Biharmonic Problems and their Applications in Engineering and Technology

Hovik A. Matevossian[1,2], Giorgio Nordo[3], and Giovanni Migliaccio[4]

[1] Federal Research Centre "Computer Science & Control" & Steklov Mathematical Institute, Russian Academy of Sciences, Vavilov str., 40-42, R 374, Moscow 119333, Russia
[2] Moscow Avaiation Institute (National Research University), Volokolomskoe shosse, 4, Moscow 125993, Russia
(E-mail: hmatevossian@graduate.org)
[3] MIFT – Università di Messina, Viale Ferdinando Stagno d'Alcontres, 31 - 98166 Messina, Italy
(E-mail: giorgio.nordo@unime.it)
[4] Università di Pisa, Largo Luccio Lazzarino 2, Pisa 56122, Italy
(E-mail: giovanni.migliaccio.it@gmail.com)

Abstract. In the present paper we study some properties of solutions of biharmonic problems. Namely, we study the Steklov, Steklov-type and Neumann boundary value problems for the biharmonic equation. For solving these biharmonic problems with application, in particular, to radar imaging, we need to solve the Dirichlet, Neumann and Cauchy boundary value problems for the Poisson equation using the scattering model. In order to select suitable solutions, we solve the Poisson equation with the corresponding boundary conditions, that is, some criterion function is minimized in the Sobolev norms. Under appropriate smoothness assumptions, these problems may be reformulated as boundary value problems for the biharmonic equation.
Keywords: Biharmonic Operator, Boundary Value Problems, Scattering Model, Variational Method.

1 Introduction

Let $\Omega \subset \mathbb{R}^n,\ n \geq 2$, be a bounded Lipschitz domain with connected boundary $\partial\Omega$, and $\Omega \cup \partial\Omega = \overline{\Omega}$ is the closure of Ω.

In Ω we consider the following boundary value problems for the biharmonic equation:

$$\Delta^2 u = F, \qquad x \in \Omega \tag{1}$$

with the Steklov boundary conditions

$$\begin{cases} u = g_1 & \text{on} \quad \partial\Omega, \\ \Delta u + \tau \frac{\partial u}{\partial \nu} = g_2 & \text{on} \quad \partial\Omega, \end{cases} \tag{2}$$

or the Steklov–type boundary conditions

$$\begin{cases} \frac{\partial u}{\partial \nu} = h_1 & \text{on} \quad \partial\Omega, \\ \frac{\partial \Delta u}{\partial \nu} + \tau\, u = h_2 & \text{on} \quad \partial\Omega, \end{cases} \tag{3}$$

or the Neumann boundary conditions

$$\begin{cases} Mu \equiv \sigma \Delta u + (1-\sigma)\frac{\partial^2 u}{\partial \nu^2} = f_1 & \text{on} \quad \partial\Omega, \\ Nu \equiv \frac{\partial \Delta u}{\partial \nu} + (1-\sigma)\frac{1}{2} \cdot \frac{\partial}{\partial t_{ij}} \left(\frac{\partial^2 u}{\partial \nu \partial t_{ij}} \right) = f_2 & \text{on} \quad \partial\Omega, \end{cases} \tag{4}$$

where ν is the outer unit normal vector to the domain, and $T = \{t_{ij}\}$ various tangential directions to the Lipschitz boundary $\partial\Omega$, $\tau \geq 0$, $\tau \not\equiv 0$, and $\tau > 0$ on a set of positive $(n-1)$ – dimensional measure on $\partial\Omega$. The coefficient σ is a constant known as the Poisson ratio, $\frac{1}{n-1} < \sigma < 1$. A unique solution u (modulo linear functions) is obtained in the class of solutions with non-tangential maximal function of the second-order derivatives in $L^p(\partial\Omega)$. The biharmonic Neumann problem in Lipschitz domains was studied in detail in [36].

For $n = 2$, these problems and also the Neumann problem are related to the study of the transverse vibrations of a thin plate with a free edge and which occupies at rest a planar region of shape $\partial\Omega$. The coefficient σ represents the Poisson's ratio of the material that the plate is made of. For more details on the physical interpretation of the Neumann problem and on the Poisson's ratio σ, we refer, for example, to [4]. Note the paper [5], where the author studies the dependence of the vibrational modes of a plate subject to homogeneous boundary conditions upon the Poisson's ratio $0 < \sigma < \frac{1}{2}$, providing also a perturbation formula for the frequencies as functions of the Poisson's coefficient.

Elliptic problems with parameters in the boundary conditions are called Steklov problems from their first appearance in [37]. In the case of the biharmonic operator, these conditions were first considered in [3], [10], [33], who studied the isoperimetric properties of the first eigenvalue.

The standard elliptic regularity results are available in [7]. This monograph covers higher order linear and nonlinear elliptic boundary value problems, mainly with the biharmonic (polyharmonic) operator as leading principal part. Underlying models and, in particular, the role of different boundary conditions are explained in detail. As for linear problems, after a brief summary of the existence theory and L^p and Schauder estimates, the focus is on positivity. The required kernel estimates are also presented in detail.

In [6] and [7], the spectral and positivity preserving properties for the inverse of the biharmonic operator under Steklov and Steklov–type boundary conditions are studied. These are connected with the first Steklov eigenvalue. It is shown that the positivity preserving property is quite sensitive to the parameter involved in the boundary condition.

In [34], the dependence of the eigenvalues of the biharmonic operator subject to Neumann boundary conditions on the Poisson's ratio σ is studied. In particular, it is proved that the Neumann eigenvalues are Lipschitz continuous with respect to $\sigma \in [0, 1)$ and that all the Neumann eigenvalues tends to zero as $\sigma \to 1^-$. Moreover, is showed that the Neumann problem defined by setting $\sigma = 1$ admits a sequence of positive eigenvalues of finite multiplicity that are not limiting points for the Neumann eigenvalues with $\sigma \in [0, 1)$ as $\sigma \to 1^-$ and that coincide with the Dirichlet eigenvalues of the biharmonic operator.

Boundary value problems for a biharmonic (polyharmonic) equation and for the elasticity system in unbounded domains are studied in [12]–[28] in which

the condition of the boundedness of the following weighted Dirichlet integral of solution is finite, namely

$$\int_{\Omega} |x|^a |\partial^\alpha u|^2 \, dx < \infty, \quad a \in \mathbb{R},$$

where $a \in \mathbb{R}$ is a fixed number and $|\partial^\alpha u|^2$ denotes the Frobenius norm of the Hessian matrix of u. In particular, in these papers has been studied the dimension of the space of the solutions to the boundary value problems for a biharmonic (polyharmonic) equation and for the elasticity system, providing explicit formulas which depends on n and a. This paper contains complete proofs of the results, partly presented in [29].

The behavior of solutions of the Dirichlet problem for the biharmonic equation as $|x| \to \infty$ was considered in [8], [9], where estimates for $|u(x)|$ and $|\nabla u(x)|$ as $|x| \to \infty$ were obtained under certain geometric conditions on the domain boundary.

Notation: $C_0^\infty(\Omega)$ is the space of infinitely differentiable functions in Ω with compact support in Ω; $H^m(\Omega)$ is the Sobolev space obtained by the completion of $C^\infty(\overline{\Omega})$ with respect to the norm

$$\|u; H^m(\Omega)\| = \left(\int\limits_{\Omega} \sum_{|\alpha| \le m} |\partial^\alpha u|^2 dx \right)^{1/2}, \; m = 1, 2,$$

where $\partial^\alpha \equiv \partial^{|\alpha|} / \partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}$, $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index, $\alpha_i \ge 0$ are integers, and $|\alpha| = \alpha_1 + \dots + \alpha_n$; $\overset{\circ}{H}{}^m(\Omega)$ is the space obtained by the completion of $C_0^\infty(\Omega)$ with respect to the norm $||u; H^m(\Omega)||$. $\overset{\circ}{H}{}^m_{loc}(\Omega)$ is the space obtained by the completion of $C_0^\infty(\Omega)$ with respect to the family of semi-norms

$$\|u; H^m(\Omega \cap B_0(R))\| = \left(\int\limits_{\Omega \cap B_0(R)} \sum_{|\alpha| \le m} |\partial^\alpha u|^2 \, dx \right)^{1/2}$$

for all open balls $B_0(R) := \{x : |x| < R\}$ in $\mathbb{R}^n$ for which $\Omega \cap B_0(R) \neq \emptyset$. Finally $H^{1/2}(\partial\Omega)$ is the usual trace space on the boundary and $H^{-1/2}(\partial\Omega)$ is its dual (see, for ex., [1]).

2 Definitions and auxiliary statements

If we set $\sigma = 1$, the Neumann boundary conditions reads

$$\begin{cases} \Delta u = f_1 & \text{on} \quad \partial\Omega, \\ \frac{\partial \Delta u}{\partial \nu} = f_2 & \text{on} \quad \partial\Omega. \end{cases} \tag{5}$$

Note that the differential operator associated with problem (1), (5) is not a Fredholm operator. We also note that the boundary conditions (5) do not satisfy the so–called complementing conditions (see [2] and [7] for details), which are necessary conditions for the well–posedness of a differential problems.

Definition 1. A solution of the biharmonic equation (1) in Ω is a function $u \in H^2(\Omega)$ such that, for every function $\varphi \in C_0^\infty(\Omega)$, the following integral identity holds:

$$\int_\Omega \Delta u \, \Delta\varphi \, dx = \int_\Omega F \, \varphi \, dx, \quad F \in L^2(\Omega). \tag{6}$$

Definition 2. A function u is a solution of the Steklov problem (1),(2) with $g_1 = g_2 = 0$, if $u \in H^2(\Omega) \cap \overset{\circ}{H}{}^1(\Omega)$ such that for every function $\varphi \in H^2(\Omega) \cap \overset{\circ}{H}{}^1(\Omega)$, the following integral identity holds

$$\int_\Omega \Delta u \, \Delta\varphi \, dx + \int_{\partial\Omega} \tau \, \nabla u \, \nabla\varphi \, ds = 0. \tag{7}$$

Definition 3. A function u is a solution of the Steklov-type problem (1),(3) with $h_1 = h_2 = 0$, if $u \in H^2(\Omega)$, $\partial u/\partial\nu = 0$ on $\partial\Omega$, such that for every function $\varphi \in C_0^\infty(\mathbb{R}^n)$, $\partial\varphi/\partial\nu = 0$ on $\partial\Omega$, the following integral identity holds

$$\int_\Omega \Delta u \, \Delta\varphi \, dx - \int_{\partial\Omega} \tau \, u \, \varphi \, ds = 0. \tag{8}$$

Definition 4. A function u is a solution of the Neumann problem (1),(5) with $f_1 = f_2 = 0$, if $u \in H^2(\Omega)$ such that the integral identity (6) holds for every function $\varphi \in C_0^\infty(\Omega)$.

3 A scattering model

In the section we derive the mathematical model used for describing the radar process. In our parametrization the unknown is the height function H. As will be shown the height function is determined in two steps. In the first step $\mathfrak{L}(H)$, with $\mathfrak{L}$ a certain second-order differential operator, is determined. After retrieving H the equation $\mathfrak{L}(H) = f$ must be solved. To a good approximation the operator $\mathfrak{L}$ can be replaced by the Laplacian. So the second step simply consists of solving the Poisson equation over some smooth bounded domain, usually a rectangular region in the plane. The problem here is that no natural boundary conditions are available.

Here we will briefly discuss the mathematical inverse problem to be resolved in order to recover the ground topography height function from radar data. First cylindrical coordinates (r, φ, z) are introduced according to Fig. 1, where it is understood that the aircraft is flying at a constant speed along the z-axis. Further r denotes the distance from a point on the ground surface to the z-axis and φ is the angle between radius vector and a horizontal plane through the z-axis. Then the ground surface may be described by a function $H(r, z)$ through the equation

$$\frac{H(r, z)}{r} - \varphi = 0. \tag{9}$$

When r is large, $H(r, z)$ is approximately a Cartesian height function. Fig. 2 shows a top view of the same scene. We have also indicated an aspect

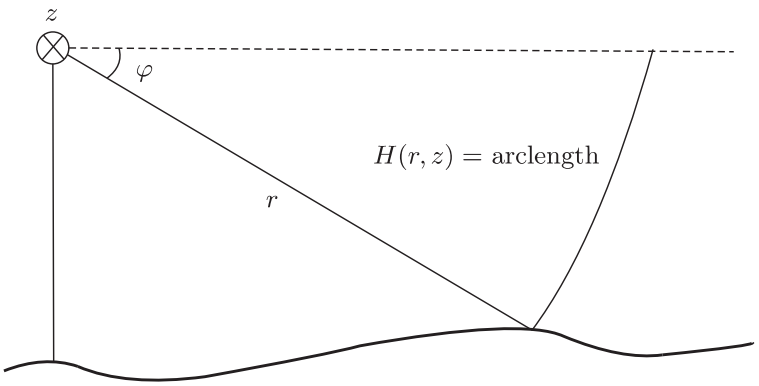


Fig. 1. The ground surface measured at a fixed aircraft position.

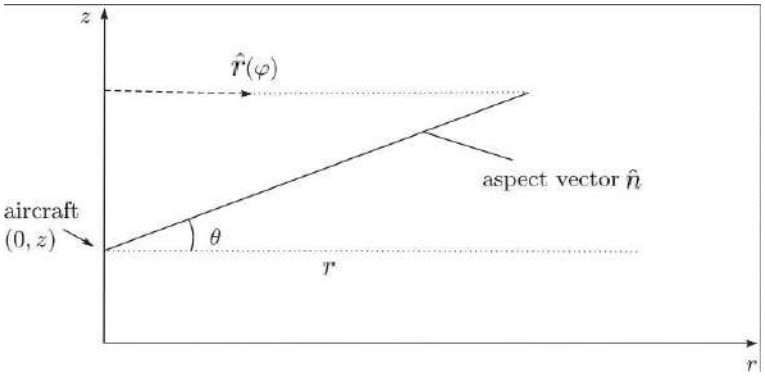


Fig. 2. The measuring geometry as seen from above.

vector from the aircraft to some point on the ground, forming an angle θ with a vertical plane through the aircraft. Normalized to unit length, the aspect vector is denoted by $\hat{n}$. Accordingly

$$\hat{n} = \cos\theta\,\hat{r}(\varphi) + \sin\theta\,\hat{z}. \tag{10}$$

Here $\hat{r}(\varphi)$ denotes the cylindrical unit basis vector corresponding to the r-coordinate for the ground point as shown in the Fig. 2. For a point on the ground surface with coordinates (r, φ, z) we obtain, from Eq. (9), the following expression for the ground surface normal $\bar{m}$,

$$\bar{m} = grad\left(\frac{H(r,z)}{r} - \varphi\right) = \frac{\partial(H/r)}{\partial r}\hat{r} + \frac{1}{r}\frac{\partial H}{\partial z}\hat{z} - \frac{1}{r}\hat{\varphi}. \tag{11}$$

Let $\hat{m}$ denote the normalized normal. Then

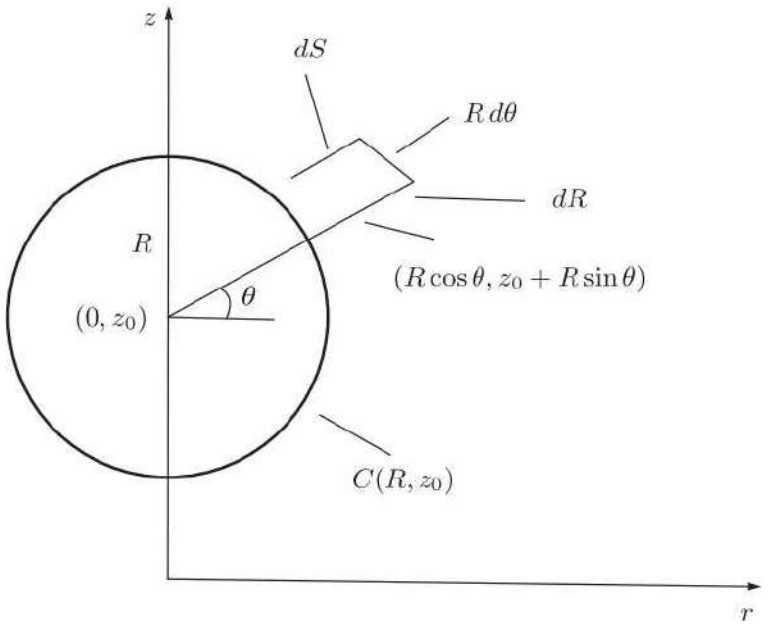


Fig. 3. The coordinate system used to describe an infinitesimal surface element, dS.

$$\hat{m} \circ \hat{n} = \left(r \cos\theta \frac{\partial (H/r)}{\partial r} + \sin\theta \frac{\partial H}{\partial z} \right) \Big/ \sqrt{1 + r^2 \left(\frac{\partial (H/r)}{\partial r} \right)^2 + \left(\frac{\partial H}{\partial z} \right)^2}. \tag{12}$$

Note that (r, φ, z) in Eq. (12) are related to the ground surface point and not to the position of the aircraft.

Let $(z_0, 0)$ be a position of the aircraft and R the distance to some point on the surface. According to Fig. 3 the coordinates (r, z) are then equal to $(R\cos\theta, z_0 + R\sin\theta)$. Next, to obtain a scattering model we will assume that the reflectivity from a ground surface element (see Fig. 4) is

$$\approx \frac{\hat{m} \circ \hat{n}}{R} dR\, d\theta. \tag{13}$$

From Fig. 4, where a vertical plane through $(z_0, 0)$ (the aircraft) and the ground point $(R\cos\theta, z_0 + R\sin\theta)$ is displayed, we conclude that the solid angle $d\Omega$ under which the surface element dS is seen from the antenna is approximately

$$\frac{dR\, \cos\alpha\, R d\, d\theta}{R^2} = -\frac{\hat{m} \circ \hat{n}}{R} dR\, d\theta.$$

In expression (13) we are consequently assuming that the local reflectivity is proportional to the solid angle occupied by the infinitesimal surface element dS. The total reflected signal $G(R, z_0)$ from all points at a distance R from the

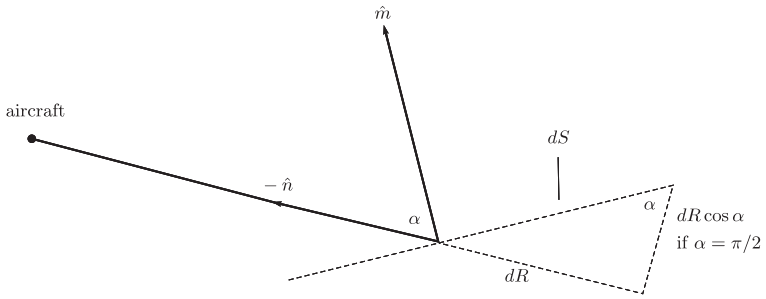


Fig. 4. The infinitesimal surface element, dS, as it is seen from the aircraft.

antenna may now be obtained by integration over the circle $C(R, z_0) = \{(r, z) : r^2 + (z - z_0)^2 = R^2\}$ in Fig. 3.

$$G(R, z_0)dR = c \int_{-\pi}^{\pi} \frac{\hat{m} \circ \hat{n}(R\sin\theta, z_0 + R\cos\theta)}{R}\, d\theta dR$$

i.e.

$$RG(R, z_0) = c \int_{-\pi}^{\pi} \hat{m} \circ \hat{n}(R\sin\theta, z_0 + R\cos\theta) d\theta. \tag{14}$$

Assuming that $\hat{m} \circ \hat{n}$ is small Eq. (12) may be replaced by

$$\hat{m} \circ \hat{n} = r\cos\theta \frac{\partial(H/r)}{\partial r} + \sin\theta \frac{\partial H}{\partial z}.$$

By inserting this into Eq. (14) we get, after multiplying by R,

$$R^2 G(R, z_0) = c \int_{-\pi}^{\pi} \left(rR\cos\theta \frac{\partial(H/r)}{\partial r} + R\sin\theta \frac{\partial H}{\partial z} \right) d\theta.$$

Using the parametrization

$$r = Rcos\theta, \quad z = z_0 + Rsin\theta,$$

this may be rewritten as a curve integral over $C(R, z_0)$, with $dz = R\cos\theta d\theta$ and $dr = -R\sin\theta d\theta$,

$$R^2 G(R, z_0) = c \int_{C(R,z_0)} \left(r \frac{\partial(H/r)}{\partial r} dz - \frac{\partial H}{\partial z} dr \right). \tag{15}$$

By applying Green's formula we get

$$R^2 G(R, z_0) = c \iint_{D(R,z_0)} \mathfrak{L}(H)(r, z)\, dz\, dr, \tag{16}$$

where D is the disc,

$$D(R, z_0) = \{(r, z) : r^2 + (z - z_0)^2 \leq R^2\}$$

and

$$\mathfrak{L}(H) = \frac{\partial}{\partial r}\left(r\frac{\partial(H/r)}{\partial r}\right) + \frac{\partial^2 H}{\partial z^2}. \tag{17}$$

The problem of finding the height function H from radar data $G(r,z)$ may now be divided into two parts:

(i) First solve the integral equation (16) for $\mathfrak{L}(H)(r,z) = f(r,z)$.

(ii) Next solve the partial differential equation

$$\mathfrak{L}(H) = f \tag{18}$$

for H. We note that if r is large and if $\hat{m} \circ \hat{n}$ is small it is reasonable to make the approximation

$$\mathfrak{L}(H) \approx \frac{\partial^2 H}{\partial r^2} + \frac{\partial^2 H}{\partial z^2} = \Delta H$$

so that Eq. (18) becomes Poisson's equation. To consider the first problem (i), both members in Eq. (16) are differentiated with respect to R. Then we get

$$\frac{1}{R}\frac{d}{dR}(R^2 G(R, z_0)) = c\int_{-\pi}^{\pi} \mathfrak{L}(H)(z_0 + R\cos\gamma, R\sin\gamma)\, d\gamma,$$

where the right-hand side is proportional to the average of $\mathfrak{L}(H)$ over the circle $C(R, z_0)$. Hence,

$$\mathfrak{L}(H)^{(F,F)}(\sigma,\omega) \sim |\omega|\left[\frac{1}{R}\frac{d}{dR}\{R^2 G(r,z)\}\right]^{(F,H_0)}(\sigma, \sqrt{\omega^2+\sigma^2}). \tag{19}$$

Here the notation (F,F) means that we have taken the Fourier transform with respect to both the variables and (F,H_0) means that we have taken Fourier transform with respect to the first variable and the Hankel-zero transform with respect to the second. After some calculations Eq. (19) may be rewritten

$$\mathfrak{L}(H)^{(F,F)}(\sigma,\omega) \sim |\omega|\sqrt{\omega^2+\sigma^2}[RG(r,z)]^{(F,H_1)}(\sigma, \sqrt{\omega^2+\sigma^2}). \tag{20}$$

Formula (20) may now be used in order to recover the function $\mathfrak{L}(H)$ in spatial coordinates. Approximating $\mathfrak{L}(H)$ by ΔH we could rewrite Eq. (20) as

$$H^{(F,F)}(\sigma,\omega) \sim |\omega|\frac{1}{\sqrt{\omega^2+\sigma^2}}[RG(r,z)]^{(F,H_1)}(\sigma, \sqrt{\omega^2+\sigma^2}), \tag{21}$$

where H_1 denotes that we have taken the Hankel-one transform with respect to the second variable. Then we could obtain H directly by a two timensional Fourier transform. However, our solution might be expected to have errors caused by, e.g. noisy radar data and errors caused by the particular numerical implementation of the inversion formula (19) (or (20)) and therefore we would rather prefer to divide the solution procedure into the two steps described above and to use the second step, the solution of Poisson's equation, so that we perform some kind of regularization of the final solution. Note also that by using (21) as our solution formula we have tacitly assumed periodic boundary conditions for the Poisson equation.

In the following we will treat part (ii) of the problem, where we wish to define a solution H to the equation

$$\Delta H = f.$$

4 Solution concepts for the Poisson equation

In this section we discuss different possibilities of defining a unique height function. Essentially our approach consists in minimizing some norm of the solution provided that it also satisfies the Poisson equation. In particular we consider the L^2- and H^1-norms. We also show how these two optimization problems may be reformulated as boundary value problems for the biharmonic equation. Note that the corresponding Poisson problem is well-posed unless $\sigma = 1$.

In the domain Ω for the Poisson equation we consider the following boundary value problems

$$\Delta u = f, \quad x \in \Omega \tag{22}$$

with the Dirichlet boundary condition

$$u = g \quad \text{on} \quad \partial\Omega, \tag{23}$$

or the Neumann boundary conditions

$$\nabla u \cdot \nu = h \quad \text{on} \quad \partial\Omega, \tag{24}$$

and the Cauchy boundary conditions

$$\begin{cases} u = g & \text{on} \quad \partial\Omega, \\ \nabla u \cdot \nu = h & \text{on} \quad \partial\Omega, \end{cases} \tag{25}$$

where ν is the outer unit normal vector to $\partial\Omega$.

The boundary operators are independent of any particular choice of orientation for the rectangular coordinate systems. Finally, for Ω a rectangular region in, e.g., the plane

$$\Omega = \{(x, y) : \, a < x < b, \, c < y < d\},$$

there may be the following boundary conditions

$$u(a, y) = u(b, y), \quad u(x, c) = u(x, d), \tag{26}$$

and with the periodic boundary conditions

$$u_x(a, y) = u_x(b, y), \quad u_y(x, c) = u_y(x, d), \tag{27}$$

$$\begin{cases} u(a, y) = u(b, y), \quad u(x, c) = u(x, d), \\ u_x(a, y) = u_x(b, y), \quad u_y(x, c) = u_y(x, d). \end{cases} \tag{28}$$

Provided g is smooth enough boundary conditions (23) define a unique solution of (22). For (24) and (25) the solution is determined up to a constant. It is also possible to use different mixtures of these three types of boundary conditions. Note that for cases (24) and (28) the following consistency conditions must hold, respectively:

$$\int_\Omega f \, dx = \int_{\partial\Omega} h \, ds \quad \text{for (24)}, \qquad \int_\Omega f \, dx = 0 \quad \text{for (25)}.$$

We now consider a different way to select a solution to (22). Here we use a criterion function and optimize this criterion over the set of solutions to the Poisson equation. Scattering model of Section 3 shows the physical interpretation of function $u(x, y)$ is a surface function. We need to pick out the smoothest surface (in some sense) that fulfills (22), using the Sobolev space norms as criterion functions. Denote by $V_{f,i}$ the following set:

$$V_{f,i} = \{u \in H^i(\Omega) : \Delta u = f,\ f \in L^2(\Omega)\},\ i = 0, 1, 2, \tag{29}$$

where $H^0(\Omega) = L^2(\Omega)$.

The equality $\Delta u = f$ is to be interpreted in the sense of distributions. i.e.,

Definition 5. A solution of the Poisson equation (22) in Ω is a function $u \in H^1(\Omega)$ such that the following integral identity holds:

$$\int_\Omega u\,\Delta\varphi\,dx = \int_\Omega f\varphi\,dx, \quad \forall\varphi \in C_0^\infty(\Omega).$$

Lemma 1. *$V_{f,i}$ is a closed, convex and nonempty set of $H^i(\Omega)$.*

Proof. The convexity is due to the linearity of Δ. To verify that $V_{f,i}$, $i = 0, 1, 2$, is nonempty it suffices to verify that $V_{f,2}$ is nonempty.

We assume that $\Omega \subset (0, 2\pi)^n$. Extend f by taking $f = 0$ in $(0, 2\pi)^n \backslash \Omega$. Then $V_{f,2}$ contains the function

$$u = f_0|x|^2/(2n) - \sum_{m\neq 0} e^{imx}/|m|^2$$

assuming $f = \sum f_m e^{imx}$ and that m denotes a multi-index. To show that $V_{f,i}$ is closed we select a sequence $\{u_n\}_1^\infty \subset V_{f,i}$, such that $u_n \to u$ in $H^i(\Omega)$. Then $u_n \to u$ in L^2 and, by Cauchy's inequality

$$\Big|\int_\Omega f\,\varphi\,dx - \int_\Omega u\,\Delta\varphi\,dx\Big| = \Big|\int (u_n - u)\,\Delta\varphi\,dx\Big| \leq$$

$$\leq \int_\Omega |u_n - u|^2\,dx \int_\Omega |\Delta\varphi|^2\,dx \to 0, \quad \forall\varphi \in C_0^\infty,$$

i.e. $\int_\Omega f\,\varphi\,dx = \int_\Omega u\,\Delta\varphi\,dx$ and $u \in V_{f,i}$

We recall some facts about Green's formula and of the normal derivatives of H^1 -functions [1]:

Preposition 1 *[1] If $v \in H^1(\Omega)$, $\Delta v \in L^2(\Omega)$ (in the sense of distribution), and $\varphi \in H^1(\Omega)$, then $(\nabla v \cdot \nu) \equiv \partial v/\partial\nu \in H^{-1/2}(\partial\Omega)$ is defined by*

$$\left(\frac{\partial v}{\partial\nu}, \varphi\right)_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} = \int_\Omega \Delta v\,\varphi\,dx + \int_\Omega \nabla v\,\nabla\varphi\,dx.$$

This definition is justified by the fact the right hand site of the last above equality defines a bounded linear functional on $H^1(\Omega)$ and by the the following lemma:

Lemma 2. *[1] If $v \in H^1(\Omega)$ and $\Delta v \in L^2(\Omega)$, then*

$$\int_\Omega \Delta v\, \varphi\, dx + \int_\Omega \nabla v\, \nabla \varphi\, dx = 0 \quad \textit{for all} \quad \varphi \in \overset{\circ}{H}{}^1(\Omega).$$

Proof. Since $C_0^\infty(\Omega)$ is dense in $\overset{\circ}{H}{}^1(\Omega)$, it suffices to prove last equality for all $\varphi \in C_0^\infty(\Omega)$. Then by Green's formula and the definition of distributional derivatives, we have

$$\int_\Omega \Delta v\, \varphi\, dx + \int_\Omega \nabla v\, \nabla \varphi\, dx = \int_\Omega \Delta v\, \varphi\, dx - \int_\Omega v\, \Delta \varphi\, dx.$$

Note also the following well known lemmas for the Dirichlet and Neumann problems in Ω [11], i.e.

Lemma 3. *[11] Suppose $g \in H^{1/2}(\partial\Omega)$, $f \in L^2(\Omega)$. Then there exists a unique function $u \in H^1(\Omega)$ such that*

$$\begin{cases} \Delta u = f & \textit{in} \quad \Omega, \\ u = g & \textit{on} \quad \partial\Omega. \end{cases}$$

Lemma 4. *[11] Suppose that $g \in H^{-1/2}(\partial\Omega)$, $f \in L^2(\Omega)$ and that*

$$(g, 1)_{H^{-1/2}(\partial\Omega), H^{1/2}(\partial\Omega)} = \int_\Omega f\, dx.$$

Then there exists a unique function $u \in H^1(\Omega)$ such that

$$\begin{cases} \Delta u = f \quad \textit{in} \quad \Omega \textit{ (in the sense of distributions)}, \\ u = g \quad \textit{on} \quad \partial\Omega \textit{ (in the sense of Prepos. 1)}, \int_\Omega u\, dx = 0. \end{cases}$$

Let α be a multi-index and $\beta_1 > 0$ a given parameter. We consider the following optimization problems:

$$I_0(u) \equiv \min_{u \in V_{f,0}} \int_\Omega |u|^2\, dx, \tag{30}$$

and

$$I_1(u) \equiv \min_{u \in V_{f,1}} \int_\Omega |u|^2\, dx + \beta_1 \int_\Omega \sum_{|\alpha|=1} |\partial^\alpha u|^2\, dx. \tag{31}$$

Theorem 1. *Problems* (30) *and* (31) *have unique solutions u_0 and u_1, respectively.*

Proof. The proof follows from Lemma 1 and the fact that we are minimizing Hilbert norms.

For problems (30) and (31) we have the following results characterizing the solutions.

Theorem 2. *Let $u_o = \Delta v$. For the solution u_0 of the problems* (30)*, where $v \in H^2(\Omega, \partial\Omega) \cap \overset{\circ}{H}{}^1(\Omega)$ is the unique solution of the Steklov biharmonic problem*

$$\begin{cases} \Delta^2 v = f & \text{in} \quad \Omega, \\ v = \Delta v + \tau \frac{\partial v}{\partial \nu} = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{32}$$

Proof. By a standard variational method, $u_0 \in L^2(\Omega)$ solves problem (30) if and only if $\Delta u_0 = f$ and

$$\int_\Omega u_0 \, \varphi \, dx = 0 \quad \text{for all} \quad \varphi \in L^2(\Omega), \quad \Delta\varphi = 0.$$

Assume first that u_0 solves problem (30). Let v be defined as the unique solution of the Dirichlet problem,

$$\begin{cases} \Delta v = u_0 & \text{in} \quad \Omega, \\ v = 0 & \text{on} \quad \partial\Omega. \end{cases}$$

If φ, $v \in H^1(\Omega)$ and $\Delta\varphi$, $\Delta v \in L^2(\Omega)$, we have the Green formula

$$\int_\Omega \Delta v \, \varphi \, dx - \int_\Omega v \, \Delta\varphi \, dx = \int_{\partial\Omega} (\nabla v \cdot \nu) \, \varphi \, ds - \int_{\partial\Omega} v \, (\nabla\varphi \cdot \nu) \, ds.$$

Now let $\varphi \in H^1(\Omega)$ be a harmonic function, $\Delta\varphi = 0$. Then we have

$$0 = \int_\Omega u_0 \, \varphi \, dx = \int_\Omega \Delta v \, \varphi \, dx =$$

$$\int_{\partial\Omega} (\nabla v \cdot \nu) \, \varphi \, ds - \int_{\partial\Omega} v \, (\nabla\varphi \cdot \nu) \, ds + \int_\Omega v \, \Delta\varphi \, dx,$$

that is

$$\int_{\partial\Omega} (\nabla v \cdot \nu) \, \varphi \, ds = 0 \quad \text{for all such} \quad \varphi.$$

Since there exists a unique function $u \in H^1(\Omega)$ such that

$$\begin{cases} \Delta u = f, \quad f \in L^2(\Omega) & \text{in} \quad \Omega, \\ u = g, \quad g \in H^{1/2}(\partial\Omega) & \text{on} \quad \partial\Omega, \end{cases}$$

and $\varphi\big|_{\partial\Omega}$ may be chosen arbitrary in $H^{1/2}(\Omega)$, we conclude that $0 = (\nabla v \cdot \nu) \in H^{-1/2}(\Omega)$. We have proved that $u_0 = \Delta v \in L^2(\Omega)$, where v satisfies the Steklov biharmonic problem

$$\begin{cases} \Delta^2 v = f & \text{in} \quad \Omega, \\ v = \Delta v + \tau \frac{\partial v}{\partial \nu} = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{33}$$

On the other hand we claim that (33) cannot have more than one solution $v \in H^1(\Omega)$ with $\Delta v \in L^2(\Omega)$. Indeed assume that (33) is satisfied and consider the function $\psi \in L^2(\mathbb{R}^n)$ defined by

$$\psi = \begin{cases} v(x) & \text{if} \quad x \in \Omega, \\ 0 & \text{if} \quad x \notin \Omega. \end{cases}$$

For arbitrary $\varphi \in C_0^\infty(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} \psi \, \Delta\varphi \, dx = \int_\Omega v \, \Delta\varphi \, dx =$$

$$= \int_{\partial\Omega} v \, (\nabla\varphi \cdot \nu) \, ds - \int_{\partial\Omega} (\nabla v \cdot \nu) \, \varphi \, ds + \int_\Omega \varphi \, \Delta v \, dx,$$

i.e.

$$\int_{\mathbb{R}^n} \psi \, \Delta\varphi \, dx = \int_\Omega \varphi \, \Delta v \, dx$$

for all $\varphi \in C_0^\infty(\mathbb{R}^n)$.

Let now $h \in C_0^\infty(\mathbb{R}^n)$ be dined by

$$h(x) = \begin{cases} \Delta v & \text{if} \quad x \in \Omega, \\ 0 & \text{if} \quad x \notin \Omega. \end{cases}$$

We have proved that

$$\Delta\psi = h$$

in the sense of distributions. Using the Fourier transformation it follows that $\psi \in H^2(\mathbb{R}^n)$. Therefore $v \in H^2(\Omega)$, and v must be the unique solution in $H^2(\Omega)$ of (33), being the unique minimizer in $\overset{\circ}{H}{}^2(\Omega)$ of the coercive quadratic functional

$$J(v) \equiv \int_\Omega \Big(\frac{1}{2} |\Delta v|^2 - f v \Big) \, dx.$$

The proof is complete.

Theorem 3. *Let $u_1 = \Delta v$. For the solution u_1 of the problems* (31)*, where $v \in H^2(\Omega), (\nabla v \cdot \nu) = 0$, is the unique solution of the Steklov-type biharmonic problem*

$$\begin{cases} \Delta^2 v = f & \text{in} \quad \Omega, \\ \frac{\partial v}{\partial \nu} = \frac{\partial \Delta v}{\partial \nu} + \tau \, v = 0 & \text{on} \quad \partial\Omega. \end{cases}$$

Proof. Assume that u_1 solves problem (31). Let v be defined as the unique solution in the class $\{\psi \in H^1(\Omega) : \Delta\psi \in H^1(\Omega)\}$ of the following biharmonic problem

$$\begin{cases} \Delta^2 v = f & \text{in} \quad \Omega, \\ v = \beta_1 \Delta v, \quad \nabla v \cdot \nu = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{34}$$

By standard variational method, $u_1 \in H^1(\Omega)$ solves problem (31), if and only if $\Delta u_1 = f$ and

$$\int_\Omega (u_1\,\varphi + \beta_1 \nabla u_1 \cdot \nabla\varphi)\,dx = 0$$

for all $\varphi \in H^1(\Omega)$ and $\Delta\varphi = 0$ in Ω. Taking $\varphi = 1$, we observe that

$$\int_\Omega u_1\,dx = 0.$$

Let $v \in H^1(\Omega)$ be any solution of the Neumann problem

$$\begin{cases} \Delta v = u_1 & \text{in} \quad \Omega, \\ \nabla v \cdot \nu = 0 & \text{on} \quad \partial\Omega. \end{cases} \tag{35}$$

Applying Green's formula we have,

$$0 = \int_\Omega \varphi\,\Delta v\,dx + \beta_1 \int_\Omega \nabla(\Delta v)\nabla\varphi\,dx =$$

$$= \int_\Omega v\,\Delta\varphi\,dx + \int_{\partial\Omega} \varphi\,(\nabla v \cdot \nu)\,ds - \int_{\partial\Omega} v\,(\nabla\varphi \cdot \nu)\,ds +$$

$$+\beta_1 \int_{\partial\Omega} \Delta v\,(\nabla\varphi \cdot \nu)\,ds - \beta_1 \int_\Omega \Delta v\,\Delta\varphi\,dx,$$

i.e.

$$\int_{\partial\Omega} (v - \beta_1 \Delta v)\,(\nabla\varphi \cdot \nu)\,ds = 0$$

for all $\varphi \in H^1(\Omega)$ and $\Delta\varphi = 0$ in Ω.

Since $(\nabla\varphi \cdot \nu) \in H^{-1/2}(\partial\Omega)$ may be chosen arbitrarily (Lemma 4) apart from the condition

$$\int_{\partial\Omega} (\nabla\varphi \cdot \nu)\,ds = 0,$$

it follows that, for some $C = \text{const}$, $v - \beta_1 \Delta v = C$ on $\partial\Omega$.

Now the solution v is uniquely defined up to an additive constant. This constant may be chosen so that $C = 0$.

We have proved that v defined by (35) satisfies (34), and that $v, \Delta v \in H^1(\Omega)$. It remains only to prove that the solution v of (34) is unique. To this end we introduce the function $\varphi = \beta_1 \Delta v - v$ and observe that v satisfies (34) if and only if $(v, \varphi) \in H^1(\Omega) \times \overset{\circ}{H}{}^1(\Omega)$ satisfies the system

$$\begin{cases} \beta_1 \Delta v - v = \varphi, \\ \beta_1 \Delta\varphi + \varphi = \beta_1^2 f - v \quad \text{with} \quad (\nabla v \cdot \nu) = 0 \quad \text{on} \quad \partial\Omega. \end{cases} \tag{36}$$

Next assume that v_1, v_2 satisfy (34), or $(v_1, \varphi_1), (v_2, \varphi_2)$ satisfy (36). From the previous argument it follows that $u_1 = \Delta v_1 = \Delta v_2$ is the unique solution of problem (31) so that $\Delta(v_1 - v_2) = 0$. From (36) we have

$$\begin{cases} \beta_1 \Delta(v_1 - v_2) - (v_1 - v_2) = \varphi_1 - \varphi_2, \\ \beta_1 \Delta(\varphi_1 - \varphi_2) + (\varphi_1 - \varphi_2) = v_1 - v_2. \end{cases}$$

This implies, $v_1 - v_2 = \varphi_1 - \varphi_2$ and $\Delta(\varphi_1 - \varphi_2) = 0$ with $\varphi_1 - \varphi_2 \in \overset{\circ}{H}{}^1(\Omega)$ whence we conclude that $\varphi_1 - \varphi_2 = 0$ and $v_1 - v_2 = 0$. The proof is complete.

We conclude this section by a theorem relating the solution of problems (30) and (31). First we recall the following definition.

Definition 6. $\Omega \subset \mathbb{R}^n$ is called *star-shaped*, if there exists $x_0 \in \Omega$ such that for all $x \in \Omega$ the set $\{t \in \mathbb{R} : \quad x_0 + t(x - x_0) \in \Omega\}$ is an interval.

Theorem 4. *Assume that $\Omega \subset \mathbb{R}^n$ is open, bounded and star-shaped. If $u_{1,\beta_1} \in H^1(\Omega)$ denotes the solution of problem* (31) *with the parameter $\beta_1 > 0$, and if $u_0 \in L^2(\Omega)$ denotes the solution of problem* (30)*, then*

$$u_{1,\beta_1} \to u_0 \quad \textit{in} \quad L^2(\Omega) \quad \textit{as} \quad \beta_1 \to 0+ .$$

Proof. For $0 < \lambda < 1$ and x_0 chosen as in the previous definition, we take

$$\Omega_\lambda = \{x \in \mathbb{R}^n : \ x_0 + \lambda(x - x_0) \in \Omega\},$$

$$u_{0,\lambda}(x) = u_0(x_0 + \lambda(x - x_0)), \quad f_\lambda = f(x_0 + \lambda(x - x_0)).$$

Then [11],

$$\Delta u_{0,\lambda} = f_\lambda \quad \text{in} \quad \Omega_\lambda, \quad \Omega_\lambda \supset \overline{\Omega}, \quad u_{0,\lambda} \in H^2_{loc}(\Omega_\lambda).$$

Since $H^2_{loc}(\Omega_\lambda) \supset H^2(\Omega)$, it follows that $u_{0,\lambda} \in H^2(\Omega)$. Further it is rather easy to see that

$$\int_\Omega |u_{0,\lambda} - u_0|^2 \, dx \to 0,$$

and

$$\int_\Omega |f_\lambda - f|^2 \, dx \to 0 \quad \text{as} \quad \lambda \to 1.$$

Next define $v_\lambda \in \overset{\circ}{H}{}^1(\Omega)$ by

$$\Delta v_\lambda = f - f_\lambda \quad \text{in} \quad \Omega_\lambda.$$

Then

$$\int_\Omega |v_\lambda|^2 \, dx \le ||v_\lambda||_{H^1(\Omega)} \le C \int_\Omega |f - f_\lambda|^2 \, dx.$$

Consequently, taking $w_\lambda = u_{0,\lambda} + v_\lambda$, we have first,

$$w_\lambda \in H^1(\Omega), \quad \Delta w_\lambda = f \quad \text{in} \quad \Omega,$$

and hence,

$$\int_\Omega |w_\lambda - u_0|^2 \, dx \to 0 \quad \text{as} \quad \lambda \to 1.$$

Now, if $\varepsilon > 0$ is given, we may choose a λ close enough to 1, so that

$$\int_\Omega w_\lambda^2 \, dx < \int_\Omega u_0^2 \, dx + \varepsilon/2.$$

Further, by definition,

$$\int_\Omega u_{1,\beta_1}^2\,dx+\beta_1\int_\Omega|\nabla u_{1,\beta_1}|^2\,dx\le\int_\Omega w_\lambda^2\,dx+\beta_1\int_\Omega|\nabla w_\lambda|^2\,dx.$$

Since

$$||w_\lambda||_{H^1(\Omega)}\le C\int_\Omega|f|^2\,dx$$

we have, for sufficiently small β_1,

$$\int_\Omega u_{1,\beta_1}^2\,dx+\beta_1\int_\Omega|\nabla u_{1,\beta_1}|^2\,dx\le\int_\Omega u_0^2\,dx+\varepsilon.$$

It follows that,

$$\lim_{\beta_1\to0^+}\sup\int_\Omega|u_{1,\beta_1}|^2\,dx\le\int_\Omega|u_0|^2\,dx.$$

Further, for some sub-sequence of β_1, we have

$$u_{1,\beta_1}\to\widetilde{u}\quad\text{in}\quad L^2(\Omega)\quad\text{(weakly)},$$

$$\Delta\widetilde{u}=f\quad\text{in}\quad\Omega,$$

and

$$\int_\Omega|\widetilde{u}|^2\,dx\le\lim_{\beta_1\to0^+}\inf\int_\Omega|u_{1,\beta_1}|^2\,dx.$$

But then $\Delta\widetilde{u}=f$ and

$$\int_\Omega|\widetilde{u}|^2\,dx\le\int_\Omega|u_0|^2\,dx$$

which, by definition of u_0, implies that $\widetilde{u}=u_0$. So,

$$u_{1,\beta_1}\to u_0\quad\text{in}\quad L^2(\Omega)\quad\text{(weakly)}.$$

Next

$$\lim_{\beta_1\to0^+}\sup\int_\Omega|u_{1,\beta_1}-u_0|^2\,dx=$$

$$=\lim_{\beta_1\to0^+}\sup\int_\Omega|u_{1,\beta_1}|^2\,dx-2\lim_{\beta_1\to0^+}\int_\Omega u_{1,\beta_1}\,u_0\,dx+$$

$$+\int_\Omega|u_0|^2\,dx\le\int_\Omega|u_0|^2\,dx-2\int_\Omega|u_0|^2\,dx+\int_\Omega|u_0|^2\,dx=0.$$

Finally, since this strong limit u_0 is uniquely defined we may conclude, by a standard argument that

$$u_{1,\beta_1}\to u_0\quad\text{in}\quad L^2(\Omega)\quad\text{as}\quad\beta_1\to0^+$$

without restriction to any subsequence. The proof is complete.

Remark 1. All convex sets are star-shaped. Rectangles Ω appearing in our applications are thus star-shaped.

5 Appendix

Many complex engineering structures, such as the rotor blades of wind turbines and helicopters, are non-prismatic beamlike structures, which may be tapered, twisted and curved in their reference unstressed state and undergo large displacements of the reference centre-line's points, as well as in- and out-of-plane warping of the transverse cross-sections. Continuous efforts to better predict the mechanical behaviour of such structures, which are aimed at improving the performance in terms of structural efficiency and costs effectiveness, offer the opportunity to address some very interesting, challenging problems in the field of continuum and solid mechanics [30].

An important point in developing rigorous yet application-oriented mathematical models for such structures is an appropriate description of their motion. In general, a non-prismatic beamlike structure can be considered as a collection of deformable plane figures (referred to as the transverse cross-sections) along a suitable three-dimensional curve (called the reference centre-line). Each cross-sectional point in the reference state can moreover move to its position in the current state through a global rigid motion on which a local warping motion can be superimposed. The description of the motion of such structures can thus be performed by introducing two kinematic maps, herein called R_A and R_B, to identify the positions of the points of the mentioned structure in the reference and current states, as discussed in [30]–[32]. Specifically, the reference map R_B can be defined as follows

$$R_B(z_i) = R_{0B}(z_1) + x_\alpha(z_i) b_\alpha(z_1)$$

where R_{0B} denotes the position of the centre-line's points in the reference state, b_α are the vectors of the reference local frame in the plane of the reference cross-section, x_α identify the position of the points in the reference cross-section relative to the reference centre-line, and finally, z_i are three independent mathematical variables which do not depend on time. In particular, z_1 is equal to the reference arch-length s, and z_α belong to a bi-dimensional mathematical domain that is used to map the position of the points, x_α, of the cross-sections. Note that in this section α and β assume values 2 and 3, i and j take values $1, 2$ and 3, while repeated indices are summed over their range.

In a similar manner, the current map R_A can be defined as follows

$$R_A(z_i, t) = R_{0A}(z_1, t) + x_\beta(z_i) a_\beta(z_1, t) + w_j(z_1, t) a_j(z_1, t)$$

where t is the considered evolution scalar real parameter (the time, for instance), R_{0A} denotes the position of the centre-line's points in the current state, while w_j are the components of the warping displacement fields with respect to the current local frames referred to as a_j.

These maps can be used to determine the gradient of transformation between the current and reference states and, successively, the Green–Lagrange strain tensor, as discussed in [30]. Given such strain tensor and a constitutive model, it is thus possible to determine the stress fields in the three-dimensional structure. The problem unknowns, such as the displacements of the centre-line's points and the warping fields, can then be determined as the solution

of a set of balance equations deduced by a stationary condition of a suitable energy functional [30]. Specifically, the result of this procedure is a mathematical problem based on partial differential equations (PDEs) with Neumann-type boundary conditions the solution of which enables obtaining all unknowns of the problem, such as the warping fields w_k, the displacements of the centreline's points, the Green–Lagrange strain fields and the corresponding stress fields as well. Further details can be found in [30] and [32].

The results of the article are presented at the 13^{th} International Conference Chaotic Modeling, Simulation and Applications (CHAOS2020, June 9–12, 2020, Florence, Italy).

References

1. Adams R. A. *Sobolev spaces*. Academic Press, New York, 1975.
2. S. Agmon, A. Douglis and L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, I, *Commun. Pure Appl. Math.*, **12**(4) (1959), 623–727.
3. Brock F. An isoperimetric inequality for eigenvalues of the Stekloff problem. *Z. Angew. Math. Mech. (ZAMM)*, **81**:1 (2001), 69–71.
4. Courant R. and Hilbert H. *Methods of Mathematical Physics*. Vol. I. Interscience Publishers, Inc.: New York, 1953.
5. Duffin R. J. The influence of Poisson's ratio on the vibrational spectrum. *SIAM J. Appl. Math.*, **17** (1969), 179–191.
6. Gazzola F. and Sweers G. On positivity for the biharmonic operator under Steklov boundary conditions. *Arch. Rational Mech. Anal.*, **188**:3 (2008), 399–427.
7. Gazzola F., Grunau H.-Ch. and Sweers G. *Polyharmonic Boundary Value Problems: Positivity Preserving and Nonlinear Higher Order Elliptic Equations in Bounded Domains*. Lecture Notes Math. **1991**, Springer-Verlag, 2010.
8. Kondratiev V. A., Kopacek I. and Oleinik O. A. On asymptotic properties of solutions of the biharmonic equation. *Diff. Equations*, **17**:10 (1981), 1886–1899.
9. Kondratiev V. A. and Oleinik O. A. Estimates for solutions of the Dirichlet problem for biharmonic equation in a neighbourhood of an irregular boundary point and in a neighbourhood of infinity Saint-Venant's principle *Proc. Royal Soc. Edinburgh*, **93A**(3–4) (1983), 327–343.
10. Kuttler J R and Sigillito V G 1968, "Inequalities for membrane and Stekloff eigenvalues", *J. Math. Anal. Appl.*, **23**(1), 148–160.
11. Lions J.-L. and Magenes E. *Non–Homogeneous Boundary Value Problems and Applications*. I., Berlin: Springer, 1972.
12. Matevosyan O. A. The exterior Dirichlet problem for the biharmonic equation: Solutions with bounded Dirichlet integral. *Math. Notes*, **70**:3 (2001), 363–377.
13. Matevossian O. A. Solutions of exterior boundary value problems for the elasticity system in weighted spaces. *Sbornik Math.*, **192**:12 (2001), 1763–1798.
14. Matevossian H. A. On solutions of mixed boundary-value problems for the elasticity system in unbounded domains. *Izvestiya Math.*, **67**:5 (2003), 895–929.
15. Matevosyan O. A. Solutions of the Robin problem for the system of elastic theory in external domains. *J. Math. Sci. (N. Y.)*, **197**:3 (2014), 367–394.
16. Matevossian H. A. On solutions of the Dirichlet problem for the polyharmonic equation in unbounded domains. *P-Adic Numbers, Ultrametric Analysis, and Appl.*, **7**:1 (2015), 74–78.

17. Matevosyan O. A. Solution of a mixed boundary value problem for the biharmonic equation with finite weighted Dirichlet integral. *Diff. Equations*, **51**:4 (2015), 487–501.
18. Matevossian O. A. On solutions of the Neumann problem for the biharmonic equation in unbounded domains. *Math. Notes*, **98** (2015), 990–994.
19. Matevosyan O. A. On solutions of the mixed Dirichlet–Navier problem for the polyharmonic equation in exterior domains. *Russ. J. Math. Phys.*, **23**:1 (2016), 135–138.
20. Matevosyan O. A. On solutions of one boundary value problem for the biharmonic equation. *Diff. Equations*, **52**:10 (2016), 1379–1383.
21. Matevossian H. A. On the biharmonic Steklov problem in weighted spaces. *Russ. J. Math. Phys.*, **24**:1 (2017), 134–138.
22. Matevossian H. A. On solutions of the mixed Dirichlet–Steklov problem for the biharmonic equation in exterior domains. *P-Adic Numbers, Ultrametric Analysis, and Appl.*, **9**:2 (2017), 151–157.
23. Matevossian H. A. On the Steklov-type biharmonic problem in unbounded domains. *Russ. J. Math. Phys.*, **25**:2 (2018), 271–276.
24. Matevossian H. A. On the polyharmonic Neumann problem in weighted spaces. *Complex Variables & Elliptic Equations*, **64**:1 (2019), 1–7.
25. Matevossian H. On the Mixed Dirichlet–Steklov-Type and Steklov-Type Biharmonic Problems in Weighted Spaces. *Math. Comput. Appl.*, **24**:1, 25 (2019), 1–9.
26. Matevossian H. A. Mixed Boundary Value Problems for the Elasticity System in Exterior Domains. *Math. Comput. Appl.*, **24**:2 (2019), 58, 1–7.
27. Matevossian H. On the Mixed Steklov–Neumann and Steklov-Type Biharmonic Problems in Unbounded Domains. *IOP Conf. Ser.: Mater. Sci. Eng.*, **683** (2019), no. 1, 012016, 10 pp.
28. Matevossian H. A. On the Mixed Neumann–Robin Problem for the Elasticity System in Exterior Domains. *Russ. J. Math. Phys.*, **27**:2 (2020), 272–276.
29. Matevossian H. A., Nordo G., Sako T. Biharmonic Problems and their Application in Engineering and Medicine. *IOP: Mater. Sci. Eng.*, (2020), 10 pp. (in press).
30. Migliaccio G., Ruta G., et al., Beamlike models for the analyses of curved, twisted and tapered horizontal-axis wind turbine (HAWT) blades undergoing large displacements. *Wind Energ. Sci.*, **5**, 685-698, 2020; https://doi.org/10.5194/wes-5-685-2020.
31. Migliaccio G., Ruta G, et al., Curved and twisted beam models for aeroelastic analysis of wind turbine blades in large displacements. *Proc. of XXIV AIMETA Conference* , 2019, Lecture Notes in Mech. Engin., Springer, Cham, 2020.
32. Migliaccio G., and Ruta G. Rotor blades as curved, twisted, and tapered beam-like structures subjected to large deflections. *Engin. Struct.*, 2020 (in press).
33. Payne L. E. Some isoperimetric inequalities for harmonic functions. *SIAM J. Math. Anal.*, **1**:3 ((1970), 354–359.
34. Provenzano L. A note of Neumann eigenvalues of the biharmonic operator, *Math. Meth. Appl. Sci.*, **41**:3, (2016), 1005-1012. DOI: 10.1002/mma.4063
35. Sobolev S. L. *Some Applications of Functional Analysis in Mathematical Physics*, 3th ed., Nauka, Moscow (1988); *Applications of Functional Analysis in Mathematical Physics*. Amer. Math. Soc., Providence 1991.
36. Verchota G. The biharmonic Neumann problem in Lipschitz domains. *Acta Math. Anal.*, **194** (1970), 217–279.
37. Stekloff W. Sur les problemes fondamentaux de la physique mathematique. *Ann. Sci. de l'E.N.S.*, 3^e serie, **19** (1902), 191-259 et 455-490.

