1\zeta_s^2f_1'(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)+D_1\zeta_sf_1(\zeta_s)f_2\left(\frac{1-\zeta_s}{D_1}\right)\Bigg).\end{aligned}$$

Then, by using MATLAB R2015b, we get the eigenvalues of J_{E_2} which are $\lambda_4 = -1$, $\lambda_5 = \frac{f_1'(\zeta_s)(\zeta_s-1)+f_1(\zeta_s)-D_1}{D_1}$ and $\lambda_6 = \frac{f_2\left(\frac{1-\zeta_s}{D_1}\right)(1+2D_1-\zeta_s-D_1\zeta_s)-1}{D_1-\zeta_s+1}$ where $D_1-\zeta_s+1 \neq 0$. We summarise the above discussion in the following theorems.

Theorem 4. *Suppose the assumptions such as*

$$\frac{f_1'(\zeta_s)(\zeta_s-1)+f_1(\zeta_s)-D_1}{D_1}<0 \text{ and } \frac{f_2\left(\frac{1-\zeta_s}{D_1}\right)(1+2D_1-\zeta_s-D_1\zeta_s)-1}{D_1-\zeta_s+1}<0$$

are satisfied, then the equilibrium point, $E_2\left(\zeta_s,\frac{(1-\zeta_s)}{D_1},0\right)$ *is locally asymptotically stable.*

Theorem 5. *If*

$$\frac{f_1'(\zeta_s)(\zeta_s-1)+f_1(\zeta_s)-D_1}{D_1}<0 \text{ and } \frac{f_2\left(\frac{1-\zeta_s}{D_1}\right)(1+2D_1-\zeta_s-D_1\zeta_s)-1}{D_1-\zeta_s+1}<0,$$

then the equilibrium point $E_2\left(\zeta_s,\frac{(1-\zeta_s)}{D_1},0\right)$ *is a hyperbolic saddle and is repels locally in y-direction. Particularly, the dimension of the stable manifold* $W^s(E_2)$ *and the unstable manifold* $W^u(E_2)$ *are given by*

$$\dim W^s(E_2) = 2 \text{ and } \dim W^u(E_2)=1.$$

Proof. The results follow from inspections of the eigenvalues of the matrix J_{E_2} and Theorem 2 (see [16, 18]). ∎

Definition 4. [17] The flow $\mathcal{F}$ will be called uniformly persistent if there exists $\varepsilon_0>0$ such that for all $x\in E^0$, $\lim_{t\to\infty} d(\pi(x,t),\partial E)\geq\varepsilon_0$.

The following theorem shows the existence of the equilibrium point E_2 using persistence analysis.

Theorem 6. *Assume that*

i. *Lemma* 1 *being holds,*

ii. E_2 *is a unique hyperbolic saddle point in* $\mathbb{R}^+_{sxy}$ *and repels locally in y-direction (as in Theorem 5),*

iii. *no existence of periodic orbits in the planes of* $\mathbb{R}^+_{sxy}$.

Then

$$\lim_{t\to+\infty}\inf s(t)>k,\ \lim_{t\to+\infty}\inf x(t)>k,\ \lim_{t\to+\infty}\inf y(t)>k,$$

where $k>0$.

Particularly, the system (3) exhibits uniform persistence and the equilibrium point $E_2\left(\zeta_s,\frac{(1-\zeta_s)}{D_1},0\right)$ *exists.*

Proof. The result follows from the Definition 4, which defines uniform persistence by Butler et al. [17] and Nani and Freedman [18]. ∎

(iii) Stability analysis of $E_3(s^*, x^*, y^*)$.

The Jacobian matrix at E_3 is

$$J_{E_3} = \begin{bmatrix} -f_1'(s^*)x^* - 1 & -f_1(s^*) & 0 \\ \frac{f_1'(s^*)x^*}{D_1} & \frac{f_1(s^*)}{D_1} - f_2(x^*)y^*\left[\left(\frac{D_1}{1+x^*+D_1by^*}\right) - \left(\frac{D_1x^*}{(1+x^*+D_1by^*)^2}\right)\right] - f_2'(x^*)y^*\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right) - 1 & \left(\frac{f_2(x^*)D_1^2bx^*y^*}{(1+x^*+D_1by^*)^2}\right) - f_2(x^*)\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right) \\ 0 & f_2(x^*)y^*\left[\left(\frac{D_1}{1+x^*+D_1by^*}\right) - \left(\frac{D_1x^*}{(1+x^*+D_1by^*)^2}\right)\right] + f_2'(x^*)y^*\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right) & f_2(x^*)\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right) - \left(\frac{f_2(x^*)D_1^2bx^*y^*}{(1+x^*+D_1by^*)^2}\right) - 1 \end{bmatrix}.$$

The eigenvalues of J_{E_3} are resulted from the characteristic equation below

$$\lambda^3 + c_4\lambda^2 + c_5\lambda + c_6 = 0, \tag{6}$$

where

$$c_4 = f_1'(s^*)x^* + Q + R + V - Z + 3 - \frac{f_1(s^*)}{D_1},$$

$$c_5 = f_1'(s^*)x^*(Q + R + V - Z + 2) + \frac{f_1(s^*)(Z - R - 2)}{D_1} + 2Q + 2R + 2V - 2Z + 3,$$

$$c_6 = f_1'(s^*)x^*(Q + R + V - Z + 1) + \frac{f_1(s^*)(Z - R - 1)}{D_1} + Q + R + V - Z + 1,$$

and

$$Q = f_2(x^*)y^*\left[\left(\frac{D_1}{1+x^*+D_1by^*}\right) - \left(\frac{D_1x^*}{(1+x^*+D_1by^*)^2}\right)\right],$$

$$R = \left(\frac{f_2(x^*)D_1^2bx^*y^*}{(1+x^*+D_1by^*)^2}\right),$$

$$V = f_2'(x^*)y^*\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right),$$

$$Z = f_2(x^*)\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right).$$

From (6), we obtain the eigenvalues; $\lambda_7 = -1,$ and λ_8, λ_9 are

$$\lambda_8 = -\frac{1}{2D_1}\left(D_1\left(f_1'(s^*)x^* + Q + R + V - Z + 2\right) - f_1(s^*)\right.$$
$$+\left(\sqrt{D_1\left(\begin{array}{r} Q^2 + 2QR + 2QV - 2QZ - 2Qf_1(s^*) - f_1'(s^*)x^*\left(2Q + 2R + 2V - 2Z - f_1'(s^*)x^*\right) + R^2 + \\ 2RV - 2RZ + V^2 - 2VZ + Z^2 \end{array}\right) + f_1(s^*)\left(2D_1R - 2D_1V - 2D_1Z - f_1'(s^*)x^*\right) + f_1(s^*)^2}\right)$$
$$\left.+f_1'(s^*)D_1x^*\right),$$

$$\lambda_9 = -\frac{1}{2D_1}\left(D_1\left(f_1'(s^*)x^* + Q + R + V - Z + 2\right) - f_1(s^*)\right.$$
$$-\left(\sqrt{D_1\left(\begin{array}{r} Q^2 + 2QR + 2QV - 2QZ - 2Qf_1(s^*) - f_1'(s^*)x^*\left(2Q + 2R + 2V - 2Z - f_1'(s^*)x^*\right) + R^2 + \\ 2RV - 2RZ + V^2 - 2VZ + Z^2 \end{array}\right) + f_1(s^*)\left(2D_1R - 2D_1V - 2D_1Z - f_1'(s^*)x^*\right) + f_1(s^*)^2}\right)$$
$$\left.+f_1'(s^*)D_1x^*\right).$$

These results lead to the following theorem.

Theorem 7. *Suppose that $\lambda_8 < 0$ and $\lambda_9 < 0$, then the equilibrium point E_3 is locally asymptotically stable.*

2.4. Global stability and uniform persistence analysis

Now we will analyse the global stability of the equilibrium point $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right)$ and the existence of positive interior equilibrium point $E_3(s^*, x^*, y^*)$. For global stability analysis, we use the similar technique as in [3, 18].

(i) Global stability analysis of $E_2\left(\zeta_s, \frac{1-\zeta_s}{D_1}, 0\right)$

Consider system (3) restricted to $\mathbb{R}^+_{sx}$ as in (5). Let $\mathcal{N}$ be a neighbourhood of equilibrium point $E_2\left(\zeta_s, \frac{1-\zeta_s}{D_1}, 0\right)$ in $\mathbb{R}^+_{sx}$. To analyse the global stability of the equilibrium point E_2, a suitable Lyapunov function $L = \frac{1}{2}\left(n_1(s-\widehat{s})^2 + n_2(x-\widehat{x})^2\right)$ is used, where $\widehat{s}$ and $\widehat{x}$ denote the components of $E_2(\widehat{s}, \widehat{x})$, i.e., $\widehat{s} = \zeta_s$ and $\widehat{x} = \frac{1-\zeta_s}{D_1}$, while n_1 and n_2 are positive constants. Note that L is a positive definite function with respect to E_2 in $\mathbb{R}^+_{sx}$ and a Lyapunov function for system (5) in $\mathcal{N}$. By differentiating L with respect to time t, we get

$$\frac{dL}{dt} = n_1(s-\widehat{s})s' + n_2(x-\widehat{x})x', \tag{7}$$

where $\widehat{x} = \frac{1-\widehat{S}}{D_1}$. From (5), we have

$$1 = \widehat{s} + \widehat{x} f_1(\widehat{s}) \text{ and } D_1 = f_1(\widehat{s}).$$

Hence (7) can be written as

$$\begin{aligned} L' &= n_1(s-\widehat{s})\big(\widehat{s} + \widehat{x} f_1(\widehat{s}) - s - f_1(s)x\big) + n_2(x-\widehat{x})\left(\frac{f_1(s)}{f_1(\widehat{s})} - 1\right)x \\ &= -n_1(s-\widehat{s})^2 + n_1(s-\widehat{s})\big(f_1(\widehat{s})\widehat{x} - f_1(s)x\big) + n_2(x-\widehat{x})x\left(\frac{f_1(s)}{f_1(\widehat{s})} - 1\right) \\ &= n_{11}(s-\widehat{s})^2 + \frac{1}{2}n_{12}(s-\widehat{s})(x-\widehat{x}) + \frac{1}{2}n_{21}(s-\widehat{s})(x-\widehat{x}) + n_{22}(x-\widehat{x})^2, \end{aligned}$$

where

$$\begin{aligned} n_{11} &= -n_1 < 0, \\ n_{12} = n_{21} &= n_1 \frac{f_1(\widehat{s})\widehat{x} - f_1(s)x}{x-\widehat{x}}, \\ n_{22} &= n_2 \frac{x\big(f_1(s) - f_1(\widehat{s})\big)}{(x-\widehat{x})f_1(\widehat{s})}. \end{aligned}$$

Clearly that L' can be written as $L' = X^T N X,$ which T denotes the transpose and the matrix N is particularly a real, symmetric 2×2 matrix, where X and N can be represented by

$$X = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} s-\widehat{s} \\ x-\widehat{x} \end{pmatrix} \text{ and } N = \begin{pmatrix} n_{11} & \frac{1}{2}n_{12} \\ \frac{1}{2}n_{21} & n_{22} \end{pmatrix}.$$

Thus, it leads to the following theorem.

Theorem 8. *The equilibrium point E_2 is global asymptotically stable with respect to solution trajectories are initiated from* $\text{int}\,\mathbb{R}^+_{sx}$ *if the assumptions $n_{22} < 0$ and $\det N > 0$ are satisfied.*

Proof. By using the Frobenius Theorem in ([18], Lemma 6.2), we can see that n_{22} and $\det N$ are the leading principal minors of the matrix N. It is shown that matrix N is negative definite if

$$n_{22} < 0, \text{ and } \det N = \det \begin{pmatrix} n_{11} & \frac{1}{2}n_{12} \\ \frac{1}{2}n_{21} & n_{22} \end{pmatrix} > 0.$$

This completes the proof of the theorem. ∎

(ii) Existence of positive interior equilibrium point E_3

In this subsection, we present some results of persistence analysis, including uniform persistence and state the necessary conditions for the existence of positive equilibrium point E_3. The following lemma from Ref. [19] is applied to obtain the persistence results.

Lemma 2. [19] Let G be an isolated hyperbolic equilibrium point in the omega limit set, $\omega(X)$ of the orbit $\mathcal{O}(X)$. Then either $\omega(X) = G$ or there exist points J^+ and J^- in $\omega(X)$ with $J^+ \in W^s(G)$ and $J^- \in W^u(G)$, where $W^s(G)$ and $W^u(G)$ denote stable and unstable manifolds of G.

Theorem 9. *Assume that*

i. $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right)$ *is a hyperbolic saddle point and is repels locally in y direction (as in Theorem* 5*),*

ii. *system (3) is dissipative and all solutions with initial values in* $\mathbb{R}^+_{sxy}$ *are uniformly bounded and attracted into region H (as in Theorem* 1*), and*

iii. *equilibrium point* $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right)$ *is globally asymptotically stable with respect to* $\mathbb{R}^+_{sx}$.

Then, the system (3) *is uniformly persistence.*

Proof. This proof strictly depends on Lemma 2. Suppose H is the region as stated in Theorem 1. It showed that region H is positive invariant set and any solutions of system (3) emanating at a point in $\mathbb{R}^3_+$ is uniformly bounded. Despites that, the only compact invariant set on $\partial\mathbb{R}^3_+$ is $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right)$. Let $P = E_3(s^*, x^*, y^*)$ belongs to the interior of $\mathbb{R}^3_+$, i.e., $P \in \operatorname{int}\mathbb{R}^3_+$. We shall show that there is no points $J_i \in \partial\mathbb{R}^3_+$ where $\partial\mathbb{R}^3_+$ belongs to $\omega(P)$, the omega limit set of P. Now, we prove that the equilibrium point $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right) \notin \omega(P)$. Assume that $E_2 \in \omega(P)$ is true. Then, there is a point $J_1^+ \in W^s(E_2)\setminus\{E_2\}$ such that $J_1^+ \in \omega(P)$ by Lemma 2. But $W^s(E_2) \cap \left(\mathbb{R}^3_+ \setminus \{E_2\}\right)$ is empty which is a contradiction for the positive invariance property of $H \subset \mathbb{R}^3_+$. Thus, the equilibrium point E_2 is not in the omega limit set of P; $E_2 \notin \omega(P)$. Next, we shall show $\partial\mathbb{R}^3_+ \cap \omega(P) = \varnothing$. Assume that $\partial\mathbb{R}^3_+ \cap \omega(P) \neq \varnothing$, and let $J \in \partial\mathbb{R}^3_+$ and $J \in \omega(P)$. Then the closure of the orbit of the point J, i.e., $\mathrm{cl}(O(J))$ must either contains E_2 or unbounded. This is a contradiction, and hence it is proved that $\partial\mathbb{R}^3_+ \cap \omega(P) = \varnothing$. We deduce that if $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right)$ is unstable, then, for stable manifold $W^s(E_2)$; $W^s(E_2) \cap \left(\operatorname{int}\left(\mathbb{R}^3_+\right)\right) = \varnothing$, and for unstable manifold $W^u(E_2)$; $W^u(E_2) \cap \left(\operatorname{int}\left(\mathbb{R}^3_+\right)\right) \neq \varnothing$. Therefore, the result of uniform persistence follows since the omega limit set of P, $\omega(P)$ must be in int $\left(\mathbb{R}^3_+\right)$. This completes the proof. ∎

Remark 2. Global stability of equilibrium point $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0\right)$ with respect to $\mathbb{R}^+_{sx}$ indicates that the boundary flow is isolated and a cyclic with respect to region H. Thus, the system (3) undergoes uniform persistence and implies that a positive interior equilibrium point $E_3(s^*, x^*, y^*)$ exists (see [20]).

2.5. Hopf bifurcation

In this section, we investigate Hopf bifurcation on the system (3) with a bifurcation real parameter, σ. Particularly, σ is selected in such a way that the growth rate function f_2 is a function of x and σ. Therefore, system (3) takes of the form

$$\begin{cases} s' = 1 - s - f_1(s)x, \\ x' = \left(\dfrac{f_1(s)}{D_1} - 1\right)x - f_2(x;\sigma)y\left(1 + \dfrac{D_1 x}{1 + x + D_1 by}\right), \\ y' = f_2(x;\sigma)y\left(1 + \dfrac{D_1 x}{1 + x + D_1 by}\right) - y, \end{cases} \tag{8}$$

where $s(0) > 0,\ x(0) > 0,\ y(0) > 0$. Next, we do linearization on the system (8). First, let

$$\begin{cases} S = s - h_1; \\ X = x - h_2; \\ Y = y - h_3; \end{cases} \Rightarrow \begin{cases} s = S + h_1; \\ x = X + h_2; \\ y = Y + h_3; \end{cases}$$

where (h_1, h_2, h_3) is the non-trivial equilibrium point. Then, we obtain the following differential equations

$$\begin{cases} S' = 1 - S - h_1 - f_1(S + h_1)(X + h_2), \\ X' = \left(\dfrac{f_1(S + h_1)}{D_1} - 1\right)(X + h_2) - f_2(X + h_2;\sigma)(Y + h_3)\left(1 + \dfrac{D_1(X + h_2)}{1 + (X + h_2) + D_1 b(Y + h_3)}\right), \\ Y' = f_2(X + h_2;\sigma)(Y + h_3)\left(1 + \dfrac{D_1(X + h_2)}{1 + (X + h_2) + D_1 b(Y + h_3)}\right) - (Y + h_3). \end{cases} \tag{9}$$

The Jacobian matrix of system (8) is given by

$$J_\sigma = \begin{bmatrix} -f_1'(s)x - 1 & -f_1(s) & 0 \\ \dfrac{f_1'(s)x}{D_1} & \dfrac{f_1(s)}{D_1} - f_2(x;\sigma)y\left[\left(\dfrac{D_1}{1 + x + D_1 by}\right) - \left(\dfrac{D_1 x}{(1 + x + D_1 by)^2}\right)\right] - {f_2}'(x;\sigma)y\left(\left(\dfrac{D_1 x}{1 + x + D_1 by}\right) + 1\right) - 1 & f_2(x;\sigma)\left[\left(\dfrac{D_1^2 bxy}{(1 + x + D_1 by)^2}\right) - \left(\dfrac{D_1 x}{1 + x + D_1 by} + 1\right)\right] \\ 0 & f_2(x;\sigma)y\left[\left(\dfrac{D_1}{1 + x + D_1 by}\right) - \left(\dfrac{D_1 x}{(1 + x + D_1 by)^2}\right)\right] + f_2'(x;\sigma)y\left(1 + \dfrac{D_1 x}{1 + x + D_1 by}\right) & f_2(x;\sigma)\left[\left(1 + \dfrac{D_1 x}{1 + x + D_1 by}\right) - \left(\dfrac{D_1^2 bxy}{(1 + x + D_1 by)^2}\right)\right] - 1 \end{bmatrix}.$$

Thus, the Jacobian matrix of system (8) about $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0; \sigma\right)$ is

$$J_{\sigma(E_2)} = \begin{pmatrix} -\frac{f_1'(\zeta_s)(1-\zeta_s)}{D_1} - 1 & -f_1(\zeta_s) & 0 \\ \frac{f_1'(\zeta_s)(1-\zeta_s)}{D_1^2} & \frac{f_1(\zeta_s)}{D_1} - 1 & -f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)\left(1+\left(\frac{1-\zeta_s}{1+\left(\frac{1-\zeta_s}{D_1}\right)}\right)\right) \\ 0 & 0 & f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)\left(1+\left(\frac{1-\zeta_s}{1+\left(\frac{1-\zeta_s}{D_1}\right)}\right)\right) - 1 \end{pmatrix}. \tag{10}$$

The characteristic equation of (10) is given as

$$\lambda^3 + c_1\lambda^2 + c_2\lambda + c_3 = 0, \tag{11}$$

where

$$\begin{aligned} c_1 = {} & \frac{1}{D_1(D_1-\zeta_s+1)}\left(f_1'(\zeta_s)\left(1+D_1-2\zeta_s+\zeta_s^2-D_1\zeta_s\right)+f_1(\zeta_s)(\zeta_s-D_1-1)\right. \\ & \left.+D_1 f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)(\zeta_s+D_1\zeta_s-2D_1-1)+3D_1(1+D_1-\zeta_s)\right), \end{aligned}$$

$$\begin{aligned} c_2 = {} & \frac{1}{D_1(D_1-\zeta_s+1)}\left(f_1'(\zeta_s)\left(2\zeta_s^2+2+2D_1-4\zeta_s-2D_1\zeta_s\right.\right. \\ & \left.-f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)\left(1+2D_1-2\zeta_s+\zeta_s^2-3D_1\zeta_s+D_1\zeta_s^2\right)\right) \\ & -f_1(\zeta_s)\left(2+2D_1-2\zeta_s+f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)(\zeta_s-2D_1-1+D_1\zeta_s)\right) \\ & \left.-f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)\left(2D_1+4D_1^2-2D_1\zeta_s-2D_1^2\zeta_s\right)+3D_1\zeta_s-3D_1^2-3D_1\right), \end{aligned}$$

$$\begin{aligned} c_3 = {} & \frac{1}{D_1(D_1-\zeta_s+1)}\left(f_1'(\zeta_s)\left(1-2\zeta_s+D_1+\zeta_s^2-D_1\zeta_s\right.\right. \\ & \left.-f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)\left(1-2\zeta_s+\zeta_s^2-D_1+D_1\zeta_s^2\right)\right) \\ & -f_1(\zeta_s)\left(f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)(\zeta_s-1-2D_1+D_1\zeta_s)+1-\zeta_s\right) \\ & \left.-f_2\left(\frac{1-\zeta_s}{D_1};\sigma\right)(D_1-D_1\zeta_s)+D_1\zeta_s-D_1-D_1^2\right). \end{aligned}$$

We applied Routh Hurwitz criterion (see [21, 22]) onto the characteristic equation (11) and obtain the matrices

$$M_1 = [c_1]; M_2 = \begin{bmatrix} c_1 & 1 \\ c_3 & c_2 \end{bmatrix}; \; M_3 = \begin{bmatrix} c_1 & 1 & 0 \\ c_3 & c_2 & c_1 \\ 0 & 0 & c_3 \end{bmatrix}.$$

Thus, the characteristic equation (11) has all negative real parts of λ if and only if

$$\left.\begin{aligned} c_1 > 0 \\ c_3 > 0 \\ c_1c_2 - c_3 > 0. \end{aligned}\right\} \tag{12}$$

When the assumptions on functional responses $f_1(s) = f_1(\zeta_s)$ and $f_2(x) = f_2\left(\frac{1-\zeta_s}{D_1}; \sigma\right)$ as in (2) for the system (3), together with the hypotheses H_1 until H_3;

$$H_1 : D_1 - \zeta_s + 1 > 0,$$

$$\begin{aligned} H_2 : f_1'(\zeta_s)\left(1 + D_1 - 2\zeta_s + \zeta_s^2 - D_1\zeta_s\right) + f_1(\zeta_s)(\zeta_s - D_1 - 1) \\ + D_1 f_2\left(\frac{1-\zeta_s}{D_1}; \sigma\right)(\zeta_s + D_1\zeta_s - 2D_1 - 1) + 3D_1(1 + D_1 - \zeta_s) > 0, \end{aligned}$$

$$\begin{aligned} H_3 : f_1'(\zeta_s)\left(1 - 2\zeta_s + D_1 + \zeta_s^2 - D_1\zeta_s - f_2\left(\frac{1-\zeta_s}{D_1}; \sigma\right)\left(1 - 2\zeta_s + \zeta_s^2 - D_1 + D_1\zeta_s^2\right)\right) \\ - f_1(\zeta_s)\left(f_2\left(\frac{1-\zeta_s}{D_1}; \sigma\right)(\zeta_s - 1 - 2D_1 + D_1\zeta_s) + 1 - \zeta_s\right) \\ - f_2\left(\frac{1-\zeta_s}{D_1}; \sigma\right)(D_1 - D_1\zeta_s) + D_1\zeta_s - D_1 - D_1^2 > 0, \end{aligned}$$

hold, we will have $c_1 > 0$ and $c_3 > 0$. We shall obtain two pure imaginary roots for the characteristic equation (11) if and only if $c_1c_2 = c_3$ for some values of σ, say, σ_1^*.

Since at $\sigma = \sigma_1^*$, there exists an interval $\left(\sigma_1^* - \varepsilon, \sigma_1^* + \varepsilon\right)$ containing σ_1^* for some $\varepsilon > 0$ such that $\sigma \in \left(\sigma_1^* - \varepsilon, \sigma_1^* + \varepsilon\right)$. Thus, for $\sigma \in \left(\sigma_1^* - \varepsilon, \sigma_1^* + \varepsilon\right)$, the characteristic equation (11) cannot have positive real roots. For $\sigma = \sigma_1^*$, we acquire (see [3, 10])

$$\left(\lambda^2 + c_2\right)(\lambda + c_1) = 0, \tag{13}$$

that consist of three roots; $\lambda_1 = \sqrt{c_2}i, \lambda_2 = -\sqrt{c_2}i$, and $\lambda_3 = -c_1$. As for $\sigma \in \left(\sigma_1^* - \varepsilon, \sigma_1^* + \varepsilon\right)$, all the roots are in general of the form;

$$\begin{aligned} \lambda_1(\sigma) &= \alpha(\sigma) + \beta(\sigma)i, \\ \lambda_2(\sigma) &= \alpha(\sigma) - \beta(\sigma)i, \\ \lambda_3(\sigma) &= -\mathrm{c}_1(\sigma). \end{aligned}$$

In order to apply the Hopf bifurcation theorem as stated in [9, 23] towards the system (8), we must verify the transversality condition

$$\operatorname{Re}\left[\frac{d\lambda_j}{d\sigma}\right]_{\sigma=\sigma_1^*} \neq 0, \qquad j = 1, 2, 3. \tag{14}$$

By substituting $\lambda_1(\sigma) = \alpha(\sigma) + \beta(\sigma)i$ and $\lambda_2(\sigma) = \alpha(\sigma) - \beta(\sigma)i$ into characteristic equation (11) and calculating the implicit derivative, we obtain the following equations

$$\left.\begin{aligned} K(\sigma)\alpha'(\sigma) - L(\sigma)\beta'(\sigma) + M(\sigma) = 0, \\ L(\sigma)\alpha'(\sigma) + K(\sigma)\beta'(\sigma) + N(\sigma) = 0, \end{aligned}\right\} \tag{15}$$

where

$$K(\sigma) = 3\alpha^2(\sigma) + 2c_1(\sigma)\alpha(\sigma) + c_2(\sigma) - 3\beta^2(\sigma);$$

$$M(\sigma) = \alpha^2(\sigma)c_1'(\sigma) + c_2'(\sigma)\alpha(\sigma) - c_1'(\sigma)\beta^2(\sigma) + c_3'(\sigma);$$

$$L(\sigma) = 6\alpha(\sigma)\beta(\sigma) + 2c_1(\sigma)\beta(\sigma);$$

$$N(\sigma) = 2\alpha(\sigma)\beta(\sigma)c_1'(\sigma) + c_2'(\sigma)\beta(\sigma).$$

Since $K(\sigma_1^*)M(\sigma_1^*) + L(\sigma_1^*)N(\sigma_1^*) \neq 0$, we have

$$\operatorname{Re}\left[\frac{d\lambda_j}{d\sigma}\right]_{\sigma=\sigma_1^*} = \frac{K(\sigma_1^*)M(\sigma_1^*) + L(\sigma_1^*)N(\sigma_1^*)}{K^2(\sigma_1^*) + L^2(\sigma_1^*)} \neq 0, \qquad j = 1, 2, 3,$$

and $\lambda_3(\sigma_1^*) = -c_1(\sigma_1^*) \neq 0$. We conclude the details above in the following theorem.

Theorem 10. *Suppose that the equilibrium point* $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0; \sigma\right)$ *exists and those assumptions similar as in* (2) *for the system* (3) *together with hypothesis* H_1 *until* H_3 *hold. Then the system* (8) *undergoes Hopf bifurcation in the first octant, which leads to a family of periodic solutions bifurcating from* $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0; \sigma\right)$ *for some values of* σ *in the neighbourhood of* σ_1^*.

Next, we determine the Hopf bifurcation at the equilibrium point $E_3(s^*, x^*, y^*; \sigma)$. The Jacobian matrix of the system (8) about the equilibrium point E_3 is given by

$$J_{\sigma(E_3)} = \begin{bmatrix} -f_1'(s^*)x^* - 1 & -f_1(s^*) & 0 \\ \frac{f_1'(s^*)x^*}{D_1} & \frac{f_1(s^*)}{D_1} - f_2(x^*;\sigma)y^*\left[\left(\frac{D_1}{1+x^*+D_1by^*}\right) - \left(\frac{D_1x^*}{(1+x^*+D_1by^*)^2}\right)\right] - {f_2}'(x^*;\sigma)y^*\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right) - 1 & f_2(x^*;\sigma)\left(\frac{D_1^2bx^*y^*}{(1+x^*+D_1by^*)^2} - \frac{D_1x^*}{1+x^*+D_1by^*} - 1\right) \\ 0 & f_2(x^*;\sigma)y^*\left[\left(\frac{D_1}{1+x^*+D_1by^*}\right) - \left(\frac{D_1x^*}{(1+x^*+D_1by^*)^2}\right)\right] + f_2'(x^*;\sigma)y^*\left(\frac{D_1x^*}{1+x^*+D_1by^*}+1\right) & f_2(x^*;\sigma)\left(\frac{D_1x^*}{1+x^*+D_1by^*} - \frac{D_1^2bx^*y^*}{(1+x^*+D_1by^*)^2} + 1\right) - 1 \end{bmatrix}.$$

Hence, the characteristic equation for the Jacobian matrix $J_{\sigma(E_3)}$ is

$$\lambda^3 + c_4\lambda^2 + c_5\lambda + c_6 = 0, \tag{16}$$

where

$$c_4 = f_1'(s^*)x^* + Q + R + V - Z + 3 - \frac{f_1(s^*)}{D_1},$$

$$c_5 = f_1'(s^*)x^*(Q + R + V - Z + 2) + \frac{f_1(s^*)(Z - R - 2)}{D_1} + 2Q + 2R + 2V - 2Z + 3,$$

$$c_6 = f_1'(s^*)x^*(Q + R + V - Z + 1) + \frac{f_1(s^*)(Z - R - 1)}{D_1} + Q + R + V - Z + 1,$$

and

$$Q = f_2(x^*;\sigma)y^*\left[\left(\frac{D_1}{1 + x^* + D_1by^*}\right) - \left(\frac{D_1x^*}{(1 + x^* + D_1by^*)^2}\right)\right],$$

$$R = \left(\frac{f_2(x^*;\sigma)D_1^2bx^*y^*}{(1 + x^* + D_1by^*)^2}\right),$$

$$V = f_2'(x^*;\sigma)y^*\left(\frac{D_1x^*}{1 + x^* + D_1by^*} + 1\right),$$

$$Z = f_2(x^*;\sigma)\left(\frac{D_1x^*}{1 + x^* + D_1by^*} + 1\right).$$

By applying the Routh-Hurwitz criterion (see [21, 22]) towards the characteristic equation (16), we obtain the following Hurwitz matrices

$$M_4 = [c_4];\ M_5 = \begin{bmatrix} c_4 & 1 \\ c_6 & c_5 \end{bmatrix};\ M_6 = \begin{bmatrix} c_4 & 1 & 0 \\ c_6 & c_5 & c_4 \\ 0 & 0 & c_6 \end{bmatrix}.$$

Thus, the characteristic equation (16) has all negative real parts of λ if and only if

$$\left.\begin{aligned} c_4 &> 0 \\ c_6 &> 0 \\ c_4c_5 - c_6 &> 0. \end{aligned}\right\} \tag{17}$$

Suppose the assumptions of functional response $f_1(s^*)$ and $f_2(x^*;\sigma)$ similar as in (2) for the system (3), together with the hypotheses H_4 until H_6;

$$H_4 : f_1'(s^*)x^* + Q + R + V + 3 > \frac{f_1(s^*)}{D_1} + Z,$$
$$H_5 : Q + R + V + 1 > Z,$$
$$H_6 : Z > R + 1,$$

hold, then clearly $c_4 > 0$ and $c_6 > 0$. In particular, we shall have two pure imaginary roots for the characteristic equation (16) if and only if $c_4c_5 = c_6$ for some values of σ, say, σ_2^*. Since at $\sigma = \sigma_2^*$, there exists an interval $(\sigma_2^* - \varepsilon, \sigma_2^* + \varepsilon)$ containing σ_2^* for some $\varepsilon > 0$. Then, for $\sigma \in (\sigma_2^* - \varepsilon, \sigma_2^* + \varepsilon)$, the characteristic Eq. (16) cannot have positive real roots. For $\sigma = \sigma_2^*$, we get (see [3, 10])

$$\left(\lambda^2 + c_5\right)(\lambda + c_4) = 0, \tag{18}$$

Consisting of three roots; $\lambda_1 = \sqrt{c_5}i, \quad \lambda_2 = -\sqrt{c_5}i, \quad \lambda_3 = -c_4$. As for $\sigma \in \left(\sigma_2^* - \varepsilon, \sigma_2^* + \varepsilon\right)$, all roots are in general of the form;

$$\begin{aligned} \lambda_1(\sigma) &= \alpha(\sigma) + \beta(\sigma)i, \\ \lambda_2(\sigma) &= \alpha(\sigma) - \beta(\sigma)i, \\ \lambda_3(\sigma) &= -c_4(\sigma), \end{aligned}$$

To establish Hopf bifurcation towards system (8), we must show that

$$\mathrm{Re}\left[\frac{d\lambda_j}{d\sigma}\right]_{\sigma=\sigma_2^*} \neq 0, \qquad j = 1, 2, 3. \tag{19}$$

By substituting $\lambda_1(\sigma) = \alpha(\sigma) + \beta(\sigma)i$ and $\lambda_2(\sigma) = \alpha(\sigma) - \beta(\sigma)i$ into characteristic equation (18) and calculating the implicit derivative, we get the following equations;

$$\left.\begin{aligned} A_1(\sigma)\alpha'(\sigma) - A_2(\sigma)\beta'(\sigma) + B_1(\sigma) = 0, \\ A_2(\sigma)\alpha'(\sigma) + A_1(\sigma)\beta'(\sigma) + B_2(\sigma) = 0, \end{aligned}\right\} \tag{20}$$

where

$$\begin{aligned} A_1(\sigma) &= 3\alpha^2(\sigma) + 2c_1(\sigma)\alpha(\sigma) + c_2(\sigma) - 3\beta^2(\sigma); \\ A_2(\sigma) &= 6\alpha(\sigma)\beta(\sigma) + 2c_1(\sigma)\beta(\sigma); \\ B_1(\sigma) &= \alpha^2(\sigma)c_1'(\sigma) + c_2'(\sigma)\alpha(\sigma) - c_1'(\sigma)\beta^2(\sigma) + c_3'(\sigma); \\ B_2(\sigma) &= 2\alpha(\sigma)\beta(\sigma)c_1'(\sigma) + c_2'(\sigma)\beta(\sigma). \end{aligned}$$

Since $A_1\left(\sigma_2^*\right)B_1\left(\sigma_2^*\right) + A_2\left(\sigma_2^*\right)B_2\left(\sigma_2^*\right) \neq 0$, we have

$$\mathrm{Re}\left[\frac{d\lambda_j}{d\sigma}\right]_{\sigma=\sigma_2^*} = \frac{A_1\left(\sigma_2^*\right)B_1\left(\sigma_2^*\right) + A_2\left(\sigma_2^*\right)B_2\left(\sigma_2^*\right)}{\left(A_1\left(\sigma_2^*\right)\right)^2 + \left(A_2\left(\sigma_2^*\right)\right)^2} \neq 0, \qquad j = 1, 2, 3,$$

and $\lambda_3\left(\sigma_2^*\right) = -c_4\left(\sigma_2^*\right) \neq 0$. We summarise the discussion above in the following theorem.

Theorem 11. *Suppose that the equilibrium point*

$$E_3(s^*, x^*, y^*; \sigma) = E_3\left(s^*, \frac{1 - s^*}{f_1(s^*)}, \frac{(2x^* + 1)f_2(x^*) - x^* - 1}{bD_1\left(1 - f_2(x^*)\right)}; \sigma\right)$$

exists and those assumptions similar as in (2) *for the system* (3) *together with hypotheses* H_4 *until* H_6 *hold. Then the system* (8) *undergoes Hopf bifurcation in the first octant, which leads to a family of periodic solutions bifurcating from* $E_3(s^*, x^*, y^*; \sigma)$ *for some values of* σ *in the neighbourhood of* σ_2^*.

2.6. Discussion

We have proposed and analysed a simple nutrient-predator-prey model in a chemostat with intratrophic predation. This system consisted of the nutrient s, prey organisms x and predator organisms y. We conclude that intratrophic predation denoted as b does not affected the local stability and existence of E_1 and E_2 when no predator organisms involved. Next, we have

shown that the washout equilibrium point $E_1(1,0,0)$ (no prey and predator organisms present) is locally asymptotically stable if $\lambda_3 = \frac{f_1(1)}{D_1} - 1 < 0$ holds. For stability of E_2 and E_3 of the system (3), some sufficient criteria or conditions are derived and satisfied. The points E_2 and E_3 are said to be asymptotically stable if all of their eigenvalues are less than zero.

In particular, we investigated the global stability analysis for E_2. A suitable Lyapunov function L is defined and E_2 is globally asymptotically stable if and only if the conditions in Theorem 8 holds. Global stability of E_2 indicates that predator organisms y might be washout in the chemostat despites the initial prey and predator organisms' density levels. Next, in the study of the existence of positive interior equilibrium point E_3, we presented some results regarding uniform persistence analysis. It has shown that the system (3) is uniformly persistence and thus, the positive interior equilibrium point E_3 exists.

In the analysis of occurrence of Hopf bifurcation, Hopf bifurcation theorems in Hassard et al. [9] are applied. We have shown that the system (8) undergoes Hopf bifurcation in the first octant, which leads to a family of periodic solutions bifurcating from $E_2\left(\zeta_s, \frac{(1-\zeta_s)}{D_1}, 0; \sigma\right)$ and $E_3(s^*, x^*, y^*; \sigma)$ for some values of σ in the neighbourhoods of σ_1^* and σ_2^*, respectively. The method used to obtain these results is similar to the method used in [3, 10].

Author details

Zabidin Salleh* and Liyana Abd Rahim

*Address all correspondence to: zabidin@umt.edu.my

School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, Kuala Nerus, Terengganu, Malaysia

References

[1] Smith HL, Waltman P. The Theory of the Chemostat: Dynamics of Microbial Competition. New York: Cambridge University Press; 1995. pp. 1-7

[2] Li B, Kuang Y. Simple food chain in a chemostat with distinct removal rates. Journal of Mathematical Analysis and Applications. 2000;**242**:75-92. DOI: 10.1006 jmaa.1999.6655

[3] El-Sheikh MMA, Mahrouf SAA. Stability and bifurcation of a simple food chain in a chemostat with removal rates. Chaos, Solitons & Fractals. 2005;**23**:1475-1489. DOI: 10.1016/j.chaos.2004.06.079

[4] Hamra MA, Yadi K. Asymptotic behavior of a chemostat model with constant recycle sludge concentration. Acta Biotheoretica. 2017;**65**:233-252. DOI: 10.1007/s10441-017-9309-4

[5] Yang J, Tang G, Tang S. Modelling the regulatory system of a chemostat model with a threshold window. Mathematics and Computers in Simulation. 2017;**132**:220-235. DOI: 10.1016/j.matcom.2016.08.005

[6] Pitchford J, Brindley J. Intratrophic predation in simple predator-prey models. Bulletin of Mathematical Biology. 1998;**60**:937-953; bu980053

[7] Mada Sanjaya WS, Salleh Z, Mamat M. Mathematical model of three species food chain with Holling Type-III functional response. International Journal of Pure and Applied Mathematics. 2013;**89**(5):647-657. DOI: 10.12737/ijpam.v89i5.1

[8] Tee LS, Salleh Z. Hopf bifurcation of a nonlinear system derived from Lorenz system using normal form theory. International Journal of Applied Mathematics and Statistics. 2016;**55**(3):122-132

[9] Hassard BD, Kazarinoff ND, Wan YH. Theory and Applications of Hopf Bifurcation. USA: Cambridge University Press; 1981

[10] Freedman HI, Ruan S. Hopf bifurcation in three-species food chain models with group defense. Mathematical Biosciences. 1992;**111**:73-87

[11] Rana SMS, Nasrin F, Hossain MI. Dynamics of a predator-prey interaction in chemostat with variable yield. Journal of Sustainability Science and Management. 2015;**10**(2):16-23

[12] Solomon ME. The natural control of animal populations. The Journal of Animal Ecology. 1949;**18**:1-32

[13] Hale JK, Kocak H. Dynamics and Bifurcations. New York: Springer; 1991

[14] Robinson CR. An Introduction to Dynamical Systems: Continuous and Discrete. Pearson Prentice Hall, New Jersey: Pearson Education, Inc; 2004

[15] Hirsch MW, Smale S, Devaney RL. Differential Equations, Dynamical Systems & an Introduction to Chaos. 2nd ed. Elsevier Academic Press, San Diego California: Elsevier; 2004

[16] Freedman HI, Mathsen RM. Persistence in predator-prey systems with ratio-dependent predator influence. Bulletin of Mathematical Biology. 1993;**55**(4):817-827

[17] Butler GJ, Hsu SB, Waltman P. Coexistence of competing predators in a chemostat. Journal of Mathematical Biology. 1983;**17**:133-151

[18] Nani F, Freedman HI. A mathematical model of cancer treatment by immunotherapy. Mathematical Biosciences. 2000;**163**:159-199

[19] Freedman HI, Waltman P. Persistence in models of three interacting predator-prey populations. Mathematical Biosciences. 1984;**68**:213-231

[20] Freedman HI, Ruan S. Uniform persistence in functional differential equations. Mathematical Biosciences. 1995;**115**:173-192

[21] Gantmacher FR. The Theory of Matrices. USA: Chelsea Publishing Company; 1964

[22] Tee LS, Salleh Z. Hopf bifurcation analysis of Zhou system. International Journal of Pure and Applied Mathematics. 2015;**104**(1):1-18

[23] Marsden JE, McKracken M. The Hopf Bifurcation and its Applications. New York: Springer-Verlag; 1976

Chapter 4

Sensitivity Analysis: A Useful Tool for Bifurcation Analysis

Raheem Gul and Stefan Bernhard

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.72345

Abstract

Sensitivity analysis and bifurcation analysis are closely related to each other. In sensitivity analysis, especially global sensitivity analysis the effects of input parameter spaces on output quantities of interest are studied. On the other hand, in bifurcation analysis the critical points within feasible regions of parameters are detected where the long-term dynamics changes qualitatively. Prior to bifurcation analysis, it is important to identify the bifurcation parameters. In complex and computationally expensive problems which consist plenty of uncertain parameters, it is essential to find a set of bifurcation parameters before bifurcation analysis. Global sensitivity analysis is a powerful tool to identify the bifurcation parameters which contribute most on output uncertainty. Global sensitivity analysis is the first step toward bifurcation analysis which helps in dimension reduction during bifurcation analysis. As an example, in this chapter, a multi compartment, lumped-parameter model of an arm artery is considered and global sensitivity analysis (Sobol's method) is applied to identify the bifurcation parameters of the arm arteries.

Keywords: lumped parameter model, arm arteries, sensitivity analysis, bifurcation analysis, bifurcation parameters, Sobol's method

1. Introduction

Sensitivity analysis and bifurcation analysis are closely related to each other. In sensitivity analysis, we study how the uncertainty in the output of a mathematical model or system (numerical or otherwise) can be apportioned to different sources of uncertainty in its inputs [1]. On the other hand, in bifurcation analysis the critical points within the feasible regions of parameters are detected where the long-term dynamics changes qualitatively [2]. Prior to the bifurcation analysis, it is important to identify the bifurcation parameters in complex and

computationally expensive problems that consist plenty of uncertain parameters. Sensitivity analysis is a powerful tool to identify the bifurcation parameters which contribute most on output uncertainty. Also, sensitivity analysis helps in dimension reduction during the bifurcation analysis by fixing less influential parameters on their nominal values.

Sensitivity analysis can be divided into two categories, local sensitivity analysis (LSA) and global sensitivity analysis (GSA). In LSA a parameter value is perturbed around its nominal values at a time, keeping other parameters fixed on their nominal values. The procedure is repeated for all parameters one by one to study their impact on output variables. LSA techniques are simple, easy to implement and computationally less expensive. On the other hand, LSA is not suitable for non-linear models and does not explore the impact of entire parameter spaces and their interactions effects on output variables [3, 4]. In order to overcome the limitations associated with the LSA, GSA can be used. In GSA, the analysis is performed over entire feasible regions of the input parameters and quantifies the impact of parameter interactions on output variables. The only deficiency related to the GSA is its computational cost [5–12] (**Figure 1**).

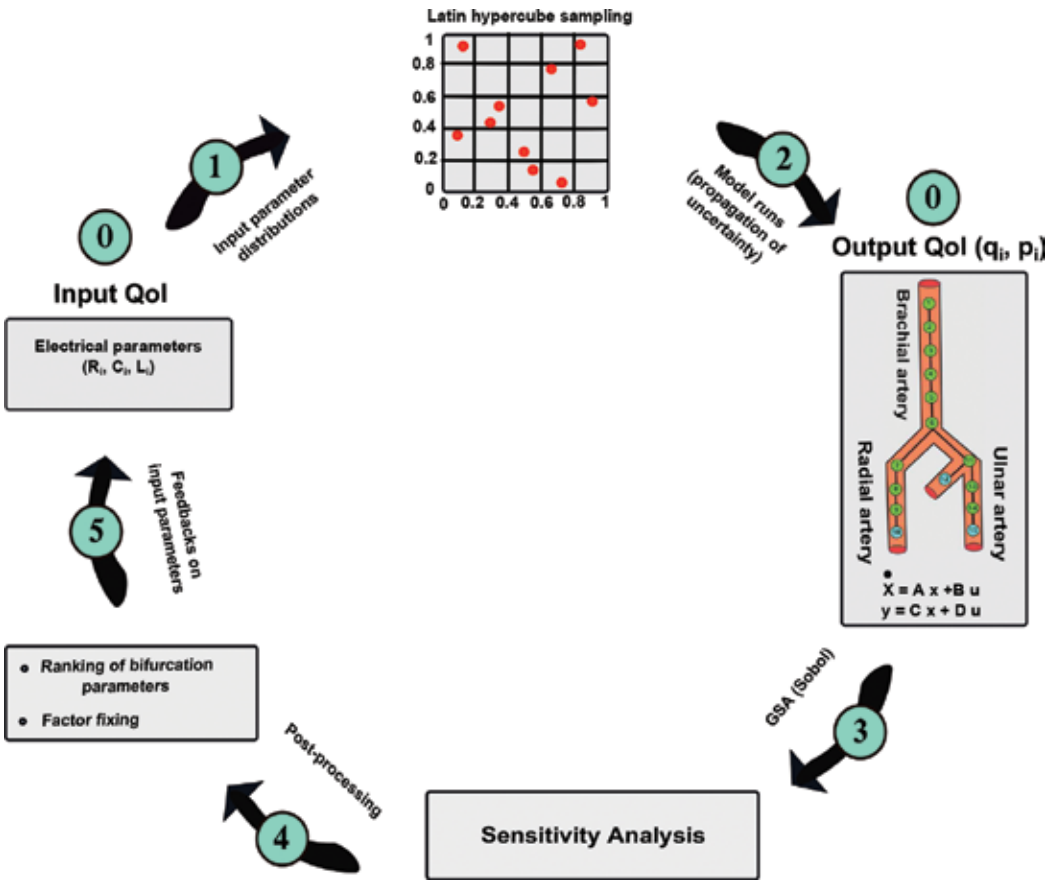

Figure 1. A simplified 5-step procedure to identify the bifurcation parameters using global sensitivity analysis.

In this chapter, the main questions of interest are:

1. How to identify the bifurcation parameters in a model having plenty of input parameters?

2. Which parameters could be exempted from the bifurcation analysis (dimension reduction)?

This chapter seeks to answers the above-mentioned questions using a simplified 5-steps procedure of uncertainty and sensitivity analysis. As an example, a multi-compartment, lumped-parameter model of arm arteries is considered [4] and global sensitivity analysis (Sobol's method) is applied to identify the bifurcation parameters (electrical) of the arm arteries.

2. Lumped-parameter model of the arm arteries

In this section, the major arteries of the arm are divided in to number of non-terminal and terminal arterial segments (nodes). The total number of arterial segments, $N_s = 15$ including 12 non-terminal and 3 terminal segments. Each non-terminal and terminal arterial segment is represented by its corresponding non-terminal and terminal electrical circuit.