On the Exterior Biharmonic Problem with the Steklov and Steklov-type Boundary Conditions

Hovik A. Matevossian[1,2] and Giovanni Migliaccio[3]

[1] Federal Research Centre "Computer Science & Control" & Steklov Mathematical Institute, Russian Academy of Sciences, Vavilov str., 40-42, R 374, Moscow 119333, Russia
[2] Moscow Avaiation Institute (National Research University), Volokolomskoe shosse, 4, Moscow 125993, Russia
(E-mail: hmatevossian@graduate.org)
[3] Università di Pisa, Largo Luccio Lazzarino 2, 56122 Pisa, Italy
(E-mail: giovanni.migliaccio.it@gmail.com)

Abstract. We study the properties of generalized solutions in unbounded domains and the asymptotic behavior of solutions of elliptic boundary value problems at infinity. Moreover, we study the unique solvability of the mixed biharmonic problem with the Steklov and Steklov-type conditions on the boundary in the exterior of a compact set under the assumption that generalized solutions of this problem has a bounded Dirichlet integral with weight $|x|^a$. Depending on the value of the parameter a, we obtained uniqueness (non-uniqueness) theorems of this problem or present exact formulas for the dimension of the space of solutions.
Keywords: Biharmonic Operator, Steklov and Steklov-type Boundary Conditions, Dirichlet Integral, Weighted Spaces.

1 Introduction

Let Ω be an unbounded domain in $\mathbb{R}^n$, $n \geq 2$, $\Omega = \mathbb{R}^n \setminus \overline{G}$ with the boundary $\partial\Omega \in C^2$, where G is a bounded simply connected domain (or a union of finitely many such domains) in $\mathbb{R}^n$, $0 \in G$, $\overline{\Omega} = \Omega \cup \partial\Omega$ is the closure of Ω, $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ and $|x| = \sqrt{x_1^2 + \dots + x_n^2}$.

In Ω we consider the following problems for the biharmonic equation

$$\Delta^2 u = 0 \tag{1}$$

with the Steklov boundary condition on Γ_1 and the Steklov–type boundary condition on Γ_2

$$u|_{\Gamma_1} = \left(\Delta u + \tau \frac{\partial u}{\partial \nu}\right)\Big|_{\Gamma_1} = 0, \quad \frac{\partial u}{\partial \nu}\Big|_{\Gamma_2} = \left(\frac{\partial \Delta u}{\partial \nu} + \tau\, u\right)\Big|_{\Gamma_2} = 0, \tag{2}$$

where $\overline{\Gamma}_1 \cup \overline{\Gamma}_2 = \partial\Omega$, $\Gamma_1 \cap \Gamma_2 = \emptyset$, $\operatorname{mes}_{n-1} \Gamma_1 \neq 0$, $\nu = (\nu_1, \dots, \nu_n)$ is the outer unit normal vector to $\partial\Omega$, $\tau \in C(\partial\Omega)$, $\tau \geq 0$, $\tau \not\equiv 0$, and $\tau > 0$ on a set of positive $(n-1)$ – dimensional measure on $\partial\Omega$.

Elliptic problems with parameters in the boundary conditions have been called Steklov or Steklov-type problems since their first appearance in [30]. For the biharmonic operator, these conditions were first considered in [1], [10]

and [28], whose authors the isoperimetric properties of the first eigenvalue were studied.

Note that standard elliptic regularity results are available in [4]. The monograph covers higher order linear and nonlinear elliptic boundary value problems, mainly with the biharmonic or polyharmonic operator as leading principal part. The underlying models and, in particular, the role of different boundary conditions are explained in detail. As for linear problems, after a brief summary of the existence theory and L^p and Schauder estimates, the focus is on positivity. The required kernel estimates are also presented in detail.

In [3] and [4], the spectral and positivity preserving properties for the inverse of the biharmonic operator under Steklov and Navier boundary conditions are studied. These are connected with the first Steklov eigenvalue. It is shown that the positivity preserving property is quite sensitive to the parameter involved in the boundary condition. Moreover, positivity of the Steklov boundary value problem is linked with positivity under boundary conditions of Dirichlet and Navier type.

In [2], the boundary value problems for the biharmonic equation and the Stokes system are studied in a half space, and, using the Schwartz reflection principle in weighted L^q -space, the uniqueness of solutions of the Stokes system or the biharmonic equation is proved.

As is well known, if Ω is an unbounded domain, one should additionally characterize the behavior of the solution at infinity. As a rule, to this end, one usually poses either the condition that the Dirichlet (energy) integral is finite or a condition on the character of vanishing of the modulus of the solution as $|x| \to \infty$. Such conditions at infinity are natural and were studied by several authors (e.g., [8], [9]).

In the present note, this condition is the boundedness of the weighted Dirichlet integral:

$$D_a(u, \Omega) \equiv \int_\Omega |x|^a \sum_{|\alpha|=2} |\partial^\alpha u|^2 \, dx < \infty, \quad a \in \mathbb{R}.$$

In various classes of unbounded domains with finite weighted Dirichlet (energy) integral, one of the author [11]–[24] studied uniqueness (non–uniqueness) problem and found the dimensions of the spaces of solutions of boundary value problems for the elasticity system and the biharmonic (polyharmonic) equation.

By developing an approach based on the use of Hardy type inequalities [6], [8], [9], in the present note, we obtain a uniqueness (non–uniqueness) criterion for a solution of the mixed biharmonic problem with the Steklov and Steklov-type boundary conditions. To construct the solution, we use a variational method, that is, we minimize the corresponding functional in the class of admissible functions.

Notation: $C_0^\infty(\Omega)$ is the space of infinitely differentiable functions in Ω with compact support in Ω.

We denote by $H^m(\Omega, \Gamma)$, $\Gamma \subset \overline{\Omega}$, the Sobolev space of functions in Ω obtained by the completion of $C^\infty(\overline{\Omega})$ vanishing in a neighborhood of Γ with

respect to the norm

$$||u; H^m(\Omega, \Gamma)|| = \left(\int_\Omega \sum_{|\alpha| \le m} |\partial^\alpha u|^2 dx \right)^{1/2}, \quad m = 1, 2,$$

where $\partial^\alpha \equiv \partial^{|\alpha|}/\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}$, $\alpha = (\alpha_1, \dots, \alpha_n)$ is a multi-index, $\alpha_i \ge 0$ are integers, and $|\alpha| = \alpha_1 + \dots + \alpha_n$; if $\Gamma = \emptyset$, we denote $H^m(\Omega, \Gamma)$ by $H^m(\Omega)$.

$\overset{\circ}{H}{}^m(\Omega)$ is the space obtained by the completion of $C_0^\infty(\Omega)$ with respect to the norm $||u(x); H^m(\Omega)||$;

$\overset{\circ}{H}{}^m_{loc}(\Omega)$ is the space obtained by the completion of $C_0^\infty(\Omega)$ with respect to the family of semi-norms

$$\|u; H^m(\Omega \cap B_0(R))\| = \left(\int\limits_{\Omega \cap B_0(R)} \sum_{|\alpha| \le m} |\partial^\alpha u|^2 \, dx \right)^{1/2}$$

for all open balls $B_0(R) := \{x : |x| < R\}$ in $\mathbb{R}^n$ for which $\Omega \cap B_0(R) \ne \emptyset$.

Let $\binom{n}{k}$ be the (n, k) - binomial coefficient, $\binom{n}{k}$=0 for $k > n$.

2 Definitions and auxiliary statements

Definition 1. A solution of the homogenous biharmonic equation (1) in Ω is a function $u \in H^2_{loc}(\Omega)$ such that, for every function $\varphi \in C_0^\infty(\Omega)$, the following integral identity holds:

$$\int_\Omega \Delta u \, \Delta\varphi \, dx = 0.$$

Lemma 1. *Let u be a solution of equation (1) in Ω such that $D_a(u, \Omega) < \infty$. Then*

$$u(x) = P(x) + \sum_{\beta_0 < |\alpha| \le \beta} \partial^\alpha \Gamma(x) C_\alpha + u^\beta(x), \quad x \in \Omega, \tag{3}$$

where $P(x)$ is a polynomial, $\operatorname{ord} P(x) < m_0 = \max\{2, 2 - n/2 - a/2\}$, $\beta_0 = 2 - n/2 + a/2$, *$\Gamma(x)$ is the fundamental solution of equation (1), $C_\alpha =$* const, *$\beta \ge 0$ is an integer, and the function u^β satisfies the estimate:*

$$|\partial^\gamma u^\beta(x)| \le C_{\gamma\beta} |x|^{3-n-\beta-|\gamma|}, \; C_{\gamma\beta} = \text{const},$$

for every multi-index γ.

Remark 1. As is known [29], the fundamental solution $\Gamma(x)$ of the biharmonic equation has the form

$$\Gamma(x) = \begin{cases} C|x|^{4-n}, \; if \, 4 - n < 0 \text{ or } n \text{ is odd}, \\ C|x|^{4-n} \ln|x|, \; if \, 4 - n \ge 0 \text{ and } n \text{ is even}. \end{cases}$$

Proof of Lemma 1 Consider the function $v(x) = \theta_N(x)u(x)$, where $\theta_N(x) = \theta(|x|/N), \theta \in C^\infty(\mathbb{R}^n),\ 0 \le \theta \le 1,\ \theta(s) = 0$ for $s \le 1,\ \theta(s) = 1$ for $s \ge 2$, while $N \gg 1$ and $G \subset \{x : |x| < N\}$. We extend v to $\mathbb{R}^n$ by setting $v = 0$ on $G = \mathbb{R}^n \setminus \overline{\Omega}$.

Then the function v belongs to $C^\infty(\mathbb{R}^n)$ and satisfies the equation

$$\Delta^2 v = f,$$

where $f \in C_0^\infty(\mathbb{R}^n)$ and $\operatorname{supp} f \subset \{x : |x| < 2N\}$. It is easy to see that $D_a(v, \mathbb{R}^n) < \infty$.

We can now use Theorem 1 of [7] since it is based on Lemma 2 of [7], which imposes no constraint on the *sign* of σ. Hence, the expansion

$$v(x) = P(x) + \sum_{\beta_0 < |\alpha| \le \beta} \partial^\alpha \Gamma(x) C_\alpha + v^\beta(x),$$

holds for each a, where $P(x)$ is a polynomial of order $\operatorname{ord} P(x) < m_0 = \max\{2, 2 - n/2 - a/2\}$, $\beta_0 = 2 - n/2 + a/2$, $C_\alpha = \text{const}$ and

$$|\partial^\gamma v^\beta(x)| \le C_{\gamma\beta}|x|^{3-n-\beta-|\gamma|}, \quad C_{\gamma\beta} = \text{const}\,.$$

Therefore, by the definition of v, we obtain (3). The proof of Lemma 1 is complete.

3 Main Results

Definition 2. By a solution of the mixed boundary value problem (1), (2) we mean a function $u \in H^2_{loc}(\Omega) \cap \overset{\circ}{H}{}^1_{loc}(\Omega, \Gamma_1)$, $\partial u/\partial\nu = 0$ on Γ_2, such that, for every function $\varphi \in C_0^\infty(\mathbb{R}^n) \cap \overset{\circ}{H}{}^1_{loc}(\Omega, \Gamma_1)$, $\partial\varphi/\partial\nu = 0$ on Γ_2, the following integral identity holds:

$$\int_\Omega \Delta u\, \Delta\varphi\, dx + \int_{\Gamma_1} \tau\, \nabla u\, \nabla\varphi\, ds - \int_{\Gamma_2} \tau\, u\, \varphi\, ds = 0. \tag{4}$$

Theorem 1. *The mixed problem (1),(2) with the condition $D(u, \Omega) < \infty$ has $n + 1$ linearly independent solutions.*

Proof. For any nonzero vector A in $\mathbb{R}^n$, we construct a generalized solution u_A of the biharmonic equation (1) with the boundary conditions

$$\begin{gathered} u_A(x)\big|_{\Gamma_1} = (Ax)\big|_{\Gamma_1},\ \left(\Delta u_A + \tau\frac{\partial u_A(x)}{\partial\nu}\right)\Big|_{\Gamma_1} = \tau\frac{\partial(Ax)}{\partial\nu}\Big|_{\Gamma_1}, \\ \frac{\partial u_A}{\partial\nu}\Big|_{\Gamma_2} = \left(\frac{\partial\Delta u_A}{\partial\nu} + \tau\, u_A\right)\Big|_{\Gamma_2} = 0, \end{gathered} \tag{5}$$

and the condition

$$\chi(u_A,\Omega)\equiv\begin{cases}\displaystyle\int_\Omega\left(\frac{|u_A|^2}{|x|^4}+\frac{|\nabla u_A|^2}{|x|^2}+|\nabla\nabla u_A|^2\right)dx<\infty\\ \qquad\qquad\qquad\qquad\qquad\qquad\text{for } n>4,\\ \displaystyle\int_\Omega\left(\frac{|u_A|^2}{||x|^2\ln|x||^2}+\frac{|\nabla u_A|^2}{||x|\ln|x||^2}+|\nabla\nabla u_A|^2\right)dx<\infty\\ \qquad\qquad\qquad\qquad\qquad\qquad\text{for } 2\le n\le 4,\end{cases}\tag{6}$$

for $A,x\in\mathbb{R}^n$, where Ax denotes the standard scalar product of A and x.

Such a solution of problem (1), (5) can be constructed by the variational method [29], minimizing the functional

$$\Phi(v)=\frac{1}{2}\int_\Omega|\Delta v|^2\,dx$$

in the class of admissible functions $\{v\colon v\in H^2(\Omega),\ v(x)\big|_{\Gamma_1}=(Ax)\big|_{\Gamma_1},\ \left(\Delta v+\tau\frac{\partial v(x)}{\partial\nu}\right)\Big|_{\Gamma_1}=\tau\frac{\partial(Ax)}{\partial\nu}\Big|_{\Gamma_1},\ v$ is compactly supported in $\overline{\Omega}\}$.

The validity of condition (6) as a consequence of the Hardy inequality follows from the results in [8], [9].

Now, for any arbitrary number $e\neq 0$, we construct a generalized solution u_e of equation (1) with the boundary conditions

$$u_e\big|_{\Gamma_1}=e,\ \left(\Delta u_e+\tau\frac{\partial u_e}{\partial\nu}\right)\Big|_{\Gamma_1}=0,\ \frac{\partial u_e}{\partial\nu}\Big|_{\Gamma_2}=\left(\frac{\partial\Delta u_e}{\partial\nu}+\tau\,u_e\right)\Big|_{\Gamma_2}=0,\tag{7}$$

and the condition

$$\chi(u_e,\Omega)\equiv\begin{cases}\displaystyle\int_\Omega\left(\frac{|u_e|^2}{|x|^4}+\frac{|\nabla u_e|^2}{|x|^2}+|\nabla\nabla u_e|^2\right)dx<\infty\\ \qquad\qquad\qquad\qquad\qquad\qquad\text{for } n>4,\\ \displaystyle\int_\Omega\left(\frac{|u_e|^2}{||x|^2\ln|x||^2}+\frac{|\nabla u_e|^2}{||x|\ln|x||^2}+|\nabla\nabla u_e|^2\right)dx<\infty\\ \qquad\qquad\qquad\qquad\qquad\qquad\text{for } 2\le n\le 4.\end{cases}\tag{8}$$

The solution of problem (1), (7) also is constructed by the variational method with the minimization of the corresponding functional in the class of admissible functions $\{v:v\in H^2(\Omega),\ v\big|_{\Gamma_1}=e,\ \left(\Delta v+\tau\frac{\partial v}{\partial\nu}\right)\big|_{\Gamma_1}=0,\ \ v$ is compactly supported in $\overline{\Omega}\}$.

The condition (8) as a consequence of the Hardy inequality follows from the results in [8],[9].

Consider the function $v=(u_A-Ax)-(u_e-e)$.

Obviously, v is a solution of problem (1), (2):

$$\Delta^2 v=0,\quad x\in\Omega,$$

$$v\big|_{\Gamma_1}=\left(\Delta v+\tau\frac{\partial v}{\partial\nu}\right)\Big|_{\Gamma_1}=0,\quad\frac{\partial v}{\partial\nu}\Big|_{\Gamma_2}=\left(\frac{\partial\Delta v}{\partial\nu}+\tau\,v\right)\Big|_{\Gamma_2}=0.$$

One can easily see that $v \not\equiv 0$ and $D(v, \Omega) < \infty$.

To each nonzero vector $\mathbf{A} = (A_0, A_1, \dots, A_n)$ in $\mathbb{R}^{n+1}$, there corresponds a nonzero solution $v_{\mathbf{A}} = (v_{A_0}, v_{A_1}, \dots, v_{A_n})$ of problem (1), (2) with the condition $D(v_{\mathbf{A}}, \Omega) < \infty$, and moreover,

$$v_{\mathbf{A}} = u_A - u_e - Ax + e.$$

Let $A_0, A_1, \dots, A_n$ be a basis in $\mathbb{R}^{n+1}$. Let us prove that the corresponding solutions $v_{A_0}, v_{A_1}, \dots, v_{A_n}$ are linearly independent. Let

$$\sum_{i=0}^{n} C_i v_{A_i} \equiv 0, \qquad C_i = \text{const}\,.$$

Set $W \equiv \sum_{i=1}^{n} C_i A_i x - C_0 e$. We have

$$W = \sum_{i=1}^{n} C_i u_{A_i} - C_0 u_e,$$

$$\int_\Omega |x|^{-2} |\nabla W|^2 \, dx < \infty, \quad n > 4,$$

$$\int_\Omega ||x| \ln |x||^{-2} |\nabla W|^2 \, dx < \infty, \quad 2 \le n \le 4.$$

Let us show that

$$W \equiv \sum_{i=1}^{n} C_i A_i x - C_0 e \equiv 0.$$

Let $T = \sum_{i=0}^{n} C_i A_i = (t_0, \dots, t_n)$, where $A_0 = -e$. Then

$$\int_\Omega |x|^{-2} |\nabla W|^2 \, dx = \int_\Omega |x|^{-2} (t_1^2 + \dots + t_n^2) \, dx = \infty, \quad n > 4,$$

$$\int_\Omega ||x| \ln |x||^{-2} |\nabla W|^2 \, dx = \int_\Omega ||x| \ln |x||^{-2} (t_1^2 + \dots + t_n^2) \, dx = \infty, \quad 2 \le n \le 4,$$

if $T \neq 0$.

Consequently, $T = \sum_{i=0}^{n} C_i A_i = 0$, and since the vectors $A_0, A_1, \dots, A_n$ are linearly independent, we obtain $C_i = 0$, $i = 0, 1, \dots, n$.

Thus, the mixed problem (1), (2) with the condition $D(u, \Omega) < \infty$ has at least $n + 1$ linearly independent solutions.

Let us prove that each solution u of problem (1), (2) with the condition $D(u, \Omega) < \infty$ can be represented as a linear combination of the functions $v_{A_0}, v_{A_1}, \dots, v_{A_n}$, i.e.

$$u = \sum_{i=0}^{n} C_i v_{A_i}, \qquad C_i = \text{const}\,.$$

Since $A_0, A_1, \dots, A_n$ is a basis in $\mathbb{R}^{n+1}$, it follows that there exists constants $C_0, C_1, \dots, C_n$ such that

$$A = \sum_{i=0}^{n} C_i A_i.$$

We set

$$u_0 \equiv u - \sum_{i=0}^{n} C_i v_{A_i}.$$

Obviously, the function u_0 is a solution of problem (1), (2), and $D(u_0, \Omega) < \infty$, $\chi(u_0, \Omega) < \infty$.

Let us show that $u_0 \equiv 0$, $x \in \Omega$. To this end, we substitute the function $\varphi(x) = u_0(x)\theta_N(x)$ into the integral identity (4) for the function u_0, where $\theta_N(x) = \theta(|x|/N)$, $\theta \in C^\infty(\mathbb{R})$, $0 \le \theta \le 1$, $\theta(s) = 0$ for $s \ge 2$ and $\theta(s) = 1$ for $s \le 1$; then we obtain

$$\begin{aligned} \int_\Omega (\Delta u_0)^2 \theta_N(x)\, dx + \int_{\Gamma_1} \tau\, |\nabla u_0|^2 \theta_N(x)\, ds - \int_{\Gamma_2} \tau\, |u_0|^2 \theta_N(x)\, ds \\ = -J_1(u_0) - J_2(u_0) - J_3(u_0), \end{aligned} \tag{9}$$

where

$$J_1(u_0) = 2\int_\Omega \Delta u_0\, \nabla u_0\, \nabla\theta_N(x)\, dx, \quad J_2(u_0) = \int_\Omega u_0\, \Delta u_0\, \Delta\theta_N(x)\, dx,$$

$$J_3(u_0) = \int_{\Gamma_1} u_0\, \nabla u_0\, \nabla\theta_N(x)\, ds.$$

By applying the Cauchy–Schwarz inequality and by taking into account the conditions $D(u_0, \Omega) < \infty$ and $\chi(u_0, \Omega) < \infty$, one can easily show that $J_1(u_0) \to 0$, $J_2(u_0) \to 0$ and $J_3(u_0) \to 0$ as $N \to \infty$. Consequently, by passing to the limit as $N \to \infty$ in (9), we obtain

$$\int_\Omega (\Delta u_0)^2\, \theta_N(x)\, dx + \int_{\Gamma_1} \tau\, |\nabla u_0|^2 \theta_N(x)\, ds - \int_{\Gamma_2} \tau\, |u_0|^2 \theta_N(x)\, ds \to 0.$$

Using the integral identity

$$\int_\Omega (\Delta u_0)^2\, dx + \int_{\Gamma_1} \tau\, |\nabla u_0|^2\, ds - \int_{\Gamma_2} \tau\, |u_0|^2\, ds = 0,$$

we find that if u_0 is a solution of the homogeneous problem (1), (2), then $\Delta u_0 = 0$. Therefore, we have

$$\Delta u_0 = 0, \quad x \in \Omega,$$

$$u_0\big|_{\Gamma_1} = \left(\Delta u_0 + \tau \frac{\partial u_0}{\partial \nu}\right)\bigg|_{\Gamma_1} = 0, \quad \frac{\partial u_0}{\partial \nu}\bigg|_{\Gamma_2} = \left(\frac{\partial \Delta u_0}{\partial \nu} + \tau\, u_0\right)\bigg|_{\Gamma_2} = 0.$$

Hence, it follows [5, Ch.2] that $u_0 = 0$ in Ω. The relation

$$\int_{\partial\Omega} \tau\, (|\nabla u_0|^2 + |u_0|^2)\, ds = 0$$

implies that $u_0 \equiv 0$ on a set of a positive measure on $\partial\Omega$. The proof of the theorem is complete.

Theorem 2. *The mixed problem (1), (2) with the condition* $D_a(u,\Omega) < \infty$ *has:*

(i) the trivial solution for $n-2 \le a < \infty$, $n > 4$*;*

(ii) n *linearly independent solutions for* $n-4 \le a < n-2$, $n > 4$*;*

(iii) $n+1$ *linearly independent solutions for* $-n \le a < n-4$, $n > 4$*;*

(iv) $k(r,n)$ *linearly independent solutions for* $-2r+2-n \le a < -2r+4-n$, $r > 1$, $n > 4$, *where*

$$k(r,n) = \binom{r+n}{n} - \binom{r+n-4}{n}.$$

The proof of Theorem 2 is based on Lemma 1 about the asymptotic expansion of the solution of the biharmonic equation and the Hardy type inequalities for unbounded domains [8],[9]. In case (iv), we need to determine the number of linearly independent solutions of the biharmonic equation (1), the degree of which not exceed the fixed number.

It is well know that the dimension of the space of all polynomials in $\mathbb{R}^n$ of degree $\le r$ is equal $\binom{r+n}{n}$ [27]. Then the dimension of the space of all biharmonic polynomials in $\mathbb{R}^n$ of degree $\le r$ is equal to

$$\binom{r+n}{n} - \binom{r+n-4}{n},$$

since the biharmonic equation is the vanishing of some polynomial of degree $r-4$ in $\mathbb{R}^n$. If we denote by $k(r,n)$ the number of linearly independent polynomial solutions of equation (1) whose degree do not exceed r and by $l(r,n)$ the number of linearly independent homogeneous polynomials of degree r, that are solutions of equation (1), then

$$k(r,n) = \sum_{s=0}^{r} l(s,n),$$

where

$$l(s,n) = \binom{s+n-1}{n-1} - \binom{s+n-5}{n-1}, \quad s > 0.$$

Further, we prove that the mixed problem (1), (2) with the condition $D_a(u,\Omega) < \infty$ for $-2r+2-n \le a < -2r+4-n$ has equally $k(r,n)$ of linearly independent solutions.

4 Appendix

Many complex engineering structures, such as the rotor blades of wind turbines and helicopters, are non-prismatic beamlike structures, which may be tapered, twisted and curved in their reference unstressed state and undergo large displacements of the reference centre-line's points, as well as in- and out-of-plane warping of the transverse cross-sections. Continuous efforts to better predict the mechanical behaviour of such structures, which are aimed at improving the

performance in terms of structural efficiency and costs effectiveness, offer the opportunity to address some very interesting, challenging problems in the field of continuum and solid mechanics [25].

An important point in developing rigorous yet application-oriented mathematical models for such structures is an appropriate description of their motion. The description of the motion of such structures can thus be performed by introducing two kinematic maps, herein called R_A and R_B, to identify the positions of the points of the mentioned structure in the reference and current states, as discussed in [25], [26].

These maps can be used to determine the gradient of transformation between the current and reference states and, successively, the Green–Lagrange strain tensor, as discussed in [25]. Given such strain tensor and a constitutive model, it is thus possible to determine the stress fields in the three-dimensional structure. The problem unknowns, such as the displacements of the centre-line's points and the warping fields, can then be determined as the solution of a set of balance equations deduced by a stationary condition of a suitable energy functional [25]. As shown in [25]and [26], the results of this procedure is a mathematical problem based on partial differential equations (PDEs) with Neumann-type boundary conditions the solution of which enable obtaining all unknowns of the problem, such as the warping fields w_k, the displacements of the centre-line's points, the Green–Lagrange strain fields and the corresponding stress fields.

The results of the article are presented at the 13^{th} International Conference Chaotic Modeling, Simulation and Applications (CHAOS2020, June 9–12, 2020, Florence, Italy).

References

1. Brock F. An isoperimetric inequality for eigenvalues of the Stekloff problem. *Z. Angew. Math. Mech. (ZAMM)*, **81**:1 (2001), 69–71.
2. Farwig R. A note on the reflection principle for the biharmonic equation and the Stokes system. *Acta Appl. Math.*, **34** (1994), 41–51.
3. Gazzola F. and Sweers G. On positivity for the biharmonic operator under Steklov boundary conditions. *Arch. Rational Mech. Anal.*, **188**:3 (2008), 399–427.
4. Gazzola F., Grunau H.-Ch. and Sweers G. *Polyharmonic Boundary Value Problems: Positivity Preserving and Nonlinear Higher Order Elliptic Equations in Bounded Domains*. Lecture Notes Math. **1991**, Springer-Verlag, 2010.
5. Gilbarg D. and Trudinger N. *Elliptic Partial Differential Equations of Second Order*. Berlin: Springer-Verlag, 1977.
6. Egorov Yu. V. and Kondratiev V. A. *On Spectral Theory of Elliptic Operators*. Basel: Birkhauser, 1996.
7. Kondratiev V. A. and Oleinik O. A. On the behavior at infinity of solutions of elliptic systems with a finite energy integral. *Arch. Rational Mech. Anal.*, **99**:1 (1987), 75–99.
8. Kondrat'ev V. A. and Oleinik O. A. Boundary value problems for the system of elasticity theory in unbounded domains. Korn's inequalities. *Russ. Math. Surveys*, **43**(5) (1988), 65–119.
9. Kondratiev V. A. and Oleinik O. A. Hardy's and Korn's inequality and their application. *Rend. Mat. Appl.* Serie VII. **10** (1990), 641–666.

10. Kuttler J. R. and Sigillito V. G. Inequalities for membrane and Stekloff eigenvalues. *J. Math. Anal. Appl.*, **23**(1) (1968), 148–160.
11. Matevosyan O. A. The exterior Dirichlet problem for the biharmonic equation: Solutions with bounded Dirichlet integral. *Math. Notes*, **70**:3 (2001), 363–377.
12. Matevossian O. A. Solutions of exterior boundary value problems for the elasticity system in weighted spaces. *Sbornik Math.*, **192**:12 (2001), 1763–1798.
13. Matevossian H. A. On solutions of mixed boundary-value problems for the elasticity system in unbounded domains. *Izvestiya Math.*, **67**:5 (2003), 895–929.
14. Matevossian H. A. On solutions of the Dirichlet problem for the polyharmonic equation in unbounded domains. *P-Adic Numbers, Ultrametric Analysis, and Appl.*, **7**:1 (2015), 74–78.
15. Matevosyan O. A. Solution of a mixed boundary value problem for the biharmonic equation with finite weighted Dirichlet integral. *Diff. Equations*, **51**:4 (2015), 487–501.
16. Matevossian O. A. On solutions of the Neumann problem for the biharmonic equation in unbounded domains. *Math. Notes*, **98** (2015), 990–994.
17. Matevosyan O. A. On solutions of the mixed Dirichlet–Navier problem for the polyharmonic equation in exterior domains. *Russ. J. Math. Phys.*, **23**:1 (2016), 135–138.
18. Matevosyan O. A. On solutions of one boundary value problem for the biharmonic equation. *Diff. Equations*, **52**:10 (2016), 1379–1383.
19. Matevossian H. A. On the biharmonic Steklov problem in weighted spaces. *Russ. J. Math. Phys.*, **24**:1 (2017), 134–138.
20. Matevossian H. A. On the Steklov-type biharmonic problem in unbounded domains. *Russ. J. Math. Phys.*, **25**:2 (2018), 271–276.
21. Matevossian H. A. On the polyharmonic Neumann problem in weighted spaces. *Complex Variables & Elliptic Equations*, **64**:1 (2019), 1–7.
22. Matevossian H. A. Mixed Boundary Value Problems for the Elasticity System in Exterior Domains. *Math. Comput. Appl.*, **24**:2 (2019), 58, 1–7.
23. Matevossian H. On the Mixed Steklov–Neumann and Steklov-Type Biharmonic Problems in Unbounded Domains. *IOP Conf. Ser.: Mater. Sci. Eng.*, **683** (2019), no. 1, 012016, 10 pp.
24. Matevossian H. A. On the Mixed Neumann–Robin Problem for the Elasticity System in Exterior Domains. *Russ. J. Math. Phys.*, **27**:2 (2020), 272–276.
25. Migliaccio G., Ruta G., et al., Beamlike models for the analyses of curved, twisted and tapered horizontal-axis wind turbine (HAWT) blades undergoing large displacements. *Wind Energ. Sci.*, **5**, 685-698, 2020; https://doi.org/10.5194/wes-5-685-2020.
26. Migliaccio G., and Ruta G. Rotor blades as curved, twisted, and tapered beamlike structures subjected to large deflections. *Engineering Structures*, 2020 (in press).
27. Mikhlin S. G. *Linear Partial Differential Equations*. Vyssaya Shkola, Moscow, 1977 (in Russian).
28. Payne L. E. Some isoperimetric inequalities for harmonic functions. *SIAM J. Math. Anal.*, **1**(3), 354–359, 1970.
29. Sobolev S. L. *Some Applications of Functional Analysis in Mathematical Physics*, 3th ed., Nauka, Moscow, 1988; *Applications of Functional Analysis in Mathematical Physics*. Amer. Math. Soc., Providence 1991.
30. Stekloff W. Sur les problemes fondamentaux de la physique mathematique. *Ann. Sci. de l'E.N.S.*, 3^e serie, **19**, 191–259 et 455–490, 1902.

Prediction of qualitative dynamics in population models through Holling's functional responses

Ashutosh Maurya* and Anupam Priyadarshi

Department of Mathematics, Institute of Science, Banaras Hindu University, Varanasi, India
(E-mail: ashunano@gmail.com)

Abstract. Natural ecosystems are complex network of biotic and abiotic interactions of species and their biological, physical and chemical constituents. Over the century, mathematical models have played primary roles in understanding the mystery behind the ecosystems processes and interesting dynamics of natural ecosystems. The di-trophic predator-prey interactions are the basic building blocks for complex food web models (multiple trophic interactions). The pioneering work of Lotka-Volterra (1926), explaining the abrupt deviations in species abundance and existence of oscillations in a simple predator-prey interaction. Studies in previous decades, it has been assumed that the functional responses are the main cause for chaotic and non-chaotic behavior. In this paper, we investigate how functional responses affect the system dynamics by using its different combinations in a simple two prey-one predator population model. Based on our present investigation, we concluded that the stabilizing properties of functional responses dominate oscillatory behavior.

Keywords: Controlling Limit cycles, Two parameter bifurcation comparison, Holling type Functional Responses, stabilizing property, dominating oscillatory property.

1 Introduction

Food webs in ecosystems play a vital role to regulate species coexistence, species interactions and carbon, nitrogen cycles in natural ecosystems. They often assumed to be responsible for species control and ecosystem balance. The complex networks of biotic (predator-prey) and abiotic interactions in food webs are ubiquitously exist within trophic levels. Among the biotic interactions, predator-prey interactions exhibit very complex dynamics and hence, attracted the attention of theoretical, experimental and field ecologists from the last century. Experiments and field observations are the basis for the development of theoretical concepts of ecological process through mathematical models. These mathematical models interpret the model dynamics using differential equations and predict the future happening based on their results [16]. The theoretical development of mathematical modeling assumed to be started from the pioneering work of Lotka and Volterra (1926) [17,22], which explained the abrupt deviations in species abundance and predicted the existence of oscillations in a simple predator-prey model. The di-trophic predator-prey interactions are the basic building blocks for complex food web models of multiple trophic interactions [16]. The food-web systems exhibit more complex dynamics (quasi-periodicity, chaos) on increasing the number of species in its chain

[1–3] which cannot be obtained by di-trophic food chain models [8].
Growth rates (prey increment with time [3,4] and *functional responses* (the relationship between prey-capturing by per predator in unit time with prey density), frequently used in theoretical predator-prey interactions, are the two important factors to control the overall dynamics of food webs. The term functional response (FR) was first acknowledged by Soloman (1949)[12] and extensively deliberated by C.S. Holling using the *Disc equations* based on predator's handling (capturing, eating and digesting) ability [1–3]. Based on terrestrial experiments and artificial lab experiments [1–3], C.S. Holling proposed three types of functional responses namely Holling types I, II and III. Later on, Holling type IV FR is coined and identified by J.F. Andrews [7] in the culture of micro-organisms which is nonlinear and density dependent function. There are several other FRs such as Leslie-Gower, Ivlev etc., which are also frequently used in predator-prey interactions besides Holling type FR. The Holling type I (H_I) FR is the simplest, in which predator's capture rate increases directly proportional to prey density till saturation. H_IFR has been used in Lotka-Volterra's predator-prey model which produced neutral stable limit cycles [17,22]. Experimentally, it is observed that the sea star's predation traits on juvenile scallops in aquatic region are H_I [10,11]. It is also used in the filter-feeding zooplankton harvesting models. It's mathematical description is $k(x, y) = min(w_1x, w_1)$. The Holling type II (H_{II}) is also similar in the sense that the rate of capture prey increases with increasing prey density but its saturation reached out slowly in comparison of H_I. It is frequently used in the population estimation of insects and parasitoids [17,6]. It's mathematical description is $k(x, y) = \frac{w_1x}{1+w_1x+w_2y}$. The Holling type III ($H_{III}$) exhibits S-shaped in which at low prey density, the capture rate exceeds and goes to saturation gradually like (H_{II}). It is widely used on population estimation of vertebrates [17,6]. Its mathematical description is $k(x, y) = \frac{w_1x^2}{1+w_1x^2+w_2y^2}$. The Holling type IV ($H_{IV}$) is dome-shaped and non-linear which is further proposed and modified by Sokol and Howell (1980) [23] and used for the population estimation on micro-organisms and mice estimations [6]. It's mathematical description is $k(x, y) = \frac{w_1x}{1+w_1x^2+w_2y^2}$. The predator behaviour in the natural ecosystem is strongly associated with prey availability and associated functional response by which predator predates on prey [3,5,8,14,19,20]. The functional response may induce oscillations in di-trophic predator-prey interaction (or chaos) in multi-trophic systems [6,13,21]. In food web systems, the predation rate associated with prey may vary due to change in physical habitat of species.
Studies of Holling's FR convey that the H_I gives the point stability and extinction while H_{II} and H_{IV} give periodic solution along with point stability and extinction with suitable parameter choices. H_{III} gives stability with persistence throughout parameter choice (see fig 1). We have extended their studies with the combined effects of different function responses in food web systems. We consider a simple food web consisting of two bottom prey and one predator on which combination of different functional responses are applied to infer the qualitative behaviour. We have studied the following questions: (1) Is the combination of functional responses gives oscillatory solutions with point stability and extinction? (2) What are the basic properties of this type of model? (3)

What are the detailed of the dynamical behaviour of these systems? (3) What are the two-parameter bifurcation analysis of these system?

2 Background (Terminology)

Assuming x and y prey densities and z predator density, a simple food web model of one predator and two prey can be expressed with the following system of differential equations:

$$\begin{aligned} x' &= xg(x) - zk(x,y) \\ y' &= yi(y) - zj(x,y) \\ z' &= h(x,y,z) \end{aligned} \tag{1}$$

Here, prey x grows logistically in the absence of predator z as follows:

$$g(x) = (1-x) \text{ such that } g(0) = 1 > 0, g_x = -1, \forall x \geq 0 \text{ and } g(1) = 0$$
$$i(y) = w_4(1-y) \text{ such that } i(0) = w_4, i_y = -w_4, \forall y \geq 0 \text{ and } i(1) = 0$$

Where w_4 represent the maximum growth coefficient of prey y. The growth of predator z depends on several factors including prey availability, predator catching and handling ability, prey searching ability etc. Here, we investigate the model dynamics when predator grows according to Leslie-Gower type FR [8,15] $h(x,y,z) = s_4 z(1 - \frac{s_3 z}{1+s_1 x+s_2 y})$. Where s_1, s_2 and s_3 are the coefficients of environmental carrying capacity which reduce predator population while is the intrinsic growth rate of the predator z. The identical participation of male and female in the growth of predator [9,18,19] gives $h(x,y,z) = s_4 z^2(1 - \frac{s_3}{1+s_1 x+s_2 y})$.

3 Model Formulations

The model formulations are described based on the following assumptions:

3.1 Case 1: when both preys have the same functional response:

Using identically H_I , H_{II} H_{III} and $H_{IV}FR$ in both bottom prey and keeping Leslie-Gower type predator, following food web models are formulated in equation 2- 5 respectively. Many Authors studies propertise of these FR in their studies [6,13,21]. The detailed one parameter analysis of functional response related model 2- 5 is given in Figure 6.

$$\begin{aligned} x' &= x(1-x) - w_1 xz \\ y' &= y(1-y)w_4 - w_5 yz \\ z' &= w_6 z^2(1 - \frac{w_7}{1 + w_8 x + w_9 y}) \end{aligned} \tag{2}$$

$$
\begin{aligned}
x' &= x(1-x) - \frac{w_1 xz}{1 + w_2 x + w_3 y} \\
y' &= y(1-y)w_4 - \frac{w_5 yz}{1 + w_2 x + w_3 y} \\
z' &= w_6 z^2 (1 - \frac{w_7}{1 + w_8 x + w_9 y})
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
x' &= x(1-x) - \frac{w_1 x^2 z}{1 + w_2 x^2 + w_3 y^2} \\
y' &= y(1-y)w_4 - \frac{w_5 y^2 z}{1 + w_2 x^2 + w_3 y^2} \\
z' &= w_6 z^2 (1 - \frac{w_7}{1 + w_8 x + w_9 y})
\end{aligned} \tag{4}
$$

$$
\begin{aligned}
x' &= x(1-x) - \frac{w_1 xz}{1 + w_2 x^2 + w_3 y^2} \\
y' &= y(1-y)w_4 - \frac{w_5 yz}{1 + w_2 x^2 + w_3 y^2} \\
z' &= w_6 z^2 (1 - \frac{w_7}{1 + w_8 x + w_9 y})
\end{aligned} \tag{5}
$$

Where parameter w_1 & w_5 are maximum attack rate at prey x and prey y respectively by predator z. Parameter w_2 & w_3 are half-saturation constant for prey x and prey y in absence of other prey respectively. Parameter w_4 is the distinuished logistic factor for prey y . Parameter w_6 is the growth rate for the predator z. Parameter w_7 is reduction in predator in the severe scarcity of prey x and prey y. Parameter w_8 & w_9 is prey preferences for predation of prey x and prey y respectively by the predator.

3.2 Case 2: Different functional responses in preys

Due to different predation behaviour is happened for different prey species for common predator and also it is documented that if the physical habitat is changed then the same predator predation rate is altered on the same prey. Assuming H_I and H_{II} in the prey equation, the Model 1 becomes the Model 6. Similarly, the functional response H_I with combination H_{III} and $H_{IV}$is taken then Model 1 becomes Model 7- 8 respectively. The FR H_{II} with combination H_{III} and H_{IV} are taken in the Model 1 becomes Model 9- 10 and our last model 11 has taken combination of the H_{II} and H_{IV} FR .

$$
\begin{aligned}
x' &= x(1-x) - w_1 xz \\
y' &= y(1-y)w_4 - \frac{w_5 yz}{1 + w_3 y} \\
z' &= w_6 z^2 (1 - \frac{w_7}{1 + w_8 x + w_9 y})
\end{aligned} \tag{6}
$$

$$
\begin{aligned}
x' &= x(1-x) - w_1 xz \\
y' &= y(1-y)w_4 - \frac{w_5 y^2 z}{1 + w_3 y^2} \\
z' &= w_6 z^2 (1 - \frac{w_7}{1 + w_8 x + w_9 y})
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
x' &= x(1-x) - w_1xz \\
y' &= y(1-y)w_4 - \frac{w_5yz}{1+w_3y^2} \\
z' &= w_6z^2(1 - \frac{w_7}{1+w_8x+w_9y})
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
x' &= x(1-x) - \frac{w_1xz}{1+w_2x+w_3y} \\
y' &= y(1-y)w_4 - \frac{w_5y^2z}{1+w_2x^2+w_3y^2} \\
z' &= w_6z^2(1 - \frac{w_7}{1+w_8x+w_9y})
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
x' &= x(1-x) - \frac{w_1xz}{1+w_2x+w_3y} \\
y' &= y(1-y)w_4 - \frac{w_5yz}{1+w_2x^2+w_3y^2} \\
z' &= w_6z^2(1 - \frac{w_7}{1+w_8x+w_9y})
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
x' &= x(1-x) - \frac{w_1x^2z}{1+w_2x^2+w_3x^2} \\
y' &= y(1-y)w_4 - \frac{w_5yz}{1+w_2x^2+w_3y^2} \\
z' &= w_6z^2(1 - \frac{w_7}{1+w_8x+w_9y})
\end{aligned} \tag{11}
$$

4 Analysis

Theorem 1. *All the model formulated in equations 2- 11 are bounded and dissipative. Further, all* **models 2- 11** *are persistent.*

Proof. The proof for individual models **models 2- 11** can be expressed on similar lines. Taking $h(x, y, z)$ explicitly as Leslie Gower type FR, the following set of differential equations represent the model 1 as:

$$
\begin{aligned}
x' &= x(1-x) - zf(x,y) \\
y' &= y(1-y)w_4 - zg(x,y) \\
z' &= w_6z^2(1 - \frac{w_7}{1+w_8x+w_9y})
\end{aligned} \tag{12}
$$

From the above expression, we have $x' \leq x(1-x) \implies x(t) \leq \frac{1}{1+ke^{-t}} \forall t \geq 0$ Here $k = \frac{1}{x(0)} - 1$ is the constant of integration. Now, taking $t \to \infty$ yield $\lim_{x\to\infty} x(t) \leq 1 \forall t \geq 0$ similarly, we have $y'(t) \leq y(1-y)w_4 \implies y(t) \leq \frac{w_4}{w_4+k_1 exp(-w_4t)} \forall t \geq 0$ with $k_1 = w_4(\frac{1}{y(0)} - 1)$ is the constant of integration. Now, taking $t \to \infty, \lim_{y\to\infty} y(t) \leq 1 \forall t \geq 0$ For predation functional response, Let $\phi(t) = x(t) + y(t) + \frac{z(t)}{\kappa_3}; \phi(0) \geq 0$, Then $\frac{d\phi}{dt} + k_4\phi(t) \leq x(1-x+\kappa_4) + y(1-y)w_4 + \zeta(x, y, z)$ where $\zeta(x, y, z) = w_6z^2(1 - \frac{w_7}{1+w_8x+w_9y}) + \frac{\kappa_4 z(t)}{\kappa_3}$. Using the

maximum property of function, we get $\frac{d\phi}{dt} + k_4\phi(t) \leq \frac{(1+w_4)}{4} + \kappa_4 + \zeta_0$ provided $w_6 a > w_7$ and $k_3 = ak_4^2 \forall t \geq 0$,

$$\phi(t) = \frac{(1 + w_4 + 4\kappa_4 + 4\zeta_0)}{4} - \frac{(1 + w_4 + 4k_4 - \phi_0)exp(-k_4 t)}{4}$$

$$\implies \lim_{t\to\infty} \phi(t) = \frac{(1 + w_4 + 4\kappa_4 + 4\zeta_0)}{4k_4}$$

This implies that solutions of system (12) are uniformly bounded for any initial value of ${\mathbb{R}_+}^3$. To show the system is dissipative, let there exist such that $(\gamma_1, \gamma_2, \gamma_3) > 0$ such that $\Omega(x_0, y_0, z_0) \subset {\mathbb{R}_+}^3 = \{(x, y, z) : 0 \leq x \leq \gamma_1, 0 \leq y \leq \gamma_2, 0 \leq z \leq \gamma_3\}$ for all $(x_0, y_0, z_0) \geq 0$ where $\Omega(x_0, y_0, z_0)$ is the omega-limit set of the orbit initiating at (x_0, y_0, z_0). Thus, the general form of the model 1 is dissipative and bounded. Therefore, for any positive solution of the model 1 can persist for longer time. Hence all **models 2- 11** are dissipative and persist.

Equilibrium points & Models Behavior

Theorem 2. *All* **models 2- 11** *has the trivial equilibrium point* $E_0(0, 0, 0)$ *and the axial equilibrium points* $E_1(1, 0, 0), E_2(0, 1, 0)$*, and the planar equilibrium point* $E_3(1, 1, 0)$*. The axial singularity* $E(0, 0, 1)$ *is biologically unfeasible and hence doesn't exist due to absence of prey.*

Theorem 3. *In the absence of one prey in* **model 2**, *the following equilibrium points exist:* $E_{11}(\frac{w_7-1}{w_8}, 0, \frac{1-\bar{x}}{w_1})$ & $E_{12}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4}{w_5})$ *and in the positive octant, the non-trivial equilibrium point* $E_{13}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{y} = \frac{(1-\bar{y})w_4}{w_5}$ & $\bar{z} = \frac{1-\bar{x}}{w_1}$

Theorem 4. *In the absence of one prey in the* **model 3**, *the following equilibrium points exist:* $E_{21}(\frac{w_7-1}{w_8}, 0, \frac{1+w_2\bar{x}(1-\bar{x})}{w_1})$ & $E_{22}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y})}{w_5})$ *and in the positive octant, the non-trivial equilibrium point* $E_{23}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})(1+w_2\bar{x}+w_3\bar{y})}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^2 - w_3\bar{y} - (1 - \bar{y})(1 + w_2\bar{x}) + \frac{w_5\bar{z}}{w_4} = 0$.

Theorem 5. *In the absence of one prey in the* **model 4**, *the following equilibrium points exist:* $E_{31}(\frac{w_7-1}{w_8}, 0, \frac{1+w_2\bar{x}^2(1-\bar{x})}{w_1\bar{x}})$ & $E_{32}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5\bar{y}})$ *and in the positive octant, the non-trivial equilibrium point* $E_{33}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})(1+w_2\bar{x}^2+w_3\bar{y}^2)}{w_1\bar{x}}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 + \bar{y}(1 + w_2\bar{x}^2) + \frac{w_5\bar{z}}{w_4} - 1 - w_2\bar{x}^2 = 0$

Theorem 6. *In the absence of one prey in the* **model 5**, *the following equilibrium points exist:* $E_{41}(\frac{w_7-1}{w_8}, 0, \frac{1+w_2\bar{x}^2(1-\bar{x})}{w_1})$ & $E_{42}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5})$ *and in the positive octant, the non-trivial equilibrium point* $E_{43}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})(1+w_2\bar{x}^2+w_3\bar{y}^2)}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 - (1 - \bar{y})(1 + w_2\bar{x}^2) + \frac{w_5\bar{z}}{w_4} - 1 = 0$

Theorem 7. *In the absence of one prey in the* **model 6**, *the following equilibrium points exist:*$E_{51}(\frac{w_7-1}{w_8}, 0, \frac{(1-\bar{x})}{w_1})$ & $E_{52}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y})}{w_5}$ *and in the positive octant, the non-trivial equilibrium point* $E_{53}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^2 + (1-w_3)\bar{y} + \frac{w_5\bar{z}}{w_4} - 1 = 0$

Theorem 8. *In the absence of one prey in the* **model 7**, *the following equilibrium points exist:*$E_{61}(\frac{w_7-1}{w_8}, 0, \frac{(1-\bar{x})}{w_1})$ & $E_{62}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5\bar{y}}$ *and in the positive octant, the non-trivial equilibrium point* $E_{63}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 + \frac{w_5\bar{z}\bar{y}}{w_4} + \bar{y} - 1 = 0$

Theorem 9. *In the absence of one prey in the* **model 8**, *the following equilibrium points exist:*$E_{71}(\frac{w_7-1}{w_8}, 0, \frac{(1-\bar{x})}{w_1})$ & $E_{72}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5}$ *and in the positive octant, the non-trivial equilibrium point* $E_{73}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 + \frac{w_5\bar{z}}{w_4} + \bar{y} - 1 = 0$

Theorem 10. *In the absence of one prey in the* **model 9**, *the following equilibrium points exist:*$E_{81}(\frac{w_7-1}{w_8}, 0, \frac{(1-\bar{x})(1+w_2\bar{x})}{w_1})$ & $E_{82}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5\bar{y}}$ *and in the positive octant, the non-trivial equilibrium point* $E_{83}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})(1+w_2\bar{x})}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 + \frac{w_5\bar{z}\bar{y}}{w_4} + \bar{y} - 1 = 0$

Theorem 11. *In the absence of one prey in the* **model 10**, *the following equilibrium points exist:*$E_{91}(\frac{w_7-1}{w_8}, 0, \frac{(1-\bar{x})(1+w_2\bar{x})}{w_1})$ & $E_{92}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5\bar{y}}$ *and in the positive octant, the non-trivial equilibrium point* $E_{93}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})(1+w_2\bar{x})}{w_1}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 + \frac{w_5\bar{z}}{w_4} + \bar{y} - 1 = 0$

Theorem 12. *In the absence of one prey in the* **model 11**, *the following equilibrium points exist:*$E_{101}(\frac{w_7-1}{w_8}, 0, \frac{(1-\bar{x})(1+w_2\bar{x}^2)}{w_1\bar{x}})$ & $E_{102}(0, \frac{w_7-1}{w_9}, \frac{(1-\bar{y})w_4(1+w_3\bar{y}^2)}{w_5}$ *and in the positive octant, the non-trivial equilibrium point* $E_{103}(\bar{x}, \bar{y}, \bar{z})$ *exists, where* $\bar{x} = \frac{w_7-1-w_9\bar{y}}{w_8}, \bar{z} = \frac{(1-\bar{x})(1+w_2\bar{x}^2)}{w_1\bar{x}}$ *and* $\bar{y}$ *can be calculated by a quadratic equation* $w_3\bar{y}^3 - w_3\bar{y}^2 + \frac{w_5\bar{z}}{w_4} + \bar{y} - 1 = 0$

The local behaviour of the above equilibrium points are stated in the following theorems:

Theorem 13. *The trivial equilibrium* $E_0(0, 0, 0)$ *in all* **model 2 - 11** *is always non-hyperbolic and unstable. There exist unstable subspace along xy plane and center subspace along z plane.*

Theorem 14. *The axial singularity* $E_1(1, 0, 0)$ *in all* **model 2 - 11** *is non-hyperbolic and saddle point. The y plane of the system is an unstable subspace, x plane of the system is stable subspace and z plane of the system is center subspace.*

Theorem 15. *The axial singularity* $E_2(0,1,0)$ *in all* **model 2 - 11** *is non-hyperbolic, saddle point. The x plane of the system is an unstable subspace, y plane of the system is stable subspace and z plane of the system is center subspace.*

Theorem 16. *The equilibrium* $E_3(1,1,0)$ *in all* **model 2 - 11** *is non-hyperbolic and stable. The system has stable subspace along* $x-y$ *plane and center subspace along z plane.*

Theorem 17. *The equilibrium points* $E_4(\bar{x},0,\bar{z}), E_5(0,\bar{y},\bar{z})$ *and* $E_6(\bar{x},\bar{y},\bar{z})$ *are asymptotically stable in each* **model 2 - 11** *individually, if the Jacobian of each model at the equilibrium point satisfies the following conditions*

$$\begin{aligned} A_1 &= -(A_{11}+A_{22}+A_{33}) > 0 \\ A_2 &= A_{22}A_{33}+A_{11}A_{33}+A_{11}A_{22}-A_{12}A_{21}-A_{32}A_{33}-A_{31}A_{13} > 0 \\ A_3 &= A_{11}A_{23}A_{32}+A_{12}A_{21}A_{33}+A_{13}A_{31}A_{22}-A_{11}A_{22}A_{33}-A_{12}A_{31}A_{23} \\ &\quad -A_{32}A_{21}A_{13} > 0. \end{aligned} \tag{13}$$

The proofs of the above theorems are quite easy and hence left. Here, proof of Theorem 17 is stated:

Proof. the general equilibrium point $E_6(\bar{x},\bar{y},\bar{z})$ put in RHS of model 1 and taking first order differentiation. We found the jacobian matrix $J_{x,y,z}$ for the model system (1) is

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix}$$

where $A_{11} = 1-2x-w_1z, A_{12} = 0, A_{13} = -w_1x, A_{21} = 0, A_{22} = w_4 - 2w_4y - w_5z$, $A_{23} = -w_5y, A_{31} = \frac{w_6w_7w_8z^2}{1+w_8x+w_9z}^2, A_{32} = \frac{w_6w_7w_9z^2}{1+w_8x+w_9z}^2$, $A_{33} = 2w_6z(1-\frac{w_7}{1+w_8x+w_9z})$ The characteristic polynomial for the Jacobean Matrix $\lambda^3+A_1\lambda^2+A_2\lambda+A_3=0$ where A_1, A_2, A_3 are mentenioned in the equation 13. The system is asymptotically stable if the eigenvalues are negative and $A_1 > 0, A_2 > 0, A_3 > 0$ and $A_1A_2 - A_3 > 0$

5 Numerical Simulations and results

To observe the impact of different functional responses on one prey two predator interaction **model 2 - 11**, the same parameters are used in each model simulations. Simulations results suggest that the qualitative behaviour of models are structurally robust (testing with other various parameter combinations), therefore following parameter combinations have been used to observe the global dynamics:

$$\begin{aligned} &w_1 = 3.6, w_2 = 1.7, w_3 = 1.8, w_4 = 1.5, w_5 = 3.2, w_6 = 1, w_7 = 2.1, w_8 = 1.7, \\ &w_9 = 1.8 \end{aligned} \tag{14}$$

6 One-Parameter Bifurcation diagrams

We extensively carried out the numerical simulations and using the continuation algorithm, we drew one-parameter bifurcation diagrams with respect to the parameter w_4 for **model 2 - 11.** The **model 2 - 4** are studied by many authors and our findings are in line with them [6,13,21] (see figure 1). We carried out the simulation of the **model 2 - 11** and found non- oscillation properties are found with persistency along with extinctions. Details bifurcation point and persistency range of all model are given in table 1.