Applying Kirchhoff's current and voltage laws on electrical representation of arm arteries, the following mathematical equations for pressure and flow are obtained:

Pressure and flow equations at non-terminal nodes:

Flow equation:

$$\begin{aligned} \dot{q}_i &= \frac{p_{i-1} - p_i - R_i q_i}{L_i}, \quad i = 1, 2, 3, \ldots 15 \text{ and } i \neq 11, 13 \\ \dot{q}_{11} &= \frac{p_6 - p_{11} - R_{11} q_{11}}{L_{11}} \\ \dot{q}_{13} &= \frac{p_{11} - p_{13} - R_{13} q_{13}}{L_{13}} \end{aligned} \tag{1}$$

Pressure equation:

$$\begin{aligned} \dot{p}_i &= \frac{q_i - q_{i+1}}{C_i}, \quad i = 1, 2, 3, \ldots 15 \text{ and } i \neq 6, 11 \\ \dot{p}_6 &= \frac{q_6 - q_{11} - q_7}{C_6}, \quad \dot{p}_{11} = \frac{q_{11} - q_{12} - q_{13}}{C_{11}} \qquad (\text{at bifurcation}) \end{aligned} \tag{2}$$

Pressure and flow equations at terminal nodes:

$$\begin{aligned} \dot{q}_{in} &= \frac{2p_{in} - 2p_i - R_i q_{in}}{L_i} \\ \dot{p}_i &= \frac{q_{in} - q_{out}}{C_i} \\ \dot{q}_{out} &= \frac{2p_i - 2p_{out} - 2R_b q_{out}}{L_i}, \quad i = 10, 12, 15 \end{aligned} \tag{3}$$

Nodes	E	l	d	h	R	C	L
units	$kgm^{-2}s^{-2} \times 10^5$	$m \times 10^{-2}$	$m \times 10^{-3}$	$m \times 10^{-4}$	$kgs^{-1}m^{-4} \times 10^6$	$kg^{-1}s^2m^4 \times 10^{-11}$	$kgm^{-4} \times 10^6$
1	4	6.1	7.28	6.2	3.539	7.454	1.539
2	4	5.6	6.28	5.7	5.868	4.778	1.898
3	4	6.3	5.64	5.5	10.15	4.035	2.648
4	4	6.3	5.32	5.3	12.82	3.514	2.976
5	4	6.3	5	5.2	16.43	2.974	3.369
6	4	4.6	4.72	5	15.10	1.9	2.76
7	8	7.1	3.48	4.4	78.90	0.667	7.838
8	8	7.1	3.24	4.3	105	0.531	9.042
9	8	7.1	3	4.2	142.9	0.448	10.55
10	8	2	2.84	4.1	55.11	0.1207	3.647
11	8	2	4.3	4.9	31.94	1.067	4.844
12	16	6.7	1.82	2.8	1173	0.0834	31.88
13	8	7.9	4.06	4.7	40.19	0.9366	5.434
14	8	6.7	3.48	4.6	50.22	0.80	6.075
15	8	3.7	3.66	4.5	33.60	0.3958	3.693

The value of boundary resistance (R_b) on three terminal nodes is $3.24 \times 10^9\ kgs^{-1}m^{-4}$, $\nu = 0.004\ kgs^{-1}m^{-1}$ and $\rho = 1050\ kgm^{-3}$ [4, 5, 13, 14].

Table 1. Numerical values of parameters for each node of the arm arteries (shown in **Figure 2**).

where, R_i, C_i and L_i is the blood flow resistance, compliance of the vessel and blood inertia of i^{th} segment of the arm arteries respectively. The electrical parameters (R_i, C_i, L_i) of i^{th} segments are related with structural parameters (E_i, l_i, d_i, h_i) as,

$$R_i = \frac{8\nu l_i}{\pi\left(\frac{d}{2}\right)^4}, \quad C_i = \frac{\rho l_i}{\pi\left(\frac{d}{2}\right)^2}, \quad L_i = \frac{2\pi\left(\frac{d}{2}\right)^2 l_i}{E_i l_i} \tag{4}$$

where, E_i is the Young modulus, l_i denotes length of the vessel, d is the diameter of the vessel and h_i represents the wall thickness of i^{th} segment of the vessel. Moreover, ν (0.004 Pa s) is the blood viscosity and ρ (1050 kgm^{-3}) is the blood density. The nominal values of all parameters of arm segments are given in **Table 1**. The geometry along with the values of the parameters is taken from [13, 14].

3. Uncertainty and sensitivity analysis

Uncertainty analysis (UA) and sensitivity analysis (SA) are closely related; however they represent two different disciplines. Uncertainty analysis assesses the uncertainty in model

http://dx.doi.org/10.5772/intechopen.72345

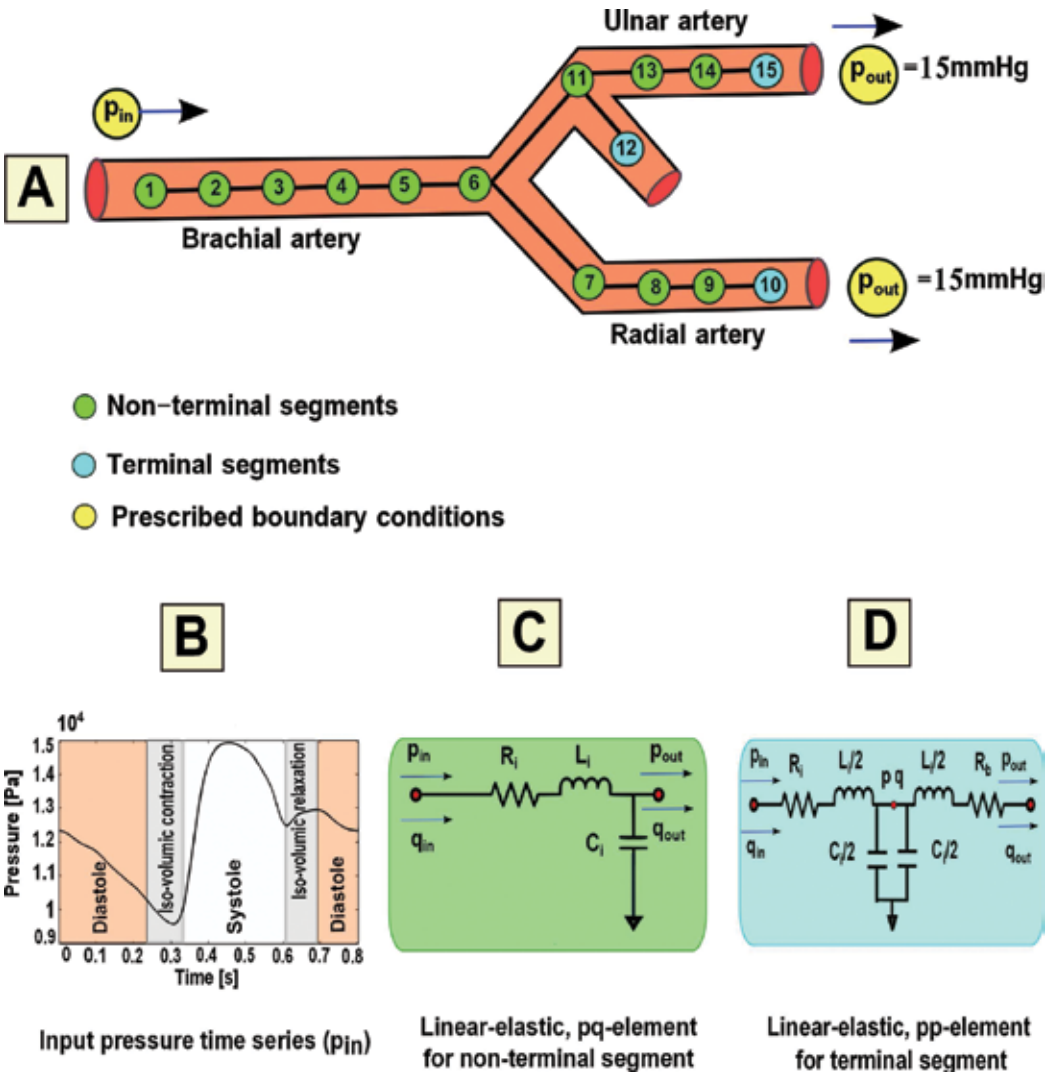

Figure 2. Model geometry of arm artery (A), with total number of arterial segments, Ns = 15, including 12 non-terminal and 3 terminal segments. Each non-terminal and terminal segment is represented by its corresponding non-terminal (C) and terminal electrical circuits (D). Pressure waves is used as an input boundary condition (B) and pout = 15 mmHg which is mean venous pressure used to calculate boundary outflow. The parameter values of each arterial segment are given in **Table 1**.

outputs caused by uncertainty of its inputs. Whereas, sensitivity analysis study the impact of input quantities of interest (QoI) on output quantities of interest (QoI). In this study, the input (QoI) are electrical parameters (R_i, C_i, L_i) and output (QoI) are pressure and flow at each node of the arm arteries. Further, for uncertainty analysis Latin hypercube sampling (LHS) is used and variance-decomposition method (Sobol's method) is used for global sensitivity analysis (GSA).

Compared to the high-dimensional cardiovascular models (3D, 2D, 1D), lumped-parameter models of the cardiovascular system (CVS) are computationally less expensive, therefore they are suitable for GSA. In our previous studies, we found that for lumped-parameter models of

the CVS, the Sobol's method is computationally less expensive as compared to the other variance-decomposition methods, like FAST and sparse grid stochastic collocation method based on Smolyak algorithm [5, 12].

3.1. The method of Sobol

The method of Sobol is the variance-decomposition method used for global sensitivity analysis. The method decomposes the output variance of a system or model into fractions and assigns them to the inputs factors. For example, given a model of the form $Y = f(X) = f(x_1, x_2, \ldots, x_k)$, where X is the vector of K uncertain parameters, which are independently generated within a unit hypercube i.e. $x_i \in [0,1]^k$ for $i = 1, 2, 3, \ldots, K$. Compared to the other GSA methods, the Sobol's method is one of the most commonly used variance-decomposition method, because of its ease of implementation. The method is primarily based on the decomposition of output Y into summands of elementary functions in terms of increasing dimensionality [1, 8],

$$f(x_1, x_2, \ldots, x_k) = f_0 + \sum_{i}^{k} f_i(x_i) + \sum_{i}^{k} \sum_{i<j}^{k} f_{ij}(x_i, x_j) + \ldots + f_{1,2,3,\ldots,k}(x_1, x_2, x_3, \ldots, x_k) \tag{5}$$

In Eq. (5), f is integrable, f_0 is a constant, f_i is a function of x_i, f_{ij} is a function of x_i and x_j and so on. Furthermore, all the terms in the functional decomposition are orthogonal, which leads toward the following definitions of the terms of the functional decomposition in term of conditional expected values.

$$\begin{aligned} f_0 &= E(Y) \\ f_i(x_i) &= E_{x\sim i}(Y|x_i) - f_0 \\ f_{ij}(x_i, x_j) &= E_{x\sim ij}(Y|x_i, x_j) - f_0 - f_i - f_j \\ &\ldots \end{aligned} \tag{6}$$

where, E describes the mathematical expectation and $x_{\sim i}$ denotes all parameters except x_i and so on. The total unconditional variance can be obtained by,

$$V = \int_{\Omega^K} f^2(X)dx - f_0^2 \tag{7}$$

From Eq. (7), the total unconditional variance can be decomposed in a similar manner like in Eq. (5) as,

$$V = \sum_{i}^{k} V_i(x_i) + \sum_{i}^{k} \sum_{i<j}^{k} V_{ij}(x_i, x_j) + \ldots + f_{1,2,3,\ldots,k}(x_1, x_2, x_3, \ldots, x_k) \tag{8}$$

where, V is the variance operator. The relationship between functions and partial variance are given by,

$$\begin{aligned} V_i &= V_{x_i}(E_{x\sim i}(Y|x_i)) = V\big(f_i(x_i)\big) \\ V_{ij} &= V_{x_i,x_j}\big(E_{x\sim ij}\big(Y|x_i,x_j\big)\big) - V_i - V = V\Big(f_{ij}\big(x_i,x_j\big)\Big) \\ &\dots \end{aligned} \tag{9}$$

Dividing both sides of the Eq. (8) by V, we get:

$$1 = \sum_{i}^{k} S_i(x_i) + \sum_{i}^{k}\sum_{i<j}^{k} S_{ij}\big(x_i,x_j\big) + \dots + S_{1,2,3,\dots,K}(x_1,x_2,x_3,\dots,x_K) \tag{10}$$

Where,

$$\begin{aligned} S_i &= \frac{V_{ij}}{V}, \text{ and} \\ S_{ij} &= \frac{V_{ij}}{V} \end{aligned} \tag{11}$$

where, S_i is the main effect (first order sensitivity index) of the i^{th} parameter on output uncertainty and S_{ij} is the interaction effect of i^{th} and j^{th} parameters on output uncertainty. Further, the total sensitivity index, S_{T_i} can be calculated as,

$$S_{T_i} = \frac{E_{x\sim i}\big(V_{x_i}(Y|x\sim i)\big)}{V} = 1 - \frac{V_{x\sim i}\big(E_{x_i}(Y|x\sim i)\big)}{V} \tag{12}$$

In general, the main effect is used identify the most influential parameters (bifurcation parameters) and the total effect is taken into account for those parameters which are exempted from bifurcation analysis (factor fixing). The total effect, S_{T_i} of the i^{th} parameter means main effect plus higher-order effect due to interactions of the i^{th} parameter. In this study, the interaction effects of parameters on the output (QoI) are negligible, therefore the main effects are used for factor fixing and ranking of bifurcation parameters.

3.2. Algorithm to compute sensitivity indices

In this section, a detailed working algorithm is presented to compute the main effect, S_i using the Monte Carlo simulations, we follow the steps, given in [1, 15].

1. Generate a random numbers matrix of row dimension $2K$ and column length N (the sample size) and split into two independent sampling matrices, $A(\mathrm{N},K)$ and $B(\mathrm{N},K)$by using LHS. Where, K is the number of uncertain model parameters.

$$A(\mathrm{N},K)=\begin{bmatrix} x_{11} & x_{12} & \dots & x_{1K} \\ x_{21} & x_{22} & \dots & x_{2K} \\ \dots & \dots & \dots & \dots \\ x_{\mathrm{N}1} & x_{\mathrm{N}2} & \dots & x_{\mathrm{N}K} \end{bmatrix} \tag{13}$$

$$B(\mathrm{N},K)=\begin{bmatrix} x_{1(K+1)} & x_{1(K+2)} & \dots & x_{1(2K)} \\ x_{2(K+1)} & x_{2(K+2)} & \dots & x_{2(2K)} \\ \dots & \dots & \dots & \dots \\ x_{\mathrm{N}(K+1)} & x_{\mathrm{N}(K+2)} & \dots & x_{\mathrm{N}(2K)} \end{bmatrix} \tag{14}$$

2. Define matrix C_i, which is matrix A except the i^{th} column of matrix B.

$$C_i(\mathrm{N},K)=\begin{bmatrix} x_{11} & x_{12} & \dots & x_1(K+i) & \dots & x_{1K} \\ x_{21} & x_{22} & \dots & x_2(K+i) & \dots & x_{2K} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{\mathrm{N}1} & x_{\mathrm{N}2} & \dots & x_{\mathrm{N}}(K+i) & \dots & x_{\mathrm{N}K} \end{bmatrix} \tag{15}$$

3. Compute and save model runs for all parameter spaces using matrices A, B and C_i i.e. $Y_A(t,T_s,\mathrm{N})=f(A)$, $Y_B(t,T_s,\mathrm{N})=f(B)$ and $Y_{C_i}(t,T_s,\mathrm{N},K)=f(C_i)$, where, t are the time points for one heart beat with period $t_p=0.8s$, T_s represents the state variables (pressure and flow time series at six locations of arm artery $(\mathrm{N}_{T_s}=15)$ and Nis the total number of model runs $(\mathrm{N}=4000)$.

4. For the time dependent model outputs, we compute the time dependent main sensitivity index, of each parameter at each time-point of the pressure and flow waves, using the estimator offered by Jansen [15–17].

$$\begin{aligned} S_{t_i}=\frac{V_i}{V}=\frac{V_{xi}(E_{x\sim i}(Y|x_i))}{V}&=\frac{V-\frac{1}{2\mathrm{N}}\sum_{n=1}^{\mathrm{N}}\left(Y_B^{(n)}-Y_{C_i}^{(n)}\right)^2}{V} \\ &=1-\frac{1}{2\mathrm{N}\times V}\sum_{n=1}^{\mathrm{N}}\left(Y_B^{(n)}-Y_{C_i}^{(n)}\right)^2 \end{aligned} \tag{16}$$

where,

$$V=\frac{1}{\mathrm{N}}\sum_{n=1}^{\mathrm{N}}\left(Y_B^{(n)}\right)^2-E^2 \tag{17}$$

and

$$E=\left(\frac{1}{\mathrm{N}}\sum_{n=1}^{\mathrm{N}}Y_B^{(n)}\right)^2 \tag{18}$$

The total variance (V) and the expectation (E) are also calculated at each time-point of pressure and flow waves with respect to each parameter.

5. Finally, the main effect, Si of each parameter on the state variables is calculated.

$$S_i = \frac{1}{\mathrm{N}_{T_s}} \frac{1}{\mathrm{N}_t} \sum_{j=1}^{\mathrm{N}_{T_s}} \sum_{t=0}^{\mathrm{N}_t} S_{t_i}(t, j, t), i = 1, 2, \ldots, K \tag{19}$$

In Eq. (19), N_{T_s} is the number of output variables (pressure and flow time series at all locations) and N_t is the number of time-points [12].

3.3. Input parameters distribution

The results of the UA and SA are greatly affected by the choice of input parameters distributions. In principle, the parameters distributions should be estimated using medical data. Unfortunately, the medical data is not easy to obtained. The input parameters distributions could be chosen according to the expert opinion or using the data from the literature. Due to limited data availability, here in this work the input parameters are randomized within ±10% range of their base (nominal) values using Latin hypercube sampling (LHS).

3.4. Convergence of sensitivity indices

The method of Sobol requires $\mathrm{N}(K + 2)$ number of model simulations to compute S_i. The main effect, S_i is computed for N = [500, 1000, 2000, 3000, 4000] model runs. It is observed that, when the total number of simulations run N increases from 3000 then the sensitivity indices (S_i) become stable [18]. Therefore, the minimum number of simulations for each parameter to achieve convergence of sensitivity indices is around 3000.

4. Results and discussion

In this section, the sensitivity results based on main effect S_i are presented. In order to calculate sensitivity time series, the method of Sobol is applied on each time point of the output QoI i.e. pressure and flow waves at each location of the arm arteries. For each parameter, there are two sensitivity time series at each segment of the arm arteries, one for the pressure and one for the flow. In total, $K \times \mathrm{N}_{T_s} = 45 \times 33 = 1485$ sensitivity time series are obtained. In order to represent the sensitivity results in a compact way, mean absolute values of each pressure and flow sensitivity time series per parameter is taken. In this way, a matrix of dimension 45×33 is acquired, where each entry of the matrix represents the mean absolute values of pressure and flow sensitivity time series per parameter, see **Figure 3**. The numbers in the boxes show the impact (%) on the output (pressure and flow) when input parameters (R_i, C_i, L_i) are randomized within the feasible ranges of ±10%. The parameters having main effect, $S_i > 10\%$ on output QoI are not shown in the **Figure 3**. Each row in **Figure 3** represents the ranking of influential (bifurcation) parameters. For convenience, the electrical parameters $(R_i, C_i, L_i), i = 1, 2, 3, \ldots, 15$ that have impact greater than 10% on pressure and flow are considered as bifurcation parameters which further can be used in bifurcation analysis. For example, for pressure

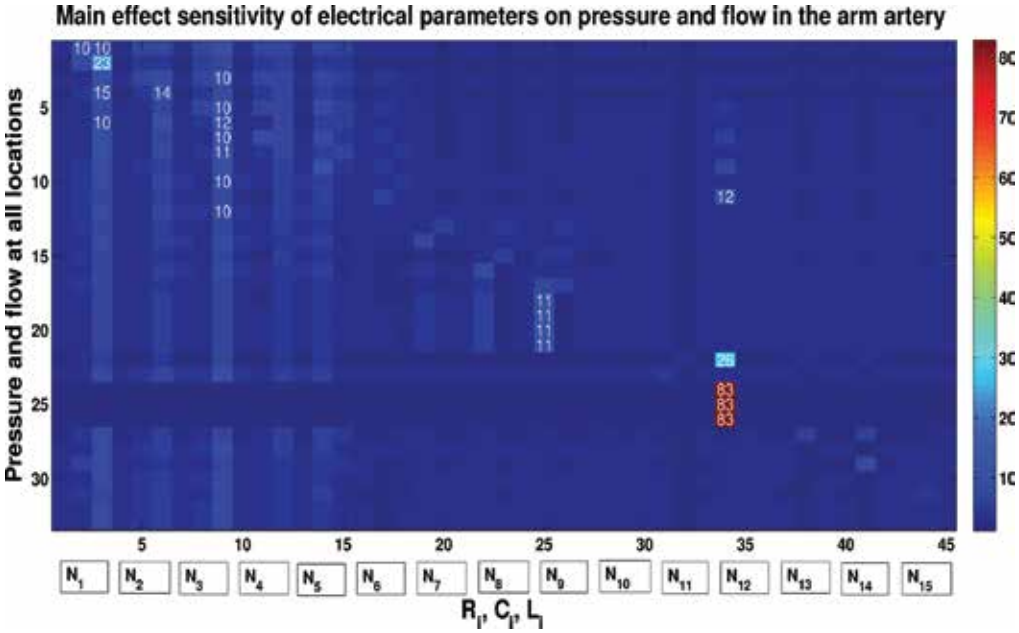

Figure 3. Main effect sensitivity of 45-electrical parameters (R_i, C_i, L_i) on pressure and flow time series (p_i, q_i) at 15-segments $(N_1, N_2, \ldots, N_{15})$ of the arm arteries with number of simulations run per parameter, N is 4000. In total N(K + 2) = 4000(45 + 2) = 188,000 = 0.188 million of simulations run are required to compute the main and total sensitivity indices. The total time taken to compute the sensitivity indices is approximately 3 hours.

at node-2, L_1 and L_2 are the bifurcation parameters, see in **Figure 4** (top). Whereas, for flow at node -2, L_1 and L_2 are considered as bifurcation parameters, see **Figure 4** (bottom).

In a similar fashion, each row of **Figure 3** represents the ranking of bifurcation parameters which further can be used in bifurcation analysis. The parameters which have main effect $S_i < 10\%$ can be exempted from the bifurcation analysis. The criteria for factor fixing vary from problem to problem.

5. Conclusion

In this chapter, a 5-step procedure of global sensitivity analysis is presented to identify the bifurcation parameters in a lumped-parameter model of the arm arteries. Moreover, the proposed procedure can be applied on any morphology or structure of the systemic circulation (carotid bifurcation, aorta or complete systemic circulation). The results of sensitivity analysis are useful to identify and rank the bifurcation parameters, as well as help which parameters could be exempted from the bifurcation analysis. In this particular example of the arm arteries, 23 out of 45 parameters can be excluded from the bifurcation analysis. Whereas, 22 identified as bifurcation parameters, which further can be used/studied in the bifurcation analysis.

http://dx.doi.org/10.5772/intechopen.72345

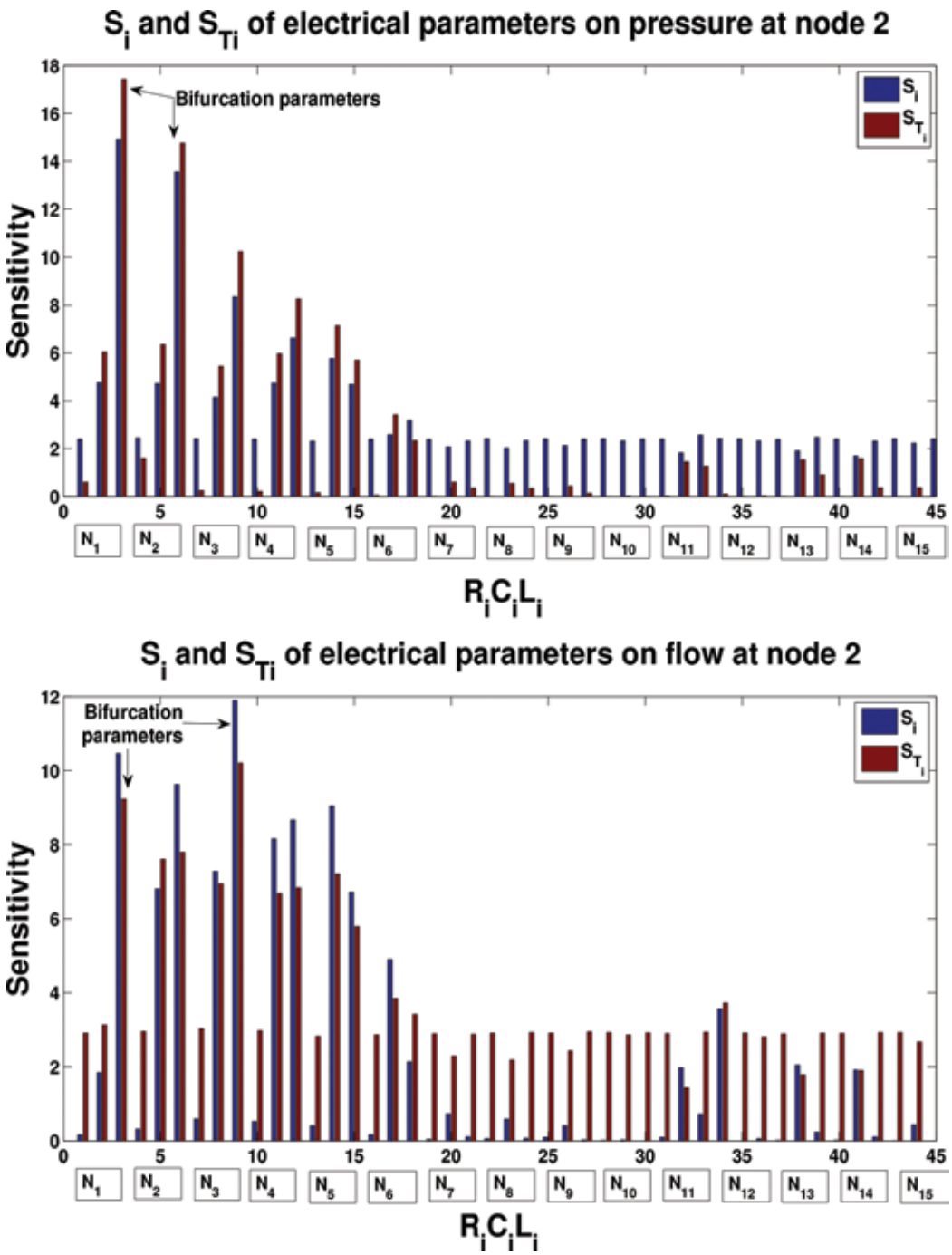

Figure 4. Ranking of bifurcation parameters (R_i, C_i, L_i) in complete arm arteries for pressure (top) and flow (bottom) at node-2. It can be clearly seen that, L_1, L_2 and L_1, L_3 are considered as bifurcation parameters for pressure and flow at node-2 respectively.

Author details

Raheem Gul[1]* and Stefan Bernhard[2]

*Address all correspondence to: gulrehman@ciit.net.pk

1 COMSATS Institute of Information Technology, Abbottabad, Pakistan

2 Department of Electrical Engineering and Information Technology, Pforzheim University of Applied Sciences, Germany

References

[1] Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S. Global Sensitivity Analysis. The Atrium, Chichester, West Sussex, England: John Wiley & Sons Ltd; 2008

[2] Kuznetsov YA. Elements of Applied Bifurcation Theory. Applied Mathematical Sciences. Vol. 112. New York: Springer Verlag; 2004

[3] Zi Z. Sensitivity analysis approaches applied to systems biology models. IET Systems Biology. 2011;**5**(6):336-346

[4] Gul R, Schütte, Bernhard S. Mathematical modeling and sensitivity analysis of arterial anastomosis in arm arteries. Applied Mathematical Modeling. 2016;**40**:7724-7738

[5] Gul R. Mathematical modeling and sensitivity analysis of lumped-parameter model of the human cardiovascular system. PhD Thesis. Germany: FU Berlin; 2016

[6] Saltelli A, Tarantola S, Chan KPS. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics. 1999;**41**(1):39-56

[7] Morris MD. Factorial sampling plans for preliminary computational experiments. Technometrics. 1991;**33**:161-174

[8] Sobol I. Sensitivity estimates for nonlinear mathematical models. Matematicheskoe Modelirovanie. 1990;**2**:112-118. In Russian, translated in English

[9] Homma T, Saltelli A. Importance measures in global sensitivity analysis of nonlinear models. Reliability Engineering and System Safety. 1996;**52**:117

[10] Chen P, Quarteroni A, Rozza G. Simulation-based uncertainty quantification of human arterial network hemodynamics. International Journal for Numerical Methods in Biomedical Engineering. 2013:1-24

[11] Gul R, Bernhard S. Optimal measurement locations for diagnosis of aortic stenoses and aneurysms in a lumped-parameter model of the systemic circulation using sensitivity analysis. International Journal of Biomathematics. 2017;**10**(8):1750116

[12] Gul R, Bernhard S. Parametric uncertainty and global sensitivity analysis in a model of the carotid bifurcation: Identification and ranking of most sensitive model parameters. Mathematical Biosciences. 2015;**269**:104-116

[13] Noordergraaf A, Verdouw PD, Boom HBK. The use of an analog computer in a circulation model. Progress in Cardiovascular Diseases. 1963;**5**(5):419-439

[14] Westerhof N, Noordergraaf A. Arterial viscoelasticity: A generalized model, effect of input impedance and wave travel in the systemic tree. Journal of Biomechanics. 1970;**3**: 357-379

[15] Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto F, Tarantola S. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Computer Physics Communications. 2010;**181**:259-270

[16] Jansen MJW. Analysis of variance designs for model output. Computer Physics Communications. 1999;**117**:35-43

[17] Jansen MJW, Rossing WAH, Daamen RA. Monte Carlo estimation of uncertainty contributions from several independent multivariate sources. In: Predictability and Nonlinear Modeling in Natural Sciences and Economics. Dordrecht: Kluwer Academic Publishers; 1994. pp. 334-343

[18] Yang J. Convergence and uncertainty analyses in Monte-Carlo based analysis. Environmental Modeling and Software. 2011;**26**:444-457

Chapter 5

Biological Hypercomputation and Degrees of Freedom

Carlos Eduardo Maldonado

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.73179

Abstract

This chapter presents the idea of biological hypercomputation (BH) and discusses how and why it entails degrees of freedom. Crossing a biological and computational point of view, the claim is made that living beings cannot be considered as machines in any sense of the word, the arguments are provided, and the consequence is drawn: the complexity of life is the very process by which living beings gain degrees of freedom. The leading thread for the analysis here is the relationship between matter, energy and information.

Keywords: life, computation, freedom, nature, complexity

1. Introduction

Biological hypercomputation (BH) is the title that summarizes the fact that living systems are not machines, and hence, do not process information like any machine, no matter what. The leading thread, thereafter, is to understanding the very way in which living systems process information. For a living being processing information is after all a matter of life or death.

A normal understanding of the issue would claim that living systems receive information from the environment and then process it, so that they would be able to learn and adapt to the changing environment. This text claims that it is a wrong take. Instead, it will be argued that there is no information before the very act of processing, and as a consequence, there is no information after the processing. Reality is the outcome of the very process of information processing.

To be sure, ranging from the various Turing Machines (TM)—u-TM, o-TM, d-TM, and many others, up to the various levels of life, say, from bacteria to living cells to plants to animals, etc., there is a growing process by which new and previously unforeseen degrees of freedom are reached. Reality, nature and the world can be seen as an increasing process of degrees of

freedom. This view, however, should not be understood hierarchically. Rather, the architecture is to be seen as a fractal organization. As the growing architecture of life increases, an amplification of previous lower layers is produced, that does not obliterate or supersedes the lower levels.

To be sure, the process by which the living systems gain degrees of freedom coincides with the entire weave of networks, fractal architectures, and interdependencies among layers, planes, and contexts. As a result, life and nature can be seen as a fantastic interweaving of times, layers, and interdependencies.

A number of works have been done that permit the approach of this text. The various authors, takes, and approaches will be discussed along the text, which brings a fresh state-of-the arte about biological hypercomputation and the weave of information processing in nature.

Throughout this chapter the basic argument is that life can be seen as a large cooperative game by which the increasing complexity of life lies on the very increasing processes of information processing.

2. A serious problem: what is information (one more time!)

In order to explain nature band the universe, the eighteenth century invented or created a physical concept, namely mas or matter. Thanks to it, the entire universe could be explained in terms of three basic but elegant laws. This was I. Newton's achievement. Anything that is or happens is the outcome of what happens to mass, thus: relationships of action-and-reaction, the relationship between two bodies based on a force equal in magnitude and opposite in direction with each other, and the heaviness or mass of a body so that the heavier body will always attract the lighter body to itself, i.e. gravity.

In the framework of the study of heat a brand new science originating both in physics and chemistry arose in the nineteenth century that introduced a quite different new concept to understand and explain nature and reality, namely energy. Whereas mass or matter is a univocal concept, energy shows a manifold of ways of existence: caloric energy, kinetic energy, potential energy, and others. The contributions of scientists such as Fourier, Carnot, Boltzmann, Lord Kelvin, and others, had as a result the identification of three laws of thermodynamics. Those three laws were sufficient to explain not only the dynamics of the universe and nature, but also the arrow of time and the challenges it arose.

Now, the twentieth century discovered still another physical concept highly more complex and different form the two original ones. This was the concept of information, originally brought out by Shannon and Weaver, which, however, was to know still further developments manly due to the research on cryptography, quantum theory, quantum computation and quantum science, mainly. The crux here is that information is a physical and yet non-tangential, non-material concept—in contrast with mass and energy [1].

Now, the relation among the three physical concepts can be adequately identified as follows:

http://dx.doi.org/10.5772/intechopen.73179

$$\text{Matter} \in \text{Energy} \in \text{Information} \tag{1}$$

Meaning that energy explains better and deeper what matter is supposed to explain, and that information explains still much better and deeper what energy was supposed to explain. Thus, the story from classical mechanics to thermodynamics to information theory is the very story by which a higher and more achieved understanding of reality is accomplished. This, however should not imply by any means a linear and necessarily accumulative process, but a story of improving, gaining more knowledge and wisdom about the universe, nature, and society. The story of knowledge can very appropriately be grasped in terms of bifurcations [2].

The amount of information of a system depends on how probable an event is. In other words, the higher the probability that an event takes place, the lower the amount—and the quality—of the system. And vice versa: the lower the probability that an event happens, the higher—and better—the information of that event. Briefly said, the information is inversely proportional to the probability—of an event. This means that information is a measure of how surprising something is [3].

As it can be seen, the characterization I have provided about information does not necessarily coincide with the standard view—say, Shannon- that identifies information and entropy. Rather, entropy can be adequately grasped as the quantification of randomness—of a system.

Therefore, it can be easily stated that the amount of information in the universe as a whole can only increase, which means that what was considered as an expanding universe in terms of the transition from energy to matter is really an expansion of information [4]. This coincides with Formula (1) presented above. Information is the concept or reality that encompasses and explains energy and matter (mass). Whereas matter presents a view of the world rather disconnected (indeed, it is the laws that gather and unify matter, according to Newton), and energy unfolds and transforms as a set of expressions and processes explained in terms of Fourier's and Boltzmann's laws, information set out a convergent, yet in-process, reality that allows for evolution [5]. Formula (2) expresses the unity of matter, energy, and information, showing hence that there are not three things but only one—read thereafter in terms of information, according to Formula (1).

$$\text{Matter} - \text{Energy} - \text{Information} \tag{2}$$

Formula (2) serves as a synthesis that claims that matter, energy, and information are one and the same thing that were historically discovered or brought out by stages. That story is exactly the story that leads from classical mechanics to thermodynamics to information theory.

Now, summarized, information theory is cryptography, quantum information, entanglement, and teleportation [6]. These are three different faces of one and the same development—an achievement that crosses the most conspicuous technologies nowadays, as well as spearhead science in any domain you wish. By and large information pervades both science and culture, currently.

A short survey to the history of computation may be very helpful here. Indeed, information processing encounters five main levels of development, thus:

- Linear, sequential, top-down information processing. Amply, this is the largely predominant paradigm in computation and computing science. Ever since the first computers, Eniac I, Eniac II, etc., on until now, computers and computations are defined by the Von Neumann architecture, and the Church-Turing thesis. A programmer programs a program and the computer follows the indications and directions. Design, parametrization, control are the key operations and ways of working and dealing with computers and computation.
- Dynamic and evolving information processing. Genetic algorithms, time series, and the first steps in artificial intelligence (AI) such as the works by Ray, Conway, and neural networks, can be said to define the second layer of information processing. The shift toward bottom-up concerns and interests is produced that largely enhance computation as a dynamic process. Not eventually modeling and simulation come along with literal explosion of programming languages aimed at a variety of goals and abilities.
- Biologically inspired information processing. More recently bio-inspired computation takes as model or guide the behavior of living systems—say, DNA, proteins, ant colonies, etc.—and develops programs that come closer to the way living beings behave. This kind of information processing has shed so far brand new fresh lights on to the very way in which computing science had being previously conceived. Thus, computation and biology at large come closer and learn from each other making computation much more flexible and robust at the same time. Adaptive computation comes to the fore.
- Biological hypercomputation. On the basis of the third layer, trying to understand the way in which living beings process information clearly entails a distance from classical computation, no matter what. Such a computation has been named biological hypercomputation. This is the core of this text.

It is clear that throughout the four mentioned layers, the very technological capabilities are developed and enhanced, for example via the so-called Moore law (after G. Moore, then president of Intel), that deals with the speed of change of the processors in computers. Undoubtedly, information processing has been changing the very understanding of computing and information, as well as it has been the subject—ultimately of software engineering as well as from hardware engineering—the two basic domains of computation, roughly said.

From the standpoint of quantum theory, information sets out the ground to unify epistemology and ontology. The link that allows such a connection is the wave function. Quantum science does not allow for the distinction, and even less the hierarchy, between ontology and epistemology, on either side. Instead, physics is about what we can say about nature—not about what nature is, any longer.

To be sure, the concept-behavior that allows for the unity of ontology and epistemology is the concept of entanglement, a concept firstly introduced by Schrödinger (*Vershränkung*), but truly implemented by J. Bell. When entangled, the unity of two or more particles becomes more important than the particles individually considered. Entanglement is not really entanglement

of entities but of information [7]. Exactly in this sense, too, teleportation is not so much the transportation of a physical entity, what of what makes that entity as such, namely information. It is information, indeed, what is teleported, not the thing as such, in its materiality. Matter is information, which brings us back to Formulas (1) and (2).

Rightly understood, we can safely say that there is information processing in nature. Briefly said, nature can be understood in terms of information.

3. Living beings process information: a complexity understanding

Living beings do not just read the environment; moreover, they accordingly unceasingly create brand new information that was non-existing before. In other words, living beings read the environment but they also, at the same time, write on the environment they are reading on—if the analogy is permitted here.

Whereas any Turing Machine (TM) processes information top-down, sequentially, linearly, and mechanistically, according to the prevailing Von Neumann architecture of TM-computers, living beings, to say the least, process information in-non classical way if the framework is the Turing-Church thesis.

This text argues that the processing of information among living beings consists not only—and not so much—in rightly reading and interpreting the environment and the surroundings, but also—and mainly—in introducing or creating brand new information into the world, correspondingly.

Thus, over against a popular understanding of the issue, information processing does not have anything to do with things such as "following a thread," "analyzing," nor even "understanding." Processing information can rather be grasped in biological or medical terms as metabolizing, i.e., changing one thing—say A—into another, B, for instance -. Good metabolizing, hence, transforms a nice meal into a poem or a scientific paper, or also a nice rest in the evening into insight and strength, for example.

For the living beings, processing information is a matter of life or death—in that a bad information processing may entail danger, peril or death. Biologically or culturally speaking, information processing might entail identifying a good male or female a god territory, the presence of a friend or an enemy, who truly loves you and who just pretends to love you, which food is healthy and which poisonous, for instance. Living beings that process information rightly may encounter better circumstances for adaptation and learning, and be literally the fittest. Evolution and adaptation are after all, it appears, a matter of good information processing.

Yet, life is not just an emergent property of the universe. A number of authors have claimed that the universe herself is alive and conscious [8–10]. The claim is hence about consciousness not as an epiphenomenon of the universe, and life as an emergent stance. In this take, life is an essential feature of the universe that can be traced back to the very processing of information, precisely. For the sake of brevity I shall put such a claim here into brackets and focus on the way in which living systems do process information.

Whereas at the same time living beings—from bacteria to cells, from organisms to biomes, for example—continuously read the environment and the surroundings they create brand new information that is brought into the world in a variety of forms, thus: as acts and actions, as behaviors, as creation in same cases of tools and rods, as the use of tools as devices aimed for a certain goal, definitely as language and communication, and as forms of organization, whether individual or collective [11]. It is the very creation of new information into the world, which is usually grasped as adaptation. Adaptation is the biological concept that originally is the outcome of information processing.

As it can be easily seen, living beings compute. Moreover, life without computation is inconceivable.

Computation is therefore a concept that truly means interaction—with the environment, and with other living beings. Natural computation is much more than a metaphor, but the distinguishing feature of living beings. Life is a large weave of computation that takes place in a manifold of ways simultaneously, as follows:

There is classical computation in nature in either form of a Turing Machine. At the same time there is parallel computation, and multilevel computation, non-local computation as well as distributed computation, fuzzy and random computation very much as also quantum computation and emergent computation, not to mention interactive computation. A typology of computation and computational models can be seen in [12]. A metaphor can be introduced here, namely a variety of computations corresponds to the diversity of life and living beings. It appears that ecology and biology go hand in hand with information theory and computing science, even though the two latter are more recent that the two before [13].

To be sure, biological hypercomputation is the way in which living beings process information—is anything but machines. However, there are a number of explanations within the health sciences and the living sciences that still operate mechanically. Thus, for instance, the functioning of the brain many times is explained in terms of "on" and "off" switches or operations, and the very functioning of the heart is usually conceived as a pumping machine, period. Many other cases can be introduced here. The crux is that those explanations still owe a big deal to the past and remain very short vis-à-vis more contemporary and spearhead explanations. Computing sciences is such one of those frontline explanations, under the proviso that it be not a Turing Machine one, in any concern.

Biological hypercomputation (BH) has been introduced [15] to mean that life and the living processes cannot be understood from lower stances: in other words, life is to be explained as it happens: life is an uncompressible "program." The most crucial and basic problems of living beings are certainly not tackled in terms of mathematical functions, whatsoever. Instead, the problems living beings faced are normally solved in terms of non-classical logics, a field that has not deserved as much attention as its own value.

More fundamentally still, computationally speaking, the difference between software and hardware is irrelevant. More exactly, it does not exist. The phenotype expresses the genotype according to evolution and the environment. Epigenetics, for example behavioral epigenetics

or even symbolic epigenetics, becomes central here, for it most adequately expresses the idea that very much as there is no difference between software and hardware, in the same tenure, there is no difference between nature and culture. They are closely intertwined and cannot be divided or split.

As a consequence, computing (= BH) is for life one and the same thing with evolution. Evolving and computing are two faces of one and the same token. Evolution is the process through which living beings gain degrees of freedom.

4. Living beings gain degrees of freedom

Information processing among living beings takes place as an unceasing creation of information that is introduced into the world. Thus, information processing is not just reading the environment and adequately interpreting it, but, moreover, correspondingly, bringing new information into the world in the form of actions, processes, behaviors, or achievements.

There is no information before the information processing by living beings, very much as there is no information after the information processing by living beings. Information exists in so far as it is processed, that is, both created and changed. In other words, there is no reality previous to the processing of information, but neither is there any reality "outside" the information processing. Yet, the universe becomes increasingly complex precisely thanks to the unceasing processing of information, i.e. the reduction of entropy in the universe. Life is that system that creates order and exhibits the best form of order possible.

As a consequence, the universe is alive it appears. A number of interpretations in this sense can be mentioned that challenge the traditional view of the universe that goes as follows:

$$\text{Physics} \rightarrow \text{Inorganic Chemistry} \rightarrow \text{Organic Chemistry} \rightarrow \text{Biology} \rightarrow \text{Culture} \quad (3)$$

I argue that Formula (3) is erroneous because it is set on the physical predominance of physics over biochemistry or biology, for example. Formula (3) is really a translation of Formula (1), and thus, it sets out the ground for a reductionist view of nature and the universe. A clear reductionist approach serves as ground for that formula.

The increasing complexity of reality and the universe is one and the same process as the increasing complexity in the information processes [13, 14]. We have been gaining an enormous field in understanding what life does and how it is made possible. Thus, for instance, regarding the realm of plants, Refs. [15, 16] have shed solid lights about how they think, literally. Down the scale, [17] did steadily study the way in which information processing takes place among bacteria. Furthermore, information processing has been studied additionally among social insect networks [18]. The range of studies and the cope is always wider and deeper, encompassing animals at large [19], until of course we reach the level of human beings, and computers and computation.

The concept "degrees of freedom" can be understood from a wide take that says that it consists in the number of parameters a system exhibits or has that may vary independently (Wikipedia) until the calculation of the various scenarios in a system or the number of dimensions that are needed to determine a full vector. In complexity science or theory the concept is pivotal and allows for crossing different approaches, sciences, discipline, and interpretations aiming at making explicit what complexity is all about.

We can safely say that the higher the degrees of freedom a system has the more complex it is. And vice versa, the lower the degrees of freedom, the less complex it is. Formula (4) synthetizes this:

$$> \text{Degrees of Freedom} > \text{Complexity} \tag{4}$$

Now, living beings are what they do. And the best they do is information processing, exactly. The better a living system processes information, the fittest it is. As a consequence, the more degrees of freedom it gains. Hence, living is a matter of gaining degrees of freedom. Death, it seems, is the complete loss of degrees of freedom. All possibilities are then closed.

Therefore the claim can be made that information plays an "active" role in the living systems, and a "passive" role in classical physics. Analogously as it happens with any Turing Machine, namely information is something that "happens" in the processor that is conceived as a black box. The audience or the researcher knows what enters (input) and knows what comes out (output), but ignores what and how happens in the black box. Classical computation seems to reduce the degrees of freedom by accepting or assessing the processing of information as a "black box." On the contrary, (BH) is the process by which new degrees of freedom are gained in so far as new adaptations, networks and behaviors are made possible, for instance?

Information processing does not stand as a condition for evolution and adaptation. On the contrary, it is the very process of evolution both of the living beings and the environment. Coevolution is the name that best suits both stances in their entanglement. Living is nothing else that coping with the information of the surroundings and accordingly changing that information in terms of brand new information brought into the world.

The better the living system processes information, the better it copes with challenges, solves problems, adapts to a continuously changing environment. Life ends when it "overdoses" of information and the living being is unable to process it. A saturation point is reached and the system can no longer process any further information [20].

Summarizing, life consists in the process by which information processing is carried out and hence unceasingly new degrees of freedom are reached. The system can be said to be "young." Contrarily, death happens when the processing of information is not possible any longer—it becomes slow, saturated, new information cannot be processed accordingly, and the "screen" slows down in bringing out the information required.

Life is a matter of "controlling" the saturation of information, and thereafter, of information processing. Software and hardware are one and the same thing.

http://dx.doi.org/10.5772/intechopen.73179

5. The weave of life

Information is a continuous process that emerges as a result of intricate non-linear interactions of living beings among themselves and with their environment that both creates nature and the world as it happens, and at the same time changes it unceasingly. This is exactly the story of learning, adaptation and coevolution.

As it is well known, living beings evolve in rugged adaptive landscapes—a concept originally introduced by Ch. Darwin in 1859, originally named as fitness landscape. Life is a wonderful complex weave that has no center, but many bubs, clusters and nodes, continuously changing on the ever-changing environment on earth. Only that the timescales in geology are vastly deeper than those of the living beings, not to mention human beings. Indeed, whereas the basic time-scale of life is the decade or even the month for some species, the basic time-scale in geology is 1000.000 years. Against all odds, this is the ultimate time-scale of life on earth.

Life is a large and robust weave of cooperation, commensalism, and mutualism, rather than competition and predation. [20] has made on this subject a great contribution, namely the importance of symbiogenesis: a large network (instead than a chain) of mutual interdependency among living beings so that each organism and species benefits from others, and vice versa.

The story of life is a story of increasing complexity, i.e. biodiversity. Such diversity is said to be at three levels, thus: genetic, natural, and cultural. The countries that have the three kinds of biodiversity are called "megadiverse" and they are 17 countries, to-date—all of them place on or very near the equator.

The story of life pivots around six fantastic moments of increasing diversity followed by six massive extinctions. At each stage, life has made of herself a wonderful asset of possibilities, forms, shapes, structures, behaviors, and characteristics that correspond with dynamics through which, as a whole, amazing degrees of freedom are been attained or reached. And yet, that story is non-teleological. In other words, life has no ends or goals, and evolution has no purpose whatsoever. Such was indeed the scandal that Darwin's *The origins of species by means of natural selection* (1859) meant particularly vis-à-vis the cornerstone of the western civilization, namely teleology [21]. The western world needs believe that there are *telos*, and that *telos* are necessary.

To the discomfort of those who believe in goals end, whether their own or imposed or suggested, the theory of evolution—namely the best theory ever developed to think about change and transformation—introduces the idea of the absence of goals or ends (*telos*). Well, the theory of symbiogenesis comes to complete, so to speak, the theory of evolution by showing that the various stages of complexity of life, firstly, and secondly that the whole weave of life naturally tends toward increasing complexity and hence to cooperation and interdependency.

As a consequence, living beings are symbiont: not only it is a cooperative interaction but also a prolonged one among different species and/or organisms. Mutualism and not competition is the rule in nature, it appears. Nature is the realm of freedom par excellence.

6. Conclusions

We have three concepts: matter, energy and information. However, a right understanding of them brings to the fore the fact that they are not three, but one and the same concept: matter-energy-information. Moreover, information can be rightly understood as fundamental (= grounding) energy and mass.

What happens to matter is truly a matter of energy and relations among various types of energy. And still the very processes, dynamics and configurations of energy are dynamics in and of information. Formula (4) expresses this understanding.

$$\text{Information} \leq \text{Energy} \leq \text{Matter (Mass)} \tag{5}$$

Formula (5) simply means that matter explains less (and worst) what energy does explain, and furthermore, information allows for a deeper and better understanding of what energy means to explain.

Now, when looked under the light of life and how life is possible, the three concepts are transformed into one process, namely biological hypercomputation. In the processing of information among the living beings matter, energy and information become one single unity: life. Thus, life allows for the overcoming three different layers of reality (physics, thermodynamics, and information); such overcoming is also the non-differentiation between software and hardware. Processing information for life consists in gaining degrees of freedom.

Living beings process information much more than (just) matter and energy, which they definitely do. Evolution can be seen as the processing of information processing by living beings through which the world and nature are complexified. Information allows for the rejection of a reductionist approach of life and nature.

Author details

Carlos Eduardo Maldonado

Address all correspondence to: maldonado.carloseduardo@gmail.com

School of Medicine, Universidad El Bosque, Bogotá, Colombia

References

[1] Gleick J. The Information: A History, a Theory, a Flood. New York: Vintage Books; 2012

[2] Serres M. Éléments d'histoire des sciences. Bordas; 1993

[3] Vedral V. Decoding Reality. The Universe as Quantum Information. Oxford: Oxford University Press; 2010

[4] Roederer JG. Information and Its Role in Nature. Springer Verlag; 2005

[5] Chaisson E. Cosmic Evolution. The Rise of Complexity in Nature. Harvard: Harvard University Press; 2001

[6] Zeilinger A. Dance of the Photons. From Einstein to Quantum Teleportation. New York: Farrar, Straus and Giroux; 2010

[7] Gilder L. The Age of Entanglement. When Quantum Physics was Reborn. New York: Vintage Books; 2008

[8] Kafatos M, Nadeau R. The Conscious Universe. Part and Whole in Modern Physical Theory. Springer Verlag; 1990

[9] Gribbin J. In the Beginning: The Birth of the Living Universe. Little Brown & Co.; 1995

[10] Kauffman S. Humanity in a Creative Universe. Oxford: Oxford University Press; 2016

[11] Maldonado CE. Hipercomputación biológica y comunicación entre los seres vivos. In: Arboleda LC, editor. Un Festschrift para José Luis Villaveces. Bogotá: Academia de Ciencias Exactas, Físicas y Naturales; 2017. pp. 109-124

[12] Burgin M, Dodig-Crnkovic G. Typologies of Computation and Computational Models; 2013. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.446.5721&rep=rep1&type=pdf

[13] Linewaever CH, Davies PCW, Ruse M, editors. Complexity and the Arrow of Time. Cambridge: Cambridge University Press; 2013

[14] Barrow JD, Davies PCW, Harper CL Jr, editors. Science and Ultimate Reality. Quantum Theory, Cosmology and Complexity. Cambridge: Cambridge University Press; 2005

[15] Maldonado CE, Gómez-Cruz N. Biological hypercomputation: A new research problem in complexity theory. Complexity. 2015;**20**(4):8-18. ISSN 1099-0526

[16] Mancuso S, Viola A. Brilliant Green. The Surprising History and Science of Plant Intelligence. Island Press; 2015

[17] Baluska F. Signaling in Plants (Signaling and Communication in Plants). Springer Verlag; 2010

[18] Ben-Jacob E. Learning from bacteria about natural information processing. Annals of the New York Academy of Sciences. Oct 2009;**1178**:78-90

[19] Waters JS, Fewell JH. Information processing in social insect networks. PLoS One. 2012;**7**(7):e40337

[20] Gross A, Vallely A, editors. Animals and the Human Imagination. A Companion to Animal Studies. New York: Columbia University Press; 2012

[21] Margulis L, Sagan D. Acquiring Genomes. A Theory of the Origins of Species. New York: Vintage Books; 2003

Section 2

Complexity in Physical Systems

Chapter 6

Self-Organization, Coherence and Turbulence in Laser Optics

Vladimir L. Kalashnikov and Evgeni Sorokin

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71478

Abstract

In the last decades, rapid progress in modern nonlinear science was marked by the development of the concept of dissipative soliton (DS). This concept is highly useful in many different fields of science ranging from field theory, optics, and condensed matter physics to biology, medicine, and even sociology. This chapter aims to present a DS appearance from random fluctuations, development, and growth, the formation of the nontrivial internal structure of mature DS and its breakup, in other words, a full life cycle of DS as a self-organized object. Our extensive numerical simulations of the generalized cubic-quintic nonlinear Ginzburg-Landau equation, which models, in particular, dynamics of mode-locked fiber lasers, demonstrate a close analogy between the properties of DS and the general properties of turbulent and chaotic systems. In particular, we show a disintegration of DS into a noncoherent (or partially coherent) multisoliton complex. Thus, a DS can be interpreted as a complex of nonlinearly coupled coherent "internal modes" that allows developing the kinetic and thermodynamic theory of the nonequilibrious dissipative phenomena. Also, we demonstrate an improvement of DS integrity and, as a result, its disintegration suppression due to noninstantaneous nonlinearity caused by the stimulated Raman scattering. This effect leads to an appearance of a new coherent structure, namely, a dissipative Raman soliton.

Keywords: optical turbulence, dissipative solitons, chaos in nonlinear optical systems, generalized cubic-quintic nonlinear Ginzburg-Landau equation, dissipative Raman soliton

1. Introduction

Coherent and partially coherent structures emerging in nonlinear systems far from the thermodynamic equilibrium play an important role in different research areas ranging from hydrodynamics and plasma physics to biophysics and sociology. Nontrivial dynamics of such

structures including chaos and turbulence is a challenge for modern nonlinear science and one may assume that "the problem of turbulence is one of the central problems in theoretical physics" [1]. The reasonable approach to this issue, which can translate some contra-intuitive and obscure ideas in this area into explicit and verifiable concepts, is a realization of simpler dynamics in quite different material context. Such an approach can be named *metaphorical* or *analog* modeling [2], and a rapid progress of modern laser technology provides an *ideal playground* for such enterprise due to high controllability, relative simplicity, and unique potential of statistic gathering. Such progress was marked by the development of the concept of a dissipative soliton (DS) [3]. The existence of DS under nonequilibrium conditions requires a well-organized energy exchange with an environment so that this energy flow forms a nontrivial internal structure of DS, which provides the energy redistribution inside it and can distort the soliton coherence. Such a DS with nontrivial internal structure can develop in lasers, and the DS dynamics can become chaotic and turbulent [3–5]. For instance, such emergent structures can be considered as a classical analog of Bose-Einstein condensate in low dissipative limit and, contrariwise, as a primitive analog of cell in the case of extensive and well-structured energy exchange with an environment. Formally, these inherently nonHamiltonian entities mimic some features of Hamiltonian systems that remain an obscure and insufficiently explored topic regarding the fundamental properties of coherent dissipative structures. The range of turbulence, noise, and rogue wave phenomena emulated by the optical DS is so broad that it turns them into a universal testbed for studies in the fields of nonlinear dynamical systems and nonequilibrium thermodynamics.

In this work, we conjecture a spectacular analogy between the spectral structures of DS and strong Langmuir turbulence. Such close relation leads to chaotization of DS dynamics with the energy growth. This analogy is deepened by analysis of energy flows inside DS so that a DS can be represented as a "glass of boiling water" or, mathematically, as an ensemble of interacting quasi-particles or "nonlinear modes." The phase decoupling of these "modes" leads to turbulence or DS dissolving. Such a representation open the door for building the *kinetic theory* of open (dissipative) semi-coherent structures which mimics, in particular, a quantum Bose-Einstein condensate in a dissipative environment. Moreover, our preliminary investigations demonstrated a mechanism of *turbulence control* provided by noninstantaneous nonlinearity (stimulated Raman scattering in optical case) [6]. This phenomenon is especially interesting because an inherently noisy process (Raman scattering) suppresses a turbulence under some conditions that is the manifestation of *stochastic resonance*, which can be significant for a dissipation control in coherent quantum systems (particularly, a quantum computer and a quantum cryptography device).

2. Analogy between DS and turbulence

The phenomenon of turbulence appears in many areas of our experience ranging from atmospheric and oceanic rogue events, aero- and hydrodynamics, optics to cardiology and neurophysiology [1, 5, 7–13]. Such a broad class of phenomena cannot be grasped by some single and simple model. However, there are some comparatively simple equations which allow

http://dx.doi.org/10.5772/intechopen.71478

describing an extremely broad class of phenomena. It is possible that the most known one is the famous nonlinear Schrödinger equation (NSE) which describes an evolution of slowly varying wave in a nonlinear medium and can be considered as a "metaphoric" simulation tool for a study of nonlinear phenomena far from equilibrium [14, 15]:

$$\frac{\partial \Psi}{\partial T} + \sum_{j=1}^{d} \frac{\partial \omega}{\partial k_j} \frac{\partial \Psi}{\partial x_j} - \frac{i}{2} \sum_{j,l=1}^{d} \frac{\partial^2 \omega}{\partial k_j \partial k_l} \frac{\partial^2 \Psi}{\partial x_j \partial x_l} + i \left(\frac{\partial \omega}{\partial |\Psi|^2} \right) |\Psi|^2 \Psi = 0. \tag{1}$$

The dimensionality of this equation is relative: the evolutional coordinate can be a time T or a propagation distance z ($T \leftrightarrow z$), the transverse coordinate can be transverse multidimensional spatial x_j ($j=1\ldots d$) one or a local time t ($x \leftrightarrow t$, $d=1$). The Fourier representations of a "field" slowly varying envelope Ψ are interchangeable between frequency and momentum domains ($\omega \leftrightarrow k$, $d=1$). Eq. (1) may describe the propagation of optical pulses in a nonlinear medium (then Ψ is a complex field amplitude and $|\Psi|^2$ is proportional to a field power), the capillary waves on a fluid surface, the Langmuir waves in plasma, or the weakly nonlinear Bose-gas in classic limit (in the last case Eq. (1) represents the famous time-dependent Gross-Pitaevskii equation [16]).

Here $\Psi(x, t)$ is a slowly varying amplitude of wave propagating in dispersive ($\sum_{j,l=1}^{d} \frac{\partial^2 \omega}{\partial k_j \partial k_l} \frac{\partial^2 \Psi}{\partial x_j \partial x_l}$ - term; let us $\beta \equiv \frac{\partial^2 \omega}{\partial k_j \partial k_l}$) and nonlinear ($\left(\frac{\partial \omega}{\partial |\Psi|^2} \right) |\Psi|^2 \Psi$ - term; let us $\gamma \equiv \frac{\partial \omega}{\partial |\Psi|^2}$) medium. The nonlinear term in Eq. (1) can have the different forms; in particular, a nonlinear response can be non-instantaneous.

The notion of turbulence is fuzzy in some sense. Here, the turbulence will be treated as a phenomenon related to the excitation of a sufficiently large number of degrees of freedom that causes a loss of their mutual phase information [15]. As a consequence, a wave package decouples into a set of individual modes ("particles") which interaction can be described in the framework of kinetic theory as many-particle collisions in Bose-gas. In other words, as some *degrees* of freedom become very large for sufficiently large energies, phase information becomes irrelevant, and the waves *decohere* [8, 15]. Thus, a wave can be considered as a set of decoupled "modes" n_k in a spectral (or wave-number) space:

$$\langle \Psi(k) \Psi(k') \rangle = n_k \delta(k - k'). \tag{2}$$

Thus, we come to a *"kinetic"* theory of turbulence, for example, to a model of four-boson interaction described by the nonlinear Schrödinger equation:

$$\begin{aligned} \frac{\partial n_k}{\partial t} \propto & \int (n_{k_1} n_{k_2} n_{k_3} + n_k n_{k_2} n_{k_3} - n_k n_{k_1} n_{k_2} - n_k n_{k_1} n_{k_3}) \\ & \times \delta\left(\vec{k} + \vec{k_1} - \vec{k_2} - \vec{k_3} \right) \delta(\omega + \omega_1 - \omega_2 - \omega_3) d\vec{k_1}\ d\vec{k_2}\ d\vec{k_3}. \end{aligned} \tag{3}$$

Such an equation becomes nontrivial in a dissipative environment [17, 18]. A simple generalization of NSE (1) taking into account the dissipative effects includes a saturable gain (energy "source") σ, dissipative nonlinearity (self-amplitude modulation, SAM) $F(|\Psi|^2)$, and spectral dissipation (spectral in the sense of dissipation in the Fourier space) $\sum_{j,l=1}^{d} \alpha(x_j, x_l) \frac{\partial^2 \Psi}{\partial x_j \partial x_l}$:

$$\text{RHS of Eq.}(1) = \sigma\Psi + F\left(|\Psi|^2\right)\Psi + \sum_{j,\,l=1}^{d} \alpha\left(x_j, x_l\right)\frac{\partial^2\Psi}{\partial x_j \partial x_l}. \tag{4}$$

Eqs. (1) and (4) called the generalized complex nonlinear Ginzburg-Landau equation have the strongly localized (in x-space) steady-state (in T-space) solutions which are named *dissipative solitons* (DS) [3]. A classical (nondissipative) soliton, which possesses the quite specific mathematical properties [18–20], develops due to mutual compensation of dispersive spreading and self-compression caused by the phase nonlinearity under the condition of $\frac{\partial^2\omega}{\partial k_j \partial k_l} \times \frac{\partial\omega}{\partial|\Psi|^2} > 0$ and is stable in a (1 + 1)-dimensional (i.e., T plus $d=1$ in Eq. (1)) case.[1] The parameters of such soliton are not fixed but only interrelated. One may say that a soliton "lives in solitude" ("pratyekabuddha," **Figure 1**).

Dissipation adds new bounds on the soliton parameters and fixes them so that one may say that the DS lives in "the heart of nonlinear world" ("bodhisattva," **Figure 2**).

The mutual balance of dispersion and phase nonlinearity remains a crucial factor for DS formation, but its physical meaning differs substantially from that for nondissipative soliton. The crucial factor here is a *resonance* between dispersive (linear) waves and DS: equality of their wave-numbers defines the frequency window where DS can exist. Indeed, a wave-number of DS is $q = \gamma P_0$ ($P_0 \equiv \max(|\Psi|^2)$) [23]. The dispersion relation providing the resonance with linear waves is $k(\omega) = \beta\omega^2/2$. To be stable (i.e., nonradiating), the DS spectrum has to be localized within a frequency window $\pm\Delta$: $k(\pm\Delta) = q$, where $\Delta = \sqrt{2\gamma P_0/\beta}$ (**Figure 3**).[2] Formation of these "domain walls" [24–26] due to phase effects in a dissipative system results in natural frequency cut-off, which is essential for inherent analogy between DS and a turbulent entity.

However, sole dispersive balance is not sufficient for the DS stability. The spectral dissipation $\sim\alpha\Delta^2$ plays a crucial role cutting the spectrum and defining the DS width (**Figure 4**). As

Figure 1. Soliton exists under a balance between phase nonlinearity and dispersion [21, 22].

Figure 2. DS parameters are fixed by both nondissipative and dissipative factors [21, 22].

[1] Further, namely one-dimensional ($d=1$) systems will be under consideration that is a quite precise approximation for solid-state and fiber laser dynamics [21].

[2] One has to remind the $x \leftrightarrow t$ and $k \leftrightarrow \omega$ dualities in Eq. (1).

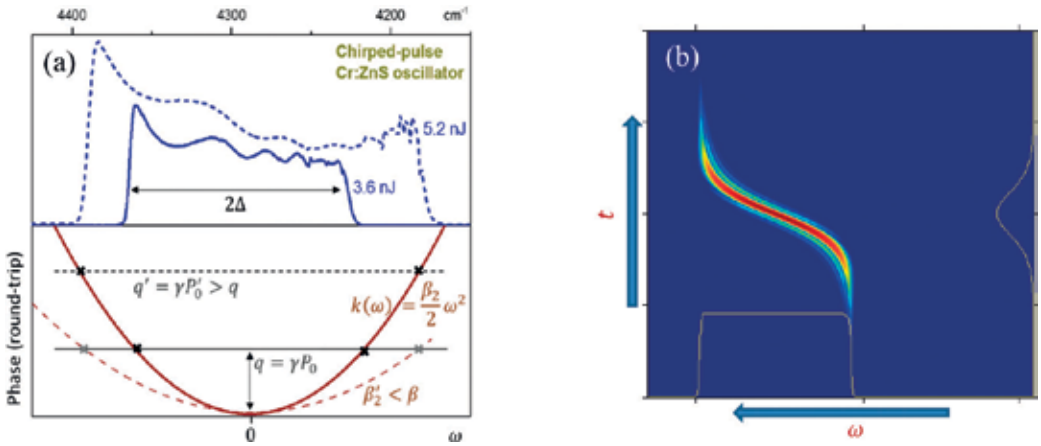

Figure 3. Resonance conditions for DS and linear waves and DS spectra in dependence on DS energy (a); and the Wigner (time-spectral) diagram of DS (b) [23].

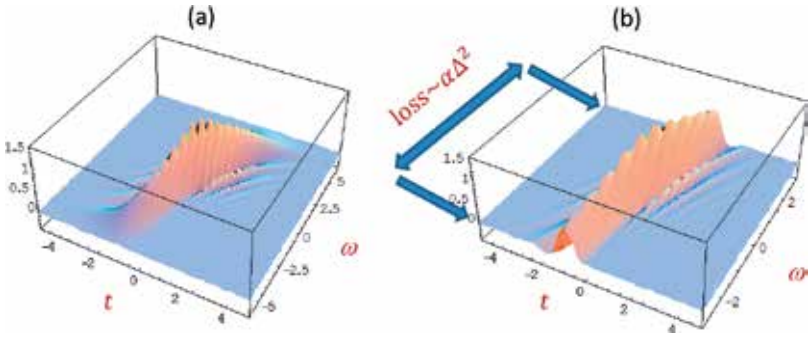

Figure 4. Wigner representation of DS (a) and cut-off due to spectral dissipation defining the DS width (b).

will be seen, this factor is crucial for dissipative soliton turbulence. A spectral dissipation must be balanced by a nonlinear gain $\sim\kappa P_0$ (we assume $_F(|\Psi|^2) \approx \kappa|\Psi|^2 - \zeta|\Psi|^4 + \ldots$, where the first term is leading) that results in the additional relation between soliton spectral width and its peak power: $\Delta = \sqrt{\kappa P_0/\alpha}$. In combination with the dispersive relation, it gives the condition for the soliton existence which combines the dissipative and nondissipative factors: $\frac{\alpha\gamma}{\beta\kappa} \leq 1/2$.[3] One has to note, that both considered mechanisms of DS formation act in the spectral domain and, as was shown in [28], a transition to spectral domain is fruitful for developing a DS theory.

The key feature of DS is its nontrivial internal structure revealing itself in the phase inhomogeneity[4] and the internal energy redistribution (E is an energy flow):

---

[3] More precise analysis [27] gives the conditions of asymptotical stability: $\frac{\alpha\gamma}{\beta\kappa} \leq \begin{matrix} 1/3 \text{ if } E \to \infty \\ 1/2 \text{ if } E \to 0 \end{matrix}$, where E is a DS energy.

[4] The measure of this inhomogeneity is a so-called *chirp* $\Theta \propto \frac{\partial^2 \arg(\Psi)}{\partial t^2}$.

$$E \equiv \frac{i}{2}\frac{\partial}{\partial t}\left(\Psi\frac{\partial\Psi^*}{\partial t} - \Psi^*\frac{\partial\Psi}{\partial t}\right) = 2\sigma|\Psi|^2 + 2\kappa|\Psi|^4 - 2\alpha\left|\frac{\partial\Psi}{\partial t}\right| + \alpha\frac{\partial^2|\Psi|^2}{\partial t^2}. \quad (5)$$

The third term in RHS of Eq. (5) is phase-sensitive and, thus, there is an energy flow from DS center, where spectral components with minimal relative frequencies are located, to the DS wings, where frequency components with maximum relative frequencies are located (**Figure 5**). Here, energy dissipates. Such nontrivial internal "life" of DS intensifies with the growth of phase inhomogeneity Θ. Simultaneously, DS becomes an energy scalable coherent concentrate with the energy ("concentrate mass") ∝Θ [5].

As a result of phase inhomogeneity and intensive internal energy flows, the internal coherence of DS can become partially broken. Then, DS splits into partially coherent "internal modes" which interact with each other as the independent "sub-solitons." [29–31] Thus, DS becomes a strongly localized "cloud" of interacting "quasi-particles" or "glass of boiling water" (**Figure 6**).

These figures demonstrate an affinity between the structures of DS and turbulence [8, 23]. Both spectral structures are defined by dispersion relations: between soliton and dispersive waves for the former and Langmuir dispersion relation for the latter (**Figure 7**). Secondly, both high-energy DS and turbulence are characterized by spectral condensation at zero frequency (wavenumber) with subsequent scattering to higher frequencies confined by cut-off at ±Δ.

Such an analogy between DS and turbulence opens a door for building the kinetic and quantum [5, 32–34] theory of open (dissipative) semi-coherent structures which mimics, in particular, a quantum Bose-Einstein condensate in a dissipative environment.

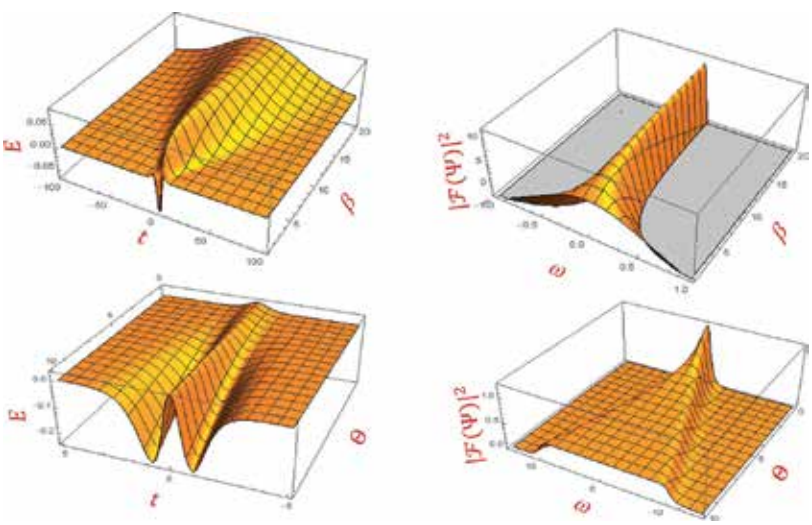

Figure 5. Energy flows (left column) and corresponding spectral profiles (right column, $\mathcal{F}$ is a Fourier image of Ψ) in dependence on dispersion β and chirp Θ for a DS with the profile $\Psi \propto \mathrm{sech}(t)^{1+i\Theta}$ [4].

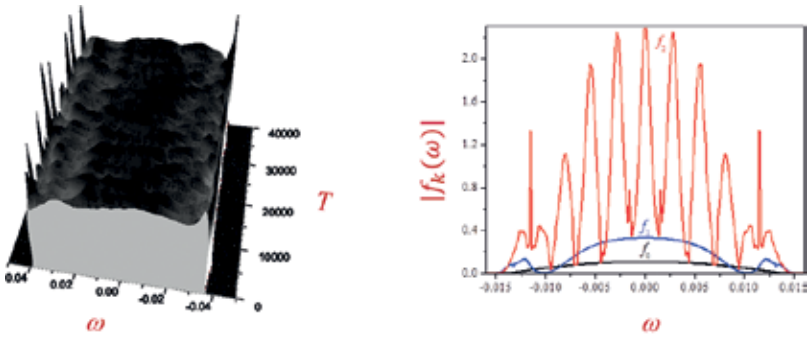

Figure 6. DS spectrum as a "glass of boiling water" (left) [30] demonstrating the dynamics of internal perturbation modes (right) [31]. The last picture is obtained by a perturbation analysis in spectral domain with the Neumann series expansion.

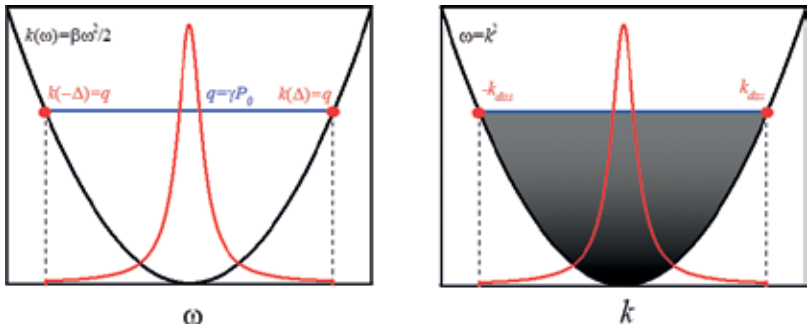

Figure 7. DS spectrum (central red curve in the left picture) where the cut-off frequency Δ is defined by the resonance condition between linear waves with a wave-number k and DS with a wave-number q (parabolic black curve in the left picture; P_0 is a DS peak power) [23]. The turbulence in the wave-number space is defined by the Langmuir dispersion relation (parabolic black curve in the right picture). Spectral condensation at $k = 0$ is illustrated by shading and forms a characteristic turbulence spectrum (central red curve; right picture) with the cut-off wave-number $\pm k_{diss}$ defined by a dissipation (adapted from [8]).

3. Transition to a DS turbulence

The mechanism of transition to turbulence for DS can be associated with the time/spectral duality (**Figure 8**). When the energy increases (i.e., $E \to \infty$ that corresponds to a system with "infinite capacity" [36]), the spectrum condensates around $\omega = 0$ within a diapason of $\Xi \to 0$ (**Figure 7**). Simultaneously, DS broadens in time domain $\propto 1/\Xi$ by analogy with a growth of Bose-Einstein condensate "mass."[5] The DS peak power tends to some constant value $P_0 \propto 1/\zeta$ defined by a saturation of dissipative nonlinearity (see above), and, thereby, the cut-off frequency $\Delta = \sqrt{\kappa P_0/\alpha}$ tends to be constant. The last value defines the coherence scale $\propto 1/\Delta$ (few picoseconds for a typical DS).[6] As a result, DS becomes "decoupled," and even small perturbations can

[5]The value $\propto 1/\Xi$ can be treated as a measure of "long-range" correlation scale.
[6]The value $\propto 1/\Delta$ can be treated as a measure of "short-range" correlation scale.

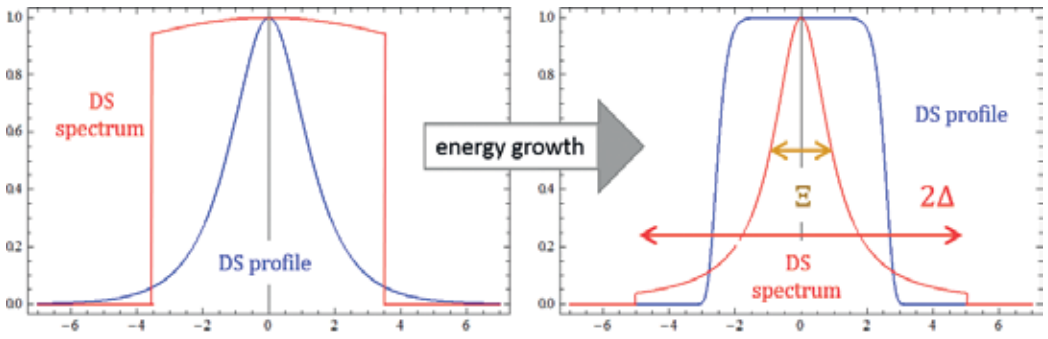

Figure 8. DS spectrum (red curves) which cut-off frequency Δ defines the correlation scale ∝1/Δ. It tends to some constant value with the DS energy growth. Simultaneously, DS broadens in time domain ∝1/Ξ (blue curves) in parallel with the concentration of energy around $\omega = 0$ in spectral domain [35].

destroy its internal coherence and split ("nucleate") it into a set of "internal modes" shown in **Figure 6**.

More close insight into this mechanism can be provided by the adiabatic theory of DS in spectral domain presented in [28]. As was shown, the DS spectrum can be expressed as follows:

$$p(\omega) = \frac{\Upsilon}{\omega^2 + \Xi^2} H\left(\Delta^2 - \omega^2\right), \tag{6}$$

where $p(\omega)$ is a DS spectral power, and H is the Heaviside function. Eq. (6) represents the spectra shown in **Figures 7** and **8**, and can be interpreted by analogy with the Rayleigh-Jeans distribution, so that Ξ^2 plays a role of negative "chemical potential" [8, 36, 37]. The parameter $\Upsilon = 6\pi\gamma/\kappa\zeta$ is an analog of "temperature" and is defined by both dissipative and nondissipative nonlinear parameters.

Since the "chemical potential" Ξ^2 decays with the energy growth (**Figure 8**), a system tends to the state of "soliton gas" [38] with the characteristic "soliton size" ∝1/Δ. Thereby, a coherent "condensate" with minimum entropy becomes a state of the decomposed "quasi-particles" with the chaotically modulated powers because the required entropy growth is provided by such modulation[7] [40]. Thus, the energy growth (i.e., the growth of "condensate mass" ∝1/Ξ) leads to extra-sensitivity to *quantum-level* noises [40, 41] that urges the quantum theory of coherent and semi-coherent dissipative structures, which would weave largest and smallest scales in the DS dynamics.

The example of such "DS decomposition" through a turbulence is shown in **Figure 9**. This figure is obtained by numerical simulation of Eqs. (1) and (4) with taking into account of the gain saturation in the form of $\sigma = \delta(1 - E/E_s)$ [27]. **Figure 9** demonstrates clearly two stages of DS evolution. The first stage corresponds to an incoherent and strongly turbulent DS, which is

[7]Here, one may draw an analogy with Hamiltonian systems, where the gradient of field is a measure of the amount of fluctuations [39].

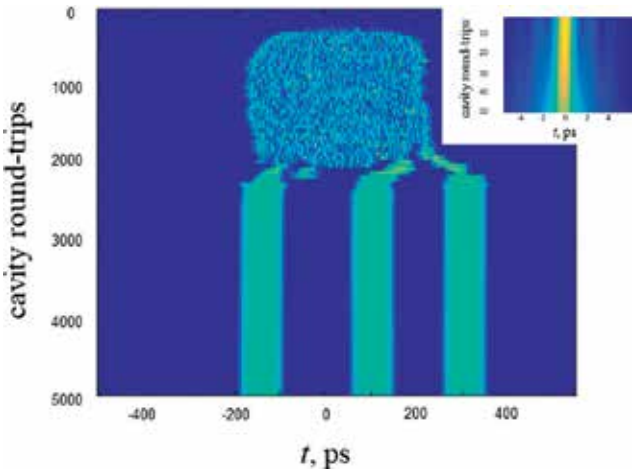

Figure 9. Contour plot of the DS power evolution and the corresponding autocorrelation function on the turbulent stage. The evolution coordinate is measured in the laser cavity round-trips, and the transverse coordinate corresponds to the local time measured in picoseconds. The energy parameter E_s normalized on $\gamma/\Delta t$ is of 3×10^5, where γ corresponds to a fused silica nonlinear coefficient and Δt=1 femtosecond is a time discretization step. Other parameters are: $\kappa = 0.1\gamma$, $\zeta = 0.05\gamma$, $\delta = 0.05$, and $\alpha = 40$ nm (Yb-fiber laser) [6].

characterized by the short-range correlation time about of 1 ps. In the process of evolution, an incoherent DS splits into three almost identical coherent solitons, which widths are lower substantially and the corresponding long-range/short-range correlation times become smaller/larger, respectively. Small long-range correlation time prevents the DSs merging and larger short-range correlation time provides DS coherence.

An analysis of turbulent DS demonstrates its complicate internal structure which can be interpreted as the complex of strongly interacting bright, dark, and gray DSs on a finite but strongly self-localized background concentrating almost the entire part of the energy.[8] In some sense, an appearance of DS turbulence resembles the laminar-turbulent transition in a fiber laser when a macroscopically coherent field ($\Xi \to 0$) becomes chaotically self-modulated [45].

Nevertheless, such a scenario is not unique. The turbulent dynamics can result from the strong interaction between "individual" DSs forming a "soliton gas" or turbulent "soliton cluster" (**Figure 10**) [1, 5, 47]. Interaction of such cluster with a low-intensity background field can result in permanent radiation or absorption of DSs in the form of so-called "soliton rains" [48, 49].

Separately, one may note the chaotization of DS dynamics caused by resonant interaction with the dispersive waves in the presence of higher-order corrections to the dispersion term in Eq. (1). In this case, the collisions between DS and dispersive wave, which radiates by it, results in a chaotic dynamic preserving, nevertheless the DS integrity (**Figure 11**) [23].

[8] One has to distinguish such a structure from the breather-like structures on a continuous-wave background. Such structures can demonstrate chaotic and rogue waves dynamics, as well (e.g., see [43, 44]).

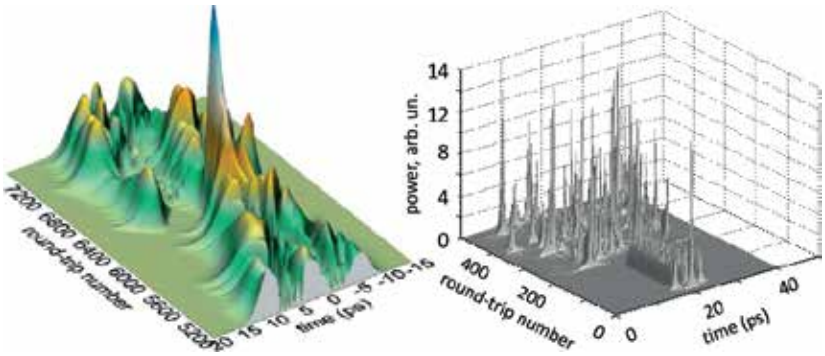

Figure 10. DS clusters in the form of a "persistent and coherent quasi-soliton" (left) [46] and a "sporadic rogue waves events that emerge from turbulent fluctuations" (right) [41].

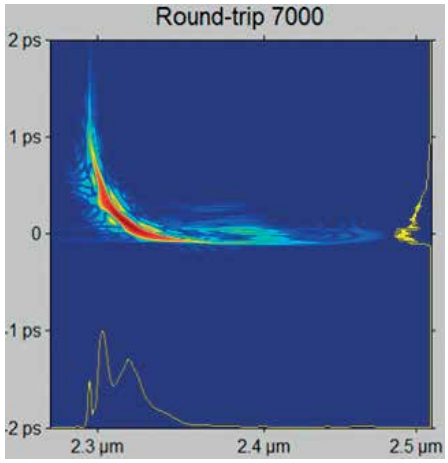

Figure 11. Wigner (time-spectral) diagram of the chaotic DS in the presence of third-order dispersion [23]. DS (dark-red region around 2.3 μm) radiates a dispersive wave (blue tail around 2.4 μm). As a result of the difference between their group velocities, DS collides permanently with a dispersive wave that causes chaotization of dynamics and modulation of both DS spectrum and time-profile.

4. Coherence of DS in the presence of nonlinearity with nonlocal/noninstantaneous response

A nonlinearity with the nonlocal/noninstanteneous response, which is of interest in optical context, can be taken into account by inclusion in Eq. (1) of the following term [50]

$$-i\gamma\Psi\int U(x-x')|\Psi|^2(T,x')dx'. \tag{7}$$

In the case of nonstationarity (i.e., $x \to t$ replacement), this equation describes the stimulated Raman scattering (SRS), for instance. Then, the response function is [6, 51]:

http://dx.doi.org/10.5772/intechopen.71478

$$U(t) = \frac{T_1^2 + T_2^2}{T_1 T_2^2} \exp\left(-\frac{t}{T_2}\right) \sin\left(\frac{t}{T_1}\right), \tag{8}$$

Where T_2 and T_1 define the effective relaxation time and resonant frequency of phonons in a nonlinear medium.

The simulations demonstrate [6, 35] that SRS suppresses the DS turbulence for the sufficiently large dispersions β. The first scenario is formation of uncoupled complex of DS and the dissipative Raman soliton (DRS) [6, 35, 42, 52] (**Figure 12**). One may assume, that such "energy discharging" is like the turbulence decay shown in **Figure 9**.

DRS is characterized by large chirp Θ and frequency down-shift. The last results from intra-pulse SRS which is possible due to a broad spectrum, which is a common characteristic of DS and results from its large Θ. A sole DRS develops with growing β (**Figure 13**) [6, 35]. It is

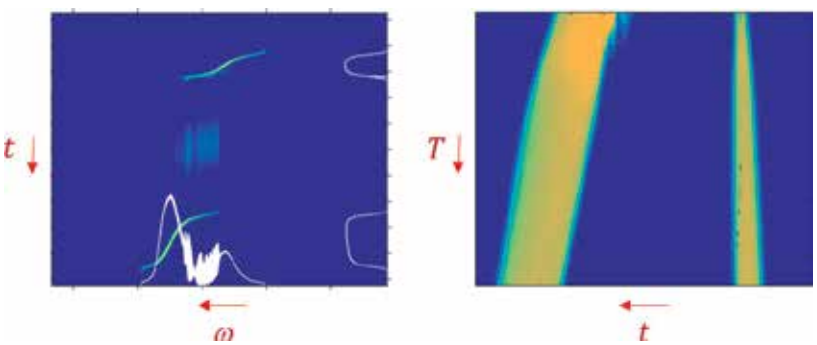

Figure 12. Wigner (time-spectral) diagram of the DS + DRS complex (left) and its evolution (contour plot of the field power; right) [35].

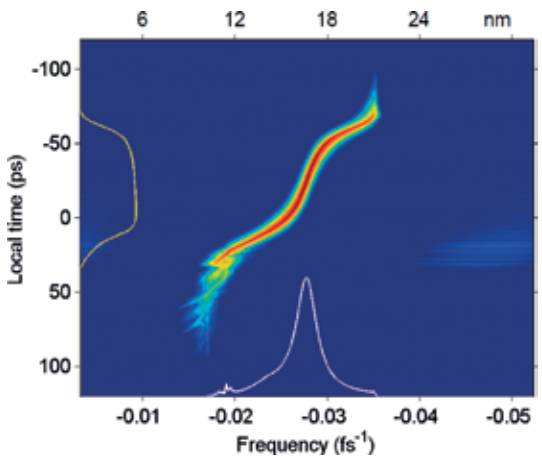

Figure 13. Wigner (time-spectral) diagram of a sole DRS developing for large β [6].

turbulence-free and characterized by perturbed anti-Stokes component, which is clearly visible on the Wigner diagram. Such perturbation induces a chaotic vibration of the DRS power [35]. Nevertheless, DRS exists within the parametric range, where an ordinary DS cannot develop in the turbulence-free regime. One may assume that the DRS stability results from passive negative feed-back based on interplay between nonlinear down-frequency shift due to SRS and spectral dissipation.

Another spectacular manifestation of the effect of a noninstantaneous nonlinearity on an incoherent field is an appearance of the spectral incoherent solitons (SIS) [53]. The spectrally localized soliton-like structures appear without any time-localization due to the causality property inherent to SRS so that a field cannot reach thermal equilibrium [54]. Formally, the corresponding evolution equation in the Langmuir turbulence limit has soliton-like solutions in the spectral domain [55, 56]. As a result, such structure localized in spectral domain possesses main properties of solitons including the property of elastic scattering.

5. Conclusions

The problem of DS coherence, chaotic, and turbulent dynamics has been outlined. A nontrivial internal structure of DS caused by it intensive energy exchange with dissipative environment allows conjecturing a close analogy with turbulent structure forming far from equilibrium. The existence of long-range correlation scale provides the DS energy scaling (or mass scaling for Bose-Einstein condensate). However, such "macroscopic" scaling is provided by strong phase inhomogeneity (chirp) so that internal coherence of DS defined by short-range correlation scale breaks and DS becomes a "cloud" of interacting "quasi-particles" or "glass of boiling water." Such structure is very sensitive to perturbation of even quantum level. Such extra-sensitivity combines macro- and micro-scales that raises an issue of the quantum theory of the macroscopic coherent, partially, and incoherent dissipative structures.

In the context of this work, such DS "decomposition" leads to turbulent dynamics and DS fragmentation. In particular, interactions inside such "soliton cluster" can result in the rogue waves' formation. An additional source of soliton destabilization is resonant interaction with a dispersive wave that results in chaotization of dynamics and formation of "soliton rains."

Nonlinearity with noninstanteneous response (e.g., SRS) leads to new interesting effect. In particular, SRS suppresses the DS turbulence due to the formation of DS + DRS pairs or sole DRS. Although DRS is turbulence-free within a broad parametric range, it has a perturbed anti-Stokes component, which causes chaotic vibrations of DRS parameters.

Another and spectacular manifestation of the noninstantaneous response of nonlinearity is the formation of SIS. This structure is a soliton in the spectral domain but incoherent and delocalized in the time domain.

The above-considered phenomena and conjectures are of interest in the context of the development of approaches to the self-consistent theory of nonequilibrium dissipative structures in classical and quantum aspects, which would use the optical DSs as a testbed.

Acknowledgements

The author acknowledges the support from Austrian Science Fund (FWF Project No. P24916-N27). Computational results have been achieved using the Vienna Scientific Cluster (VSC).

Author details

Vladimir L. Kalashnikov* and Evgeni Sorokin

*Address all correspondence to: vladimir.kalashnikov@tuwien.ac.at

Institute of Photonics, Vienna University of Technology, Vienna, Austria

References

[1] Zakharov V, Diasc F, Pushkarev A. One-dimensional wave turbulence. Physics Reports. 2004;**398**:1-65. DOI: 10.1016/j.physrep.2004.04.002

[2] Editorial. The power of analogies. Nature Photonics. 2014;**8**:1. DOI: 10.1038/nphoton.2013.359

[3] Akhmediev N, Ankiewicz A, editors. Dissipative Solitons. Berlin: Springer; 2005. DOI: 10.1007/b11728

[4] Kalashnikov VL, Sergeyev SV. Dissipative solitons in fibre lasers. In: Paul MC, editor. Fiber Laser. Rijeka: InTechOpen; 2016. p. 165-210

[5] Kalashnikov VL. Optics and chaos: Chaotic, rogue and noisy optical dissipative Solitons. In: Skiadas CH, Skiadas C, editors. Handbook of Applications of Chaos Theory. London: Chapman & Hall; 2016. p. 587-626

[6] Kalashnikov VL, Sorokin E. Dissipative Raman solitons. Optics Express. 2014;**22**:30118-30126. DOI: 10.1364/OE.22.030118

[7] Shats M, Punzmann H, editors. Turbulence and Coherent Structures in Fluids, Plasmas and Nonlinear Media. New Jersey: World Scientific; 2006

[8] Robinson PA. Nonlinear wave collapse and strong turbulence. Reviews of Modern Physics. 1997;**69**:507-574. DOI: 10.1103/RevModPhys.69.507

[9] Laurie J, Bortolozzo U, Nazarenko S, Residori S. One-dimensional optical wave turbulence: Experiment and theory. Physics Reports. 2012;**514**:121-175. DOI: 10.1016/j.physrep.2012.01.004

[10] Shrira V, Nazarenko S, editors. Advances in Wave Turbulence. New Jersey: World Scientific; 2013

[11] Nazarenko S. Wave Turbulence. Berlin: Springer; 2011

[12] Manneville P. Dissipative Structures and Weak Turbulence. San Diego: Academic Press; 1990

[13] Ilday FÖ. Turbulent times. Nature Photonics. 2013;**7**:767-769

[14] Benney DJ, Newell AC. The propagation of nonlinear wave envelopes. Studies in Applied Mathematics. 1967;**46**:133-139. DOI: 10.1002/sapm1967461133

[15] Dyachenko S, Newell AC, Pushkarev A, Zakharov VE. Optical turbulence: Weak turbulence, condensates and collapsing filaments in the nonlinear Schrödinger equation. Physica D: Nonlinear Phenomena. 1992;**57**:96-160. DOI: 10.1016/0167-2789(92)90090-A

[16] Kevrekidis PG, Frantzeskakis DJ, Carretero-González R, editors. Emergent Nonlinear Phenomena in Bose-Einstein Condensates. Berlin: Springer; 2008

[17] Churkin DV, Kolokolov IV, Podivilov EV, Vatnik ID, Nikulin MA, Vergeles SS, Terekhov S, Lebedev VV, Falkovich G, Babin SA, Turitsyn SK. Wave kinetics of random fibre lasers. Nature Communications. 2015;**6**:6214. DOI: 10.1038/ncomms7214

[18] Zakharov VE, editor. What is Integrability? Berlin: Springer; 1991

[19] Zakharov VE, Shabat AB. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. Soviet Physics—JETP. 1972;**34**: 62-69

[20] Akhmediev NN, Ankiewicz A. Solitons: Nonlinear Pulses and Beams. London: Chapman & Hall; 1997

[21] Grelu P, Akhmediev N. Dissipative solitons for mode-locked lasers. Nature Photonics. 2012;**6**:84-92. DOI: 10.1038/NPHOTON.2011.345

[22] Akhmediev N, Ankiewicz A, editors. Dissipative Solitons: From Optics to Biology and Medicine. Berlin: Springer; 2008. DOI: 10.1007/978-3-540-78217-9

[23] Sorokin E, Tolstik N, Kalashnikov VL, Sorokina IT. Chaotic chirped-pulse oscillators. Optics Express. 2013;**21**:29567-29577. DOI: 10.1364/OE.21.029567

[24] Rumpf B, Newell AC. Localization and coherence in nonintegrable systems. Physica D: Nonlinear Phenomena. 2003;**184**:162-191. DOI: 10.1016/S0167-2789(03)00220-3

[25] Picozzi A, Haeltermann M, Pitois S, Millot G. Incoherent solitons in instantaneous response nonlinear media. Physical Review Letters. 2004;**92**:143906. DOI: 10.1103/PhysRevLett.92.143906

[26] Aschieri P, Garnier J, Michel C, Doya V, Picozzi A. Condensation and thermalization of classical optical waves in a waveguide. Physical Review A. 2011;**83**:033838. DOI: 10.1103/PhysRevA.83.033838

[27] Kalashnikov VL. Chirped-pulse oscillators: Route to the energy scalable femtosecond pulses. In: Al-Khursan AH, editor. Solid-State Laser. Rijeka: InTechOpen; 2012. p. 145-184

[28] Podivilov E, Kalashnikov VL. Heavily-chirped solitary pulses in the normal dispersion region: New solutions of the cubic-quintic Ginzburg-Landau equation. JETP Letters. 2005;**82**:467-471

[29] Akhmediev N, Królikowski W, Snyder AW. Partially coherent solitons of variable shape. Physical Review Letters. 1998;**81**:4632-4635. DOI: 10.1103/PhysRevLett.81.4632

[30] Kalashnikov VL, Chernykh A. Spectral anomalies and stability of chirped-pulse oscillators. Physical Review A. 2007;**75**:033820. DOI: 10.1103/PhysRevA.75.033820

[31] Kalashnikov VL. Dissipative solitons: Perturbations and chaos formation. In: III Chaotic Modelling and Simulation International Conference; June 1–4, 2010; Chania, Greece; 2010.

[32] Yoon B, Negele JW. Time-dependent approximation for a one-dimensional system of bosons with attractive δ-function interactions. Physical Review A. 1977;**16**:1451-1457. DOI: 10.1103/PhysRevA.16.1451

[33] Lai Y, Haus HA. Quantum theory of solitons in optical fibers. I. Time-dependent Hartree approximation. Physical Review A. 1989;**40**:844-853. DOI: 10.1103/PhysRevA.40.844

[34] Lai Y, Haus HA. Quantum theory of solitons in optical fibers. II. Exact solution. Physical Review A. 1989;**40**:854-866. DOI: 10.1103/PhysRevA.40.854

[35] Kalashnikov VL, Sorokin E. Turbulence of optical dissipative solitons. In: X Chaotic Modelling and Simulation International Conference; May 30–June 2, 2017; Barcelona, Spain; 2017.