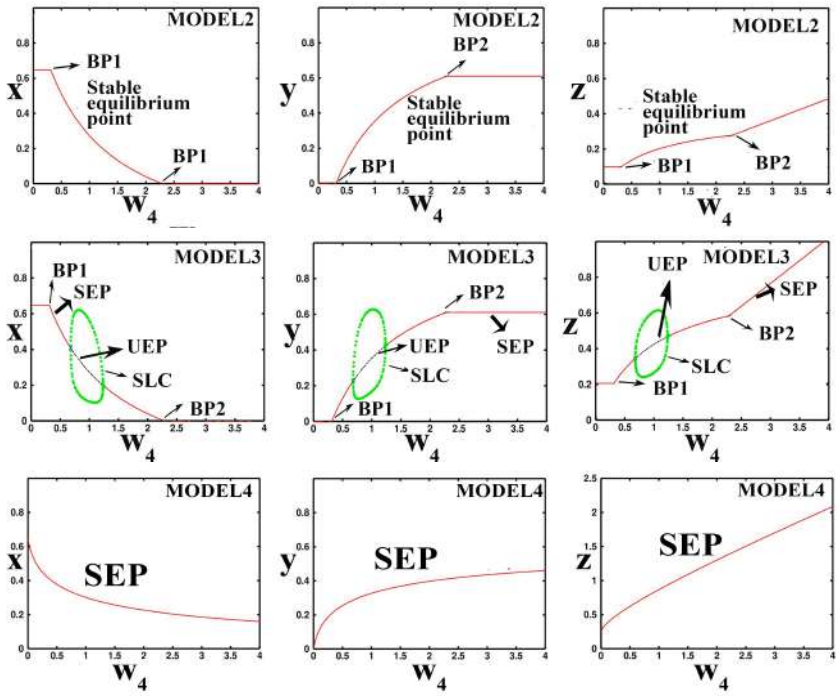


Fig. 1. One parameter bifurcation diagram for prey x, y and predator z with respect to the parameter w_4 for **model 2 - 4** are shown. Here red line is stable equilibrium point (SEP), green circle is stable limit cycle (SLC), Black colour is unstable equilibrium point(UEP). BP1= first branch point, BP2= second branch point, Model 2 is H_IFR alone, Model 3 is $H_{II}FR$ alone, Model 4 is $H_{III}FR$ alone. Model 3 and Model 5 have similar qualitaive behaviour. Combination of FR (in Model 6 - 11) are similar qualitative dynamics like Model 2.

7 Two-Parameter Bifurcation Diagrams

The parameter region for the species coexistence, extinction of prey is observed in two-parameter bifurcation diagram in figure 2 - 3. Model 4 with H_{III} FR

exhibit stabilizing effect on system behaviour and a possibility of avoiding the extinction of species. The extinction of prey or predator may not be possible in the Model 4 only which is a significant criterion of using these functional responses in predator-prey interactions. In the two-parameter bifurcation diagram for the Model 3, there are four regions, two regions are extinction region where preys are extinct while in between these two regions, all species survive and co-exists. In the middle of coexists region, a periodic region exists where all the populations are periodic. Similar behaviour is given by H-IV FR (Model 5) see in figure 2. While, the Model 2 shows three regions, the first region and third region gives extinction of one of the prey and center region gives coexistence. Similar behaviour is shown by Model 6, 8, 10. Model 7, 9, 11 give only two regions one is coexistence and other is extinction.

Model	**BP1**	**BP2**	**HB1**	**HB2**	**Persistence range**
2	0.3137	2.286	N.A.	N.A.	(0.3137 - 2.286)
3	0.3138	2.286	0.6805	1.235	(0.3138 - 2.286)
4	N.A.	N.A.	N.A.	N.A.	(0 - 4)
5	0.3138	2.286	0.6692	1.233	(0.3138 - 2.286)
6	0.3137	1.088	N.A.	N.A.	(0.3137 - 1.088)
7	0.8353	N.A.	N.A.	N.A.	(0 - 0.8353)
8	0.3137	1.367	N.A.	N.A.	(0.3137 - 1.367)
9	0.8354	N.A.	N.A.	N.A.	(0 - 0.8354)
10	0.6588	1.367	N.A.	N.A.	(0.6588 - 1.367)
11	0.8300	N.A.	N.A.	N.A.	(0.8300 - 4)

Table 1. List of all branch point (First branch point (BP1) and Second branch point (BP2)), Hopf bifurcation point (First Hopf point (HB1) and Second Hopf point (HB2)), and persistence range found in the respective model.

8 Discussion and conclusions

For understanding the significant effect of different functional response in two prey and one predator, we applying a combination of functional response in the Model 1 and found there are only three significant qualitative behaviours which are well explained in the result section. We have applied the combination of functional responses and found co-existence is possible for a wide range of parameter in every model but care should be taken as some parameter combination may lead to the extinction of some species except $H_{III}FR$ (Model 4) . The oscillatory behaviour is observed in $H_{II}FR$ (Model 3) and $H_{IV}FR$ (Model 5) with system persistence, species co-existence in the form of limit cycles and extinction of one of the prey in these systems. We observed that $H_{II}FR$ (*Model 3*) and $H_{IV}FR$ (*Model 5*) most likely to produce periodic solution through supercritical or subcritical Hopf-bifurcation. Simple extinction and coexistence are observed in H_IFR (Model 2).
The previous studies on FR in predator-prey models, provide an insight into

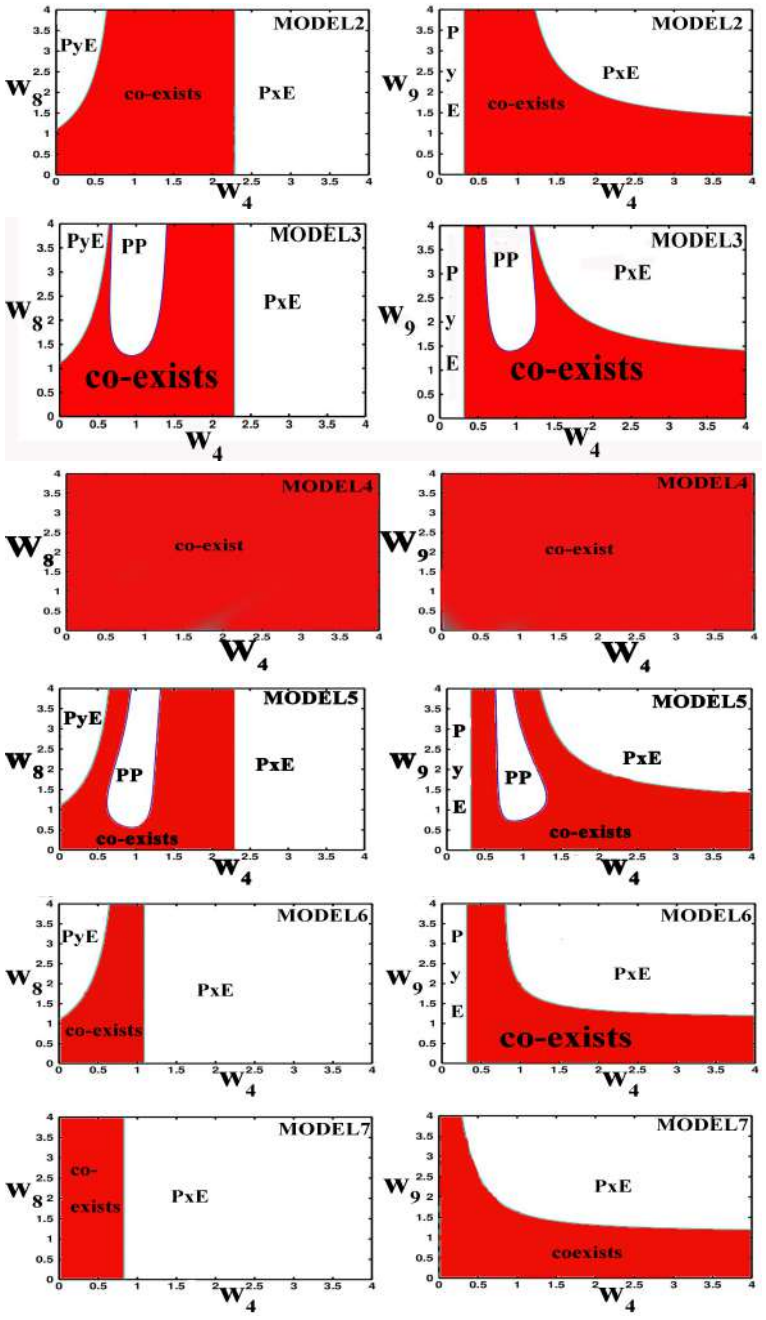


Fig. 2. Two-parameter bifurcation diagram with respect to w_8 & w_4 (left) and w_9 & w_4 right) for the Model 2- 7 respectively. Here, PyE=prey y extinct, PP= periodic populations,PxE= prey x extinct.

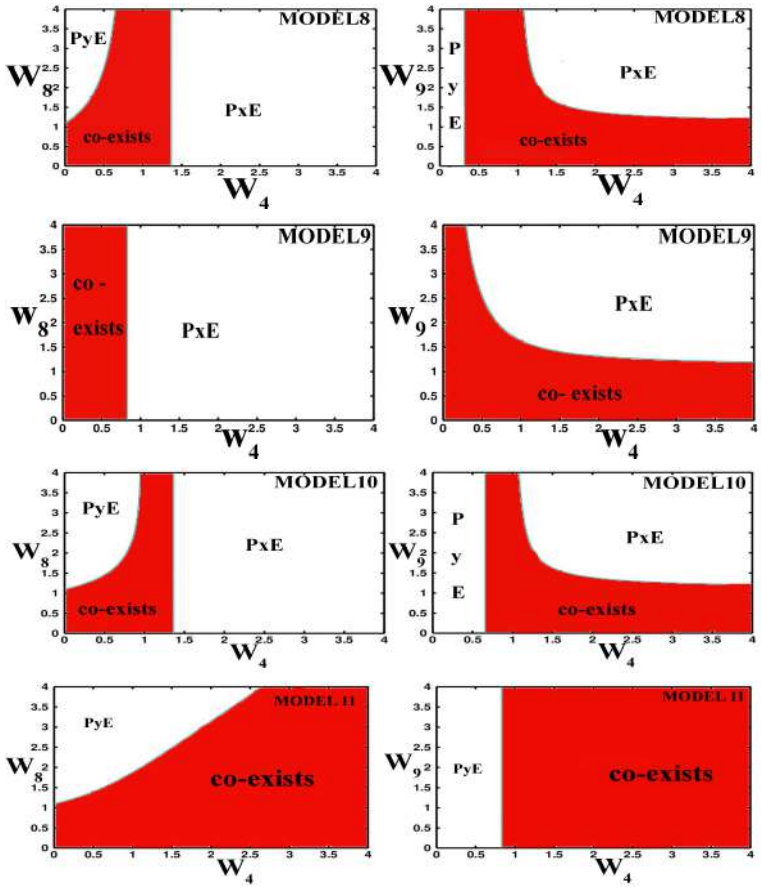


Fig. 3. Two-parameter bifurcation diagram with respect to w_8 & w_4 (left) and w_9 & w_4 right) for the Model 8- 11 respectively. Here, PyE=prey y extinct, PxE= prey x extinct.

system dynamics with the individual functional response and many experimental results require the use of multiple FR in their Mathematical Models. Therefore, it is essential to observe the impact of the combination of functional responses on simple predator-prey models and to compare the results. So, we have applied the combination of FRs in the same model and found that other FRs have stabilize the system by dampening the oscillations. Any FR is combined with $H_{II}FR$ (or $H_{IV}FR$), the oscillatory behaviour of the Model is disappeared and extinction of one of the prey occurred through the branch point. When H_IFR (or $H_{III}FR$) is taken with $H_{II}FR$ (or $H_{IV}FR$) then two threshold value is found with coalescing of both Hopf bifurcation point. Below the branch point, one prey y is extinct while above the branch point another prey x is extinct, in between these, species co-existence is found. Some distinguished differences like Hopf bifurcation points, branch points and specific population with parameter value are shown in table 1. Here, we observe

non- linear functional response $H_{II}FR$ (or $H_{IV}FR$) alone taken in system, model dynamics grants oscillations with extinction while the different combination of FR offering only stability and extinctions with the coalescence of Hopf bifurcation points. From this argument, we can say that non- linearity of functional responses are not always giving periodic solutions/ limit cycles. The FR H_I, H_{II} or H_{IV} alone produce two branch point while the combination with $H_{III}FR$ destroy one branch point and enhance the stability of the system. In overall systems, the prey's equations are varied through different functional response and predator's equations are same but we found the quantitative value of predator are varied rather than that of prey. i.e. shaping of system dynamics is main rooted by FR. FR is also the main cause of a different range of persistence value.
System equations deliberated above were kept very simple to make the effect of functional responses easier to analyze. In our two prey and predator model, we find H_I, H_{II}, H_{III} FR give distinguished qualitative behaviour and H_{II} and $H_{IV}FR$ dynamical behaviour are similar. The combination of Holling FR in model produce similar behaviour like Model 2. The combination of different FR in one model is not studied by others (in my knowledge) but found in experimental studies and we found non- linear FR have counter their oscillatry property and the model shows only co-existence and extinction.

9 Acknowledgement:

The first author also acknowledged to CSIR, New Delhi for providing Junior Research fellowship wide $fileno. - 9/013(0822)/2018 - EMR - 1$.

References

1. C. S. Holling, Some characteristics of simple types of predation and parasitism. *Canadian Entomologist* 91, 385- 398,1959a.
2. C. S. Holling, The components of predation as revealed by a study of small mammal predation of the European pine sawfly. *Canadian Entomology* 91, 293–320, 1959b.
3. C. S. Holling, The functional response of invertebrate predators to prey density,*Memoirs of the Entomological Society of Canada,* 45, pp 3–60,1965.
4. G. F. Gause, *The struggle for existence* Williams and Wilkins, Baltimore, Maryland, USA, 1934.
5. J. A. J. Metz, M. W. Sabelis, and J. H. Kuchlein. Sources of variation in predation rates at high prey densities: an analytic model and a mite example.*Experimental and Applied Acarology*5, 187–205, 1988.
6. J. B. Collings, The effects of the functional response on the bifurcation behaviour of a mite predator-prey interaction model. *Journal of Mathematical Biology* 36 ,149-168, 1997.
7. J. F. Andrews, A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates, *Biotechnology and Bioengineering,* 10, pp. 707–723,, 1968.
8. L. A. Real, The kinetics of functional response. *American Naturalist* 111 289–300, 1977.

9. M. A. Aziz-Alaoui, Study of a Leslie–Gower-type tritrophic population model,*Chaos Solitons Fractals*; 14, 1275–93, 2002.
10. M. A. Barbeau, R. E. Scheibling, Behavioural mechanisms of prey size selection by sea stars and crabs preying on juvenile sea scallops. *J Exp Mar BIolEcol* 180, 103-136, 1994.
11. M. A. Barbeau, R. E. Scheibling, B. G. Hatcher, Behavioural responses of predatory crabs and sea stars to varying density of juvenile sea scallops. *Aquaculture*, 169, 97-98, 1998.
12. M. E. Solomon, The Natural Control of Animal Populations, *Journal of Animal Ecology* 18, **1** , 1-35, 1949.
13. N. Bairagi, P. K. Roy, J. Chattopadhyay, Role of infection on the stability of a predator–prey system with several response functions—A comparative study, *Journal of Theoretical Biology,* 248, *1*, 10-25,2007.
14. P. A. Abrams, Functional responses of optimal foragers. *American Naturalist*, 120, 382-390. 1982.
15. P. H. Leslie and J. C. Gower, the properties of a stochastic model for the predator – prey type of intersection between two species, *Biometrica* 47, 219-234, 1960.
16. P. Turchin, *Complex population dynamics, A theoretical/empirical synthesis, Monograph in Population Biology* Princeton University Press, 35 2003.
17. R. J. Taylor, *Predation*, Chapman and Hall, New York, 1984.
18. R. K. Upadhayay, V. Rai, S. R. K. Iyengar, Species extinction problem: genetic vs ecological factors. *Appl Math Model* 11, **13** , 937–51, 2001.
19. S. Gakkhar, B. Singh, Complex dynamic behaviour in a food web consisting of two preys and a predator. *Chaos, Solitons and Fractals* 24, 789–801, 2005.
20. T. Royama, A comparative study of models for predation and parasitism. *Researches on Population Ecology* S1, 1–90, 1971.
21. Vance R.R. Predation and resource partitioning in one predator – two prey model Communities. *The American Naturalist* 112,987,1978.
22. Volterra, Varizione e fluttuazionidelnumerod'individui in specie animaliconventi, *Mem Accad. Nazi. LaIncci.* 2, 31-113, 1926.
23. W. Sokol and J. A. Howell, Kineties of phenol oxidation by washed cell, *Biotechnol Bioeng*, 23, 2039–2049,1980.

Fractal Dimension of Braided Rivers from Detailed Two-Dimensional Hydrodynamic Simulations

Matteo Nicolini[1]

[1] Polytechnic Department of Engineering and Architecture, via del Cotonificio ,114, University of Udine, 33100 Udine, Italy
(E-mail: matteo.nicolini@uniud.it)

Abstract. As opposed to meandering or channelized fluvial beds, braided rivers are characterized by a morphological activity starting at very low flows, since fluxes are concentrated in a limited number of small channels. With increasing discharge, more channels are involved, up to the situation in which the complete alluvial plain is flooded. As a consequence, there is an intermediate range of flows for which pattern complexity is maximum and braided indices are highest, representing essential conditions for the coexistence of a large variety of habitats and for ecosystems prosperity. In this paper, a new methodology for a quantitative assessment of the complexity of braided rivers at a reach scale is introduced. It is based on the application of the box-counting algorithm to flooded areas identified through a two-dimensional (shallow water) hydrodynamic simulation model, in order to derive an estimate of the fractal dimension with varying flow rate. The identification of the range of discharges for which the fractal dimension is highest is of particular importance in river restoration projects. An application to the River Tagliamento (North-East Italy) is illustrated.

.

Keywords: Morphodynamics, River restoration, Formative discharge, Box-counting algorithm, River Tagliamento.

1 Introduction

Braided rivers are complex, non-linear systems characterized by chaotic dynamics. Their local properties, such as the solid transport and the number of channels in a given section, are spatially and temporally variable, their prediction being precluded in the mid-to-long term (Redolfi[16]).

Among others, the process of bifurcation plays a fundamental role in the formation and development of a network of braided channels. Gravel bed rivers with small width-to-depth ratio are characterized by stable equilibrium, while width-to-depth ratios above a critical threshold trigger an initial formation of alternate bars, which determine a sinuous pattern and then a first bifurcation. In these last decades, such processes have been investigated with analytical approaches, numerical simulations and laboratory experiments (Parker[15]; Ikeda[9]; Jaeggi[10]; Colombini et al.[4]; Tubino et al.[21]; Lanzoni[13]).

In particular, experimental studies have demonstrated the importance of the presence of alternate bars in the formation of a braiding pattern (Federici and Paola[7]; Bertoldi and Tubino[2]; Jang and Shimizu[11]).

The morphodynamics of such systems is generally described by a statistical approach, which identifies the dependence of some reach-averaged properties

(like width, total and active channels, solid transport) on main controlling factors, like discharge, bed slope, sediment size and total river width (Ashmore[1]).
In the definition of reach-averaged properties, the most important issue is the identification of a proper scale length. In line with previous research, some authors have proposed at least ten times the river width (Egozi and Ashmore[6]), which is a value much higher than those related to meandering or channelized fluvial beds, typically of the order of the river width (Ikeda[9]; Seminara and Tubino[19]; Lanzoni[13]).
The identification of a characteristic length is difficult due to the fact that there are a lot of spatial and temporal scales coexisting together (Sapozhnikov and Foufoula-Georgiou[18]). Accordingly, some researchers (Sapozhnikov and Foufoula-Georgiou[17]; Walsh and Hicks[23]; Lane[12]) have proposed that fluvial patterns in braided rivers resemble those of self-similar fractals (or self-affine systems, if an isotropic character is not present).
One problem still open is the quantitative description of the pattern complexity: to this end, specific indices have been introduced, like the average number of channels in the characteristic length (Egozi and Ashmore[5]). Besides, not all channels are simultaneously morphologically active, since solid transport takes place only in a fraction of them. Bertoldi et al.[3] have shown that the total number of channels is well correlated with the dimensionless discharge, while the number of active channel is more dependent on the dimensionless stream power.
Another important controlling factor is the formation and growth of vegetation, whose influence on the planimetric configuration depends on the ratio between the growth time and the average interval of morphologically relevant inundations (Paola[14]). Gurnell et al.[8] have shown that such role is much more important in relatively low energy systems.
In the present paper we aim at a different approach for a quantitative description of such complex fluvial patterns, starting from detailed numerical simulations of the flow field and then analyzing the braiding patterns with conventional fractal analysis algorithms. An application to the River Tagliamento (North-East Italy) is presented. The results show that the values for which the fractal dimension is highest correspond to a narrow range of formative discharges.

2 Materials and Methods

2.1 Materials

The River Tagliamento is located in the North-East part of Italy, with a drainage basin of 2780 km^2 and a length of 178 km. Its upper part is characterized by the most extensive and connected length of dynamic and morphologically intact braided pattern within the Alps, leading to the most important braided reference system of the Alpine region (Tockner et al.[20]).
The area under investigation (Figure 1) is between Venzone (where a historically relevant gage station is present) and Pinzano gorge (where the corridor width shrinks from 1 km down to 130 m), for a total length of 22.5 km and an area of nearly 30 km^2. The 100-year discharge is 4500 m^3/s.

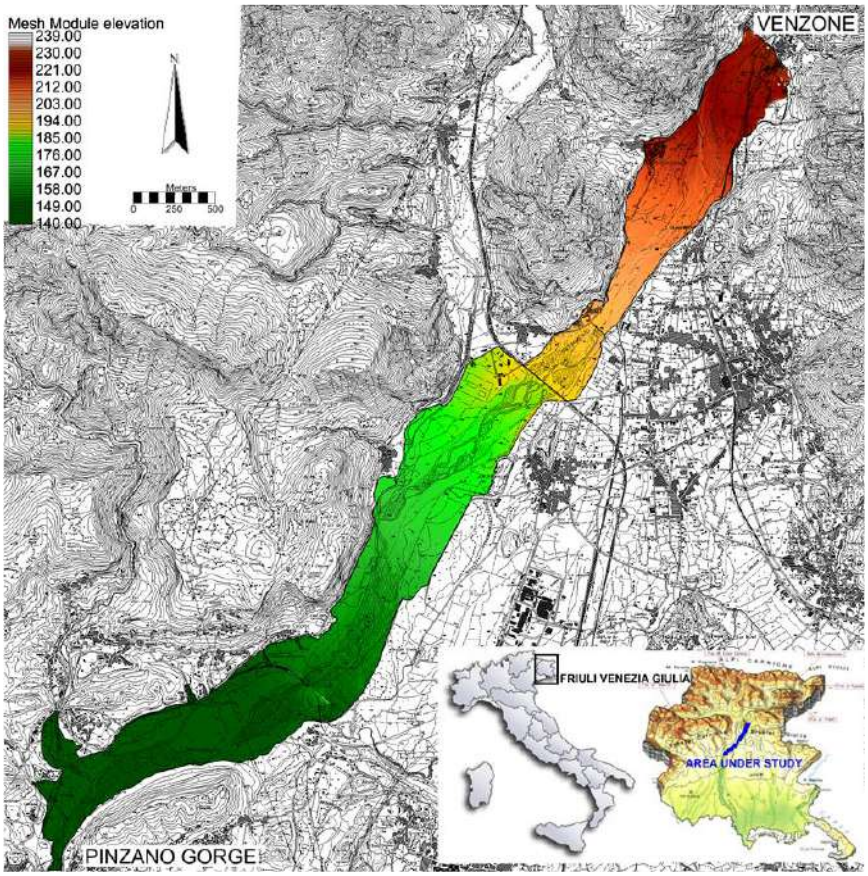


Fig. 1. General view of the area under analysis, with color map indicating terrain elevations

Lidar data available from Regione Autonoma Friuli Venezia Giulia with a resolution of 2 points/m^2 have been used for assigning elevations at mesh nodes, while aerial photographs (0.2 m resolution) have been analyzed in order to determine an estimate of water depths in those areas where water was present at the time of survey. It is important to put into evidence that lidar data do not actually discriminate between surface water elevation and underlying bottom elevation, thus not properly describing the geometry of flooded channels.
Several field campaigns and investigations allowed to determine local geometric characteristics of some defense structures (like groins, jetties and retaining walls) as well as bridge piers, which have been included in the computational mesh.

2.2 Methodology

The software SRH-2D has been adopted for the implementation of the hydrodynamic model. The software is free and can be downloaded from https://www.usbr.gov/. Last version is 3.2.4 (June 2019).

The software SMS (Surface Modeling System, distributed by Aquaveo™) has been used for the construction of the mesh, which is formed by 1372586 nodes and 1557718 cells, due to the fine discretization of the computational domain (cell sides range between 1 and 8 meters). The inflow (upstream) boundary condition has been kept constant in each hydrodynamic simulation, and results have been saved after a sufficient transient time guaranteed steady-state conditions throughout the computational domain.
The freeware software Fractalyse has been chosen for the fractal analysis of simulated flooded areas (downloadable from http://www.fractalyse.org/). In particular, the procedure adopted is the well-known box-counting algorithm (Turcotte[22]), which allows to determine the fractal dimension of the braided patterns.
The overall methodology can be summarized in the following steps:

1) Mesh generation and elevation assignment at nodes.
2) Hydrodynamic simulations with very low discharges (typically 10-20 m^3/s), in order to identify the areas occupied by low flows.
3) Subdivision of each aerial photograph image in RGB bands, and estimation of the water depth, H, obtained as: $H = a \cdot \ln(\lambda_R/\lambda_G)$, in which a is a calibration parameter and λ_R and λ_G are the red and green band intensities, respectively.
4) Creation of a 'mask' covering only the areas occupied by low flows; in this way, mesh nodes affected by new elevation assignments are limited to those calculated in step 2 (this step is necessary because the algorithm in step 3 identifies fictitious water depths in some dry, vegetated areas).
5) Assignment of new elevation at mesh nodes identified by the 'mask' of step 4; this is achieved by subtracting the water depths calculated in step 3 to the elevation originally present. In this way, only the wet areas as obtained by numerical simulations are altered by elevation changes.
6) Hydrodynamic simulations with updated mesh node elevations and varying water discharges (from very low flows up to the condition in which the flood plain is flooded).
7) Analysis of the planimetric flow patterns with box-counting algorithm and estimation of the fractal dimension of the flooded areas.

In particular, the calibration parameter a of step 3 has been determined starting from measured water depths just upstream a water intake for irrigation purposes (managed by Friulian Plain Bureau of Reclamation).

2.3 Box-counting algorithm and derivation of fractal dimension

In Euclidean geometry, a point is zero-dimensional, a line is one-dimensional, a plane is two-dimensional, and so on. The traditional meaning of the dimension of an object is that of giving the number of values needed to specify the position of a point on the object. Thus, one value needs to be given in order to specify the position of a point on a line, two values need to be specified for obtaining the position of a point on a plane, and so on.

Another meaning of 'dimension of an object' can be based on the idea of self-similarity. Of course, this must reproduce the traditional values when applied to classical Euclidean objects such as lines and planes. Consider how to give a self-similar description of a one-dimensional object, that is, a line segment: one way to do this is to say that a line segment of length, ℓ, consists of two copies of itself, each characterized by a length $\ell/2$. With the same reasoning, a filled-in square can be thought as four copies of itself, each having a side of length $\ell/2$.
In this way, two quantities characterize each of the self-similar shapes of the examples above:

- the number of self-similar copies, N;
- the edge length of the original relative to each copy, ϵ.

The following formula can be used to define the dimension D of an object:

$$D = \frac{\log N}{\log \epsilon} \tag{1}$$

Applying eq. (1) to the examples introduced before, in the case of a line segment: $N = 2$, $\epsilon = 2$, $D = 1$; and for the filled square: $N = 4$, $\epsilon = 2$, $D = 2$.
Equation (1) gives a formula for calculating the dimension also of a fractal object: it is sufficient to know the number of self-similar copies, N, and the size of the original relative to each copy, ϵ. For complicated objects, like the braiding pattern of rivers, the box-counting algorithm can be summarized in the following steps:

1) 'Cover' all the points in the object with boxes of edge-length ϵ_0, and count the number of these boxes, denoted as $N(\epsilon_0)$. In the case of braiding patterns, the boxes are squares.
2) Repeat step (1) using boxes with edge-length $\epsilon_1 = \epsilon_0/2$. Then repeat again using $\epsilon_2 = \epsilon_1/2$, $\epsilon_3 = \epsilon_2/2$, and so on. Obviously, for each ϵ_i, there is also the corresponding $N(\epsilon_i)$.
3) Theoretically, the dimension D is the number for which
$$\lim_{\epsilon \to 0} N(\epsilon) = A\epsilon^{-D} \tag{2}$$
where A is a constant. In practice, D may be estimated as
$$D = \frac{\log N(\epsilon_{i+1})/N(\epsilon_i)}{\log \epsilon_i/\epsilon_{i+1}} \tag{3}$$

The only difficulty is in selecting the value of i, and generally it is selected as large as possible in order to approximate the limit $\epsilon \to 0$. However, for real objects the boxes cannot be infinitely small, like in the present case, since it is inappropriate to make the covering boxes smaller than the size of a computational cell.

3 Results and Discussion

Several simulations have been run, with the discharge varying from 50 m^3/s up to 3000 m^3/s, this last value determining the inundation of all the floodplain and adjacent terraces.

For each discharge value, the simulation has been carried out for a sufficient time in order to establish steady state conditions. Figure 2 shows some results of the hydrodynamic model in terms of water depths obtained, respectively, for the discharge of 50, 300, 500 and 900 m^3/s.

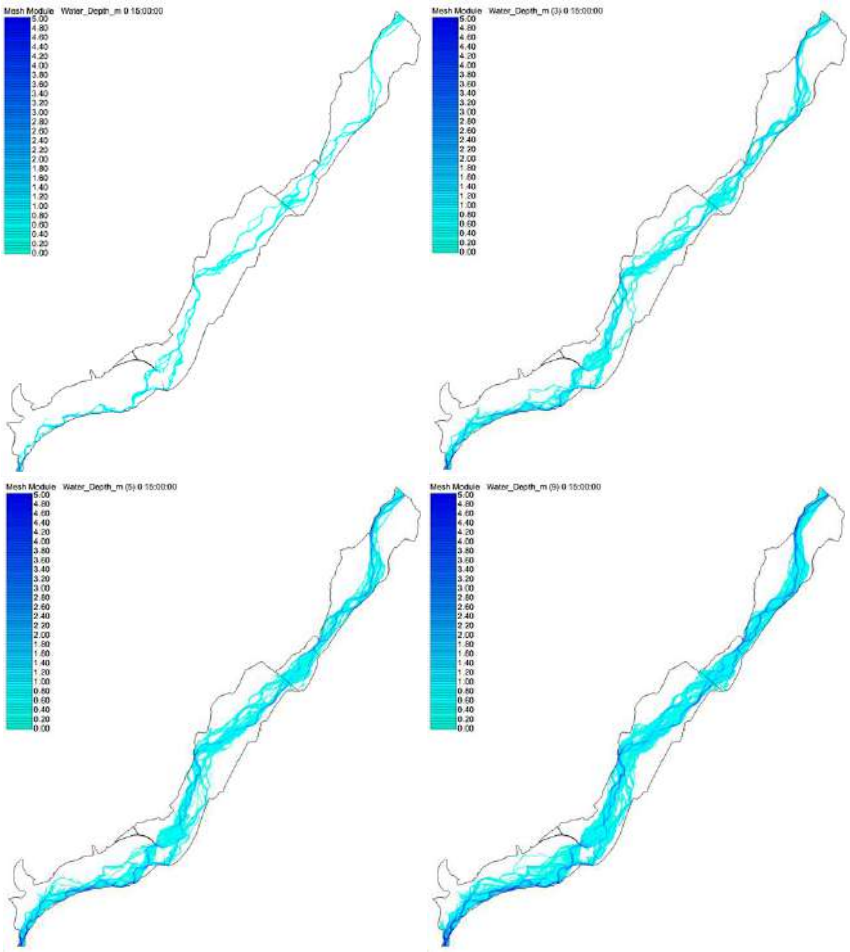


Fig. 2. Maps representing water depths for different discharge values, respectively 50 m^3/s (top left), 300 m^3/s (top right), 500 m^3/s (bottom left) and 900 m^3/s (bottom right)

Some level measurements allowed model calibration in the range of medium-to-high discharges, considering that a stage-discharge relationship is available at an intermediate section of the area under analysis. A previous physical model study identified such relationship for a control section where a bridge is present. In this way, Manning's roughness coefficients have been defined in order to reproduce the measured water levels within an acceptable tolerance.

From Figure 2, it is evident how the number of flooded channels at low flows is quite limited, and the morphology of braiding is not very articulated. For higher discharges, the fluxes subdivide into more channels, and braiding complexity increases accordingly, until it reaches a maximum. The further increase in discharge rapidly determines a situation for which all the floodplain is interested by waters, and braiding complexity decreases.
Such behavior is well described by the fractal dimension depicted in Figure 3, as a function of river discharge. In this case, the maximum value of nearly 1.29 is obtained for 500 m^3/s, corresponding to the return period of 4 months.

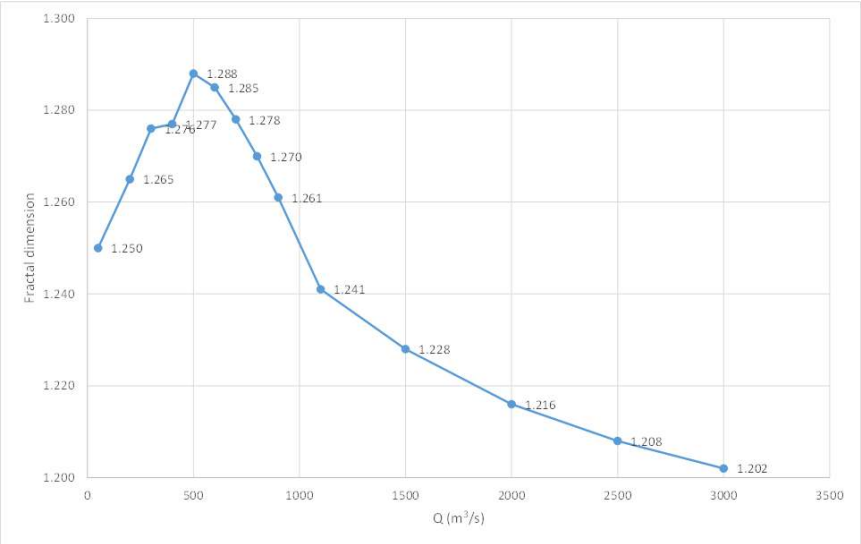


Fig. 3. Fractal dimension calculated through the application of the box-counting algorithm, as a function of varying river discharge

Figure 4 shows water depth maps for the southern zone of the area under study. The phenomenon described above is clearly captured by detailed two-dimensional hydrodynamic modelling.

Conclusions

The paper has presented a methodology for quantifying the pattern complexity of steady-state planimetric configuration of braided rivers. Starting from the results obtained from detailed two-dimensional hydrodynamic simulations, the box-counting algorithm allows to determine the fractal dimension of the braided pattern.
The procedure is relevant for the a-priori evaluation of river management works, in particular those aimed at excavation, extraction and widening. The maintenance of a range of formative discharges at low-to-medium flows is very

important for preserving ecological stability and the possibility of habitat evolution, which could be potentially endangered by failures in the restoration projects.

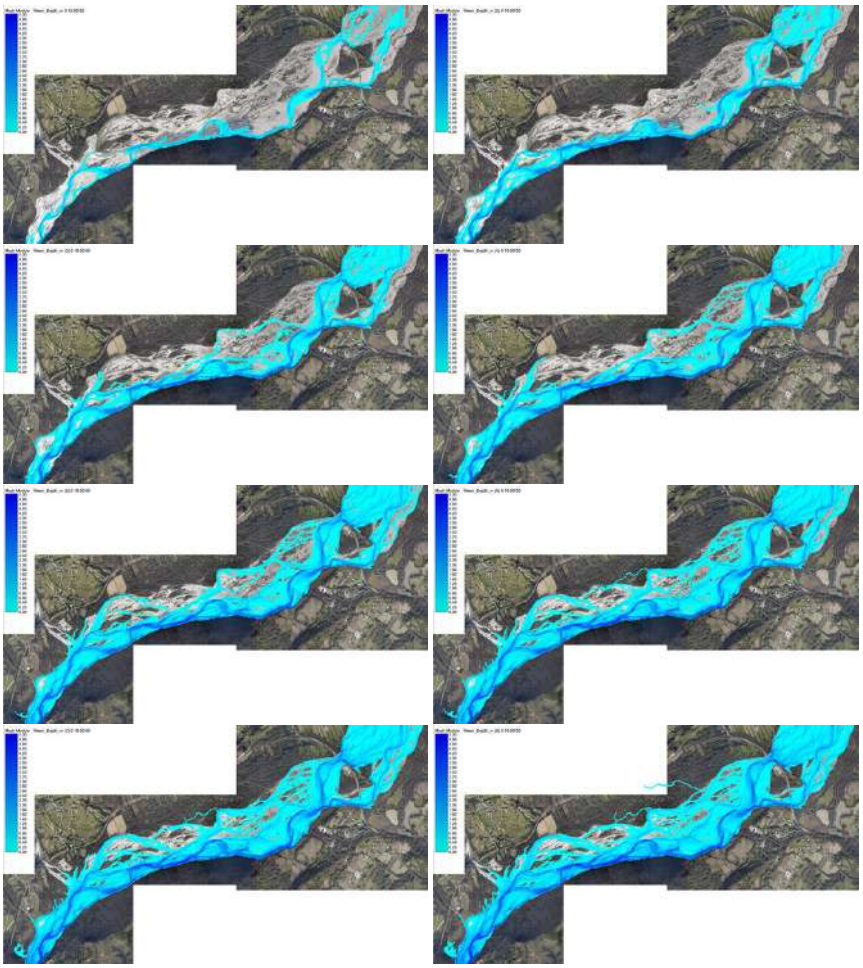


Fig. 4. Water depths for the southern zone of the computational area for varying discharges: from top to bottom, left to right: 50 m^3/s; 200 m^3/s; 300 m^3/s; 400 m^3/s; 500 m^3/s; 600 m^3/s; 700 m^3/s; 800 m^3/s

Aknowledgements

The author wishes to thank Regione Autonoma Friuli Venezia Giulia (www.regione.fvg.it), the Friulian Plain Bureau of Reclamation (Consorzio di bonifica Pianura Friulana, www.bonificafriulana.it) and the International Centre

of Mechanical Sciences (www.cism.it) for their support and the provision of materials and data related to River Tagliamento.

References

1. P. Ashmore. Morphology and dynamics of braided rivers. In: J. Shroder and E. Wohl (Eds.), Treatise on Geomorphology, Volume 9: Fluvial Geomorphology. Academic Press, San Diego, CA, 289–312, 2013.
2. W. Bertoldi and M. Tubino. Bed and bank evolution of bifurcating channels. Water Resources Research, 41, 2005.
3. W. Bertoldi, L. Zanoni and M. Tubino. Planform dynamics of braided streams. Earth Surface Processes and Landforms, 34, 547–557, 2009.
4. M. Colombini, G. Seminara and M. Tubino. Finite amplitude alternate bars. Journal of Fluid Mechanics, 181, 213–232, 1987.
5. R. Egozi and P. Ashmore. Defining and measuring braiding intensity. Earth Surface Processes and Landforms, 33, 14, 2121–2138, 2008.
6. R. Egozi and P. Ashmore. Experimental analysis of braided channel pattern response to increased discharge. Journal of Geophisical Research, 114, 2009.
7. B. Federici and C. Paola. Dynamics of channel bifurcations in noncohesive sediments. Water Resources Research, 39, 6, 2003.
8. A.M. Gurnell, W. Bertoldi and D. Corenblit. Changing river channels: the roles of hydrological processes, plants and pioneer fluvial landforms in humid temperate, mixed load, gravel bed rivers. Earth-Science Reviews, 111, 129–141, 2012.
9. S. Ikeda. Prediction of alternate bar wavelength and height. Journal of Hydraulic Engineering, 110, 371–386, 1984.
10. M.N.R. Jaeggi. Formation and effects of alternate bars. Journal of Hydraulic Engineering, 110, 2, 1984.
11. C.L. Jang and Y. Shimizu. Bed and bank evolution of bifurcating channels. ASCE Journal of Hydraulic Engineering, 131, 7, 2005.
12. S. Lane. Approaching the system-scale understanding of braided river behavior. In: G.H. Sambrook Smith, J.L. Best, C.S. Bristow and G.E. Petts (Eds.), Braided Rivers: Process, Deposits, Ecology and Management. Association of Sedimentologists Special Publication (volume 36), Blackwell, Oxford, 339–359, 2006.
13. S. Lanzoni. Experiments on bar formation in a straight flume 1. Uniform sediment. Water Resources Research, 36, 11, 3337–3349, 2000.
14. C. Paola. Modelling stream braiding over a range of scales. In: M.P. Mosley (Ed.), Gravel Bed Rivers V, New Zealand Hydrological Society, Wellington, 11–46, 2001.
15. G. Parker. On the cause and characteristic scales of meandering and braiding in rivers. Journal of Fluid Mechanics, 76, 3, 457–479, 1976.
16. M. Redolfi. Sediment transport and morphology of braided rivers: steady and unsteady regime. PhD Thesis, University of Trento and Queen Mary University of London, 2014.
17. V.B. Sapozhnikov and E. Foufoula-Georgiou. Self-affinity in braided rivers. Water Resources Research, 32, 5, 1429-1439, 1996.
18. V.B. Sapozhnikov and E. Foufoula-Georgiou. Validation of braided-stream models: spatial state-space plots, self-affine scaling, and island shapes. Water Resources Research, 34, 9, 2353-2364, 1998.
19. G. Seminara and M. Tubino. Weakly nonlinear theory for regular meanders. Journal of Fluid Mechanics, 244, 257–288, 1992.

20. K. Tockner, J.V. Ward, D.B. Arscott, P.J. Edwards, J. Kollmann, A.M. Gurnell, G.E. Petts and B. Maiolini. The Tagliamento River: a model ecosystem of European importance. Aquatic Sciences, 65, 239–253, 2003.
21. M. Tubino, R. Repetto and G. Zolezzi. Free bars in rivers. Journal of Hydraulic Research, 37(6), 759–775, 1999.
22. D.L. Turcotte. Fractal and chaos in geology and geophysics. Cambridge University Press, 1992.
23. J. Walsh and D.M. Hicks. Braided channels: Self-similar or self-affine? Water Resources Research, 38, 2002.

Modelling mechano-electric feedback and arrhythmia in a simplified multiscale cardiac model

Nicholas F Pearce and Eun-jin Kim

Fluid and Complex Systems Research Centre, Coventry University, Coventry, CV1 5FB, UK
(E-mail: pearcen5@Uni.Coventry.ac.uk)

Abstract. Mechano-electric feedback (MEF) is the mechanism by which mechanical changes in cardiac muscles influence the electrical excitation of the heart (and vice-versa). It is a fundamental feedback process in normal behaviour, helping to maintain the heart's response to the demands placed on it. Indeed, a disruption to the pattern of electrical signals triggered by pathological conditions in the mechanical environment, can lead to cardiac alternans and arrhythmia. In this paper, MEF is investigated using a low-order, lumped-parameter model which incorporates the mechanical, electrical and chemical action across both sarcomere (cardiac muscle fibres) and heart scales. Feedback is simulated by coupling the mechanical model to both the fast and slow variables of the electrical activity, thereby allowing the effects of mechanical stress and stretch on electrical patterns to be observed. Although simple, the model allows significant qualitative behaviour to be observed without significant cost. It is found that disruption to the MEF causes interesting bifurcations in the mechanical-electrical system, leading to various pathological conditions such as periodic doubling in heart beats, irregular (quasi-periodicity) frequencies.

Keywords: Nonlinear Dynamics; Multiscale model; Cardiac cycle; Mechano-electric Effect; Arrhythmias; Feedback; Lumped-parameter Model; Biological Complexity; Self-organization.

1 Introduction and Background

Regular function of the heart is ensured through several feedback and control mechanisms, allowing this complex organ to function across multiple scales and domains; from the macro-scale mechanical regulation down to the micro-scale electro-chemical activity. Failure of these mechanisms is known to lead to heart problems and even death. Indeed, heart diseases and failure remain the leading causes of death worldwide [1]. One of the most important feedback mechanisms is the mechano-electric feedback (MEF), which is the influence of the macro-scale mechanical activity on the micro-electric behaviour of the individual muscle fibres (sarcomere).

The aim of this study is to investigate MEF using a modified form of the newly proposed model of cardiac function previously described in Kim and Capoccia [12] and Kim and Capoccia [13]. This low-order, lumped-parameter model integrates the macro-scale mechanics of the heart and circulation with the micro-scale electro-chemical behaviour of the sarcomere without reference

to the time-varying elastance method. The time-varying elastance concept is often drawn upon by similar low-order models (Sagawa [9], Simaan [10]) to simulate the heart ventricle pressure-volume relationship, but it's foundations have been questioned (Claessens *et al.* [11]) and is unsuitable for modelling MEF. As will be shown, this new relatively modest representation yields meaningful qualitative information without the considerable computational expense of similar models involving far greater detail.

Some examples of previous attempts to model the MEF include the 'abstract' analysis by Knudsen *et al.* [2], who present a low order, lumped parameter model, whose excitable system is of Fitzhugh-Nagumo type (Fitzhugh [3]). This is coupled to an ODE for the ventricular contraction. In contrast, Collet *et al.* [4] include spatial dynamics as well as temporal in their one-dimensional analysis of MEF. Temperature variations are also incorporated. Dysfunction of cardiac spatial dynamics are known to lead to complex non-linear behaviour and arrhythmia in many settings (Qu *et al.* [5], Franzone *et al.* [6]). Amar *et al.* [7] employ their previously developed three-dimensional model of a contracting ventricle, into which an electrophysiological model of sarcomere membrane potential is introduced. This electrophysiological model is a first order ODE for the sarcomere membrane potential and describes 13 ionic currents, one of which is a stretch-activated current (SAC), used to simulate the MEF.

In the next section, our numerical model is described and the governing equations given. Then, the results of the MEF investigation are presented and discussed followed by a brief conclusion.

2 Numerical Method

Our model is presented below and as described above, is adapted from Kim and Capoccia [12,13] and differs, notably in equations (5) and (7) along with the values of some parameters used.

2.1 Microscale mechano-chemical model

At the micro-scale, the mechano-electric behaviour of the sarcomere is modelled by adopting the Bestel-Clement-Sorine method (Bestel [14], Sainte-Marie *et al.* [16], Bestel *et al.* [15]). In this approach, the active stress τ_c, stiffness k_c, strain ϵ_c, and velocity $v_c = \frac{d\epsilon_c}{dt}$, are obtained using the following set of differential equations:

$$\frac{dv_c}{dt} = -\chi\tau_c - \omega_0^2\epsilon_c - a\tau_c d_0(\epsilon_c) + b\left(\sqrt{\frac{V}{V_0}} - 1\right) \tag{1}$$

$$\frac{d\epsilon_c}{dt} = v_c \tag{2}$$

$$\frac{d\tau_c}{dt} = k_c v_c - (a_l|v_c| + |u|)\tau_c + \sigma_0 u_+ \tag{3}$$

$$\frac{dk_c}{dt} = -(a_l|v_c| + |u|)k_c + k_0 u_+ \tag{4}$$

$$d_0(\epsilon) = -e^{-\beta_0(\epsilon_c - 0.12)^2} \tag{5}$$

where subscript $_+$ is used here to denote only positive values of the preceding variable are used. Otherwise, 0 is used instead. In equation (1), χ, a, b, α_l are positive constants, χv_c is a damping force, $\omega_0^2 \epsilon_c$ is a micro-scale harmonic force with ω_0 the oscillation frequency. $a\tau_c d_0(\epsilon_c)$ is an active force, and $b(\sqrt{V/V_0}-1)$ is a passive force. Equation (1) is derived from a simplified cylinder with constant height; hence, the strain ϵ_c and ventricle volume V are related as $\epsilon_c \propto \sqrt{V}$. u represents a 'summary' for the chemical activity (Bestel [14]) from which the contractile force derives. $d_0(\epsilon)$ is the length-tension relationship of cardiac muscle cells. Here, a shifted-Gaussian function is used, which differs from Kim and Capoccia [12] [13] and is more representative of length-tension relationships observed in experiments. Finally, σ_0 and k_0 are the maximum sarcomere tension and elastance, respectively.

2.2 Macroscale Dynamics

For the macroscale dynamics, the ventricular pressure, P_V, evolves according to

$$P_V = \gamma \frac{V_0}{V}[d_0(\epsilon_c)\tau_c + \sigma_p] \tag{6}$$

$$\sigma_p = \frac{k_2}{k_1}[\exp(k_1\sqrt{V/V_0}) - 1]_+ \tag{7}$$

where σ_p is the muscle fibre passive stress, k_1 and k_2 are positive constants for the passive tension, and γ is the ratio of ventricular left ventricular wall thickness to radius (for explanation, see Bestel [14]). Here, its value is tuned to the model rather than be physiologically precise. From equation (6), one can see how the macroscale is coupled to the microscale. Equation (7) differs from [12] [13] in that σ_p is forced to remain positive, making it more realistic.

2.3 Electro-Chemical Activity

The electro-chemical response of the microscale sarcomere to deformation behaves as follows:

$$\frac{dp}{dt} = 0.1(q - p + \mu_1\tau_c) \tag{8}$$

$$\frac{dq}{dt} = 10q(1-q^2) - 10(2\pi)^2 p + \mu_2 V_+ + 10\cos(2\pi t) \tag{9}$$

$$u = \alpha_u q \tag{10}$$

Equations (8) and (9) are adapted from the Fitzhugh-Nagumo model and describe the evolution of a slow-electric response p and a rapid-electric response q. $10\cos(2\pi t)$ is an oscillating force given a value of 1 Hz for a normal heart rate. Equation (10) relates the chemical to the rapid-electrical activity with a positive proportional constant α_u. The mechano-electric feedback is represented here in the two positive constants μ_1 and μ_2. It is apparent, then, how electrical variables q and p are coupled to the mechanical variables τ_c and V

directly through these MEF parameters. Kim and Cappoccia [13] discuss the meaning of these two values in greater detail, but in summary; μ_1 represents the influence of the stress on the action potential (excitation voltage) during ventricle contraction and is therefore associated with the effect of systolic or ejection, stretch. μ_2 is the effect of stetch during diastole (filling), since it is only active when the ventricle volume exceeds the reference value V_0 (through $_+$)

2.4 Basic Circulation Model

It is left to represent the macroscale circulation changes resulting from the above. Two models for the arterial circulation are used: a 'basic' and 'full'. For a more detailed discussion of the these models, see Kim and Capoccia [12] and Kim and Capoccia [13]. For the basic model, the system arterial circulation is ignored, and instead a focus is placed on V, the ventricular volume, and m, the aortic pressure.

$$\frac{dV}{dt} = \frac{1}{R_M}(P_R - P_V)_+ - \frac{1}{R_A}(P_V - m)_+ - \delta_p n \tag{11}$$

$$\frac{dm}{dt} = \frac{1}{C_S R_C}(m - m_0) + \frac{(P_V - m)_+}{C_A R_A} + \delta_p \frac{n}{C_A} \tag{12}$$

$$\frac{dn}{dt} = \frac{\delta p}{L*}[P_V - m - R_* n + \beta\omega^2] \tag{13}$$

Equation (13) is included to represent the blood flow, n, from a left-ventricular assist device (LVAD); a pump to assist ventricle function. δ_p is a pump parameter and ω the pump frequency. The atrial pressure parameter P_R, and the arterial pressure m_0 are taken as constants. Changing these changes the preload or afterload. R_S, R_C, R_M and R_A are the systemic, characteristic, mitral valve, and aortic valve resistances, respectively, and R_* the total pump resistance. C_S and C_A are the systemic and aortic compliances. L_* is the total pump inertance. For further discussion of the above, see [12] and [13].

2.5 Full Circulation Model

In the full circulation model, the arterial circulation from Simaan (2000) is employed to incorporate the systemic arterial circulation, resulting in the following:

$$\frac{dm}{dt} = -\frac{F_a}{C_A} + \frac{(P_V - m)_+}{C_A R_A} + \delta_p \frac{n}{C_A} \tag{14}$$

$$\frac{dF_a}{dt} = \frac{m - P_S}{L_S} - \frac{R_C F_a}{L_S} \tag{15}$$

$$\frac{dP_R}{dt} = \frac{-P_R + P_S}{R_S C_R} - \frac{(P_R - P_V)_+}{C_R R_M} \tag{16}$$

$$\frac{dP_S}{dt} = \frac{P_R - P_S}{R_S C_S} + \frac{F_a}{C_S} \tag{17}$$

where Equation (12) in the basic model has been simplified. The equations 15 to 17 describe the aortic and artrial pressures F_a and P_R respectively, and P_S the aortic flow. The new parameters R_S and L_S are the vascular resistance and aortic blood inertance. C_R is the atrial compliance.

2.6 Parameter Values and Basic Model Control Case

The MEF parameters μ_1 and μ_2 were tuned to the values 0.0024 kpa^{-1} and 0 (s ml)$^{-1}$ in Kim and Cappoccia [12], which are used here for a control case. For brevity, variables and parameter values are not provided here but are listed in tables 1 and 2 in Kim and Cappoccia [12], [13] along with their units and physiological meaning. Some changes have been made, however: equations (5) and (7) have been modified and as a result, $\beta_0 = 20\frac{2}{3}$.

3 Results and Discussion

The control case for both the basic and full models employs the values given in Table 2 in Kim and Capoccia [13] and without pump support ($\delta_p = 0$). In what follows, the MEF parameters μ_1 and μ_2 are changed independently and the noteworthy cases are described herein. Particular attention is paid on bifurcations and changes in period. The influence of an axial pump assist device ($\delta_p = 1$) on conditions of a failing heart is then investigated.

3.1 MEF in basic model

Starting with the basic model, the results of the control case using the parameter values $\mu_1 = 0.0024$ kpa^{-1}, $\mu_2 = 0$ (s ml)$^{-1}$, $P_R = 9$ mmHg, $m_0 = 70$ mmHg, and initial left ventricular volume $V(0) = 0.5V_0$ are presented in figure 1. Initial transients in the figure and all figures that follow are removed unless otherwise stated. From left to right, figure 1 shows the left ventricular pressure-volume

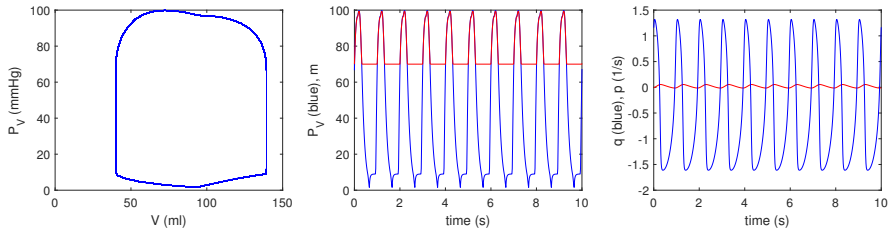


Fig. 1: Basic model: control case.

(P-V) loop, the time evolution of the aortic pressure m (red) with ventricular pressure P_V (blue), and the time evolution of the slow (red) and rapid (blue) electrical variables p and q. The P-V loop is an important expression of the global cardiac pump function. The output stroke volume (SV) is clearly visible as the difference between End-Systolic volume (ESV) and End Diastolic volume (EDV).