[36] Newell AC, Zakharov VE. Optical turbulence. In: Tabeling P, Cardoso O, editors. Turbulence: A Tentative Dictionary. New York: Plenum Press; 1993. p. 59-66

[37] Picozzi A. Towards a nonequilibrium thermodynamic description of incoherent nonlinear optics. Optics Express. 2007;**15**:9063-2083

[38] D'yachenko AI, Zakharov VE, Pushkarev AN, Shvets VF, Yan'kov VV. Soliton turbulence in nonintegrable wave systems. Soviet Physics—JETP. 1990;**69**:1144-1147

[39] Düring G, Picozzi A, Rica S. Breakdown of weak-turbulence and nonlinear wave condensation. Physica D: Nonlinear Phenomena. 2009;**238**:1524-1549. DOI: 10.106/j.physd.2009.04.014

[40] Pomeau Y. Asymptotic time behavior of nonlinear classical field equations. Nonlinearity. 1992;**5**:707-720

[41] Kalashnikov VL. Dissipative solitons in presence of quantum noise. Chaotic Modeling and Simulation (CMSIM). 2014:29-37

[42] Babin SA, Podivilov EV, Kharenko DS, Bednyakova AE, Fedoruk MP, Shtyrina OV, Kalashnikov VL, Apolonski AA. SRS-driven evolution of dissipative solitons in fiber lasers. In: Grelu P, editor. Nonlinear Optical Cavity Dynamics: From Microresonators to Fiber Optics. Weinheim: Wiley-VCH; 2016. p. 277-346

[43] Onorato M, Residori S, Baronio F, editors. Rogue and Shock Waves in Nonlinear Dispersive Media. Switzerland: Springer; 2016. DOI: 10.1007/978-3-319-39214-1

[44] Genty G, Närhi M, Amiot C, Jacquet M. Supercontinuum generation in optical fibers. In: Faccio D, Dudley J, Clerici M, editors. Frontiers in Modern Optics (Proceedings of the International School of Physics). Amsterdam: IOS Press; 2016. p. 233-262

[45] Turitsyna EG, Smirnov SV, Sugavanam S, Tarasov N, Shu X, Babin SA, Podivilov EV, Churkin DV, Falkovich G, Turitsyn SK. The laminar-turbulent transition in a fibre lase. Nature Photonics. 2013;**7**:783-786. DOI: 10.1038/NPHOTON.2013.246

[46] Zavyalov A, Egorov O, Iliew R, Lederer F. Physical Review A. 2012;**85**:013828

[47] Hammani K, Kibler B, Finot C, Picozzi A. Emergence of rogue waves from optical turbulence. Physics Letters A. 2010;**374**:3585-3589

[48] Chouli S, Grelu P. Rains of solitons in a fiber laser. Optics Express. 2009;**17**:11776-11781

[49] Chouli S, Grelu P. Soliton rains in a fiber laser: An experimental study. Physical Review A. 2010;**81**:063829

[50] Garnier J, Lisak M, Picozzi A. Toward a wave turbulence formulation of statistical nonlinear optics. Journal of the Optical Society of America B: Optical Physics. 2012;**29**: 2229-2242

[51] Agrawal G. Nonlinear Fiber Optics. Amsterdam: Elsevier; 2013

[52] Babin SA, Podivilov EV, Kharenko DS, Bednyakova AE, Fedoruk MP, Kalashnikov VL, Apolonski A. Multicolour nonlinearly bound chirped dissipative solitons. Nature Communications. 2014;**5**:4653. DOI: 10.1038/ncomms5653

[53] Picozzi A, Pitois S, Millot G. Spectral incoherent solitons: A localized soliton behavior in the frequency domain. Physical Review Letters. 2008;**101**:093901. DOI: 10.1103/PhysRevLett. 101.093901

[54] Garnier J, Pikozzi A. Unified kinetic formulation of incoherent waves propagating in nonlinear media with noninstantaneous response. Physical Review A. 2010;**81**:033831. DOI: 10.1103/PhysRevA.81.033831

[55] Musher SL, Rubenchik AM, Zakharov VE. Weak Langmuir turbulence. Physics Reports. 1995;**252**:177-274. DOI: 10.1016/0370-1573(94)00071-A

[56] Michel C, Kibler B, Picozzi A. Discrete spectral incoherent solitons in nonlinear media with noninstantaneous response. Physical Review A. 2011;**83**:023806. DOI: 10.1103/PhysRevA. 83.023806

Chapter 7

Interaction of Solitons with the Electromagnetic Field in Classical Nonlinear Field Models

Jon C. Luke

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71215

Abstract

Solitary waves and solitons are briefly discussed in their historical context, especially regarding the motivation to use solitons for particle models in the nonlinear Klein-Gordon (NKG) equation. Conservation equations for charge, energy, and momentum follow from Noether's theorem in the usual way, and the NKG equation can be coupled to Maxwell's equations in the standard, gauge invariant manner. A recently proposed model is summarized in which two NKG equations are coupled to the electromagnetic field. In that model, solitons mimic the dynamical behavior of electrons and protons. A new result is then presented that follows from that model. Although the model is purely classical, it turns out that the arc spectrum of hydrogen is emitted into the electromagnetic field when small oscillations of one Klein-Gordon equation occur in the vicinity of a proton-like soliton. It is perhaps unexpected that a purely classical model can exhibit behavior suggestive of a phenomenon that is generally presumed to occur only in a quantum-mechanical context. Because of the way in which that result occurs, however, it is not clear whether there is any possible relevance to the actual physical phenomenon.

Keywords: nonlinear systems and models, solitons, nonlinear Klein-Gordon-Maxwell equations, particle models, hydrogen spectrum, nonlinear field theories

1. Introduction

When a term that acts as a restoring force is added to the usual linear wave equation, the equation

$$\frac{\partial^2 \psi}{\partial t^2} - \nabla^2 \psi + \psi = 0 \tag{1}$$

is obtained. Schweber ([1], p. 54) points out "When Schrödinger wrote down the nonrelativistic equation now bearing his name, he also formulated the corresponding relativistic

equation." Within the year of 1926, Klein, Gordon, and at least three others independently proposed the use of Eq. (1) as the relativistic generalization of Schrödinger's equation. Eq. (1) has since become known as the Klein-Gordon equation. Instead of the original time and space variables $\hat{t}, \hat{x}, \hat{y}, \hat{z}$ that might be used in a quantum-mechanical context, we have chosen to write Eq. (1) using dimensionless variables t, x, y, z defined as

$$t = \frac{m c^2}{\hbar}\hat{t}, \quad x = \frac{mc}{\hbar}\hat{x}, \quad y = \frac{mc}{\hbar}\hat{y}, \quad z = \frac{mc}{\hbar}\hat{z}, \tag{2}$$

where $\hat{t}, \hat{x}, \hat{y}, \hat{z}$ are time and space coordinates in customary units (e.g., seconds and meters). Similarly, m is the mass of the electron, c is the speed of light, and $\hbar$ is Planck's constant divided by 2π, also in customary units. Unit distance in our dimensionless coordinates corresponds to $\hbar/(mc)$, which is about 3.86×10^{-13} m. That distance is known as the Compton wavelength of the electron (divided here by 2π), which is a fundamental unit of length for quantum-mechanical phenomena that involve electrons.

When one attempts to use the Klein-Gordon equation to describe the quantum-mechanical problem of an electron in the Coulomb field of a nucleus of atomic number Z, Eq. (1) becomes

$$\left(\frac{\partial}{\partial t} + i\,\frac{Z\,\alpha_{fs}}{r}\right)\left(\frac{\partial}{\partial t} + i\,\frac{Z\,\alpha_{fs}}{r}\right)\psi - \nabla^2\psi + \psi = 0, \tag{3}$$

where α_{fs} is the fine structure constant (approximately 1/137), where $i = \sqrt{-1}$, and where $r^2 = x^2 + y^2 + z^2$. Unit distance for the dimensionless quantity r is thus also $\hbar/(mc)$. Although Eq. (3) reduces properly to Schrödinger's equation in the nonrelativistic approximation, it was realized almost immediately that Eq. (3) was in fact not appropriate for the quantum-mechanical description of the electron, and the Klein-Gordon equation fell even into a degree of disrepute, especially after the remarkable insight that led to Dirac's equation in 1928 [1, 2].

The Klein-Gordon equation is, nevertheless, a simple example of a relativistic partial differential equation and, as such, it is second in importance only to the wave equation. The relativistic invariance is preserved even when the ψ term is replaced with a term that depends nonlinearly on ψ, and the resulting equation can be used, for example, in the study of nonlinear waves [3, 4]. This chapter will deal primarily with the nonlinear Klein-Gordon (NKG) equation in the form

$$\frac{\partial^2\psi}{\partial t^2} - \nabla^2\psi + W'\left(\psi\bar{\psi}\right)\psi = 0, \tag{4}$$

where ψ depends on time t and spatial coordinates x, y, z. Overbar denotes the complex conjugate since ψ will usually be taken as complex-valued. The real-valued function W is used to introduce an appropriate nonlinearity, and $W'(\psi\bar{\psi})$ denotes the derivative of $W(\psi\bar{\psi})$ with respect to its argument $\psi\bar{\psi}$. Let us take $W(0) = 0$, and with suitable scaling we can also take $W'(0) = 1$, so that Eq. (4) will agree with Eq. (1) when ψ is small.

Here, we are interested in the soliton solutions that occur in the NKG equation, so Section 2 is a brief historical overview of solitary waves and solitons, followed by the historical context in Section 3 that motivates the use of solitons in classical field theories. Before coupling the NKG equation to Maxwell's equations in Section 8, we need to examine the NKG equation

itself in more detail in Sections 5 and 6. An explicit example of a soliton solution is given in Section 7. In Sections 9–11, we summarize a recently proposed model in which a second NKG equation is introduced so as to model solitons that can be thought of as protons as well as electrons. In that model, the proton and electron-like solitons attract and repel each other in the desired way. Finally, in Section 12, we present new results from the same model where it is shown that small oscillations in the neighborhood of a proton-like soliton emit the well-known frequencies of the hydrogen spectrum into the electromagnetic field. Since emission of the hydrogen spectrum is a phenomenon that would normally be expected to occur only in a quantum-mechanical setting and not in the present classical field model, we interpret this result in Section 13 and suggest the future direction of the present research.

2. Historical background of solitary waves and solitons

The study of solitary waves has a long history [4], which originated with observations by Scott Russell ([5], [6, p. 118]), of an isolated wave moving in a canal. Following along on horseback, he was able to observe that the wave traveled a great distance with little change in form and only gradually died out. The linear theories of water waves available at that time predicted, however, that such waveshapes would necessarily disperse, so for some time Scott Russell's observations were not taken seriously. The full equations for water waves, even in the irrotational, inviscid case, present quite a formidable boundary value problem, so various simple model equations have been examined in much greater detail. Strangely in contrast to the linear theory, a certain nonlinear model of water waves for shallow water predicted that waves of all shapes would steepen and ultimately break. Eventually, various simple, nonlinear model equations such as the equation of Korteweg and deVries (often known as the KdV equation)

$$\frac{\partial \eta}{\partial t} + (c_0 + c_1 \eta) \frac{\partial \eta}{\partial x} + \nu \frac{\partial^3 \eta}{\partial x^3} = 0, \tag{5}$$

were formulated, where both the dispersion and steepening effects are included. In these equations, the two effects balance each other and do, in fact, allow an isolated waveshape to propagate without change of form.

The term soliton was later introduced to emphasize that solitary waves, like particles, tend to maintain their identity. When a nonlinear term is introduced into the usual linear Schrödinger equation, the nonlinear Schrödinger equation

$$i \frac{\partial \psi}{\partial t} - \nabla^2 \psi + W'(\psi \bar{\psi}) \psi = 0, \tag{6}$$

is obtained. It has soliton solutions and is of particular interest in relation to the ideas of de Broglie and Bohm [7, 8]. Soliton solutions occur in the NKG Eq. (4) and also in the related equation

$$\frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial x^2} + \sin \psi = 0 \tag{7}$$

known (a bit flippantly) as the sine-Gordon equation.

Remarkable methods, such as the inverse scattering method, have been devised for use with certain model equations. With these methods, it is possible to find exact solutions that show solitons colliding and passing through each other while still maintaining their identity. The subject of solitary waves and solitons has been extensively described in books and review articles [9–13]. For more recent work, see [14–24] and references therein.

3. Use of solitons as particle models

The motivation for the use of solitons as particle models can best be understood by comparison with the historical development of physical theories, especially electromagnetism and quantum mechanics. Despite the successes of Maxwell's theory of electromagnetism in the late 1800s and the clarified understanding of its transformation properties according to Einstein's special theory of relativity, classical physics seemed to reach an impasse at the beginning of the twentieth century. The notion of the electron as a point particle was immediately inconsistent with electromagnetism since the electromagnetic energy around a charged point particle is easily calculated to be infinite. Moreover, if an electron were to orbit a proton like a planet around the sun, it would radiate into the electromagnetic field at a frequency based on the orbital speed and quickly spiral inward toward the nucleus.

Early in the twentieth century, however, Planck, Einstein, and Bohr advanced daring new hypotheses to describe blackbody radiation, the photoelectric effect, and the hydrogen atom, respectively. Dirac, Heisenberg, and others later developed various versions of a new quantum mechanics more sophisticated but certainly no more intuitively comprehensible. Quantum mechanics has been perpetually troubled by divergences, as when Bethe [25] somewhat jokingly referred to a quantity that "…comes out infinite in all existing theories, and has therefore always been ignored." Remarkable progress was later achieved in quantum electrodynamics by Tomonaga, Schwinger, Feynman, and others, but still only with the aid of a rather arbitrary (some would say procrustean) renormalization procedure. Dirac [2] concludes his book with the thought that "It would seem that we have followed as far as possible the path of logical development of the ideas of quantum mechanics as they are at present understood. The difficulties, being of a profound character, can be removed only by some drastic change in the foundations of the theory, probably a change as drastic as the passage from Bohr's orbit theory to the present quantum mechanics."

There have been various attempts, notably by de Broglie [7, 26], to gain insight into a possible "drastic change in the foundations of the theory" as envisioned by Dirac, but so far such attempts have had little success. One possible approach, which apparently occurred to many people, is to ask whether solitons could be used to model actual particles as localized regions where the field is large. If point particles were replaced by solitons one could hope for a theory along more classical lines where problems with infinite, divergent integrals would be avoided in a natural way. The line of research described here is meant to be a step, if only a tiny step, in that direction. A recently proposed model is summarized in Sections 9–11. A new result is presented in Section 12, where a phenomenon suggestive of quantum-mechanics occurs, but in an unexpected context.

http://dx.doi.org/10.5772/intechopen.71215

4. Topological and nontopological solitons

It was found early on that if one tries to use the NKG Eq. (4) to form a localized, particle-like solution, where ψ is a static, real-valued function of spatial coordinates x, y, z, that such a solution turns out to be wildly unstable. Even the linear stability analysis shows that there is a mode that grows in time, so the solution either tends to collapse or to grow arbitrarily large. Hobart [27] and Derrick [28] concluded from energy considerations that a wide class of field equations would turn out to be unstable in this way. One possible way around this difficulty is for the solution to have a suitable topological property. Scott [29] and Rubinstein [30] have examined a simple example of such a topological soliton in the sine-Gordon Eq. (7). Scott's interpretation by means of a mechanical model is especially easy to visualize, where ψ represents an angle in the y, z plane, and the solution loops around the x-axis.

Here, however, we will be concerned with another approach to stability that uses the NKG Eq. (4) but with a complex-valued variable ψ that has a time-dependent phase. Thus, we will be interested in possible solutions of the form

$$\psi = U(r)\, e^{i\omega t}, \tag{8}$$

where $r^2 = x^2 + y^2 + z^2$ and where the real-valued function $U(r)$ is exponentially small for large r. The solution is stationary in the sense that U does not depend on t, but it is not static since ψ is in effect rotating with angular frequency ω in the complex plane. Some of the early works along these lines were done by Glasko et al. [31], Zastavenko [32], and Rosen [33]. It turns out that orbital stability can be achieved in suitable circumstances. Roughly speaking, the rotation allows stability to occur just as a spinning gyroscope can be stable in a situation that would otherwise be unstable from energetic considerations. Solitons of this nature are often called nontopological solitons, in contrast to the topological solitons that occur in examples such as the sine-Gordon Eq. (7). The term soliton is sometimes reserved for particle-like solutions that have been shown to be orbitally stable; but here, we will allow the use of the term for particle-like solutions in general. One difficulty with the use of nontopological solitons as particle models is that a whole family of particle-like solutions typically occurs, as will be seen in the example of Section 7. This problem will be discussed further in Section 10. Existence and stability questions have been extensively studied for the NKG equation [15, 34]. More recently, soliton solutions with nonzero angular momentum have also been of interest [18, 19].

5. Conservation equations for the NKG equation

To understand soliton solutions for the NKG equation, it is important first to see that the NKG equation can be derived from a variational principle. The Lagrangian density is

$$L = \frac{1}{2}\frac{\partial\psi}{\partial t}\frac{\partial\bar{\psi}}{\partial t} - \frac{1}{2}\,\nabla\psi\cdot\nabla\bar{\psi} - \frac{1}{2}\,W(\psi\bar{\psi}), \tag{9}$$

and the corresponding Euler operator (with respect to ψ) is

$$Op_{\psi} = -\frac{\partial^2\bar{\psi}}{\partial t^2} + \nabla^2\bar{\psi} - W'(\psi\bar{\psi})\bar{\psi}. \tag{10}$$

When we set the Euler operator (10) to zero (and take the negative of the complex conjugate), we obtain the NKG Eq. (4). It is well known from Noether's theorem [4, 35, 36] that each symmetry of the variational principle leads to an equation in what is referred to as conservation form. Once the equation is in conservation form, it is straightforward to integrate in order to show that a certain quantity is conserved. Symmetry with respect to time translation thus leads to conservation of energy, and symmetry with respect to translation in x, y, z space leads to conservation of momentum.

Since the variational principle that follows from Eq. (9) is unchanged under time translation, it follows that energy is conserved. Specifically, we can rewrite the expression

$-\left(Op_{\psi}\partial\psi/\partial t + \overline{Op_{\psi}}\partial\bar{\psi}/\partial t\right)/2$ in the following conservation form:

$$\begin{aligned} -\frac{1}{2}\left(Op_{\psi}\frac{\partial\psi}{\partial t} + \overline{Op_{\psi}}\frac{\partial\bar{\psi}}{\partial t}\right) &= \frac{\partial}{\partial t}\frac{1}{2}\left(\frac{\partial\psi}{\partial t}\frac{\partial\bar{\psi}}{\partial t} + \nabla\psi\cdot\nabla\bar{\psi} + W(\psi\bar{\psi})\right) \\ &\quad -\frac{1}{2}\nabla\cdot\left(\frac{\partial\psi}{\partial t}\ \nabla\bar{\psi} + \frac{\partial\bar{\psi}}{\partial t}\nabla\psi\right). \end{aligned} \tag{11}$$

Let

$$\mathscr{E} = \int\frac{1}{2}\left(\frac{\partial\psi}{\partial t}\frac{\partial\bar{\psi}}{\partial t} + \nabla\psi\cdot\nabla\bar{\psi} + W(\psi\bar{\psi})\right)d^3x, \tag{12}$$

where the integral is to be taken over a certain region in x, y, z space and where d^3x denotes a volume element. Then, $\mathscr{E}$ is defined as the energy in that region. Let us suppose that Eq. (4) is satisfied, so that the right-hand side of Eq. (11) must be equal to zero. Then, if Eq. (11) is integrated over the region in question, it is clear from the divergence theorem that the energy within that region will remain constant in time, except for any energy that enters or leaves the region through the outer surface.

Similarly, since the variational principle is invariant under translation in the x direction, a conservation equation can be found for momentum in the x direction. We can thus rewrite the expression $\left(Op_{\psi}\partial\psi/\partial x + \overline{Op_{\psi}}\partial\bar{\psi}/\partial x\right)/2$ in the following conservation form:

$$\begin{aligned} \frac{1}{2}\left(Op_{\psi}\frac{\partial\psi}{\partial x} + \overline{Op_{\psi}}\frac{\partial\bar{\psi}}{\partial x}\right) &= \frac{\partial}{\partial t}\left(-\frac{1}{2}\frac{\partial\psi}{\partial t}\frac{\partial\bar{\psi}}{\partial x} - \frac{1}{2}\frac{\partial\bar{\psi}}{\partial t}\frac{\partial\psi}{\partial x}\right) \\ &\quad + \frac{\partial}{\partial x}\left(\frac{1}{2}\frac{\partial\psi}{\partial t}\frac{\partial\bar{\psi}}{\partial t} - \frac{1}{2}\ \nabla\psi\cdot\nabla\bar{\psi} - \frac{1}{2}W(\psi\bar{\psi})\right) \\ &\quad + \nabla\cdot\left(\frac{1}{2}\frac{\partial\bar{\psi}}{\partial x}\ \nabla\psi + \frac{1}{2}\frac{\partial\psi}{\partial x}\ \nabla\bar{\psi}\right). \end{aligned} \tag{13}$$

Then

$$P_x = \int\left(-\frac{1}{2}\frac{\partial\psi}{\partial t}\frac{\partial\bar{\psi}}{\partial x} - \frac{1}{2}\frac{\partial\bar{\psi}}{\partial t}\frac{\partial\psi}{\partial x}\right)d^3x \tag{14}$$

is defined as the momentum in the x direction (as contained within the region in question). Again, the use of the divergence theorem shows that momentum within the region

is conserved, except for any momentum that enters or leaves through the outer surface. Conservation of momentum in the y and z directions is similar.

Next, because the variational principle is invariant when ψ is rotated in the complex plane, the quantity $i\left(\psi\, O\, p_\psi - \overline{\psi\, O\, p_\psi}\right)/2$ can be put in the conservation form

$$\frac{i}{2}\left(\psi\, O\, p_\psi - \overline{\psi\, O\, p_\psi}\right) = \frac{\partial}{\partial t}\left(\frac{i\bar{\psi}}{2}\frac{\partial\psi}{\partial t} - \frac{i\psi}{2}\frac{\partial\bar{\psi}}{\partial t}\right) + \nabla\cdot\left(\frac{i\psi}{2}\,\nabla\bar{\psi} - \frac{i\bar{\psi}}{2}\,\nabla\psi\right). \tag{15}$$

The quantity

$$Q = \int\left(\frac{i\bar{\psi}}{2}\frac{\partial\psi}{\partial t} - \frac{i\psi}{2}\frac{\partial\bar{\psi}}{\partial t}\right) d^3x \tag{16}$$

will be defined as the charge (contained within the region in question), for reasons that will become clear when the NKG equation is coupled to Maxwell's equations. Again, by the divergence theorem, the charge within the region must remain constant in time except for any charge that enters or leaves through the outer surface.

6. Rotating nonlinear solutions for the NKG equation

Now let us look for solutions of the NKG equation that are rotating in the complex plane with angular frequency ω, that is

$$\psi = U e^{i\omega t}, \tag{17}$$

where U is a real function that depends on x, y, z, but is independent of t. Then, U needs to satisfy

$$\nabla^2 U + \omega^2 U - W'(U^2)U = 0. \tag{18}$$

Generally, we will be interested in localized solutions, and U will typically be exponentially small when r is large. Substitution of expression (17) in Eqs. (16) and (12) shows that the charge of such a solution is

$$Q = -\omega\int U^2\, d^3x, \tag{19}$$

and the energy is

$$\mathscr{E} = \int\tfrac{1}{2}(\omega^2 U^2 + \nabla U\cdot\nabla U + W(U^2))\, d^3x, \tag{20}$$

where the integrals are now taken over all of x, y, z space. From Eq. (19), it is clear that the charge will be positive or negative depending on the sign of ω. We are not interested here in the case where Q is zero since we want U to be nontrivial, and ω needs to be nonzero when we try for an orbitally stable soliton solution.

If we look for a radially symmetric solution $U(r)$, where $r^2 = x^2 + y^2 + z^2$, we note that

$$\nabla^2 U = \frac{1}{r}\frac{d^2(rU)}{dr^2} \tag{21}$$

so that Eq. (18) can be rewritten as

$$\frac{d^2\mathcal{U}}{dr^2} + \omega^2 \mathcal{U} - W'(U^2)\mathcal{U} = 0, \tag{22}$$

with $\mathcal{U} = rU$.

7. An explicit soliton solution for the NKG equation

To fix ideas, let us derive the following standard result in somewhat more than the usual detail. An expedient that has been frequently used [37–40] is to let W' be a step function, in which case a simple example of a soliton solution can be worked out explicitly. Let us suppose that W' has a fixed value a_W when r is greater than a certain r_0. Then, we want $a_W - \omega^2 > 0$, so we can find a solution of Eq. (22) of the form $\mathcal{U} = B\,e^{-b(r-r_0)}$ for $r > r_0$, where B is a constant and $b^2 = a_W - \omega^2$. To satisfy an appropriate boundary condition at infinity, we have chosen the solution where $\mathcal{U}$ approaches 0 for large r. Next, let us suppose that W' has a somewhat smaller fixed value $a_W - b_W$ when $r < r_0$. We want the constant b_W to be a suitable value so that $a_W - b_W - \omega^2$ will be negative. Then, we can find a solution of Eq. (22) in the form $\mathcal{U} = A\sin(ar)$ for $r < r_0$, where A is a constant and $a^2 = \omega^2 - a_W + b_W$. We have chosen the sine rather than the cosine solution so as to satisfy an appropriate boundary condition at $r = 0$. Then, we note that

$$a^2 + b^2 = b_W. \tag{23}$$

We need $A\sin(ar_0) = B$ to make $\mathcal{U}$ continuous at r_0, and we need $Aa\cos(ar_0) = -Bb$ to make $d\mathcal{U}/dr$ continuous at r_0, so, altogether, we want

$$a\cos(a\,r_0) = -b\sin(a\,r_0). \tag{24}$$

There will be a whole family of rotating solutions for Eq. (18), as was mentioned in Section 4. It is convenient to use $\theta_0 = ar_0$ as a parameter for this family. Using Eqs. (23) and (24), we find that $a^2 = b_W \sin^2\theta_0$ and $b^2 = b_W \cos^2\theta_0$. We will be looking for solutions with θ_0 between $\pi/2$ and π (where $\sin\theta_0$ is positive but $\cos\theta_0$ is negative) so we find

$$a = \sqrt{b_W}\sin\theta_0, \quad b = -\sqrt{b_W}\cos\theta_0. \tag{25}$$

If we wish to have U take the value U_W at r_0, then

$$B = \frac{U_W\theta_0}{a}, \quad A = \frac{U_W\theta_0}{a\sin\theta_0}. \tag{26}$$

Collecting the above results, we have

$$U = \frac{\mathcal{U}}{r} = \frac{\theta_0 U_W \sin(\hat{r}\sin\theta_0)}{\hat{r}\sin^2\theta_0} \tag{27}$$

for $\hat{r} \le \hat{r}_0$ and

$$U = \frac{\mathcal{U}}{r} = \frac{\theta_0 U_W e^{(\hat{r}-\hat{r}_0)\cos\theta_0}}{\hat{r}\sin\theta_0} \tag{28}$$

for $\hat{r} > \hat{r}_0$. The result has been written in terms of scaled quantities

$$\hat{r} = \sqrt{b_W}\, r, \quad \hat{r}_0 = \sqrt{b_W}\, r_0 = \theta_0/\sin\theta_0. \tag{29}$$

Then, we need $W'(U^2) = a_W$ for $U^2 < U_W^2$ and $W'(U^2) = a_W - b_W$ for $U^2 > U_W^2$. Given values of a_W and b_W, as well as a value of the parameter θ_0, we find that ω is to be determined from the equation

$$\omega^2 = a_W - b_W \cos^2\theta_0. \tag{30}$$

The shape of the solution, as given by the dependence of U on $\hat{r}$ in Eqs. (27) and (28), depends on θ_0 but not on a_W or b_W. For now, let us take $a_W = 1$. It is easy to obtain U in terms of the original variable r using Eq. (29). Eq. (29) shows that the physical size of the soliton, i.e., its size in x, y, z space, will be larger when b_W is smaller.

The charge Q and energy $\mathscr{E}$ of such a soliton solution follow from the integrals (19) and (20). It turns out that the integral for the charge can be worked out explicitly as

$$\begin{aligned} Q &= -\omega \int_0^\infty U^2\, 4\pi\, r^2\, dr \\ &= -\sqrt{1 - b_W\cos^2\theta_0}\, \frac{2\pi(\theta_0\cos\theta_0 - \sin\theta_0)\, b_W^{-3/2}\, U_W^2\, \theta_0^2}{(\sin\theta_0)^5\cos\theta_0}. \end{aligned} \tag{31}$$

If a graph of charge Q is plotted (say for $b_W = 0.1$) with respect to the parameter θ_0, it is seen that a minimum value of $|Q|$ occurs. Expressed in another way, the family of solutions bifurcates as $|Q|$ is increased.

8. Coupling of the NKG equation to the electromagnetic field

Next, we want to couple the NKG equation to electromagnetism to obtain the nonlinear Klein-Gordon-Maxwell (NKGM) system of equations, as has been done by Rosen [37], Morris [41], and others. Again it is important to express the coupling of the NKG equation to the electromagnetic field in terms of a variational principle so that the conservation laws of interest can be obtained from Noether's theorem. Although Maxwell's equations are usually expressed in terms of the quantities **E**, **D**, **H**, **B**, it turns out that, to get a suitable variational principle, we need to work instead with the scalar potential A_0 and the vector potential **A**, which are defined in terms of the relations

$$\nabla A_0 - \frac{\partial \mathbf{A}}{\partial t} = \mathbf{E}, \quad \nabla \times \mathbf{A} = \mathbf{B}. \tag{32}$$

Here, the notation A_0 is used for the scalar potential to emphasize that it can be thought of as the time component of a Lorentz 4-vector, while $\mathbf{A} = (A_1, A_2, A_3)$ gives the three spatial components. The gauge condition can then be taken in the form

$$\frac{\partial A_0}{\partial t} - \nabla \cdot \mathbf{A} = 0. \tag{33}$$

It should be noted that $A_0 = -\Phi$ has the opposite sign from the symbol Φ that is often used to designate the scalar potential [42]. Eq. (4) can be coupled to the electromagnetic field in the usual gauge invariant way by replacing $\partial/\partial t$ by $\partial/\partial t - iA_0$ and by replacing ∇ by $\nabla - i\mathbf{A}$, where $i = \sqrt{-1}$. The resulting equation for ψ is

$$\left(\frac{\partial}{\partial t} - i A_0\right)\left(\frac{\partial}{\partial t} - i A_0\right)\psi - (\nabla - i\mathbf{A}) \cdot (\nabla - i\mathbf{A})\, \psi + W'(\psi\bar{\psi})\psi = 0. \tag{34}$$

When oscillations of ψ are small so that $W'(\psi\bar{\psi})$ is effectively 1, we want Eq. (34) to reduce to the linear Klein-Gordon equation. In particular, for the example of the Coulomb field of a nucleus of atomic number Z, where $\mathbf{A}$ will be 0, we find that Eq. (34) reduces to

$$\left(\frac{\partial}{\partial t} - i A_0\right)\left(\frac{\partial}{\partial t} - i A_0\right)\psi - \nabla^2 \psi + \psi = 0. \tag{35}$$

Then by comparison with Eq. (3), it is clear that a nucleus of atomic number Z should generate a potential well

$$A_0 = -\frac{Z\alpha_{fs}}{r}. \tag{36}$$

The Lagrangian density for the electromagnetic field can be expressed in terms of A_0 and $\mathbf{A}$, as mentioned, but it involves a tensor quantity that is not convenient to express in terms of the usual vector notation. Details are given in [24]. When the equations for ψ, A_0, and $\mathbf{A}$ are derived as Euler equations of a variational principle, as in [24], it turns out that A_0 and $\mathbf{A}$ are to satisfy

$$\frac{\partial^2 A_0}{\partial t^2} - \nabla^2 A_0 = \frac{i\psi}{2}\frac{\partial \bar{\psi}}{\partial t} - \frac{i\bar{\psi}}{2}\frac{\partial \psi}{\partial t} - A_0 \psi\bar{\psi} \tag{37}$$

and

$$\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2 \mathbf{A} = \frac{i\psi}{2}\nabla\bar{\psi} - \frac{i\bar{\psi}}{2}\nabla\psi - \mathbf{A}\psi\bar{\psi}, \tag{38}$$

respectively. Substantial work has been done regarding existence of solitons in the NKGM system [43–45], in some cases even with nonzero angular momentum [16]. Stability of solitons in the NKGM system is a difficult subject, but some progress has been made [17, 21, 44], especially in the case of small coupling of the NKG equation to the electromagnetic field. The effect of an external field on a soliton has also been examined in the NKGM context by several authors [20, 46, 47].

9. A model with interacting solitons

Recently, a model (call it Model One) has been proposed [24] in which a second NKG equation

$$\frac{\partial^2 \phi}{\partial t^2} - \nabla^2 \phi + M^2 W'\left(\frac{\phi\bar{\phi}}{M^2}\right)\phi = 0, \tag{39}$$

is also coupled to the electromagnetic field, but scaled with the constant M so that solitons in the ϕ field can be thought of as protons, whereas solitons in the ψ field are to be regarded as electrons. It must be stated at the outset that Model One is a very primitive model that clearly does not represent the universe in which we live; nevertheless, as an example in applied mathematics it has interesting features that seem worth examining. Although Model One is purely classical in nature, we will show in Section 12 that it exhibits behavior suggestive of a phenomenon that is generally presumed to occur only in a quantum-mechanical context.

In addition to Eq. (34) for ψ, we now have an equation

$$\left(\frac{\partial}{\partial t} - iA_0\right)\left(\frac{\partial}{\partial t} - iA_0\right)\phi - (\nabla - i\mathbf{A})\cdot(\nabla - i\mathbf{A})\phi + M^2 W'\left(\frac{\phi\bar{\phi}}{M^2}\right)\phi = 0 \tag{40}$$

for ϕ, and Eqs. (37) and (38) are to be replaced by

$$\frac{\partial^2 A_0}{\partial t^2} - \nabla^2 A_0 = \frac{i\psi}{2}\frac{\partial\bar{\psi}}{\partial t} - \frac{i\bar{\psi}}{2}\frac{\partial\psi}{\partial t} - A_0\psi\bar{\psi} + \frac{i\phi}{2}\frac{\partial\bar{\phi}}{\partial t} - \frac{i\bar{\phi}}{2}\frac{\partial\phi}{\partial t} - A_0\phi\bar{\phi} \tag{41}$$

and

$$\frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla^2\mathbf{A} = \frac{i\psi}{2}\nabla\bar{\psi} - \frac{i\bar{\psi}}{2}\nabla\psi - \mathbf{A}\psi\bar{\psi} + \frac{i\phi}{2}\nabla\bar{\phi} - \frac{i\bar{\phi}}{2}\nabla\phi - \mathbf{A}\phi\bar{\phi}, \tag{42}$$

respectively.

10. Electron-like solitons

Let us first investigate solutions of the system (34), (40)–(42) where ψ rotates in the complex plane and $\phi = 0$. Also, we will take the scalar field A_0 to be constant in time and set $\mathbf{A}$ to zero, that is

$$\psi = U(x, y, z)\, e^{i\omega t}, \quad \phi = 0, \quad A_0 = V(x, y, z), \quad \mathbf{A} = 0, \tag{43}$$

where U and V are real-valued. Then, Eqs. (34) and (40)–(42) reduce to

$$\nabla^2 U + (\omega - V)^2 U - W'(U^2)U = 0, \tag{44}$$

$$\nabla^2 V = (-\omega + V)\, U^2. \tag{45}$$

We want to find a soliton solution that satisfies both Eqs. (44) and (45) and can be thought of as an electron. Suppose that we start with a soliton solution for Eq. (18) of the kind that was previously discussed. We want the coupling to the electromagnetic field in Eqs. (44) and (45) to result in only a small change, so that U will still be spherically symmetric and localized, with U exponentially small for large r. The right-hand side of Eq. (45) is then nearly zero for large r, so the solution for V will be nearly proportional to $1/r$ for large r. If we can arrange for V to approach α_{fs}/r for large r we see by comparison with Eq. (36) that the soliton will have one negative unit of elementary charge, so we can think of it as an electron-like soliton.

In the present NKGM context, it turns out Eq. (16) is to be replaced by

$$Q = \int\left(\frac{i\bar{\psi}}{2}\frac{\partial\psi}{\partial t} - \frac{i\psi}{2}\frac{\partial\bar{\psi}}{\partial t} + \psi\bar{\psi}A_0\right) d^3x \tag{46}$$

so for a solution of the form (43), the charge is now

$$Q = \int(-\omega + V)\, U^2\, d^3x. \tag{47}$$

The divergence theorem together with Eq. (45) shows that for large r, where U is small, V is approximately $-Q/(4\pi r)$, and so, by comparison with Eq. (36), $Q = 4\pi\alpha_{fs}$ corresponds to one unit of elementary charge. Positive ω in a solution of the form (43), which gives counterclockwise rotation in the complex plane, thus gives negative Q in agreement with the usual convention for electron charge. There will be a corresponding solution with negative ω and clockwise rotation which can then be thought of as a positron.

As mentioned in Section 4, there is a fundamental difficulty with the use of nontopological solitons as particle models. A whole family of soliton solutions typically exists, and the different solutions will have various values of charge, as was illustrated in the example of Section 7. Thus, it is not yet clear whether a preferred solution occurs in practice that could be regarded as defining one unit of elementary charge. Morris [41] has made a suggestion in this regard, but it seems that the resolution of the question is beyond the scope of present research on models of this nature since it may well involve complex, possibly chaotic, interactions of many solitons. A second problem is that the stability proofs currently available apply only in the limit of small coupling, so it is not clear whether orbital stability is achieved when the coupling is sufficient to correspond to a meaningful physical case. Future research should resolve this question, although perhaps in a somewhat nonrigorous manner based at least in part on numerical calculation. Despite these two difficulties let us assume for present purposes that a suitable orbitally stable solution of the form (43) is available which can then be thought of as an electron.

11. Proton-like solitons

Now let us look for solutions of the form

$$\psi = 0, \quad \phi = \tilde{U}(x, y, z)\, e^{-i\tilde{\omega}t}, \quad A_0 = \tilde{V}(x, y, z), \quad \mathbf{A} = 0, \tag{48}$$

for real-valued functions $\tilde{u}$ and $\tilde{V}$. Then $\tilde{u}$ and $\tilde{V}$ need to satisfy

$$\nabla^2\tilde{U} + (\tilde{\omega} + \tilde{V})^2\tilde{U} - M^2 W'\left(\frac{\tilde{U}^2}{M^2}\right)\tilde{U} = 0, \quad \nabla^2\tilde{V} = (\tilde{\omega} + \tilde{V})\,\tilde{U}^2. \tag{49}$$

If a solution $U = f(r)$ and $V = g(r)$ has been found for Eqs. (44) and (45) with a certain value of ω, then

$$\tilde{U} = M f(Mr), \quad \tilde{V} = -M g(Mr) \tag{50}$$

gives a solution for Eq. (49) with $\tilde{\omega} = M\omega$. We are now looking for a positively charged soliton, so we have purposely introduced the minus sign in the solution form (48) so as to obtain clockwise rotation in the complex plane but still allow $\tilde{\omega}$ to be taken as positive. The new solution has one positive unit of elementary charge, but its energy is larger by a factor of M than that of the electron-like soliton (43). We want to interpret a solution of the form (48) as a proton, so we will refer to ϕ as the proton field, and take M to be the appropriate value, approximately 1836, to give the desired mass ratio between the proton and electron.

Solitons in the electron field ψ and the proton field ϕ interact at a distance through the electromagnetic field. It turns out that, as desired, like charges repel and opposite charges attract, and the magnitude of the interaction agrees with Coulomb's law to a good approximation when W satisfies an appropriate condition. Details are available in [24]. By contrast, some early attempts to investigate interaction of solitons [48, 49] led to obviously undesired results such as attraction of like charges.

12. An unexpected result of the model

Although Model One was set up from purely classical considerations to describe the dynamical interaction of electron and proton-like solitons, a startling and perhaps unexpected result occurs. When small oscillations in the electron field occur in the vicinity of a proton-like soliton, it turns out that, because of the nature of the coupling to the electromagnetic field, only the difference frequencies are radiated, and it will be seen that the radiation is just the familiar arc spectrum of hydrogen.

Because of the scaling (50) needed for Eq. (39), the spatial size of a proton-like soliton is much smaller (by a factor of 1836) than that of an electron-like soliton even though the magnitude of the charge is the same. Consequently, in the vicinity of a proton-like soliton, Eq. (36) with $Z = 1$ will be a good approximation, and the electric potential V will be nearly equal to $-\alpha_{fs}/r$ except for very small values of r.

Now let us suppose that small oscillations are excited in the ψ field in the vicinity of a proton-like soliton, so that various modes of oscillation occur. Since the oscillations are small, we can set $W'(U^2) = 1$ and factor Eq. (44) to get

$$\nabla^2 U + (\omega - V - 1)(\omega - V + 1)U = 0, \tag{51}$$

where V now represents the potential well caused by the ϕ field of the proton-like soliton. Any contribution to the V field caused by the ψ field can be neglected if the oscillations in the ψ field are sufficiently small. Let us suppose that $|V|$ is typically much smaller than 1 and that ω is near to 1. Then, we can approximate $\omega - V + 1$ as 2 in the Klein-Gordon Eq. (51) to obtain the usual Schrödinger equation approximation:

$$-\frac{1}{2}\nabla^2 U + VU = (\omega - 1)U. \tag{52}$$

Here, V describes a potential well, but the quantity $\omega - 1$ is related to an actual oscillation frequency ω, which is unlike the traditional use of Schrödinger's equation, where the right-hand side of the equation is regarded as related to an energy.

The main point of interest here is that solutions of the soliton equations lead in a natural way to Eq. (52), which of course results in frequencies of the hydrogen spectrum. For the $V = -\alpha_{fs}/r$ potential well of a proton-like soliton, it is then straightforward to solve Eq. (52) according to the well-known solution of Schrödinger's equation for a hydrogen atom [50]. Eigenfunctions occur in the form

$$U = R(r)\, P_l^m(\cos\theta)(\cos(m\varphi) + \sin(m\varphi)), \tag{53}$$

where the spherical harmonics are written here in terms of the associated Legendre polynomials P_l^m and the usual spherical coordinates θ and φ. In Eq. (53), the usual factor $e^{im\varphi}$ has been replaced by sinusoids since a real-valued formulation for U is desired. A real-valued solution for the radial function $R(r)$ can be obtained in the standard way. The corresponding frequencies are

$$\omega = 1 - \frac{\alpha_{fs}^2}{2n^2}, \quad n = 1, 2, \ldots, \tag{54}$$

with $l = 0, 1, \ldots, n-1$ and $m = -l, \ldots, l$.

When small oscillations in the electron field ψ are allowed in the vicinity of a proton-like soliton, the overall solution (48) will be slightly changed, so we need to investigate how the electron field source terms

$$S_0 = \frac{i\psi}{2}\frac{\partial\bar{\psi}}{\partial t} - \frac{i\bar{\psi}}{2}\frac{\partial\psi}{\partial t} - A_0\,\psi\bar{\psi} \tag{55}$$

and

$$\mathbf{S} = \frac{i\psi}{2}\,\nabla\bar{\psi} - \frac{i\bar{\psi}}{2}\,\nabla\psi - \mathbf{A}\psi\bar{\psi} \tag{56}$$

in Eqs. (41) and (42) affect the electromagnetic field A_0, $\mathbf{A}$. First, let us try $\psi = U(x, y, z)e^{i\omega t}$, where U is a solution of the form (53) and ω is the corresponding frequency given by Eq. (54). Since the oscillation is to be small, we will assume that an appropriate small constant is absorbed in $R(r)$. Substitution in Eqs. (55) and (56) shows that $S_0 = (\omega - V)U^2$ and $\mathbf{S} = 0$. Since the time dependence has canceled out, we conclude that no radiation into the electromagnetic field occurs

when the ψ field is restricted to a single mode of oscillation. The nonzero value of S_0 causes a change in the static V potential field, but the change in the potential well is small since the ψ oscillation is taken as small (in the evaluation of Eqs. (55) and (56) for present purposes, we have used the original values of A_0 and **A** from the proton solution (48) since any change in A_0 and **A** caused by the ψ oscillation will be of higher order).

Next, let us consider the case where small oscillations occur in two different modes in the ψ field, so that

$$\psi = U(x, y, z)\, e^{i\omega t} + \hat{U}(x, y, z)\, e^{i\hat{\omega} t}. \tag{57}$$

Substitution in Eqs. (55) and (56) shows

$$S_0 = (\omega - V)\, U^2 + (\hat{\omega} - \mathrm{V})\, \hat{U}^2 + (\omega + \hat{\omega} - 2V)\, U\hat{U} \cos(\omega - \hat{\omega})t, \tag{58}$$

$$\mathbf{S} = (\hat{U}\ \nabla U - U\ \nabla \hat{U}) \sin(\omega - \hat{\omega})t. \tag{59}$$

Since S_0 and **S** are now time-varying it is clear that radiation into the electromagnetic field is a possibility, but from the form of the sine and cosine terms it turns out that the radiation can only occur at the difference frequency $\omega - \hat{\omega}$.

It should be noted that charge is not necessarily quantized in Model One, and that the small oscillations in the ψ field indeed represent a small amount of charge. In this example, the V field is not to be noticeably changed from that of the proton-like soliton, however, so the small oscillations in the ψ field must correspond to only a minuscule amount of charge, much less than one elementary charge. Then, Eqs. (41) and (42) are to be solved using the source terms (55), (56) in the form (58), (59). Since charge cannot be radiated into the electromagnetic field, small amounts of charge must be transferred between the trapped modes in the ψ field, and the charge must go to those modes in which the ratio of energy to charge is smaller. It is expected that there will be a selection rule since the two modes in question will not effectively radiate into the electromagnetic field in some cases because of their relative symmetry.

Thus, we find in Model One that radiation is emitted into the electromagnetic field when small oscillations in the ψ field occur in the potential well created by a proton-like soliton. That radiation turns out to be the well-known spectrum of hydrogen, namely, difference frequencies of the form $\omega - \hat{\omega}$, where ω and $\hat{\omega}$ are given by Eq. (54) with different values of n.

13. Interpretation and summary

Although Model One, Eqs. (34) and (40)–(42), was set up to let solitons mimic the dynamical behavior of electrons and protons, it turns out that small oscillations of the electron field ψ in the vicinity of a proton-like soliton have modes of oscillation that correspond to the various terms in the Grotrian diagram for hydrogen. Thus, a quantum-like phenomenon automatically occurs even though the model is of a purely classical nature. It has seemed

in the past that such a result was beyond the realm of a possible classical model since physical oscillation of a charge would presumably radiate away energy at the actual oscillation frequency, whereas only difference frequencies are in fact observed in the hydrogen spectrum.

Here, it should be noted that three circumstances intrinsic to Model One contribute to allow this quantum-like result to occur. First, the standard, i.e., gauge invariant, coupling of the ψ and ϕ fields to the electromagnetic field dictates that the electric potential created by the proton-like soliton is experienced by the ψ field as a potential well that allows trapped oscillations. Secondly, the great difference in scale (by a factor of 1836) needed in Eq. (50) assures that the ψ field automatically experiences almost a pure $V = -\alpha_{fs}/r$ potential well, which leads to frequencies that correspond to the well-known terms in the Grotrian diagram for hydrogen. Finally, standard coupling of the ψ field to the electromagnetic field again automatically implies that only the difference frequencies are radiated.

The emission of the hydrogen spectrum in Model One has some interesting aspects but it occurs in an unexpected context. When oscillations of the ψ field occur in the vicinity of a proton-like soliton, only small oscillations are allowed or the $V = -\alpha_{fs}/r$ potential well created by the proton-like soliton would be substantially changed and the spectrum typical of hydrogen would not be observed. Thus, in the situation in question, an electron-like soliton need not, indeed must not, be in the vicinity at all. When we investigate Model One, then, whether we like it or not, within that model the hydrogen spectrum is emitted from the vicinity of an isolated proton-like soliton. In other words, in Model One, the hydrogen arc spectrum is emitted from the equivalent of a hydrogen *ion* rather than a neutral hydrogen atom, which is in startling contrast to the accepted understanding of the actual physical phenomenon. It remains to be seen whether such a result might have any possible physical significance, but it seems of interest, nevertheless, to see how a purely classical model can behave in a way suggestive of a quantum mechanical phenomenon.

Relativistic effects have not been considered in the above treatment of Model One since it is well-known that the Klein-Gordon equation gives wrong answers for fine-structure corrections, especially those that involve the anomalous Zeeman effect. It is contemplated that the NKG equations might be replaced with nonlinear Dirac equations in a more advanced Model Two. Such a model will be comparatively difficult to examine, however, so it seems that further study of Model One will still be of interest.

Author details

Jon C. Luke

Address all correspondence to: jcluke@iupui.edu

Department of Mathematical Sciences, Indiana University-Purdue University at Indianapolis, Indianapolis, Indiana, USA

References

[1] Schweber SS. An Introduction to Relativistic Quantum Field Theory. New York: Harper & Row; 1961. p. 913

[2] Dirac PAM. The Principles of Quantum Mechanics. 4th ed. Oxford: Clarendon; 1958. p. 312

[3] Luke JC. A perturbation method for nonlinear dispersive wave problems. Proceedings of the Royal Society, A. 1966;**292**:403-412

[4] Whitham GB. Linear and Nonlinear Waves. New York: Wiley; 1974. p. 636

[5] Russell SR. Report on waves. Report of the British Association for the Advancement of Science; 1845

[6] Lee TD. Particle Physics and Introduction to Field Theory. New York: Harwood; 1981. p. 865

[7] de Broglie L. The Current Interpretation of Wave Mechanics: A Critical Study. New York: Elsevier; 1964. p. 94

[8] Abbondandolo A, Benci V. Solitary waves and Bohmian mechanics. Proceedings of the National Academy of Sciences. 2002;**99**:15257-15261

[9] Scott AC, Chu FYF, McLaughlin DW. The soliton: A new concept in applied science. Proceedings of the IEEE. 1973;**61**:1443-1483

[10] Makhankov VG. Dynamics of classical solitons (in non-integrable systems). Physics Reports. 1978;**35**:1-128

[11] Makhankov V. Computer experiments in soliton theory. Computer Physics Communications. 1980;**21**:1-49

[12] Kivshar YS, Malomed BA. Dynamics of solitons in nearly integrable systems. Reviews of Modern Physics. 1989;**61**:763-915

[13] Lee TD, Pang Y. Nontopological solitons. Physics Reports. 1992;**221**:251-350

[14] Badiale M, Benci V, Rolando S. Solitary waves: Physical aspects and mathematical results. Rendiconti del Seminario Matematico Universitá Politecnico Torino. 2004;**62**:107-154

[15] Bellazzini J, Benci V, Bonanno C, Micheletti AM. Solitons for the nonlinear Klein-Gordon equation. Advanced Nonlinear Studies. 2010;**10**:481-499

[16] Benci V, Fortunato D. Spinning Q-balls for the Klein-Gordon-Maxwell equations. Communications in Mathematical Physics. 2010;**295**:639-668

[17] Benci V, Fortunato D. On the existence of stable charged Q-balls. Journal of Mathematical Physics. 2011;**52**:093701

[18] Badiale M, Rolando S. A note on vortices with prescribed charge. Advanced Nonlinear Studies. 2012;**12**:703-716

[19] Bellazzini J, Benci V, Bonanno C, Sinibaldi E. On the existence of hylomorphic vortices in the nonlinear Klein-Gordon equation. Dynamics of Partial Differential Equations. 2013;**10**:1-23

[20] Babin A, Figotin A. Relativistic point dynamics and Einstein formula as a property of localized solutions of a nonlinear Klein-Gordon equation. Communications in Mathematical Physics. 2013;**322**:453-499

[21] Benci V, Fortunato D. Hylomorphic solitons and charged Q-balls: Existence and stability. Chaos, Solitons & Fractals. 2014;**58**:1-15

[22] Brihaye Y, Diemer V, Hartmann B. Charged Q-balls and boson stars and dynamics of charged test particles. Physical Review D. 2014;**89**:084048

[23] Nugaev E, Smolyakov MN. Particle-like Q-balls. Journal of High Energy Physics. 2014;**2014**:1-9. DOI: 10.1007/JHEP07(2014)009

[24] Luke JC. A proposed model in which solitons exhibit electron and proton-like behavior. Advanced Nonlinear Studies. 2016;**16**:301-314

[25] Bethe HA. The electromagnetic shift of energy levels. Physical Review. 1947;**72**:339-341

[26] de Broglie L. New Perspectives in Physics. New York: Basic Books, Inc.; 1962. p. 291

[27] Hobart RH. On the instability of a class of unitary field models. Proceedings of the Physical Society. 1963;**82**:201-203

[28] Derrick GH. Comments on nonlinear wave equations as models for elementary particles. Journal of Mathematical Physics. 1964;**5**:1252-1254

[29] Scott AC. A nonlinear Klein-Gordon equation. American Journal of Physics. 1969;**37**:52-61

[30] Rubinstein J. Sine-Gordon equation. Journal of Mathematical Physics. 1970;**11**:258-266

[31] Glasko VB, Leriust F, Terletskii IaP, Shushurin SF. An investigation of particle-like solutions of a nonlinear scalar field equation. Journal of Experimental and Theoretical Physics (U.S.S.R.). 1958;**35**:452-457, translated: Soviet Physics JETP. 1959;**35**:312-315

[32] Zastavenko LG. Particle-like solutions of a nonlinear wave equation. Prikladnaya Matematika i Mechanika. 1965;**29**:430-439, translated: Journal of Applied Mathematics and Mechanics (PMM). 1965;**29**:497-509

[33] Rosen G. Charged particlelike solutions to nonlinear complex scalar field theories. Journal of Mathematical Physics. 1968;**9**:999-1002

[34] Shatah J. Stable standing waves of nonlinear Klein-Gordon equations. Communications in Mathematical Physics. 1983;**91**:313-327

[35] Noether E. Invariante Variationsprobleme. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse. 1918;**1918**:235-257

[36] Gelfand IM, Fomin SV. Calculus of Variations. Englewood Cliffs, New Jersey: Prentice-Hall; 1963. p. 232

[37] Rosen G. Particlelike solutions to nonlinear complex scalar field theories with positive-definite energy densities. Journal of Mathematical Physics. 1968;**9**:996-998

[38] Deumens E, Warchall H. Explicit construction of all spherically symmetric solitary waves for a nonlinear wave equation in multiple dimensions. Nonlinear Analysis. 1988;**12**:419-447

[39] Luke JC. A note on particle-like solutions for a nonlinear complex-valued Klein-Gordon equation. SIAM Journal of Applied Mathematics. 1991;**51**:1194-1204

[40] Luke JC. On particle-like solutions for a nonlinear complex-valued Klein-Gordon equation. In: Debnath L, editor. Nonlinear Dispersive Wave Systems. River Edge, New Jersey: World Scientific; 1992. p. 611-623

[41] Morris TF. Is the electron a soliton? Hadronic Journal. 1980;**3**:1360-1392

[42] Jackson JD. Classical Electrodynamics. New York: John Wiley & Sons; 1962. p. 641

[43] Benci V, Fortunato D. Solitary waves of the nonlinear Klein-Gordon equation coupled with the Maxwell equations. Reviews in Mathematical Physics. 2002;**14**:409-420

[44] Long E. Existence and stability of solitary waves in non-linear Klein-Gordon-Maxwell equations. Reviews in Mathematical Physics. 2006;**18**:747-779

[45] Benci V, Fortunato D. Existence of hylomorphic solitary waves in Klein-Gordon and in Klein-Gordon-Maxwell equations. Rendiconti Lincei-Matematica e Applicazioni. 2009;**20**:243-279

[46] Bodurov TG. Solitary waves interacting with an external field. International Journal of Theoretical Physics. 1996;**35**:2489-2499

[47] Long E, Stuart D. Effective dynamics for solitons in the nonlinear Klein-Gordon-Maxwell system and the Lorentz force law. Reviews in Mathematical Physics. 2009;**21**:459-510

[48] Rosen G. Equations of motion in classical nonlinear field theories. Journal of Mathematical Physics. 1967;**8**:573-575

[49] Derrick GH, Kay-Kong W. Particle motion and interaction in nonlinear field theories. Journal of Mathematical Physics. 1968;**9**:232-240

[50] Dicke RH, Wittke JP. Introduction to Quantum Mechanics. Reading, Massachusetts: Addison-Wesley; 1960. p. 369

Chapter 8

A Perturbation Theory for Nonintegrable Equations with Small Dispersion

Georgy Omel'yanov

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71030

Abstract

We describe an approach called the "weak asymptotics method" to construct multisoliton asymptotic solutions for essentially nonintegrable equations with small dispersion. This paper contains a detailed review of the method and a perturbation theory to describe the interaction of distorted solitons for equations with small perturbations. All constructions have been realized for the gKdV equation with the nonlinearity u^μ, $\mu \in (1, 5)$.

Keywords: generalized Korteweg-de Vries equation, soliton, interaction, perturbation, weak asymptotics method

2010 Mathematics Subject Classification: 35D30, 35Q53, 46F10

1. Introduction

We consider the problem of propagation and interaction of soliton-type solutions of nonlinear equations. Our basic example is the nonhomogeneous version of the generalized KdV equation

$$\frac{\partial u}{\partial t} + \frac{\partial u^\mu}{\partial x} + \varepsilon^2 \frac{\partial^3 u}{\partial x^3} = f\left(u, \varepsilon \frac{\partial u}{\partial x}\right), \quad x \in R^1, \quad t > 0, \tag{1}$$

where $\mu \in (1, 5)$, $\varepsilon \ll 1$, $f(u, z)$ is a known smooth function such that $f(0, 0) = 0$. Note that the restriction on μ implies both the soliton-type solution and the stability of the equation with respect to initial data (see, for example [1, 2]).

In the special case $f \equiv 0$ and $\mu = 2$ ($\mu = 3$), Eq. (1) is the famous KdV (modified KdV) equation. It is well known that KdV (mKdV) solitons are stable and interact in the elastic manner: after

the collision, they preserve the original amplitudes and velocities shifting the trajectories only (see [3] and other bibliographies devoted to the inverse scattering transform (IST) method). In the case of $\mu = 2$ ($\mu = 3$) but with $f \neq 0$, Eq. (1) is a nonintegrable one. However, using the smallness of ε (or of f for other scaling), it is possible to create a perturbation theory that describes the evolution of distorted solitons (see the approaches by Karpman and E. Maslov [4] and Kaup and Newell [5] on the basis of the IST method, and the "direct" method by V. Maslov and Omel'yanov [6]). Moreover, the approach by V. Maslov and Omel'yanov [6] can be easily extended to essentially nonintegrable equations ($\mu \neq 2, 3$), but for a single soliton only. In fact, it is impossible to use any direct method in the classical sense for the general problem of the wave interaction. To explain this proposition, let us consider the homogeneous gKdV equation

$$\frac{\partial u}{\partial t} + \frac{\partial u^{\mu}}{\partial x} + \varepsilon^2 \frac{\partial^3 u}{\partial x^3} = 0, \quad x \in R^1, \quad t > 0. \tag{2}$$

It is easy to find the explicit soliton solution of (2),

$$u(x, t, \varepsilon) = A\omega\big(\beta(x - Vt)/\varepsilon\big), \quad \omega(\eta) = \cosh^{-\gamma}\big(\eta/\gamma\big), \tag{3}$$

$$\gamma = 2/(\mu - 1), \quad V = \beta^2, \quad A^{\mu-1} = V(\mu + 1)/2. \tag{4}$$

Next let us consider two-soliton initial data

$$u\,|_{t=0} = \sum_{i=1}^{2} A_i\, \omega\Big(\beta_i\big(x - x_{(i,0)}\big)/\varepsilon\Big), \tag{5}$$

where $x_{(1,0)} > x_{(2,0)}$ and $A_2 > A_1$. Obviously, since $(x_{(2,0)} - x_{(1,0)})/\varepsilon \to \infty$ as $\varepsilon \to 0$, the sum of the waves (3)

$$u = \sum_{i=1}^{2} A_i\, \omega\Big(\beta_i\big(x - V_i t - x_{(i,0)}\big)/\varepsilon\Big) \tag{6}$$

approximates the problem (2), (5) solution with the precision $O(\varepsilon^{\infty})$ but for $t \ll 1$ only. Conversely, the sum (6) does not satisfy the gKdV equation for $t \sim O(1)$ in view of the trajectories $x = V_i t + x_{(i,0)}$ intersection at a point (x^*, t^*).

Let us consider shortly how it is possible to analyze the problem (2), (5). There are some different cases:

1. Let $A_1 \ll A_2$. Then, one can construct an asymptotic solution

$$u = W\big(\big(x - \varphi_2(t)\big)/\varepsilon, t, x, \varepsilon, \nu\big), \tag{7}$$

where $\nu = A_1/A_2 \ll 1$ and $W((x - \varphi_2(t))/\varepsilon, t, x, \varepsilon, \nu) = A_2\omega(\beta_2(x - V_2 t - x_{(2,0)})/\varepsilon) + O(\nu + \varepsilon)$. Thus, to find the leading term of the asymptotics, we obtain an equation with nonlinear ordinary differential operator; whereas to construct the corrections, it is enough to analyze the linearization of this operator. This construction (with a little bit of other viewpoints) has been realized by Ostrovsky et al. [7].

2. Let $A_2 - A_1 \ll 1$. We write again the ansatz in the form (7), where $\nu = A_2 - A_1 \ll 1$ now, and we assume $\nu/\varepsilon \ll 1$. In fact, this case coincides with the problem considered in [7].

3. The amplitudes $A_2 > A_1$ are arbitrary numbers. Then, we should write a two-phase ansatz

$$u(x, t, \varepsilon) = W\big((x - \varphi_1(t))/\varepsilon, (x - \varphi_2(t))/\varepsilon, t, x, \varepsilon\big) \tag{8}$$

without any additional parameter. Substituting (8) into equation (2), we obtain for the leading term $W_0(\tau_1, \tau_2, t)$:

$$\hat{A}W_0 + \hat{B}W_0^{\mu} + \hat{B}^3 W_0 = 0, \quad \hat{A} = -\sum_{i=1}^{2} \dot{\varphi}_i \frac{\partial}{\partial \tau_i}, \quad \hat{B} = \sum_{i=1}^{2} \frac{\partial}{\partial \tau_i}. \tag{9}$$

Since $\dot{\varphi}_1 \neq \dot{\varphi}_2$, we can pass to new variables, $\eta = (\tau_1 - \tau_2)/(\dot{\varphi}_2 - \dot{\varphi}_1)$, $\zeta = (\dot{\varphi}_1 \tau_2 - \dot{\varphi}_2 \tau_1)/(\dot{\varphi}_1 - \dot{\varphi}_2)$, and transform equation (9) to the gKdV form (2) again

$$\frac{\partial W_0}{\partial \eta} + \frac{\partial W_0^{\mu}}{\partial \zeta} + \frac{\partial^3 W_0}{\partial \zeta^3} = 0. \tag{10}$$

Therefore, to construct two-phase asymptotics, we should solve (10) explicitly what is impossible for any essentially nonintegrable case.

This difficulty can be overcome by using the weak asymptotics method. The main point here is that solitons tend to distributions as $\varepsilon \to 0$. Thus, it is possible to pass to the weak description of the problem, ignore the actual shape of the multiwave solutions, and find only the main solution characteristics, that is, the time dynamics of wave amplitudes and velocities. The weak asymptotics method has been proposed at first for shock wave type solutions [8] and for soliton-type solutions [9] many years ago. Further generalizations, modifications, and adaptations to other problems can be found in publications by M. Colombeau, Danilov, Mitrovic, Omel'yanov, Shelkovich, and others, see, for example, [10–20] and references therein.

The contents of the paper are the following: in Section 2, we present a detailed survey of the weak asymptotics method application to the problem of multisoliton asymptotics and Section 3 contains new results, namely a perturbation theory to describe the evolution and collision of distorted solitons for equation (1).

2. Weak asymptotics method

2.1. Main definitions

Let us associate equation (2) with first two conservation laws written in the differential form:

$$\frac{\partial Q_j}{\partial t} + \frac{\partial P_j}{\partial x} = \varepsilon^2 \frac{\partial^3 R_j}{\partial x^3}, \quad j = 1, 2, \tag{11}$$

$$Q_1 = u, \quad P_1 = u^{\mu}, \quad Q_2 = u^2, \quad P_2 = 2\mu\, u^{\mu+1}/(\mu+1) - 3\,(\varepsilon\, u_x)^2, \tag{12}$$

and $R_1 = u$, $R_2 = u^2$. Next, we define smallness in the weak sense:

Definition 1. *A function $v(t, x, \varepsilon)$ is said to be of the value $O_{\mathcal{D}'}(\varepsilon^{\varkappa})$ if the relation $\int_{-\infty}^{\infty} v(t, x, \varepsilon)\Psi(x)dx = O(\varepsilon^{\varkappa})$ holds uniformly in t for any test function $\Psi \in \mathcal{D}(\mathbb{R}^1_x)$. The right-hand side here is a $\mathcal{C}^{\infty}$-function for $\varepsilon = const > 0$ and a piecewise continuous function uniformly in $\varepsilon \geq 0$.*

Following [9, 17, 18], we define two-soliton weak asymptotics:

Definition 2. *A sequence $u(t, x, \varepsilon)$, belonging to $\mathcal{C}^{\infty}\left(0, T; \mathcal{C}^{\infty}(\mathbb{R}^1_x)\right)$ for $\varepsilon = \text{const} > 0$ and belonging to $\mathcal{C}\left(0, T; \mathcal{D}'(\mathbb{R}^1_x)\right)$ uniformly in ε, is called a weak asymptotic mod $O_{\mathcal{D}'}(\varepsilon^2)$ solution of (2) if the relations (11) hold uniformly in t with the accuracy $O_{\mathcal{D}'}(\varepsilon^2)$.*

Let us consider the interaction of two solitary waves for the model (2) with the initial data (5).

Following [9, 17, 18] again, we write the asymptotic ansatz in the form:

$$u = \sum_{i=1}^{2} G_i(\tau)\omega\left(\beta_i\left(x - \varphi_i(t, \tau, \varepsilon)\right)/\varepsilon\right), \quad G_i(\tau) = A_i + S_i(\tau). \tag{13}$$

Here $\varphi_i = \varphi_{i0}(t) + \varepsilon\varphi_{i1}(\tau)$, where $\varphi_{i0} = V_i t + x_{(i,0)}$ are the trajectories of noninteracting solitary waves, $\tau = \psi_0(t)/\varepsilon$ denotes the "fast time", $\psi_0(t) = \beta_1(\varphi_{20}(t) - \varphi_{10}(t))$, and the phase and amplitude corrections φ_{i1}, S_i are smooth functions such that with exponential rates

$$\varphi_{i1}(\tau) \to 0 \quad as \quad \tau \to -\infty, \quad \varphi_{i1}(\tau) \to \varphi_{i1}^{\infty} = const_i \quad as \quad \tau \to +\infty, \tag{14}$$

$$S_i(\tau) \to 0 \quad as \quad \tau \to \pm\infty. \tag{15}$$

2.2. Two-wave asymptotic construction

To construct the asymptotics, we should calculate the weak expansions of the terms from the left-hand sides of the relations (11). It is easy to check that

$$u = \varepsilon \sum_{i=1}^{2} a_1 \frac{G_i}{\beta_i}\delta\left(x - \varphi_i\right) + O_{\mathcal{D}'}(\varepsilon^3), \tag{16}$$

where $\delta(x)$ is the Dirac delta-function. Here and in what follows, we use the notation

$$a_k \overset{\text{def}}{=} \int_{-\infty}^{\infty} \left(\omega(\eta)\right)^k d\eta, \quad k > 0, \quad a_2' \overset{\text{def}}{=} \int_{-\infty}^{\infty} \left(\omega'(\eta)\right)^2 d\eta. \tag{17}$$

At the same time for any $F(u, \varepsilon\partial u/\partial x) \in C^1$, we have

$$\begin{aligned}
&\int_{-\infty}^{\infty} F\left(\sum_{i=1}^{2} G_i\,\omega\left(\beta_i \frac{x - \varphi_i}{\varepsilon}\right), \sum_{i=1}^{2} \beta_i G_i\,\omega'\left(\beta_i \frac{x - \varphi_i}{\varepsilon}\right)\right)\psi(x)dx \\
&= \varepsilon \sum_{i=1}^{2} \frac{1}{\beta_i} \int_{-\infty}^{\infty} F\left(A_i\,\omega(\eta), \beta_i A_i\,\omega'(\eta)\right)\psi\left(\varphi_i + \varepsilon \frac{\eta}{\beta_i}\right)d\eta \\
&\quad + \frac{\varepsilon}{\beta_2} \int_{-\infty}^{\infty} \left\{ F\left(\sum_{i=1}^{2} G_i\,\omega(\eta_{i2}), \sum_{i=1}^{2} \beta_i G_i\,\omega'(\eta_{i2})\right)\right. \\
&\quad \left. - \sum_{i=1}^{2} F\left(A_i\,\omega(\eta_{i2}), \beta_i A_i\,\omega'(\eta_{i2})\right)\right\}\psi\left(\varphi_2 + \varepsilon \frac{\eta}{\beta_2}\right)d\eta,
\end{aligned} \tag{18}$$

where

$$\eta_{12} = \theta\eta - \sigma, \quad \eta_{22} = \eta, \quad \sigma = \beta_1(\varphi_1 - \varphi_2)/\varepsilon, \quad \theta = \beta_1/\beta_2. \tag{19}$$

We take into account that the second integrand in the right-hand side of (18) vanishes exponentially fast as $|\varphi_1 - \varphi_2|$ grows; thus, its main contribution is at the point x^*. We write

$$\varphi_{i0} = x^* + V_i(t - t^*) = x^* + \varepsilon V_i \tau/\dot{\psi}_0 \quad \text{and} \quad \varphi_i = x^* + \varepsilon\chi_i, \tag{20}$$

where $\dot{\psi}_0 = \beta_1(V_2 - V_1)$, $\chi_i = V_i\tau/\dot{\psi}_0 + \varphi_{i1}$. It remains to apply the formula

$$f(\tau)\delta(x - \varphi_i) = f(\tau)\delta(x - x^*) - \varepsilon\chi_i f(\tau)\,\delta'(x - x^*) + O_{\mathcal{D}'}(\varepsilon^2), \tag{21}$$

which holds for each φ_i of the form (20) with slowly increasing χ_i and for $f(\tau)$ from the Schwartz space. Moreover, the second term in the right-hand side of (21) is $O_{\mathcal{D}'}(\varepsilon)$. Thus, under the assumptions (14) and (15), we obtain the weak asymptotic expansion of $F(u, \varepsilon\partial u/\partial x)$ in the final form:

$$\begin{aligned} F(u, \varepsilon u_x) = {} & \varepsilon \sum_{i=1}^{2} \left\{ \frac{a_{F,i}^{(0)}}{\beta_i}\delta(x - \varphi_i) - \varepsilon \frac{a_{F,i}^{(1)}}{\beta_i}\delta'(x - x^*) \right\} + \frac{\varepsilon}{\beta_2}\left\{ \mathfrak{R}_F^{(0)}\delta(x - x^*) - \varepsilon\,\bar{\mathfrak{R}}_F\,\delta'(x - x^*) \right\} \\ & + O_{\mathcal{D}'}(\varepsilon^3), \quad \bar{\mathfrak{R}}_F = \chi_2\,\mathfrak{R}_F^{(0)} + \mathfrak{R}_F^{(1)}/\beta_2, \end{aligned} \tag{22}$$

where

$$a_{F,i}^{(n)} = \int_{-\infty}^{\infty} \eta^n F\big(A_i\,\omega(\eta), \beta_i A_i\,\omega'(\eta)\big)d\eta, \tag{23}$$

$$\mathfrak{R}_F^{(n)} = \int_{-\infty}^{\infty} \eta^n \left\{ F\left(\sum_{i=1}^{2} G_i\,\omega(\eta_{i2}), \sum_{i=1}^{2} \beta_i G_i\,\omega'(\eta_{i2}) \right) - \sum_{i=1}^{2} F\big(A_i\,\omega(\eta_{i2}), \beta_i A_i\,\omega'(\eta_{i2})\big) \right\} d\eta. \tag{24}$$

Here, we take into account that to define $\partial u^2/\partial t \bmod O_{\mathcal{D}'}(\varepsilon^2)$, it is necessary to calculate u^2 with the precision $O_{\mathcal{D}'}(\varepsilon^3)$. Thus, using (22) with $F(u) = u^2$ and transforming (16) with the help of (21), we obtain modulo $O_{\mathcal{D}'}(\varepsilon^3)$:

$$u = \varepsilon \sum_{i=1}^{2} a_1 K_{i0}^{(1)}\delta(x - \varphi_i) + \varepsilon \sum_{i=1}^{2} a_1 K_{i1}^{(1)}\{\delta(x - x^*) - \varepsilon\chi_i\,\delta'(x - x^*)\}, \tag{25}$$

$$u^2 = \varepsilon \sum_{i=1}^{2} a_2 K_{i0}^{(2)}\delta(x - \varphi_i) + \frac{\varepsilon}{\beta_2}\left\{ \mathfrak{R}_{u^2}^{(0)}\delta(x - x^*) - \varepsilon\bar{\mathfrak{R}}_{u^2}\,\delta'(x - x^*) \right\}, \tag{26}$$

where

$$K_i^{(n)} = G_i^n/\beta_i, \quad K_{i0}^{(n)} = A_i^n/\beta_i, \quad K_{i1}^{(n)} = K_i^{(n)} - K_{i0}^{(n)}. \tag{27}$$

Calculating weak expansions for other terms from Definition 2 and substituting them into (11), we obtain linear combinations of $\varepsilon\delta'(x - \varphi_i)$, $i = 1, 2$, $\delta(x - x^*)$, and $\varepsilon\delta'(x - x^*)$. Therefore, we pass to the system:

$$a_1 V_i K_{i0}^{(1)} - a_{P_1,i}^{(0)}/\beta_i = 0, \quad a_2 V_i K_{i0}^{(2)} - a_{P_2,i}^{(0)}/\beta_i = 0, \quad i = 1, 2, \tag{28}$$

$$\sum_{i=1}^{2} K_{i1}^{(1)} = 0, \quad \mathfrak{R}_{u^2}^{(0)} = 0, \quad i = 1, 2, \tag{29}$$

$$\dot{\psi}_0 \frac{d}{d\tau} \sum_{i=1}^{2} \{K_{i0}^{(1)} \varphi_{i1} + \chi_i K_{i1}^{(1)}\} = f, \quad \dot{\psi}_0 \frac{d}{d\tau} \left\{ \sum_{i=1}^{2} a_2 K_{i0}^{(2)} \varphi_{i1} + \bar{\mathfrak{R}}_{u^2} \right\} = F, \tag{30}$$

where

$$f = \frac{1}{a_1 \beta_2} \mathfrak{R}_{P_1}^{(0)}, \quad F = \frac{1}{\beta_2} \mathfrak{R}_{P_2}^{(0)} - a_1 \dot{\psi}_0 \sum_{i=1}^{2} \varphi_{i1} \frac{dK_{i1}^{(2)}}{d\tau}. \tag{31}$$

The first four algebraic equations (28) imply again the relation (4) among A_i, β_i, and V_i. Furthermore, there exists a number $\theta^* \in (0, 1)$ such that equations (29), (30) have the required solution S_i, φ_{i1} with the properties (14) and (15) under the sufficient condition $\theta \le \theta^*$ (see [9, 17]). It is obvious that the existence of the weak asymptotics (13) with the properties (14) and (15) implies that the solitary waves interact like the KdV solitons at least in the leading term.

Theorem 1. *Let $\theta \le \theta^*$. Then (13) describes $\mathrm{mod}\, O_{\mathcal{D}'}(\varepsilon^2)$ the elastic scenario of the solitary waves interaction for the μ-gKdV equation (2).*

Numerical simulations ([14, 15, 17]) confirm the traced analysis, see **Figure 1**. Note that a small oscillating tail appears after the soliton collision, see [15] for detail. Obviously, this effect is similar to the "radiation" appearance for the perturbed KdV [21].

2.3. Multisoliton interaction

N-wave solutions of the form similar to waves (13) contain 2*N* free functions S_i, φ_{i1}. Thus, to describe an *N*-soliton collision, we should consider *N* conservation laws. However, nonintegrability implies the existence of a finite number of conservation laws only. For this reason, we need to involve into the consideration balance laws. For the gKdV-4 equation, the first conservation and balance laws have the form

$$\frac{\partial Q_j}{\partial t} + \frac{\partial P_j}{\partial x} + \varepsilon^{-1} K_j = O_{\mathcal{D}'}(\varepsilon^2), \tag{32}$$

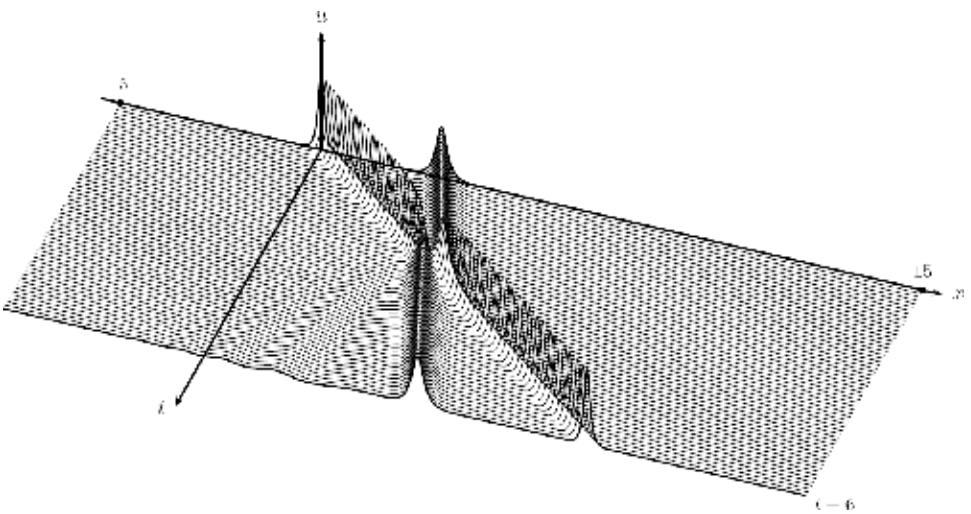

Figure 1. Evolution of two solitary waves for $\mu = 4$ and $\varepsilon = 0.1$.

http://dx.doi.org/10.5772/intechopen.71030

where Q_j, P_j, $j = 1, 2$, coincide with (12) for $\mu = 4$, $K_i = 0$, $i = 1, 2, 3$,

$$Q_3 = (\varepsilon u_x)^2 - \frac{2}{5}u^5, \quad P_3 = 16\,u^3(\varepsilon u_x)^2 - u^8 - 3\left(\varepsilon^2 u_{xx}\right)^2, \tag{33}$$

$$Q_4 = \frac{1}{2}\left(\varepsilon^2 u_{xx}\right)^2 + \frac{5}{21}u^8 - \frac{10}{3}u^3(\varepsilon u_x)^2, \quad K_4 = -(\varepsilon u_x)^5, \tag{34}$$

$$P_4 = 12u^3\left(\varepsilon^2 u_{xx}\right)^2 - 19u\,(\varepsilon\, u_x)^4 - \frac{3}{2}\left(\varepsilon^3 u_{xxx}\right)^2 + \frac{160}{231}u^{11} - \frac{100}{3}u^6(\varepsilon u_x)^2. \tag{35}$$

Note that the nondivergent "production" $\varepsilon^{-1}K_4$ has the same value $O(\varepsilon^{-1})$ (in the C-sense and for rapidly varying functions) as the first ones in (32).

The formal scheme of the asymptotic construction is similar to the one described above: we write the ansatz of the form (13) but with N summands, found weak representations for all terms in (32), and pass to a system similar to (28)–(30). The main obstacle here is the proof that this system admits a solution with the properties of (14), (15). This idea has been realized in [18, 19] for the problem of three soliton collisions for the gKdV-4 equation.

Theorem 2. *Let us denote A_i the amplitudes of the original solitons and $x_{(i,0)}$ their initial positions such that $A_{i+1} > A_i$, $x_{(i,0)} > x_{(i+1,0)}$,* and *$i = 1, 2$. Let all trajectories $x = \varphi_{i0}(t)$ have an intersection point (x^*, t^*). Then, under the assumption*

$$\beta_2/\beta_3 = \nu^3, \quad \beta_1/\beta_3 = \nu^{3(3+\alpha)/2}, \ \alpha \in [0, 1) \tag{36}$$

with sufficiently small $\nu < 1$, the three-phase asymptotic solution exists and describes $\mathrm{mod}\, O_{\mathscr{D}}(\varepsilon^2)$ the elastic scenario of the solitary waves interaction.

Figure 2 depicts the evolution of a three-wave solution [14].

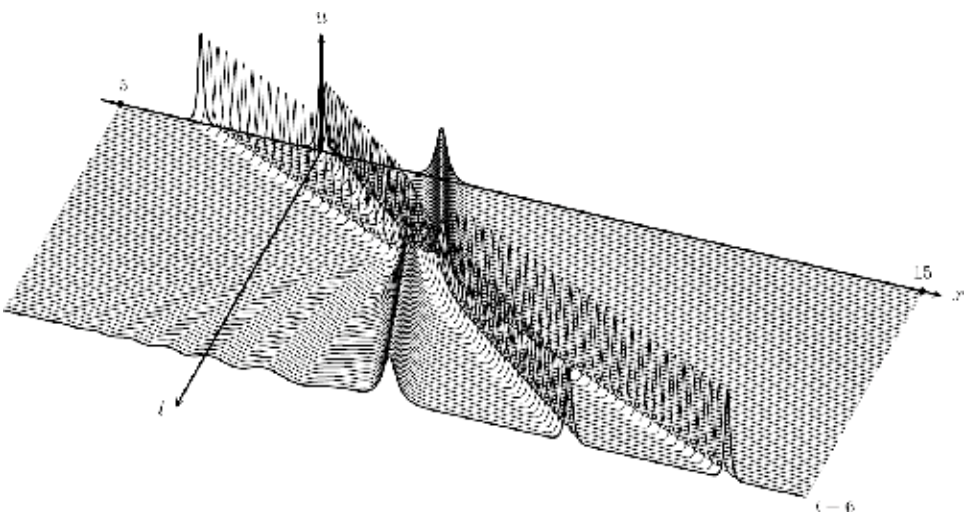

Figure 2. Evolution of the soliton triplet with $\mu = 4$, $\varepsilon = 0.1$.

2.4. Asymptotic equivalence

Let us come back to the case of two-phase asymptotics and transform the ansatz (13) to the following form:

$$\tilde{u} = \sum_{i=1}^{2}\left\{G_i(\tau)\omega\left(\beta_i\frac{x-\varphi_i(t,\tau,\varepsilon)}{\varepsilon}\right)+\mathfrak{S}_i(\tau)W\left(\beta_i\frac{x-\varphi_i(t,\tau,\varepsilon)}{\varepsilon}\right)\right\}, \tag{37}$$

where $\mathfrak{S}_i(\tau)$, $i=1,2$ are arbitrary functions from the Schwartz space,

$$W(\eta) = d^{2l+1}\omega(\eta)/d\,\eta^{2l+1}, \tag{38}$$

and $l \geq 1$ is an arbitrary integer. Calculating the weak representations for $\tilde{u}$ and $\tilde{u}^2$, we obtain

$$\begin{aligned}\tilde{u} &= u + O_{\mathcal{D}'}(\varepsilon^{2l+2}),\\ \tilde{u}^2 &= \varepsilon\sum_{i=1}^{2}a_2K_{i0}^{(2)}\delta(x-\varphi_i)+\frac{\varepsilon}{\beta_2}\mathfrak{R}_{\tilde{u}^2}^{(0)}\delta(x-x^*)+O_{\mathcal{D}'}(\varepsilon^2),\end{aligned} \tag{39}$$

where

$$\mathfrak{R}_{\tilde{u}^2}^{(0)} = \int_{-\infty}^{\infty}\left\{\left(\sum_{i=1}^{2}\left(G_i\,\omega(\eta_{i2})+\mathfrak{S}_i\,W(\eta_{i2})\right)\right)^2-\sum_{i=1}^{2}\left(A_i\,\omega(\eta_{i2})\right)^2\right\}d\eta, \tag{40}$$

and u in the right-hand side in (39) is the representation (25). Thus, the difference between u of the forms (13) and (37) is arbitrarily small in the sense $\mathcal{D}'(\mathbb{R}_x)$. At the same time, instead of (29), (30), we obtain

$$\sum_{i=1}^{2}K_{i1}^{(1)} = 0, \quad \mathfrak{R}_{\tilde{u}^2}^{(0)} = 0, \quad i = 1,2, \tag{41}$$

$$\dot{\psi}_0\frac{d}{d\tau}\sum_{i=1}^{2}\left\{K_{i0}^{(1)}\varphi_{i1}+\chi_iK_{i1}^{(1)}\right\} = \tilde{f},\ \dot{\psi}_0\frac{d}{d\tau}\left\{\sum_{i=1}^{2}a_2K_{i0}^{(2)}\varphi_{i1}+\overline{\mathfrak{R}}_{\tilde{u}^2}\right\} = \tilde{F}, \tag{42}$$

where $\tilde{f}$, $\tilde{F}$ differ from f, F in the same manner as $\mathfrak{R}_{\tilde{u}^2}^{(0)}$ differs from $\mathfrak{R}_{u^2}^{(0)}$. The system (41) and (42) have again a solution with the properties (14) and (15) [9, 12]; however, it differs from the solution of Eqs. (29) and (30) with the value $O(1)$ in the C-sense. Moreover, the asymptotic solutions (13) and (37) differ with the precision $O_{\mathcal{D}'}(\varepsilon)$ in the sense of Definition 1. This implies the principal impossibility to describe explicitly neither the real shape of the waves at the time instant of the collision nor the real ε-size displacements of the trajectories after the interaction. However, the nonuniqueness of the value $O(\varepsilon)$ is concentrated within $O(\varepsilon^{1-\nu})$-neighborhood of the time instant t^* of the interaction, $\nu > 0$. Thus, it is small in the $\mathcal{D}'(R^2_{x,t})$ sense. We set

Definition 3. *Functions $u_1(x,t,\varepsilon)$ and $u_2(x,t,\varepsilon)$ are said to be asymptotically equivalent if for any test function $\psi \in \mathcal{D}(R^2)$*

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\{u_1(x,t,\varepsilon)-u_2(x,t,\varepsilon)\}\psi(x,t)dx\,dt = O(\varepsilon^2). \tag{43}$$

In this sense, the solutions (13) and (37) are asymptotically equivalent.

We now focus attention on another question: how to choose, from the set of all possible conservation and balance laws, those that allow to construct a multiphase asymptotic solution? It seems

that there is not any rule and it is possible to use arbitrary combination of the laws. Thus, there appears the next question: what is the difference between such solutions? This problem has been discussed in [20] for two-phase asymptotic solutions of the gKdV-4 equation. Let us define two-phase asymptotics in the following manner:

Definition 4. *Let* $1 \le k_0 < k_1 \le 4$ *and let a sequence* $u_{k0,k1} = u_{k0,k1}(t, x, \varepsilon)$ *belong to the same functional space as* $u(t, x, \varepsilon)$ *in Definition 2. Then,* $u_{k0,k1}$ *is called a weak asymptotic mod* $O_{\mathcal{D}'}(\varepsilon^2)$ *solution of (2) if the relations (32) hold for* $j = k_0$ *and* $j = k_1$ *uniformly in t.*

A detailed analysis implies the assertion [20].

Theorem 3. *Let* θ *be sufficiently small. Then, the weak asymptotic solutions* $u_{1,k1}$ *and* $u_{1,k_1'}$ *of the problem (2), (5) exist and they are asymptotically equivalent for all* $k_1, k_1' \in \{2, 3, 4\}$.

3. Collision of distorted solitons

We consider now the nonhomogeneous version of the gKdV equation (1). It is easy to verify that, in the case of rapidly varying solutions, the right-hand side f can be treated as a "small perturbation."

An approach to construct one-phase self-similar asymptotic solutions for (1) had been created in [6] (see also [17]). Let us generalize this approach to the multiphase case. From the beginning, we state that equation (1) is associated with balance laws, the first two of which are

$$\frac{\partial Q_j}{\partial t} + \frac{\partial P_j}{\partial x} + K_j = O_{\mathcal{D}'}(\varepsilon^2), \quad j = 1, 2, \tag{44}$$

where Q_j and P_j coincide with ones described in (12),

$$K_1 = -f(u, \varepsilon u_x), \quad K_2 = -uf(u, \varepsilon u_x). \tag{45}$$

Note that, in contrast to K_j in (32), productions here are regularly degenerating functions with the value $O(1)$ in the C-sense.

Let us first construct a two-phase version of self-similar asymptotics, which assumes a special initial data for (1) and discuss afterward how to treat it for more realistic initial data. By analogy with Definition 2, we write:

Definition 5. *Let a sequence* $u = u(t, x, \varepsilon)$ *belong to the same functional space as in Definition 2. Then u is called a weak asymptotic mod* $O_{\mathcal{D}'}(\varepsilon^{\overline{q}})$ *solution of (1) if the relation (44) hold uniformly in* $t \in (0, T)$, $\overline{q} = \min\{\mu, 2\}$.

Generalizing one-phase asymptotics, we write the ansatz as

$$u = \sum_{i=1}^{2} \left\{ G_i(\tau, t)\omega(\eta_i) + \varepsilon\left(z_i(x, t)\mathfrak{H}(\eta_i) + \mathfrak{G}_i(\tau)\omega(\eta_i)\right)\right\}, \tag{46}$$

$$G_i(\tau, t) = A_i(t) + S_i(\tau), \ \eta_i = \beta_i(t)\left(x - \varphi_i\right)/\varepsilon, \ \varphi_i(t, \tau, \varepsilon) = \varphi_{i0} + \varepsilon\varphi_{i1}. \tag{47}$$

Here $A_i(t)$, $\varphi_{i0} = \varphi_{i0}(t)$, $\beta_i^2(t) = \gamma A^{\mu-1}(t)$, $\omega(\eta)$, $S_i(\tau)$, $\varphi_{i1}(\tau)$ are the same as in (13); $\tau = \psi_0(t)/\varepsilon$ with $\psi_0(t) = \varphi_{20}(t) - \varphi_{10}(t)$ denotes the "fast time" again; $z_i(x,t) \in \mathcal{C}^\infty$; and $\mathfrak{G}_i$, $\mathfrak{H}$ are smooth functions such that

$$\mathfrak{G}_i(\tau) \to 0 \quad as \quad \tau \to -\infty\,, \quad \mathfrak{G}_i(\tau) \to \mathfrak{G}_i^\infty = const_i \quad as \quad \tau \to +\infty\,, \tag{48}$$

$$\mathfrak{H}(\eta) \to 1 \quad as \quad \eta \to -\infty\,, \quad \mathfrak{H}(\eta) \to 0 \quad as \quad \eta \to +\infty \tag{49}$$

with exponential rates. We assume also the intersection of the trajectories $x = \varphi_{i0}(t)$, $i = 1, 2$ at a point $x^* = \varphi_{i0}(t^*)$ namely,

$$\exists t^* > 0 \text{ such that } \varphi_{10}(t^*) = \varphi_{20}(t^*),\ \dot\psi_0 \stackrel{\text{def}}{=} \frac{d}{dt}\big(\varphi_{20}(t) - \varphi_{10}(t)\big)\Big|_{t=t^*} \neq 0. \tag{50}$$

It is easy to verify the weak representations with the precision $O_{\mathcal{D}'}(\varepsilon^2)$:

$$u = \varepsilon \sum_{i=1}^{2} \Big\{ a_1 K_{i0}^{(1)}\, \delta\big(x - \varphi_i\big) + a_1 K_{i1}^{(1)*}\, \delta(x - x^*) + z_i(x,t) H\big(\varphi_i - x\big) \Big\}, \tag{51}$$

$$\begin{aligned}
\frac{\partial u}{\partial t} &= a_1 \dot\psi_0 \frac{\partial}{\partial \tau} \sum_{i=1}^{2} K_{i1}^{(1)*}\, \delta(x - x^*) + \varepsilon \sum_{i=1}^{2} \frac{\partial z_i}{\partial t} H(\varphi_i - x) \\
&+ \varepsilon\, a_1 \sum_{i=1}^{2} \left\{ \frac{dK_{i1}^{(1)}}{dt} + z_i(\varphi_i, t) \frac{d\varphi_{i0}}{dt} \right\} \delta(x - \varphi_i) - \varepsilon\, a_1 \sum_{i=1}^{2} K_{i0}^{(1)} \frac{d\varphi_{i0}}{dt}\, \delta'(x - \varphi_i) \\
&+ \varepsilon \sum_{i=1}^{2} \left\{ \dot\psi_0 \frac{\partial}{\partial t} \left(a_1 \frac{\mathfrak{S}_i}{\beta_i} + z_i(x^*, t)\, \varphi_{i1} \right) - \frac{a_1}{\beta_i^2} S_i \frac{d\beta_i}{dt} \right\} \Big|_{t=t^*} \delta(x - x^*) \\
&- \varepsilon\, a_1 \sum_{i=1}^{2} \left\{ K_{i1}^{(1)} \frac{d\varphi_{i0}}{dt} + \dot\psi_0 \left(K_i^{(1)} \frac{\partial \varphi_{i1}}{\partial \tau} + \frac{\chi_i}{\beta_i} \frac{\partial K_i^{(1)}}{\partial \tau} \right) \right\} \Big|_{t=t^*} \delta'(x - x^*),
\end{aligned} \tag{52}$$

where $H(x)$ is the Heaviside function, $H(x) = 0$ for $x < 0$ and $H(x) = 1$ for $x > 0$; $g^* \stackrel{\text{def}}{=} g(\tau, t)\big|_{t=t^*}$, $\chi_i \stackrel{\text{def}}{=} \varphi_i(\Phi(\varepsilon\tau, t^*), \tau, \varepsilon) - x^*$, and $\Phi(\varepsilon\tau, t^*)$ is the solution of the equation $\varphi_{20}(t^* + \Phi) - \varphi_{10}(t^* + \Phi) = \varepsilon\tau$, which exists in accordance with (50).

Next, the existence of nonsoliton summands in (51) implying a correction of formula (22), namely

$$\begin{aligned}
F(u, \varepsilon u_x) &= \varepsilon \sum_{i=1}^{2} a_{F,i}^{(0)} \beta_i^{-1}\, \delta\big(x - \varphi_i\big) + \varepsilon \beta_2^{-1}\, \mathfrak{R}_F^{(0)}\, \delta(x - x^*) \\
&+ \varepsilon\, F_u'(0,0) \sum_{i=1}^{2} z_i(x,t) H\big(\varphi_i - x\big) + O_{\mathcal{D}'}(\varepsilon^{\overline{2}}),
\end{aligned} \tag{53}$$

where $a_{F,i}^{(0)}$ and $\mathfrak{R}_F^{(0)}$ are defined in (23), (24), $F_u'(0,0) = \partial F(u,0)/\partial u\big|_{u=0}$.

Repeating the same calculations as above, we obtain linear combinations of $\varepsilon\delta(x - \varphi_i)$, $\varepsilon\delta'(x - \varphi_i)$, $\varepsilon H(\varphi_i - x)$, $i = 1, 2$; $\delta(x - x^*)$, $\varepsilon\delta(x - x^*)$, and $\varepsilon\delta'(x - x^*)$. Equating zero, the coefficients of $\varepsilon\delta(x - \varphi_i)$ and $\varepsilon\delta'(x - \varphi_i)$ yield

$$a_1 A_i \frac{d\varphi_{i0}}{dt} = a_{P,i}^{(0)}, \quad a_2 \frac{d}{dt} \frac{A_i^2}{\beta_i} = \frac{a_{K,i}^{(0)}}{\beta_i}, \tag{54}$$

$$a_2 A_i^2 \frac{d\varphi_{i0}}{dt} = a_{P_{2^*}i}^{(0)}, \quad a_1 \frac{d}{dt}\frac{A_i}{\beta_i} + z_i(\varphi_i, t)\frac{d\varphi_{i0}}{dt} = \frac{a_{f,i}^{(0)}}{\beta_i}. \tag{55}$$

Equation (54) forms the closed system to define $A_i(t)$ and $\varphi_{i0}(t)$. To simplify it, let us use the equalities (4) and rewrite the model equation for $\omega(\eta)$ as follows:

$$\frac{d}{d\eta}\left\{-\omega + \frac{\mu+1}{2}\omega^{\mu} + \frac{d^2\omega}{d\eta^2}\right\} = 0. \tag{56}$$

Simple manipulations with (56) allow us to find relations between structural constants:

$$a_1 = (\mu+1)a_{\mu}/2, \quad a_2 = (\mu+3)a_{\mu+1}/4, \quad a_2' = (\mu-1)a_{\mu+1}/4. \tag{57}$$

Next, we use (57), the equality $\beta_i^2 = \gamma A_i^{\mu-1}$, add the initial conditions, and obtain from (54) the Cauchy problem

$$\frac{dA_i}{dt} = -c_1 a_{K_{2^*}i}^{(0)} A_i^{-1}, \quad \frac{d\varphi_{i0}}{dt} = \frac{a_{\mu}}{a_1}A_i^{\mu-1}, \quad t > 0, \tag{58}$$

$$A_i\big|_{t=0} = A_i^0, \quad \varphi_{i0}\big|_{t=0} = x_{(i,0)}, \tag{59}$$

where $c_1 = 2/(a_2(5-\mu))$; $A_i^0 > 0$ and $x_{(i,0)}$ are arbitrary numbers; and $i = 1, 2$. Note also that the first equalities in equations (54) and (55) are equivalent.