3.2 Influence of μ_1 ($\mu_2 = 0$)

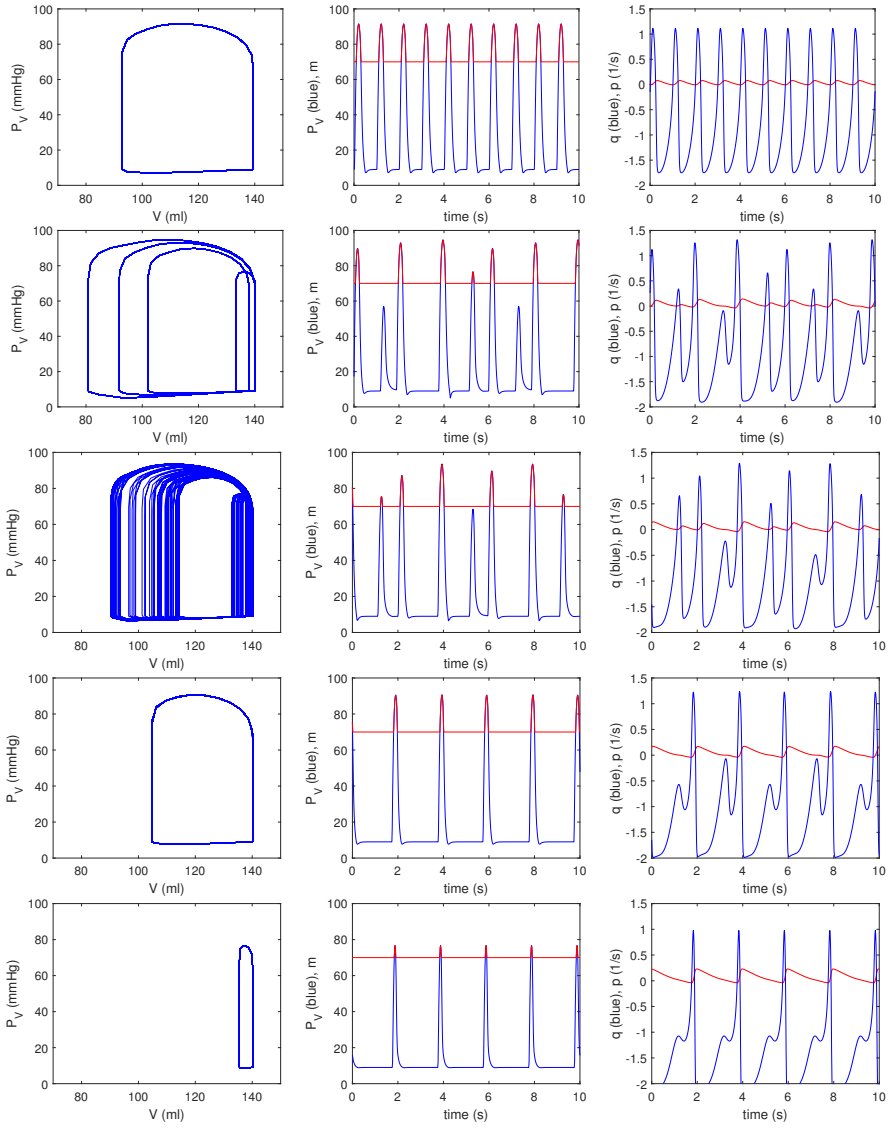


Fig. 2: Basic model: $\mu_1 = 0.0008 \times [9, 17, 21, 31, 100]$ from top to bottom; $\mu_2 = 0$.

The effect of the MEF parameter μ_1 is now investigated in the basic model. Figure 2 shows μ_1 increasing from top to bottom with the same figures given from left to right as figure 1. In general, an increase in μ_1 reduces the stroke volume of the P-V loop, as the end-systole pressure rises. The last plot in each

row shows that the slow electrical variable p (red) rises further and more rapidly, then falls more gradually. The correspondence to Amar *et al.* [7] and Khol *et al.* [17] becomes evident if the slow excitation variable is associated with the action potential, in that the MEF tends to prolong the action potential through a slower, late excitation phase of the sarcomere membrane potential. Knudsen *et al.* [2] also find an increase in MEF coupling accompanies an increase in action potential duration. Until decline in action potential is sufficient, a further excitation cannot take place.

The increase in action potential duration accompanies system bifurcation in the second row. The ventricular pressure time trace (blue) in that row shows weaker rises and in some instances, no rise in the aortic pressure occurs (red trace) resulting in missed beats. These missed beats are caused by the fall in the peak of the fast electrical activity. By counting the number of beats in ten seconds and comparing this to the first row, one finds a reduction in beats overall. Further rises in μ_1 sees a transition to quasi-periodic ($2 \geq$ unequal frequencies) bifurcations. The Fourier spectrum (not shown) shows strong peaks around multiples of 1/2 and in-between. This is followed by further reductions in the rapid-electric variable q, triggering reductions in P_V and m, accompanied by missed beats and reduced stroke volume. In the last row, the reduced value of q is just large enough to trigger contractions. Further rises in μ_1 result in a complete breakdown of the system.

3.3 Influence of μ_2 ($\mu_1 = 0.0024$)

With μ_1 set back to its control case value, the second MEF parameter μ_2 is now increased and the results are displayed in figure 3. In the top row, $\mu_2 = 0.18 \times 0.87$ and the system still appears regular. The rapid-electric variable q, however, shows a slight difference around the minimum value. In the second row, despite a small change in μ_2, the system again becomes quasi-periodic. Additional, ectopic contractions appear, with 11 beats taking place in the 10 second sample. As argued in Kim and Capoccia [13], these ectopic beats are caused by additional 'incommensurate' frequencies. The extra beats grow in number as μ_2 increases, with 15 contractions clearly visible in the fourth row. The rapid-electric variable also grows in strength and frequency. The slow variable, in contrast, shows a slight reduction and a degradation in the stroke volume, although not as pronounced as that seen above, results.

3.4 MEF in Full Circulation Model

Turning to the full model, the control case behaviour with no pump support and the initial conditions $P_R(0) = 10$ mmHg, $P_S(0) = 70$ mmHg, $m(0) = 70$ mmHg, $F_a(0) = 90$ mL/s, and $V(0) = 0.9V_0$, is shown in figure 4. Recall, the full model incorporates the dynamics of the system arterial circulation.

3.5 Influence of μ_1 ($\mu_2 = 0$)

Keeping $\mu_2 = 0$, μ_1 is increased and the results are displayed in figure 5. From top to bottom, figure 5 shows μ_1 increasing as $\mu_1 = 0.0008 \times [15.5, 25, 35, 500, 1100]$.

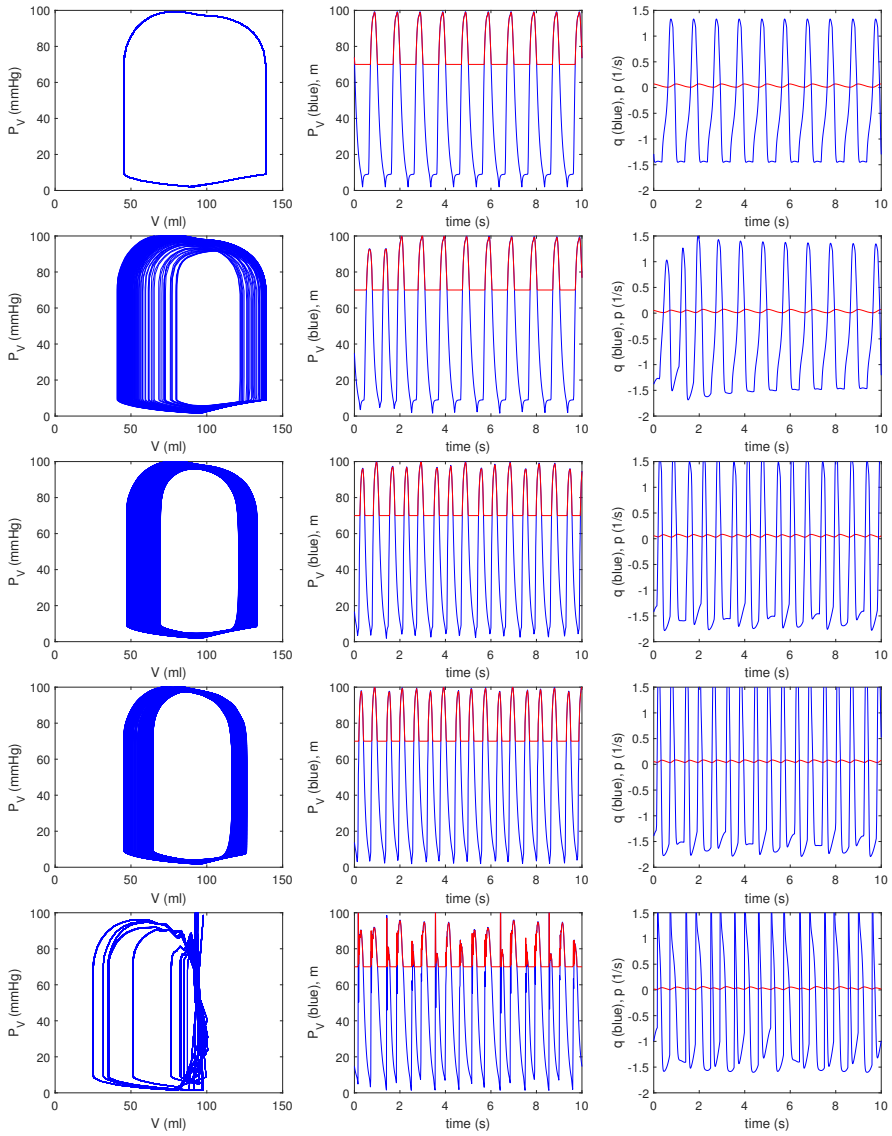


Fig. 3: Basic model: $\mu_1 = 0.0024$; $\mu_2 = 0.18 \times [0.87, 0.88, 2, 3, 50]$ from top to bottom.

As μ_1 increases, the P-V loop shifts to the right, though the stroke volume remains roughly the same. When μ_1 reaches 0.0008×15.5, period doubling occurs. A reduction in the rapid-electric activity occurs in a similar way to the basic model, along with the increase in the slow electric activity peak and a gentler decline.

As μ_1 increases further, period three and four are reached. The changes noted above for the electric activity continue and as might be expected from

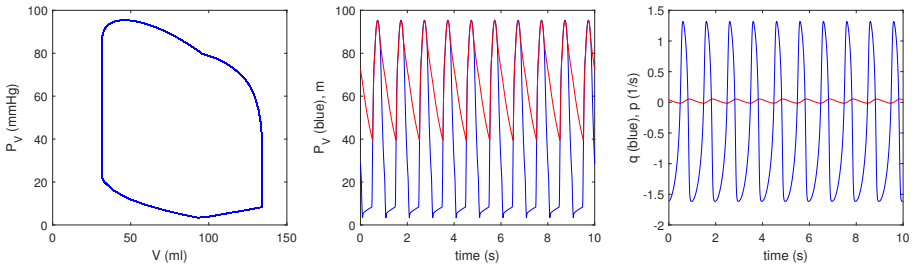


Fig. 4: Full model: control case.

the experience in the basic model, fewer beats result. Indeed, the changes that occur in the full model are similar to the basic model, though only qualitatively. The stroke volume decreases along with a reduction in the ventricular pressure. The reduction in ventricular pressure is more severe than the basic model.

3.6 Influence of μ_2 ($\mu_1 = 0.0024$)

With μ_1 set back to its control value, μ_2 is now increased and the results are shown in figure 6. In a similar pattern to the basic model, an augmentation of fast-excitation activity is accompanied by a degradation of slow activity. The reduction of the slow activity implies that the duration of the action potential is also reduced, allowing a further excitation cycle to take place sooner. Consequently, as already seen in the basic model, the excitation pace increases and additional ectopic beats appear, growing in number as μ_2 is increased. By the time μ_2 reaches 0.18×10, shown in the bottom row of figure 6, 25 beats take place in the 10 second sample. The P-V loop representation becomes complex due to the number of oscillations and the stroke volume falls.

3.7 Introduction of a Pump Device

The introduction of a ventricular pump assist device is a well-established method of compensating for defective heart function. The ability of a rotary pump to reverse the conditions seen in the previous sections, where arrythmia result from abnormal MEF, is now investigated. Starting with the basic model, the quasi-periodic scenario $\mu_1 = 0.0008 \times 21$, $\mu_2 = 0$, shown in the third row of figure 3, is employed and δ_p in equation (13) is set to 1. Figure 7 shows the results using a pump speed $\omega = 8000$ rpm. The results show a restoration of regular heart rhythm and an improvement in the stroke volume and pumping power of the heart. The aortic pressure also increases. The number of beats remains low however, caused by the fast-electric excitations q, half of which remain too small to generate contraction. Nevertheless, the introduction a heart assist device is seen to have a positive effect on heart arrythmia, induced through a malfunction of the MEF. Faster pump speeds were investigated, though conditions are only marginally improved. Indeed, as pump speeds reach $\omega = 12$ krpm, arrythmia return. Introduction of a ventricular pump is now investigated in the

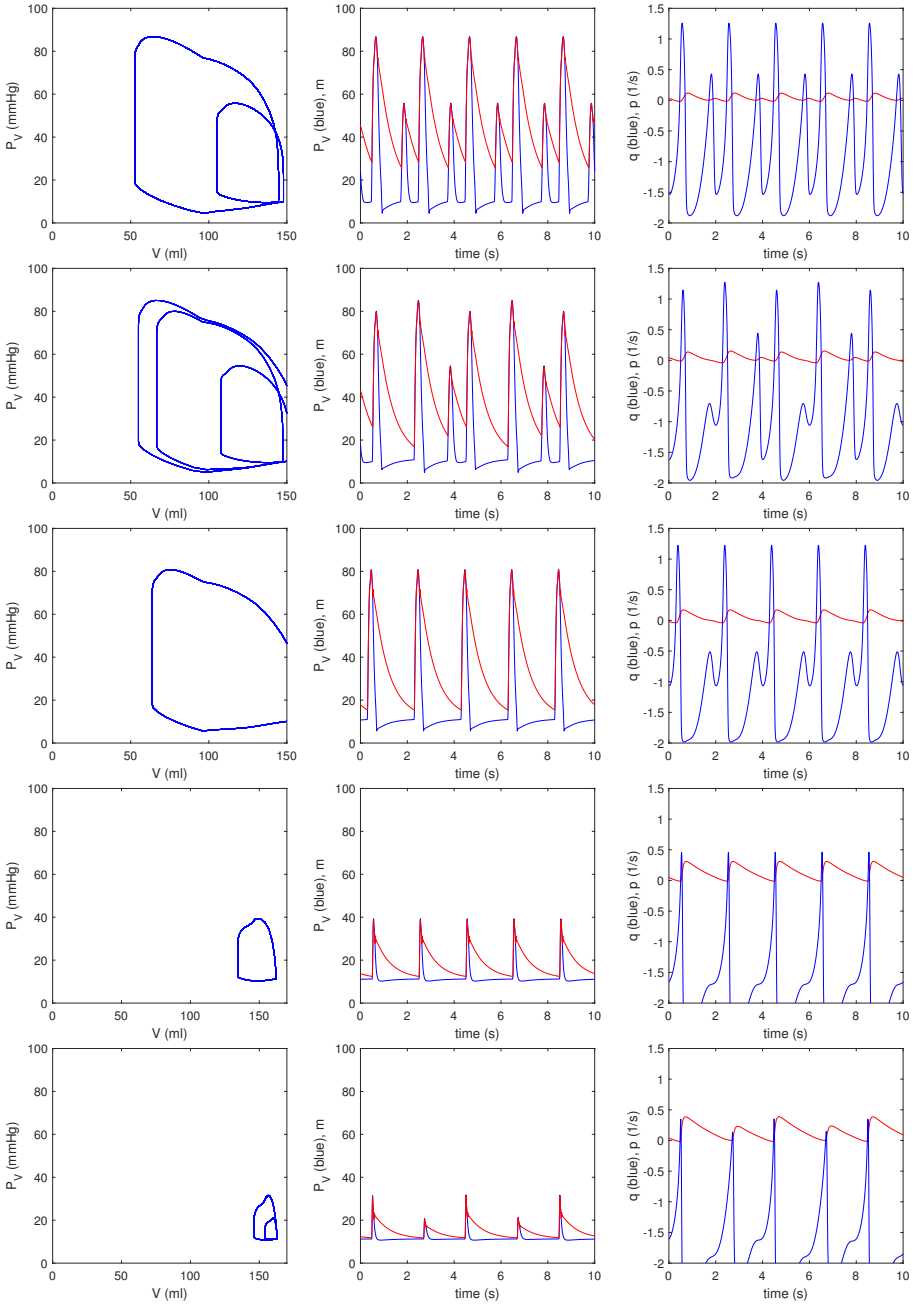


Fig. 5: Full model: $\mu_1 = 0.0008 \times [15.5, 25, 35, 500, 1100]$ from top to bottom; $\mu_2 = 0$.

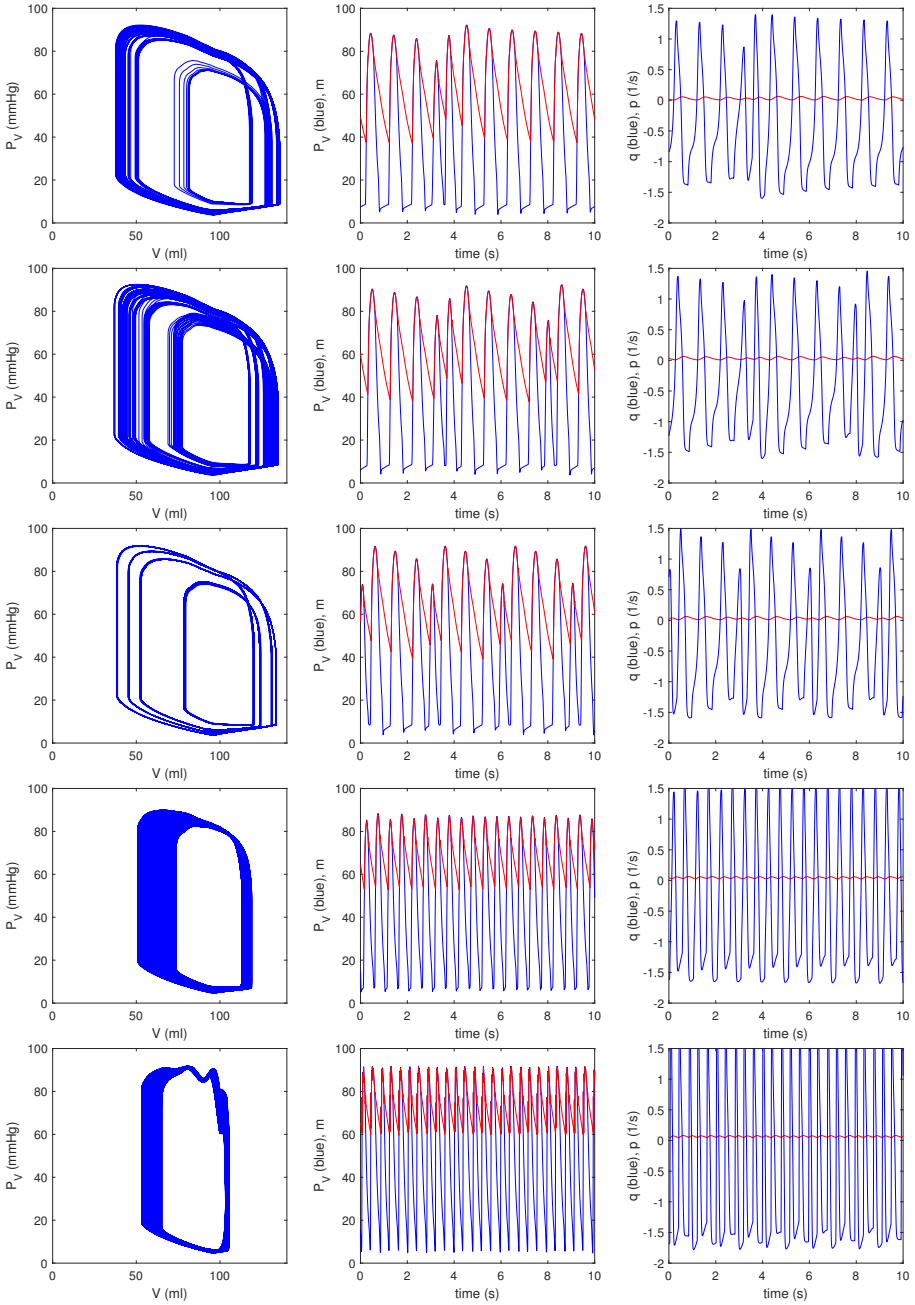


Fig. 6: Full model: $\mu_1 = 0.0024$; $\mu_2 = 0.18 \times [0.85, 0.9, 0.974, 2, 10]$ from top to bottom.

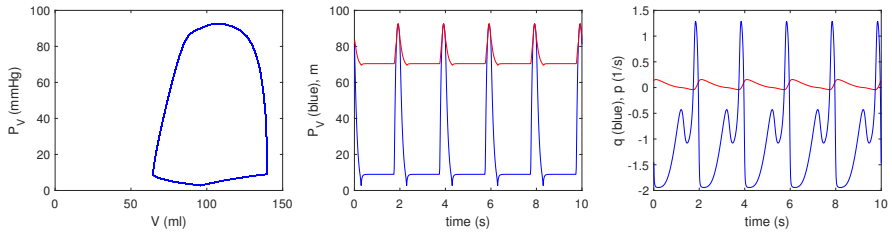


Fig. 7: Basic model with pump support: $\mu_1 = 0.0008 \times 21$; $\mu_2 = 0$, $\omega = 8000$ rpm.

full model. Choosing the scenario shown in the fourth row of figure 5, that is $\mu_1 = 0.0008 \times 500$, $\mu_2 = 0$, the results using a pump speed $\omega = 13.3$ krpm are shown in figure 8.

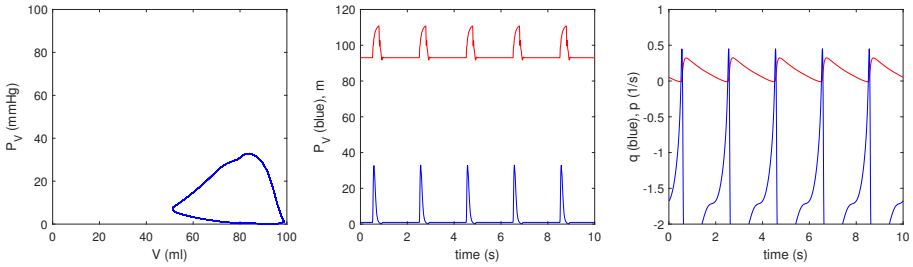


Fig. 8: Full model with pump support: $\mu_1 = 0.0008 \times 500$; $\mu_2 = 0$; $\omega = 13.3$ krpm.

Though the restoration of cardiac function is not as good as in the basic model, a slight improvement in stroke volume is experienced, the ventricular volume reduces overall, and the aortic pressure increases. The heart rhythm is not improved though. Higher pump speeds again, see a return of arrhythmia.

4 Conclusion

MEF has been investigated using a low-order, lumped parameter model which links the microscale electro-chemical dynamics to the macroscale mechanical action thereby negating the time-varying elastance concept often relied up to model the sarcomere compliance. A dysfunction of MEF is seen to lead to pathological conditions and arrythmia through period doubling bifurcations. Dysfunction during systolic stretch is found to increase the action potential duration by increasing the late stage 'recovery' of the excitation cycle. The initial rise in action potential falls, eventually leading to missed contractions. feedback dysfunction during diastole leads to additional ectopic beats as the initial rise in action potential increases. Stoke volume and pumping power of

the heart are reduced in both cases. Systemic conditions can be improved using a rotary pump of specified frequency.

References

1. World Health Organization. *Global Health Estimates 2016: Deaths by Cause, Age, Sex, by Country and by Region, 2000-2016*, Geneva, 2018.
2. Z. Knudsen, A. Holden, J. Brindley, *Qualitative modeling of mechano-electrical feedback in a ventricular cell*, Bull. Math. Biol., 115–181 **6**, 1997.
3. R. FitzHugh, *Impulses and physiological states in theoretical models of nerve membrane*, Biophysical J., 445-466 **1**, 1961.
4. A. Collet, J Bragard, P.C. Dauby *Temperature, geometry, and bifurcations in the numerical modeling of the cardiac mechano-electric feedback*, Chaos Interdiscip. J. Nonlinear Sci, **27**, 2017.
5. Z. Qu, G. Hu, A. Garfinkel, J. N. Weiss, Nonlinear and stochastic dynamics in the heart, Phys. Rep., 61–162 **543**, 2014.
6. C. Franzone, L. F. Pavarino, S. Scacchiet, *Effects of mechanical feedback on the stability of cardiac scroll waves: A bidomain electro-mechanical simulation study*, Chaos, 093905 **27**, 2017.
7. A. Amar, S. Zlochiver, O. Barnea, Mechano-electric feedback effects in a three-dimensional (3D) model of the contracting cardiac ventricle, PLoSONE, e0191238, **13**, 2018.
8. M. Radszuweit, E. Alvarez-Lacalle, M. Bar, B. Echebarria, *Cardiac contraction induces discordant alternans and localized block*, Phys. Rev. E, 022703, **91**, 2015.
9. K. Sagawa, *The ventricular pressure-volume diagram revisited*, Circ. Res., 677–687, **43**, 1978.
10. M. A. Simaan, A. Ferreira, S. Chen, J. F. Antaki, D. G. Galati, *A dynamical state space representation and performance analysis of feedback-controlled rotary left ventricular assist device*, IEEE Trans, Control Syst. Tech.,, 15–28, **17** 2019.
11. T. E. Claessens, D. Georgakopoulos, M. Afanasyeva, S. J. Vermeersch, H. D. Millar, N. Stergiopulos, et al., *Nonlinear isochrones in murine left ventricular pressure-volume loops: how well does the time-varying elastance concept hold?*, Am. J. Physiol. Heart Circ. Physiol., **290**, 2006
12. E. Kim, M. Capoccia, *Synergistic model of cardiac function with a heart assist device*, Bioengineering, 1–16, **7** 2019.
13. E. Kim, M. Capoccia, *Mechano-electric effect and a heart assist device in the synergistic model of cardiac function*, Mathematical Biosciences and Engineering, 17(5): 5212-52331–16, 2020.
14. J. Bestel, *Modele differentiel de la contraction musculaire controlee: Application au systeme cardio-vasculaire*, Ph.D thesis, Universit Paris 9, 2000.
15. J. Bestel, F. Clément, M. Sorine, *A biomechanical model of muscle contraction*, International Conference on Medical Image Computing and Computer-Assisted Intervention; Goos, G., Hartmanis, J., van Leeuwen, J., Eds.; Springer: Berlin, Germany, 1159–1161, 2001.
16. J. Sainte-Marie, D. Chapelle, R. Cimrman, M. Sorine, *Modeling and estimation of the cardiac electromechanical activity*, Computers And Structures, 1743-1759, **84**, 2006.
17. P. Kohl, P. Hunter, D. Noble, *Stretch-induced changes in heart rate and rhythm: clinical observations, experiments and mathematical models*, Biophysics and Molecular Biology, 91–138 **71** , 1999.

18. Kohl P, Bollensdorff C, Garny A. *Effects of mechanosensitive ion channels on ventricular electrophysiology: experimental and theoretical models*. Exp Physiol. Mar; 91(2):307-21, 2006 .
19. Link MS, Wang PJ, Pandian NG, Bharati S, Udelson JE, Lee MY, Vecchiotti MA, VanderBrink BA, Mirra G, Maron BJ, Estes NA 3rd. *An experimental model of sudden death due to low-energy chest-wall impact (commotio cordis)*. N Engl J Med. Jun 18;338(25):1805-11. 1998

Simulations on the peridynamic equation in continuum mechanics

Sabrina Francesca Pellegrino[1]

Department of Mathematics, Università degli Studi di Bari Aldo Moro, Via E. Orabona, 4, 70125 Bari, Italy
(E-mail: sabrina.pellegrino@uniba.it)

Abstract. The peridynamic equation of motion consists in a second order in time partial integro-differential equation and is largely used in elastodynamics as it is able to model cracks avoiding the spatial partial derivatives. In this paper, we consider the linear model of peridynamics in a one-dimensional spatial domain. We review some numerical techniques to solve this equation and propose some new computational methods of higher order in space. Several numerical tests are given in order to validate our results.
Keywords: Peridynamics, Non-local models, Quadrature formula, Trigonometric scheme.

1 Introduction

Modeling fractures and damages is one of the major issue in the framework of continuum mechanics. The classical theory uses spatial derivatives to model the motion of a material subject to elastic stresses. So, it is not able to describe discontinuous phenomena like cracks and fractures, as partial derivatives are not defined on discontinuities, and, moreover, it cannot predict where the crack is located. Therefore, there is the need to develop a non-local theory able to use a unique equation both on or off a crack, see [14,23,9,8,3].

Recent studies show that differential operators of fractional orders may be introduced in order to depict the nature of such phenomena, see for instance [6,15,16,18,17,20,19].

The peridynamic theory is a non-local generalization of the elasticity theory introduced by Silling in [34], and has attracted the attention of a growing number of researchers, as it addresses discontinuous problems. He proposed to model the motion of a material body using integro-differential partial equations, without involving spatial derivatives. The main assumption of the theory concerns the presence of an interacting force f between the particle x and the particle $\hat{x}$ belonging to V_x, which represents the peridynamic neighborhood of x. This basic assumption also suggests that peridynamics could be suitable for multiscale material modeling ([24,32]).

We fix $[0,T]$, for some $T>0$, as the time domain under consideration, and let $V \subset \mathbb{R}$ be the rest configuration of a material body having mass density $\rho : V \times [0,T] \to \mathbb{R}_+$. Then, the peridynamic equation is given by

$$\rho(x)u_{tt}(x,t) = \int_{V_x} f(\hat{x}-x, u(\hat{x},t)-u(x,t))d\hat{x} + b(x,t), \quad x \in V, \quad T \in [0,T], \tag{1}$$

usually enriched by the initial conditions

$$u(x,0) = u_0(x), u_t(x,0) = v(x), \qquad x \in V, \tag{2}$$

where u is the displacement field and b describes all the external forces acting on the material body. The integrand f represents the force density that the particle $\hat{x}$ exerts on the particle x and is called **pairwise force function**, see for instance [10,34]. The interaction between x with all particle in the peridynamic neighborhood V_x is called **bond**.

We set

$$\xi = \hat{x} - x, \qquad \eta = u(\hat{x},t) - u(x,t),$$

which denotes the relative position of two particles in the reference configuration and the relative displacement, respectively. Thus $\xi + \eta$ represents the current relative position vector, and we can observe that the pairwise force function f has to satisfy Newton's third law and the conservation of the angular momentum:

$$f(-\xi,-\eta) = -f(\xi,\eta), \qquad \eta \times f(\xi,\eta) = 0. \tag{3}$$

It is reasonable to require the existence of a positive constant δ, called **horizon**, such that there are no interactions among particles having relative distance greater than δ, namely

$$f(\xi,\eta) = 0, \qquad \text{for } |\xi| > \delta \text{ and for every } \eta.$$

In what follows, we restrict our attention to the case of an homogeneous bar of infinite length, and in particular we focus on the **linear peridynamic** model

$$\rho(x)u_{tt}(x,t) = \int_{x-\delta}^{x+\delta} C(\hat{x}-x)\,(u(\hat{x},t) - u(x,t))\,d\hat{x} + b(x,t), \qquad x \in \mathbb{R}, \quad t \in [0,T], \tag{4}$$

where the function C is a non-negative even function, i.e. $C(-\xi) = C(\xi)$, called **micromodulus function**.

In this paper, we survey some numerical techniques for the model and we propose an accurate spatial discretization accompanied to a trigonometric scheme for time integration. Additionally, we extend the methods to the non linear case.

The paper is organized as follows. Section 2 collects the main analytic results for the problem. In Section 3 we presents some quadrature formula for the space discretization of the model. Section 4 is devoted to the time integration techniques. In Section 5 we extend the proposed methods to the non linear case. Section 6 shows the numerical tests, and, finally, Section 7 concludes the paper.

2 Analytical results

In this section we recall the main theoretical results concerning the peridynamic equation. The well-posedness of the non linear model depends on the assumptions on the pairwise force function f, see [6,10,11].

Theorem 1 ((see [10])). *Assume that* u_0, $v \in \mathcal{C}(\overline{V})$ *and* $b \in \mathcal{C}([0,T];\mathcal{C}(\overline{V}))$. *If the pairwise force function* $f : \overline{B_\delta(0)} \times \mathbb{R}^d \to \mathbb{R}^d$ *is a continuous function, such that there exists a nonnegative function* $\ell \in L^1(B_\delta(0))$ *such that*

$$|f(\xi,\hat{\eta}) - f(\xi,\eta)| \leq \ell(\xi)|\hat{\eta} - \eta|, \quad \text{for all } \xi \in \mathbb{R}^d \text{ and } \eta,\ \hat{\eta} \in \mathbb{R}^d,$$

then, the integral operator in (1) *is well-defined and Lipschitz continuous, and the initial-value problem* (1)-(2) *is globally well-posed with solution* $u \in \mathcal{C}^2([0,T];\mathcal{C}(\overline{V}))$.

For a **microelastic** material (see [34]), it is possible to derive the pairwise force function $f(\xi,\eta)$ from a scalar-valued function $w(\xi,\eta)$ called **pairwise potential function** (see [13]), such that

$$f(\xi,\eta) = \nabla_\eta w(\xi,\eta), \tag{5}$$

and the peridynamic equation (1) follows from the variational problem: find

$$u = \arg\min J(u)\ , \qquad J(u) = \int_0^T \int_V e(x,u(x,t),t)dxdt, \tag{6}$$

where $e = e_{kin} - e_{el} - e_{ext}$ is the Lagrangian density, and incorporates the kinetic energy density, the elastic energy density and the density due to the external force density, given respectively by

$$\begin{aligned} e_{kin} &= \frac{1}{2}\rho(x)\ u_t^2(x,t), \\ e_{el} &= \frac{1}{2}\int_V w(\hat{x}-x, u(\hat{x},t) - u(x,t))d\hat{x}, \\ e_{ext} &= -b(x,t)u(x,t). \end{aligned}$$

In particular, in the one-dimensional linear peridynamic model (4), the pairwise force function is given by

$$f(\xi,\eta) = C(\xi)\ \eta, \tag{7}$$

and the potential function is given by

$$w(\xi,\eta) = \frac{1}{2}C(\xi)\eta^2. \tag{8}$$

The following well-posedness result holds.

Theorem 2 ((see [13])). *Let* u_0, $v \in \mathcal{C}^0(\mathbb{R})$ *be the initial conditions. If the micromodulus function* C *belongs to* $\mathcal{C}^2(\mathbb{R})$, *then, for any* $T > 0$, *the initial value problem* (4)-(2) *is well-posed and the unique solution* u *belongs to* $\mathcal{C}^2([0,T];\mathcal{C}(\mathbb{R}))$.

Moreover, if b is autonomous, namely $b(x,t) \equiv b(x)$, we can prove the conservation of the total energy of the system.

Theorem 3. *If the external force b does not depend on time, the total energy associated to* (4)*, given by*

$$E(t) = \frac{1}{2}\int_V \rho\ |u_t(x,t)|^2\,dx + \frac{1}{2}\int_V\int_V w(\hat{x}-x, u(\hat{x},t)-u(x,t))dx\,d\hat{x}, \quad (9)$$

is preserved, namely

$$\frac{d}{dt}E(t) = 0.$$

Proof. We set $\xi = \hat{x} - x$. Then, using (3) and (1) we have

$$\begin{aligned}
\frac{d}{dt}\,E(t) &= \int_V \rho\ u_t(x,t)u_{tt}(x,t)\,dx + \frac{1}{2}\int_V\int_V f(\xi, u(x,t)-u(x-\xi,t))u_t(x,t)\,dx\,d\xi\\
&\quad - \frac{1}{2}\int_V\int_V f(\xi, u(x,t)-u(x-\xi,t))u_t(x-\xi,t)\,dx\,d\xi\\
&= \int_V \rho u_t(x,t)u_{tt}(x,t)\,dx + \frac{1}{2}\int_V\int_V f(\xi, u(x,t)-u(x+\xi,t))u_t(x+\xi,t)\,dx\,d\xi\\
&\quad - \frac{1}{2}\int_V\int_V f(\xi, u(x,t)-u(x-\xi,t))u_t(x-\xi,t)\,dx\,d\xi\\
&= \int_V \rho u_t(x,t)u_{tt}(x,t)\,dx + \frac{1}{2}\int_V\int_V f(\xi, u(x,t)-u(x-\xi,t))u_t(x,t)\,dx\,d\xi\\
&\quad + \frac{1}{2}\int_V\int_V f(\xi, u(x,t)-u(x-\xi,t))u_t(x,t)\,dx\,dy\\
&= \int_V u_t(x,t)\left(\rho u_{tt}(x,t) + \int_V f(\xi, u(x,t)-u(x-\xi,t))\,d\xi\right)dx = 0.
\end{aligned}$$

The following result is related to the case of non autonomous external force.

Theorem 4 ((see [13])). *If the external force is not autonomous, then, the Lagrangian density associated to the linear problem* (4) *satisfies the following inequality*

$$\begin{aligned}
&e_{kin}(t) + e_{el}(t) + \nu\int_0^t e^{\nu(t-s)}e_{ext}(s)ds\\
&\qquad \le e^{\nu t}(e_{kin}(0) + e_{el}(0)) + \frac{1}{2\nu}\int_0^t\int_{-\infty}^{\infty}\frac{e^{\nu(t-s)}}{\rho}|b(x,t)|^2dxds,
\end{aligned}$$

for all $\nu > 0$ and $t > 0$.

Additionally, in [6], the authors proved the well-posedness of the nonlinear peridynamic equation assuming very general constitutive assumptions in the framework of fractional Sobolev spaces.

Moreover, we can think to the linear one-dimensional peridynamic equation (4) as a non local version of the classical linear one-dimensional wave equation, (see for instance [12,4]). Indeed, if we choose $u_0(x) = U\exp[(-x/L)^2]$, $v(x) = 0$, with U and L suitable constants, as initial conditions and the following micromodulus function

$$C(\hat{x}-x) = 4E\exp[-(\hat{x}-x)^2/l^2]/(l^3\sqrt{\pi}), \qquad \hat{x}, x \in \mathbb{R}\ , \quad (10)$$

where E denotes the Young modulus, and $l > 0$ a length-scale parameter, then for $l \to 0$, (4) becomes the wave equation of the classical elasticity theory:

$$\rho\ u_{tt}(x,t) = Eu_{xx}(x,t) + b(x,t), \qquad x \in \mathbb{R},\ t \geq 0\ , \tag{11}$$

As a consequence, the parameter l can be seen as a degree of non locality.

3 Spatial discretization of the peridynamics

In this section we discretize in space the equation (4) by means of a quadrature formula.

Let $N > 0$ be an even integer and $\Delta x > 0$ be the spatial step size. We discretize the spatial domain $\mathbb{R}$ by a compact set $[-D, D]$, for some positive large constant D, and such interval by means of the points $x_j = -D + jh = -D + j\frac{2D}{N}$, for $j = 0, \ldots, N$. We consider a quadrature formula of order s on these points:

$$\int_{-\infty}^{\infty} C(\hat{x} - x)(u(\hat{x},t) - u(x,t))d\hat{x} \approx \Delta x \sum_{j=0}^{N} w_j\ C(x_j - x)\left(u(x_j,t) - u(x,t)\right), \tag{12}$$

where w_j are the weights of the formula.

Then, at the collocation points $x = x_i$ for $i = 0, \ldots, N$, we approximate the equation (4) by

$$\rho u_{tt}(x_i,t) \approx h \sum_{j=0}^{N} w_j C(x_j - x_i)(u(x_j,t) - u(x_i,t)) + b(x_i,t), \qquad t \geq 0. \tag{13}$$

We define the **stiffness matrix** $K = (k_{ij})$, for i, $j = 0, \ldots, N$ by

$$k_{ij} = \alpha_i \delta_{ij} - w_j C_{ij},$$

where $C_{ij} = C(x_j - x_i)$, $\alpha_i = \sum_{k=0}^{N} w_k C_{ik}$, and δ_{ij} is the Kronecker Delta.

The stiffness matrix K is not symmetric, in general, unless the weights are constant with respect to j, namely $w_j = w$ for every $j = 0, \ldots, N$. In this case we obtain the **composite midpoint rule**: we approximate the spatial domain $(-\infty, \infty)$ by the interval $[-(N+1)h/2, (N+1)h/2]$ and the points of the discretization x_j^{MR} are taken as the midpoints of the sub-intervals $[-(N+1)h/2 + jh, -(N-1)h/2 + jh]$, for $j = 0, \ldots, N$. For sufficiently smooth assumptions on C and u, this formula is of the second order of accuracy in space and the constant weights are equal to 1, (see for instance [13,33]).

The **composite Gauss two points formula** is of the fourth order of accuracy and provides a symmetric stiffness matrix K. We fix $M > 0$ and consider a partition of the interval $[-D, D]$ given by the sequence $\tilde{x}_j = -D + j\Delta x$ for $j = 0, \ldots, M$, where $\Delta x = 2\ D/M = (\tilde{x}_M - \tilde{x}_0)/M$. Then on each sub-interval $[\tilde{x}_{j-1}, \tilde{x}_j]$, the formula uses two points where the function $\psi(x)$ is evaluated:

$$\int_{\tilde{x}_0}^{\tilde{x}_M} \psi(x)dx \approx \frac{\Delta x}{2} \sum_{j=1}^{M} \left[\psi(m_j^-) + \psi(m_j^+)\right], \tag{14}$$

where

$$m_j = \frac{\tilde{x}_{j-1} + \tilde{x}_j}{2}, \qquad m_j^- = m_j - \frac{\Delta x}{2\sqrt{3}}, \qquad m_j^+ = m_j + \frac{\Delta x}{2\sqrt{3}}, \qquad j = 1, \ldots, M.$$

Setting

$$x_j = \begin{cases} m^-_{\frac{j+1}{2}}, & \text{if } j \text{ is even,} \\ m^+_{\frac{j+1}{2}}, & \text{if } j \text{ is odd,} \end{cases} \qquad j = 0, \ldots, N,$$

with $N = 2M - 1$, then we can rewrite the quadrature formula (14) in the following way

$$\int_{x_0}^{x_M} \psi(x)dx \approx \frac{\Delta x}{2} \sum_{j=1}^{M} \left[\psi(m_j^-) + \psi(m_j^+)\right] = \frac{\Delta x}{2} \sum_{j=0}^{N} \psi(x_j),$$

so, in this case the constant weights w_j are equal to $\frac{1}{2}$, for $j = 0, 1, \ldots, N$.

Remark 1. Using the composite midpoint rule, or the composite Gauss two points formula, the stiffness matrix K is a positive and semi-definite with non-negative eigenvalues. In general K is not sparse because of the infinite horizon, however, in case of finite horizon $\delta > 0$ (see [5,33]), that is $C(x - \hat{x}) = 0$, when $|\hat{x} - x| > \delta$, then K has a banded structure, and, in particular, the size of the band r depends on the horizon δ and on the space step Δx and is given by $r = \lfloor \delta/h \rfloor$.

3.1 The semidiscretized problem

Let x_j be the spatial nodes for $j = 0, \ldots, N$ and

$$U(t) = [U_0(t), U_1(t), \ldots, U_N(t)],$$

be an approximation of the solution, where $U_j(t) \approx u(x_j, t)$ for $j = 0, \ldots, N$.

We set

$$B(t) = \frac{1}{\rho}[b(x_0, t), \ldots, b(x_N, t)]^T.$$

Then, we can approximate the peridynamic equation (4) by the following second order differential system

$$U''(t) + \Omega^2 \, U(t) = B(t), \tag{15}$$

with $\Omega^2 = \dfrac{\Delta x}{\rho} K$ (or $\Omega^2 = \dfrac{\Delta x \; w}{\rho} K'$, for K' depending only on the micro-modulus function C), where K is a positive semi-definite matrix, and with the initial conditions

$$U_0 = [u_0(x_0), \ldots, u_0(x_N)]^T \quad \text{and} \quad V_0 = [v(x_0), \ldots, v(x_N)]^T.$$

The system (15) is equivalent to the following first order differential system

$$\begin{pmatrix} U' \\ V' \end{pmatrix} = \begin{pmatrix} 0 & I \\ -\Omega^2 & 0 \end{pmatrix} \begin{pmatrix} U \\ V \end{pmatrix} + \begin{pmatrix} 0 \\ B(t) \end{pmatrix}, \tag{16}$$

where $V = U'$, with the initial conditions U_0 and V_0. Therefore, we can write the exact solution of (16) as

$$\begin{pmatrix} U(t) \\ V(t) \end{pmatrix} = \exp(tA) \begin{pmatrix} U_0 \\ V_0 \end{pmatrix} + \int_0^t \exp[(t-s)A] \begin{pmatrix} 0 \\ B(s) \end{pmatrix} ds, \tag{17}$$

with $A = \begin{pmatrix} 0 & I \\ -\Omega^2 & 0 \end{pmatrix}$.

Remark 2. In order to avoid computational problems, particularly, when we will consider trigonometric schemes where the square root Ω of Ω^2 is required or the inverse of Ω is necessary, we can regularize the matrix Ω^2 by adding a diagonal matrix of the form $(\Delta x)^s\ I$, where s is the order of accuracy of the quadrature formula used. In this way, the matrix Ω^2 is symmetric and positive definite, so it admits a unique symmetric and definite-positive square root Ω.

4 Time discretization of the peridynamics

In this section we consider the time discretization of the semidiscretized system (16) obtained by applying a quadrature formula to the original problem. Let $\Delta t > 0$ be the time step and $t_n = n\Delta t$ be the partition of the time interval $[0, T]$, for $n = 0, \ldots, N_T$, where $N_T = \left\lfloor \frac{T}{\Delta t} \right\rfloor$, and let $U_n \approx U(t_n)$ and $V_n \approx U'(t_n)$.

We describe standard time discretization schemes, such as the Störmer-Verlet scheme, the implicit midpoint method, and a new procedure based on a **trigonometric** approach.

4.1 Störmer-Verlet scheme

This is a symplectic, second order in time, explicit scheme:

$$\begin{cases} V_{n+\frac{1}{2}} = V_n + \frac{\Delta t}{2}[-\Omega^2 U_n + B(t_n)], \\ U_{n+1} = U_n + \Delta t\ V_{n+\frac{1}{2}}, \\ V_{n+1} = V_{n+\frac{1}{2}} + \frac{\Delta t}{2}[-\Omega^2 U_{n+1} + B(t_{n+1})]. \end{cases} \tag{18}$$

Because of its geometric properties, Störmer-Verlet method is widely used in the context of partial differential equations of wave propagation or peridynamic problems, (see [7,25,36]).

In [30], the authors perform the von Neumann analysis to study the stability of the Störmer-Verlet scheme.

Theorem 5 ((see [7])). *Let Δx, $\Delta t > 0$ be the space and the time steps, respectively, and let $N > 0$, even, be the points' number used to discretize in space the linear model* (4)*. If*

$$\Delta t < \sqrt{\frac{\rho}{\Delta x \sum_{j=-N/2}^{N/2} C_{ij}}}, \tag{19}$$

where $C_{ij} = C(x_i - x_j)$, then the Störmer-Verlet method (18) *is numerically stable.*

4.2 Implicit Midpoint scheme

This is a symplectic implicit second order scheme:

$$\begin{cases} U_{n+1} = U_n + \frac{\Delta t}{2}(V_{n+1} + V_n), \\ V_{n+1} = V_n + \frac{\Delta t}{2}[-\Omega^2(U_{n+1} + U_n) + (B(t_n) + B(t_{n+1}))]. \end{cases} \tag{20}$$

Since this scheme is implicit, it allows us to consider larger time step values, and, as a consequence, it is linearly unconditionally stable.

4.3 Trigonometric schemes

Thanks to the Variation of Constants formula, the solution in (17) can be rewritten as

$$\begin{cases} U(t) = \cos(t\Omega)U_0 + t\ \mathrm{sinc}(t\Omega)V_0 + \displaystyle\int_0^t (t-s)\mathrm{sinc}((t-s)\Omega)B(s)ds, \\ V(t) = -\Omega \sin(t\Omega)U_0 + \cos(t\Omega)V_0 + \displaystyle\int_0^t \cos((t-s)\Omega)B(s)ds, \end{cases} \tag{21}$$

where Ω is the unique positive definite square root of Ω^2, see Remark 2, and $\mathrm{sinc}(x) = \frac{\sin x}{x}$.

Therefore, applying a discretization of the Variation of Constants formula to the system (21), we find the following explicit numerical procedure

$$\begin{cases} U_{n+1} = \cos(\tau\Omega)U_n + \tau\ \mathrm{sinc}(\tau\Omega)V_n + \displaystyle\int_0^\tau (\tau - s)\ \mathrm{sinc}((\tau-s)\Omega)B(t_n+s)ds, \\ \\ V_{n+1} = -\Omega \sin(\tau\Omega)U_n + \cos(\tau\Omega)V_n + \displaystyle\int_0^\tau \cos((\tau-s)\Omega)B(t_n+s)ds, \end{cases} \tag{22}$$

enriched by the initial conditions U_0 and V_0.

The scheme (22) requires the computation of the matrix functions $\cos(\tau\Omega)$ and $\mathrm{sinc}(\tau\Omega)$. The evaluation of $\cos(\tau\Omega)$ can be done by using a MATLAB routine, while, the computation of the $\mathrm{sinc}(\tau\Omega)$ matrix funtion is more delicate. A way to overcome this difficulty is to employ the series expression for $\mathrm{sinc}(\tau\Omega)$ but this could be too expensive or inaccurate [22]. So, one can try first to diagonalize the matrix function. Moreover, the computation of products of functions of matrices by vectors could be efficiently done by means of Krylov subspace methods, (see for instance [26,27,21]).

In [7], the authors show that the trigonometric method is unconditionally stable.

5 The one-dimensional non linear peridynamic model

We propose a numerical approach to study the one-dimensional non linear model (1) for an homogeneous bar of infinite length. This approach allows us to extend to the non linear case the numerical methods proposed in the previous sections.

For an **isotropic** material, the general form of the pairwise force function is given by

$$f(\xi,\eta) = c\ s(|\xi|,|\eta|)\ \frac{\eta}{|\eta|}, \tag{23}$$

where c is a positive constant depending on the material and the horizon. The function

$$s(|\xi|,|\eta|) = \frac{|\eta|-|\xi|}{|\xi|},$$

describes the relative change of the Euclidean distance of the particles. Since f is discontinuous in the first argument, the order of accuracy of the implemented numerical schemes will reduce.

In order to apply the results of the previous sections, we assume that $|\eta| << 1$ and $f(\xi,\eta)$ is sufficiently smooth. We consider the integral form of the function $f(\xi,\cdot)$:

$$f(\xi,\eta) = f(\xi,0) + \int_0^\eta \frac{\partial f(\xi,s)}{\partial \eta}(\eta - s)ds,$$

and then we apply an accurate quadrature formula

$$f(\xi,\eta) \approx f(\xi,0) + \sum_{r=1}^{m} w_r \frac{\partial f(\xi,s_r)}{\partial \eta}(\eta - s_r),$$

where w_r are the weights while s_r are the nodes of this formula. In general this approach leads to implicit methods, in fact, if we use the trapezoidal formula

$$f(\xi,\eta) \approx f(\xi,0) + \frac{\eta}{2}\left[\frac{\partial f(\xi,0)}{\partial \eta} + \frac{\partial f(\xi,\eta)}{\partial \eta}\right], \tag{24}$$

we derive a second order implicit method. Instead, if $f(\xi,\eta)$ is sufficiently smooth, we can derive an explicit scheme by using a Taylor expansion

$$f(\xi,\eta) \approx f(\xi,0) + C_1(\xi)\eta + \ldots + C_s(\xi)\eta^s, \tag{25}$$

where

$$C_i(\xi) = \frac{\partial^i f(\xi,0)}{\partial \eta^i}\ , \qquad i = 1,\ldots,s.$$

6 Simulations

In this section we present some numerical tests to confirm our results. We start with the linear model (4), assuming that the material body is not subject to

external forces, namely $b(x,t) = 0$. For simplicity, we consider the case of a constant density $\rho(x) = 1$. We take (10) as micromodulus function and we choose $u_0(x) = e^{-(x/L)^2}$ $x \in \mathbb{R}$ and $v = 0$ as initial conditions.

The choice of this micromodulus function is justified by the fact that its decay at infinity makes possible to consider a bounded computation domain. Moreover, in this setting the exact solution for (4) is given by

$$u^*(x,t) = \frac{2}{\sqrt{\pi}} \int_0^\infty \exp\left(-s^2\right) \cos\left(2sx\right) \cos\left(2t\sqrt{1-\exp\left(-s^2\right)}\right) ds, \qquad (26)$$

see for instance [35].

We denote by $\mathbf{u}^*(t) = (u^*(x_0,t), ..., u^*(x_N,t))^T$ the exact solution vector at time t and at the points of the spatial discretized domain.

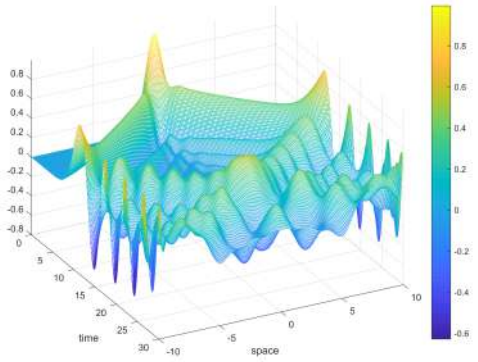


Fig. 1. With reference to Test 1: the numerical solution obtained by the MSV method. The parameters for the simulations are $\Delta x = \Delta t = 0.1$, $N = 200$, $N_T = 300$.

To order to perform an error study and to show the orders of accuracy of the decribed methods, we define $\mathbf{e}_k$ as

$$\mathbf{e}_k = \|\mathbf{u}(t_k) - \mathbf{u}^*(t_k)\|_\infty := \max\Big\{|u(x_i,t_k) - u^*(x_i,t_k)| : i = 0,\ldots,N,\Big\},$$

and then, for each method, we take the maximum error in the time interval $[0,T]$, namely

$$||\mathbf{e}||_\infty := \max\left\{\mathbf{e}_k : k = 1,\ldots,N_T\right\}.$$

We denote by MT, MSV, MMI and GT the methods consisting of the Midpoint+Trigonometric method, the Midpoint+Störmer-Verlet method, the Midpoint+Implicit Midpoint method and the Gauss two points+Trigonometric method, respectively.

6.1 Test 1: Comparison between MT, MSV, MMI and GT methods

In this section we study the performance of the MT, MSV, MMI and GT methods, by varying the time and space steps, by computing the error between the exact and the numerical solution and studying the rate of convergence.

Figure 1 shows the numerical solution computed by MSV method, while Table 1 summarizes the errors of the different methods by varying the spatial and time discretization steps. The term R_n denotes the ratio between the errors corresponding to Δx and $\Delta x/2$, thus, $\log_2(R_n)$ represents the convergence order of the methods. The last column of Table 1 confirms that the methods

Methods	$\Delta x = \Delta t$	N	N_T	$\|\|\mathbf{e}\|\|_\infty$	$\log_2(R_n)$
MSV	0.100	200	30	1.2911×10^{-3}	-
	0.050	400	60	3.2340×10^{-4}	1.9971
	0.025	800	120	8.0821×10^{-5}	2.0004
MT	0.100	200	30	5.9276×10^{-3}	-
	0.050	400	60	1.1126×10^{-3}	2.3959
	0.025	800	120	2.1350×10^{-4}	2.3992
MMI	0.100	200	30	2.5754×10^{-3}	-
	0.050	400	60	6.4621×10^{-4}	1.9946
	0.025	800	120	1.6106×10^{-4}	2.0043
GT	0.100	400	30	1.4940×10^{-4}	-
	0.050	800	60	9.3380×10^{-6}	3.9998
	0.025	1600	120	5.8300×10^{-7}	4.0015

Table 1. With reference to Test 1: the comparison among MSV, MT, MMI and GT methods by varying Δx, Δt, N and N_T.

MSV, MT, MMI are of the second order of accuracy while GT is of the fourth order.The method MSV is computationally less expensive than the others, but it has a bounded stability region, see Table 2.