Next, equating zero the coefficients of the Heaviside functions, we obtain the equations

$$\frac{\partial z_i}{\partial t} = f_u'(0,0)z_i, \quad x < \varphi_{i0}(t), \quad t > 0, \quad i = 1, 2. \tag{60}$$

In view of (58) $d\varphi_{i0}/dt > 0$, so we use the second equality in (55) to state the correct initial condition for (60)

$$z_i(x,t)\big|_{x=\varphi_{i0}(t)} = \sqrt{\gamma}\, a_{f,i}^{(0)} A_i^{(3-4\mu)/2}(t) + c_2 a_{K_{2^*}i}^{(0)} A_i^{(1-3\mu)/2}(t), \quad t > 0, \tag{61}$$

$$z_i(x,t)\big|_{t=0} = z_i^0(x), \quad x \le x_{(i,0)}, \tag{62}$$

where $c_2 = a_1(3-\mu)(1+\mu)/(2a_2(5-\mu))$, $z_i^0(x)$ is an arbitrary smooth function, which satisfies the consistency condition

$$z_i^0(x_{(i,0)}) = \left\{\sqrt{\gamma}\, a_{f,i}^{(0)} A_i^{(3-4\mu)/2} + c_2 a_{K_{2^*}i}^{(0)} A_i^{(1-3\mu)/2}\right\}\Big|_{t=0}. \tag{63}$$

We should note that the nonlinearity u^{μ} in (1) can require the inequality $u \ge 0$. To this end, we will assume

$$A_i(t) > 0, \quad z_i(x,t)\big|_{x=\varphi_{i0}(t)} \ge 0 \quad \text{for} \quad t \ge 0. \tag{64}$$

Furthermore, equating zero the coefficients of $\delta(x-x^*)$ and $\varepsilon\delta'(x-x^*)$ yield (29), (30) again. Consequently, the condition $\theta \le \theta^*$ guaranties the existence of S_i, φ_{i1} with the properties of (14), (15). In particular

$$S_1 = -\theta S_2 \quad S_1 = \theta\beta_1^{\gamma}\lambda(\sigma)(1+O(\theta^q)), \tag{65}$$

where

$$q = min\{1,\gamma\}, \quad \lambda(\sigma) = a_2^{-1}\int_{-\infty}^{\infty}\omega(\eta)\omega(\eta_{12})d\eta. \tag{66}$$

The last step of the construction is the determination of $\mathfrak{G}_i(\tau)$, $i=1,2$. By setting the coefficients of $\varepsilon\delta(x-x^*)$ zero, we obtain

$$\frac{\partial}{\partial\tau}\sum_{i=1}^{2}\frac{\mathfrak{G}_i}{\beta_i^*} = \mathfrak{F}_1, \quad \frac{\partial}{\partial\tau}\sum_{i=1}^{2}\left(\frac{G_i}{\beta_i}+\frac{G_{\bar{i}}}{\beta_2}\lambda(\sigma)\right)\Big|_{t=t^*}\mathfrak{G}_i = \mathfrak{F}_2, \tag{67}$$

where $\bar{i} = 2$ for $i=1$ and $\bar{i} = 1$ for $i=2$,

$$\mathfrak{F}_1 = -\sum_{i=1}^{2}\left\{\frac{z_i(x^*,t)}{a_1}\frac{d\varphi_{i1}}{d\tau}+\dot{\psi}_0^{-1}\frac{\partial}{\partial t}\frac{S_i}{\beta_i}\right\}\Big|_{t=t^*}+\frac{1}{\beta_2}\mathfrak{R}_f^{(0)}\Big|_{t=t^*}, \tag{68}$$

$$\mathfrak{F}_2 = -\frac{\partial}{\partial\tau}\sum_{i=1}^{2}\left\{a_H z_i(x^*,t)\frac{S_i}{\beta_i}+z_{\bar{i}}(x^*,t)\frac{G_i}{\beta_2}\lambda_{H,i}(\sigma)\right\}\Big|_{t=t^*}+\frac{\mathfrak{R}_{K_2}^{(0)}}{c_H}\Big|_{t=t^*}, \tag{69}$$

$$a_H = a_1/(2a_2), \quad c_H = 2a_2\dot{\psi}_0\beta_2, \quad \lambda_{H,i}(\sigma) = a_2^{-1}\int_{-\infty}^{\infty}\omega(\eta_{i2})\mathfrak{H}(\eta_{\bar{i}2})d\eta.$$

Calculating the determinant Δ of the matrix in the left-hand part of (67) and using (65), we conclude

$$\Delta = \left(G_2-G_1+\lambda(G_1-\theta G_2)\right)\Big|_{t=t^*} = \beta_2^{\gamma}(1-\theta^{\gamma}-\lambda(\theta-\theta^{\gamma})-\lambda^2\theta^{\gamma}(1+O(\theta^q)))\Big|_{t=t^*}. \tag{70}$$

Obviously, $\Delta\neq 0$ for sufficiently small θ. Since the right-hand sides $\mathfrak{F}_i$ belong to the Schwartz space, the functions $\mathfrak{G}_i$ exist and satisfy the assumption (48).

Henceforth, we pass to the final result:

Theorem 4. *Let θ be sufficiently small and let the assumptions (50), (63), and (64), if it is necessary, be fulfilled. Then, the self-similar two-wave weak asymptotic mod $O(\varepsilon^{\bar{q}})$ solution of the equation (1) exists and has the form (46).*

Let us finally stress that the self-similarity implies a special choice of the initial data: for the classical asymptotics in the *C*-sense, there appears a very restrictive condition for small correction of the soliton $A(0)\omega((x-x_0)/\varepsilon)$ (see [6, 17]), and for weak asymptotics, there appears the restriction (63). If it is violated, then the perturbed soliton generates a rapidly oscillating tail of the amplitude $o(1)$ ("radiation") instead of the smooth tails $\varepsilon u^-(x,t)$ (see [21] and numerical results [14, 15, 17]). Nowadays, this radiation phenomenon can be described analytically only for integrable equations, so that we should use self-similar approximation for essentially nonintegrable equations. However, the smooth tail $\varepsilon u^-(x,t)$, which can be treated

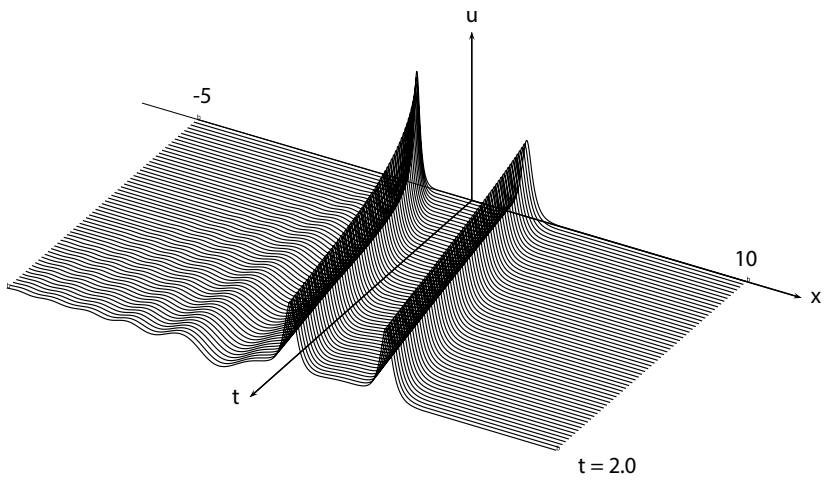

Figure 3. Example of noninteracting solitary waves, $\mu = 4$, $f = u(1 - u)$.

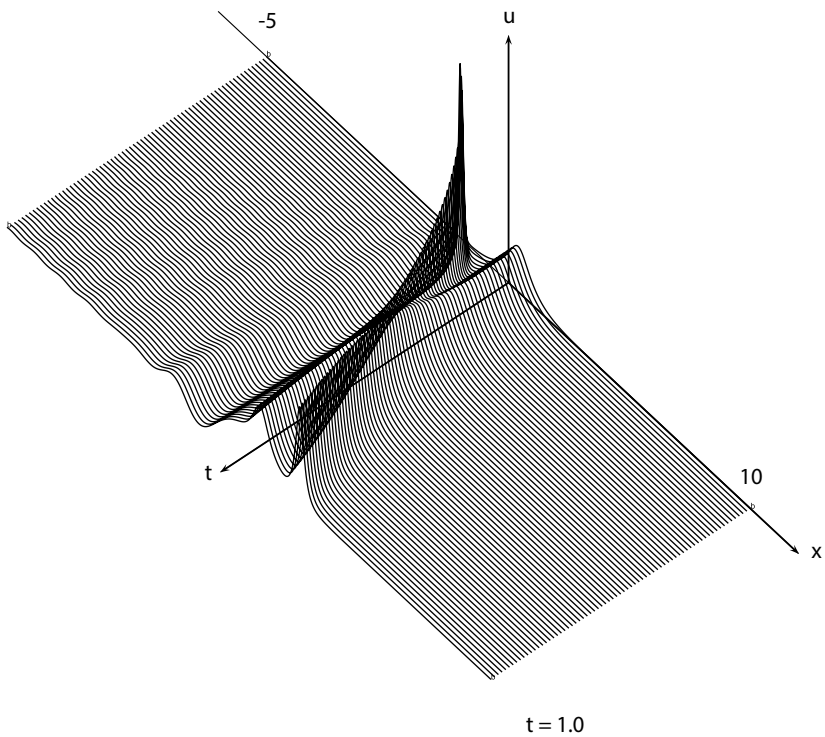

Figure 4. Example of interacting solitary waves, $\mu = 4$, $f = u(1 - u)$, $\varepsilon = 0.1$.

as an average of the radiation, describes sufficiently well the tendency of the radiation amplitude behavior, see graphics depicted in **Figures 3** and **4**, and other numerical results in [15, 17]).

Acknowledgements

The research was supported by SEP-CONACYT under grant no. 178690 (Mexico).

Author details

Georgy Omel'yanov

Address all correspondence to: omel@mat.uson.mx

University of Sonora, Hermosillo, Mexico

References

[1] Bona J, Souganidis P, Strauss W. Stability and instability of solitary waves of Korteweg-de Vries type. Proceedings of the Royal Society of London A. 1987;**411**(1841):395-412. DOI: 10.1098/rspa.1987.0073

[2] Faminskii A. Cauchy problem for the Korteweg-de Vries equation and its generalizations. Journal of Soviet Mathematics. 1990;**50**(1):1381-1420. Available from: https://link.springer.com/article/10.1007/BF01097030

[3] Ablowitz M, Segur H. Solitons and the Inverse Scattering Transform. Philadelphia: SIAM; 1981. Available from: http://dx.doi.org/10.1137/1.9781611970883.bm

[4] Karpman V, Maslov E. Perturbation theory for solitons. Journal of Experimental and Theoretical Physics. 1977;**46**(2):281-291. Available from: http://jetp.ac.ru/cgi-bin/dn

[5] Kaup D, Newell A. Solitons as particles, oscillators and in slowly changing media: A singular perturbation theory. Proceedings of the Royal Society of London Series A. 1978;**361**:413-446. DOI: 10.1098/rspa.1978.0110

[6] Maslov V, Omel'yanov G. Asymptotic soliton-form solutions of equations with small dispersion. Russian Mathematical Surveys. 1981;**36**:73-149. Available from: http://iopscience.iop.org/0036-0279/36/3/R02

[7] Gorshkov K, Ostrovsky L. Interaction of solitons in nonintegrable systems: Direct perturbation method and applications. Physica D. 1981;**3**:428-438. Available from: https://doi.org/10.1016/0167-2789(81)90146-9

[8] Danilov V, Shelkovich V. Generalized solutions of nonlinear differential equations and the Maslov algebras of distributions. Integral Transformations and Special Functions. 1997;**6**:137-146. Available from: http://dx.doi.org/10.1080/10652469808819161

[9] Danilov V, Omel'yanov G. Weak asymptotics method and the interaction of infinitely narrow delta-solitons. Nonlinear Analysis: Theory Methods & Applications. 2003;**54**:773-799 Available from: https://doi.org/10.1016/S0362-546X(03)00104-4

[10] Abreu E, Colombeau M, Panov E. Weak asymptotic methods for scalar equations and systems. Journal of Mathematical Analysis and Applications. 2016;**444**(2):1203-1232. Available from: https://doi.org/10.1016/j.jmaa.2016.06.047

[11] Danilov V, Mitrovic D. Shock wave formation process for a multidimensional scalar conservation law. Quarterly of Applied Mathematics. 2011;**69**(4):613-634. Available from: doi: 10.1090/S0033-569X-2011-01234-9

[12] Danilov V, Omel'yanov G, Shelkovich V. Weak asymptotics method and interaction of nonlinear waves. In: Karasev M, editor. Asymptotic Methods for Wave and Quantum Problems. Trans. Ser. 2. Vol. 208. Providence, RI: American Mathematical Society; 2003. p. 33-164. Available from: https://goo.gl/aPPhqG

[13] Danilov V, Shelkovich V. Dynamics of propagation and interaction of delta-shock waves in conservation law systems. Journal of Differential Equations. 2005;**211**(2):333-381. Available from: https://goo.gl/Y6g4vv

[14] Garcia M, Omel'yanov G. Interaction of solitary waves for the generalized KdV equation. Communications in Nonlinear Science and Numerical Simulation. 2012;**17**(8):3204-3218. Available from: https://doi.org/10.1016/j.cnsns.2011.12.001

[15] Garcia M, Omel'yanov G. Interaction of solitons and the effect of radiation for the generalized KdV equation. Communications in Nonlinear Science and Numerical Simulation. 2014;**19**(8):2724-2733. Available from: https://doi.org/10.1016/j.cnsns.2013.06.024

[16] Kalisch H, Mitrovic D. Singular solutions of a fully nonlinear 2 × 2 system of conservation laws. Proceedings of the Edinburgh Mathematical Society II. 2012;**55**:711-729. DOI: https://doi.org/10.1017/S0013091512000065

[17] Omel'yanov G, Valdez-Grijalva M. Asymptotics for a C^1-version of the KdV equation. Nonlinear Phenomena in Complex Systems. 2014;**17**(2):106-115. Available from: http://elib.bsu.by/handle/123456789/116849

[18] Omel'yanov G. Soliton-type asymptotics for non-integrable equations: A survey. Mathematical Methods in The Applied Sciences. 2015;**38**(10):2062-2071. DOI: 10.1002/mma.3198 2014

[19] Omel'yanov G. Propagation and interaction of solitons for nonintegrable equations. Russian Journal of Mathematical Physics. 2016;**23**(2):225-243. DOI: 10.1134/S10619 20816020084

[20] Omel'yanov G. Multi-soliton Collision for essentially nonintegrable equations. In: Oberguggenberger M et al., editors. Generalized Functions and Fourier Analysis, Series: Operator Theory: Advances and Applications. Birkhäuser, Cham. 2017;**260**:153-170. DOI: https//doi.org/10.1007/978-3-319-51911-1_9

[21] Kalyakin L. Perturbation of the Korteweg-de Vries soliton. Theoretical and Mathematical Physics. 1992;**92**:736-747. Available from: https://goo.gl/aPPhqG

Chapter 9

Weakly Nonlinear Stability Analysis of a Nanofluid in a Horizontal Porous Layer Using a Multidomain Spectral Collocation Method

Osman A.I. Noreldin, Precious Sibanda and Sabyasachi Mondal

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71066

Abstract

In this chapter, we present a weakly nonlinear stability analysis of the flow of a nanofluid in a porous medium with stress-free boundary conditions. Some previous studies have investigated cross-diffusion in a nanofluid layer although in most cases these studies mostly deal with linear stability analysis. It is important to study the nonlinear stability in flows subject to cross-diffusion due to the wide range of applications where such flows arise such as in hydrothermal growth, compact heat exchanges, the solidification of binary mixtures, geophysical systems, solar pond, etc. Here we consider flow between parallel plates with an applied magnetic field and zero nanoparticle flux at the boundaries. A truncated Fourier series is introduced reducing the flow equations to a Lorenz-type system of nonlinear evolution equations. The multidomain spectral method is used to solve the equations that describe the growth of the convection amplitudes. The solutions are obtained as sets of trajectories in the phase space. Some interesting spiral trajectories and their sensitivity to the Rayleigh number are given.

Keywords: nonlinear instability, nanofluid flow, porous medium, multidomain spectral collocation method

1. Introduction

The enhancement of thermal conductivity of a fluid is a matter of supreme interest to engineers due to the important applications of fluids in heat transfer processes. Natural and forced convection plays an important role in heat transfer processes due to continuous molecular movements in fluid. Recent studies show that the suspension of solid nanoparticles in a fluid can substantially improve the fluid's thermophysical properties, including thermal conductivity.

The term nanofluid describes a liquid containing a suspension of nanometer sized 1–100 nm solid particles [1]. Examples of commonly used nanoparticles include metallic particles such as Al, Cu and Ag, and oxides such as Al_2O_3 and CuO. The base fluid is often a common liquid such as water, ethylene, glycol, or oil. The enhancement of thermochemical properties of a fluid due to the addition of nanoparticles has been observed in experimental studies such as in [2, 3]. Researchers have investigated the influence of seven slip mechanisms, namely, inertia, Brownian diffusion, thermophoresis, diffusiophoresis, magnus effect, fluid drainage, and gravity in nanofluids. It has been shown that, in the absence of turbulence, the most significant among these mechanisms are the Brownian diffusion and thermophoresis.

The classical Rayleigh-Benard convection problem in a heated horizontal layer has been extensively studied in the literature. Among recent studies on nanofluids, Tzou [4] studied the thermal instability and natural convection in nanofluid flow using an eigenfunction expansion method. Narayana et al. [5, 6] studied convection and the stability of a Maxwell fluid in a porous medium. Yadav et al. [7] investigated thermal instability of a rotating nanofluid layer. The studies by Kuznetsov and Nield [8–11] focused on thermal instability in a porous layer saturated with a nanofluid. They investigated the onset of instability in a horizontal porous layer using a model for the nanofluid that incorporated particle Brownian motion and thermophoresis. Related studies with various assumptions on the geometry and flow structure have been made by [12–15]. In the last few decades, researchers have also investigated thermal instability in a horizontal nanofluid layer subject to an applied magnetic field [16, 17]. The effects of a magnetic field on convection and the onset of instability have important applications in problems such as in cooling systems, pumps, magnetohydrodynamics and generators. The experimental study by Heris et al. [18] showed that thermal efficiency could be achieved by subjecting the flow to a magnetic field. The studies by Ghasemi et al. [19] and Hamad et al. [20] focused on the flow behavior and heat transfer in an electrically conducting nanofluid under the influence of a magnetic field and subject to Brownian diffusion and thermophoresis. They used a water-based nanofluid containing different types of nanoparticles such as copper, alumina and silver in their numerical simulations. Related studies of interest include [21–24]. Rana et al. [25] studied thermal convection in a Walters (Model B) fluid in a porous medium. They showed that a magnetic field may introduce oscillatory instability modes and acts to stabilize the system.

In this chapter, we give a weakly nonlinear stability analysis of a nanofluid layer with an applied magnetic field, stress free boundary conditions and under the assumption of zero nanoparticle flux at the boundary. The studies by Kuznetsov and Nield [9] and Nield and Kuznetsov [10, 11] investigated cross-diffusion in a nanofluid layer. However, these studies mostly presented a linear stability analysis. It is important to study the nonlinear regime for a nanofluid flow subject to cross-diffusion due to the wide range of applications where such flows may arise. Typical examples may be found in hydrothermal growth, compact heat exchanges, solidification of binary mixtures, geophysical systems, and so on. Hence, with this in mind, we studied the finite amplitude convection in a nanofluid flows subject to cross-diffusion. By introducing a truncated Fourier series, a Lorenz-type system of seven nonlinear differential equations is obtained. The recent multidomain spectral method is used to solve the nonlinear equations. This method is accurate and very easy to implement compared to older methods such as finite difference methods. An analysis of heat and mass transfer for different parameters such as the Prandtl number, the Dufour and thermophoresis is presented.

http://dx.doi.org/10.5772/intechopen.71066

2. Mathematical formulation

Consider viscous incompressible MHD nanofluid flow in an infinitely extended horizontal porous layer, confined between two boundaries at $z = 0$ and $z = h$. The layer is heated from below and cooled from above, see **Figure 1**. A Cartesian frame of reference is chosen in which the z-axis is vertically upward. The boundaries are perfectly conducting. The temperature at the lower and upper walls is T_c and T_h, respectively with $T_h > T_c$. The Oberbeck-Boussinesq approximation and the Darcy law are assumed to be applicable. The continuity equation, momentum equation, energy equation, concentration equation and volumetric fraction nanoparticle equation, which describe the above configuration in dimensionless form, are given as

$$\nabla \cdot V = 0, \tag{1}$$

$$\frac{Da}{\mathrm{Pr}}\frac{\partial V}{\partial t} = -\nabla P + Da\nabla^2 V - V + Q\widehat{Ve}_z - Rm\widehat{e}_z + RaT\widehat{e}_z + RsC\widehat{e}_z - Rn\phi\widehat{e}_z, \tag{2}$$

$$\frac{\partial T}{\partial t} + V \cdot \nabla T = \nabla^2 T + \frac{N_B}{Les}\nabla\phi \cdot \nabla T + \frac{N_A N_B}{Les}\nabla T \cdot \nabla T + Du\nabla^2 C, \tag{3}$$

$$\frac{\partial C}{\partial t} + V \cdot \nabla C = \frac{1}{Le}\nabla^2 C + Sr\nabla^2 T, \tag{4}$$

$$\frac{1}{\sigma}\frac{\partial \phi}{\partial t} + \frac{1}{\varepsilon}V \cdot \nabla\phi = \frac{1}{Les}\nabla^2\phi + \frac{N_A}{Les}\nabla^2 T, \tag{5}$$

subject to the boundary conditions

$$V = 0, T = 1, C = 1\frac{\partial \phi}{\partial z} + N_A\frac{\partial T}{\partial z} = 0 \text{ at } z = 0, \tag{6}$$

$$V = 0, \quad T = 0, \quad C = 0 \qquad \frac{\partial \phi}{\partial z} + N_A\frac{\partial T}{\partial z} = 0 \quad \text{at} \quad z = 1, \tag{7}$$

where V is the fluid velocity, T is the temperature, C is the solute concentration and ϕ is the volumetric fraction of nanoparticles. The dimensionless parameters are the Darcy number (modified by the viscosity ratio) Da, Prandtl number Pr, Hartmann-Darcy number Q, thermal Rayleigh-Darcy number Ra, nanoparticle Rayleigh number Rn and the basic density Rayleigh number Rm. The parameter N_A is a modified diffusivity ratio, Le is the Lewis number, Rs is solutal Rayleigh number, N_B is a modified nanoparticle density increment and Du is a modified Dufour parameter. The parameter Les is the thermo-nanofluid Lewis number, ν is the kinematic viscosity and Sr is a modified Soret parameter. These parameters have the form

$$Da = \frac{\tilde{\mu}K}{\mu h^2}, \quad \mathrm{Pr} = \frac{\mu}{\rho_f \alpha_m}, \quad Q = \frac{\delta \mathbf{B}_0^2 K}{\mu}, \quad Ra = \frac{\rho_f \beta K h g\left(T_h^* - T_c^*\right)}{\mu\alpha_m}, \quad Les = \frac{\alpha_m}{D_B}, \tag{8}$$

$$Rn = \frac{\left(\rho_p - \rho_f\right)\left(\phi_1^* - \phi_0^*\right)gKh}{\mu\alpha_m}, \quad Rm = \frac{\rho_p\phi_0^* + \left(1 - \phi_0^*\right)\rho_f gKh}{\mu\alpha_m}, \quad Le = \frac{\alpha_m}{D_S}, \tag{9}$$

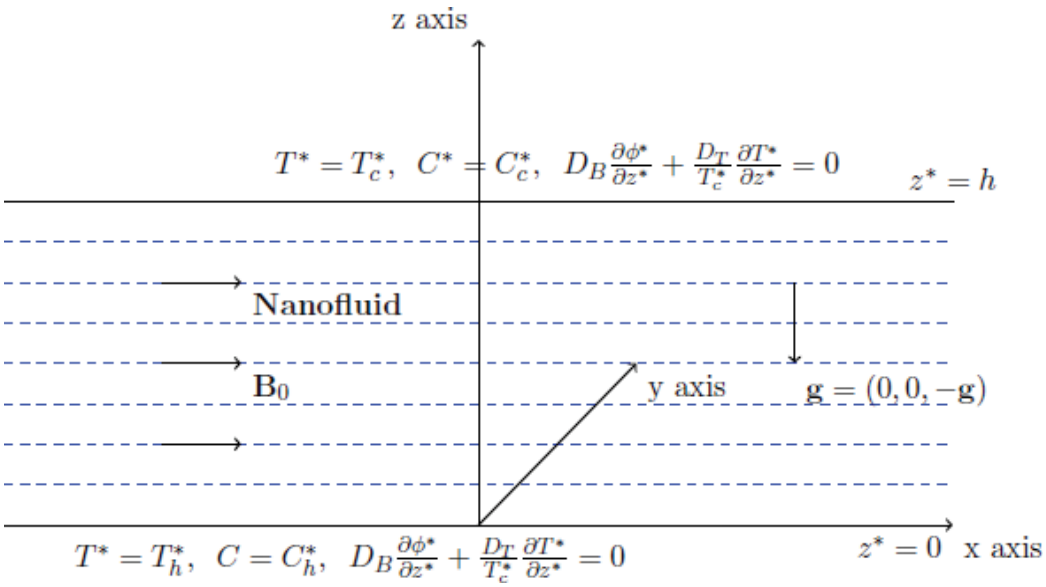

Figure 1. A schematic diagram of the problem.

$$N_B = \frac{\varepsilon(\rho c)_p}{(\rho c)_f}\left(\phi_1^* - \phi_0^*\right), \quad N_A = \frac{D_T\left(T_h^* - T_c^*\right)}{D_B T_c^*\left(\phi_1^* - \phi_0^*\right)}, \quad Rs = \frac{\rho_f \beta K h g\left(C_h^* - C_c^*\right)}{\mu \alpha_m}, \tag{10}$$

$$Du = \frac{\sigma D_{TC}\left(C_h^* - C_c^*\right)}{\alpha_m\left(T_h^* - T_c^*\right)}, \quad Sr = \frac{\sigma D_{CT}\left(T_h^* - T_c^*\right)}{\alpha_m\left(C_h^* - C_c^*\right)}, \tag{11}$$

where ρ_f, ρ_p, $\tilde{\mu}$, β_1, β_2, κ_m, δ, ε and K are the fluid density, nanoparticle density, effective viscosity of porous medium, thermal volumetric expansion coefficient of the fluid, solutal volumetric expansion coefficient, the thermal conductivity of porous medium, the electrical conductivity, the porosity, and permeability of porous medium, respectively. The gravitational acceleration is denoted by g and D_B is the Brownian diffusion coefficient, D_T is the thermophoresis diffusion coefficient, D_S is the solutal diffusion coefficient, D_{TC} is the Dufour parameter and D_{CT} is the Soret parameter. The heat capacity of the fluid is $(\rho c)_f$, $(\rho c)_p$ is the effective heat capacity of the nanoparticle, $(\rho c)_m$ is the effective heat capacity of the porous medium and $\mathbf{B}_0$ is the uniform magnetic field strength.

The basic state is the time independent solution of Eqs. (1)–(5). Solving these equations with boundary conditions, we obtain

$$T_b = 1 - z, \quad C_b = 1 - z, \quad \phi_b = \phi_0 + N_A z. \tag{12}$$

3. Weakly nonlinear stability analysis

In this section, we restrict the analysis to the case of two-dimensional disturbances. We define the stream function Ψ by the equations

$$u = \frac{\partial \Psi}{\partial z}, \quad w = -\frac{\partial \Psi}{\partial x}.$$

Eqs. (1)–(5) may now be simplified by introducing the truncated Fourier series

$$\Psi' = A_{11} \sin \alpha x \sin \pi z, \quad T' = B_{11} \cos \alpha x \sin \pi z + B_{02} \sin 2\pi z, \tag{13}$$

$$C' = C_{11} \cos \alpha x \sin \pi z + C_{02} \sin 2\pi z, \quad \phi' = -N_A(D_{11} \cos \alpha x \sin \pi z + D_{02} \sin 2\pi z), \tag{14}$$

where A_{11}, B_{11}, B_{02}, C_{11}, C_{02}, D_{11} and D_{02} are amplitudes that depend on time. This leads to the Lorenz-type system of nonlinear ordinary differential equations

$$\dot{Y}_1 = \frac{\mathrm{Pr}}{Da}\left\{-BY_1 - N(Y_2 + Y_4) + \frac{N_A Rn}{R} Y_6\right\} \tag{15}$$

$$\dot{Y}_2 = RY_1 - Y_2 - DuY_4 - Y_1Y_3 \tag{16}$$

$$\dot{Y}_3 = \frac{1}{2} Y_1Y_2 - G(Y_3 + DuY_5) \tag{17}$$

$$\dot{Y}_4 = RY_1 - \frac{1}{Les} Y_4 - SrY_2 - Y_1Y_5 \tag{18}$$

$$\dot{Y}_5 = \frac{1}{2} Y_1Y_2 - G\left(\frac{1}{Les} Y_5 + SrY_3\right) \tag{19}$$

$$\dot{Y}_6 = \frac{\sigma R}{\varepsilon} Y_1 - \frac{\sigma}{Le}(Y_6 - Y_2) - \frac{N_A \sigma}{\varepsilon} Y_1Y_7 \tag{20}$$

$$\dot{Y}_7 = \frac{N_A \sigma}{2\varepsilon} Y_1Y_7 - \frac{G}{Le}(Y_7 - Y_3) \tag{21}$$

subject to $Y_n(0) = Y_n^0$ for $n = 1, 2, \cdots, 7$. The following variables have been introduced in the equations above:

$$Y_1 = \frac{\alpha\pi}{\gamma} A_{11}, \quad Y_2 = -\pi R B_{11}, \quad Y_3 = -\pi R B_{02}, \quad Y_4 = -\pi R C_{11}, \quad Y_5 = -\pi R C_{20},$$

$$Y_6 = -\pi R D_{11}, \quad Y_7 = -\pi R D_{20}, \quad t^* = \gamma t, \quad R = \frac{\alpha^2}{\gamma^3} Ra, \quad G = \frac{4\pi^2}{\gamma} \quad \text{and} \quad N = \frac{Rs}{Ra},$$

$$B = \frac{Da\gamma^2 + \gamma - \alpha^2 Q}{\gamma^2}.$$

Eqs. (15)–(21) give an approximate description of the full dimensional nonlinear system. An analytical solution of the system of nonlinear ordinary differential Eqs. (15)–(21) is not possible for the general time variable t. However, it is possible to discuss the stability of the nonlinear system of equations. The system of equations is uniformly bounded in time and dissipative in the phase space. We can easily show that

$$\sum_{i=1}^{7} \frac{\partial \dot{Y}_i}{\partial Y_i} = -\left[\frac{DaB}{\Pr} + 1 + G + Les^{-1} + GLes^{-1} + \sigma Le^{-1} + GLe^{-1}\right]. \tag{22}$$

This is always true if $B \geq 0$. As has been shown in previous studies, the trajectories may be attracted to a fixed point, limit cycle or other attractor. For a set of initial points in the phase space occupying a region $V(0)$ at time $t = 0$, after a time $t > 0$, the end point of the corresponding trajectories fills a volume

$$V(t) = V(0)\exp\left\{-\left[\frac{DaB}{\Pr} + 1 + G + Les^{-1} + GLes^{-1} + \sigma Le^{-1} + GLe^{-1}\right]t\right\}. \tag{23}$$

Eq. (23) shows that the volume decays exponentially with time. Further, it can be noted that the system of Eqs. (15)–(21) are invariant under the transformation

$$S(Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7) \rightarrow -S(Y_1, Y_2, Y_3, Y_4, Y_5, Y_6, Y_7). \tag{24}$$

We obtain the possible stationary points of the nonlinear system of equations by setting $\dot{Y}_i = 0$ for $i = 1, 2, \cdots, 7$. One of these stationary points is $Y_i = 0$ and by linearizing about this point, we obtain the Jacobian matrix

$$A = \begin{pmatrix} -\frac{PrB}{Da} & -\frac{\Pr}{Da} & 0 & \frac{PrN}{Da} & 0 & \frac{PrN_A Rn}{DaR} & 0 \\ R & -1 & 0 & -Du & 0 & 0 & 0 \\ 0 & 0 & -G & 0 & -GDu & 0 & 0 \\ R & -Sr & 0 & -Les^{-1} & 0 & 0 & 0 \\ 0 & 0 & -GSr & 0 & -GLes^{-1} & 0 & 0 \\ \frac{\sigma R}{\varepsilon} & \sigma Le^{-1} & 0 & 0 & 0 & -\sigma Le^{-1} & 0 \\ 0 & 0 & GLe^{-1} & 0 & 0 & 0 & -GLe^{-1} \end{pmatrix}. \tag{25}$$

The eigenvalues of the above matrix depend on the various parameters. For the specific parameters $R = 10^3$, $Da = 20$, $\Pr = 10$, $N = 25$, $Du = 0.2$, $Sr = 3$, $Les = 10$, $Le = 5$, $\sigma = 0.05$, $G = 3$ and $\varepsilon = 0.04$, the characteristic polynomial is

$$P(\lambda) = \lambda^7 + 21.41\lambda^6 + 651.5\lambda^5 + 5391.7\lambda^4 + 12772.232\lambda^3 - 370.962\lambda^2 - 2996.712\lambda + 545.8$$

with eigenvalues

$$\lambda_1 = 0.2955056985,\ \lambda_2 = 0.2402382976,\ \lambda_3 = -0.6139990637,\ \lambda_4 = -4.886000936,$$
$$\lambda_5 = -5,\ \lambda_6 = -5.7228719981 - 21.9033954659i,\ \lambda_7 = -5.7228719981 + 21.9033954659i.$$

This stationary point is a saddle point. Nonetheless, because the eigenvalues depend on various parameters, we cannot make general conclusions as to the stability of the system. We note, however, that if we denote the trace of the matrix A by T and the determinant d, then

$$T = -\left(\frac{DaB}{\mathrm{Pr}} + 1 + G + Les^{-1} + GLes^{-1} + \sigma Le^{-1} + GLe^{-1}\right), \tag{26}$$

and

$$d = \frac{\sigma \mathrm{Pr} G^3}{DaLeLes}\left(-\frac{DuNRSr}{Le} - \frac{DuNRSr}{\varepsilon} + \frac{BDuSr}{Le} + \frac{DuRSr}{Le} + \frac{NR}{LeLes} + \frac{NR}{Les\varepsilon} - \frac{B}{LeLes} - \frac{R}{LeLes}\right). \tag{27}$$

The trace is always negative, but the sign of determinant depends on the parameter values. If $d < 0$ then

$$(1-N)\varepsilon DuSrRLes + (\varepsilon B - NRLe)DuSrLes + (\varepsilon + Le)NLe < (B+R)\varepsilon, \tag{28}$$

suggesting a saddle point.

4. Method of solution

To study the influence of various physical parameters on the average Nusselt and Sherwood numbers, we solved the nonlinear system of Eqs. (15)–(21) numerically using the multidomain spectral collocation method. This is a novel technique for solving nonlinear initial value problems and parabolic equations with large time domains. It has been suggested in the literature that the method gives better accuracy compared to other methods such as finite difference and Runge-Kutta methods [26]. To apply the multidomain spectral collocation to the nonlinear system of equations, we first divide the interval $[0, T]$ into subintervals $\Omega_i = [t_{i-1}, t_i]$ for $i = 1, 2, \cdots, p$. The transformation

$$t = \frac{t_i - t_{i-1}}{2}\tau + \frac{t_i + t_{i-1}}{2} \tag{29}$$

is used to transform each subinterval Ω_i into the interval $[-1, 1]$. The system of Eqs. (15)–(21) can be written in the form

$$\frac{dY_1^i}{dt} = \frac{\mathrm{Pr}}{Da}\left\{-BY_1^i - N\left(Y_2^i + Y_4\right) + \frac{N_A Rn}{Ra}Y_6^i\right\}, \tag{30}$$

$$\frac{dY_2^i}{dt} = RY_1^i - Y_2^i - DuY_4^i - Y_1^i Y_3^i, \tag{31}$$

$$\frac{dY_3^i}{dt} = \frac{1}{2}Y_1^i Y_2^i - G\left(Y_3^i + DuY_5^i\right), \tag{32}$$

$$\frac{dY_4^i}{dt} = RY_1^i - \frac{1}{Les}Y_4^i - SrY_2^i - Y_1^i Y_5^i, \tag{33}$$

$$\frac{dY_5^i}{dt} = \frac{1}{2}Y_1^i Y_2^i - G\left(\frac{1}{Les}Y_5^i + SrY_3^i\right), \tag{34}$$

$$\frac{dY_6^i}{dt} = \frac{\sigma R}{\varepsilon} Y_1^i - \frac{\sigma}{Le}\left(Y_6^i - Y_2^i\right) - \frac{N_A \sigma}{\varepsilon} Y_1^i Y_7^i, \tag{35}$$

$$\frac{dY_7^i}{dt} = \frac{N_A \sigma}{2\varepsilon} Y_1^i Y_7^i - \frac{G}{Le}\left(Y_7^i - Y_3^i\right), \tag{36}$$

subject to

$$Y_n^i(t_{i-1}) = Y_n^{i-1}(t_{i-1}) \quad \text{for} \quad n = 1, 2, \cdots, 7. \tag{37}$$

The first step in using the multidomain spectral collocation method (MDSCM) concerns the quasilinearization of Eqs. (30)–(36) leading to a system of equations in the form

$$\sum_{n=1}^{7} a_{(j,n)r}^i Y_{n,r+1}^i - \frac{dY_{j,r+1}^i}{dt} = R_{jr}^i, \tag{38}$$

subject to

$$Y_{n,r+1}^i(t_{i-1}) = Y_{n,r+1}^{i-1}(t_{i-1}) \quad \text{for} \quad n = 1, 2, \cdots, 7. \tag{39}$$

where $a_{(j,n)r}^i$ and R_{jr}^i for $j = 1, 2, \ldots, 7$ are given in the Appendix. Having linearized the equations, the second step is to integrate Eqs. (30)–(36). To this end, we use the Gauss-Lobatto nodes

$$\tau_j^i = \cos\frac{\pi j}{Nc}, \quad \text{for } j = 0, 1, \cdots, Nc. \tag{40}$$

We approximate the derivatives of the unknown functions $Y_{n,r+1}^i(t)$ at the collocation points by

$$\frac{dY_{n,r+1}^i}{dt}\left(\tau_j^i\right) = \sum_{k=0}^{Nc} \mathbf{D}_{jk} Y_{n,r+1}^i\left(\tau_j^j\right) = \left[\mathbf{D}\mathbf{U}_{n,r+1}^i\right]_j, \tag{41}$$

where $\mathbf{D} = 2D/(t_i - t_{i-1})$, D is the Chebyshev differentiation matrix and

$$\mathbf{U}_{n,r+1}^i = \left(Y_{n,r+1}^i(\tau_0^i), \cdots Y_{n,r+1}^i(\tau_{Nc}^i)\right)^T,$$

is a vector of the unknown functions at the collocation points. Substituting Eq. (41) into Eqs. (38) and reducing the result into matrix form, we obtain

$$\begin{aligned} \mathbf{A}\mathbf{U}_{n,r+1}^i &= \mathbf{R}_n^i, \\ \mathbf{U}_{n,r+1}^i\left(\tau_{N_c}^{i-1}\right) &= \mathbf{U}_n^i\left(\tau_{N_c}^{i-1}\right), \quad n = 1, 2, \cdots, 7. \end{aligned} \tag{42}$$

where the matrices $\mathbf{A} = \left[A_{ij}\right]$ and $\mathbf{R}_n^i$ are given in the Appendix.

http://dx.doi.org/10.5772/intechopen.71066

5. Heat and mass transfer

The study of heat and mass transfer in a horizontal nanofluid layer heated from below and cooled from above has important engineering applications. We define the rate of heat transfer by the average Nusselt number $Nu(t)$ where

$$Nu(t) = 1 + \left[\frac{\frac{\alpha}{2\pi}\int_{0}^{\frac{2\pi}{\alpha}}\frac{\partial T}{\partial z}dx}{\frac{\alpha}{2\pi}\int_{0}^{\frac{2\pi}{\alpha}}\frac{\partial T_b}{\partial z}dx}\right]_{z=0} + Du\left\{1 + \left[\frac{\frac{\alpha}{2\pi}\int_{0}^{\frac{2\pi}{\alpha}}\frac{\partial C}{\partial z}dx}{\frac{\alpha}{2\pi}\int_{0}^{\frac{2\pi}{\alpha}}\frac{\partial C_b}{\partial z}dx}\right]_{z=0}\right\}. \tag{43}$$

Substituting Eqs. (12) and (13) into Eq. (43), we obtain

$$Nu(t) = 1 + \frac{2}{R}Y_3 + Du\left(1 + \frac{2}{R}Y_5\right). \tag{44}$$

Similarly, the rate of mass transfer stated in terms of the average Sherwood number is

$$Sh(t) = 1 + \frac{2}{R}Y_5 + Sr\left(1 + \frac{2}{R}Y_3\right) \tag{45}$$

6. Results and discussion

We have studied the weakly nonlinear instability of nanofluid flow in a horizontal layer with stress free boundary conditions. For numerical simulations, the parameter values were chosen from the literature on nanofluid flow such as [4, 7]. In the literature, the critical Rayleigh number is found when the Darcy number is very large. In this study, we investigated the critical Rayleigh number for low Darcy numbers.

The method of solution described in Section 4 was used to solve Eqs. (15)–(21). All computations are carried out up to a value of maximum time $t_{max} = 1$, and solutions are obtained using initial conditions selected in the neighborhood of stationary points. Periodic solution sets were obtained for the system of nonlinear equations. We determined the rate of heat and mass transfer as functions of time for different parameter values. The results are shown in **Figures 2–4**. **Figure 2** shows the effect of the Dufour and Soret parameters on the Nusselt and Sherwood numbers with time t. **Figure 2(a)** shows how the heat transfer coefficient changes with both the Dufour parameter and time. The heat transfer coefficient increases with the Dufour parameter but eventually settles to a steady value with time. In **Figure 2(b)**, the Soret parameter is similarly shown to enhance the mass transfer coefficient. We investigated the effect of the Prandtl and Lewis numbers (see **Figures 3** and **4**). An increase in the Lewis number enhances both heat and mass transfer in a nanofluid layer heated from below. However, **Figure 3** shows that increasing the Prandtl number reduces the amplitude of oscillatory heat and mass transfer. The Prandtl number can lead to both positive and negative contributions to the Nusselt and Sherwood numbers. It is interesting to note that our investigation shows that the magnetic field parameter has very little effect on the heat and mass transfer for this type of flow.

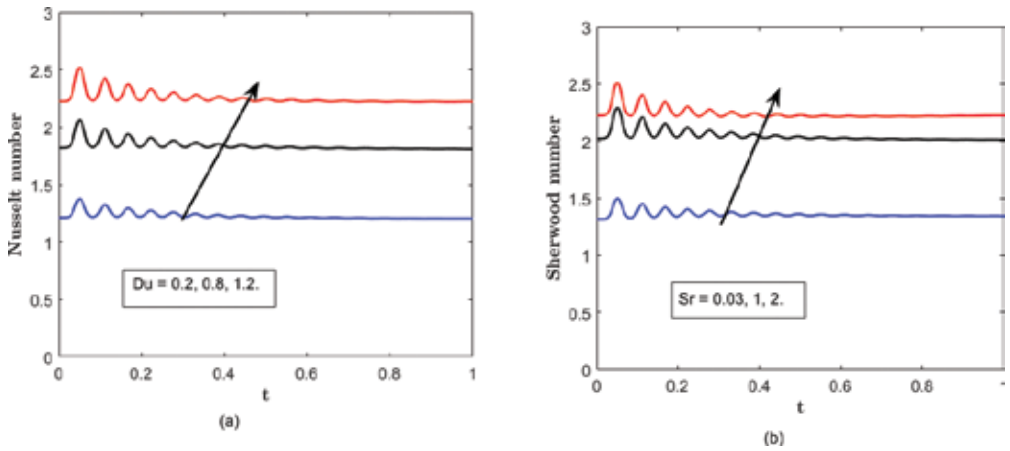

Figure 2. The effect of cross-diffusive parameters on (a) the Nusselt number Nu and (b) the Sherwood number Sh for $Da = 0.05$, $Le = 2$, $Du = 0.2$, $\varepsilon = 0.04$, $\sigma = 0.05$, $Les = 100$, $Rn = 5$, $Ra = 1000$ and various values of the Dufour and Soret parameters.

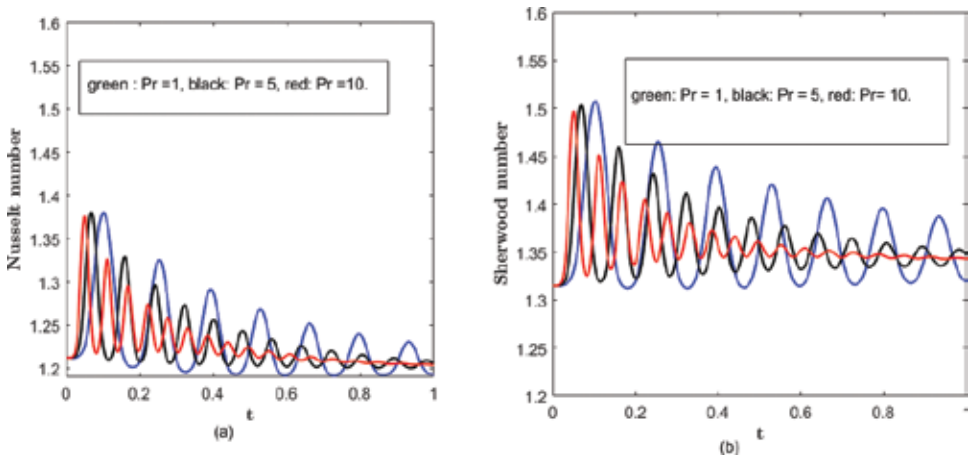

Figure 3. The effect of Prandtl number Pr on (a) the Nusselt number Nu and (b) the Sherwood number Sh when $Da = 0.05$, $Le = 2$, $Du = 0.2$, $\varepsilon = 0.04$, $\sigma = 0.05$, $Les = 100$, $Rn = 5$ and $Ra = 1000$.

Figures 5–11 show the effect of the Rayleigh number on the trajectories projected onto the (Y_i, Y_j) phase planes. The solution sets provide a visual representation of the system's behavior with every phase point on the phase space representing the physical state of the system. The convective solution sets for different values of R have been presented with the trajectories projected onto the (Y_i, Y_j) phase planes. These trajectories spiral toward the fixed point for Rayleigh numbers from 10^2 to 10^4. The solution sets give spiral phase portraits as R increases and for the high Rayleigh numbers, the trajectories spiral many times before they reach a fixed point.

Figures 5–8 show the phase portraits projected onto the (Y_i, Y_j)- plane correspond to a simple spiral for $R = 100$. As R is increased to 10^4, the complexity of the trajectories

http://dx.doi.org/10.5772/intechopen.71066

increases leading to certain chaotic forms. **Figures 8–11** show the trajectories in the three-dimensional phase space. Here, we observe similar solution sets as in the two-dimensional phase portraits.

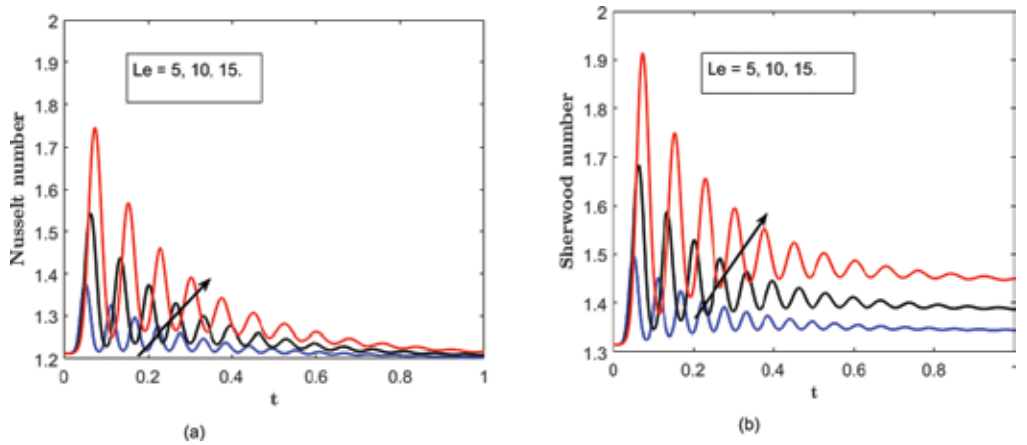

Figure 4. The effect of Lewis number on (a) the Nusselt number Nu and (b) the Sherwood number Sh when $Da = 0.05$, $Du = 0.2$, $\varepsilon = 0.04$, $\sigma = 0.05$, $Les = 100$, $Rn = 5$ and $Ra = 1000$.

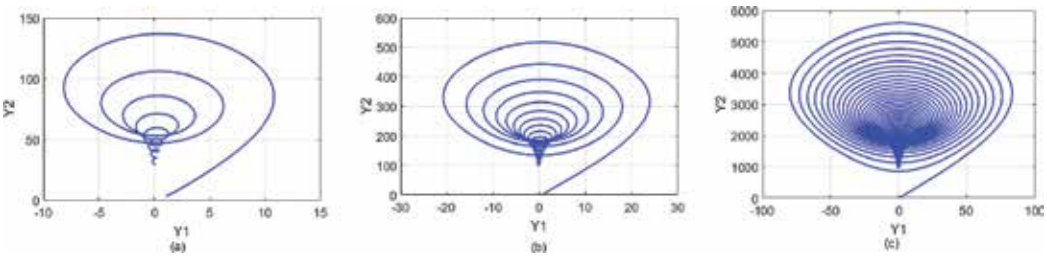

Figure 5. The trajectories of the system of nonlinear equations projected on the Y_1, Y_2-plane when revised Rayleigh number (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$ when $Da = 0.05$, $Les = 100$, $\sigma = 0.05$, $\varepsilon = 0.04$, $Q = 10$, $Du = 0.2$, $Sr = 0.3$, $Le = 2$, $N = 25$ and $Rn = 5$.

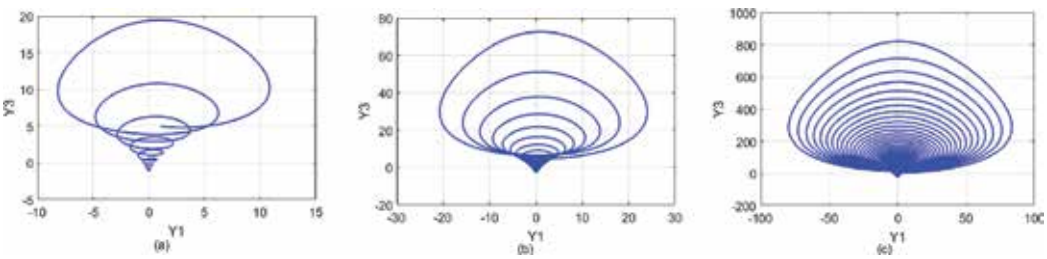

Figure 6. Trajectories of the system of nonlinear equations projected on the (Y_1, Y_3) plane when the revised Rayleigh number (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$ when $Da = 0.05$, $Les = 100$, $\sigma = 0.05$, $\varepsilon = 0.04$, $Q = 10$, $Du = 0.2$, $Sr = 0.3$, $Le = 2$, $N = 25$ and $Rn = 5$.

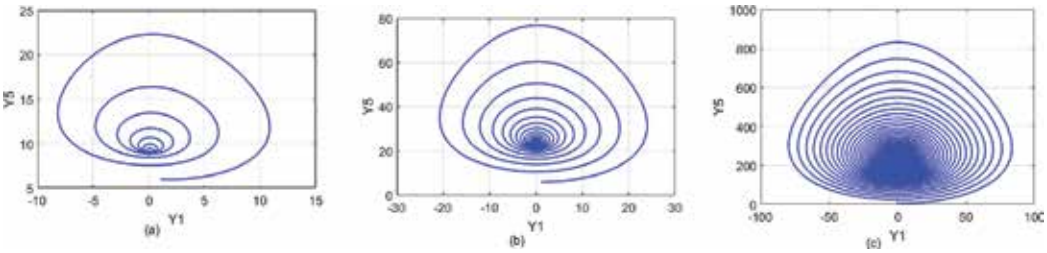

Figure 7. Trajectories of the system of nonlinear equations projected on the (Y_1, Y_5)-plane showing the sensitive dependence of the trajectories on the revised Rayleigh number for (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$ when $Da = 0.05$, $Les = 100$, $\sigma = 0.05$, $\varepsilon = 0.04$, $Q = 10$, $Du = 0.2$, $Sr = 0.3$, $Le = 2$, $N = 25$ and $Rn = 5$.

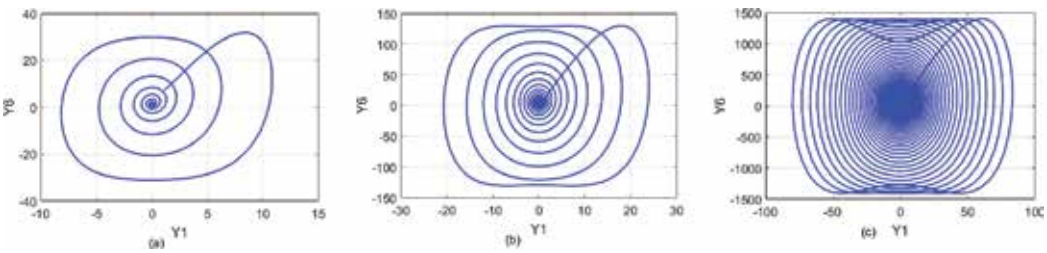

Figure 8. Trajectories of the system of nonlinear equations projected on the (Y_1, Y_6)-plane showing the sensitive dependence of the trajectories on the revised Rayleigh number for (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$ when $Da = 0.05$, $Les = 100$, $\sigma = 0.05$, $\varepsilon = 0.04$, $Q = 10$, $Du = 0.2$, $Sr = 0.3$, $Le = 2$, $N = 25$ and $Rn = 5$.

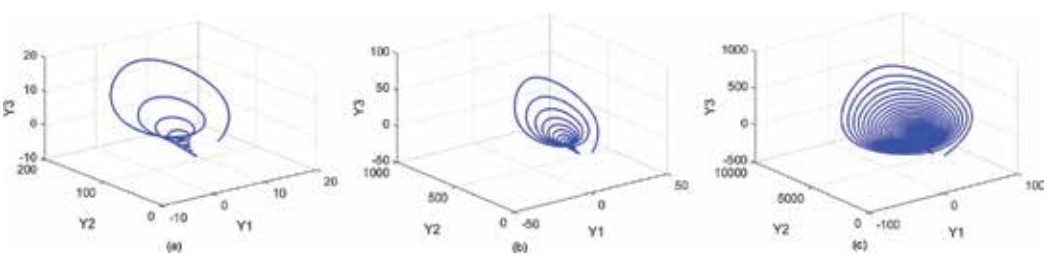

Figure 9. The bifurcations in the three-dimension solution space (Y_1, Y_2, Y_3) for (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$ when $Da = 0.05$, $Les = 100$, $\sigma = 0.05$, $\varepsilon = 0.04$, $Q = 10$, $Du = 0.2$, $Sr = 0.3$, $Le = 2$, $N = 25$ and $Rn = 5$.

Figures 12 and **13** show the streamline, isotherm and isoconcentration contours in the nanofluid flow for different values of the Darcy number and buoyancy ratio. **Figure 12** displays the streamlines for various values of the buoyancy ratio term. Two different eddies are observed. The clockwise and anticlockwise flows are shown via negative and positive stream function values, respectively. The anticlockwise rotating flow occupies the largest area of the nanofluid layer.

For low buoyancy ratio parameters, the flow structure is significantly influenced by the buoyancy within the whole enclosure. Increasing the buoyancy ratio causes the boundary layer thickness to

http://dx.doi.org/10.5772/intechopen.71066

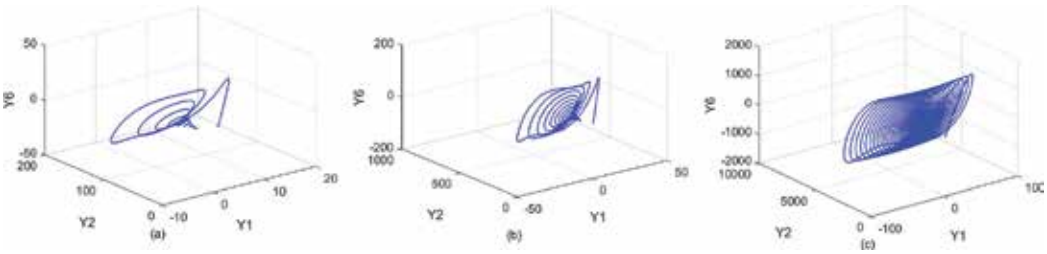

Figure 10. Flow trajectories and bifurcations in the three-dimensional space (Y_1, Y_2, Y_6) for Rayleigh numbers (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$.

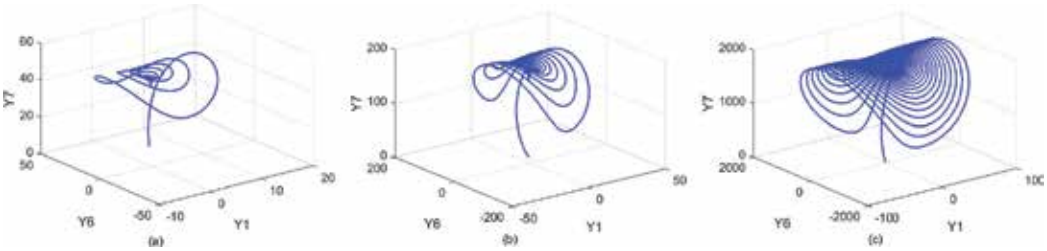

Figure 11. Flow trajectories and bifurcations in the three-dimensional space (Y_1, Y_6, Y_7) for Rayleigh numbers (a) $R = 10^2$, (b) $R = 10^3$ and (c) $R = 10^4$.

become thinner. Also, a high buoyancy ratio changes the flow structure, and this impacts significantly on the concentration field, which builds up a vertical stratification in the enclosure. It is interesting to note that for $N = -25$, the effect of the solutal buoyancy force is in the opposite direction of the thermal buoyancy force. The isothermal and isoconcentration profiles are situated toward the left wall, while for $N = 1$, the thermal and solutal buoyancy forces are equal. For $N = 25$, the effect of solutal buoyancy force is in the same direction as the thermal buoyancy force. In such cases, the isothermal and isoconcentration contours are mostly toward the right wall.

We observe that when $N = -25$, the stream function values in the central eddies increase because the thickness of the boundary layer increases with the buoyancy ratio. The streamlines and the flow behavior are affected by the change in the buoyancy ratio, but the flow pattern remains unaltered. As N decreases from 1 to -25, the streamlines become very dense to the left side of nanofluid layer while when N increases from 1 to 25, the streamlines are less so. The buoyancy forces that drive the nanofluid motion are mainly due to the temperature gradient.

Three different types of eddies are observed for the isoconcentration contours when $N = 25$. Of these, two have a clockwise rotation and one is anticlockwise. It is seen that the small eddy at the right bottom edge is diminished as N decreases from 1 to -25. Here, the concentration boundary layer decreases due to increasing N values, hence the buoyancy ratio has a significant influence on the concentration gradient. As the buoyancy ratio N increases from 1 to 25 the isoconcentrations become very dense at the bottom of nanofluid layer.

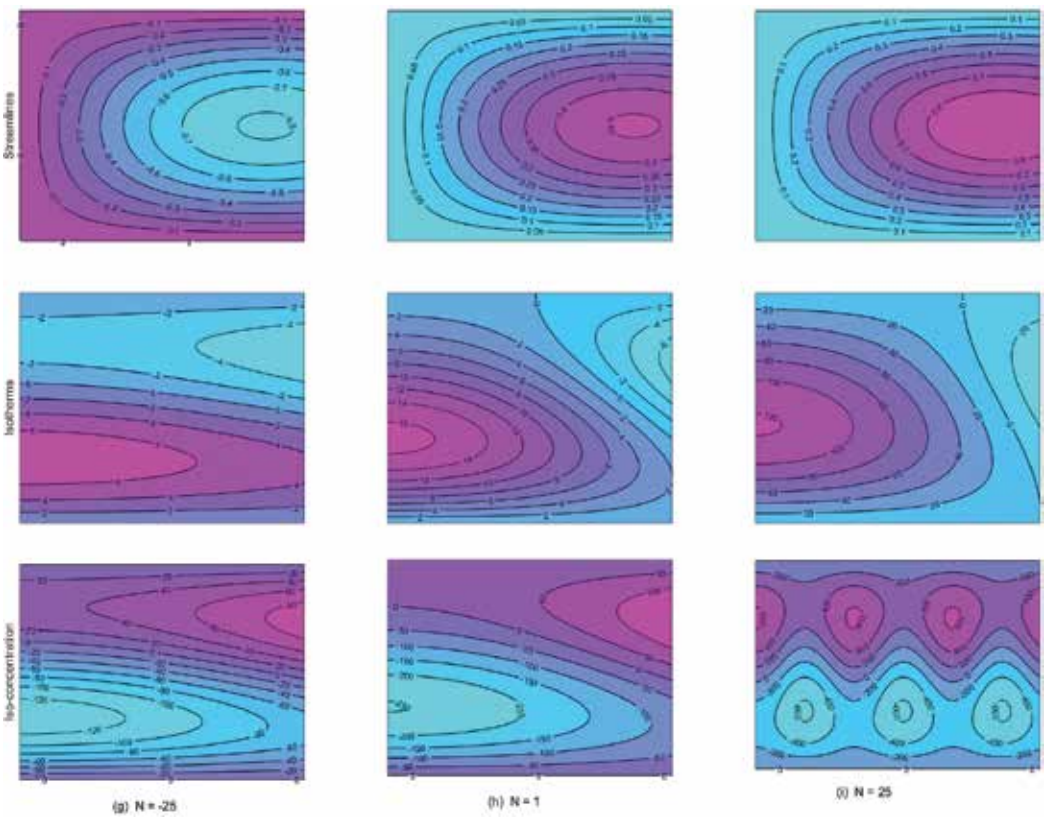

Figure 12. The pattern of streamlines (top), isotherms (middle) and isoconcentration (bottom) for different values of the buoyancy ratio *N*.

The effect of the Darcy number on the nanofluid flow in the porous medium is shown in detail in **Figure 13**. The streamline patterns are similar to those depicted in **Figure 12**. However, as *Da* increases from 0.05 to 0.07, the rotation of the streamlines changes. Similarly, the isotherm patterns change with increasing Darcy numbers. The value of the center eddies increases with increasing *Da*. Increasing *Da* has the effect of increasing the effective fluid viscosity and reducing the thermal and solutal boundary layers.

7. Conclusion

We have investigated the onset of thermal instability in a horizontal porous layer of infinite extent in a cross-diffusive nanofluid flow. The focus of the study has been on stress free boundary conditions with zero nanoparticle flux at the wall. A multidomain spectral collocation method was used to solve the system of nonlinear evolution equations. As the Rayleigh

http://dx.doi.org/10.5772/intechopen.71066

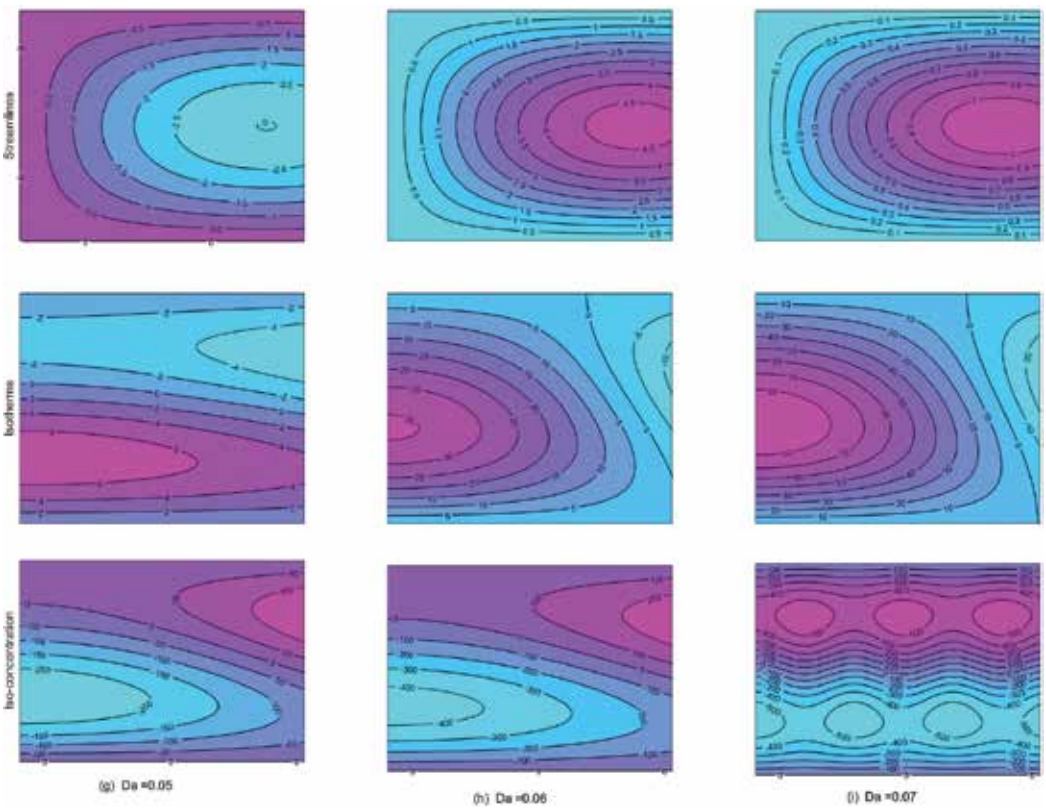

Figure 13. The streamlines (top), isotherms (middle) and isoconcentration (bottom) for different values of the Darcy number *Da*.

number increases to 10^4, the trajectories spiral many times before reaching a fixed point. The nanofluid convection regime is complex for Rayleigh numbers higher than $R = 10^4$, and the flow pattern presents difficulties in interpreting correctly.

Additionally, a change in system parameters, such as an increase in the flow Lewis number, improves the rate of heat and mass transfer in the nanofluid saturated porous media. The Dufour parameter has the effect of increasing heat transfer, while increasing the Soret parameter increases the rate of mass transfer.

Acknowledgements

The authors are grateful to the University of KwaZulu-Natal and the Claude Leon Foundation, South Africa for financial support.

A. Appendix

The terms $a^i_{(j,n)r}$ and R^i_{jr} for $j = 1, 2, \ldots, 7$ in Eq. (38) are given by

$$
\begin{aligned}
a^i_{(1,1)r} &= -\gamma_3 B,\ \ a^i_{(1,2)r} = -\gamma_3,\ \ a^i_{(1,4)r} = -\gamma_3 N,\ \ a^i_{(1,6)r} = \frac{\gamma_3 N_A Rn}{Ra},\\
a^i_{(1,3)r} &= a^i_{(1,5)r} = a^i_{(1,7)r} = 0,\ \ a^i_{(2,1)r} = R - Y^i_{3,r},\quad a^i_{(2,2)r} = -1,\ \ a^i_{(2,3)r} = -Y^i_{1,r},\\
a^i_{(2,4)r} &= -Du,\ \ a^i_{(2,5)r} = a^i_{(2,6)r} = a^i_{(2,7)r} = 0 a^i_{(3,1)r} = -\frac{\alpha\pi}{2} u^i_{2,r},\ \ a^i_{(3,2)r} = \frac{1}{2} Y^i_{1,r},\\
a^i_{(3,3)r} &= -G,\ \ a^i_{(3,5)r} = -GDu a^i_{(3,4)r} = a^i_{(3,6)r} = a^i_{(3,7)r} = 0, a^i_{(4,1)r} = R - Y^i_{5,r},\\
a^i_{(4,2)r} &= -Sr,\ \ a^i_{(5,3)r} = -GSr\\
a^i_{(4,4)r} &= -\frac{1}{Les}, a^i_{(4,5)r} = -Y^i_{1,r},\ \ a^i_{(4,3)r} = a^i_{(4,6)r} = a^i_{(4,7)r} = 0,\ \ a^i_{(5,1)r} = \frac{1}{2} Y^i_{4,r},\\
a^i_{(5,4)r} &= \frac{1}{2} Y^i_{1,r}, a^i_{(5,5)r} = -\frac{G}{Les}, a^i_{(5,2)r} = a^i_{(5,6)r} = a^i_{(5,7)r} = 0, a^i_{(6,1)r} = \gamma_1 - \gamma_2 Y^i_{7,r},\\
a^i_{(6,2)r} &= \frac{\sigma}{Le}, a^i_{(6,6)r} = -\frac{\sigma}{Le},\ \ a^i_{(6,7)r} = -\gamma_2 Y^i_{1,r},\ \ a^i_{(6,3)r} = a^i_{(6,4)r} = a^i_{(6,5)r} = 0,\\
a^i_{(7,1)r} &= \frac{\gamma_2}{2} Y^i_{6,r}, a^i_{(7,3)r} = \frac{G}{Le}, a^i_{(7,6)r} = \frac{\gamma_2}{2} Y^i_{1,r},\ \ a^i_{(7,7)r} = -\frac{G}{Le},\\
a^i_{(7,2)r} &= a^i_{(7,4)r} = a^i_{(7,5)r} = 0,\\
R^i_{1r} &= 0, R^i_{2r} = -Y^i_{1,r} Y^i_{3,r} R^i_{3r} = \frac{1}{2} Y^i_{1,r} Y^i_{2,r},\ \ R^i_{4r} = -Y^i_{1,r} Y^i_{5,r},\ \ R^i_{5r} = \frac{1}{2} Y^i_{1,r} Y^i_{4,r},\\
R^i_{6r} &= -\frac{\sigma N_A}{\varepsilon} Y^i_{1,r} Y^i_{7,r},\ \ R^i_{7r} = \frac{\sigma N_A}{2\varepsilon} Y^i_{1,r} Y^i_{6,r},
\end{aligned}
$$

where $\gamma_1 = \frac{\sigma R}{\varepsilon}$, $\gamma_2 = \frac{N_A \sigma}{\varepsilon}$ and $\gamma_3 = \frac{\mathrm{Pr}}{Da}$.

B. Appendix

The matrices A_{ij} in Eq. (42) are given by

$$
\begin{aligned}
&A_{nn} = \mathbf{diag}\left(a_{n,n}) r^i\right) - \mathbf{D},\ \ A_{12} = \mathbf{diag}\left(a^i_{(1,2)r}\right),\ \ A_{16} = \mathbf{diag}\left(a^i_{(1,6)r}\right),\\
&A_{13} = A_{14} = A_{15} = A_{17} = \mathbf{O},\\
&A_{21} = \mathbf{diag}\left(a^i_{(2,1)r}\right),\quad A_{23} = \mathbf{diag}\left(a^i_{(2,3)r}\right),\quad A_{24} = \mathbf{diag}\left(a^i_{(2,4)r}\right),\quad A_{25} = A_{26} = A_{27} = \mathbf{O},\\
&A_{31} = \mathbf{diag}\left(a^i_{(3,1)r}\right),\quad A_{32} = \mathbf{diag}\left(a^i_{(3,2)r}\right),\quad A_{35} = \mathbf{diag}\left(a^i_{(3,5)r}\right),\quad A_{34} = A_{36} = A_{37} = \mathbf{O},\\
&A_{41} = \mathbf{diag}\left(a^i_{(4,1)r}\right),\quad A_{42} = \mathbf{diag}\left(a^i_{(4,2)r}\right),\quad A_{45} = \mathbf{diag}\left(a^i_{(4,5)r}\right),\quad A_{43} = A_{46} = A_{47} = \mathbf{O},\\
&A_{51} = \mathbf{diag}\left(a^i_{(5,1)r}\right),\quad A_{53} = \mathbf{diag}\left(a^i_{(5,3)r}\right),\quad A_{54} = \mathbf{diag}\left(a^i_{(5,4)r}\right),\quad A_{52} = A_{56} = A_{57} = \mathbf{O},\\
&A_{61} = \mathbf{diag}\left(a^i_{(6,1)r}\right),\quad A_{62} = \mathbf{diag}\left(a^i_{(6,2)r}\right),\quad A_{67} = \mathbf{diag}\left(a^i_{(6,7)r}\right),\quad A_{63} = A_{64} = A_{65} = \mathbf{O},
\end{aligned}
$$

$$A_{71} = \mathbf{diag}\left(a^i_{(7,1)r}\right), \quad A_{73} = \mathbf{diag}\left(a^i_{(7,3)r}\right), \quad A_{76} = \mathbf{diag}\left(a^i_{(7,6)r}\right), \quad A_{72} = A_{74} = A_{75} = \mathbf{O},$$

where $\mathbf{O}$ is an $(N+1) \times (N+1)$ matrix of zeros and **diag** is an $(N+1) \times (N+1)$ diagonal matrix.

Author details

Osman A.I. Noreldin[1], Precious Sibanda[1]* and Sabyasachi Mondal[2]

*Address all correspondence to: sibandap@ukzn.ac.za

1 School of Mathematical Sciences, University of KwaZulu-Natal, South Africa

2 Department of Mathematics, Amity University, Kolkata, West Bengal, India

References

[1] Choi SU, Eastman JA. Enhancing Thermal Conductivity of Fluids with Nanoparticles, Technical Report. IL, United States: Argonne National Lab; 1995

[2] Buongiorno J. Convective transport in nanofluids. Journal of Heat Transfer. 2006;**128**(3): 240-250

[3] Masuda H, Ebata A, Teramae K. Alteration of thermal conductivity and viscosity of liquid by dispersing ultra-fine particles. Netsu Bussei. 1993;**4**:227-233

[4] Tzou D. Thermal instability of nanofluids in natural convection. International Journal of Heat and Mass Transfer. 2008;**51**(11):2967-2979

[5] Narayana M, Sibanda P, Motsa S, Lakshmi-Narayana P. Linear and nonlinear stability analysis of binary maxwell fluid convection in a porous medium. Heat and Mass Transfer. 2012;**48**(5):863-874

[6] Narayana M, Sibanda P, Siddheshwar PG, Jayalatha G. Linear and nonlinear stability analysis of binary viscoelastic fluid convection. Applied Mathematical Modelling. 2013;**37**(16): 8162-8178

[7] Yadav D, Bhargava R, Agrawal G. Numerical solution of a thermal instability problem in a rotating nanofluid layer. International Journal of Heat and Mass Transfer. 2013;**63**:313-322

[8] Kuznetsov A, Nield D. Effect of local thermal non-equilibrium on the onset of convection in a porous medium layer saturated by a nanofluid. Transport in Porous Media. 2010;**83**(2): 425-436

[9] Kuznetsov A, Nield D. Thermal instability in a porous medium layer saturated by a nanofluid: Brinkman model. Transport in Porous Media. 2010;**81**(3):409-422

[10] Nield D, Kuznetsov A. The onset of double-diffusive convection in a nanofluid layer. International Journal of Heat and Fluid Flow. 2011;**32**(4):771-776

[11] Nield D, Kuznetsov A. Thermal instability in a porous medium layer saturated by a nanofluid: A revised model. International Journal of Heat and Mass Transfer. 2014;**68**:211-214

[12] Bhadauria B. Double diffusive convection in a porous medium with modulated temperature on the boundaries. Transport in Porous Media. 2007;**70**(2):191-211

[13] Bhadauria B, Siddheshwar P, Kumar J, Suthar OP. Weakly nonlinear stability analysis of temperature/gravity-modulated stationary rayleigh–bénard convection in a rotating porous medium. Transport in Porous Media. 2012;**92**(3):633-647

[14] Gresho P, Sani R. The effects of gravity modulation on the stability of a heated fluid layer. Journal of Fluid Mechanics. 1970;**40**(4):783-806

[15] Kiran P, Bhadauria B. Chaotic convection in a porous medium under temperature modulation. Transport in Porous Media. 2015;**107**(3):745-763

[16] Chandrasekhar S. Hydrodynamic and Hydromagnetic Stability. Oxford University Press; 1961

[17] Thompson W. Cxliii. Thermal convection in a magnetic field. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science. 1951;**42**(335):1417-1432

[18] Heris SZ, Salehi H, Noie S. The effect of magnetic field and nanofluid on thermal performance of two-phase closed thermosyphon (tpct). International Journal of Physical Sciences. 2012;**7**(4):534-543

[19] Ghasemi B, Aminossadati S, Raisi A. Magnetic field effect on natural convection in a nanofluid-filled square enclosure. International Journal of Thermal Sciences. 2011;**50**(9): 1748-1756

[20] Hamad M, Pop I, Ismail AM. Magnetic field effects on free convection flow of a nanofluid past a vertical semi-infinite flat plate. Nonlinear Analysis: Real World Applications. 2011;**12**(3):1338-1346

[21] Bergman MI, Fearn DR. Chimneys on the earth's inner-outer core boundary? Geophysical Research Letters. 1994;**21**(6):477-480

[22] Furmanski P, Banaszek J. Modelling of the mushy zone permeability for solidification of binary alloys. Materials Science Forum. Trans Tech Publ. 2006;**508**:411-418

[23] Gupta U, Ahuja J, Wanchoo R. Rayleigh-bénard convection of nanofluids with magnetic field and permeability effects. Procedia Engineering. 2015;**127**:325-332

[24] Imomnazarov KK. Modified darcy laws for conducting porous media. Mathematical and Computer Modelling. 2004;**40**(1–2):5-10

[25] Rana G, Kango S, Chadha K. Magneto-thermal convection in Walters' (model b') elastico-viscous fluid saturated by a Darcy-Brinkman porous medium. Engineering Mechanics. 2014;**21**:425-435

[26] Motsa S, Dlamini P, Khumalo M. A new multistage spectral relaxation method for solving chaotic initial value systems. Nonlinear Dynamics. 2013;**72**(1–2):265-283

Chapter 10

Small-Angle Scattering from Mass and Surface Fractals

Eugen Mircea Anitas

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.70870

Abstract

The concepts of mass and surface fractals are introduced, and the corresponding small-angle scattering (SAS; X-rays, neutrons) intensities are computed. It is shown how to resolve the fractal structure of various complex systems from experimental scattering measurements, and how obtained data are related to specific features of the fractal models. We present and discuss various mass and surface fractal structures, including fractals generated from iterated function systems and cellular automata. In addition to the fractal dimension and the overall fractal size, the suggested analysis allows us to obtain the iteration number, the number of basic units which form the fractal and the scaling factor.

Keywords: small-angle scattering, form factor, structure factor, fractals, iterated function system, cellular automata

1. Introduction

A great number of natural systems provide us with examples of nano- and micro- structures, which appear similar under a change of scale. These structures are called fractals [1], and can be observed in various disordered materials, rough surfaces, aggregates, metals, polymers, gels, colloids, thin films, etc. Quite often, the physical properties (mechanical, optical, statistical, thermodynamical, etc.) depend on their spatial configurations, and a great deal of activity has been performed in elucidating such correlations [2–4]. To this aim, in the last decade, important steps have been performed in three directions: instrumentation [5–7], computer programs [8, 9], and the development of new methods for sample preparation [10–14]. The first two have been proved to be very useful for the physical execution and data analysis, and the third one for preparation of nano- and micro-materials with the pre-defined structure and functions.

These achievements have stimulated a great interest in the development of theoretical investigations for structural modeling of self-similar objects at nano- and micro-scale. In particular, models based on deterministic or exact self-similar fractals (i.e., fractals that are self-similar at every point, such as the Koch snowflake, Cantor set, or Mandelbrot cube) have been frequently

used, since this type of fractals allows an analytical representation of various geometrical parameters (radius of gyration) or of the scattering intensity spectrum. Although for most fractals generated by natural processes, this is only an approximation, in the case of deterministic nano- and micro-materials obtained recently such as 2*D* Sierpinski gaskets [15] and Cantor sets [16], or 3*D* Menger sponge [17] and octahedral structures [18], this approximation becomes exact.

Small-angle scattering of X-rays (SAXS) and/or neutrons (SANS) are well established techniques for probing the nano/micro scale structure in disordered materials [19–21]. While in the case of X-rays, the scattering is mostly determined by the interaction of the incident radiation with electrons, in the case of neutrons, the scattering is determined by their interaction with the atomic nuclei and with the magnetic moments in magnetic materials. Since the wavelengths of X-rays (0.5–2 Å) are of the same order of magnitude as those of thermal neutrons (1–10 Å), often, the data analysis and interpretation procedures for SANS can be interchanged with SAXS, and the developed theoretical models can be applied, generally, to both techniques [22]. However, using neutrons is very important in studying magnetic properties of materials as well as in emphasizing or concealing certain features of the investigated sample [23]. In the later case, the possibility of wide variations in the neutron scattering lengths (which can be negative sometimes) is exploited, and this is a unique feature of SANS, which makes it a preferred method over SAXS in structural analysis of biological materials [24, 25].

As compared with other methods of structural investigations, SAXS/SANS have the advantage that they are noninvasive, the physical quantities of interest (specific surface, radius of gyration, volume, or the fractal dimension) are averaged over a macroscopic volume and they can be extracted with almost no approximation [26]. In particular, for self-similar objects (either exact or statistical), the most important advantage is that SAXS/SANS can distinguish between mass [27] and surface fractals [28]. Experimentally, the difference is revealed through the value of the scattering exponent τ in the region where the scattering intensity $I(q)$ decays as a power-law, i.e., $I(q) \propto q^{-\tau}$, where $q = (4\pi/\lambda)\sin\theta$ is the scattering vector, λ is the wavelength of the incident radiation, and 2θ is the scattering angle. For mass fractals $\tau = D_m$, where D_m is the mass fractal dimension with $0 < D_m < 3$. For surface fractals $\tau = 6 - D_s$, where D_s is the surface fractal dimension with $2 < D_s < 3$. Thus, in practice if the absolute value of the measured scattering exponent is smaller than 3, the sample is a mass fractal with fractal dimension τ (in the measured q-range), and if the exponent is between 3 and 4, the sample is a surface fractal with fractal dimension $6 - \tau$.

Besides the fractal dimension (either mass or surface one), traditionally from SAXS/SANS patterns, we can also obtain the overall size of the fractal, as well as the size of the basic structural units composing the fractal. The last years have brought a breakthrough in the theoretical analysis of SAXS/SANS experimental data, allowing for the extraction of additional structural information and detailed modeling of fractals using a deterministic approach [29–40]. This progress was stimulated by recent advances in nanotechnology, which allows preparation of both mass and surface deterministic fractals at sub-micrometer scale [15–18, 41, 42], as well as by instrumentation which allows novel structural features to be recorded in experimental data [43].

This chapter focuses on the interpretation of SAXS/SANS data from deterministic mass and surface fractals. First, a brief theoretical background on the basics of SAS theory, and on description of mass and surface fractals, is presented. Here, we also include the theory of some well-known methods of generating fractals, such as iterated function system of cellular automata. Novel data analysis methods for extracting additional structural information are presented and illustrated by applications to various models of mass and surface fractals.

2. Theoretical background

In this section, some important concepts for the analysis of SAXS/SANS are reviewed, and analytical and numerical procedures for calculating the scattering intensity from some basic geometrical shapes are described. As we will see in the next section, these geometrical shapes will form the "scattering units" of the fractals. Then, the basic notions of fractal theory including fractal dimension, mass, and surface fractals are presented and defined in a rigorous manner, and two general methods for generating fractal structures are presented. These concepts are then applied in calculating the SAXS/SANS patterns from several theoretical models of mass and surface fractals.

2.1. Small-angle scattering

2.1.1. General remarks

In a SAS experiment, a beam of X-rays or neutrons is emitted from a source and strikes the sample. A small fraction is scattered by the sample and is recorded by the detector. In **Figure 1,** the incident beam has a wave vector k_i, and the scattered beam with the wave vector k_s makes the angle 2θ with the direction of the incoming or transmitted beam.

To describe the scattering from assemblies of objects with scattering length b_j, we write the scattering length density *SLD* as $\rho(\boldsymbol{r}) = \sum_j b_j(\boldsymbol{r} - \boldsymbol{r}_j)$ [21], where $\vec{r}_j$ is the object positions. Then, the total scattering amplitude is defined by the Fourier transform of $\rho\left(\vec{r}\right)$:

$$A(\boldsymbol{q}) \equiv \int_v \rho(\boldsymbol{r}) e^{-iq.r} \mathrm{d}^3 r, \tag{1}$$

where v is the total volume irradiated by the beam. In the following, we consider scattering occurring in a particulate system where particles of density ρ_m are dispersed in a uniform solid matrix of density ρ_p. Then, the excess scattering *SLD* is defined by $\Delta\rho = \rho_m - \rho_p$. We also consider that the objects are fractals that are randomly distributed and with uncorrelated positions and orientations. Thus, the scattering intensities of each object are added, and the intensity from the entire sample can be obtained from a single object averaged over all orientations, according to [29]:

$$I(q) = n\left|\Delta\rho\right|^2 V^2 \left\langle |F(\boldsymbol{q})|^2 \right\rangle, \tag{2}$$

where n is the concentration of objects, V is the volume of each object, and $F\left(\vec{q}\right)$ is the normalized scattering amplitude given by:

$$F(\boldsymbol{q}) = \frac{1}{V}\int_V e^{-iq.r}\mathrm{d}^3 r. \tag{3}$$

The symbol $\langle\cdots\rangle$ stands for ensemble averaging over all orientations, and for an arbitrary function f, it is calculated according to:

$$\left\langle f\left(q_x, q_y, q_z\right)\right\rangle = \frac{1}{4\pi}\int_0^{\pi} \mathrm{d}\theta \sin\theta \int_0^{2\pi} \mathrm{d}\phi f(q,\theta,\phi), \tag{4}$$

where, in spherical coordinates $q_x = q\cos\phi\sin\theta$, $q_y = q\sin\phi\sin\theta$, $q_z = q\cos\theta$.

Another useful form of the scattering intensity, as a function of the correlation function, is the following [21]:

$$I(q) = 4\pi\int_0^D \gamma(r)\frac{\sin qr}{qr} r^2 \mathrm{d}r, \tag{5}$$

where $\gamma(r) \equiv \langle \rho(\boldsymbol{r}) * \rho(-\boldsymbol{r})\rangle$ is the correlation function of the object, with $\gamma(r) = 0$ for $r > D$ and D is the largest dimension in the object. The symbol $*$ denotes a convolution, and thus, the correlation function can be seen as an averaged self-convolution of density distribution.

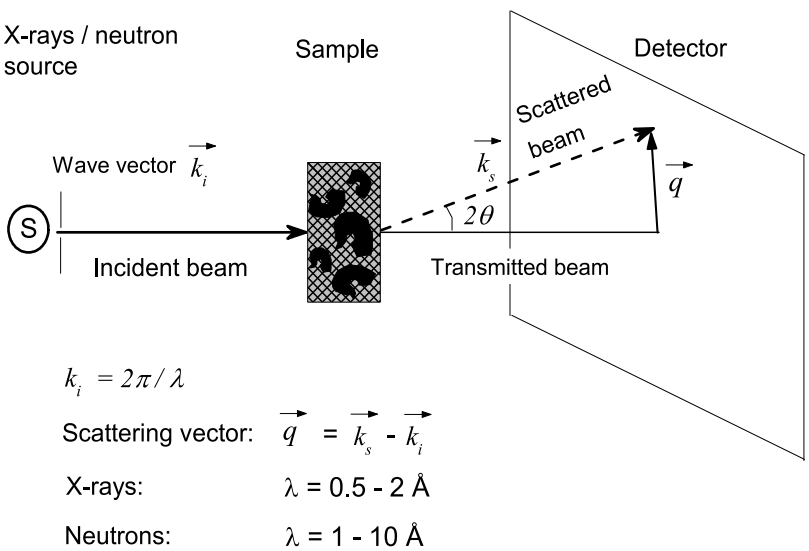

Figure 1. Schematic representation of a small-angle scattering experiment. 2θ is the scattering angle and λ is the wavelength of incident beam. The sample shown is a two-phase system of polydisperse scatterers with the same shape and random orientations, embedded in a matrix or solution.

At low values of the scattering vector ($q \lesssim 2\pi/D$), the above expression can be further exploited. By considering first the MacLaurin series

$$\frac{\sin(qr)}{qr} \simeq 1 - \frac{q^2r^2}{6} + \frac{q^4r^4}{120} - \cdots, \tag{6}$$

and then using only the first two terms of this approximation into Eq. (5), the Guinier equation is obtained:

$$I(q) = I(0)\left(1 - \frac{q^2R_g^2}{3} + \cdots\right), \tag{7}$$

where $I(0) = 4\pi \int_0^D \gamma(r) r^2 \mathrm{d}r$, and

$$R_g = \frac{1}{2}\int_0^D \gamma(r) r^4 \mathrm{d}r \Big/ \int_0^D \gamma(r) r^2 \mathrm{d}r, \tag{8}$$

is the radius of gyration of the object. In practice, a plot of $\log I(q)$ vs. q^2 is used to obtain the slope $R_g^2/3$, and then an overall size of the object. For a ball of radius R, it is known that $R = R_g\sqrt{5/3}$ [21]. If a scattering experiment is performed on an absolute scale, molecular weight can also be obtained.

However, as we shall see in the next sections, inside a fractal, the scattering units have defined positions and correlations, and the interference among rays scattered by different units may no longer be ignored, so that the scattering amplitudes of individual units have to be added. By considering that the fractal is composed of N balls of size R, its form factor becomes [29]:

$$F(\boldsymbol{q}) = \rho_q F_0(\boldsymbol{q}R)/N, \tag{9}$$

where $\rho_{\boldsymbol{q}} = \sum_j e^{-i\mathbf{q}\cdot\mathbf{r}_j}$ is the Fourier component of the density of units centers, F_0 is the form factor of each scattering unit composing the fractal, and $\vec{r}_j$ are their positions. By introducing Eq. (9) into Eq. (2), the scattering intensity can be written as [29]:

$$I(q) = I(0)S(q)|F_0(qR)|^2/N, \tag{10}$$

where $I(0) = n|\Delta\rho|^2V^2$, and $S(q) \equiv \left\langle \rho_{\boldsymbol{q}}\rho_{-\mathbf{q}} \right\rangle/N$ is the structure factor and it describes the correlations between the scattering units inside the fractal.

A physical sample almost always consists of fractals that have different sizes, which is called polydispersity. An exception to this rule is protein solutions, in which all have the same size and shape. Thus, the corresponding scattering intensity from polydisperse fractals can be regarded as the sum of each individual form factor weighted with the corresponding volume V and contrast $\Delta\rho$. We consider here a continuous distribution $D_N(l)$ of fractals with different sizes l, defined in such a way that $D_N(l)dl$ gives the probability of finding a fractal with

dimension l lying in the range $(l, l+dl)$. Although any kind of broad distribution can be used, we take here, as an application, a log-normal distribution of fractal sizes, such as:

$$D_N(l) = \frac{1}{\sigma l (2\pi)^{1/2}} e^{-\frac{\left(\log\left(l/\mu_0\right)+\sigma^2/2\right)^2}{2\sigma^2}}, \tag{11}$$

where $\sigma = \left(\log\left(1+\sigma_r^2\right)\right)^{1/2}$, $\mu_0 = \langle l \rangle_D$ is the mean length, $\sigma_r = \left(\left\langle l^2 \right\rangle - \mu_0^2\right)^{1/2}/\mu_0$ is the relative variance, and $\langle \cdots \rangle_D = \int_0^\infty \cdots D_N(l) \mathrm{d}l$. Since for a polydisperse fractal dispersion, the volume of each fractal has a continuous variation with its size, the polydisperse scattering intensity becomes:

$$I(q) = n\left|\Delta\rho\right|^2 \int_0^\infty \left\langle \left|F(\boldsymbol{q})\right|^2 \right\rangle V^2(l) D_N(l) \mathrm{d}l, \tag{12}$$

where the form factor $F(\mathbf{q})$ is given by Eq. (9). The effect of polydispersity is to smooth the scattering curves [20, 21] (see also **Figure 2a**).

2.1.2. Debye-Pantos formula

In the next section, we shall make use of chaos game representation (CGR) and cellular automata (CA) to generate positions of the N scattering units/points. Thus, we can start with the Debye formula [44]

$$I^D(q) = NI_s(q) + 2F_s(q)^2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{\sin qr_{ij}}{qr_{ij}}, \tag{13}$$

where $I_s(q)$ is the intensity scattered by each fractal unit, and r_{ij} is the distance between units i and j. When the number of units exceeds few thousands, the computation of the term $\sin(qr_{ij})/(qr_{ij})$ is time consuming, and thus it is handled via a pair-distance histogram $g(r)$, with a bin-width commensurate with the experimental resolution [45]. Thus, Eq. (13) becomes

$$I^D(q) = NI_s(q) + 2F_s^2(q) \sum_{i=1}^{N_{bins}} g(r_i) \frac{\sin qr_i}{qr_i}, \tag{14}$$

where $g(r_i)$ is the pair-distance histogram at pair distance r_i. For determining fractal properties, we can neglect the form factor, and consider $I_s(q) = F_s^2(q) = 1$. Thus, Eq. (14) gives the structure factor:

$$I^D(q) \equiv S^D(q) = N + 2 \sum_{i=1}^{N_{bins}} g(r_i) \frac{\sin qr_i}{qr_i}. \tag{15}$$

2.1.3. Scattering from a ball and from a triangle

As a first example, we derive the scattering intensity of a ball with unit density, radius R, and volume $V = (4\pi/3)R^3$. To do so, we can rewrite the normalized scattering amplitude given by Eq. (3), in spherical coordinates, such as:

$$F_0(\boldsymbol{q}) = \frac{1}{V}\int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{r=0}^{\infty} e^{-i\mathbf{q}\cdot\mathbf{r}} r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi. \tag{16}$$

Note that since balls represent here the basic units of the fractal, we have chosen the notation $F_0(\boldsymbol{q})$ instead of $F(\boldsymbol{q})$, in spirit of Eq. (9). We can choose the polar axis to coincide with the direction of $\boldsymbol{q}$, and therefore $\mathbf{q}\cdot\mathbf{r} = qr\cos\theta$. By denoting $u = \cos\theta$, in the new variable, Eq. (16) becomes:

$$F_0(\boldsymbol{q}) \equiv \frac{1}{V}\int_{\phi=0}^{2\pi}\int_{u=-1}^{1}\int_{r=0}^{\infty} e^{-iqru} r^2 \sin\theta \mathrm{d}r\mathrm{d}\theta\mathrm{d}\phi = \frac{1}{V}\int_{r=0}^{R} 4\pi r^2 \frac{\sin(qr)}{qr}\mathrm{d}r. \tag{17}$$

By performing an integration by parts of the last expression, the normalized scattering amplitude of the ball of radius R becomes:

$$F_0(qR) = \frac{3(\sin(qR) - (qR)\cos(qR))}{(qR)^3}. \tag{18}$$

Thus, the total scattering intensity (see also Eq. (2)) becomes:

$$I(q) \equiv [F_0(qR)]^2 = \frac{3^2(\sin(qR) - (qR)\cos(qR))^2}{(qR)^6}. \tag{19}$$

Figure 2a shows the scattering intensity of a ball of radius $R = 10$ nm. The scattering is represented on a double logarithmic scale and shows the presence of two main regions. At low values of the scattering vector ($q \lesssim \pi/R$), we have the Guinier region, which is a plateau with $I(q) \propto q^0$, and from which one can obtain the radius of gyration, as described in the previous section. At higher values (i.e., $q \gtrsim \pi/R$), there is a power-law decay of the type $I(q) \propto q^{-4}$ and with many minima. This is called the Porod region and generally it gives information about the specific surface of the investigated object. The main feature here is that by increasing the relative variance σ_r, the scattering curve

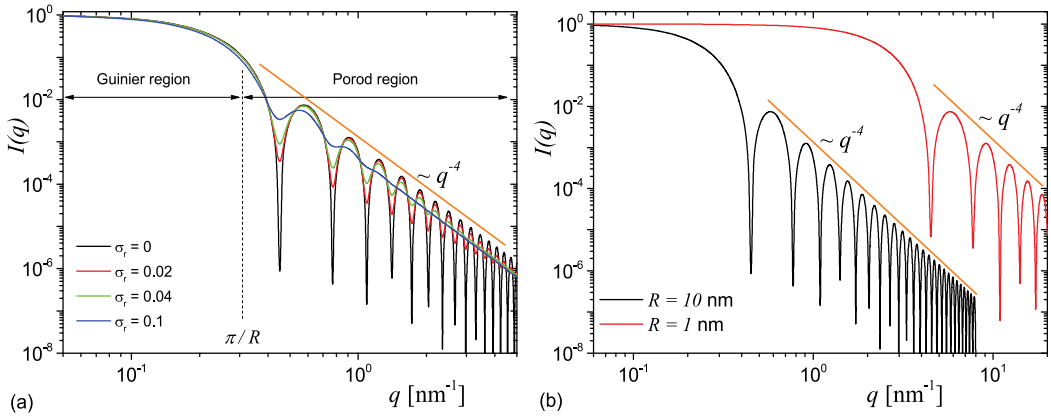

Figure 2. (a) SAS intensity from a ball of radius $R = 10$ nm from Eq. (19) (lowest curve). The higher the relative variance, the smoother the curve. Polydisperse SAS intensity, according to Eq. (12) at various values of relative variances. (b) SAS intensities from a ball of radius $R = 10$ nm (left curve), and $R = 1$ nm (right curve).

becomes smoother, and the value of the scattering exponent is preserved. **Figure 2b** shows that by decreasing the size of the ball, the corresponding scattering curve has the same characteristics and the Porod region is shifted to the right with the corresponding factor (here by 10).