6.2 Test 2: Comparison between MSV and MMI in the nonlinear case

Now we focus on the non linear case. In particular, we assume that f has the following form

$$f(\xi,\eta) = \begin{cases} c\frac{|\xi+\eta|-|\xi|}{|\xi|}\frac{\xi+\eta}{|\xi+\eta|}, & \text{if } 0 < |\xi| \le \delta, \\ 0, & \text{if } |\xi| > \delta, \end{cases} \qquad c > 0,$$

which has a singularity in $\xi = 0$. One can find an exact solution for this problem in [29].

Methods	Δx	Δt	N	N_T	$\|\|\mathbf{e}\|\|_\infty$
MSV	0.100	0.100	200	300	1.0543
	0.050	0.200	400	150	2.6300×10^{168}
	0.025	0.400	800	75	4.3600×10^{131}
MT	0.100	0.100	200	300	1.0941
	0.050	0.200	400	150	1.1081
	0.025	0.400	800	75	1.2987
MMI	0.100	0.100	200	300	1.0923
	0.050	0.200	400	150	1.0925
	0.025	0.400	800	75	8.2060×10^{-1}

Table 2. With reference to Test 1: the maximum error for the methods MSV, MT and MMI for different choices of Δx, Δt, N and N_T.

Table 3 depicts the maximum errors by varying the spatial and time discretization steps. We can see how all methods reduce their order of accuracy to 1. The reason of such reduction relies on the singularity of the pairwise force function f.

Methods	Δx	Δt	N	N_T	$\|\|\mathbf{e}\|\|_\infty$	$\log_2(R_n)$
MSV	0.1000	0.0100	10	1000	5.4590×10^{-2}	-
	0.0500	0.0050	20	2000	2.7285×10^{-2}	1.0007
	0.0250	0.0025	40	4000	1.3605×10^{-2}	1.0007
MMI	0.1000	0.0100	10	1000	5.3895×10^{-2}	-
	0.0500	0.0050	20	2000	2.7281×10^{-2}	0.9819
	0.0250	0.0025	40	4000	1.3603×10^{-2}	1.0036

Table 3. With reference to Test 2: the comparison among the performance of MSV and MMI methods in the nonlinear case by varying Δx, Δt, N and N_T.

7 Conclusions and perspectives

In this paper, we have reviewed numerical spatial discretization of higher order together with time integration techniques applied to a linear peridynamic model. Moreover, we have extended such techniques to the nonlinear model.

In future we would apply spectral techniques to both the linear and the nonlinear model following the results obtained in [7,25] and we will extend the

results to space domains of dimension greater than 1, using finite element or volume methods or mimetic finite difference methods, see [28,1,31,2].

Acknowledgements

This paper has been partially supported by PRIN2017. The author is member of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

References

1. L. Beirao Da Veiga, L. Lopez, and V. Vacca. Mimetic finite difference methods for Hamiltonian wave equations in 2D. *Computers and Mathematics with Applications*, 74(5):1123–1141, 2017.
2. B. Berardi, F. Difonzo, and L. Lopez. A mixed MoLTMoL for the numerical solution of the 2D Richards equation in layered soils. *Computers & Mathematics with Applications*, 79(7):1990 – 2001, 2020.
3. M. Berardi, F. Difonzo, M. Vurro, and L. Lopez. The 1D Richards equation in two layered soils: a Filippov approach to treat discontinuities. *Advances in Water Resources*, 115:264 – 272, 2018.
4. H. R. Beyer, B. Aksoylu, and F. Celiker. On a class of nonlocal wave equations from applications. *Journal of Mathematical Physics*, 57:062902, 2016.
5. F. Bobaru, M. Yang, S. Alves, F.and Silling, E. Askari, and J. Xu. Convergence, adaptive refinement, and slaning in 1D peridynamics. *Int. J. Numer. Mech. Eng.*, 77:852 – 877, 2009.
6. G. M. Coclite, S. Dipierro, F. Maddalena, and E. Valdinoci. Wellposedness of a nonlinear peridynamic model. *Nonlinearity (to appear)*.
7. G. M. Coclite, A. Fanizzi, L. Lopez, F. Maddalena, and S. F. Pellegrino. Numerical methods for the nonlocal wave equation of the peridynamics. *Applied Numerical Mathematics*, 155:119 – 139, 2020.
8. G. M. Coclite, F. Paparella, and S. F. Pellegrino. On a salt fingers model. *Nonlinear Analysis*, 176:100 – 116, 2018.
9. E. Dal Santo, C. Donadello, S. F. Pellegrino, and M. D. Rosini. Representation of capacity drop at a road merge via point constraints in a first order traffic model. *ESAIM: M2AN*, 53(1):1–34, 2019.
10. E. Emmrich and D. Puhst. Well-posedness of the peridynamic model with Lipschitz continuous pairwise force function. *Commun. Math. Sci.*, 11(4):1039–1049, 2013.
11. E. Emmrich and D. Puhst. Survey of existence results in nonlinear peridynamics in comparison with local elastodynamics. *Comput. Methods Appl. Math.*, 15(4):483–496, 2015.
12. E. Emmrich and O. Weckner. Numerical simulation of the dynamics of a nonlocal, inhomogeneous, infinite bar. *Journal of Computational and Applied Mechanics*, 6(2):311–319, 2005.
13. E. Emmrich and O. Weckner. The peridynamic equations and its spatial discretization. *Mathematical Modelling And Analysis*, 12(1):17–27, 2007.
14. A. C. Eringen and D. G. B. Edelen. On nonlocal elasticity. *Int. J. Eng. Sci.*, 10(3):233–248, 1972.
15. R. Garrappa. Exponential integrators for time-fractional partial differential equations. *Eur. Phys. J. Special Topics*, 222:1915–1927, 2013.

16. R. Garrappa. On some generalizations of the implicit Euler method for discontinuous fractional differential equations. *Mathematics and Computers in Simulation*, 95:213–228, 2014.
17. R. Garrappa. Numerical evaluation of two and three parameter Mittag-Leffler Functions. *SIAM Journal on Numerical Analysis*, 53(3):1350–1369, 2015.
18. R. Garrappa. Trapezoidal methods for fractional differential equations: theoretical and computational aspects. *Mathematics and Computers in Simulation*, 110:96–112, 2015.
19. R. Garrappa, Moret. I., and M. Popolizio. Solving the time-fractional Schrödinger equation by Krylov projection methods. *Journal of Computational Physics*, 293:115 – 134, 2015.
20. R. Garrappa and M. Popolizio. Computing the Matrix Mittag-Leffler Function with applications to Fractional Calculus. *Journal of Scientific Computing*, 77:129–153, 2018.
21. V. Grimm and M. Hochbruck. Rational approximation to trigonometric operators. *BIT*, 48(2):215–229, 2008.
22. Nicholas J. Higham and Matthew I. Smith. Computing the matrix cosine. *Numer. Algorithms*, 34(1):13–26, 2003.
23. E. Kröner. Elasticity theory of materials with long range cohesive forces. *Int. J. Solids Structures*, 3:731–742, 1967.
24. R. Lipton. Dynamic brittle fracture as a small horizon limit of peridynamics. *J. Elasticity*, 117:21–50, 2014.
25. L. Lopez and S. F. Pellegrino. A spectral method with volume penlization for a nonlinear peridynamic model. *Submitted*.
26. L. Lopez and V. Simoncini. Analysis of projection methods for rational function approximation to the matrix exponential. *SIAM Journal on Numerical Analysis*, 44(2):613–635, 2006.
27. L. Lopez and V. Simoncini. Preserving geometric properties of the exponential matrix by block Krylov subspace methods. *BIT Numerical Mathematics*, 46(4):813–830, Dec 2006.
28. L. Lopez and V. Vacca. Spectral properties and conservation laws in mimetic finite difference methods for PDEs. *Journal of Computational and Applied Mathematics*, 292(15):760–784, 2016.
29. E. Madenci and E. Oterkus. *Peridynamic theory and its applications*. Springer New York, 2013.
30. K.W. Morton and D.F. Mayers. *Numerical Solution of Partial Differential Equations*. Cambridge University Press, Cambridge, 1994.
31. S. F. Pellegrino. On the implementation of a finite volumes scheme with monotone transmission conditions for scalar conservation laws on a star-shaped network. *Applied Numerical Mathematics*, 155:181 – 191, 2020.
32. D. Qiang, T. Yunzhe, and T. Xiaochuan. A peridynamic model of fracture mechanics with bond-breaking. *J. Elasticity*, 2017.
33. S. Silling and E. Askari. A meshfree based on the peridynamic model of solid mechanics. *Computer & Structures*, 83(17–18):1526–1535, 2005.
34. S.A. Silling. Reformulation of elasticity theory for discontinuities and long-range forces. *J. Mech. Phys. Solids*, 48(17–18):175–209, 2000.
35. O. Weckner and R. Abeyaratne. The effect of long-range forces on the dynamics of a bar. *Journal of the Mechanics and Physics of Solids*, 53(3):705 – 728, 2005.
36. M. Zaccariotto, T. Mudric, D. Tomasi, A. Shojaei, and U. Galvanetto. Coupling of FEM meshes with Peridynamic grids. *Computer Methods in Applied Mechanics and Engineering*, 330:471 – 497, 2018.

Mechanism of Formation for Fluctuation Phenomena

E. Prozorova[i]
Mathematical-Mechanical Department St. Petersburg State
University Av. 28, Peterhof, 198504, Russia
(E-mail: e.prozorova@spbu.ru)

Abstract. The paper proposes a deterministic mechanism for the formation of Brownian motion. The movement of molecules, even under equilibrium conditions, leads to a constant displacement of the center of inertia, which creates a moment and leads to the emergence of additional force. The result is a possible accumulation of molecules in separate areas. Having arisen as a result of collisions of slower molecules than average ones, they are forced to move together for some time, m6ellely no changing their position. The action of the moment manifests itself in all known processes. In the kinetic theory based on the Lagrange and Liouville equations, the motion of the axis of inertia in the process of rearrangement and motion of particles is neglected. The concept of a derivative in terms of finite values of such quantities as the mean free path, time between collisions, etc. has features. At small mean free paths, the motion of the center of inertia contributes to the equation of state. A new algorithm is proposed for calculating the force entering the Langevin equation and the equation of state for a liquid. The listed issues are
Keywords: Conference, CHAOS, Chaotic Modeling, CMSIM Style

1. Introduction.

Fluctuations are called random deviations of physical systems from their equilibrium state (or physical processes - from their steady flow). Fluctuations exist both in no equilibrium states and in unsteady processes; in their absence, relaxation would be a "smooth" process and they could be described by single-valued functions of time. The presence of thermal fluctuations causes random deviations of real processes from such a "smooth" flow [1-7]. Diffusion and Brownian motion occur due to the chaotic thermal motion of molecules, and as a result are described by similar mathematical rules.

The difference between them is that during diffusion, a molecule always moves in a straight line until it collides with another molecule, after which it changes its trajectory. A Brownian particle does not “fly free”, but undergoes very small and frequent, as it were, “tremors”, as a result of which it randomly moves here and there. Fluctuation effect was first explained by A. Einstein and then by M. Smolukhovsky. The theoretical studies of M. Smoluchowski, which differed from the works of A. Einstein only by a slightly less rigor, but greater clarity. It consisted in the fact that the diffusion force should be equal to the viscous value Stokes drag force.

$$\frac{\mathrm{n_p}\,\mathrm{F_s}}{\mathrm{c_s}} = \mathrm{D_E}\,\frac{\partial \mathrm{n_p}}{\partial \mathrm{x}}$$

Here F_s force, n_p - particle concentration, c_s - mass concentration, $D_e = \frac{kT}{(6\pi\eta_2 R_0)}$, R_0 −

- particle radius, η_2—viscosity. Later the theory was developed on the basis of the Langevin and Fokker-Planck equations. The evolution of a Brownian particle (fluctuation) is determined by its interaction with the environment, which is always collective. In the kinetic representation, the evolution of a system of Brownian particles is described by a nonlocal equation for the n-particle distribution function. Now the Langevin and Fokker-Planck equations are obtained from the Liouville equation for specially selected models of integral kernels using of phenomenological conservation laws [6,7]. The Langevin and Fokker – Planck equations refer to the stochastic approximation of particle motion. Since the equation contains a resistance force arising during the translational individual movement of particles, then the speed of movement of the particle in time τ will change. The visible part of the particle motion will depend on the characteristics of the instruments. In addition, collective effects will play a major role. Like the Liouville equation, the Langevin equation does not take into account particle collisions. The possibilities of calculating the motion of particles by these equations are limited.

In the general case, the connection between macroscopically observable quantities and fluctuations of the corresponding dynamical variables is established by solving the dynamic Liouville equation. However, the equation does not take into account possible dissipative processes associated with particle collisions. At the macrolevel, equations are more consistently obtained for the distribution function (the Fokker-Planck and Boltzmann equations).

In the kinetic theory based on the Lagrange and Liouville equations, the motion of the axis of inertia in the process of rearrangement and motion of particles is neglected. This means that the contribution of the angular momentum (force), which leads to collective processes, is not taken into account [8-13} The influence of the moment is confirmed by the work [14], which is devoted to the calculation of the conductivity of a nonideal fully ionized plasma under the assumption that there are no straight sections of the electron trajectory. The experimental data turned out to be possiblesatisfy for the velocity correlator only with the involvement of the angular momentum.

The exact molecular theory, which gives results that are in satisfactory agreement with experiment, can be applied only in the special case when the potential corresponding to the force arising from the interaction of neighboring molecules depends only on the distance between them. But this assumption is valid only for liquids consisting of monoatomic molecules.

In recent years, the molecular modeling method has been widely used [7,15,16]. The main version of modern theory is random disturbances. The theory of Brownian motion is approximate. And although in most practically important noted that cases the existing theory gives satisfactory results, in some cases it may require refinement. It should be the classical Boltzmann equation does not comply with the law of conservation of angular momentum as and another equations of kinetic theory and statistic mechanics. This is clearly seen if we multiply equation of speed on radius-vector of the particle to get angular momentum. Even with central interaction we get different values for the non-equilibrium conditions. In numerical calculations by the difference scheme using the grid pitch that is smaller the mean free path and with the ideology of a closed volume not obtain the influence of the angular momentum due to the absence of collisions. A mechanism for the occurrence of fluctuations has not been proposed. The work is limited to the case of "simple" media, by which we mean gases and liquids, consisting of point molecules that do not contain internal degrees of freedom. Now we suggest the reason of fluctuation effect.

2. Lagrangian function for the collective interaction

In classical mechanics, kinetic theory and statistical mechanics, the role of the angular momentum and, therefore, the moment of force is underestimated. Let us consider three interacted among themselves (Fig.1) particles

$$\boldsymbol{r_c} = \frac{\boldsymbol{m_1 r_1 + m_2 r_2 + m_3 r_3}}{\boldsymbol{m_1 + m_2 + m_3}}$$

$$r_{c+\Delta c} = \frac{m_1(r_1 + \dot{r_1}\Delta t) + m_2(r_2 + \dot{r_2}\Delta t) + m_3(r_3 + \dot{r_3}\Delta t)}{m_1 + m_2 + m_3}$$

At the next moment in time, the position will change under the action of the force arising in connection with the new position of the center of inertia. Thus, the new position of the molecules will create a new force. The same result can be obtained by counting the angular momentum, from which to determine the effective force and speed of the center of inertia.

At equilibrium, or at small strains, but under no equilibrium thermodynamic effects and perturbations lead to an uneven distribution of the physical parameters and the role of collective effects, that determined by the growing influence of the angular momentum. In addition, when these strains change position of the center of mass of elementary volume, that is sign for changing Lagrangian function

$$\frac{\mathrm{dL}}{\mathrm{dt}} = \sum_i \left[\frac{\partial L}{\partial q_i}\dot{q}_i + \frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i\right] + \sum_i \left[\frac{\partial L}{\partial (q_i - a)}(\dot{q}_i - \dot{a}) + \frac{\partial L}{\partial (\dot{q}_i - \dot{a})}(\ddot{q}_i - \ddot{a})\right],$$

$\boldsymbol{a} = \sum_i \frac{m_i \boldsymbol{r}_i}{m_i}$, for electrical interaction $\boldsymbol{a} = \sum_i \frac{e_i \boldsymbol{r}_i}{e_i}$. q_i – generalized coordinate, $\dot{q}_\iota$ –generalized speed. The classical Liouville equation $\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\iota} - \frac{\partial L}{\partial q_i} = 0 \quad (i = 1,2 \ldots)$, $L = L(q_i, \dot{q}_\iota, t)$ – Lagrange function. The whole theory is developed for a force of the form $\mathbf{F} = -\frac{\partial \mathbf{U}}{\partial \mathbf{R}}, \boldsymbol{U} = \boldsymbol{U}(\boldsymbol{R})$. $\boldsymbol{R}$ - radius vector

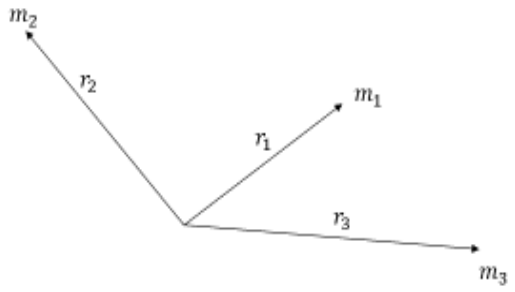


Fig. 1 Interaction of three parts

In view of the time we are invited to consider force formula $\boldsymbol{F} = \boldsymbol{F_0} + \nabla\left((\boldsymbol{R} - \boldsymbol{a}) \times \frac{\partial \boldsymbol{U}}{\partial \boldsymbol{R}}\right), \boldsymbol{R}$ – the current radius. This formula is transformed with the permutability derivatives and directions of forces in the formula

$$\boldsymbol{F} = \boldsymbol{F_0} + \nabla\left((\boldsymbol{R} - \boldsymbol{a}) \cdot \frac{\partial \mathrm{U}}{\partial \boldsymbol{R}}\right).$$

For the N partial distribution function, it was [7]

$\frac{\partial F_N}{\partial t} + \sum_{i=1}^{N} \dot{\xi}_i \cdot \frac{\partial F_N}{\partial x_i} + \frac{1}{m} \sum_{i=1}^{N} \frac{\partial}{\partial \xi_i} \cdot (X_i F_N) = 0.$ Now

$\frac{\partial F_N}{\partial t} + \frac{\partial \boldsymbol{L}}{\partial \boldsymbol{a}} \dot{\boldsymbol{a}} + \sum_{i=1}^{N} (\dot{\boldsymbol{\xi}}_{\boldsymbol{i}} - \dot{\boldsymbol{a}}) \cdot \frac{\partial F_N}{\partial (x_i - \boldsymbol{a})} + \frac{1}{m} \sum_{i=1}^{N} \frac{\partial}{\partial (\xi_i - \dot{\boldsymbol{a}})} \cdot (X_i F_N) = 0.$

Usually, however, such as the Hamiltonian system of two interacting molecules after separation of the center of mass is represented as the sum of the Hamiltonians of isolated molecules $H_0 = H_A + H_B$ operator and their electrostatic interaction [17]

$$H = H_0 + H_B$$

$$H_B = -\sum_{a=1}^{n_A}\sum_{j=1}^{N_B}\frac{Z_a}{r_{aj}} - \sum_{b=1}^{n_B}\sum_{j=1}^{N_B}\frac{Z_b}{r_{bj}} + \sum_{i=1}^{N_A}\sum_{j=1}^{N_B}\frac{1}{r_{ij}} + \sum_{a=1}^{n_A}\sum_{b=1}^{n_B}\frac{Z_a Z_b}{R_{ab}},$$

where the indices A, B numbered core indices i, j - the electrons of molecules A, B, respectively, the atomic units. In the rarefied gas unusual situation arises when for describing the derivatives we use the limit of the ratio of the increment function to the increment argument. It turns out that for recording the time derivative of the final terms. We have mean free path (rarefied gas) taking into account only the high-speed components, as slow collisions do not have time to occur.

It is interesting to compare the derivatives for discrete and continuous descriptions. First, consider the relaxation process (in time).

Here and then $\boldsymbol{\xi_i}$ is the velocity of the molecule, u is the velocity of the elementary volume $\boldsymbol{p_i} = \boldsymbol{\xi_i} - \boldsymbol{u}$, is the intrinsic velocity of the molecule, r is the coordinate. The velocity $\boldsymbol{\xi_i}$ of molecules is included in the definition of the Boltzmann equation and in the calculations for the models used as an independent variable.

$f = f\,(t, \boldsymbol{r}(t), \boldsymbol{\xi}(t))$. We represent the distribution function as

$f = \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r})}$, that is

$f = \frac{n}{N}$, where n is the number of molecules in an elementary volume, N is the number of molecules in a perturbed volume. We consider $n \le N$. Then

$\frac{\partial f}{\partial t}\Big|_{\boldsymbol{r}=const} = \frac{\partial}{\partial t} \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r})}$.

Let us consider the dependence $\delta(\boldsymbol{r_i} - \boldsymbol{r})$- on t only as $(\boldsymbol{r_i} - \boldsymbol{r})$.

Consideration gets more complicated when there are cross-border flows

If there are no streams

$$\frac{\boldsymbol{F_1}}{\boldsymbol{F_3}} - \frac{\boldsymbol{F_2}}{\boldsymbol{F_4}} = \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r}) + \sum_{i}^{n} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + \cdots}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r}) + \sum_{i}^{N} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + \cdots} - \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r})} \approx$$

$$\approx \left(\frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r}) + \sum_{i}^{n} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + \cdots}{\sum_{i=1}^{N} \delta(r_i - r)} \left(1 - \frac{\sum_{i}^{N} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + \cdots}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r})}\right)\right.$$

$$\left. - \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\sum_{i=1}^{N} \delta(r_i - r)}\right) \approx \frac{\sum_{i}^{n} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + O\left((\Delta t)^2\right.}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r})}.$$

Thus, when solving the Boltzmann equation, we obtain the dependence of the time derivative only through the derivatives of the macroparameters. This hypothesis is used in the theory of rarefied gas when constructing a solution to the Boltzmann equation by the Chapman-Enskiy method.

2.Taking into account flows across the border. We have

$$\frac{\boldsymbol{F_1}}{\boldsymbol{F_3}} - \frac{\boldsymbol{F_2}}{\boldsymbol{F_4}} = \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r}) + \sum_{i}^{n} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + \sum_{j} \frac{\boldsymbol{p_j}}{\boldsymbol{m}} \delta(\boldsymbol{r_j} - \boldsymbol{r}) + \sum_{j} \frac{\boldsymbol{p_j}}{\boldsymbol{m}} \Delta t \frac{\partial \delta(\boldsymbol{r_j} - \boldsymbol{r})}{\partial t} + \cdots}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r}) + \sum_{i}^{N} \Delta t \frac{\partial \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\partial t} + \sum_{j} \frac{\boldsymbol{p_j}}{\boldsymbol{m}} \delta(\boldsymbol{r_j} - \boldsymbol{r}) + \sum_{j} \frac{\boldsymbol{p_j}}{\boldsymbol{m}} \Delta t \frac{\partial \delta(\boldsymbol{r_j} - \boldsymbol{r})}{\partial t} + \cdots} - \frac{\sum_{i=1}^{n} \delta(\boldsymbol{r_i} - \boldsymbol{r})}{\sum_{i=1}^{N} \delta(\boldsymbol{r_i} - \boldsymbol{r})}.$$

1/Δt, the time derivative begins to prevail. Thus, the role of borders is increasing. The distribution function does not provide correct parameter values. Here, only the molecular dynamics method with a very small time step is correct.

3. Calculate the derivatives with respect to space in a rarefied gas

$$(F(t\ +\Delta t)-\boldsymbol{F}(t))/\Delta t \approx$$
$$\approx \frac{-div\left(\sum_{j=1}^{n_2}\frac{p_j}{m}\ \delta(r_i-r)\ +\sum_{i=1}^{n}\frac{p_j}{m}\ \delta(r_i-r)\ +\cdots\right)}{(\sum_{i=1}^{N}\delta(r_i-r)+\sum_{j=1}^{N_1}\frac{p_j}{m}\ \Delta_k\delta\left(r_j(t)-r\right)+\cdots)}.$$

$$\sum_{i=N}^{n+\Delta n}\delta\left(r_j-r\right))\ =\ -div\sum_{j=1}^{n_2}\frac{p_j}{m}\ \delta(r_j-r).$$

3.Effect of Angular Momentum for a Discrete Environment.

Knowledge of virial coefficients is necessary in various practical problems [25]. In previous works, the effect of angular momentum on physical parameters in the kinetic theory and continuum mechanics was discussed. It can be assumed that density fluctuations are associated with inhomogeneity of the velocity distribution at temperatures greater than zero Kelvin degrees, which in turn is associated with the movement of the inertia axis of elementary volumes. Here we consider the algorithm for calculating the additional force associated with the action of the angular momentumin a discrete medium. The formula for determining the center of gravity of the system of material points.

$$r_c=\ \frac{\sum_{i=1}^{k}m_i r_i}{\sum_{i=1}^{k}m_i}\ .$$

Center of inertia at different times for identical molecules in an elementary volume

$$\frac{\sum_{i=1}^{n}m\,\delta(r_i-r_c)\,r_i}{Nm}\ =\ r_c,\quad \frac{\sum_{i=1}^{n}m\,\delta(r_i+\Delta r_i-r_c+\Delta r_c)\,(r_i+\Delta r_i)}{(N+\Delta N)\,m}\ \approx\ \widetilde{r_c}\ .$$

Angular momentum is

$$\widetilde{p_i}\times(r_i+\Delta r_i-\ \widetilde{r_c}\ )-\ p_i\times\ r_i=L_i\ ,$$
$$\frac{(p_i+\Delta p_i)\times(r_i+\Delta r_i-\ r_c-\ \Delta r_c\ )-\ p_i\times\ (r_i-r_c)}{\Delta t}\ =\frac{dL_i}{dt}.$$
$$\frac{dL}{dt}=M,\quad F=\ \frac{dM}{dr}\ .$$
$$\frac{dM_i}{dr}=\frac{d}{dr}\frac{p_i(\Delta r_i-\ \Delta r_c)}{\Delta t}=\ \frac{d}{dr}\ p_i(\Delta p_i-\Delta p_c\ ).$$

Then we get an additional force

$$F_i^{ad}\ \approx\ \frac{\sum_{i=1}^{n}((\Delta r_i-\Delta r_c)\frac{\partial\delta}{\partial r}r_i\ +\Delta r_i\,\delta(r_i-r_c)r_i)}{N},\qquad \Delta p_i^k=\ F_i^{ad}\Delta t.$$

$p_i^{k2}\Delta t^2+p_j^{k2}\Delta t^2\le(r+\sigma)_i^{k2}-(r+\sigma)_j^{\ k2}$. The number of molecules that form dimers.

$p_i^{k2}\Delta t^2+p_j^{k2}\Delta t^2\ \ge E$. The number of dimers formed.

Here p_j^k is the contribution of the component k. Here we determine the number of dimers that form equilibrium conditions. Pressure

$$\sum_{i=1}^{n}\frac{(p_i+\Delta p_i)^2}{m}+\ \sum_{N}^{n+\Delta n}\frac{p_i^{\ 2}}{m}\ =\ \text{p.}$$

The first term is responsible for the usual thermodynamic pressure, taking into account the change in speed from the action of the moment, the second term is responsible for the pressure caused by dimers. From this it can be seen that the pressure change is non-monotonic. Langevin equation taking into account the influence of the angular momentum is

$\frac{dV}{dt}\ =\ -\frac{\zeta_{\hat{v}}}{m}\mathrm{V}+\frac{1}{m}\frac{dM}{dr}$, where $\zeta_{\hat{v}}$ - coefficient of friction of the selected particle, m is the mass of the particle, M is the moment of force acting on the particle.

The classical Langevin equation for one particle

$$\frac{dV}{dt}\ =\ -\frac{\zeta_{\hat{v}}}{m}\mathrm{V}+\frac{1}{m}F(t)\ ,\quad \text{where } F\ (t) \text{ is a random force.}$$

A Markov Gaussian process is considered with the condition that the average for an ensemble of particles ⟨F = 0. In our version, this condition is fulfilled by virtue of the fulfillment of the theorem on the conservation of the moment in a closed volume. For equallibrium condition, this is true.

In conclusion, we note that in the construction of statistical theories of an equilibrium liquid in the Clausius theory, for pressure,

$$p - nkT = \; - \; \frac{1}{6}\,\rho^2 \int v(\boldsymbol{r})\; g(\boldsymbol{r}\,) d\boldsymbol{r}\,,$$

where $v(r)$ is the power of the intermolecular interaction force $(\,d\varphi(r))dln(r)\;,\;\varphi(r)\,-$potential, $g(r\,) = \;\; d\varphi\,(r)\,/\,(dln\,(r))$, and $g(r\,)$ is the pair distribution function; i.e. the value of the same structure (dimensional) as the term in the equations with allowance for the angular momentum, that is $M = (r - \; r_0) \times \frac{d\varphi(\boldsymbol{r})}{dr}, r_0.$ center of mass position. So

$$p - nkT = \; - \; \frac{1}{6}\,\rho^2 \int (v(\boldsymbol{r}) + r\frac{dM}{dr})\; g(\boldsymbol{r}\,) d\boldsymbol{r}\,.$$

The theoretical calculation of the binary correlative function for liquids by the methods of statistical mechanics is associated with great difficulties that have not yet been overcome. Calculation based on experimental studies of X-ray scattering is also possible only for a small circle of simple liquids, the molecular structure of which has been well studied. Nevertheless, using the existing relations of the statistical theory of fine-structure fluctuations and the results of X-ray studies of simple liquids, one can get an idea of a number of features that distinguish fine-structure concentration fluctuations from thermodynamic ones.

Additional force should contribute at high temperatures. For inert gases at medium and low temperatures, the contribution is negligible. For rarefied gas the angular momentum is value of first order. For the water molecules we have the potential for interaction of dipole and so it have some maximum and minimum, but angular momentum is main correlation effect for point molecules and another interaction is values of smaller. Then it seems probable that for pressure is to be sufficient if liquefied gases are simple, for example Ar, of two virial coefficients and binary interaction of particles. The effect of a small distortion of circular orbits will be small due to the incommensurability the time of the rotation and displacement times of the molecules. For water, additional components will arise in connection with the asymmetry of the molecules and the perturbation of the basic potential by the interaction of the hydrogen parts with each other and with the nucleus of the second molecule.

4.Boltzmann kinetic equation

For nonequilibrium states, the Boltzmann kinetic equation is widely used. The equation is derived in two ways. One - based on conservation laws, for the second, the starting point is a chain of coupled equations proposed by Bogolyubov [18]. As a first approximation, the Boltzmann equation is obtained. In this case considering fluctuations, it is theoretically possible to take into account both collisions between particles and collisions between particles and gas molecules, considering a mixture of two components: particles and gas. Solving the equation even for a gas presents significant difficulties. In addition, the equilibrium distributin function does not satisfy the equation [19] and it does not preserve the angular momentum [9]. The latter is verified by vector multiplication of the moment for the velocity equation by the radius vector.

The classical derivation of the Boltzmann equation consists in recording the balance of particles through the relation for the single-particle distribution function

$$f\big(t + dt, \boldsymbol{r} + \boldsymbol{\xi_j}\,dt, \boldsymbol{\xi}_j + \; \boldsymbol{F_j} dt\big)\mathrm{d}\mathrm{r}\mathrm{d}\xi_\mathrm{j} = f\big(\boldsymbol{r}, \boldsymbol{\xi_j}, \mathrm{t}\,\big)d\boldsymbol{r}\; d\boldsymbol{\xi_j} + (\frac{\partial f}{\partial t})_{coll}\; dt.$$

Often the latter is written in the form

$$f\big(t + dt, \boldsymbol{r} + \boldsymbol{\xi_j}\,dt, \boldsymbol{\xi_j} + \; \boldsymbol{F_j} dt\big) dr d\xi_j = \mathrm{f}\big(\boldsymbol{r}, \boldsymbol{\xi_j}, \mathrm{t}\,\big) + (\widetilde{\frac{\partial \mathrm{f}}{\partial \mathrm{t}}})_{\mathrm{coll}}\; \mathrm{dt}.$$

That is r, the radius vector; x- coordinate of the point; ξ- the velocity of the point, m - the molecular weight, and, according to the definition of the distribution function f_N, the probability of finding the system at the points (x, ξ) in the intervals $dx_i\; d\,\xi_i$ is

$$f_N\;(t, x_1\,, \ldots, x_N\,, \xi_1\,,, \xi_N)\; dx_1 \ldots dx_N\;\; d\xi_1 \ldots d\xi_N.$$

The new Boltzmann equation can be written as follows with momentum

$$\frac{df}{dt} = \;\; \frac{\partial f}{\partial t} + \boldsymbol{\xi}_{\,i} \cdot \Big[\frac{\partial f}{\partial \boldsymbol{r_i}}\Big] + \boldsymbol{\xi_i} \cdot \frac{\partial}{\partial \boldsymbol{r_i}}[\,\boldsymbol{r_j}\frac{\partial f}{\partial \boldsymbol{r_j}}\,] - \frac{F}{m}\frac{\partial f}{\partial \boldsymbol{\xi_i}} = I.$$

Where $(\frac{\partial f}{\partial t})_{coll}, (\widetilde{\frac{\partial f}{\partial t}})_{coll}$ – are the collision integrals that record in different phase spaces. Externally, these equalities are identical, but the second relation is satisfied on the interaction times of the molecules and all interactions are correlated. For gas dynamic problems, the characteristic length of an elementary volume for which equality is written is equal to10^{-8} cm, and the requirement of a large number of particles in an elementary volume is not

satisfied for altitudes of 120-300 km in the terrestrial atmosphere. Indeed, the required minimum size is 10^{-3}cm. Since,$N = \pi R^2 \cdot \xi \cdot \tau \cdot n$ where $R$is the radius of the cylinder of elementary volume; τ is the mean time of free movement, then for statistical independence the number$N$ of particles must be at least 100.cm. Then $\pi R^2 \cdot 10^4 \cdot 10^{12} \cdot 10^{-5} = 10^2$, that is $R = 10^{-3}$ cm.

Functionally, the Boltzmann equation is invariant with respect to the choice of macro parameters of the distribution function. It is necessary to compare the equilibrium distribution function with macroparameters taken from the Euler and Navier-Stokes equations. The difference will give us a small increment functions. We find that for the Euler equations (zero approximation of the Chapman-Enskog) the difference is zero. There are differences to the first approximation. The first approximation is responsible for the tangential component (p_{ij} tensor of viscous stresses). Euler equations are obtained with the use of locally-equilibrium distribution function. Consequently, they are responsible for the normal component of the velocity values regardless of macroparameters. Upon receipt of the first order correction of the terms included in the final decision of the Chapman-Enskog leave only after integration over the phase velocity ξ. The integrals are taken from f function, i.e. for (ρu). Consider

$$\frac{Df_0}{dt} = \frac{1}{n} f_0 \frac{\partial n}{\partial t} + \frac{3}{2}\frac{1}{T} f_0 \frac{\partial T}{\partial t} + \frac{mc^2}{2kT^2} f_0 \frac{\partial T}{\partial t} + f_0 \left(\frac{m}{kT}(\boldsymbol{\xi} - \boldsymbol{u})\frac{\partial \boldsymbol{u}}{\partial t}\right) + \boldsymbol{\xi} \cdot \left\{\frac{1}{n} f_0 \frac{\partial n}{\partial x} + \left(-\frac{3}{2}\right)\frac{1}{T} f_0 \frac{\partial T}{\partial x} + \frac{mc^2}{2kT^2} f_0 \frac{\partial T}{\partial x} + f_0 \left(\frac{m}{kT}(\boldsymbol{\xi} - \boldsymbol{u})\frac{\partial \boldsymbol{u}}{\partial x}\right)\right\} =$$

$$= 2J(f_0, f_0\varphi^k) = \int f_0 f_1^0 \left(\varphi_1^{(k)'} + \varphi^{(k)'} - \varphi_1^{(k)} - \varphi^{(k)}\right) g\, b\, dbd\, \epsilon d\xi_1 \xi = 0.$$

In classic case

$$\frac{\partial f_0}{\partial t}\Big|_{t=0} = f_0 \left\{\frac{m}{kT}\left(c_i c_j - \frac{1}{3}c^2 \delta_{ij}\right)\frac{\partial u_i}{\partial t} + \frac{1}{2T}\frac{\partial T}{\partial t} c_i \left[\left(\frac{m}{kT}\right)c^2 - 5\right]\right\}.$$

The Boltzmann equation was wrote for full function and have the local equilibrium function and addition item. The tangent velocity component is obtained because off ξ have arbitrary direction of velocity relative position of coordinate axes

$$\int \boldsymbol{n} \cdot (\boldsymbol{\tau} \cdot f\boldsymbol{\xi})\boldsymbol{ds}\, \boldsymbol{d\xi} = \int div\, (\boldsymbol{\tau} \cdot f\boldsymbol{\xi})\boldsymbol{dx}\, \boldsymbol{d\xi}$$

τf give us addition item. Besides local equilibrium function f_0 we have addition item $\frac{\boldsymbol{p_{ij}}}{\boldsymbol{2p}}\left(\frac{\boldsymbol{m}}{\boldsymbol{2T}}\right)\boldsymbol{c_i}\,\boldsymbol{c_j} - \frac{\boldsymbol{q_i}}{\boldsymbol{p}}\left(\frac{\boldsymbol{m}}{\boldsymbol{kT}}\right)\left(\boldsymbol{1} - \frac{\boldsymbol{c^2}}{\boldsymbol{5}}\frac{\boldsymbol{m}}{\boldsymbol{kT}}\right)\boldsymbol{c_i}]$. Main account gives derivatives of local equilibrium function. These items definite the self-diffusion and thermo-diffusion which were foretold by S. V. Vallander [20]. The second derivative is result item $c_i \cdot \frac{\partial f}{\partial r_i}$.

Examples from kinetic theory a) the problem of kinetic theory. Gas in stationary force field with potential φ (analogue of the problem [21]):

$$\xi_i \frac{\partial f}{\partial x_i} + \xi_i \frac{\partial}{\partial x_i} x_i \frac{\partial f}{\partial x_i} - \frac{1}{m}\frac{\partial \varphi}{\partial x_i}\frac{\partial f}{\partial \xi_i} = J(f, f).$$

ξ_i the phase velocity of the coordinates $x, y. z$; f - distribution function, $J(f, f) -$ the collision integral. Classic distribution is $f = A(x)e^{-B(x)\xi^2}$. In this case, we have the old results, B = Const. For A (x) we have the equation

$$\frac{dA}{dx_i} + \frac{d}{dx_i} x_i \frac{dA}{dx_i} + 2\frac{A \cdot B}{m}\frac{\partial \varphi}{\partial x_i} = 0.$$

Then we have $f = n_0 \left(\frac{m}{2\pi kT}\right)^{3/2} e^{-\frac{\varphi}{kT}} e^{-\frac{m}{2kT}\xi^2}$.

General Maxwell distribution has the form. $f = n\left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \exp\left\{-\frac{m}{2kT}c^2\right\}$, $c = \xi - u$.

The modified Boltzmann equation

$$\xi_i \frac{\partial f}{\partial x_i} + \xi_i \frac{\partial}{\partial x_i} x_i \frac{\partial f}{\partial x_i} - g_i \frac{\partial f}{\partial \xi_i} = J(f, f).$$

$$g = \frac{X}{m} - \text{acceleration of molecules}$$

Apply to the solution of the old algorithm $\ln f = \gamma_0 + \gamma_i \xi_i + \gamma_4 \xi^2$. .

Then we get the equation of the old and the new equation.

$$\frac{\partial \gamma_0}{\partial t} + g_i \gamma_i = 0,$$

$$\frac{\partial \gamma_i}{\partial t} + 2g_i\gamma_4 + \frac{\partial \gamma_0}{\partial x_i} + \frac{\partial \gamma_0}{\partial x_i} + \frac{1}{2} x_i \frac{\partial \gamma_0{}^2}{\partial x_i^2} + \frac{\partial}{\partial x_i} x_i \frac{\partial \gamma_0}{\partial x_i} = 0,$$

$$\frac{\partial \gamma_4}{\partial t}\delta_{ij} + \frac{1}{2}\left(\frac{\partial \gamma_i}{\partial x_j} + \frac{\partial \gamma_j}{\partial x_i}\right) + \frac{1}{2}\left(\frac{\partial \gamma_i}{\partial x_j} + \frac{\partial \gamma_j}{\partial x_i}\right) +$$

$+\frac{1}{2}*\frac{1}{2}(x_i+x_j)\left(\frac{\partial\gamma_i}{\partial x_j}+\frac{\partial\gamma_j}{\partial x_i}\right)\frac{\partial\gamma_0}{\partial x_i}+\frac{1}{2}(x_i+x_j)\frac{1}{2}\left(\frac{\partial}{\partial x_j}\left(\frac{\partial\gamma_i}{\partial x_j}+\frac{\partial\gamma_j}{\partial x_i}\right)+\frac{\partial}{\partial x_i}\left(\frac{\partial\gamma_i}{\partial x_j}+\frac{\partial\gamma_j}{\partial x_i}\right)\right)=0,$

as before $\frac{\partial\gamma_4}{\partial x_i}=0,\ T=\text{const}.$

Thus, an exact solution of the modified Boltzmann equation was received.

5.Conclusion

Brownian motion (fluctuations) is a consequence and evidence of the existence of thermal motion and is involved in many physical processes. The mathematical theory of describing such a motion is currently probabilistic in nature and says nothing about the causes and mechanism of the phenomenon. Now we account for reason this effect. On the example of the interaction of three particles, a new position of the center of inertia is established. It is proposed to calculate the corresponding driving force using the moment, which makes it possible to apply the procedure for calculating the force in the interaction of many particles. A model is proposed for including this force to calculate the virial coefficient and to calculate the force in the Langevin equation.

Reference

1. R. Balescu. Equilibrium and nonequilibrium statistic mechanics. A Wiley-Intersciences Publication John Willey and Sons. New-yourk-London…1975.
2. P.Resibois, M.De Lener. Classical kinetic theory of fluids. John Wiley and Sons. New-York, London, … 1977.
3. N.G. Van Kampen. Stochastic processes in physics and chemistry. North-Holland. 1984.
4. Physics of Simple Liquids. Edited by H.N.V. Temperley, I.S. Rowlinson, G.S. Rushbrooke. Amsterdam, 1068. 308.
5. L5. Boltzmann. Selected Works. Moscow: Nauka, 1984.
6. D.N. Zubarev. Nonequilibrium Statistical Thermodynamics), Moscow: Nauka, 1971.
7. V. Ya. Rudyak. Statistical aerohydromechanics of homogeneous and heterogeneous media. Vol.2 Novosibirsk: NGASU. 2005
8. E. V. Prozorova. Features of the rarefied gas description in terms of a distribution function. APhM2018.IOP Conf. Series: Journal of Physics: Conf. Series 1250 (2019) 012023.IOP Publishing. doi:10.1088/1742-6596/1250/1/012023 2.
9.Evelina Prozorova. Influence the form of writing conservation laws in computation. JP Journal of Heat and Mass Transfer, October 28, 2019; Accepted: November 23, 2019
10. Evelina Prozorova. The Effect of Angular Momentum and Ostrogradsky-Gauss Theorem in the Equations of Mechanics WSEAS TRANSACTIONS on FLUID MECHANICS DOI: 11.37394/232013.2020.15.2
12.Evelina Prozorova The influence of the no symmetric stress tensor on the flow separation WSEAS TRANSACTIONS on APPLIED and THEORETICAL MECHANICS DOI: 10.37394/232011.2020.15.9
13.Evelina Prozorova Consequences of the Ostrogradsky-Gauss theorem for numerical simulation in aeromechanics. Elsever Granthaalayah, Volume 8 Issue 6 June 2020
DOI: https://doi.org/10.29121/granthaalayah.v8.i6.2020.549
14. A. L. Khomkin, A. S. Shumikhin. Plasma frequency, parabolic paths and conductivity of non-ideal fully ionized plasma. High Temperature Thermal Physics RAS, vol. 58, No. 3.2020
15. Morozov A.N., Skripkin A.V. Application of integral transforms to a description of the Brownian motion by a non-Markovian random process // Russian Physics Journal. Volume 52, N. 2 .2009
16. Morozov A.N., Skripkin A.V. Spherical particle Brownian motion in viscous medium as non-Markovian random process // Physics Letters A. . Vol. 375. 2011
17. I.G. Kaplan. Introduction to the theory of intermolecular interactions. M .: Science. 1982.
18. N.N. Bogolyubov. Problems of dynamic theory in statistical physics. Kiev.T.2.1970.
19. V.V. Vedgnyapin, S.Z. Adjiev. Entropy according to Boltzmann and Poincaré. Advances in Mathematical Sciences. T.69 N 6. 2014.
20. S. V. Vallander. The equations for movement viscosity gas.// DAN SSSR V. LXX III , N 1 1951.
21. M.N. Kogan, The dynamics of the rarefied gases. M.: Nauka, 1967.

Analysis of the logistic and skew tent map for smart coupling over a finite field

Zongchao Qiao[1], Ina Taralova[1], Mazen Saad[2], and Safwan El Assad[3]

[1] Ecole Centrale de Nantes, LS2N UMR CNRS 6004, Nantes, France
(E-mail: zongchao.qiao@ec-nantes.fr; ina.taralova@ec-nantes.fr)
[2] Ecole Centrale de Nantes, Laboratoire de Mathématiques Jean Leray, UMR CNRS 6629, Nantes, France
(E-mail: mazen.saad@ec-nantes.fr)
[3] Université de Nantes/Polytech Nantes, IETR UMR CNRS 6164, Nantes, France
(E-mail: safwan.elassad@univ-nantes.fr)

Abstract. Chaotic maps have been proved to be efficient in the design of pseudo-random number generator (PRNG). However, the great majority of chaotic maps use real numbers. Due to the high sensitivity of chaos and the finite precision nature of digital device, when a PRNG based on real numbers is numerically implemented, quantization and round-off errors may occur and consequently lead to security breach. Besides, initial conditions and parameters of chaotic maps constitute the seed of a PRNG and pseudo-chaotic behavior has to be guaranteed for all initial conditions. Logistic map and skew tent map are supposed to exhibit good chaos with well defined parameters, but in some particular initial conditions, their trajectories will be trapped into fixed points and lose the chaos quality. For this, we analyze logistic and skew tent map from the perspective of inverse maps in order to find all these unexpected seeds (fixed points and their preimages). To overcome the drawbacks caused by real numbers, a robust PRNG scheme based on a smart coupling of integer chaotic maps over a 32−bit finite field is proposed in this paper. The coupling method can improve the nonlinear dynamics and enhance the randomness efficiently. Simulation results have demonstrated that the proposed PRNG can produce pseudo-random numbers and this PRNG is suitable for encryption systems or other engineering applications.
Keywords: Chaotic map coupling, Logistic map, Skew tent map, Finite field, Pseudo-random number, Encryption.

1 Introduction

Pseudo-random number generators (PRNGs) are vital components for a plethora of applications, from noise simulation in statics and control, to secure information transmission and cryptography[1].

The pseudo-random feature implies that randomness is combined with perfect reproducibility. The seed of a PRNG determines uniquely its output sequence, and guarantees its reproducibility: the same seed will generate the same output sequence, and a different seed is supposed to generate another

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

uncorrelated sequence with features close to true random numbers[2]. This property is crucial for most applications and it makes the PRNG appear on the list of the most used tools in engineering, economics and physics, etc. As an example, considering the application in control domain, the PRNG can be used to simulate erroneous sensor measurements (external noise), un-modeled dynamics (internal noise). The robustness of different control laws should be compared for identical "noise" provided by the PRNG, otherwise the obtained results may be biased and eventually unreliable. Equivalently, in cryptography, decryption must use the identical key stream (pseudo-random numbers) as the encryption in order to recover the same original message.

Numerical methods used in PRNG design are considered to be insecure and have heavy calculation burden[3]. Nowadays, the good news is that the deterministic feature, random-like behavior combined with the extreme sensitivity to the initial conditions have rendered chaotic maps perfect candidates for PRNG design[4]. The initial conditions and parameters constitute the seed of a PRNG.

Designing a chaotic PRNG requires a careful and wise choice of the individual chaotic maps. One-dimensional (1-D) chaotic maps have advantages of good chaos with simple structure and lower computational cost[5]. But they cannot be used alone as PRNG owing to their not long enough periods, attainable map functions and uneven distributions, etc. Also, in digital devices, it is inevitable that finite precision will cause dynamical degradation in chaotic systems. Hence, an efficient algorithm to make the most use of the chaotic maps to design PRNGs is another important issue. In the open literature, based on multiple chaotic maps, some effective methods have been proposed to overcome the dynamical degradation and enhance the chaos property, such as coupling different chaotic maps [6,7], integrating chaotic maps[4,8], multiplexing mechanism[9,10], permutation approaches[1,11], linear feedback shift register operations[12].

However, most of the proposed PRNGs are defined by floating-point notation, which have the defects of slow data transfer and inefficient resource utilization from a hardware perspective[3]. Furthermore, because of the high sensitivity of the chaotic systems to the initial conditions and parameters, the chaos properties are also strongly affected by the data type of the chaotic systems when applied to hardware implementation. Therefore, due to the finite nature of the machine number set, these systems may lose the chaotic characteristics because quantization, truncation or round-offs are required when they are numerically realized under a finite precision[13]. Thus, they are not reliable enough to be applied into practical situations.

To overcome these drawbacks, Elmanfaloty and Abou-Bakr[3] brought out a solution of using fixed-point precision notation and proposed a binary chaotic PRNG with 32-bit fraction length. Considering the integer arithmetic with reduced resources utilization is more hardware friendly than the decimals arithmetic, we aim to design PRNG using finite integer numbers.

Pseudo-random numbers play an important role in cryptosytems which demand not only the randomness, but also a high sensitivity to its seed (secret key for a cryptosystem), especially in the stream cipher whose security depends

mainly on its key stream (pseudo-random numbers provided by a PRNG)[14]. Thus, we will propose a new PRNG from the perspective of cryptosystem, but note that the PRNG is not limited to this kind of application.

In our previous works, we have proposed reliable PRNGs based on the 1-D integer chaotic maps for stream cipher and block cipher cryptosystems[15,16]. They all have avoided the degradation security problem and have achieved high security and reliability. They also can be used in other engineering applications.

In this paper, to pursue a new efficient and general coupling method to improve the randomness of the 1-D logistic map and skew tent map, we introduce a smart coupling algorithm based on integer chaotic maps and design a robust PRNG scheme over an N-bit (N=32) integer finite field. It is known that a chaotic attractor possesses an infinity of unstable dense periodic orbits[17,18]. In particular, the fixed points may represent a problem if the "randomly chosen" initial conditions coincide with a periodic point, even though the latter is unstable. This is clearly the case, if starting from any point of the unstable orbit, the trajectory will remain locked at the same periodic point. Both logistic and skew tent maps are non invertible maps. But to overcome this unexpected case and ensure pseudo-chaotic behavior for all initial conditions in PRNG design, we can analyze their inverse maps just to reveal all the possible initial conditions and their preimages that lead to the fixed points. Then, based on this analysis, the logistic map and the skew tent map over an N-bit integer finite-state space are reformulated. After that, a coupling matrix, the kernel of the PRNG design, is applied to break the original orbits of the 1-D chaotic maps for avoiding undesirable dynamic behavior and enhancing the scheme complexity. Finally, statistical and security tests are applied to evaluate the cryptographic properties of the PRNG.

The paper is organized as follows. Section 2 analyzes the logistic map and skew tent map from the inverse function point of view, reformulates their expressions over an N-bit integer field and briefly investigates their qualities for PRNG design. The proposed PRNG and the coupling performance are discussed in Section 3. Section 4 analyzes the cryptographic properties of the proposed PRNG. Section 5 gives the conclusion.

2 1-D chaotic map over integer finite field

2.1 Logistic map

Logistic map is a well-known classical chaotic map defined over a real number domain ranging from 0 to 1, which is given as below:

$$x(n+1) = \mu x(n)\,(1 - x(n)) \tag{1}$$

where $\{x(n), n = 1, 2, 3...\}$ represents the iteration state and $x(0) \in (0, 1)$ is the initial condition; the parameter $\mu \in (0, 4]$ controls the chaotic behavior. The Lyapunov exponent is the largest when $\mu = 4$, indicating the logistic map reaches the complete chaos.

The delayed phase space of the logistic map when $\mu = 4$ is shown in Fig.1 (in blue), where the solid red line means $x(n+1) = x(n)$. The intersections are

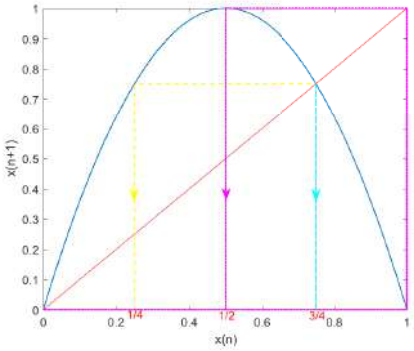


Fig. 1. Delayed phase space and preimages of logistic map over a real domain ($\mu = 4$)

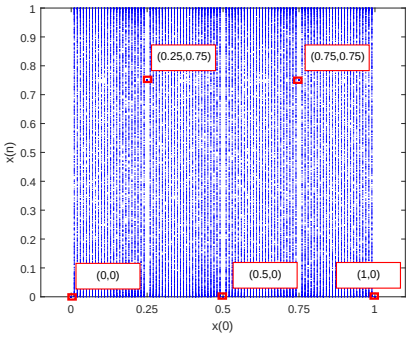


Fig. 2. Iterations $x(n)$ of logistic map versus its initial conditions $x(0)$ over a real domain ($\mu = 4$)

two unstable fixed points: $0, \frac{\mu-1}{\mu}$, which are 0 and $\frac{3}{4}$ when $\mu = 4$. If the initial value is $\frac{3}{4}$, even though the parameter μ equals 4, all the following iterations will be trapped into the fixed point $\frac{3}{4}$. This is an undesirable case when one intends to use the chaos features to design PRNG or encryption purposes. In addition to this, preimages (backward iterates) of the fixed points can cause the fixed point problem as well, which can be seen from Fig.1. The preimages of 0 are $0, \frac{1}{2}$ and 1, while the preimages of $\frac{3}{4}$ are $\frac{1}{4}$ and $\frac{3}{4}$. These values all lead to the fixed points. It also can be observed from Fig.2, which plots 800 iterations $x(n)$ $(n = 1, 2, ...800)$ versus the different initial conditions $x(0)$: for specific $x(0)(0, 0.25, 0.5, 0.75, 1)$, the corresponding iterations are locked into the fixed points, thus there exists no chaos even for $\mu = 4$. To unearth all preimages of the fixed points, we analyze the inverse map.

In the literature, plenty of papers have investigated the logistic map, but very few of them analyze the inverse map. However, it is very important to ascertain chaotic behavior so that the trajectories don't get locked into the fixed point. Because the unstable fixed points for the iterated map behave as stable fixed points by the inverse map. In other words, the preimages of the unstable fixed points converge towards the fixed points under forward iterations. Therefore, not only the fixed points, but also their preimages have to be avoided as a seed in order to guarantee chaotic behavior for $\mu = 4$.

The preimages of logistic map can be obtained by Eq.(2)

$$x(n-1) = \frac{\mu \pm \sqrt{\mu^2 - 4\mu x(n)}}{2\mu} \tag{2}$$

The preimages $(x(n-1), x(n-2), \dots)$ of the fixed points $(x(n) = 0, \frac{3}{4})$ are summarized in Table 1. We can find that, if ignoring the irrational numbers in the range of $(0, 1)$, the initial conditions to be avoided are $\frac{1}{4}, \frac{1}{2}$ and $\frac{3}{4}$.

Logistic map redefined over the N-bit integer finite filed when $\mu = 4$ is given by Eq.(3).

$x(n)$	$x(n-1)$	$x(n-2)$	$x(n-3)$	$\cdots$
0	0	0	0	$\cdots$
			1	$\cdots$
		1	$\frac{1}{2}$	$\cdots$
			1	$\cdots$
	1	$\frac{1}{2}$	$\frac{2\pm\sqrt{2}}{4}$ (irrational value)	-
		1	$\frac{1}{2}$	$\cdots$
			1	$\cdots$
$\frac{3}{4}$	$\frac{1}{4}$	$\frac{2\pm\sqrt{3}}{4}$ (irrational value)	-	-
	$\frac{3}{4}$	$\frac{1}{4}$	$\frac{2\pm\sqrt{3}}{4}$ (irrational value)	-
		$\frac{3}{4}$	$\frac{1}{4}$	$\cdots$
			$\frac{3}{4}$	$\cdots$

Table 1. Preimages of the fixed points of logistic map

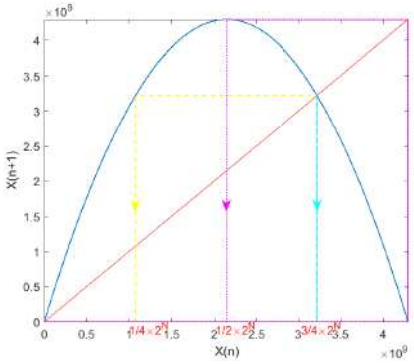


Fig. 3. Delayed phase space and preimages of logistic map over the 32-bit integer field ($\mu = 4$)

$$X(n+1) = \begin{cases} 2^N - 1, & X(n) = \frac{3}{4} \times 2^N \; or \; 2^N \\ \left\lfloor \frac{X(n) \times \left(2^N - X(n)\right)}{2^{N-2}} \right\rfloor, & otherwise \end{cases} \tag{3}$$

where $\{X(n), n = 1, 2, 3...\}$ is the produced chaotic sequence by iterations and all values are integers ranging in $\left[1, 2^N - 1\right]$; symbol $\lfloor \cdot \rfloor$ means that each element in it rounds to the nearest integer less than or equal to the element.

The delayed phase space of Eq.(3) is displayed in Fig.3, where there obviously exist two unstable fixed points: $X(n) = 0, \frac{3}{4} \times 2^N$. Similar to the analysis of Fig.1, the preimages of the fixed points in the range of $[1, 2^N - 1]$ are $\frac{1}{4} \times 2^N$, $\frac{1}{2} \times 2^N$ and $\frac{3}{4} \times 2^N$ that should be avoided carefully. Considering $\frac{1}{4} \times 2^N$ produces $\frac{3}{4} \times 2^N$ after one iteration, Eq.(3) just needs to deal with the value $\frac{3}{4} \times 2^N$.

2.2 Skew tent map

Skew tent map is derived from the classical tent map but it achieves better statistical performances. Skew tent map defined in real domain $(0, 1)$ is given by Eq.(4).

$$x(n+1) = \begin{cases} \frac{x(n)}{p}, & 0 < x(n) < p \\ \frac{1-x(n)}{1-p}, & p \leqslant x(n) < 1 \end{cases} \tag{4}$$

where $\{x(n), n = 1, 2, 3...\}$ represents the iteration state and $p \in (0, 1)$ is the control parameter.

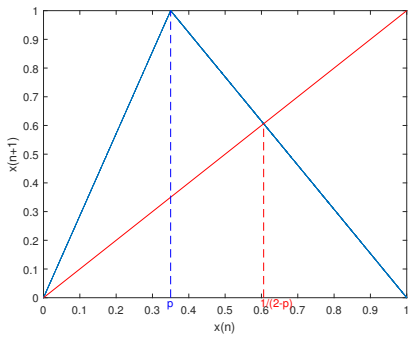


Fig. 4. Delayed phase space of skew tent map over a real domain

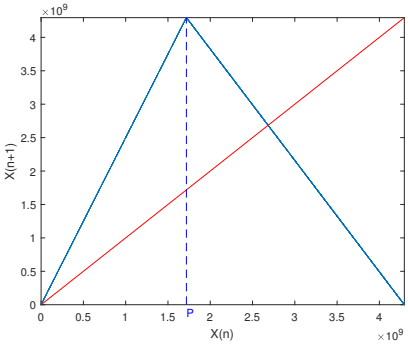


Fig. 5. Delayed phase space of skew tent map over the 32-bit integer field

The delayed phase space diagram of the skew tent map is shown in Fig.4 where the solid red line reveals the unstable fixed points: 0 and $\frac{1}{2-p}$. From the inverse function point of view, the preimages have the following iteration relation:

$$x(n-1) = \{p \times x(n), 1 - (1-p)\, x(n)\} \tag{5}$$

The skew tent map over the N-bit integer field is given by Eq.(6):

$$X(n+1) = \begin{cases} \left\lfloor 2^N \times \frac{X(n)}{P} \right\rfloor, & 0 < X(n) < P \\ \left\lfloor 2^N \times \frac{2^N - X(n)}{2^N - P} \right\rfloor, & P < X(n) < 2^N \\ 2^N - 1, & otherwise \end{cases} \tag{6}$$

where the iterated state is $X(n) \in \mathbf{N}_+$ and $X(n) \in [1, 2^N - 1]$; $P \in \mathbf{N}_+$ is the control parameter and $P \in [1, 2^N - 1]$.

The unstable fixed point 0 does not belong to the region of definition, while the other fixed point in the mapping shown in Fig.5 must satisfy one condition: the equation of $X_{fixedpoint} = 2^N \times \frac{2^N - X_{fixedpoint}}{2^N - P}$ should have integer solutions $X_{fixedpoint}$. It is a bit complicated to analyze the preimages of the fixed point without restricting the parameter P. We analyze the parameter P in the range of $[1, 2^{32} - 1]$ in turn and find when using specific P, there exist preimages

that can lead to the fixed point, for instance, if $P = 262140$, $X_{fixedpoint} = 2147549185$. In order to prevent the trajectory from being trapped into the fixed point, although this rarely happens, we add the following statement in the algorithm:

$$X(n+1) = X(n+1) - 1, if\ X(n+1) = X(n) \tag{7}$$

2.3 Analysis for encryption purposes

Lyapunov exponent The Lyapunov exponent characterizes the stability of a chaotic motion by measuring the average exponential divergence between two nearby trajectories. If the Lyapunov exponents have a positive value, the chaotic map shows chaotic behavior and the larger this value is, the better the chaotic performances are[19].

The Lyapunov exponent of the logistic map ($\mu = 4$) is 0.6931. Fig.6 demonstrates that the estimated Lyapunov exponents of the skew tent map are always positive for the range of interest of parameter P, exhibiting the maximum 0.6939 for $P = \frac{1}{2} \times 2^N$. At this parameter value, the chaoticity of skew tent map is equivalent to that of the logistic map.

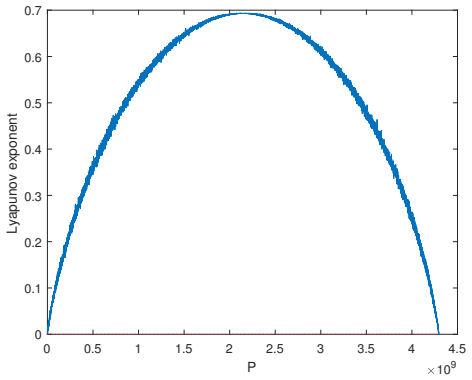


Fig. 6. Estimated Lyapunov exponents of skew tent map

Histogram Histograms of the chaotic maps Eq.(3) and Eq.(6),(7) defined over 32-bit integer field are plotted in Fig.7 and Fig.8 in 1000 classes, where 2×10^6 values are generated for each map but the first 10^6 are considered transient and removed; the red lines mean the average values in every 10 classes.

Uniformity is one of the important criteria for randomness. According to Fig.7 and Fig.8, we can observe that the piece-wise linear map (skew tent map) shows much better uniform distribution features than the logistic map.

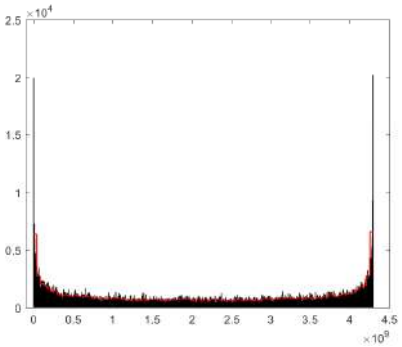


Fig. 7. Histogram of logistic map

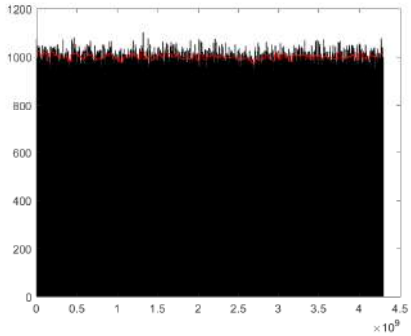


Fig. 8. Histogram of skew tent map

Key space contribution In encryption applications, a large secret key space of PRNG is necessary to resist the brute-force attack and it is considered to be secure if the key space is greater to 2^{128}[20]. Unlike the statistical and security performances, which only can be tested after completing the design of PRNG, key space needs to be taken into account when we are conceiving a new PRNG scheme for cryptosystems.

Initial conditions and parameters form the key space. For logistic map, only N-bit initial value can be taken into the key space, while for the skew tent map, besides the N-bit initial value, N-bit parameter P can be counted into the key space as well. From this point of view, skew tent is able to provide N more bit of key space than the logistic map.

3 Proposed PRNG

In this section, we first give the proposed PRNG scheme whose core is a smart coupling structure. This coupling is inspired by the idea of weak coupling over a real number domain in our previous work[21,22]. Then, we analyze the coupled effect using logistic map and skew tent map. Considering the key space issue in encryption system, two skew tent maps are chosen to construct the PRNG.

3.1 Proposed PRNG scheme

The proposed PRNG scheme is shown in Fig.9.

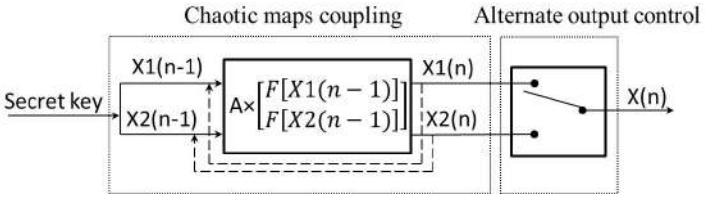


Fig. 9. The proposed PRNG scheme

It mainly contains two operations: first, a new proposed smart chaotic maps coupling is used to break the original chaotic orbits, enhance the complexity and improve the chaotic property; the alternate output control aims to mix the coupled numbers and increase the unpredictability.

The coupling matrix A is defined as follows:

$$A = \begin{bmatrix} 17 - e & e \\ 2e & 31 - 2e \end{bmatrix} \tag{8}$$

$e \in [1, 2^4 - 1]$ is a coupling control parameter.

The coupling process is described as below:

$$\begin{bmatrix} X1\,(n) \\ X2\,(n) \end{bmatrix} = A \times \begin{bmatrix} F\,[X1\,(n-1)] \\ F\,[X2\,(n-1)] \end{bmatrix} \tag{9}$$

where F represents the chaotic functions and it can be a similar type of chaotic map or two different kinds of maps; $X1(n-1)$and $X2(n-1)$ are the previous states of the current states $X1(n)$and $X2(n)$.

The final output chaotic sequence X is controlled by selecting the intermediate outputs $X1$ and $X2$ alternately (another switching law is also possible):

$$X\,(n) = \begin{cases} X1\,(n)\,, when\ mod\,(n,2) = 1 \\ X2\,(n)\,, when\ mod\,(n,2) = 0 \end{cases} \tag{10}$$

3.2 Coupling performance

The behavior of the final output sequence depends highly on the coupling performance. Here, we first couple two different maps: logistic map (Eq.(3)) and skew tent map (Eq.(6),(7)), and analyze the coupling performance in terms of the statistical histogram distribution and delayed phase space behavior.

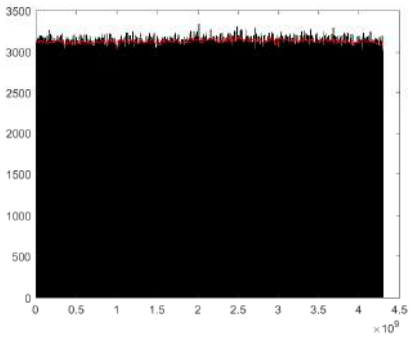


Fig. 10. Histogram of $X1$

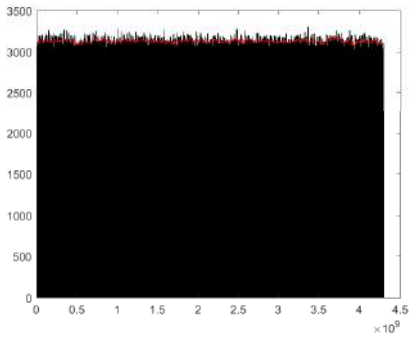


Fig. 11. Histogram of $X2$

From the histograms of intermediate outputs $X1$ and $X2$ with length of 3125000 (distributed in 1000 classes) shown in Fig.10 and Fig.11, we can observe that the sequence after the coupling algorithm is able to achieve a uniform

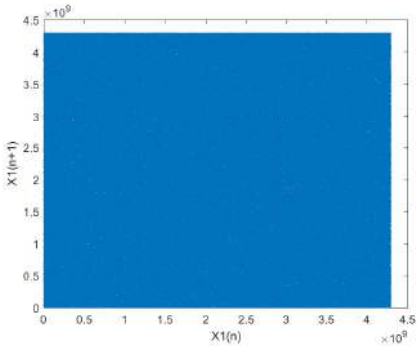


Fig. 12. Delayed phase space of $X1$

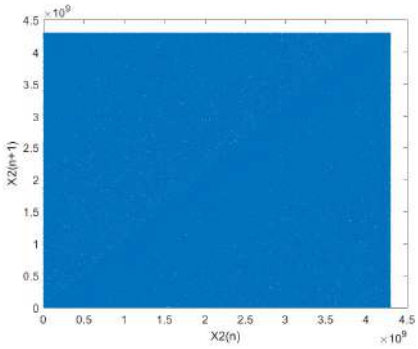


Fig. 13. Delayed phase space of $X2$

distribution. According to the delayed phase space diagrams shown in Fig.12 and Fig.13, the coupling matrix can hide the generating function effectively, which is required for most applications in security.

If logistic map and skew tent map are used to design the PRNG, the key space contains the initial values of these two maps (each is in 32 bits), a parameter P (32 bits) for the skew tent map and a control parameter e (4 bits) for the coupling matrix A. Thus, the key space is 2^{100} in total, which is not large enough for encryption purposes. Hence, this coupling combination can be used to design PRNG, but needs to parallel the coupling scheme to expand the key space for encryption applications.

Skew tent map contributes more key space than logistic map. Apart from this advantage, the skew tent map has an approximately uniform distribution, which outperforms most of the well-known chaotic maps. Thus, we use two skew tent maps to design the PRNG for cryptosystem.

The secret key of this PRNG contains the initial conditions $(Xs1(0), Xs2(0))$, the parameters $(P1, P2)$ for skew tent maps, and the coupling control parameters e. Thus, the key size is :

$$|K| = |Xs1\,(0)| + |P1| + |Xs2\,(0)| + |P2| + |e| = 132\ bits \tag{11}$$

where $|Xs1\,(0)| = |P1| = |Xs2\,(0)| = |P2| = 32\ bits$ and $|e| = 4\ bits$.

Therefore, the key space of this proposed PRNG is 2^{132}, which is large enough to make the brute-force attack infeasible if the PRNG is used for cryptography.

The coupling performances of $X1$,$X2$ when using skew tent maps are as good as when coupling the logistic map and skew tent map. The final output is the one which will be exploited, so in the following analyses, we just give the performance test results of the final output chaotic sequence X.

4 Performance analysis

The seed and the output of the PRNG are also called secret key and key stream respectively in encryption applications. To guarantee high security, the key

stream must be random enough to ensure no statistical information is exposed to hackers so that unauthorized ones cannot deduce the inner states or even recover the secret key. Thus, the PRNG should have random statistical performance and good security property. This section analyzes these performances by delayed phase space graph, histogram and χ^2 test, key sensitivity and NIST test. In these tests, each test sequence has 3125000 values (3125000 $\times$ 32 bits $= 100 \times 10^6$ bits). All simulations are conducted in MATLAB (R2017b) and each secret key is randomly created.

4.1 Delayed phase space

The delayed phase space of the final output sequence X has been drawn in Fig.14, where 3125000 values are generated and the last 10^6 values are plotted. Contrary to the easily identified mapping function of the original chaotic map shown in Fig.5, the final output sequence is distributed randomly in the delayed phase space and shows more complex dynamical behavior thanks to the coupling and alternate output control operations. Thus, it is impossible for potential attackers to analyze the iteration trajectory and they cannot find a hint of which chaotic map we use in the PRNG.

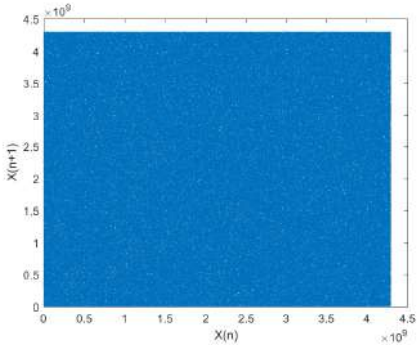


Fig. 14. Delayed phase space of X

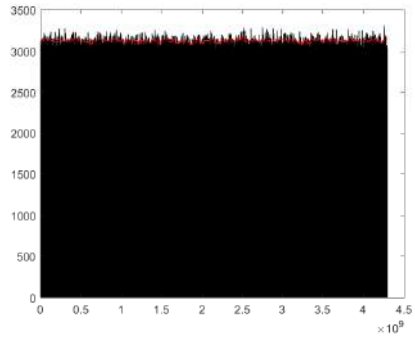


Fig. 15. Histogram of X

4.2 Histogram

A basic requirement for a robust PRNG is that the generated chaotic sequence have a uniform distribution. The histogram of the chaotic sequence X with length of 3125000 is drawn in Fig.15 in 1000 classes, which shows visually that the generated sequence is uniformly distributed in the whole definition field.

4.3 χ^2 test

To analyze the uniformity more precisely, the χ^2 test is applied. The experimental value χ^2_{exp} is calculated by Eq.(12):

$$\chi^2_{\text{exp}} = \sum_{i=0}^{K-1} \frac{(O_i - E_i)^2}{E_i} \tag{12}$$

where $K = 1000$ is the number of classes, O_i is the number of observed values in the $i-th$ class and E_i is the expected number in a uniform distribution. The theoretical value $\chi^2_{theo}(K, \alpha)$ equals to 1073.64 which is obtained for a threshold $\alpha = 0.05$. If $\chi^2_{exp} < \chi^2_{theo}(K, \alpha)$, the test sequence can be considered to have a uniform distribution.

Here, we use 100 different secret keys to produce 100 chaotic sequences. Each contains 3125000 values, hence, $E_i = 3125000/1000$. χ^2_{exp} is calculated for each sequence. The average $\chi^2_{exp} = 1004.49$ that is smaller than $\chi^2_{theo}(K, \alpha)$. Thus, χ^2_{exp} test has confirmed the uniformity of the output chaotic sequence.

4.4 Key sensitivity

The generated chaotic sequence should show high sensitivity to the secret key (seed). This property is necessary and important for resisting differential attack and chosen-plaintext attack. The key sensitivity can be measured by Hamming Distance (HD) given as follows:

$$HD\,(X, Y) = \frac{1}{Nb} \times \sum_{k=1}^{Nb} (X\,(k) \oplus Y\,(k)) \tag{13}$$

where X and Y are two output chaotic sequences from the proposed PRNG whose secret keys are just one bit (randomly chosen) different; Nb is the bit length in a sequence and $\oplus$ represents the XOR operator.

Here, we use 100 different secret keys to produce 100 pairs of X and Y. Then, 100 HDs are computed by Eq.(13). The average HD is 49.9989 which is very close to the optimal HD value 50% (bit change probability). This result means the high secret key sensitivity is achieved.

4.5 NIST test

NIST (National Institute of Standard and Technology) test is a suite of tests which is widely used to measure sequences for randomness. We apply the NIST test on the produced sequence ($3125000 * 32\ bits = 100 \times 10^6\ bits$). The results in Table 2 have demonstrated that the generated chaotic sequence has passed the NIST test successfully, which has verified the output sequence of the PRNG is pseudo-random.

Test	P-value	Proportion	Results
Frequency test	0.798	99.000	Passed
Block-frequency test	0.290	97.000	Passed
Cumulative-sums test	0.765	98.500	Passed
Runs test	0.679	99.000	Passed
Longest-run test	0.494	99.000	Passed
Rank test	0.475	100.000	Passed
FFT test	0.658	100.000	Passed
Nonperiodic-templates	0.502	98.973	Passed
Overlapping-templates	0.924	97.000	Passed
Universal	0.658	99.000	Passed
Approximty entropie	0.964	100.000	Passed
Random-excursions	0.441	98.182	Passed
Random-excursions-variant	0.328	98.788	Passed
Serial test	0.906	99.500	Passed
Linear-complexity	0.154	99.000	Passed

Table 2. Results of NIST test

5 Conclusion

In this paper, a smart coupling based on the reformulated logistic and skew tent maps has been proposed to design a robust PRNG. These two maps have been redefined using integers over the 32-bit finite field. On one hand, this has avoided the iteration states being locked into a fixed point (or its preimages), and on the other hand, this can solve the security problems caused by applying the real domain defined chaotic maps into finite precision hardware implementations. In addition, the smart coupling has overcome the dynamical degradation existed in chaotic maps with finite precision. Composed of the coupling and the output control operation, the proposed PRNG structure can enhance the nonlinear dynamics and increase the complexity effectively.

Conducted simulations results have demonstrated that the proposed PRNG is able to produce pseudo-random numbers with good randomness and cryptographic properties. Therefore, this PRNG can be used in the design of cryptosystems or any other pseudo-random generator required applications.

References

1. Lambi D, Nikoli M. Pseudo-random number generator based on discrete-space chaotic map. Nonlinear Dynamics. 2017 Oct 1;90(1):223-232.
2. Jallouli O, El Assad S, Taha MA, Chetto M, Lozi R, Caragata D. An efficient pseudo chaotic number generator based on coupling and multiplexing techniques. InInternational Conference on Emerging Security Information, Systems and Technologies (SECURWARE 2016) 2016 Jul 24 (pp. 35-40).
3. Elmanfaloty RA, Abou-Bakr E. Random property enhancement of a 1D chaotic PRNG with finite precision implementation. Chaos, Solitons & Fractals. 2019 Jan 1;118:134-144.

4. Wang XY, Li ZM. A color image encryption algorithm based on Hopfield chaotic neural network. Optics and Lasers in Engineering. 2019 Apr 1;115:107-118.
5. Murillo-Escobar MA, Cruz-Hernndez C, Cardoza-Avendao L, Mndez-Ramrez R. A novel pseudo random number generator based on pseudo randomly enhanced logistic map. Nonlinear Dynamics. 2017 Jan 1;87(1):407-425.
6. Hua Z, Jin F, Xu B, Huang H. 2D logistic-Sine-coupling map for image encryption. Signal Processing. 2018 Aug 1;149:148-161.
7. Sahari ML, Boukemara I. A pseudo-random numbers generator based on a novel 3D chaotic map with an application to color image encryption. Nonlinear Dynamics. 2018 Oct 1;94(1):723-744.
8. Wang M, Wang X, Zhang Y, Zhou S, Zhao T, Yao N. A novel chaotic system and its application in a color image cryptosystem. Optics and Lasers in Engineering. 2019 Oct 1;121:479-494.
9. Qiao Z, Taralova I, El Assad S. A robust pseudo-chaotic number generator for cryptosystem based on chaotic maps and multiplexing mechanism. International Conference for Internet Technology and Secured Transactions (ICITST'2019), Dec 2019, London, United Kingdom. [Qiao et al] 4 p. hal-02430572
10. Jallouli O, El Assad S, Chetto M, Lozi R. Design and analysis of two stream ciphers based on chaotic coupling and multiplexing techniques. Multimedia tools and applications. 2018 Jun 1;77(11):13391-13417.
11. Franois M, Grosges T, Barchiesi D, Erra R. Pseudo-random number generator based on mixing of three chaotic maps. Communications in Nonlinear Science and Numerical Simulation. 2014 Apr 1;19(4):887-895.
12. Xu H, Tong X, Meng X. An efficient chaos pseudo-random number generator applied to video encryption. Optik. 2016 Oct 1;127(20):9305-9319.
13. Yang, B., Liao, X. Period analysis of the logistic map for the finite field. Science China Information Sciences, 2017,60(2), 022302.
14. Lambi D. Security analysis and improvement of the pseudo-random number generator based on quantum chaotic map. Nonlinear Dynamics. 2018 Oct 1;94(2):1117-1126.
15. Qiao Z, Taralova I, El Assad S. Efficient pseudo-chaotic number generator for cryptographic applications. International Journal of Intelligent Computing Research(IJICR). 2020 March Volume 11, Issue 1:1041-1048.DOI: 10.20533/ijicr.2042.4655.2020.0126
16. Qiao Z, El Assad S, Taralova I. Design of secure cryptosystem based on chaotic components and AES S-Box. AEU-International Journal of Electronics and Communications. 2020 May 5:153205.
17. Burns K, Hasselblatt B. The Sharkovsky theorem: A natural direct proof. The American Mathematical Monthly. 2011 Mar 1;118(3):229-244.
18. Li TY, Yorke JA. Period three implies chaos. In The Theory of Chaotic Attractors 2004 (pp. 77-84). Springer, New York, NY.
19. Lee WK, Phan RC, Yap WS, Goi BM. SPRING: a novel parallel chaos-based image encryption scheme. Nonlinear Dynamics. 2018 Apr 1;92(2):575-593.
20. Özkaynak F. Brief review on application of nonlinear dynamics in image encryption. Nonlinear Dynamics. 2018 Apr 1;92(2):305-313.
21. Garasym O, Lozi R, Taralova I. Robust PRNG based on homogeneously distributed chaotic dynamics. InJournal of Physics: Conference Series 2016 Feb (Vol. 692, p. 012011).
22. Lozi R, Taralova I. From chaos to randomness via geometric undersampling. ESAIM: Proceedings and surveys. 2014 Nov 1;46:177-195.

Information measures and synchronization in regular ring lattices with discontinuous local dynamics

J. Leonel Rocha[1] and S. Carvalho[2]

[1] CEAUL. ADM, ISEL-Engineering Superior Institute of Lisbon, Polytechnic Institute of Lisbon, Rua Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal (E-mail: jrocha@adm.isel.pt)

[2] CEAFEL. ADM, ISEL-Engineering Superior Institute of Lisbon, Polytechnic Institute of Lisbon, Rua Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal (E-mail: E-mail:sonia.carvalho@isel.pt)

Abstract. We study information measures and synchronization in complete dynamical networks of maps, with local identical chaotic dynamical systems. The network topologies are regular ring lattices which are characterized by circulant matrices and the conditional Lyapunov exponents are explicitly determined. For discontinuous local dynamics, some properties of the mutual information rate and the Kolmogorov-Sinai entropy are established, depending on the topological entropy of the individual chaotic nodes and on the synchronization interval. It is proved that as large as the network topology is, measured by its network topological entropy and directly related with the network order, the information measures studied increase or decrease, according to the network order in relation to the synchronization interval. Some numerical studies are included.

Keywords: Mutual information rate, Kolmogorov-Sinai entropy, synchronization, complete networks, discontinuous dynamics, Lyapunov exponents, topological order, circulant matrix.

1 Introduction

In the last decades, several authors have dedicated their investigation to the study of the information theory and its applications. The amount of information produced by a network may be measured by the mutual information rate. This measure together with the Kolmogorov-Sinai entropy are expressed in terms of the conditional Lyapunov exponents. On the other hand, it is well known that chaotic systems can be synchronized. The recognized potential for communications systems has driven this phenomenon to become a distinct subfield of nonlinear dynamics.

Information theory and synchronization are directly related in a network. Motivated by the theoretical and practical connection between the information measures and the phenomenon of synchronization, our purpose in this paper is to analyze the relations between the mutual information rate, the Kolmogorov-Sinai entropy and the synchronization in a space of complete dynamical network of maps of order $N \in \mathbb{N} \setminus \{1\}$. The discontinuous local dynamics considered at each node establish the topological, metrical and chaotic complexity of the

network that is being studied. Discontinuous dynamical systems are recurrently found in physical systems, which are also used in various applications in engineering, economic, biological and ecological models, among others, see, for example, [1], [9] and [16]. The study of discontinuous dynamics in synchronization phenomena has also attracted the attention of several researchers, see [10] and the other works of this same volume and issue.

The paper is organized as follows: In Sec.2 are presented preliminar definitions and results. We start Sec.3 with the analysis of the case where the local dynamics are given by discontinuous piecewise linear maps with slope $s > 1$. We obtain explicit expressions for the synchronization interval and for the parallel and transversal Lyapunov exponents. For this case it is proved that to stabilize the synchronized states, it suffices to require that the transversal Lyapunov exponent is negative. Some properties of the mutual information rate and the Kolmogorov-Sinai entropy, depending on the slope s and the synchronization interval are established. We also study the approach to a topological invariant associated with the dynamics between the nodes of the complete network: the network topological entropy. In this context is established a topological order: it is proved that as large as the network topology, measured by its network topological entropy and directly related with the network order, the information measures studied increases or decreases, according to the network order in relation to the synchronization interval. Numerical simulations are performed to obtain more information and complement the theoretical results presented. Finally, in Sec.4, we discuss our work and provide some conclusions.

2 Preliminars

An active channel is usually described by an active network constructed using $N \in \mathbb{N} \setminus \{1\}$ elements that have some intrinsic dynamics and can be characterized by classical dynamical systems, such as chaotic oscillators, neurons, phase oscillators, and so on. Throughout this work we will consider a family of complex networks of chaotic dynamical systems defined by complete networks of order N with $\frac{N(N-1)}{2}$ edges and discontinuous local dynamics. These networks of N identical chaotic dynamical oscillators or units, are described by a connected and unoriented graph $G = (V, E)$, where V represents the vertices (nodes), and E the edges of G, with no loops and no multiple edges, where every vertex of G has degree $N-1$. The space of complete dynamical networks with N nodes will be denoted by K_N.

Consider A the adjacency matrix of K_N and $D = diag(N-1, \ldots, N-1)$, then $L = [l_{ij}] = A - D$ represents the laplacian matrix of the complete graph and is written in the following form,

$$L = \begin{bmatrix} -(N-1) & 1 & 1 & \ldots & 1 \\ 1 & -(N-1) & 1 & \ldots & 1 \\ \ldots & \ldots & \ldots\ldots & & \ldots \\ 1 & 1 & \ldots & 1 & -(N-1) \end{bmatrix}.$$

The dynamics of these N coupled oscillators can be expressed by the following system of differential equations,

$$\dot{x}_i = f(x_i) + \sigma \sum_{j=1}^{N} l_{ij} x_j, \tag{1}$$

where f is a vector-valued map describing the dynamics of the nodes, $\sigma > 0$ is the coupling strength or parameter and $i = 1, 2, ..., N$.

However, the state equations of the complex network given by Eq.(1), can be rewritten in the discretized form as,

$$x_i(k+1) = f(x_i(k)) + \sigma \sum_{j=1}^{N} l_{ij} f(x_j(k)), \tag{2}$$

which is also known as a complex dynamical network of maps, see, for example, [8] and [12]. Let f' be the derivative of f, then the jacobian matrix of this dynamical network K_N is written as follows,

$$J = \begin{bmatrix} f' - (N-1)\sigma f' & \sigma f' & \dots & \sigma f' \\ \sigma f' & f' - (N-1)\sigma f' & \dots & \sigma f' \\ \dots & \dots & \dots & \dots \\ \sigma f' & \sigma f' & \dots & f' - (N-1)\sigma f' \end{bmatrix}.$$

Every matrix associated with a complete network K_N has a certain regularity, so we are able to determine its spectra and the associated eigenspaces. Let $\mu_1 < \mu_2 \leq \ldots \leq \mu_N$ and $\lambda_1 < \lambda_2 \leq \ldots \leq \lambda_N$ be the eigenvalues of the laplacian and the jacobian matrices of K_N, respectively. Notices that the matrices A, L and J are irreducible matrices.

1. Both matrices L and J are circulant matrices, so they are diagonalizable and have the same eigenspaces. Let $x^{(N)} = (1, 1, \ldots, 1)$, this is an eigenvector of every circulant matrix, and it is associated with the eigenvalue $\mu_1 = 0$ and $\lambda_1 = f'$, respectively, that is equal to the row sum of each matrix.
2. *The other eigenvectors of a circulant matrix.*
   Let $\omega_N = \exp^{\frac{2\pi i}{N}}$ be one of the N-th complex roots of 1. It is known that, for $1 \leq k \leq N$,
   $$x^{(k)} = \left(\omega_n^{0k}, \omega_n^{1k}, \ldots, \omega_N^{(N-1)k}\right)$$
   is an eigenvector of every circulant matrix C where every row has the elements $\{c_1, \ldots, c_N\}$. In particular, if $k = N$ we obtain $x^{(N)} = (1, 1, \ldots, 1)$. Moreover, the eigenvalue associated with $x^{(k)}$, considering the regularity of C and ω_N^k, is equal to $\sum_{j=1}^{N} c_j \omega_N^{jk}$.

Let us recall an important property of the sum of the N complex roots of the unit, i.e.,

$$\sum_{j=1}^{N} \omega_N^{jk} = \begin{cases} N, & \text{if } k \equiv 0 \,(\bmod\, N) \\ 0, & \text{otherwise} \end{cases}.$$

We can also state that, the laplacian matrix L has exactly two eigenvalues $\mu_1 = 0$, a simple root, and $\mu_2 = -N$, with multiplicity $N-1$, and the jacobian matrix J has also two eigenvalues $\lambda_1 = f'$, also a simple root, and $\lambda_2 = f'(1 - N\sigma)$, with multiplicity $N-1$. Notice that, in the context of the study of information measures, the eigenvalue λ_1 measures the exponential divergence of nearby trajectories in the direction of the synchronization manifold and the eigenvalue λ_2 measures the exponential divergence of nearby trajectories in the direction transversal to the synchronization manifold, see [2] and [3].

In an active network, every pair of elements form a communication channel and the rate with which information is exchanged between these elements, a transmitter S_i and a receiver S_j, is given by the mutual information rate, represented by $I_C(S_i, S_j) = \lambda_{\|}^{+} - \lambda_{\perp}^{+}$, where $\lambda_{\|}^{+}$ denotes the positive Lyapunov exponents, associated to the synchronization manifold, and $\lambda_{\perp}^{+}$ denotes the positive Lyapunov exponents, associated to the transversal manifold, see [2]. The Kolmogorov-Sinai entropy, denoted by H_{KS}, gives a suitable way of obtaining the entropy production of a dynamical system. It also provides a global measure of the amount of information that can be simultaneously transmitted among the network. For systems with a measurable (the trajectory is bounded to a finite domain) and ergodic (average quantities can be calculated in space and time) invariant (with respect to time translations of the system and to smooth transformations) natural measure, that is smooth along the unstable manifold, the Kolmogorov-Sinai entropy is obtained by the sum of the positive Lyapunov exponents, see [2], [3], [6] and [13]. Regarding the case of complete network K_N, where every node is connected with all the others, each node is only one connection apart from any other and there is just one single transversal Lyapunov exponent. Thus, according to the dynamical network given by Eq.(2) and [3], we have the following definitions for the information measures analyzed in this paper,

$$I_C = \begin{cases} \lambda_{\|} - \lambda_{\perp}, & \text{if } \lambda_{\perp} > 0 \\ \lambda_{\|}, & \text{if } \lambda_{\perp} \leq 0 \end{cases} \tag{3}$$

and

$$H_{KS} = \begin{cases} \lambda_{\|} + \lambda_{\perp}, & \text{if } \lambda_{\perp} > 0 \\ \lambda_{\|}, & \text{if } \lambda_{\perp} \leq 0 \end{cases}. \tag{4}$$

Other central point of our investigation is related with the synchronization in the space of complete networks K_N and its relations with the information measures I_C and H_{KS}, just mentioned in Eqs.(3) and (4), respectively. Following the results presented in [8], a dynamical network given by Eq.(2), having in each node identical chaotic nodes ($\chi(f) > 0$), synchronizes in the following interval,

$$\sigma_1 = \frac{1 - e^{-\chi(f)}}{|\mu_2|} < \sigma < \frac{1 + e^{-\chi(f)}}{|\mu_N|} = \sigma_2, \tag{5}$$

where $0 = \mu_1 < |\mu_2| \leq \ldots \leq |\mu_N|$ are the eigenvalues of the laplacian matrix L and $\chi(f)$ is the Lyapunov exponent of each individual n-dimensional node, see also [5] and [6]. Notice that, if each local dynamical node is chaotic, then

the Lyapunov exponent $\chi(f)$ is positive. Throughout this work, the synchronization interval of K_N will be denoted by $I_\sigma =]\sigma_1, \sigma_2[$.

3 Local dynamics: discontinuous piecewise linear maps with slope $s > 1$

In this section we consider the space of all the complete dynamical network of maps K_N, given by Eq.(2), where the local dynamics in each node is defined by a discontinuous piecewise linear one-dimensional (1D) map, $f : I = [b_1, b_2] \subset \mathbb{R} \to I$, with $|I| = 1$ represents the amplitude of the interval I, such that there exist points $b_1 = d_0 < d_1 < \ldots < d_p < d_{p+1} = b_2$, where f has constant slope $s > 1$ everywhere in each subinterval $I_i =]d_i, d_{i+1}[$, $i = 0, \ldots, p$. Generally, the discontinuous piecewise linear map is defined by,

$$f(x) = s\ x + a_i \pmod 1,\ \forall x \in I_i \text{ and } a_i \in \mathbb{R}. \tag{6}$$

In this context the map f has $p-1$ discontinuity points: $d_1, d_2, \ldots, d_p$, see Fig.1. Thus, throughout this section we consider the following parameters space,

$$\Sigma^+ = \left\{(N, s, \sigma) \in \mathbb{R}^3 : N \in \mathbb{N} \setminus \{1\}, s > 1, \sigma > 0\right\}. \tag{7}$$

Since each complete dynamical network K_N has identical chaotic nodes and $|\mu_2| = |\mu_N| = N$, then the synchronization interval is nonempty, for all $s > 1$. Moreover, from Eq.(5), the synchronization interval may be expressed in terms of the topological entropy of f, i.e., the chaoticity of the dynamics of the node map f is measured by $h_{top}(f) = \chi(f) = \log|s|$, see [5] and [11].

Property 1. Consider the (K_N, Σ^+) space of complete dynamical networks, given by Eq.(2). Let f be the discontinuous piecewise linear map with slope $s > 1$ everywhere, given by Eq.(6). The synchronization interval of K_N, defined by Eq.(5), is given by,

$$\sigma_1 = \frac{s-1}{Ns} < \sigma < \frac{s+1}{Ns} = \sigma_2,\ \forall s > 1. \tag{8}$$

Consequently, if the dynamics of the individual nodes, defined by f, of the complete dynamical network K_N are fixed, then the amplitude of the synchronization interval $|I_\sigma|$ decrease, as larger is the order N of the complete dynamical network.

3.1 Synchronization and information measures

Establishing that the local dynamics f is a discontinuous piecewise linear map with slope $s > 1$ everywhere, given by Eq.(6), we have noticed that the jacobian matrix J has only two distinct eigenvalues, $\lambda_1 = s$ and $\lambda_2 = s(1 - N\sigma)$, with multiplicity $N - 1$. So, the parallel Lyapunov exponent is given by,

$$\lambda_\parallel = \int_I \ln|\lambda_1|\, d\bar{\mu} = \ln(s), \tag{9}$$

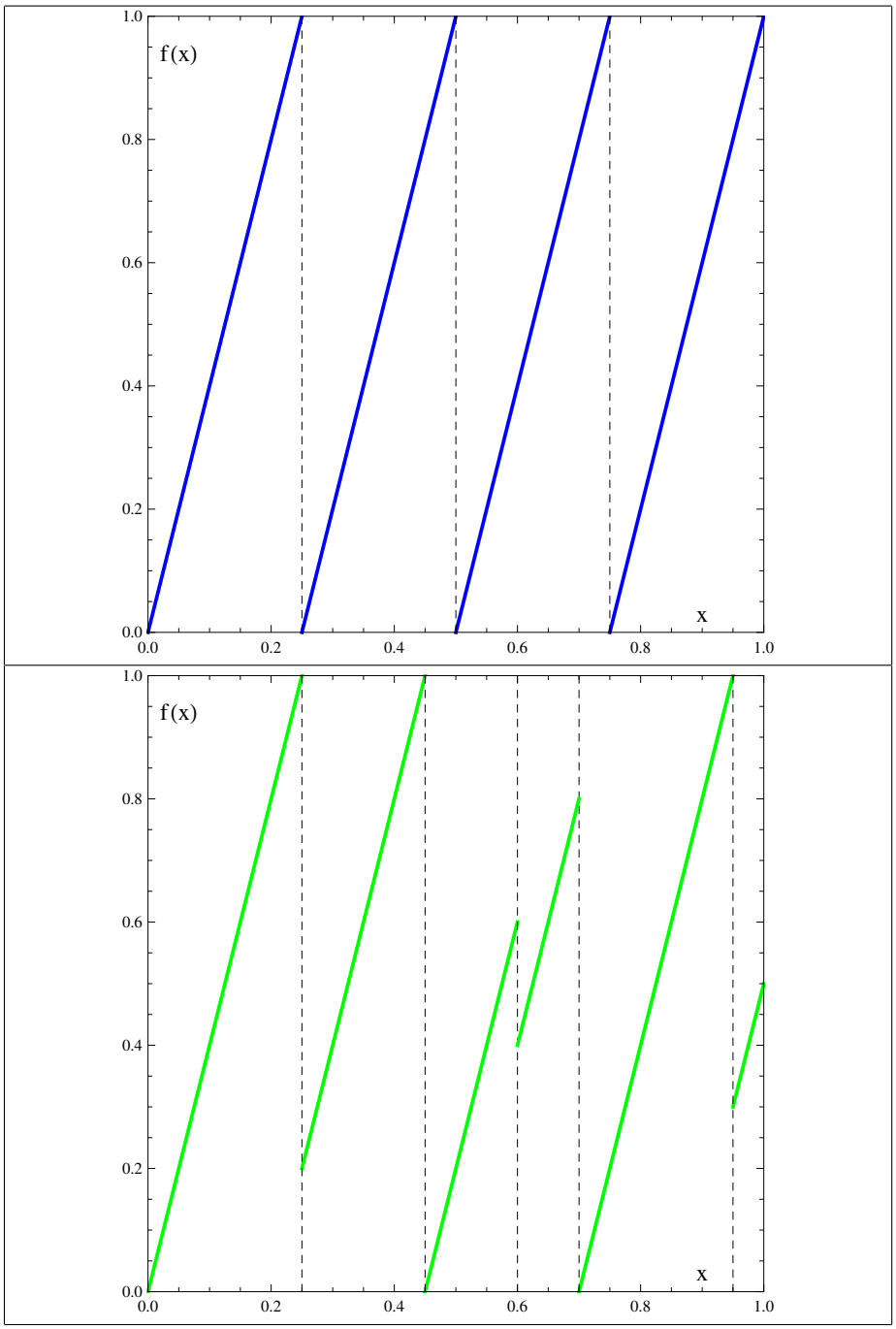


Fig. 1. Graphics of discontinuous piecewise linear 1D map $f(x) = 4x + a_i$, $x \in [0, 1]$.

where $|I| = 1$ represents the amplitude of the interval I and $\bar{\mu}$ is a measurable and ergodic invariant natural measure. The transversal Lyapunov exponent is given by,

$$\lambda_\perp = \int_I \ln|\lambda_2|\, d\bar{\mu} = \ln|s(1 - N\sigma)|\,. \tag{10}$$

Notice that for each complete dynamical network K_N, there is a single transversal Lyapunov exponent.

The following proposition stablishes that to stabilize the synchronized states, it suffices to require that the transversal Lyapunov exponent is negative, with piecewise linear maps f with slope $s > 1$ as local chaotic dynamics.

Proposition 1. *Consider the (K_N, Σ^+) space of complete dynamical networks, given by Eq.(2). Let f be the discontinuous piecewise linear map with slope $s > 1$ everywhere, given by Eq.(6), I_σ be the synchronization interval, given by Eq.(8), and $I_{\lambda_\perp^-}$ be the interval where $\lambda_\perp < 0$, with $\lambda_\perp$ given by Eq.(10). $\forall s > 1$, it is verified that:*

(i) $I_\sigma \equiv I_{\lambda_\perp^-} \neq \emptyset$;

(ii) there exists $\sigma > 0$ such that the synchronized states of Eq.(2) stabilize exponentially, i.e., $x_1(k) = x_2(k) = \ldots = x_N(k) \to s(k)$, as $k \to \infty$.

Proof. Consider $(N, s, \sigma) \in \Sigma^+$ and $\lambda_\perp \in I_{\lambda_\perp^-}$, according to Eq.(10) we have that,

$$\ln|s(1 - N\sigma)| < 0 \Leftrightarrow |s(1 - N\sigma)| < 1 \Leftrightarrow$$
$$-1 \leq s(1 - N\sigma) < 1 \Leftrightarrow \frac{s-1}{Ns} < \sigma < \frac{s+1}{Ns},\ \forall s > 1.$$

Thus, from Property 1 the result of item (i) is proved, see Fig.2.

Given that the chaoticity of the node map f is measured by $\chi(f) = \log(s) > 0$, $\forall s > 1$, from condition given by Eq.(5), it follows that the inequality

$$\frac{1 - e^{-\chi(f)}}{|\mu_2|} < \frac{1 + e^{-\chi(f)}}{|\mu_N|} \Leftrightarrow \frac{s-1}{Ns} < \frac{s+1}{Ns}$$

should be satisfied for the existence of a coupling strenght $\sigma > 0$. In this context can be defined a ratio $\frac{1}{R} := \frac{\mu_2 - \mu_N}{\mu_1 - \mu_2}$, where $0 = \mu_1 < |\mu_2| \leq \ldots \leq |\mu_N|$ are the eigenvalues of the laplacian matrix L, which measures the distance from the first eigenvalue to the main part of the spectral density $\rho(\mu)$ normalized by the extension of the main part, see [8]. From straightforward and simple calculations it is proved that,

$$\frac{2e^{-\chi(f)}}{1 - e^{-\chi(f)}} > \frac{1}{R} \Leftrightarrow \frac{2}{s-1} > 0,\ \forall s > 1.$$

Thus, considering the previous conditions, we can state that for all $s > 1$ there exists a coupling strenght $\sigma > 0$ such that the synchronized states of Eq.(2) stabilize exponentially. Proposition 1 is thus proved. ■

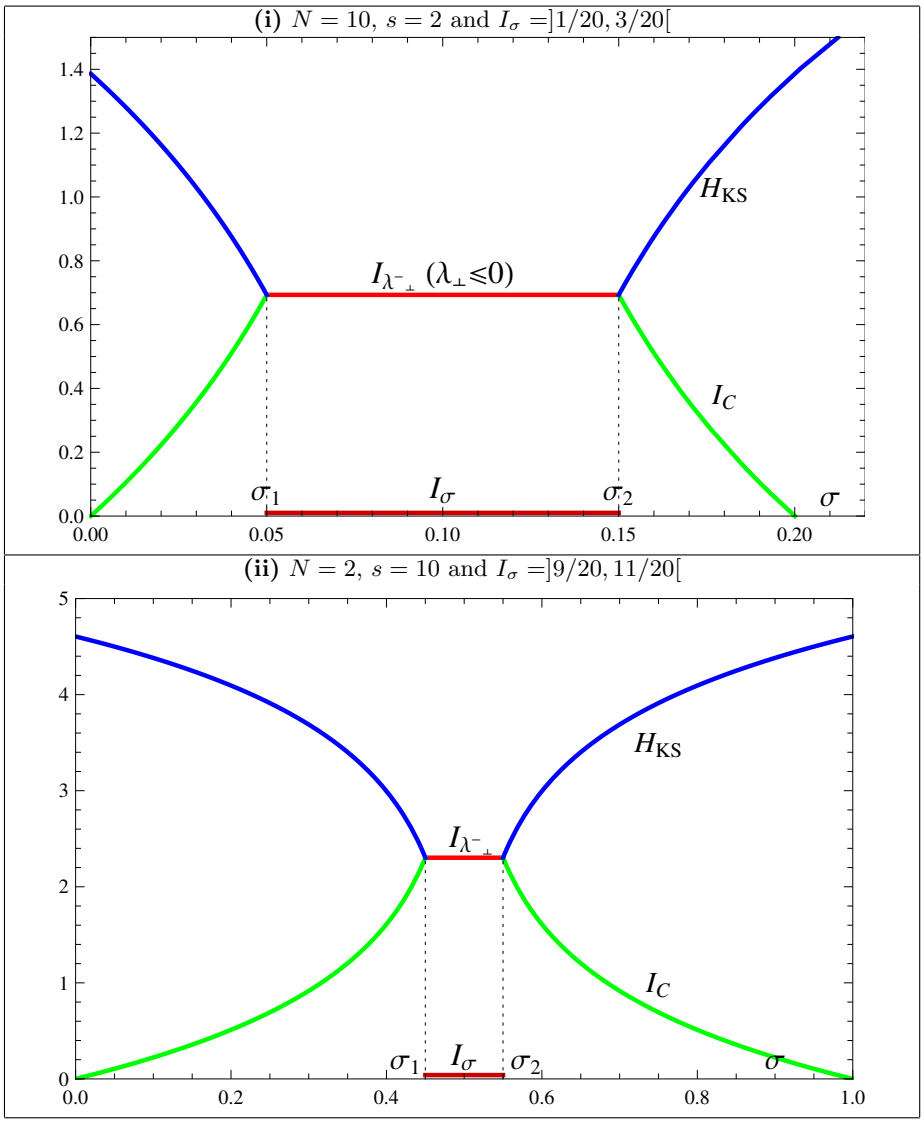


Fig. 2. Numerical simulation for Propositions 1 and 2, where to stabilize the synchronized states of the dynamical network, given by Eq.(2), it suffices to require that all transversal Lyapunov exponents be negative.

The results in Proposition 1 bring up to the discussion the complete synchronization versus the negativity of the conditional or transversal Lyapunov exponents. For more details on this classic discussion, see, for example, [6], [7], [8], [12], [15] and references therein. The negativity of the conditional Lyapunov exponents is a necessary condition for the stability of the synchronized state, see also [4]. To illustrate these results see the numerical cases shown in Fig.2.

Taking into account the expressions of the parallel Lyapunov exponent and the transversal Lyapunov exponent, given by Eqs.(9) and (10), respectively, the information measures defined by Eqs.(3) and (4) are explicitly written by the following expressions:

$$I_C = \begin{cases} \ln\left(\frac{1}{|1-N\sigma|}\right), & \text{if } \lambda_\perp > 0 \\ \ln(s), & \text{if } \lambda_\perp \leq 0 \end{cases} \tag{11}$$

and

$$H_{KS} = \begin{cases} \ln\left(s^2|1-N\sigma|\right), & \text{if } \lambda_\perp > 0 \\ \ln(s), & \text{if } \lambda_\perp \leq 0 \end{cases}. \tag{12}$$

The next proposition establishes some properties of the mutual information rate and the Kolmogorov-Sinai entropy, depending on the synchronization interval I_σ.

Proposition 2. *Consider the (K_N, Σ^+) space of complete dynamical networks, given by Eq.(2). Let f be the discontinuous piecewise linear map with slope $s > 1$ everywhere, given by Eq.(6), I_σ be the synchronization interval, given by Eq.(8), and $I_{\lambda_\perp^-}$ be the interval where $\lambda_\perp < 0$, with $\lambda_\perp$ given by Eq.(10). $\forall s > 1$, it is verified that:*

(i) if $\sigma \in I_\sigma$, then $I_C = H_{KS}$;
(ii) if $\sigma \notin I_\sigma$ and $\sigma < \sigma_1$, then I_C increases and H_{KS} decreases;
(iii) if $\sigma \notin I_\sigma$ and $\sigma > \sigma_2$, then I_C decreases and H_{KS} increases.

Proof. Considering the definitions of I_C and H_{KS}, given by Eqs.(11) and (12), respectively, it is verified that $I_C = H_{KS}$ if and only if $\lambda_\perp \leq 0$. On one hand, we have proved, in Proposition 1 (i), that $I_\sigma \equiv I_{\lambda_\perp^-} \neq \emptyset$, $\forall s > 1$. Thus, item (i) is proved.

The result of item (ii) is a consequence of the previous argumentation, i.e., if $\sigma \notin I_\sigma$ and $\sigma < \sigma_1$, then follows that $I_C \neq H_{KS}$, $\forall s > 1$. In this region, considering the monotony of logarithmic function, we also have that $I_C = \ln\left(\frac{1}{|1-N\sigma|}\right)$ increases, for all coupling strenght $\sigma > 0$ under the conditions required by hypothesis, and $H_{KS} = \ln\left(s^2|1-N\sigma|\right)$ decreases, for all σ under the same conditions. This proves the claim (ii).

The proof of item (iii) is similar to the proof of item (ii), the monotony of I_C and H_{KS} follows as discussed above, $\forall \sigma \notin I_\sigma$ and $\sigma > \sigma_2$. See also Fig.2. This completes the proof of Proposition 2. ■

Under the conditions of Proposition 2, we can establish that $I_C \leq H_{KS}$, $\forall \sigma > 0$, see also [2].

3.2 Topological order and information measures

To end this section we will approach a topological invariant associated with the dynamics between the nodes of the complete dynamical network K_N: the network topological entropy. In this work we use the network topological entropy concept used in [13] and [14]. Let G be the connected and unoriented graph associated to the complete dynamical network K_N of order $N \in \mathbb{N} \setminus \{1\}$. Considering that the adjacency matrix A of the complete dynamical network K_N is irreducible, then the Perron-Frobenius Theorem states that the network topological entropy of K_N is given by,

$$h_{top}(K_N) = h_{top}(G) = \ln(\lambda_A) = \ln(N-1), \tag{13}$$

where λ_A is the Perron eigenvalue of A. Clearly, the topological entropy of a complete dynamical network K_N is characterized by its order N.

Proposition 3. *Let (K_N, Σ^+) be the space of complete dynamical networks, given by Eq.(2), with a fixed local dynamics given by a discontinuous piecewise linear map f with slope $s > 1$ everywhere, given by Eq.(6). It is verified that if the network topological entropy $h_{top}(K_N)$ increases, such that $\sigma \notin I_\sigma$, then:*

(i) the mutual information rate I_C increases;
(ii) the Kolmogorov-Sinai entropy H_{KS} decreases.

Proof. Consider that the local dynamics of two complete dynamical networks K_N and K_{N+1} are fixed, i.e., $\chi(f) = \log(s) > 0$ is constant, $\forall s > 1$. Considering the definition of the mutual information rate I_C, given by Eq.(11), with $\sigma \notin I_\sigma$, let $u_N = \ln\left(\frac{1}{|1-N\sigma|}\right)$, with $N \in \mathbb{N} \setminus \{1\}$. The following statements holds,

$$u_{N+1} - u_N = \ln\left(\frac{1}{|1-(N+1)\sigma|}\right) - \ln\left(\frac{1}{|1-N\sigma|}\right) = \ln\left|\frac{1-N\sigma}{1-(N+1)\sigma}\right| > 0.$$

This means that the mutual information rate I_C increases, if the network topological entropy $h_{top}(K_N)$ increases. Now under the same assumptions, if we consider the definition of the Kolmogorov-Sinai entropy H_{KS}, given by Eq.(12) with $\sigma \notin I_\sigma$ and let $v_N = \ln\left(s^2|1-N\sigma|\right)$, with $N \in \mathbb{N} \setminus \{1\}$. It can be easily verified that,

$$v_{N+1} - v_N = \ln\left(s^2|1-(N+1)\sigma|\right) - \ln\left(s^2|1-N\sigma|\right) = \ln\left|\frac{1-(N+1)\sigma}{1-N\sigma}\right| < 0.$$

This proves that the Kolmogorov-Sinai entropy H_{KS} decreases, when the network topological entropy $h_{top}(K_N)$ increases. Thus, the desired results are proved. ■

It is interesting to note that this result is related with the structural complexity of the dynamical network analyzed, this means that as larger as the network topology is, measure by its network topological entropy $h_{top}(K_N)$ and directly related with the network order $N \in \mathbb{N} \setminus \{1\}$, the mutual information rate I_C and the Kolmogorov-Sinai entropy H_{KS} increases or decreases, according to order N in relation to $\sigma \notin I_\sigma$.