The second example is an equilateral triangle of edge size a. This is a slightly more complicated structure since it does not have a center of symmetry as a ball, and thus an orientational averaging is required. Thus, its height is $h = a\sqrt{3}/2$ and its surface area is:

$$Area(a) = a^2 \frac{\sqrt{3}}{4}. \tag{20}$$

We choose a Cartesian coordinate system where one edge is parallel to the x-axis and the opposite vertex coincides with the origin. Thus, Eq. (3) becomes a surface integral given by:

$$F_0(\boldsymbol{q}) = \frac{1}{Area(a)} \int_0^a \mathrm{d}y \int_{-\frac{ya}{2h}}^{\frac{ya}{2h}} \mathrm{d}x e^{-i\left(xq_x + yq_y\right)}, \tag{21}$$

which can be calculated and transformed into:

$$F_0(\boldsymbol{q}) = 2e^{-i\alpha} \frac{\beta e^{i\alpha} - \beta\cos\beta - i\alpha\sin\beta}{\beta\left(\beta^2 - \alpha^2\right)}, \tag{22}$$

where $\alpha = hq_x$ and $\beta = hq_y$. As we shall see in next sections, the scattering amplitude of a system of triangles can be obtained by properly taking into account their scaling, rotation, and translations.

The averaging over all orientations is performed by allowing the triangle to rotate in a $2D$ space, and thus the average given by Eq. (4) for $3D$ case, becomes now:

$$\left\langle f\left(q_x, q_y\right)\right\rangle = \frac{1}{2\pi} \int_0^{2\pi} \mathrm{d}\phi f(q, \phi), \tag{23}$$

with $q_x = q\cos\phi$ and $q_y = q\sin\phi$.

Similarly to scattering from a ball, the intensity curve of a triangle also shows the Guinier region at low-q ($q \lesssim 2\pi/a$), and a Porod region at high-q ($q \gtrsim 2\pi/a$) as shown in **Figure 3a**. However, for a triangle, the absolute value of the scattering exponent in the Porod region is equal to 3. This is in contrast to the value of 4 obtained for the ball in the previous example (see **Figure 3b**). The difference arises due to the fact that the triangle is a $2D$ object while the ball is a $3D$ one. In addition, due to the lack of symmetry, scattering from a triangle does not show pronounced minima as in the case of scattering from a ball.

2.1.4. Scattering from systems of triangles

In the previous section**,** we have seen that regardless of the shape and Euclidean dimension, the SAS intensity from basic geometrical structures always reproduces a Guinier region followed by a Porod one. Without loosing from generality, we will restrict in the following to calculate SAS intensity from systems of triangles. In principle, any geometrical shape can be chosen but we

http://dx.doi.org/10.5772/intechopen.70870

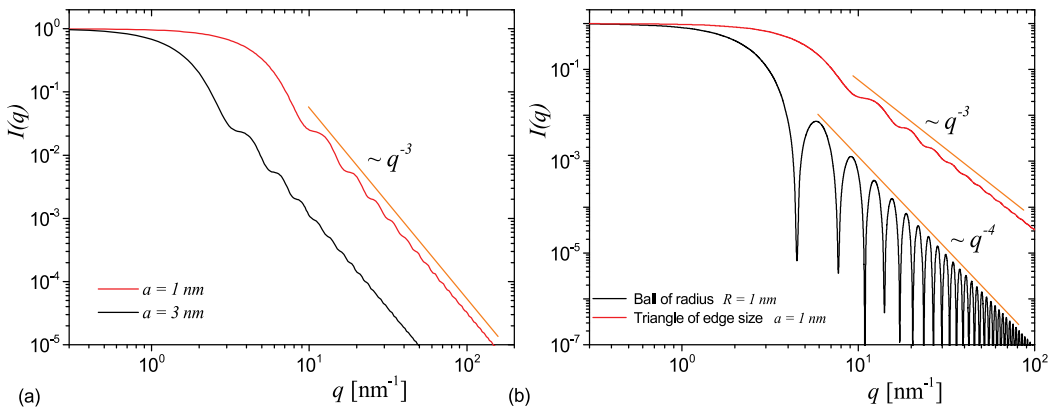

Figure 3. (a) SAS intensity from an equilateral triangle of edge size a = 1 nm (right curve), and a = 3 nm (left curve), respectively. (b) A comparison between the SAS intensity of a ball of radius R = 1 nm (left curve) and a triangle of edge size a = 1 nm (right curve).

prefer here triangles due to the fact that the both well-known techniques for generating fractal structures: iteration function systems and cellular automata, in their basic form, involve triangles in the construction process.

We start first with a simple model consisting of three triangles of edge size $a/2$, with a = 1 nm, as shown in **Figure 4a**. For this configuration, we can write [40]:

$$Area(a/3)F(\boldsymbol{q}) = \sum_{j=0}^{2} \beta_s^2 Area(a) F_0(\boldsymbol{\beta}_s \mathbf{q}) e^{-iqa_j}, \tag{24}$$

where β_s is the scaling factor, $F_0(\boldsymbol{q})$ is given by Eq. (22), and the translation vectors are given by:

$$\boldsymbol{a}_j = \frac{a\sqrt{3}}{6}\left\{\cos\frac{\pi}{3}\left(2j+\frac{3}{2}\right), \sin\frac{\pi}{3}\left(2j+\frac{3}{2}\right)\right\} \tag{25}$$

The corresponding scattering intensity is shown in **Figure 4b**, and as expected, it consists of a Guinier region followed by a Porod one with scattering exponent −3. For comparison, the same figure shows the scattering intensity of a single of edge size a = 1 nm.

The second example is a system of 6 triangles of edge sizes a/3 arranged in such a way that they form a hexagon as shown in **Figure 5a** in black. For this configuration, the translation vectors can be written as [38]:

$$\boldsymbol{b}_j = \frac{2a}{3\sqrt{3}}\left\{\cos\frac{\pi}{3}j, \sin\frac{\pi}{3}j\right\}. \tag{26}$$

The corresponding scattering intensity is shown in **Figure 5b** and shows a superposition of maxima and minima. Excepting the distribution of minima in the Porod region, there is no significant difference between this scattering curve and the one corresponding to the system of three triangles shown in **Figure 4b**.

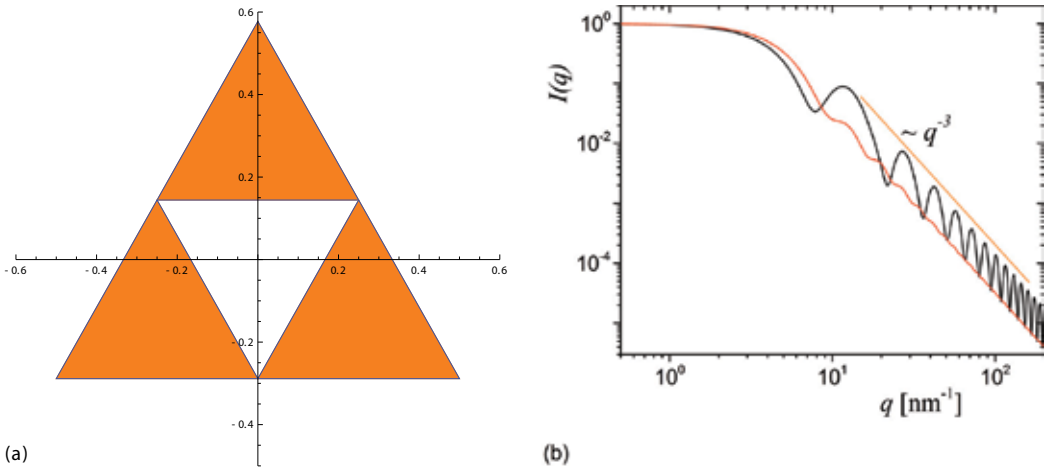

Figure 4. (a) A model of three triangles (gray) of edge size $a/2$, with $a = 1$ nm; (b) the corresponding SAS intensity (highly oscillating curve). The smoother curve is the SAS intensity corresponding to a single triangle of edge size $a = 1$ nm and whose center coincides with the center of white triangle in part (a).

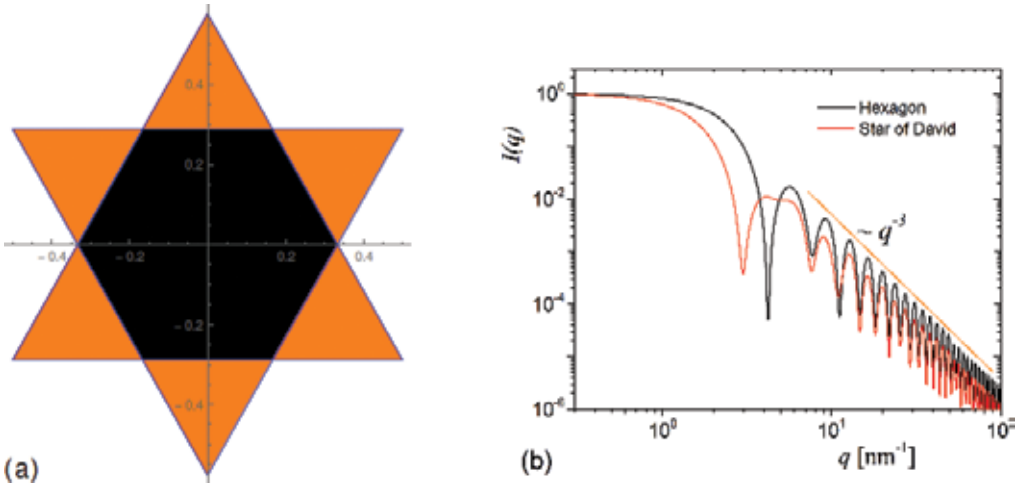

Figure 5. (a) Two models consisting only of equilateral triangles: hexagon (black color) and Star of David (black and gray colors); (b) the corresponding SAS intensities of the hexagon (continuous curve) and Star of David (dashed curve).

The last example is slightly more complicated and it consists of one hexagon of edge size $a/3$ (black) and six triangles of edge size $a/3$ (gray), with a = 1 nm arranged as in **Figure 5a**. This configuration is equivalent to a system of one triangle of edge size a and three triangles of edge size $a/3$. This is known in the literature, also as the Star of David. The translation vectors of the three triangles can be obtained in a similar way, as in the case of the first example. The corresponding scattering intensity is shown in **Figure 5b** (red). It can be seen that due to various sizes of the triangles involved in the construction, an intermediate regime emerges between the Guinier and Porod regions, approximately at $2 \lesssim q \lesssim 8$ nm^{-1}. As we shall see later, for even more complex structures, this intermediate region will evolve into a fractal one.

In particular, if the triangles composing the system are arranged in such a way that they form a Sierpinski gasket, the intermediate region will correspond to a mass fractal region [29–31]. When the triangles have a power-law distribution in their sizes, the intermediate region will correspond to a surface fractal one [37, 38]. The great advantage of the SAS technique consists in the possibility to differentiate these two types of fractal regimes, as discussed at the beginning of this chapter.

2.2. Mass and surface fractals

As it was already pointed out before, the main characteristic of fractals obtained from a SAS experiment is the fractal dimension. Mathematically, the α-dimensional Hausdorff measure is defined by [29]

$$m^{\alpha}(A) = \lim_{a \to 0} \inf_{\{V_i\}} \sum_i a_i^{\alpha}, \alpha > 0 \tag{27}$$

where A is a subset of an n-dimensional Euclidean space, $\{V_i\}$ is a covering of A with $a_i = diam(V_i) \leq a$, and the infimum is on all possible coverings. Then, the Hausdorff dimension D of the set A is given by:

$$D \equiv \inf\{\alpha_i, m^{\alpha}(A) = 0\} = \sup\{\alpha, m^{\alpha}(A) = \infty\}, \tag{28}$$

and it represents the value of α for which the Hausdorff measure changes its value from zero to infinity.

However, in practice, this definition is quite inconvenient to be used, and here we shall use the "mass-radius" relation for calculating the Hausdorff dimension of the fractals [1]. In this approach, the total fractal measure (i.e., mass, surface area, volume) of the fractal within a ball of radius r centered on the fractal is given by:

$$M(r) = A(r) r^D, \tag{29}$$

where $\log A(r)/\log r \to 0$ for $r \to \infty$.

As an application, for a deterministic mass fractal of length L, scaling factor β_s, and k structural units in the first iteration, we can write [1]:

$$M(L) = kM(\beta_s L), \tag{30}$$

and using Eq. (29), one obtains a formula for calculating the fractal dimension of the mass fractal:

$$k\beta_s^D = 1. \tag{31}$$

Thus, if a finite iteration of the fractal consists of N scattering balls of radius a, the fractal dimension is given by the asymptotic:

$$N \propto \left(\frac{L}{a}\right)^D, \tag{32}$$

in the limit of large number of iterations.

If the quantity to be measured is the *mass* $M(r)$ embedded in a disk of radius r, Eq. (30) becomes $M(r) \propto r^{D_m}$, which leads to $I(q) \propto q^{-D_m}$. The lower the value of mass fractal dimension D_m, the less compact is the structure. In a similar way, for a surface fractal of fractal dimension Ds, its surface obeys $S(r) \propto r^{2-D_s}$ and thus the scattering intensity decays as $I(q) \propto q^{-(2d-D_s)}$.

More generally, in a two-phase system where one phase is of dimension D_m and the second phase is its complement set of dimension D_p ("pores"), the "boundary" between the two phases also forms a set of dimension D_s ("surface"). Thus, for a mass fractal, we have $D_s = D_m < d$ and $D_p = d$, while for a surface fractal we have $D_m = D_p = d$ and $d - 1 < D_s < d$ [37, 38]. The possibility of differentiating between mass and surface fractals makes SAS a very convenient technique for measuring fractal dimensions of materials at nano- and micro-scales.

2.3. Iterated function system

As a first method of generating fractals, we consider an iterated function system (IFS). By definition, an IFS is a complete metric space $(\boldsymbol{X}, d)$ together with a finite set $w_n : \boldsymbol{X} \to \boldsymbol{X}$ of contraction mappings and contractivity factors $s_n, n > 1, 2, \cdots, N$. In general, a transformation $f : \mathbf{X} \to \mathbf{X}$ on a metric space (X, d) is a contraction mapping if there is a constant (contractivity factor) $0 \le s < 1$ such that

$$\mathbf{d}(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{y})) \le \mathbf{s}.\ \mathbf{d}(\mathbf{x}, \mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in \mathbf{X}. \tag{33}$$

By considering a hyperbolic IFS and if we denote by $(H(\boldsymbol{X}), h(d))$, the space of nonempty compact subsets with the Hausdorff metric $h(d)$, then the transformations $W : H(\boldsymbol{X}) \to H(\boldsymbol{X})$ defined by $W(B) = \cup_{n=1}^{N} w_n(B), \forall B \in H(\boldsymbol{X})$ is a contraction mapping on the complete metric space $(\mathcal{H}(\mathbf{X}), h(d))$ with the contractivity factor s [46], i.e.,

$$h(W(B), W(C)) \le s \cdot h(B, C) \forall B, C \in \mathcal{H}(\mathbf{X}). \tag{34}$$

The unique fix point $A \in \mathcal{H}(\mathbf{X})$ obeys $A = \cup_{n=1}^{N} w_n(A)$ and is called the attractor of the IFS, which is a deterministic fractal [46].

We generate here the attractor by using random iteration algorithm, and thus we assign the probability $p_n > 0$ to w_n for $n = 1, 2, \cdots$ where $\sum_{n=1}^{N} p_n = 1$. Then, a point $x_0 \in \boldsymbol{X}$ is chosen, and we build recursively the sequence $x_k \in \{w_1(x_{k-1}), w_2(x_{k-1}), \cdots, w_N(x_{k-1})\}$, where the probability of the event $x_k = w_n(x_{k-1})$ is p_n, and $k = 1, 2, \cdots$. This leads to the sequence $\{x_{k-1} : k = 0, 1, \cdots\}$, which converges to the attractor of IFS.

2.4. Cellular automata

Another important method to generate exact self-similar fractals is by using cellular automata (CA) [47–49]. They provide simple models for dynamical systems dealing with the emergence of collective phenomena such as chaos, turbulence, or fractals. Basically, a cellular automaton is a set of cells on a grid (rectangular, hexagonal, etc.) that evolves through a number of discrete steps according to a set of rules based on the states of neighboring cells. The rules are then

applied alternatively for as many times as needed. The grid is n-dimensional but for our purposes, we will choose $n = 1$.

In this chapter, we shall present Rule 90 [49] together with the corresponding structure factor based on Pantos formula given by Eq. (15). For particular values of the number of steps, Rule 90 generates exact shapes of the Sierpinski gasket (SG). However, for most of the number of steps, it generates intermediate structures between two consecutive iterations of SG. Thus, for Rule 90, CA extends considerably the number of structures generated, and therefore new classes of materials consisting of a "mixture" of SG at various iterations can be investigated.

More generally, SAS from CA could be used to check whether a generated structure is a mass or surface fractal (or none of them), i.e., whether there exists a power-law distribution of some entities (collections of cells). In addition, through the oscillations of the scattering curve in the fractal region, SAS from CA can shed some light on the randomness of the generated structures. This could be of particular interest since some rules like Rule 30 generates so-called pseudo-random structures. However, this is beyond the scope of this chapter. Here, we shall restrict ourselves to calculation and interpretation of SAS intensities from basic structures, such as those based on SG. This shall facilitate a quick comparison with the theoretical model based on SG presented in the next section, and thus to support the validity of the obtained results.

3. Small-angle scattering from mass fractals

As an example of a deterministic mas fractal, we calculate the scattering from a two dimensional Sierpinski gasket (SG), generated by three different methods. In order to calculate the scattering intensities of an ensemble of triangles, we use the following properties:

- When the size of a triangle is scaled as $a \to \beta_s a$, then the form factor scales as $F(\boldsymbol{q}) \to F(\beta_s \boldsymbol{q})$;
- When the triangle is translated by a vector $\boldsymbol{b}$ such as $\mathbf{r} \to \mathbf{r} + \mathbf{r}$, then $F(\boldsymbol{q}) \to F(\boldsymbol{q}) e^{-iq.\ b}$.

Zero-th iteration of SG consists of a single triangle of edge size a (here, $a = 1$ nm). First iteration ($m = 1$) consists of four smaller triangles, each of the edge length $a/2$ as shown in **Figure 3a**. At second iteration ($m = 2$), the same operation is repeated for each of the triangles of edge length $a/2$. In the limit of large number of iterations m, the total number of triangles of edge size $a_m = a/2^m$ is:

$$N = 3^m. \tag{35}$$

Therefore using Eq. (31), one obtains the fractal dimension of SG as:

$$D = \lim_{m \to \infty} \frac{\log N}{\log (a/a_m)} \approx 1.585. \tag{36}$$

At m-th iteration, the positions of the triangles forming the SG are given by:

$$P_m\left(\vec{q}\right) \equiv \prod_{i=1}^{m} G_i\left(\vec{q}\right), \tag{37}$$

where $G_i\left(\vec{q}\right)$ is known as the generative function [29, 31, 38], and is given by

$$G_1\left(\vec{q}\right) = \frac{1}{3}\sum_{j=0}^{2} e^{-i\vec{q}\cdot\vec{a}_j}, \tag{38}$$

where the translation vectors are given by Eq. (25), $G_m\left(\vec{q}\right) = G_1\left(\beta_s^{m-1}\ \vec{q}\right)$, and $\beta_s = 1/2$. **Figure 5** left part, shows first three iterations of SG.

By using the property given in Eq. (37), the fractal structure factor can be written as [29]:

$$S(q)/N = \left\langle \prod_{i=1}^{m} \left| G_i\left(\vec{q}\right)\right|^2 . \right\rangle \tag{39}$$

Thus, by introducing Eqs. (39) and (22) into Eq. (10), we obtain an analytical expression for the scattering intensity:

$$I(q)/I(0) = \left\langle \left| F_0\left(\beta_s^m a\,\vec{q}\right)\right|^2 \right\rangle S(q)/N. \tag{40}$$

By ignoring the form factor F_0 in Eq. (40), an analytical expression of the structure factor is obtained. This case is discussed in [40].

The corresponding scattering intensities are shown in **Figure 5** right part, for a triangle ($m = 0$) and for the first three iterations of SG ($m = 1, 2, 3$). At low q-values ($q \lesssim 2$ nm^{-1}), all the scattering curves are characterized by a Guinier region. A main feature of scattering from deterministic mass fractals is that after the Guinier region, it follows a fractal regime in which the absolute value of the scattering exponent equals the fractal dimension of the fractal. The length of the fractal regime increases with increasing the iteration number since the distances between the scattering units of the fractals (here triangles) decrease. In **Figure 5** right part, the fractal regime is clearly seen within the range $2 \lesssim q \lesssim 14$ nm^{-1} for $m = 3$. It is characterized by a succession of maxima and minima superimposed on a power-law decay (also known as a generalized power-law decay [29]). The number of minima in the fractal regime is equal with the fractal iteration number and from their periodicity, we can extract the value of the scaling factor [29]. Beyond the fractal regime, one obtain as expected, the Porod region where the exponent of the power-law decay is -3 (or -4 for three-dimensional objects). In **Figure 5** right part, at $m = 3$ the Porod regime begins near $q \gtrsim 14$ nm^{-1}.

By taking into account, the effect of polydispersity or the random distribution of the scattering units, the scattering curve is smoothed and thus the maxima and minima are smeared out. The "smoothness" of the curve increases with increasing the relative variance of the distribution function, and they can be completely smeared out when a threshold is reached. Experimentally,

most of the times, SAS experiments give this type of behavior, when the curve is completely smeared out. Please note that if we neglect the contribution of the form factor F_0 in Eq. (40), then the Porod region is replaced by an asymptotic region, from which the number of scattering units inside the fractal can be obtained (see **Figure 6**). This case is presented and discussed in the following, for the SG generated using CGR and IFS, respectively.

CGR representation gives directly the positions of the scattering units in the fractal. **Figure 6** left part, shows the SG generated from CGR for $N = 30, 130, 230,$ and for 430 points, respectively. The figure clearly shows that by increasing the number of points, the obtained structure approaches better the structure of SG. From the same figure, we can see that for $N = 430$ points, the second iteration of SG can be quite clearly distinguished. Thus, a convenient way to calculate the scattering intensity is to use Pantos formula given by Eq. (15), since we neglect the shape of the scattering units.

Figure 6 right part shows the scattering structure factor of SG built from the CGR, for the four structures in the left part. Generally, all scattering curves are characterized by the presence of the three main regions specific to SG, obtained using the analytic representation (**Figure 5**) right part: Guinier region at low q, fractal region at intermediate q, and here an asymptotic region instead of a Porod one, since we neglect the form factor. The same figure shows that by increasing the number of particles, the length of the fractal region also increases. This is to be expected, since increasing the number of points leads to a better approximation of the deterministic SG. However, for SG generated using CGR, a transition region appears (at $40 \lesssim q \lesssim 200$), since in this region, the pair distance distribution function does not follow a power-law decay distribution of the number of distances. Finally, in the asymptotic region, the curves are proportional to $1/N$, and thus, the number of points can be recovered. The asymptotic values of the curves are marked by horizontal dashes lines in **Figure 5** right part.

Rule 90 (**Figure 7** left part) is an elementary cellular automaton rule, and it produces SG for particular values of the number of steps. Although the system is generated on a rectangular

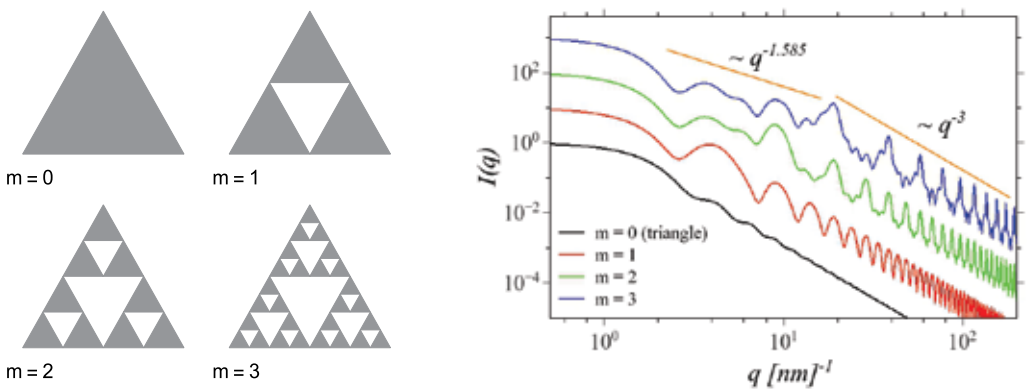

Figure 6. Left part: first four iterations of the Sierpinski gasket generated using a deterministic algorithm; right part: The corresponding SAS intensities of the Sierpinski gasket. The curves are shifted vertically by a factor 10^m, for clarity. The lowest curve corresponds to m = 0, and the highest one corresponds to m = 3.

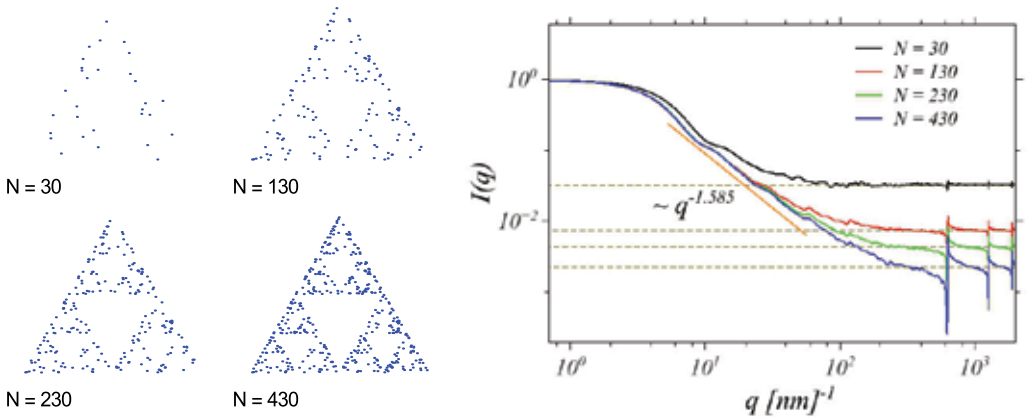

Figure 7. Left part: Sierpinski gasket generated using CGR, for different values of the number of scattering units; right part: the corresponding SAS intensities of the Sierpinski gasket from CGR. The highest curve corresponds to N = 30 and the lowest one corresponds to N = 430.

grid with cells of finite size, we consider here (as in the previous example) that the scattering units are points centered in the cells. **Figure 7** left part, shows the structure generated by Rule 90 for $p = 31, 41, 51,$ and for $p = 63$ steps, respectively. Note that $p = 31$ and $p = 63$ correspond to SG at iterations $m = 3$, and $m = 4$, respectively, while $p = 41$ and $p = 52$ correspond to some intermediate structures between the two consecutive iterations.

The corresponding structure factors are shown in **Figure 7** right part. In the scattering curve, all the three main regions are present: Guinier, fractal, and asymptotic regions. However, as opposed to the Guinier region of the SG generated deterministically, or through CGR, here its length decreases with increasing the number of steps. This is a consequence of the construction algorithm used, namely the structure increases its size at every step. Decreasing the length of the Guinier region leads to an increase of the fractal one, since the smallest distances between scattering points remain the same. This is indicated in the scattering curve, by approximately equal positions ($q \simeq 360$ nm^{-1}) of the first minima, for each of the four steps. As expected, the absolute value of the scattering exponent in the fractal region coincides with the fractal dimension of the structure, and the asymptotic behavior at large q (shown by horizontal dotted lines in **Figure 7** right part) tens to $1/N$, where N is the total number of scattering points.

4. Small-angle scattering from surface fractals

Recently, it has been shown that surface fractals can be built as a sum of mass fractals [37, 38]. As an example, we consider here a generalized version of the Koch snowflake (KS). The construction algorithm starts from an equilateral triangle with edge length a, which is the zero-th mas fractal iteration (black triangle in **Figure 8** left part). At the second step, each side is divided into three segments, each of length $a/3$, and a new triangle pointing outward is added. The base of the new triangle coincides with the central segment. This is the first mass fractal iteration (three

http://dx.doi.org/10.5772/intechopen.70870

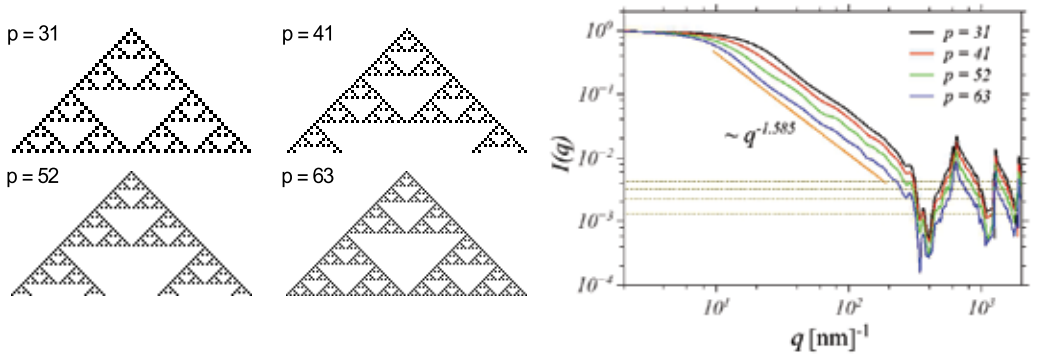

Figure 8. Left part: Rule 90 for $p = 31, 41, 52$ and $p = 63$ number of steps, respectively; right part: the corresponding scattering intensities. The leftmost curve corresponds to p = 63 and the rightmost one corresponds to p = 31.

orange triangles in **Figure 8** left part). By repeating the same procedure for each of the new triangles, one obtains the KS having the fractal dimension:

$$D_s = \lim_{m\to\infty} \frac{\log 3 \cdot 4^m}{\log (a/a_m)} = \frac{\log 4}{\log 3} \approx 1.26, \tag{41}$$

where a_m is the side length at mth iteration. In **Figure 8**, each of the triangle is scaled down by a factor of $0.6 \cdot a_m$. The case when the triangles are not scaled down corresponds exactly to the well-known KS and its scattering properties have been extensively studied in [38].

Since a surface fractal can be constructed as a sum of mass fractals, the normalized scattering amplitude is written as a sum of the scattering amplitudes of mass fractals. If we denote $A_m\left(\vec{q}\right) \equiv Area(a)F_m\left(\vec{q}\right)$, then a recurrence formula for the scattering amplitude of the surface fractal at arbitrarily iteration can be written as [38]:

$$A_m\left(\vec{q}\right) = 6G_3\left(\vec{q}\right)\left[\beta_s^2 A_{m-1}\left(\beta_s\,\vec{q}\right) - 6G_2\left(\beta_s\,\vec{q}\right)\beta_s^4 A_{m-2}\left(\beta_s^2\,\vec{q}\right)\right] + \beta_s^2 A_{m-1}\left(\beta_s\,\vec{q}\right)\left[1 + 6G_2\left(\vec{q}\right)\right]. \tag{42}$$

Here, the scaling factor is $\beta_s = 3/10$, and the generative functions are given by:

$$G_2\left(\vec{q}\right) = \frac{1}{6}\sum_{j=0}^{5} e^{-i\vec{q}\cdot\vec{c}_j}, \tag{43}$$

and respectively by:

$$G_3\left(\vec{q}\right) = \frac{1}{6}\sum_{j=0}^{5} e^{-i\vec{q}\cdot\vec{b}_j}. \tag{44}$$

The translation vectors are defined as:

$$\vec{c}_j = \frac{2a}{9}\{\cos\pi(j+1/2)/3,\ \sin\pi(j+1/2)/3\}, \tag{45}$$

and the vectors $\vec{b}_j$ are given by Eq. (26). Thus, using Eq. (2), the non-normalized scattering amplitude is $I(q) \propto \left\langle \left| A_m\left(\vec{q}\right)\right|^2 \right\rangle$, provided the amplitudes at $m=0$ and $m=1$ are known.

Figures 8 and **9** left parts show the construction of the generalized KS when the ratio between the sizes of the triangles and the distances between them are 0.6 and 0.4, respectively. **Figures 8** and **9** right parts show the corresponding scattering intensities. The general feature is that when the ratio is 0.4, in the fractal region the overall agreement between the total scattering intensity and the intensity corresponding to uncorrelated triangles is slightly better. The curve also shows that the absolute value of the scattering exponent is now $6 - D_s$, with $D_s = 1.26$ given by Eq. (41), which is a "signature" of scattering from surface fractals. The reason for such behavior is that at a given iteration, the surface fractal consists of scattering units of different sizes following a power-law distribution (see **Figure 8** left part), while mass fractals consist of scattering units of the same size (see **Figure 5** left part). A mathematical derivation of the value of the scattering exponent in the case of scattering from surface fractals can be found in [37, 38].

The centers of triangles in **Figure 8** left part coincide with the centers of triangles of the regular KS. From another hand, it has been recently shown that the scattering intensity corresponding to a system of triangles whose sizes follow a power-law distribution and whose positions are uncorrelated, approximates under certain conditions the scattering intensity of the same system but when the positional correlations are not considered [38]. **Figure 8** right part, shows also the scattering intensity of a system of triangles whose positions are uncorrelated (red curve). We can observe that in the fractal region ($20 \lesssim q \lesssim 400$ nm^{-1}), the approximation is not quite satisfactory. The reason of this behavior is that here the scaling factor is β_s, and thus the distances between the centers of the triangles is of the same order as their size. However, when the distances between

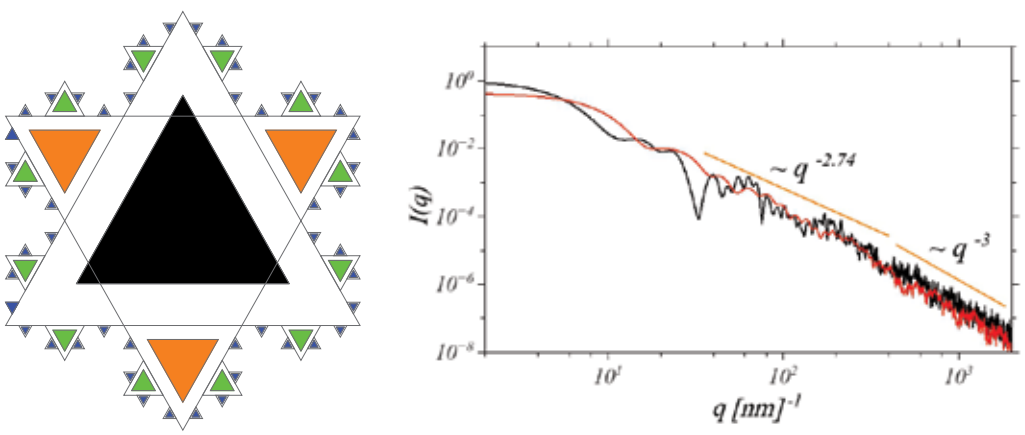

Figure 9. Left part: a generalized Koch snowflake model. Right part: highly oscillating curve—the corresponding scattering intensity, smoother curve—the scattering intensity from a system containing the same number of triangles but whose positions are randomly.

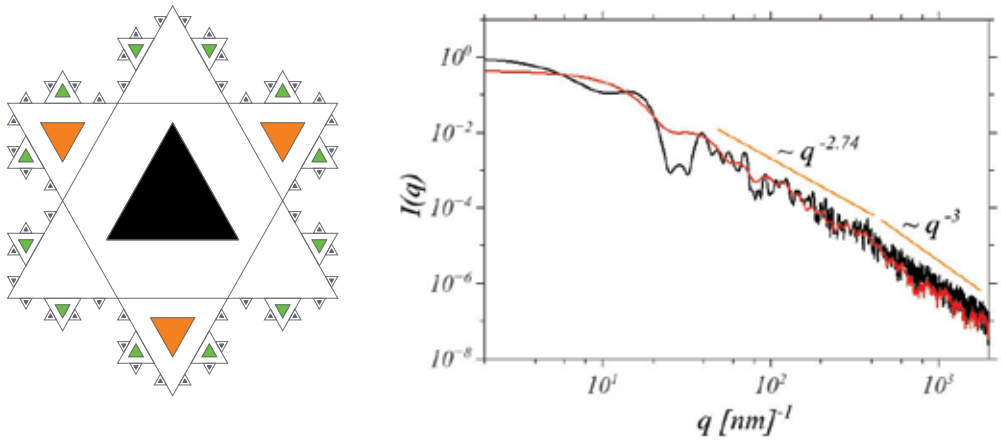

Figure 10. Left part: a generalized Koch snowflake model. Right part: highly oscillating curve—the total scattering intensity, smoother curve—the scattering intensity from a system containing the same number of triangles but whose positions are random.

the scattering units are much bigger than the sizes of the units, the approximation of the uncorrelated positions of the triangles works fairy well in the fractal region [38] (**Figure 10**).

5. Conclusions

In this chapter, we have presented and discussed some general concepts in small-angle scattering (SAS; neutrons, X-ray, light) from deterministic mass and surface fractals. To do so, we have considered the Sierpinski gasket (SG) as a model for deterministic mass fractals, and a generalized version of Koch snowflake (KS) as a model for deterministic surface fractals. The model for SG has been introduced through three main algorithms: deterministic and random iteration function system (IFS), and through cellular automata (CA). KS has been constructed in the framework of deterministic IFS.

The SAS intensities (fractal and structure factor) has been calculated from a system of 2D, monodisperse diluted (i.e., spatial correlation can be neglected) and randomly oriented fractals. They are characterized by the presence of three main regions: Guinier (at low q), fractal (at intermediate q), and Porod/asymptotic (at high q). We have shown that in the case of mass fractals, we can extract structural information about the fractal dimension (from the exponent of the SAS curve in the fractal region), scaling factor (from the periodicity of minima in the fractal region), iteration number (from the number of minima in the fractal region), and the total number of scattering units inside a fractal (from the value of the structure factor in the asymptotic region). In addition, mass fractals generated using IFS are able to reproduce fairly well the scattering curve of deterministic IFS under the proper conditions. Mass fractals generated from CA increase their size with increasing the number of steps. This growing process is also reflected by a decrease of the length of Guinier region. We have also shown that

a surface fractal can be considered as a superposition of mass fractals at various iterations and the range of structural information which can be extracted is similar to the case of scattering from mass fractals.

The IFS and CA algorithms used to generate the mass fractals models can be easily extended to more general structures and can be used to address various questions. For example, in the case of CA, SAS could be used to determine the fractal dimension of an arbitrary structure generated using one dimensional rules, it can shed some light on the randomness of some structures, or it can reveal the existence of a power-law distribution of some entities (of arbitrarily shape) generated by a specific rule.

Author details

Eugen Mircea Anitas[1,2]*

*Address all correspondence to: anitas@jinr.ru

1 Joint Institute for Nuclear Research, Dubna, Moscow Region, Russian Federation

2 Horia Hulubei National Institute of Physics and Nuclear Engineering, Bucharest-Magurele, Romania

References

[1] Mandelbrot BB. The Fractal Geometry of Nature. New York: W. H. Freeman; 1982. 468 p

[2] Feder J. Fractals. New York: Springer; 1988. 284 p

[3] Vicsek T. Fractal Growth Phenomena. 2nd ed. Singapore: World Scientific; 1992. 528 p

[4] Bunde A, Havlin S, editors. Fractals and Disordered Systems. 2nd ed. Berlin Heidelbeerg: Springer; 1995. 408 p. DOI: 10.1007/978-3-642-84868-1

[5] Willis BTM, Carlile CJ. Experimental Neutron Scattering. Great Britain: Oxford University Press; 2009. 344 p

[6] Strangl J, Mocuta C, Chamard V, Carbone D. Nanobeam X-Ray Scattering: Probing Matter at the Nanoscale. Singapore: Wiley-VCH; 2013. 284 p

[7] Monteiro WA, editor. Neutron Scattering. Croatia: InTech; 2016. 151 p. DOI: 10.5772/60694

[8] McStas—a Neutron Ray-Trace Simulation Package [Internet]. Available from: http://www.mcstas.org [Accessed: Thu Oct 05 15:03:20 2017]

[9] FullProf Suite [Internet]. Available from: https://www.ill.eu/sites/fullprof/index.html [Accessed: Thu Oct 05 15:03:20 2017]

[10] Ariga K. Manipulation of Nanoscale Materials: An Introduction to Nanoarhitectonics. Great Britain: RSC Publishing; 2012. 473 p. DOI: 10.1039/9781849735124

[11] Alguero M, Gregg JM, Mitoseriu L, editors. Nanoscale Ferroelectrics and Multiferroics. India: Wiley; 2016. 984 p

[12] Bica I, Anitas EM, Chirigiu L, Bunoiu M, Juganaru I, Tatu RF. Magnetodielectric effects in hybrid magnetorheological suspensions. Journal of Industrial and Engineering Chemistry. 2015;**22**:53-62. DOI: 10.1016/j.jiec.2014.06.024

[13] Bica I, Anitas EM, Averis LME. Influence of magnetic field on dispersion and dissipation of electric field of low and medium frequencies in hybrid magnetorheological suspensions. Journal of Industrial and Engineering Chemistry. 2015;**27**:334-340. DOI: 10.1016/j.jiec.2014.09.047

[14] Bica I, Anitas EM, Averis LME. Tensions and deformations in composites based on polyurethane elastomer and magnetorheological suspension: Effects of the magnetic field. Journal of Industrial and Engineering Chemistry. 2015;**28**:86-90. DOI: 10.1016/j.jiec.2015.02.003

[15] Newkome GR, Wang P, Moorefield CN, Cho TJ, Mohapatra PP, Li S, Hwang SH, Lukoyanova O, Echegoyen L, Palagallo JA, Iancu V, Hla SW. Nanoassembly of a fractal polymer: A molecular â€œSierpinski hexagonal Gasketâ€. Science. 2006;**312**:1782-1785. DOI: 10.1126/science.1125894

[16] Cerofolini GF, Narducci D, Amato P, Romano E. Fractal nanotechnology. Nanoscale Research Letters. 2008;**3**:381-385. DOI: 10.1007/s11671-008-9170-0

[17] Mayama H, Tsuji K. Menger sponge-like fractal body created by a novel template method. The Journal of Chemical Physics. 2006;**125**:124706. DOI: 10.1063/1.2336200

[18] Berenschot EJW, Jansen HV, Tas NR. Fabrication of 3D fractal structures using nanoscale anisotropic etching of single crystalline silicon. Journal of Micromechanics and Microengineering. 2013;**23**:055024. DOI: 10.1088/0960-1317/23/5/055024

[19] Guinier A, Fournet G. Small-Angle Scattering of X-Rays. New York: John Wiley & Sons; 1955. 263 p

[20] Brumberger H, editor. Modern Aspects of Small-Angle Scattering. New York: NATO ASI Series; 1995. 463 p

[21] Feigin LA, Svergun DI. Structure Analysis by Small-Angle X-Ray and Neutron Scattering. New York and London: Plenum Press; 1987. 335 p

[22] Franke D, Petoukhov MV, Konarev PV, Panjkovich A, Tuukkanen A, Mertens HDT, Kikhney AG, Hajizadeh NR, Franklin JM, Jeffries CM, Svergun DI. ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. Journal of Applied Crystallography. 2017;**50**:1-14. DOI: 10.1107/S1600576717007786

[23] Craus ML, Islamov AK, Anitas EM, Cornei N, Luca D. Microstructural, magnetic and transport properties of La0.5Pr0.2Pb0.3-xSrxMnO3 manganites. Journal of Alloys and Compounds. 2014;**592**:121-126. DOI: 10.1016/j.jallcom.2014.01.002

[24] Heller WT. Small-angle neutron scattering and contrast variation: A powerful combination for studying biological structures. Journal of Applied Crystallography. 2010;**D66**:1213-1217. DOI: 10.1107/S0907444910017658

[25] Gabel F, Lensink MF, Clantin B, Jacob-Dubuisson F, Villeret V, Ebel C. Probing the conformation of FhaC with small-angle neutron scattering and molecular Modeling. Biophysical Journal. 2014;**107**:185-196. DOI: 10.1016/j.bpj.2014.05.025

[26] Zemb T, Lindner P. Neutron, X-Rays and Light. Scattering Methods Applied to Soft Condensed Matter. The Netherlands: North Holland; 2002. 552 p

[27] Martin JE, Hurd AJ. Scattering from fractals. Journal of Applied Crystallography. 1987;**20**:61-78. DOI: 10.1107/S0021889887087107

[28] Schmidt PW. Small-angle scattering studies of disordered, porous and fractal systems. Journal of Applied Crystallography. 1991;**24**:414-435. DOI: 10.1107/S0021889891003400

[29] Cherny AY, Anitas EM, Osipov VA, Kuklin AI. Deterministic fractals: Extracting additional information from small-angle scattering data. Physical Review E. 2011;**83**:036203. DOI: 10.1103/PhysRevE.84.036203

[30] Anitas EM. Small-angle scattering from fat fractals. European Physical Journal B. 2014;**87**:139. DOI: 10.1140/epjb/e2014-41066-9

[31] Cherny AY, Anitas EM, Osipov VA, Kuklin AI. Small-angle scattering from multiphase fractals. Journal of Applied Crystallography. 2014;**47**:198-206. DOI: 10.1107/S1600576713029956

[32] Anitas EM. Microscale fragmentation and small-angle scattering from mass fractals. Advances in Condensed Matter Physics. 2015;**2015**:5. ID 501281. DOI: 10.1155/2015/501281

[33] Anitas EM. A structural model for scattering intensities with multiple fractal regions of variable length. Journal of Optoelectronics and Advanced Materials. 2015;**17**:1122-1127

[34] Anitas EM. Scattering structure factor from fat fractals. Romanian Journal of Physiology. 2015;**60**:647-652

[35] Cherny AY, Anitas EM, Osipov VA, Kuklin AI. The structure of deterministic mass, surface and multi-phase fractals from small-angle scattering data. Romanian Journal of Physiology. 2015;**60**:658-663

[36] Anitas EM, Osipov VA, Kuklin AI, Cherny AY. Influence of randomness on small-angle scattering from deterministic mass fractals. Romanian Journal of Physiology. 2015;**61**:457-463

[37] Cherny AY, Anitas EM, Osipov VA, Kuklin AI. Scattering from surface fractals in terms of composing mass fractals. Journal of Applied Crystallography. 2017;**50**:919-931. DOI: 10.1107/S1600576717005696

[38] Cherny AY, Anitas EM, Osipov VA, Kuklin AI. Small-angle scattering from the cantor surface fractal on the plane and the Koch snowflake. Physical Chemistry Chemical Physics. 2017;**19**:2261-2268. DOI: 10.1039/C6CP07496K

[39] Anitas EM, Slyamov A, Todoran R, Szakacs Z. Small-angle scattering from Nanoscale fat fractals. Nanoscale Research Letters. 2017;**12**:389. DOI: 10.1186/s11671-017-2147-0

[40] Anitas EM, Slyamov A. Structural characterization of chaos game fractals using small-angle scattering analysis. PLoS One. DOI: 10.1371/journal.pone.0181385

[41] Sandru A. Fractal structures for optical applications feasible with 2D nanotechnology. Journal of Modern Optics. 2012;**59**:199-204. DOI: 10.1080/09500340.2011.624646

[42] Shang J, Wang Y, Chen M, Dai J, Zhou X, Kuttner J, Hilt G, Shao X, Wu K. Assembling molecular Sierpinski triangle fractals. Nature Chemistry. 2015;**7**:389-393. DOI: 10.1038/nchem.2211

[43] Loh ND, Hampton CY, Martin AV, Starodub D, Sierra RG, Barty A, Aquila A, Schulz J, Lomb L, Steinbrener J, et al. Fractal morphology, imaging and mass spectrometry of single aerosol particles in flight. Nature. 2012;**486**:513-517. DOI: 10.1038/nature11222

[44] Debye P. Zerstreuung von Röntgenstrahlen. Annalen der Physik. 1915;**46**:809-823. DOI: 10.1002/andp.19153510606

[45] Pantos E, van Garderen HF, Hilbers PAJ, Beelen TPM, van Santen RA. Simulation of small-angle scattering from large assemblies of multi-type scatterer particles. Journal of Molecular Structure. 1996;**383**:303-306. DOI: 10.1016/S0022-2860(96)09302-7

[46] Barnsley M. Fractals Everywhere. USA: Dover Publications; 2012. 560 p

[47] Wolfram S. Statistical mechanics of cellular automata. Reviews of Modern Physics. 1983;**55**:601-644. DOI: 10.1103/RevModPhys.55.601

[48] Wolfram S. Universality and complexity in cellular automata. Physica D: Nonlinear Phenomena. 1984;**10**:1-35. DOI: 10.1016/0167-2789(84)90245-8

[49] Wolfram S. A New Kind of Science. USA: Wolfram Media; 2002. 1192 p

Edited by Ricardo López-Ruiz

Modeling and simulating biological and physical systems are nowadays active branches of science. The diversity and complexity of behaviors and patterns present in the natural world have their reciprocity in life systems. Bifurcations, solitons and fractals are some of these ubiquitous structures that can be indistinctively identified in many models with the most diverse applications, from microtubules with an essential role in the maintenance and the shaping of cells, to the nano/microscale structure in disordered systems determined with small-angle scattering techniques. This book collects several works in this direction, giving an overview of some models and theories, which are useful for the study and analysis of complex biological and physical systems. It can provide a good guidance for physicists with interest in biology, applied research scientists and postgraduate students.

Published in London, UK

IntechOpen

ISBN 978-1-83881-342-0