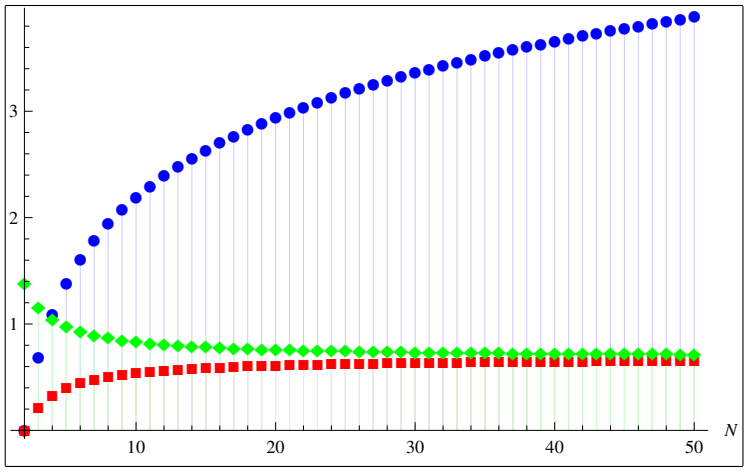


Fig. 3. Numerical simulation for Proposition 3: blue points are $h_{top}(K_N)$, red squares are I_C, green lozenges are H_{KS}, where it is considered $s = 2$ and $N \in \{2, 3, \ldots, 50\}$.

Remark 1. Under the conditions of Proposition 3 and according to Property 1, we can establish that the amplitude of the synchronization interval $|I_\sigma|$ decrease, as larger is the network topological entropy $h_{top}(K_N)$ of the complete dynamical network.

4 Conclusion and Discussion

In this paper we have considered the space of complete dynamical networks K_N, an extreme case of a ring lattice with maximal degree, of identical chaotic dynamical oscillators or nodes, in which each oscillator is coupled linearly and symmetrically with their neighbours, see Eq.(2). The topology of the networks K_N is characterized by a circulant matrix of order $N \in \mathbb{N} \setminus \{1\}$, which provides a certain regularity in the network. Furthermore, we consider discontinuous local dynamics: piecewise linear maps f with positive slope $s > 1$. The chaoticity of the local dynamics is measured by the topological entropy of f, i.e., $h_{top}(f) = \log |s|$. Several measures have been considered in the field of information theory. In our paper we have considered the mutual information rate and the Kolmogorov-Sinai entropy. However, the synchronization is vital for modern methods of digital communication that rely on the synchronous operation of many subsystems. So, our main concern was to determine explicit expressions for these measures, properties between them and relations with the synchronization interval I_σ. The results presented in Propositions 1 and 2 fully serve this purpose. Finally, in Proposition 3 it is addressed the structural complexity of the complete dynamical networks K_N, using the network topological entropy $h_{top}(K_N)$. We proved that as large as the network topological entropy $h_{top}(K_N)$ is, the mutual information rate I_C and the Kolmogorov-Sinai en-

tropy H_{KS} increase or decrease, according to the network order N in relation to $\sigma \notin I_\sigma$. Therefore, a topological order was established regarding the studied information measures.

Clearly, with the fixed coupling topology in the networks K_N and variation of the discontinuous local dynamics f in the nodes, the complexity of our analysis increases due to measure theory issues. For future work, on one hand we would like to generalize our study for the case where $|s| > 1$ dependening on the amplitudes of the subintervals with slope $s > 1$ and slope $s < -1$. On the other hand, we would like to analyze this problem in other networks of regular lattices rings, starting with the minimal degree, the cycles of order N. In this context, we finish this work with some open questions: There are sufficient conditions to guarantee the negativity of the conditional Lyapunov exponents, for different slopes of f? Under what conditions the chaotic signals transmitted through filters produce an output with higher dimension, due to the appearance of a fractal set?

Acknowledgments

Research funded by the project IPL - MISRedes, IDI&CA 2019, FCT - Fundação para a Ciência e a Tecnologia, Portugal, through the projects UIDB/00006/2020 (CEAUL), UIDB/04721/2020 (CEAFEL) and ISEL.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. V. Avrutin, L. Gardini, I. Sushko and F. Tramontana. *Continuous and Discontinuous Piecewise-Smooth One-Dimensional Maps: Invariant Sets and Bifurcation Structures*, World Scientific, Singapore, 2019.
2. M. S. Baptista and J. Kurths. Chaotic channel, *Phys. Rev. E*, **72**, 045202R, 2005.
3. M. S. Baptista, R. M. Szmoski, R. F. Pereira and S. E. Souza Pinto. Chaotic, informational and synchronous behaviour of multiplex networks, *Sci. Rep.*, **6 (22617)**, 1–9, 2016.
4. S. Boccaletti, J. Kurths, G. Osipov, D. L. Valladares and C. S. Zhou. The synchronization of chaotic systems, *Phys. Rep.*, **366**, 1–101, 2002.
5. A. Caneco, J. L. Rocha and C. Grácio. Topological entropy in the synchronization of piecewise linear and monotone maps. Coupled Duffing oscillators, *Int. J. Bifurc. Chaos*, **19 (11)**, 3855–3868, 2009.
6. A. Caneco and J. L. Rocha. Synchronization and information transmission in networks, *in Proceedings of ECIT 2012 (eds. W. Jarczyk, et al.), EDP Sciences, ESAIM: Proceedings and Surveys*, **46**, 111–124, 2014.
7. J. Cao and J. Lu. Adaptive synchronization of neural networks with or without time-varying delays, *Chaos*, **16**, 013133, 2006.
8. X. Li and G. Chen. Synchronization and desynchronization of complex dynamical network: an engineering viewpoint, *IEEE Trans. Circuits Syst. I*, **50 (11)**, 1381–1390, 2003.

9. A. C. J. Luo. *Discontinuous Dynamical Systems*, Springer, New York, 2012.
10. A. C. J. Luo. Discontinuous dynamical systems and synchronization, *Eur. Phys. J. Spec. Top.*, **228 (6)**, 1383–1384, 2019.
11. J. Milnor and W. Thurston. On iterated maps of the interval, *Lect. Notes Math.*, Springer-Verlag, **1342**, 465-563, 1988.
12. L. M. Pecora and T. L. Carroll. Driving systems with chaotic signals, *Phys. Rev. A*, **44**, 2374–2383, 1991.
13. J. L. Rocha and A. Caneco. Mutual information rate and topological order in networks, *Chaotic Modeling and Simulation, Int. J. Nonlinear Sci.*, **4**, 553–562, 2013.
14. J. L. Rocha, C. Grácio, S. Fernandes and A. Caneco. Spectral and dynamical invariants in a complete clustered network, *Applied Mathematics & Information Sciences*, **9 (5)**, 2367–2376, 2015.
15. J. W. Shuai, K. W. Wong and L. M. Cheng. Synchronization of spatiotemporal chaos with positive conditional Lyapunov exponents, *Phys. Rev. E*, **56**, 2272, 1997.
16. F. Tramontana and F. Westerhoff. Piecewise-linear maps and their application to financial markets, *Front. Appl. Math. Stat.*, **2 (10)**, 1–10, 2016.

How chaotic dynamics drive a vintage grill-room spit

Julio Rodriguez[1] and Max-Olivier Hongler[2]

[1] Swissgrid AG Bleichemattstrasse 31, 5001 Aarau, Switzerland
(E-mail: julio.rodriguez@swissgrid.ch)
[2] Ecole polytechnique fédérale de Lausanne, STI-DO, BM 5132, Station 17, 1015 Lausanne, Switzerland
(E-mail: max.hongler@epfl.ch)

Abstract. In 1943 Yves Rocard published an amazing book entitled "Dynamique Générale des Vibrations". Among a collection of mechanical devices that are studied with great care in the book, the so-called Bouasse-Sarda's *"tournebroche"* (rotary grill spit) is a fascinating two degrees of freedom device. What makes it interesting is that it stylises a wide class of forced parametric oscillators exhibiting a rich range of complex dynamical behaviours including fully deterministic chaos, a dynamic concept not popular in the forties and hence not yet discussed in the Y. Rocard's book. Today's numerical tools offer a new possibility to revisit this amazingly simple but rich dynamical system.
Keywords: Nonlinear coupled mechanical systems, Stability analysis, "Washboard"-like dynamics, Bistability, Chaotic dynamics, Feigenbaum cascade, Period–doubling bifurcations, Antimonotonicity.

Highlights: In the present paper, we propose an extensive, didactical and updated investigation of the nonlinear dynamical system studied by Y. Rocard more than 75 years ago.

1 Preliminary

Our study starts in the Spring of 2006 on a sunny Saturday, when one of the authors (MOH) rambled through Geneva's flea market. Among a chaotic pile of secondhand books, suddenly a well-stocked pile of scientific and engineering publications stuck out. The bookseller was so happy to get rid of such "boring material" that the whole stock was purchased for virtually nothing. Among them, "Optique Géométrique Supérieure" by H. Bouasse (1917), a version of the "Vibration Problems in Engineering" by S. Timoschenko (1937) and a remarkable contribution entitled "Dynamique Générale des Vibrations" by Y. Rocard published in Laval (Mayenne, France) 1943, during the French occupation. In Rocard's opus, we discovered a truly intriguing two degrees of freedom device called the *"tournebroche"* (i.e. rotating spit) de Bouasse-Sarda. More than 75 years after Rocard's publication, we revisit this pretty fascinating *"tournebroche"* dynamics and show how it naturally offers the possibility to exhibit chaotic evolution.

2 Introduction

The dynamic response of mechanical devices driven by external energy sources is often discussed under the implicit assumtion that the energy source itself is fully insensitive to behaviour of the driven device. It is assumed that the energy source is large enough to ensure that all feedback response of the driven system to the source are negligible. While such approximation is fully legitimate when a clear energy-scale distinction between the system itself and its driver can be made[1], in many other circumstances the system-driver's retro-actions may generate complex and unsuspected responses that cannot be ignored. The suggestive Bouasse-Sarda (B-S) *"tournebroche"* device sketched in Figure 1 which is also called the B-S regulator offers a rich and intuitive illustration of the underlying complexity of the clear and consistent distinction between a system and its driving unit. From Figure 1, we may view the spring as the driven system and the mass m together with the pulley as playing the role of an energy source driving the oscillatory mechanism. Depending on the size of the crankshaft a in the B-S device, there is a simple way to continuously increase the spring-pulley retro-action and hence infer the resulting evolution. As we shall discuss, this a-priori deceptively simple system uncovers a very rich variety of dynamical behaviours which include limit cycles, bistability, the now famous Feigenbaum cascade with period-doubling bifurcations ultimately leading to a fully developed deterministic chaos and antimonotonicity.

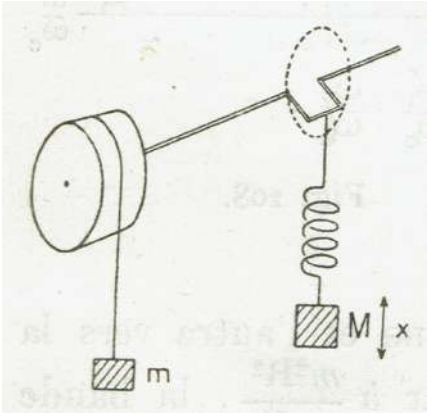


Fig. 1. The Bouasse-Sarda's rotary spit drawn by Y. Rocard in his 1943 book entitled: *"Dynamique Générale des Vibrations"*.

[1] As a paradigmatic illustration, you may invoke the thermodynamic model of someone diving into a pool: the diver's body temperature is lowered by the pool's environment, yet one usually neglects the temperature increase of the pool's content. From a mechanical perspective, this idea of the interaction between a system and the force acting on it, is studied in Roseau[17], and as an illustration to the theoretical framework, the B-S regulator is used.

3 Equations of the motion

The dynamical system sketched in Figure 1 consists of a shaft or a rotating spit with, on one side, a crank of radius a and, on the other side of the shaft, a drum of radius ρ with inertial moment I. A cable is fastened around the drum on which suspends a mass m. A mass M is attached to a spring with stiffness K. The spring is then fixed to the crank. For this two degrees of freedom device $(\theta(t), x(t)) \in \mathbb{R}^2$, the Lagrangian $\mathcal{L} = T(\dot{\theta}, \dot{x}) - V(\theta, x)$ is given by:

$$\begin{aligned} T(\dot{\theta}, \dot{x}) &= (\mathsf{I} + m\rho^2)\frac{\dot{\theta}^2}{2} + \mathsf{M}\frac{\dot{x}^2}{2}, \\ V(\theta, x) &= -(mg\rho\theta + \mathsf{M}gx) + \frac{\mathsf{K}(x - a\sin(\theta))^2}{2}, \end{aligned} \tag{1}$$

leading to the time evolution[2]:

$$\begin{cases} (\mathsf{I} + m\rho^2)\ddot{\theta} + h\dot{\theta} = mg\rho + \mathsf{K}(x - a\sin(\theta))a\cos(\theta), \\ \mathsf{M}\ddot{x} + f\dot{x} = \mathsf{M}g - \mathsf{K}(x - a\sin(\theta)). \end{cases} \tag{2}$$

Eq.(2) involves a set of nine positive control parameters $(\mathsf{I}, m, \mathsf{M}, \rho, \mathsf{K}, f, h, g, a)$, with gravitational acceleration g. Energy dissipation is implemented via a couple of viscous damping mechanisms adjusted by the parameters $f \geq 0$ and $h \geq 0$. We emphasise that in Y. Rocard's book Rocard[15], the dynamical system involves only friction on the spring (i.e. $h = 0$). The second equation in Eq.(2) is simply a $\theta(t)$ forced harmonic oscillator for which the response is explicitly calculable provided that signal $\theta(t)$ itself is given. In particular, for small damping parameter f, one expects a resonance peak at frequency $\omega_0 := \sqrt{\mathsf{K}/\mathsf{M}}$, which arises whenever ω_0 belongs to the Fourier spectrum of the signal $\theta(t)$.

To proceed, we rewrite Eq.(2) in terms of the (dimensionless) coordinate $x(t) = au(t) + (\mathsf{M}g/\mathsf{K})$ and obtain:

$$\begin{cases} \mathsf{J}\ddot{\theta} + h\dot{\theta} = \underbrace{g(\mathsf{M}a\cos(\theta) + m\rho)}_{=:H_a^{(1)}(\theta)} + \underbrace{\mathsf{K}a^2(u - \sin(\theta))\cos(\theta),}_{=:H_a^{(2)}(\theta,u)} \\ \mathsf{M}\ddot{u} + f\dot{u} = -\mathsf{K}(u - \sin(\theta)), \end{cases} \tag{3}$$

with $\mathsf{J} := \mathsf{I} + m\rho^2$.

Note that for vanishing crankshaft (i.e. $a = 0$), the system Eq.(3) degenerates into a couple of independent one degree of freedom dynamics. For a small crankshaft $a \gtrsim 0$ (i.e. up to first order in a and thus omitting the $H_a^{(2)}(\theta, u)$ term), the (θ, u) variables are partially coupled. Indeed, we have a special two degrees of freedom system for which the $\theta(t)$ evolution can be calculated independently of $u(t)$. Observe that up to first order in a, the $\theta(t)$

[2] For detailed calculations, see Appendices 7.1 and 7.2.

evolution describes either the dynamics of a point particle evolving inside a tilted washboard potential or, alternatively, an anharmonic pendulum subject to a constant external torque. Finally, for general a, the action of the feedback type term $H_a^{(2)}(\theta, u)$ drastically modifies the evolution. The u variable now affects the driving mechanism $\theta(t)$ itself in Eq.(3).

4 Study of the dynamics

4.1 Static equilibrium and associated stability issues

The fixed points $\mathcal{P}^* := (\theta^*, u^*)$ of the dynamics Eq.(3) read:

$$\begin{aligned}(\theta^*, u^*) &= \big(\cos^{-1}(-\tfrac{m\rho}{\mathsf{M}a}), \sin(\cos^{-1}(-\tfrac{m\rho}{\mathsf{M}a}))\big),\\ &= \big(\cos^{-1}(-\tfrac{m\rho}{\mathsf{M}a}), \sqrt{1-(\tfrac{m\rho}{\mathsf{M}a})^2}\big),\end{aligned} \tag{4}$$

and hence $\mathsf{M}a \geqslant m\rho$ is required for the existence of $\mathcal{P}^*$. Observe that $\mathcal{P}^*$ does not depend on g.

Stability of $\mathcal{P}^*$: The stability of $\mathcal{P}^*$ is studied via the linearisation of the dynamics in the $\mathcal{P}^*$-neighbourhood. By adding a couple of conjugate variables $(\eta, v) := (\dot{\theta}, \dot{u})$, we rewrite the coupled second-order differential system Eq.(3) as an equivalent set of four first-order differential equations, and hence the first variational equation is:

$$\begin{pmatrix}\dot{\epsilon_\theta}\\ \dot{\epsilon_\eta}\\ \dot{\epsilon_u}\\ \dot{\epsilon_v}\end{pmatrix} = \underbrace{\begin{pmatrix}0 & 1 & 0 & 0\\ -\mathsf{J}^{-1}(\mathsf{M}gas^* + \mathsf{K}(ac^*)^2) & -\mathsf{J}^{-1}h & \mathsf{J}^{-1}\mathsf{K}a^2c^* & 0\\ 0 & 0 & 0 & 1\\ \mathsf{M}^{-1}\mathsf{K}c^* & 0 & -\mathsf{M}^{-1}\mathsf{K} & -\mathsf{M}^{-1}f\end{pmatrix}}_{:=DF(\theta^*,u^*)} \begin{pmatrix}\epsilon_\theta\\ \epsilon_\eta\\ \epsilon_u\\ \epsilon_v\end{pmatrix} \tag{5}$$

with $c^* := \cos(\theta^*)$ and $s^* := \sin(\theta^*)$ and perturbations $\epsilon_\theta, \epsilon_\eta, \epsilon_u, \epsilon_v$ around $(\theta^*, 0, u^*, 0)$. The characteristic polynomial $P(\lambda)$ relevant to infer the stability property of $\mathcal{P}^*$ is given by:

$$P(\lambda) := \mathrm{Det}\left[DF(\theta^*, u^*) - \lambda \mathbb{I}_4\right] = 0, \tag{6}$$

with $\mathbb{I}_4$ standing for the (4×4) identity matrix. A direct calculation yields:

$$\begin{aligned}P(\lambda) = \lambda^4 + (\mathsf{J}^{-1}h + \mathsf{M}^{-1}f)\lambda^3 +\\ (\mathsf{J}^{-1}\mathsf{M}^{-1}hf + \mathsf{M}^{-1}\mathsf{K} - DF_{2,1}(\theta^*, u^*))\lambda^2 +\\ (\mathsf{J}^{-1}\mathsf{M}^{-1}h\mathsf{K} - \mathsf{M}^{-1}fDF_{2,1}(\theta^*, u^*))\lambda +\\ \mathsf{J}^{-1}\mathsf{K}ga\sin(\theta^*).\end{aligned}$$

Note that all coefficients of $P(\lambda)$ are positive. Provided we have:

$$(\mathsf{J}^{-1}h + \mathsf{M}^{-1}f) > 0,$$

direct application of the Routh-Hurwitz criterion shows that all real parts of the eigenvalues of the linearised dynamics around $\mathcal{P}^*$ are negative and hence

$\mathcal{P}^*$ is stable.

In summary, when $\mathsf{M}a \geqslant m\rho$, a fixed point $\mathcal{P}^*$ given by Eq.(4) exists and its stability (i.e. stable focus) is ensured by the viscous dampers (and this even if only a single damper is active).

4.2 Dissipation free dynamics - Harmonic undamped oscillations

In the absence of dissipation (i.e. $f = h = 0$), the characteristic polynomial $P(\lambda)$ is bi-quadratic, yielding a couple of eigenfrequencies (ν_-, ν_+):

$$\begin{cases} \nu_\pm = \frac{1}{2}\left[r_1 \pm \sqrt{(r_1^2 - 4r_2)}\right], \\ r_1 = \mathsf{M}^{-1}\mathsf{K} - DF_{2,1}(\theta^*, u^*), \\ r_2 = \mathsf{J}^{-1}\mathsf{K}ga\sin(\theta^*). \end{cases} \tag{7}$$

and the fixed point $\mathcal{P}^*$ is a center. For large inertial moment J (i.e. $\mathsf{J} >> \mathsf{M}$ and $\mathsf{J} >> \mathsf{K}$), one verifies that $r_2 \to 0$, and we obtain in this limiting case an oscillation with frequency $\nu_+ \cong \sqrt{\mathsf{K}/\mathsf{M}}$ of M, together with a very low frequency $\nu_- \cong 0$ oscillation of m around the stationary point $\mathcal{P}^*$ [3].

4.3 "Washboard"-like dynamics

For a small crankshaft parameter a, we may approximately neglect $H_a^{(2)}(\theta, u)$ in Eq.(3), which is of order a^2. In this subsection, $h > 0$. The resulting dynamics for the variable $\theta(t)$ describes a classical particle evolving with friction inside a tilted washboard potential. Alternatively, it can also be viewed as an anharmonic pendulum subject to constant external torque Coullet *et al.*[4]:

$$\ddot{\theta}(t) + \frac{h}{\mathsf{J}}\dot{\theta}(t) = \frac{\mathsf{M}ag}{\mathsf{J}}\cos(\theta(t)) + \frac{mg\rho}{\mathsf{J}}. \tag{8}$$

The time rescaling $\tau(t) := \sqrt{\frac{\mathsf{M}ag}{\mathsf{J}}}t$ together with the phase shift $\phi(\tau(t)) := \theta(t) - \frac{\pi}{2}$ enable us to write:

$$\phi'' + \beta\phi' + \sin(\phi) = \gamma, \tag{9}$$

where $'$ now stands for the τ-derivative and $\beta := \frac{h}{\sqrt{\mathsf{JM}ag}}$ and $\gamma := \frac{m\rho}{\mathsf{M}a}$. The dynamical system given by Eq.(9) is currently discussed in the context of Josephson's junctions dynamics[4].

[3] Remember that the linearising assumption is valid only for motions close to $\mathcal{P}^* = (\theta^*, u^*)$. For arbitrary $\mathsf{J}, \mathsf{K}, \mathsf{M}, m$, the resulting 2-D harmonic system may enter into resonance, thus precluding the possibility to use linear approximations.

[4] See, for example, the recent review Blackburn *et al.*[1].

In particular, the trade-off between the damping factor β and the external torque γ is summarised in Figure 2 [5]. Note that in Eq.(9), one has explicitally $\gamma = \frac{m\rho}{\mathrm{M}a}$ and this is coherent with the condition for the existence of $\mathcal{P}^*$) (i.e. $1 > \frac{m\rho}{\mathrm{M}a}$). Hence, when $\gamma > 1$, only periodic solutions exists, consistent with the fact that are no fixed points.

We emphasise that since Eq.(9) describes a single degree of freedom system, then up to first order in a (i.e. small crankshafts), no chaotic behaviour can possibly be observed for the $\theta(t)$-evolution. As a consequence, the $u(t)$-motion itself, which describes the evolution of a forced harmonic oscillator with damping, does not exhibit a chaotic behaviour either. This follows from the fact that the $\theta(t)$ motion results from simple quadratures[6]. Summarising, for small crankshafts a, the two degrees of freedom (θ, u) system does not exhibit chaotic behaviour. As a consequence, **up to first order in a, the largest Lyapunov exponent always remains negative**.

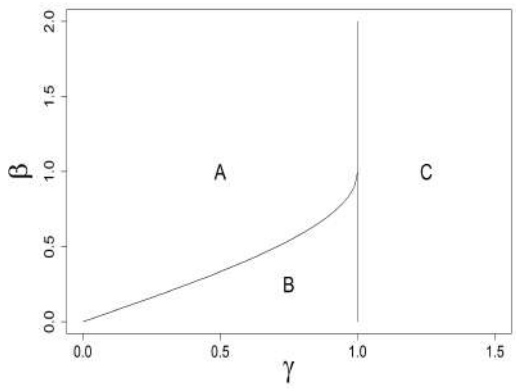


Fig. 2. Evolution regimes as a function of the damping β and the external torque γ: A = only the stationary solution is stable, B = both the stationary and the periodic solution are stable, C = only the periodic solution is stable. This diagram is reproduced from Coullet *et al.*[4].

4.4 The Rocard's quasi-uniform regime

We now focus on the general situation (i.e. also considering the influence of the retro-action $H_a^{(2)}(\theta, u)$). We follow the lines originally drawn by Y. Rocard himself and note that this quasi-uniform regime has been later re-discussed in Panovko and Gubanova[14] for $h = 0$ and in Blekhman and Dzhanelidze[2], Colombo[3], Roseau[17], Fridman[6], Fridman[7] for $h > 0$. Rocard assumed

[5] Figure 2 is directly reproduced from Figure 6 in Coullet *et al.*[4].
[6] See, for example, chapter 2 in Lakshmanan and Rajasekar[11].

that $\theta(t) \cong \omega t$, and so the $u(t)$-motion follows the permanent regime of an harmonically driven linear oscillator, namely:

$$u(t) = \frac{\frac{\mathsf{K}}{\mathsf{M}}\sin(\omega t+\varphi)}{\sqrt{(\frac{\mathsf{K}}{\mathsf{M}}-\omega^2)^2+(\frac{f}{\mathsf{M}})^2\omega^2}} =: \mathsf{A}\sin(\omega t+\varphi), \tag{10}$$

with $\tan(\varphi) = \frac{-\frac{f}{\mathsf{M}}\omega}{\frac{\mathsf{K}}{\mathsf{M}}-\omega^2}$. The pulsation[7] ω (i.e. the dominant frequency at which both drum and spring oscillate) is determined by solving the polynomial equation explicitly given by:

$$\frac{2\mathsf{M}^2(mg\rho - h\omega)}{f(\mathsf{K}a)^2} = \frac{\omega}{(\frac{\mathsf{K}}{\mathsf{M}}-\omega^2)^2+(\frac{f}{\mathsf{M}})^2\omega^2}. \tag{11}$$

For the sake of completeness, Eq.(11) is re-derived in Appendix 7.3. The solutions to Eq.(11) are the roots of a polynomial of degree 5 (for $h > 0$) or degree 4 (for $h = 0$), namely

$$\begin{aligned} &\text{for } h>0: \ 0 = h\omega^5 - mg\rho\omega^4 + h(p^2-2q)\omega^3 - (p^2-2q)mg\rho\omega^2 \\ &\qquad\qquad +(q^2(h+\tfrac{a^2 f}{2}))\omega - q^2 mg\rho, \\ &\text{and} \\ &\text{for } h=0: \ 0 = \omega^4 + (p^2-2q)\omega^2 - \tfrac{q^2a^2f}{2mg\rho}\omega + q^2, \end{aligned} \tag{12}$$

with $p = \frac{f}{\mathsf{M}}$ and $q = \frac{\mathsf{K}}{\mathsf{M}}$. Depending on the value of $a \geqslant 0$, Eq.(11) may have up to three real solutions for $h > 0$, or two for $h = 0$. To see this, start with the RHS of Eq.(11) and define the function

$$\alpha(\omega) := \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2}\Big(\frac{\omega}{(\frac{\mathsf{K}}{\mathsf{M}}-\omega^2)^2+(\frac{f}{\mathsf{M}})^2\omega^2}\Big) \tag{13}$$

for positive ω. As such, the function α is positive and exhibits a resonance-like curve shape: it takes on the value zero at $\omega = 0$, increases to its maximum value at $\omega = \omega_0$ (in the neighbourhood of $\sqrt{\frac{\mathsf{K}}{\mathsf{M}}}$), and then, as ω increases further, converges asymptotically to zero. The solutions of Eq.(11) are the intersections between the curve given by α and the line $mg\rho - h\omega$ (i.e. a decreasing line for $h > 0$ or a straight line, parallel to the ω-axis, for $h = 0$). See, for example, Figure 3 (a) (for $h > 0$) and Figure 3 (b) (for $h = 0$). One still needs to prove that there is only one point of inflection after α attained its maximum. This is done in Appendix 7.4.

For small values of a, there is only one real root for $h > 0$ and only complex roots for $h = 0$. As a increases, there exists a particular value of a for which the line $mg\rho - h\omega$ intersects α not far from its maximum value. In the case $h = 0$, the intersection is at the maximum value of α itself. This a_0 and the corresponding ω_0 are explicitly given by

$$a_0 = \sqrt{\frac{2mg\rho}{fq^2}\big(4\omega_0^3 + 2(p^2-2q)\omega_0\big)},$$

[7] Adopting Rocard's terminology (the French-English translation is the authors').

with $\omega_0 = \left(\frac{-(p^2-2q)+\sqrt{(p^2-2q)^2+12q^2}}{6}\right)^{\frac{1}{2}}$. See Appendix 7.5 for details. Table 1 (for $h > 0$) and Table 2 (for $h = 0$) present a summary of the different possible real roots for the polynomial in Eq.(12) according to the value of a.

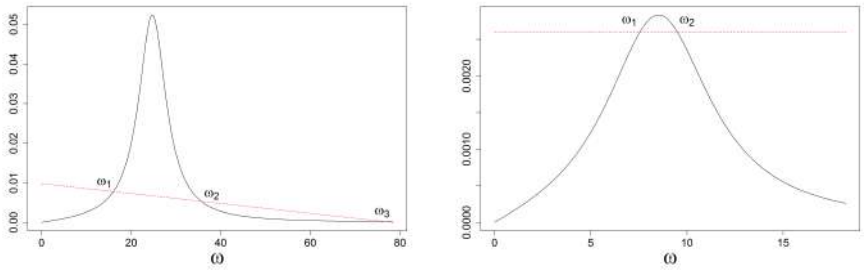
$a = 0$	$\omega_3 = \frac{mg\rho}{h} \in \mathbb{R}$
$]0, a_1[$	$\omega_3 \in \mathbb{R}$
$a = a_1$	$\omega_1 = \omega_2$ $\omega_3 \neq \omega_1, \quad \omega_j \in \mathbb{R},\ j = 1, 2, 3$
$]a_1, a_2[$	$\omega_1 \neq \omega_2, \quad \omega_2 \neq \omega_3,$ $\omega_3 \neq \omega_1, \quad \omega_j \in \mathbb{R},\ j = 1, 2, 3$
$a = a_2$	$\omega_1 \neq \omega_2$ $\omega_2 = \omega_3, \quad \omega_j \in \mathbb{R},\ j = 1, 2, 3$
$]a_2, +\infty[$	$\omega_1 \in \mathbb{R}$

Table 1. Real roots of the polynomial in Eq.(12) for $h > 0$ with respect to $a \geqslant 0$ and for positive parameters $m, \mathsf{M}, \rho, \mathsf{K}, f$ and g.

$[0, a_0[$	no real roots
$a = a_0$	$\omega_0 = \omega_1 = \omega_2 \in \mathbb{R}$
$]a_0, +\infty[$	$\omega_1 \neq \omega_2, \quad \omega_j \in \mathbb{R},\ j = 1, 2$

Table 2. Real roots of the polynomial in Eq.(12) for $h = 0$ with respect to $a \geqslant 0$ and for positive parameters $m, \mathsf{M}, \rho, \mathsf{K}, f$ and g.

Having determined how many real solutions Eq.(11) has with respect to a, the next question is: under Rocard's assumption, towards which solution of Eq.(11) does the B-S system converge?



(a) (b)

Fig. 3. In black, function α as defined in Eq.(13) with respect to ω. The red dashed line is $mg\rho - h\omega$. In Figure (a) $a = 0.05$ and in Figure (b) $a = 0.04$, while the values of the other parameters are those of the first line in Table 4 for Figure (a) and the third line in Table 4 for Figure (b).

To answer this question, let us focus on the case $h = 0$ and follow Rocard's reasoning (here $a > a_0$). For this, let $\omega_1 < \sqrt{\frac{\mathsf{K}}{\mathsf{M}}}$ and $\omega_2 > \sqrt{\frac{\mathsf{K}}{\mathsf{M}}}$ be the two real solutions of Eq.(11). Since ω_1 solves Eq.(11), we have $\alpha(\omega_1) = mg\rho$. On the other hand, from Eq.(3), the quasi-uniform assumption $\theta(t) \cong \omega t$ implies:

$$\mathsf{J}\ddot{\theta} \cong 0 = mg\rho + g\mathsf{M}a\cos(\theta(t)) + H_a^{(2)}(\theta(t), u(t)),$$

with $u(t)$ given by Eq.(10) and $\theta(t) \cong \omega t$, so that:

$$-\alpha(\omega_1) = -mg\rho = g\mathsf{M}a\cos(\theta(t)) + H_a^{(2)}(\theta(t), u(t)).$$

Hence,

$$\mathsf{J}\ddot{\theta} \cong 0 = mg\rho - \alpha(\omega_1),$$

that is, the drum's angular acceleration is approximately $mg\rho - \alpha(\omega_1)$. Suppose now that the drum is slightly accelerated (i.e. $\omega_1 \to \omega_1 + \delta$, with $\delta > 0$). By the shape of the "resonance" curve α, we have $\alpha(\omega_1) < \alpha(\omega_1 + \delta)$, implying that $mg\rho - \alpha(\omega_1 + \delta) < 0$, and therefore $\mathsf{J}\ddot{\theta} \cong mg\rho - \alpha(\omega_1 + \delta) < 0$. This implies that the angular velocity decreases and so the nominal pulsation ω_1 tends to be recovered. The same reasoning holds when the drum is slightly decelerated. A similar analysis for the solution ω_2 can be performed. Table 3 summarises the results for the four possibilities.

$\omega_1 + \delta$	$\alpha(\omega_1 + \delta) \uparrow$	$\mathsf{J}\ddot{\theta} \cong mg\rho - \alpha(\omega_1 + \delta) < 0$	ω_1 recovered
$\omega_1 - \delta$	$\alpha(\omega_1 - \delta) \downarrow$	$\mathsf{J}\ddot{\theta} \cong mg\rho - \alpha(\omega_1 - \delta) > 0$	ω_1 recovered
$\omega_2 + \delta$	$\alpha(\omega_2 + \delta) \downarrow$	$\mathsf{J}\ddot{\theta} \cong mg\rho - \alpha(\omega_2 + \delta) < 0$	ω_2 not recovered
$\omega_2 - \delta$	$\alpha(\omega_2 - \delta) \uparrow$	$\mathsf{J}\ddot{\theta} \cong mg\rho - \alpha(\omega_2 - \delta) > 0$	ω_2 not recovered

Table 3. Summary of the perturbation analysis on the theoretical pulsations ω_1 and ω_2.

Therefore, the pulsation ω_1 persists under slight perturbations, which is not the case for ω_2. A similar analysis is also presented in Panovko and Gubanova[14]. When $h > 0$ and depending on the initial conditions, the dynamical system may converge towards two different pulsations ω, thus allowing for bistability. See Roseau[17] for further details.

4.5 Route to chaotic evolution - Feigenbaum cascade of period-doubling bifurcations

In the general case where not only the $H_a^{(2)}(\theta, u)$ retro-action is active but also $\theta(t) \neq \omega t$, the previous Rocard's quasi-uniform regime becomes unstable itself. Therefore, further discussions are to be based on numerical solutions, i.e. an impossible approach at Y. Rocard's time[8]. As will be shown in the

[8] For an overview of historical events concerning chotic dynamical systems see chapter 4 in Lorenz[12], chapter 2.5 in Skiadas[18] or in chapter 10.4 in Ginoux[8].

next section, for a whole range of control parameters, the B-S dynamics may exhibit a deterministic chaotic evolution, a behaviour obviously not presented in Rocard's 1943 opus. Specifically, we shall numerically unveil that for a sequence of critical crankshaft values a_k for $k = 1, 2, 3, \ldots$, a nowadays classic Feignebaum's cascade of period-doubling bifurcations emerge from the dynamics. The sequence of $\{a_k\}$ converges to an accumulation point a_∞ with the Feigenbaum's universal constant $\mathcal{F}$:

$$\mathcal{F} := \lim_{k\to\infty} \frac{a_{k-1} - a_{k-2}}{a_k - a_{k-1}} \cong 4.6692\ldots . \tag{14}$$

5 Numerical Simulations

All simulations numerically integrate Eq.(2) with initial conditions:

$$\begin{array}{ll} x(0) = \mathsf{A}\sin(\varphi) + \frac{\mathsf{M}g}{\mathsf{K}} & \dot{x}(0) = \omega\mathsf{A}\cos(\varphi) \\ \theta(0) = 0 & \dot{\theta}(0) = \omega \end{array} \tag{15}$$

with $\mathsf{A} = \frac{\frac{\mathsf{K}a}{\mathsf{M}}}{\sqrt{(\frac{\mathsf{K}}{\mathsf{M}}-\omega^2)^2+(\frac{f}{\mathsf{M}})^2\omega^2}}$, $\tan(\varphi) = \frac{-\frac{f}{\mathsf{M}}\omega}{\frac{\mathsf{K}}{\mathsf{M}}-\omega^2}$, and ω the smallest real solution of Eq.(11), if not otherwise stated. These initial conditions correspond to Rocard's approximated solution in Eq.(10), expressed here in the coordinates for Eq.(2). The values for the parameters are in Table 4.

I [kg][m]^2	m [kg]	M [kg]	ρ [m]	K [N]/[m]	f	h	g [m]/[s]^2
0.03	0.1	0.02	0.01	12.5	0.15	0.000125	9.8
0.0001	0.5	0.5	0.1	50	0.01	0.03	9.8
0.0001	0.1	0.03	0.016	2.5	0.2	0	1.625

Table 4. Numerical values of the parameters for different numerical experiments.

Parameter a [m] is given according to the numerical experiment. The value of $g = 1.625$ corresponds to the gravity acceleration on the moon.

5.1 Numerical analysis of the quasi-uniform regime

We investigate how $\dot{\theta}$ behaves with respect to a once the B-S system converges to its equilibrium state. For this, the value of the parameters are as in the first line of Table 4 and for each $a \in \{\frac{6j}{1000} | j = 0, \ldots, 50\}$ (i.e. 51 equidistant point in the interval $[0, 0.30]$), we run the numerical simulation for a time length of $T = 500$. The max, min and mean value of $\dot{\theta}$ for the last 50 time unites are calculated. The results are presented in Figure 4: the 51 equidistant values of a (on the x-axis) are plotted with their corresponding mean value of $\dot{\theta}$ (dots in red), and max and min values (dots in gray). The underlying black line is the predicted pulsation, calculated by Rocard's method (i.e. solving Eq.(11)). In Figure 4 (a), the smallest real solution of Eq.(11) is taken for the initial

conditions in Eq.(15). On the other hand, in Figure 4 (b), the largest real solution of Eq.(11) is taken for the initial conditions, as long as Eq.(11) has three real solutions, and then, as a increases, the smallest real solution is used to determine the initial conditions.

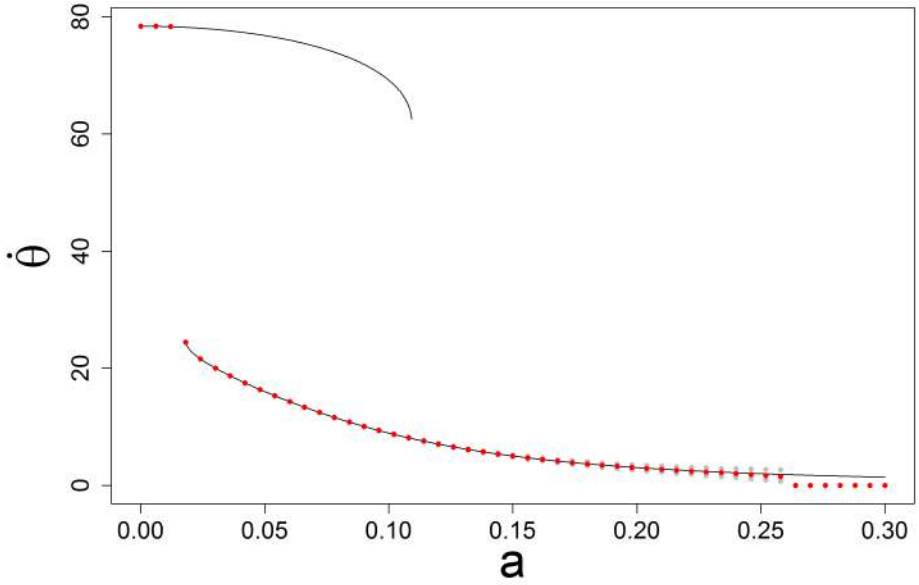
One can clearly appreciate the precision of Rocard's method to predict the pulsation ω, and this even for the case $h > 0$. The prediction is accurate before "l'accrochage" (i.e. for small a) and still valid for a large range of frequencies (form approximately 20 to 5). As a increases, $\dot{\theta}$ oscillates with increasing amplitude around the mean value, hence the increasing difference between max and min values of $\dot{\theta}$. For a larger than approximately $a = 0.26$, the B-S stops oscillating. Hence the red dots with value zero for both plots in Figure 4. Here, the predicated pulsation is no longer valid.

5.2 Period-doubling bifurcations with $h > 0$

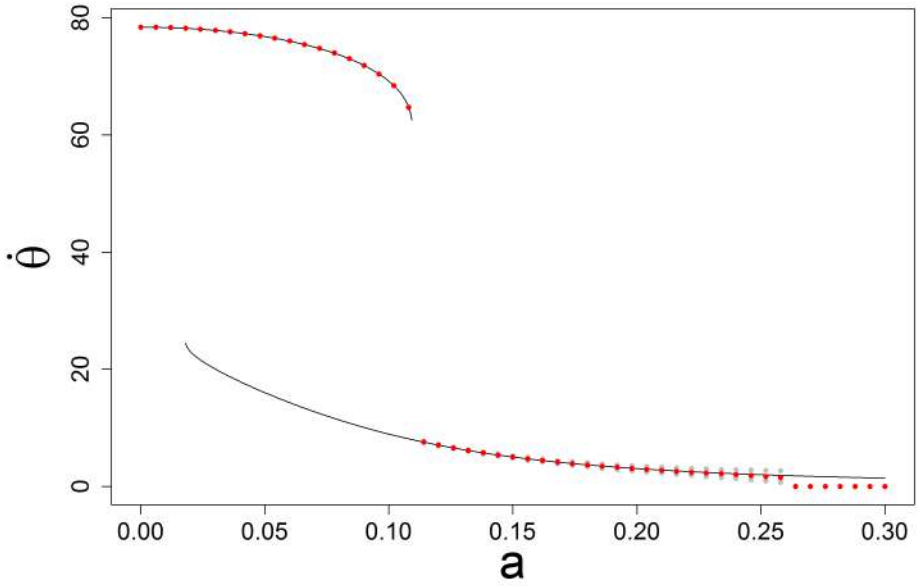
We chose here the crank's length a as control parameter and observe the period-doubling bifurcations on variable x. This case $h > 0$ has the advantage that one can clearly visualize period-doubling bifurcations with variables x and $\dot{x}$. With parameter values as in the second line of Table 4, Figure 5 shows, in its first column (Figures (a), (c), (e) and (g)), the last 7 time units of the numerical integration on [0, 40] for the variables x and $\dot{x}$. The second column in Figure 5 (Figures (b), (d), (f) and (h)) displays (also for the last 7 time units) variables x in black and $\sin(\theta)$ in dashed blue (with an appropriate amplitude and additional constant for comparison's sake). For both columns, each plot shows a specific value of a. Clearly, as a increases, the x variable bifurcates: for $a = 0.06$, x is qualitatively similar to a sin curve and the pair $(x, \dot{x})$ converges towards a limit cycle (i.e. a closed curve without any intersection on the $x - y$ plan). For $a = 0.07$, the variable x bifurcates to a period-two oscillating regime: the pair $(x, \dot{x})$ converges towards a closed curve in $\mathbb{R}^2$ with one intersection. This continues for $a = 0.08$ with a period-four oscillating regime, and for $a = 0.081$ with a period-eight oscillating regime, and so on. Note that in Figure 5 (a), the variable θ is already in a period-two oscillating regime.

As a increases further, the system becomes chaotic. This is numerically shown in Figure 6 when $a = 0.083$. The numerical integration is also on [0, 40] but, in order to appreciate the non-periodicity of this regime better, the last 14 time units are shown.

In Figure 7 (a), the Feigenbaum's cascade of period-doubling bifurcation is displayed for the B-S system with $h > 0$. The different values of the crank a are shown on the x-axis, while the y-axis shows the x values (when $\dot{x} = 0$) for which the B-S converges to (for the corresponding crank value a). Figure 7 (b) shows a zoom out from Figure 7 (a), so one can observe how the x variable (basically the amplitude) depends on a.



(a)



(b)

Fig. 4. Mean (red dots) and max & min (gray dots) of $\dot{\theta}$ against parameter a and predicted pulsation (black line) with parameters as in the first line of Table 4. The black line is the pulsation determined by solving Eq.(11).

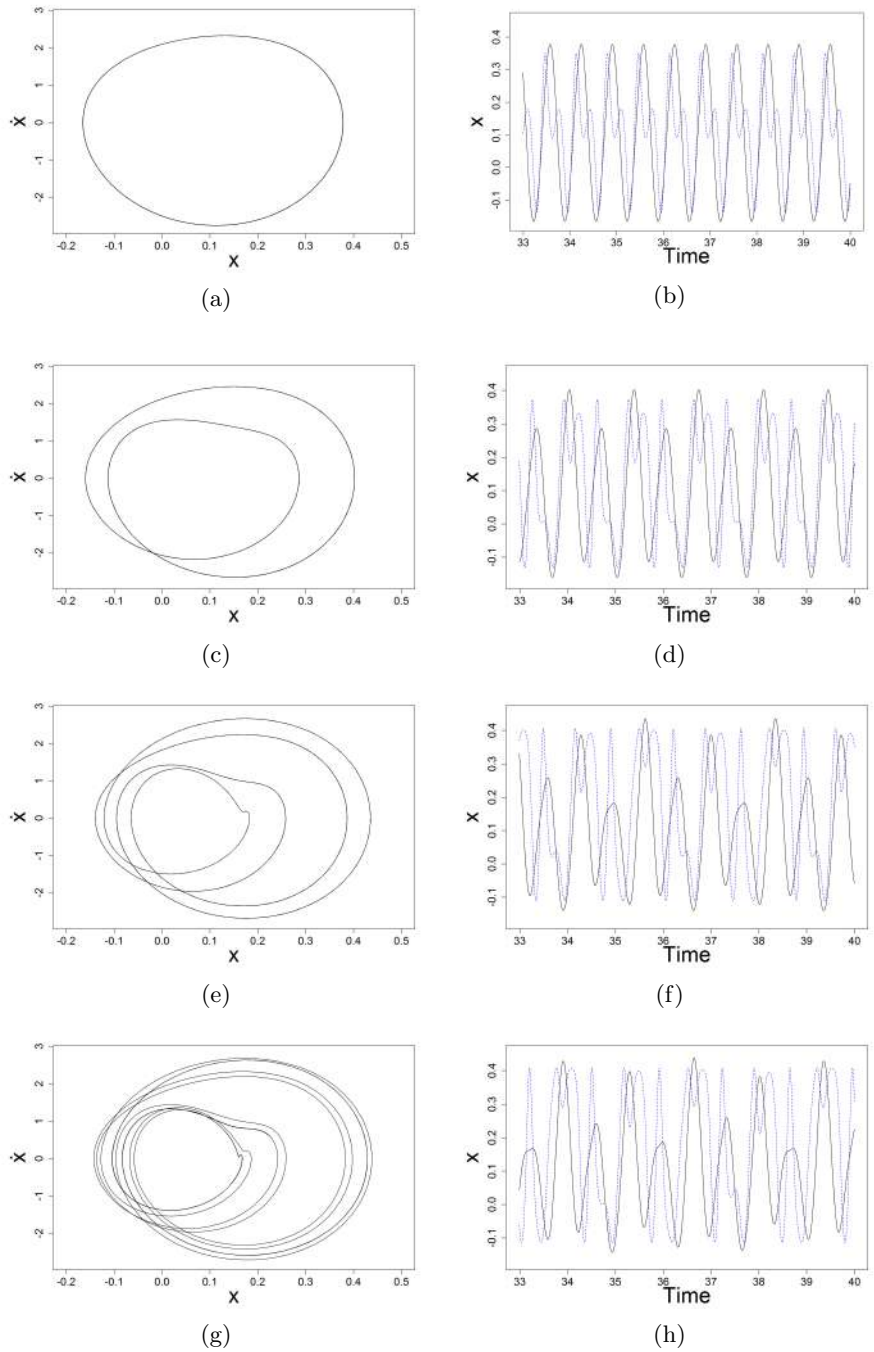
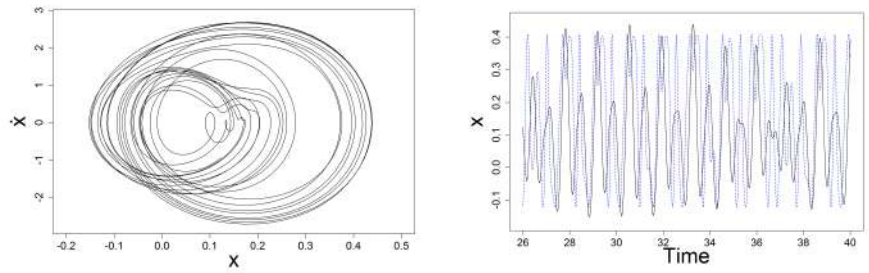


Fig. 5. Time evolution of the $x - y$ representation and of variables x (black) and $\sin(\theta)$ (dashed blue), with parameter value $\mathsf{a} = 0.06$ (Figures (a) & (b)), $\mathsf{a} = 0.07$ (Figures (c) & (d)), $\mathsf{a} = 0.08$ (Figures (e) & (f)) and $\mathsf{a} = 0.081$ (Figures (g) & (h)). The numerical values of the parameters are those of the second line in Table 4.



(a) (b)

Fig. 6. Time evolution of the $x-y$ representation and of variables x (black) and $\sin(\theta)$ (dashed blue), with parameter value $\mathsf{a} = 0.083$ (Figures (a) & (b)). The numerical integration is also on time $t \in [0, 40]$ but, in order to appreciate the non-periodicity of this regime better, the last 14 time units are shown. The numerical values of the parameters are those of the second line in Table 4.

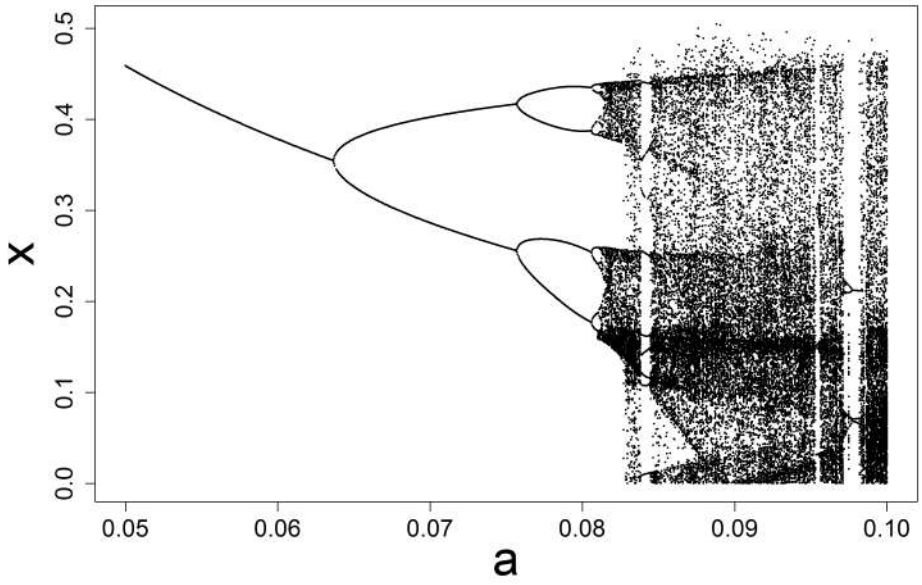
5.3 Period-doubling bifurcations with $h = 0$

As in Section 5.2, period-doubling bifurcations may also appear for B-S with $h = 0$. We again chose here the crank's length a as control parameter and observe the period-doubling bifurcations on variable x. Figure 8 (a) shows yet another Feigenbaum's cascade of period-doubling bifurcation, and here for a B-S system with $h = 0$ as given in Y. Rocard's book Rocard[15]. The numerical values of the parameters are those of the third line in Table 4. As in Figure 7, the different values of the crank a are shown on the x-axis, while the y-axis shows the x values (when $\dot{x} = 0$) for which the B-S converges to (for the corresponding crank value a). To the best of the authors' knowledge, this is the first observation of deterministic chaos in a B-S system for $h = 0$. Period-doubling bifurcations in the case for $h > 0$ (as discussed in Section 5.2) have already been observed in Rodriguez[16].

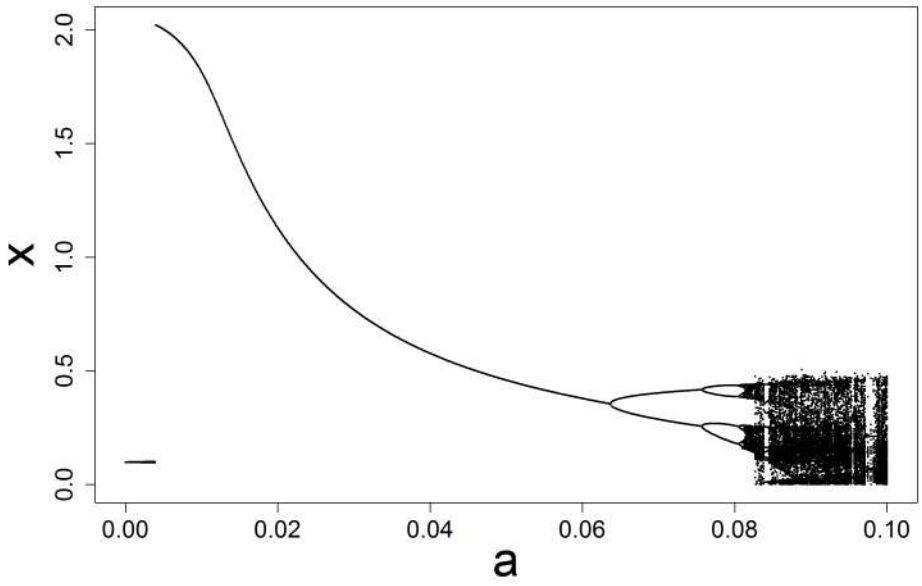
Figure 8 (b) zooms out from Figure 8 (a) and one can see how the variable x depends on a, suggest an antimonotonicity behaviour: the creation and annihilation of periodic orbits. This feature is further developed in the next section.

5.4 Antimonotonicity with $h = 0$

The phenomenon of antimonotonicity, as reported in the seminal works of Kan and Yorke[9], Dawson *et al.*[5] and Kan *et al.*[10]), is observed for a B-S with parameter values as in the third line of Table 4, and for which the inertia I is taken as the second control parameter. In Figure 9, similar bifurcation diagrams as in Figure 8 are presented, while here each plot has a different value for I. Figure 9 (a) has the largest I value and one can see a period-two bifurcation being created and then annihilated as parameter a gets larger. Further decreasing the value of I shows, in Figure 9 (b), yet another period-two

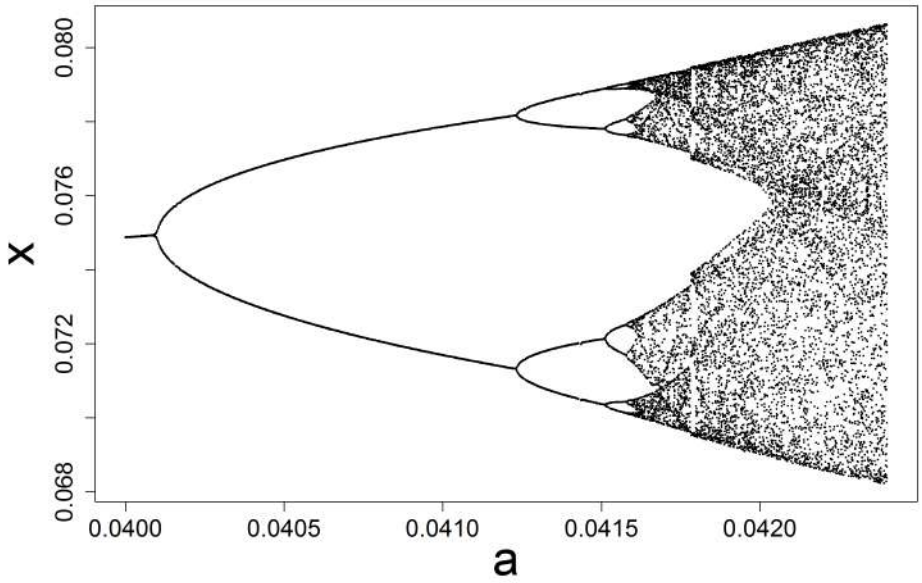


(a)

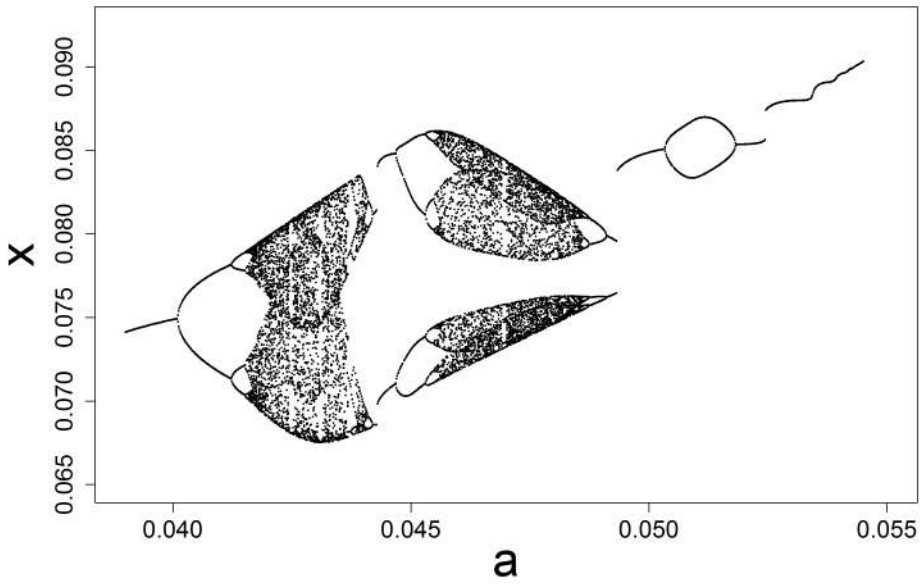


(b)

Fig. 7. In Figure (a), x (when $\dot{x} = 0$) is plotted against parameter a. Figure (b) is a zoom out to show the dynamical behaviour before the the period-doubling cascade (i.e. for smaller values of a). The numerical values of the parameters are those of the second line in Table 4.



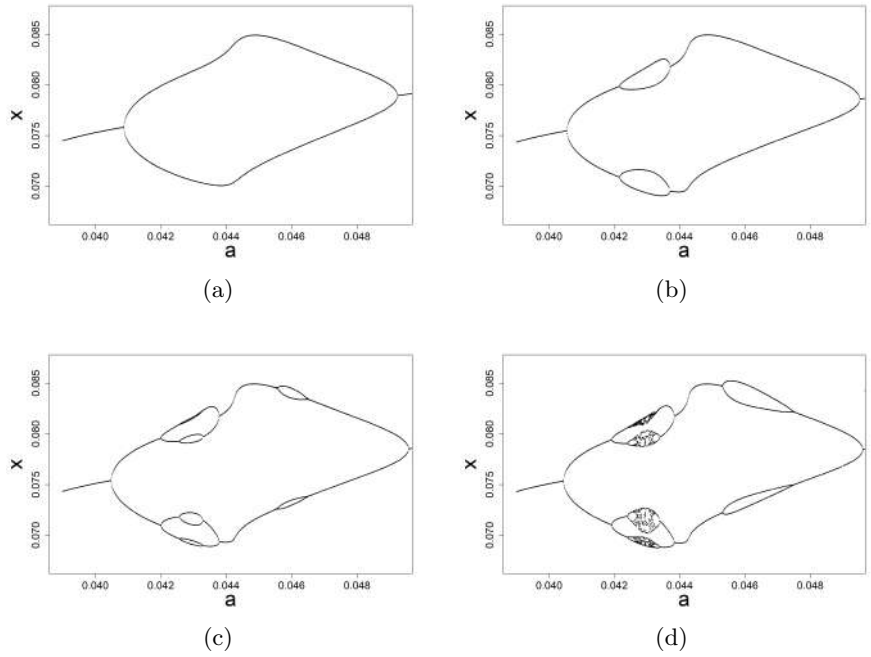
(a)



(b)

Fig. 8. In Figure (a), x (when $\dot{x} = 0$) is plotted against parameter a. Figure (b) is a zoom out to show the dynamical behaviour beyond the period-doubling cascade (i.e. for larger values of a). The numerical values of the parameters are those of the third line in Table 4.

bifurcation: that is, the x variable goes from a period-one, to a period-two, to a period-four and then back to a period-two and finally to a period-one, as a increases in value. This phenomenon continues on as I decreases as shown in Figures 9 (b) and (c).



(a) (b)

(c) (d)

Fig. 9. In all four Figures, x (when $\dot{x} = 0$) is plotted against parameter a. In Figure (a): $I = 0.000115$, in Figure (b): $I = 0.000109$, in Figure (c): $I = 0.0001078$ and in Figure (d): $I = 0.000107$. The other numerical values for the parameters are those of the third line in Table 4.

5.5 Numerical estimation of Feigenbaum's universal constant

In our numerical experiments, we numerically determine the locations of the first six bifurcations thresholds. These are presented in Table 5 (when $h > 0$ and for parameter values as in the second line of Table 4) and in Table 6 (when $h = 0$ and for parameter values as in the third line of Table 4).

6 Conclusions

Yves Rocard used to say: "La physique c'est toujours un petit peu faux"[9]. Indeed, Rocard's idea to determine the pulsation being an approximation is "a

[9] See, for example, Lurçat[13]. In English: "Physics is always a bit wrong".

n	Period	Bifurcation parameter (a_n)	Ratio $\frac{a_{k-1}-a_{k-2}}{a_k-a_{k-1}}$
1	2	0.06367045	-
2	4	0.07569875	-
3	8	0.08059755	2.455356
4	16	0.08106015	10.589710
5	32	0.08115485	4.884900
6	64	0.08117475	4.758794
7	128	0.08117895	4.738095

Table 5. Numerically determined bifurcation points for parameter a for Figure 7 (a) (when $h > 0$ and for parameter values as in the second line of Table 4)

n	Period	Bifurcation parameter (a_n)	Ratio $\frac{a_{k-1}-a_{k-2}}{a_k-a_{k-1}}$
1	2	0.04009785	-
2	4	0.04123185	-
3	8	0.04151235	4.042781
4	16	0.04157945	4.180328
5	32	0.04159425	4.533784
6	64	0.04159745	4.625000
7	128	0.04159815	4.571429

Table 6. Numerically determined bifurcation points for parameter a for Figure 8 (a) (when $h = 0$ and for parameter values as in the third line of Table 4)

bit wrong" - however, one must recognise that Rocard's method is remarkably accurate over a wide range of parameters! Let us emphasise that Rocard had the right intuition despite the fact that numerical simulations were not available in his time. Today's tools not only enable us to confirm Rocard's result, but allow us to study the dynamical system when one relaxes the assumption $\theta(t) \cong \omega t$.

In this contribution, we have numerically shown that the B-S system exhibits period-doubling bifurcations when increasing the values of parameter a and this, even when $h = 0$. Furthermore, we have shown that for small a, the B-S follows "washboard"-like dynamics, implying that it will converge only towards a periodic regime. Only once a is large enough may the B-S have chaotic behaviour.

There are several questions that could be treated as follow up: 1) for what value of a do both systems (drum and spring) converge towards a common pulsation ω for arbitrary initial conditions, and, 2), how does the accuracy of the predicated pulsation ω depend on the inertia I, in particular, in what way does friction h influence the quality of Rocard's method to predict ω.

Acknowledgements: The authors would like to thank Dr. M. Anderegg for his constructive remarks that were very helpful to this contribution.

7 Appendix

7.1 Euler-Lagrange equations

Applying the Euler-Lagrange equations to $L(\theta, x, \dot{\theta}, \dot{x}) = T(\dot{\theta}, \dot{x}) - V(\theta, x)$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \dot{\theta}}) = \frac{\partial L}{\partial \theta}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial L}{\partial \dot{x}}) = \frac{\partial L}{\partial x},$$

we have

$$\begin{array}{ll}
\frac{\partial L}{\partial \dot{\theta}}(\theta, x, \dot{\theta}, \dot{x}) = \frac{\partial T}{\partial \dot{\theta}}(\dot{\theta}, \dot{x}) & = (\mathsf{I} + m\rho^2)\dot{\theta}, \\
\frac{\partial L}{\partial \dot{x}}(\theta, x, \dot{\theta}, \dot{x}) = \frac{\partial T}{\partial \dot{\theta}}(\dot{\theta}, \dot{x}) & = \mathsf{M}\dot{x}, \\
\frac{\partial L}{\partial \theta}(\theta, x, \dot{\theta}, \dot{x}) = -\frac{\partial V}{\partial \theta}(\theta, x) & = -\big(- mg\rho + \mathsf{K}(x - a\sin(\theta))(-a\cos(\theta))\big), \\
\frac{\partial L}{\partial x}(\theta, x, \dot{\theta}, \dot{x}) = -\frac{\partial V}{\partial x}(\theta, x) & = -\big(\mathsf{K}(x - a\sin(\theta)) - \mathsf{M}g\big),
\end{array}$$

and so the dynamical system is given by

$$\begin{aligned}
(\mathsf{I} + m\rho^2)\ddot{\theta} + h\dot{\theta} &= mg\rho + \mathsf{K}(x - a\sin(\theta))a\cos(\theta), \\
\mathsf{M}\ddot{x} + f\dot{x} &= -\mathsf{K}(x - a\sin(\theta)) + \mathsf{M}g.
\end{aligned}$$

7.2 Deriving the equations of motion

We assume the reader is familiar with the modelling of a spring of stiffness K on which a mass M is attached to it:

$$\mathsf{M}\ddot{x} + f\dot{x} = -\mathsf{K}x + \mathsf{M}g,$$

and the modelling of a drum on which a constant torque $mg\rho$ is applied on it:

$$(\mathsf{I} + m\rho^2)\ddot{\theta} + h\dot{\theta} = mg\rho.$$

To derive the equations of motion of the B-S regulator, we must adapt the force of the spring and add an additional torque to the drum, both terms accounting for the coupling. Figure 10 shows the radius of the crank a on to which the spring is attached. The referential for the spring's position is given for $\theta = 0$ (i.e. the doted line for which, by definition, $x = 0$). By definition, upwards from the referential line is considered as a negative displacement and downwards from the referential line is a positive displacement. For θ, clockwise direction is defined as positive. As one can see, for different values of θ, one mus readjust x's reference position. Accordingly, the force of the spring on the mass M is given by $-\mathsf{K}(x - a\sin(\theta))$.

The drum (viewing the flat surface of the cylinder) is sketched in Figure 11. The force of the spring on mass M acts on the drum. The component of this force that acts on the drum as a torque is determined by elementary geometry, as shown in Figure 11. In Panovko and Gubanova[14], a similar derivation is presented.

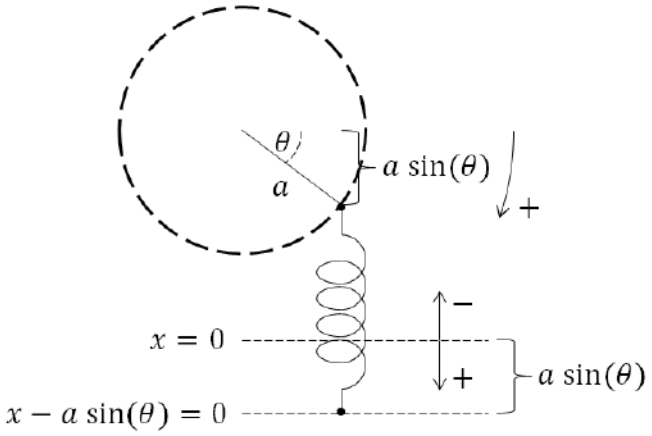


Fig. 10. Determining the force of the spring on mass M according to the value of θ.

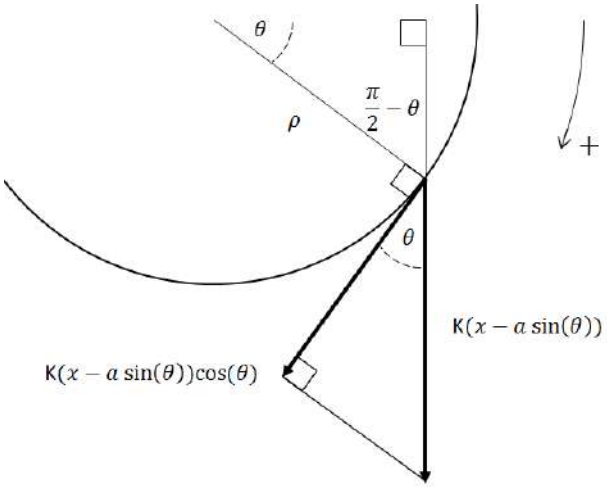


Fig. 11. The additional torque acting on the drum.

7.3 Deriving Eq.(11)

We here derive Eq.(11). By hypothesis, in this regime $\theta(t) \cong \omega t$, then $\ddot{\theta}(t) \cong 0$, enabling us write:

$$0 = H_a^{(1)}(\omega t) + H_a^{(2)}(\omega t, u(t)) - h\omega.$$

Since we are in a stationary regime, the dissipation due to damping has to be counterbalanced by the potential energy delivery over one cycle. Accordingly, over a a full cycle, we have:

$$
\begin{aligned}
0 &= \int_0^{\frac{2\pi}{\omega}} H_a^{(1)}(\omega t) + H_a^{(2)}(\omega t, u(t)) - h\omega dt, \\
&= \int_0^{\frac{2\pi}{\omega}} H_a^{(1)}(\omega t) + H_a^{(2)}(\omega t, u(t)) dt - (h\omega)\frac{2\pi}{\omega}.
\end{aligned}
$$

For the two other integrals, we have

$$
\begin{aligned}
\int_0^{\frac{2\pi}{\omega}} H_a^{(1)}(\omega t) dt &= \int_0^{\frac{2\pi}{\omega}} g(\mathsf{M}a\cos(\omega t) + m\rho) dt, \\
&= g\big(\mathsf{M}a \underbrace{[\frac{\sin(\omega t)}{\omega}]_0^{\frac{2\pi}{\omega}}}_{=0} + m\rho\frac{2\pi}{\omega}\big), \\
&= mg\rho\frac{2\pi}{\omega}.
\end{aligned}
$$

$$
\begin{aligned}
\int_0^{\frac{2\pi}{\omega}} H_a^{(2)}(\omega t, u(t)) dt &= \int_0^{\frac{2\pi}{\omega}} \mathsf{K}a^2\big(\mathsf{A}\sin(\omega t + \varphi) - \sin(\omega t)\big)\cos(\omega t) dt, \\
&= \mathsf{K}a^2\big(\int_0^{\frac{2\pi}{\omega}} \mathsf{A}\sin(\omega t + \varphi)\cos(\omega t) dt - \underbrace{\int_0^{\frac{2\pi}{\omega}} \sin(\omega t)\cos(\omega t) dt}_{[\frac{\sin(\omega t)^2}{2\omega}]_0^{\frac{2\pi}{\omega}} = 0}\big), \\
&= \mathsf{AK}a^2 \int_0^{\frac{2\pi}{\omega}} \sin(\omega t + \varphi)\cos(\omega t) dt, \\
&= \mathsf{AK}a^2 \int_0^{\frac{2\pi}{\omega}} \sin(\omega t)\cos(\varphi)\cos(\omega t) + \cos(\omega t)^2\sin(\varphi) dt,
\end{aligned}
$$

and so we have:

$$
\mathsf{AK}a^2\cos(\varphi)\int_0^{\frac{2\pi}{\omega}} \sin(\omega t)\cos(\omega t) dt = \mathsf{AK}a^2\cos(\varphi)[\frac{\sin(\omega t)^2}{2\omega}]_0^{\frac{2\pi}{\omega}} = 0,
$$

and

$$
\begin{aligned}
\mathsf{AK}a^2\sin(\varphi)\int_0^{\frac{2\pi}{\omega}} \cos(\omega t)^2 dt &= \mathsf{AK}a^2\sin(\varphi)[\frac{t}{2} + \frac{\sin(2\omega t)}{4\omega}]_0^{\frac{2\pi}{\omega}}, \\
&= \mathsf{AK}a^2\sin(\varphi)\frac{\frac{2\pi}{\omega}}{2}.
\end{aligned}
$$

Therefore, we have

$$(h\omega - mg\rho)\frac{2\pi}{\omega} = \mathsf{AK}a^2\sin(\varphi)\frac{\frac{2\pi}{\omega}}{2},$$

$$(h\omega - mg\rho) = \frac{1}{2}\mathsf{AK}a^2\sin(\varphi) = \frac{\frac{(\mathsf{K}a)^2}{\mathsf{M}}\sin(\varphi)}{2\sqrt{(\frac{\mathsf{K}}{\mathsf{M}} - \omega^2)^2 + (\frac{f}{\mathsf{M}})^2\omega^2}},$$

$$\frac{2\mathsf{M}(h\omega - mg\rho)}{(\mathsf{K}a)^2} = \frac{\sin(\varphi)}{\sqrt{(\frac{\mathsf{K}}{\mathsf{M}} - \omega^2)^2 + (\frac{f}{\mathsf{M}})^2\omega^2}}.$$

At this stage, let us note that for the above equation to be consistent (i.e. so that $\sin(\varphi)$ is defined), one must satisfy (for given m, M, ρ, K, f, h, g, a and ω)

$$\Big|\frac{2\mathsf{M}(h\omega - mg\rho)}{(\mathsf{K}a)^2}\sqrt{(\frac{\mathsf{K}}{\mathsf{M}} - \omega^2)^2 + (\frac{f}{\mathsf{M}})^2\omega^2}\Big| \leqslant 1.$$

Now, since

$$\sin(\varphi) = \frac{-\frac{f}{\mathsf{M}}\omega}{\sqrt{(\frac{\mathsf{K}}{\mathsf{M}} - \omega^2)^2 + (\frac{f}{\mathsf{M}})^2\omega^2}},$$

we finally end with

$$\frac{2\mathsf{M}^2(mg\rho - h\omega)}{f(\mathsf{K}a)^2} = \frac{\omega}{(\frac{\mathsf{K}}{\mathsf{M}} - \omega^2)^2 + (\frac{f}{\mathsf{M}})^2\omega^2}.$$

7.4 Point of inflections analysis

For $\omega \geqslant 0$, the function to analyse is (see Eq.(13))

$$\alpha(\omega) = \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2}\Big(\frac{\omega}{(q - \omega^2)^2 + p^2\omega^2}\Big)$$

with $p = \frac{\mathsf{f}}{\mathsf{M}} > 0$, $q = \frac{\mathsf{K}}{\mathsf{M}} > 0$ and $a > 0$. Proving that there is only one point of inflection after α attained its maximum is equivalent to showing that the function α'' changes sign only once in $]\omega_0, +\infty[$, where ω_0 is the unique positive number such that $\alpha(\omega_0)$ is the maximum value of α in $\mathbb{R}_{\geqslant 0}$.

The first derivative of α with respect to ω is

$$\alpha'(\omega) = \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2}\Big(\frac{-3\omega^4 - \big(p^2 - 2q\big)\omega^2 + q^2}{\big((q - \omega^2)^2 + p^2\omega^2\big)^2}\Big)$$

and with the substitution $\omega(s) := \sqrt{sq}$ (i.e. $sq = \omega^2$) and $r = \frac{p^2}{q}$, we have

$$\alpha'(\omega(s)) = \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2q^2}\Big(\frac{-3s^2 - (r-2)s + 1}{(s^2 + (r-2)s + 1)^2}\Big).$$

Define the function β as $\beta(s) := \alpha'(\omega(s))$ as well as $\beta_n(s) := -3s^2-(r-2)s+1$ and $\beta_d(s) := s^2 + (r-2)s + 1$ so that

$$\alpha'(\omega(s)) = \beta(s) = \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2q^2}\frac{\beta_n(s)}{\beta_d(s)^2}.$$

Differentiating the above expression with respect to s leads to

$$\frac{d\big(\alpha'(\omega(s))\big)}{ds} = \alpha''(\omega(s))\omega'(s) = \beta'(s)$$

where α'' is the second derivative with respect to ω, and ω' and β' are, respectfully, the first derivative with respect to s. Since $\omega'(s) = \frac{\sqrt{q}}{2\sqrt{s}}$ and

$$\begin{aligned}\beta'(s) &= \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2q^2}\frac{\big(\beta_d(s)\beta_n'(s) - 2\beta_n(s)\beta_d'(s)\big)}{\beta_d(s)^3} \\ &= \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2q^2}\Big(\frac{6s^3 + 3(r-2)s^2 + ((r-2)^2-10)s - 3(r-2)}{(s^2+(r-2)s+1)^3}\Big)\end{aligned}$$

then the second derivative of α with respect to ω is

$$\alpha''(\omega(s)) = \big(\frac{2\sqrt{s}}{\sqrt{q}}\big)\frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2q^2}\Big(\frac{6s^3 + 3(r-2)s^2 + ((r-2)^2-10)s - 3(r-2)}{(s^2+(r-2)s+1)^3}\Big).$$

With theses calculations, we show that α has a unique maximum for $\omega \geqslant 0$. Setting the first derivate of α to equal zero (i.e. $\alpha'(\omega) = 0$) and with the substitution $x := \omega^2$, we have

$$0 = 3x^2 + (p^2 - 2q)x - q^2 \tag{16}$$

giving the two roots

$$x_{1,2} = \frac{-(p^2-2q) \pm \sqrt{(p^2-2q)^2 + 12q^2}}{6}.$$

Since $\omega \geqslant 0$, we are only interested in the positive root, namely:

$$\omega_0 := \sqrt{x_1} = \big(\frac{-(p^2-2q) + \sqrt{(p^2-2q)^2 + 12q^2}}{6}\big)^{\frac{1}{2}}$$

We still need to show that $\alpha''(\omega_0) < 0$. For this, let s_0 such that $\omega(s_0) = \omega_0$. Then $\beta(s_0) = 0$ and by Eq.(16), $\beta_n(s_0) = 0$. This leads to $s_0 = \frac{-(r-2)+\sqrt{(r-2)^2+12}}{6}$. Therefore,

$$\beta'(s_0) = \frac{f(\mathsf{K}a)^2}{2\mathsf{M}^2q^2}\frac{\beta_n'(s_0)}{\beta_d(s_0)^2}$$

and since $\beta_n'(s) = -6s - (r-2)$, then $\beta_n'(s_0) = -\sqrt{(r-2)^2+12} < 0$. Hence, the function α has one maximum for $\omega \geqslant 0$.

We now analyse the inflections points. For this we need to analyse the signs of $\alpha''(\omega)$ for $\omega \geqslant 0$. This means, we need to study the signs of $\beta'(s)$. There are two cases for $r-2$ to investigate. For both cases $s \geqslant 0$.

Case: $r \geqslant 2$. Here the denominator of β' is strictly positive because $(r-2)$ is positive (i.e. $s^2+(r-2)s+1>0$). According to Descartes' rule of signs, the numerator of β' has only one positive root. Therefore, β' changes sign only once in $\mathbb{R}_{\geqslant 0}$. Since the function α has a its maximum attained by $\omega_0 > 0$ and has 0 as asymptote when ω tends towards $+\infty$, then this unique change of sign for β' corresponds to the one point of inflection of α taking place after ω_0.

Case: $0 < r < 2$. Here the denominator of β' is strictly positive because it has no real roots: the discriminant of $s^2+(r-2)s+1$ is $(r-2)^2-4<0$. According to Descartes' rule of signs, the numerator of β' may either have

a) two positive roots or
b) no roots at all.

Case b) is not possible, since the function α possesses a maximum and has 0 as asymptote (i.e. the function α must have at least one inflection point since α' is zero at ω_0, decreases in value and then converges to zero when ω tends towards $+\infty$). In case a), since there are two positive roots, these must be on either side of ω_0 (the positive number that maximises the value of of α), since at ω_0, the function α'' is strictly negative (i.e. $\alpha''(\omega_0)<0$) and the sign of the dominant term of the polynomial in the numerator is positive. Hence, one point of inflection after ω_0 (i.e. after attaining the maximum of α).

7.5 Determining the particular a_0 with $h = 0$

From Appendix 7.4, we know that the function α (see Eq.(13)) attains its maximum value for

$$\omega_0 = \sqrt{\frac{-(p^2-2q)+\sqrt{(p^2-2q)^2+12q^2}}{6}}.$$

Substituting the values a_0 and ω_0 in Eq.(11), leads to:

$$\begin{aligned}
\frac{2mg\rho\mathsf{K}^2}{f(\mathsf{K}a_0)^2} &= \frac{\omega_0}{\omega_0^4+(p^2-2q)\omega_0^2+q^2},\\
2mg\rho\mathsf{K}^2\big(\omega_0^4+(p^2-2q)\omega_0^2+q^2\big) &= \omega_0(f\mathsf{K}^2)a_0^2,\\
a_0^2 &= \frac{2mg\rho}{fq^2}\big(\frac{\omega_0^4+(p^2-2q)\omega_0^2+q^2}{\omega_0}\big).
\end{aligned}$$

Since $3\omega_0^4 + (p^2 - 2q)\omega_0^2 - q^2 = 0$ (see Eq.(16)), then

$$\begin{aligned} a_0^2 &= \frac{2mg\rho}{fq^2}\Big(\frac{\omega_0^4 + (p^2-2q)\omega_0^2 + \big(3\omega_0^4 + (p^2-2q)\omega_0^2\big)}{\omega_0}\Big), \\ &= \frac{2mg\rho}{fq^2}\Big(\frac{4\omega_0^4 + 2(p^2-2q)\omega_0^2}{\omega_0}\Big), \\ &= \frac{2mg\rho}{fq^2}\big(4\omega_0^3 + 2(p^2-2q)\omega_0\big). \end{aligned}$$

Therefore, the critical crankshaft is

$$a_0 = \sqrt{\frac{2mg\rho}{fq^2}\big(4\omega_0^3 + 2(p^2-2q)\omega_0\big)},$$

with $\omega_0 = \Big(\frac{-(p^2-2q)+\sqrt{(p^2-2q)^2+12q^2}}{6}\Big)^{\frac{1}{2}}$.

References

1. J.A. Blackburn, M. Cirillo and N. Grønbech-Jensen *A survey of classical and quantum interpretations of experiments on Josephson junctions at very low temperatures*, Physics Reports, **611**, 1–33, 2016.
2. I.I. Blekhman, G.Y. Dzhanelidze, *The Bouasse Sarda dynamic regulator*. Bulletin of the Academy of Sciences of the USSR, Division of technical science **10**, 1955.
3. G. Colombo, *Teoria del regolatore di Bouasse e Sarda*. Rendiconti del Seminario Matematico della Università di Padova **28**, 338–347, 1958.
4. P. Coullet, J.M. Gilli, M. Monticellli and N. Vandenberghe, *A damped pendulum forced with a constant torque*. American Journal of Physics, **73**, 1122–1128, 2005.
5. S.P. Dawson, C. Grebogi, J.A. Yorke, I. Kan, H. Koçak, *Antimonotonicity: inevitable reversals of period-doubling cascades*. Physics Letters A, **162(3)**, 249–254, 1992.
6. M. Fridman, *The transient and steady states of the Bouasse-Sarda regulator*. Journal of Sound and Vibration **123**, 341–356, 1988.
7. M. Fridman, *The method of multiple scales applied to the Bouasse-Sarda regulator*. Journal of Sound and Vibration **138**, 135–154, 1990.
8. J.-M. Ginoux, *History of Nonlinear Oscillations Theory in France (1880-1940)*. Springer, 2017.
9. I. Kan, J.A. Yorke, *Antimonotonicity: Concurrent creation and annihilation of periodic orbits*. Bulletin of the American Mathematical Society, **23(2)**, 469–476, 1990.
10. I. Kan, H. Koçak, J.A. Yorke, *Antimonotonicity: Concurrent creation and annihilation of periodic orbits*. Annals of Mathematics, **136(2)**, 219–252, 1992.
11. M. Lakshmanan and S. Rajasekar, *Nonlinear Dynamics, Integrability, Chaos and Patterns*. Springer, 2003.
12. E.N. Lorenz, *The Essence of Chaos*. University of Washington Press, 1995.
13. F. Lurçat, *Le Chaos*. Que sais-je?, 2002.
14. Ya.G. Panovko, I.I. Gubanova, *Stability and Oscillation of Elastic Systems: Modern Concepts, Paradoxes and Errors*. NASA Technical Translation, 1973.
15. Y. Rocard, *Dynamique Générale des Vibrations*. Masson et C^{ie}, 1943.

16. J. Rodriguez, *Networks of Self-Adaptive Dynamical Systems*. PhD thesis, Ecole Polythechnique Fédérale de Lausanne, 2011.
17. M. Roseau, *On the coupling between a vibrating mechanical system and the external forces acting on it*. International Journal of Non-Linear Mechanics, **17** 211–216, 1982.
18. Ch.H. Skiadas, *The Foundations of Chaos Revisited: From Poincaré to Recent Advancements*. Springer, 2016.

Localized waves in silicates. What we know from experiments?

F. Michael Russell, Juan F.R. Archilla, and Santiago Medina-Carrasco

Abstract Since the latest review about solitary localized waves in muscovite, called quodons, [FM Russell, Springer Ser. Mater Sci. 221 (2015) 3] there have been many developments, specially from the point of view of experiments, published in several journals. The breakthrough hypothesis that was advanced in that review that dark tracks were produced by positive electrical charge moving in a localized wave, either transported by swift particles or by nonlinear localized waves, has been confirmed by experiments in muscovite and other silicates. In this paper we review the experimental results, some already published and some new, specially the phenomenon of charge transport without an electric field, called hyperconductivity. We also consider alternative explanations as phase transitions for other tracks. We also attempt to describe numerical simulations that have confirmed the order of magnitude of quodons energy and calculations underway to determine more properties of electron and hole transport by quodons.

Key words: Layered silicates, nonlinear waves, quodons, kinks, breathers, charge transport, hyperconductivity.

F.M. Russell
School of Computing and Engineering, University of Huddersfield, HD1 3DH, United Kingdom, e-mail: mika2mike@aol.com

J.F.R. Archilla (✉)
Group of Nonlinear Physics, Universidad de Sevilla, ETSII, Avda. Reina Mercedes s/n, 41012-Sevilla, Spain, e-mail: archilla@us.es

S. Medina-Carrasco
X-Ray Laboratory, CITIUS, Universidad de Sevilla, Avda Reina Mercedes 4B, 41012-Sevilla, Spain, e-mail: sanmedi@us.es

1 Introduction

The existence of localized waves in silicates layers were first proposed in 1994 [12]. This was an important step in a long story of research about the nature of tracks in muscovite mica since 1967 [23, 24]. A scientific review [32] and a longer historical review [31] were published in 2015. The hypothesis that *quodons*, i.e., quasi on-dimensional lattice excitations, transport electric charge was proposed in those reviews but not developed.

This hypothesis was a fundamental change that led to new theory, new interpretation of previous results about tracks in muscovite, and specially to experiments that confirmed and modified the theory. Therefore, we have thought that it was time for a new review that provided a comprehensive and brief summary of the state of knowledge and the challenges in front of the research.

The research can be divided in three stages that are interconnected.

1. **Tracks**: Tracks by swift particles.
2. **Quodons**: Tracks by lattice excitations or *quodons*.
3. **Hyperconductivity**: Quodons with electric charge and *hyperconductivity*.

Here we present the beginning and end of the three stages and some of the highlights. Later, we will explain in detail some key aspects.

1.1 Tracks *by swift particles*

This stage starts in 1967 with the observation in mica of dark tracks of charged particles from neutrino interactions [23] and finishes in 1993 with an explanation of track formation by release of lattice energy [40] and the description of semi-transparent tracks in mica related with positron dark tracks [42]. Dark tracks are made out of magnetite and some shorter semi-transparent tracks are made out of the mineral epidote.

Note that tracks are also the result of experiments similar to particle tracks in a bubble chamber. They are experiments that nature has made and have been conserved as a fossil in muscovite crystals. They have been done at temperatures, pressure and specially time scales outside of the possibilities of physicists.

1.2 Tracks by lattice excitations or quodons

From the very beginning [23] it has been observed that only 0.1% of the dark tracks in muscovite were produced by swift particles, while the rest lie along the close-packed direction within the cation layers and therefore are related with the crystal structure. This second stage starts with the calculation of nonlinear forces between potassium ions and using them to obtain an approximate KdV equation for

lattice displacements. The KdV equation supports soliton solutions [18], therefore the majority dark lines in muscovite could be produced by lattice-solitons. These results were presented at a conference in 1994 [12] and extended the following year [35, 36].

Interestingly, in the same year 1994, it was attempted to observe lattice-solitons by bombarding silicon with 0.8 MeV Ar^{+} and detecting the ejection of an atom [41]. The experiment failed, perhaps among other reasons because it used silicon which is not layered and have a complicated structure for soliton propagation as the nearest neighbours do not form straight lines.

These lattice excitations were named *qodons* in 1995 [36] and later *quodons* in 1998 [21]. This was an acronym for quasi one-dimensional excitations, a descriptive term which also recognized that the actual type of excitation was not well known. It is worth noting that the term lattice-soliton was changed to *breather*. Breathers differ from solitons in having an internal vibration and smaller energy and were starting to be thoroughly studied [20, 17].

The highlight of this stage is probably the success of another experiment in 2007 [37] similar in design to the previous one [41]. In this case a mica monocrystal was bombarded with alpha particles and it was possible to detect the ejection of atoms at the opposite side of the sample along the direction of close-packed lines within the cation layers.

This stage finishes in 2015 with two comprehensive reviews, a shorter and scientifically oriented one [32] and a longer historical review oriented to the non-specialist [31]. But in these two reviews the next stage is also hinted.

1.3 Quodons with electric charge and hyperconductivity

It was well known that most tracks in muscovite were produced by the recoil of potassium atoms after beta decay [26]. In 2015, a thorough analysis of the decay modes of ^{40}K [4, 11, 22] showed that 90% of decays left a charge behind, and this charge was positive except in 0.001% of positron decays, when it was negative. Then, it was realized that dark tracks by swift particles were produced only by positive particles and that the thickness of, for example, positron tracks, at sonic speed, when they were about to stop, were similar to quodon tracks. These two observations led to the deduction that quodons have electric charge, and dark tracks, positive charge [30]. This hypothesis was already introduced at the previous reviews and it was later extended in Ref. [6].

This profound change in the quodon concept provided something to measure easily, electric current, when quodons were excited by particle bombardment without an electric field, a phenomenon called *hyperconductivity*. Experiments were successful and also were able to explain new properties of quodons [34, 39] in muscovite and other layered silicates.

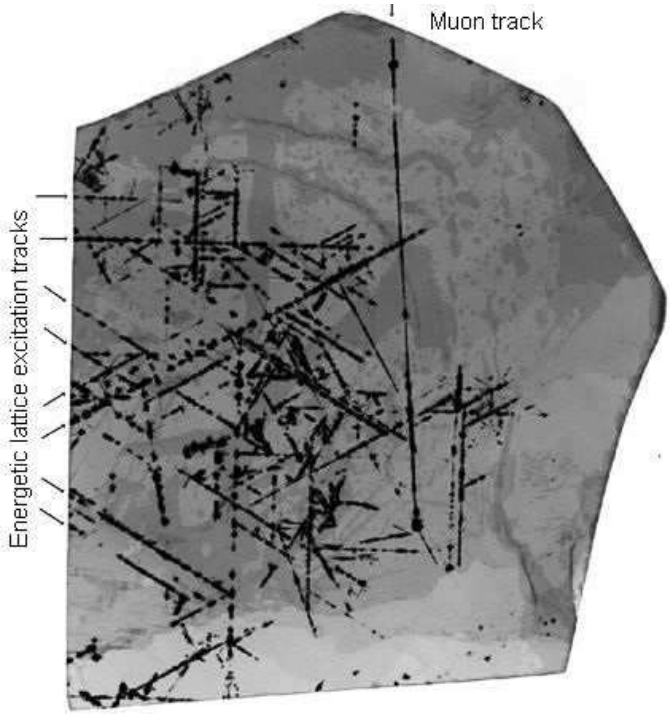


Fig. 1 A sheet of mica muscovite showing many majority tracks due to lattice excitations within the hexagonal structure of the cation layer and a muon track in an unrelated direction

2 Important points

In this section we concentrate in some important points which illustrate either fossil tracks or experimental results or experiment setup.

2.1 First encounter with dark tracks in muscovite

It is important to emphasize that the main author of this research F.M. Russell has been all his career dedicated to high energy physics, first at Harwell Laboratory[1], then at Oak Ridge National Laboratory (ORNL) in the U.S.A, and thereafter at the Rutherford Appleton Laboratory (RAL) in the U.K. In this way, when in 1963 at a museum in North Caroline[2], he found himself in front of a specimen of muscovite

[1] Atomic Energy Research Establishment near Harwell, Oxfordshire, U.K.

[2] Museum of North Carolina Minerals, Spruce Pine, North Caroline, U.S.A.

with abundant dark tracks, he recognized the striking similarity with the tracks of swift particles in bubble chambers. A similar sheet is presented in Fig. 1.

2.2 How were the swift particles identified?

There were different methods, but perhaps the clearest is the kinkiness of those dark tracks. Charged swift particles when entering in matter experience scattering with the matter ions. The probability of scattering at a given angle can be calculated by Rutherford law and the angles can be can be seen and measured with a microscope and the results compared with given particles. An example can be seen in Fig. 2, comparing the second difference, basically the scattering angle, of some track in mica with positrons in photographic film [43], taking into account the difference in mass and density of the scattering ions.

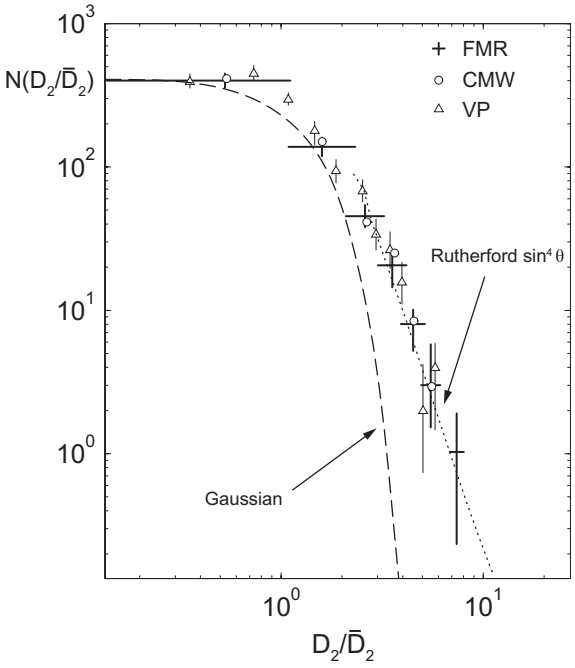


Fig. 2 Probability of scattering at given angles for tracks corresponding to positrons in muscovite [+], compared with positrons in photographic film [△] and by Wolfendale's group [○]. The results fit closely to the Rutherford Law, thus strongly supporting the hypothesis that the lines are tracks of charged particles. Data from Refs. [43, 13, 26] for VP, CMW and FMR, respectively. Reproduced with permission from: Russell, F. M. (1988)[26] Copyright © 1988, Elsevier

2.3 Which particle tracks were identified?

The particles that produce dark tracks in muscovite and could be identified were positive muons, i.e., antimuons, which are the particles that can be produced deep underground after neutrino interaction [23, 24, 25, 26], positrons from ^{40}K decay and antimuon decay [26, 27, 29, 28, 42]. Protons can be recognized by the short length of the tracks corresponding to non-relativistic speed [38, 32]. Also, alpha particles can be discriminated from the multiple scattering events, proof of their large energy and mass [25, 32].

The remarkable fact that all the particles that produce dark tracks were positive was used in 2015 to recognize that the large majority of quodons that produce dark tracks have also positive charge [30, 32, 31].

2.4 How were the tracks produced?

There is not enough energy to produce the dark tracks, this means that the source of energy is already in the lattice, in the form of a metastable state [29, 28, 41, 36].

Natural crystals of muscovite mica contain various impurities, especially iron, incorporated during their growth. It has been found that this can lead to a unique situation, as a crystal cools following growth, during which minute perturbations of the crystal can be recorded and stored indefinitely. Although muscovite is a common mineral in rocks, large crystals grow only in pegmatites associated with magmas at temperatures of about 500°C and under high pressure at about 5 km underground [14]. Inevitably, large single crystals of good quality are rare but they are of special interest because of the information they have been found to contain. A common feature of micas is their ease of cleavage, in the (001)-plane. The black material forming the patterns is the iron oxide mineral magnetite, so named because it is ferro-magnetic.

As a crystal cools slowly at high temperature it tries to reach a lower energy state by expelling the magnetite at the weakest part of the lattice, the cleavage plane. The magnetite grows epitaxially, centred in the potassium sheet and grows in the directions of structural weakness. These are the principal crystallographic directions, which are easily determined by percussion figures [31]. This has been confirmed by both optical and electron microscopy. In fact, the distortion of the lattice is readily seen by observing the of the intrusive magnetite by reflected light or by surface interferometry. Contrary to the basic assumption in of global bi-stability of structure there is no evidence for this in the observed patterns involving magnetite.

2.5 Two different recording processes

It has been found that there are two different recording processes leading to the observed patterns, involving different impurities. The dominant process leading to magnetite is triggered by passage through the crystal of a positive charge in the vicinity of the potassium sheets. This can result from a positively charged, high-energy, muon created in a neutrino interaction within the Earth or by direct penetration of a cosmic ray. Another source is from electron-positron showers arising from a high-energy gamma interaction. The flight-paths of these particles are influenced by channelling and diffraction scattering due to the pronounced layered structure of muscovite [26]. The most informative source, however, is from the rare decay channel of ^{40}K creating positrons [42]. Study of the fossil tracks of these positrons has shown that the origin of the nucleation sites for triggering magnetite growth does not involve ionization of the lattice. For relativistic positrons from this source a fossil track results even when the rate of energy loss is less than 1 eV per 10,000 atoms along the flight path. The rate of energy loss increases as a positron slows down, leading to an increase in the amount of magnetite formed. Due to anisotropy of the mechanical properties of the layered structure this increase shows as a widening of the magnetite ribbon delineating the flight-path. This suggests that the recording process is of a chemical nature, with the probability for an impurity ion migrating to the flight path increasing as the positron's speed decreases.The dominant source of the long ribbons of magnetite arising from moving positive charges is the dominant decay channel of ^{40}K, in which an electron is emitted. These energetic electrons do not initiate fossil magnetite tracks. However, they leave a positive charge at the decay site that can be trapped and carried by a mobile lattice excitation arising from the recoil motion of the decayed nucleus. These mobile, non-dissipative, highly localized excitations move at slightly sub-sonic speed, leading to magnetite ribbons of width of similar width to those due to nearly stopped positrons [32]. The last known source of swift positively-charged ions is from atomic cascades arising from nuclear scattering of relativistic particles.

The second and much rarer recording process involves formation of the mineral epidote, which requires an excess of calcium during crystal growth. These fossil tracks arise from the emission of a positron, leaving a negative charge at the decay site, which is trapped and transported by the mobile recoil excitation. This leads to a ribbon of transparent epidote that is not intrusive in the potassium sheets [42]. The formation process of the epidote is poorly understood and might involve a bistable crystal state [19]. It is hoped that this explanation of the origin of the fossil magnetite-ribbon tracks might encourage study of the formative process for the fossil epidote tracks, as this has the potential for ballistic, low-loss, transport of electrons in layered insulators [34, 42, 33, 39].

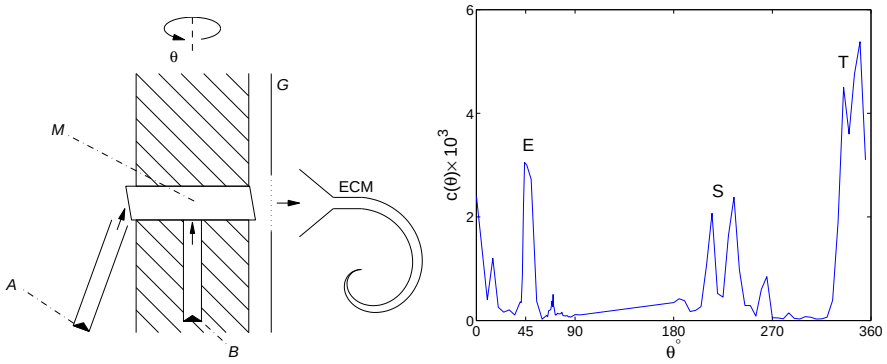


Fig. 3 **Left**: Setup of the quodon experiment. A: alpha source, M: mica monocrystal, ECM: electron channel multiplier, G: grid, B: alternative position for the alpha source. **Right**: Outcome of the quodon experiment: Plot of the angular dependence of the ECM count rate. T: test, S: sputtering from the front face, E: peak from ejected atoms at the rear face in the [0 1 0] direction. Reproduced with permission from: Russell, F.M. and Eilbeck, J.C. [37]. Copyright © 2007, EPLA.

3 How was the experiment in lattice-excitations or quodons done?

The highlight of the research on lattice excitations, sometimes called lattice-solitons, breathers or quodons in this context was the experiment in 2007 [37]. Alpha particles were sent at an angle with the muscovite sheet and therefore with the potassium layer to prevent the possibility of transmission and it was detected at the other side of the monocrystal corresponding to low Miller indexes, the ejection of an atom from the surface. The atom was detected because it was ionized by an electric field and the charge detected. Ejection of atoms from a silicate surface needs energies of 7-8 eV, however, it is not necessary that a quodon had that energy as the passage of a vibrational energy in the vicinity of the surface is enough to increase the probability of ejection [16]. Both the setup and the outcome can be seen in Fig. 3.

4 How it was demonstrated hyperconductivity, i.e., that quodons carry charge?

Hyperconductivity is defined as the transport of charge in absence of an electric field. The charge is transported by nonlinear excitations which have their own energy and momentum from the cause that created them. Due to the combination of nonlinearity and discreteness they travel long distances in atomic terms with little attenuation.

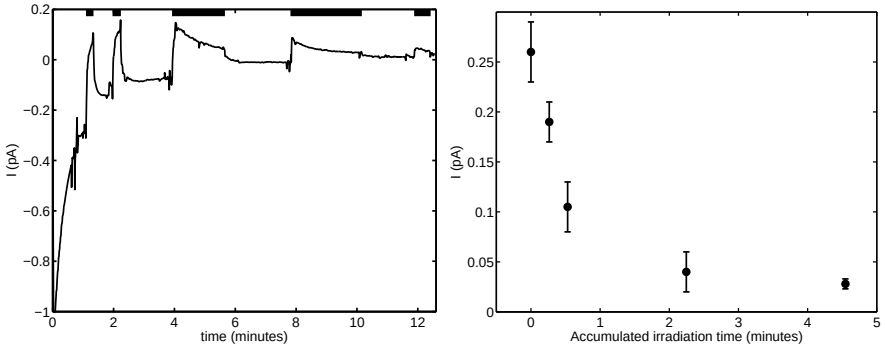


Fig. 4 Left: Plot of the hyperconductivity current, the time intervals marked in black at the top correspond to the opening of the alpha gate. **Right**: Hypercurrent as a function of the accumulated time of alpha exposure, showing an exponential decrease corresponding to the depletion of the charge reservoir from ^{40}K beta decay. Reproduced with permission from: Russell, F.M. et al. [34]. Copyright © 2017, EPLA.

An experiment was set up quite similarly to the previous ones. The way to excite lattice excitations or quodons was also by sending alpha particles, due to its simplicity. The hypothesis was that alpha particles would produce many quodons and some proportion of them would propagate to the other side of the sample and in this way a current could be measured. Muscovite is a very good insulator but there was the possibility that the surface and certainly the ionized air would transport charge. To discard this effect the two sides of the sample were connected and therefore the potential difference among both contacts would be zero and also the electric field would be zero. Lattice excitation or quodons would travel due to their initial energy and momentum.

The experiment was a success but with some unexpected results. Instead of a steady current after the alpha gate was open, the current showed a peak, but then it would diminish to a small limiting value. The phenomenon was soon explained: there are not free carriers in muscovite band structure, the available charge is the one obtained after beta decay of ^{40}K, mainly positive after β^-, i.e., the emission of an electron is the dominant branch, but also some negative charge after β^+ positron emission. This reservoir is depleted in some minutes, and the remaining current is exactly the flux of electric charge brought by the alpha flux [34]. The current peaks and their decrease can be seen in Fig. 4.

5 What properties of hyperconductivity and quodons were deduced from experiments?

More experiments in hyperconductivity [39] were able to deduce a number of facts:

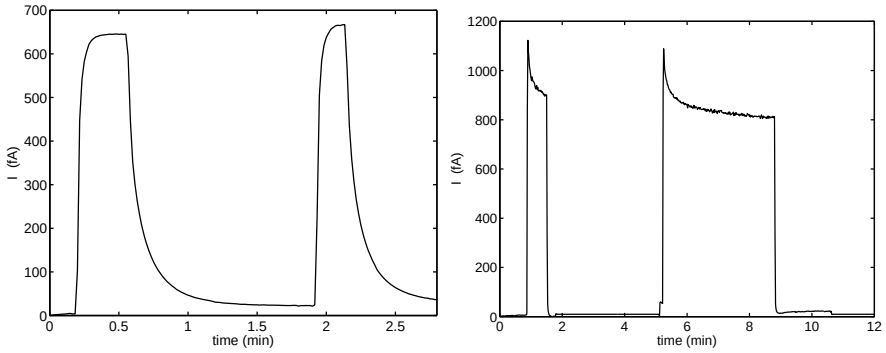


Fig. 5 Plot of the hypercurrent corresponding to two intervals of alpha exposure in a previously depleted crystal. **Left**: Crystal of lepidolite of good quality. Note the soft decay of the hypercurrent after the alpha flow is stopped. **Right**: Crystal of phlogopite of bad quality. Note the abrupt decrease of the hypercurrent after alpha irradiation is stopped. Reproduced with permission from: Russell, F.M. et al. [39]. Copyright © 2019, EPLA.

- Other layered silicates as lepidolite, phlogopite, chrysolite and both natural and synthetic fluorphlogopite supported hyperconductivity and thus the propagation of quodons. However, a layered silicate as biotite with similar structure does not support it. It was not found in unrelated materials that could be used in quodon technology as PTFE, quartz, borosilicate glass and epoxy resin.
- Hyperconductivity is not sensitive to minor crystal defects and can even anneal some of them. It can also pass through some interfaces.
- Hyperconductivity is not affected by magnetic fields up to 1.1 T.
- Quodons have very long flight paths, this can be deduced by comparing the drop in hypercurrent when the alpha bombardment is stopped. In a good crystal the hypercurrent continues to flow some seconds, while in a crystal with many defects, the hypercurrent stops almost immediately as can be seen in Fig. 5

6 What types of quodons are there?

There is no clear information from the experiments, however from the fossil tracks as seen in Fig. 6, it can be deduced:

- There are positive quodons, negative quodons and probably neutral quodons. Negative quodons can be seen as an epidote track in exactly the opposite direction from a positron track and therefore corresponding to the recoil of the nucleus of ^{40}K after β^+ decay. Neutral quodons can be deduced from intermittent dark tracks, which seem like quodons loosing an regaining positive charge.
- There are some more energetic quodons that produce straighter and thicker dark tracks and some less energetic quodons because they appear often as weaker

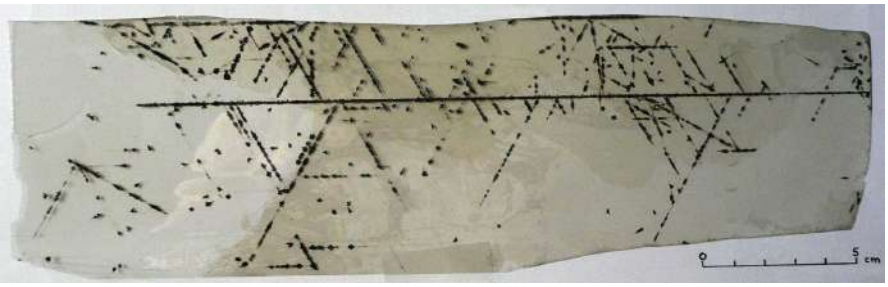


Fig. 6 A sheet of mica muscovite showing a quodon primary track and many secondary tracks scattered from it. Also it is possible to see the intermittency in the secondary tracks along the close packed direction of the cation layer. This is interpreted as a quodon loosing and regaining a positive charge. Reproduced with permission from: Russell, F.M. and Eilbeck, J.C. [38]. Copyright

dark tracks scattered from a primary track. As both types are dark, it is deduced that both have positive charge. They might have different nature, maybe primary tracks could be crowdions or kinks as they transport charge in an ionic crystal and have large energies of 20-30 eV [15, 4, 9, 5, 7]. Secondary tracks, could be interpreted as breathers, because they have good mobility in mica models with little or no radiation, with energies of 0.2-0.3 eV [21, 8, 2] and recently they have been shown to scatter in different close-packed directions [10]. However, breathers do not transport charge and if they couple to a charge their properties and physical description change completely. Certainly, breathers could correspond to neutral quodons. A model for lattice excitations coupled to a hole or electron has been constructed for muscovite, but the properties of localized excitations using it are still under study [1, 3].

7 Alternative explanations of tracks

There have not been many alternative explanations of tracks in muscovite. It was suggested that the majority of dark lines corresponds to dislocations because they lie in the close-packed directions, but without further proof [13]. Arguments against dislocations are that they should appear along crystal fractures, which does not occur [26] and that dark tracks do not continue to the edge of the crystal specimen as it should occur with dislocations [32]. Recently an interesting explanation based on phase transition in a bistable lattice has been proposed [19]. The research was based in the observation[3] that the pitch of the on-site potential and the equilibrium distance of the interatomic potential should be different in a real material. This bring about the existence of different stable configurations, and the authors found a switching

[3] J.F.R. Archilla, private communication (2019).

wave between configurations that propagates longitudinally along the direction of atomic chains. They used a Frenkel-Kontorova 2D system with morse interaction potential. There was no attempt to relate their findings with physical magnitudes and to explain the coloration of lines or the kinkiness of the swift particle tracks. Also, the hyperconductivity experiments were not explained and the charge of the ions in the cation layer were not taken into account as explained in the article. Nevertheless, it opens a new path to understand some of the phenomena observed in muscovite and other layered silicates, particulary epidote tracks, which are not produced by swift particles.

8 Summary

In this article we have tried to present an updated review of the research in nonlinear waves in layered silicates, particularly, but not only, in mica muscovite. We have attempted to make clear for the non specialist which are the main experimental facts and their interpretation, leaving many details to the references. The main results are that some dark tracks in muscovite can be related to swift positive particles, that many other tracks along atom chain direction of the cation layers can be interpreted as lattice excitations, called quodons. Most quodons carry positive charge although some may have negative charge or none. This was demonstrated by hyperconductivity experiments, that is, the transport of charge in the absence of an electric field. Variants of hyperconductivity experiments allowed for the deduction of many properties of quodons. Other interpretations of dark tracks may be complementary and be useful to understand some of the tracks.

Acknowledgments

JFRA thanks a travel grant from VIPPITUS 2020 and projects PAIDI 2019/FQM-280 and MICINN PID2019-109175GB-C22. SM-C acknowledges grant from PPITUS-2018.

References

1. Archilla, J.F.R., Doi, Y., Kimura, M.: Energy and charge transport in a silicate layer. In: Proceedings of NOLTA 2019: the 2019 International Symposium in Nonlinear Theory and its Applications, Kuala Lumpur, Malaysia, December 2-6, 2019, p. 453. IEICE, Malaysia (2019)
2. Archilla, J.F.R., Doi, Y., Kimura, M.: Pterobreathers in a model for a layered crystal with realistic potentials: Exact moving breathers in a moving frame. Phys. Rev. E **100**(2), 022206 (2019)

3. Archilla, J.F.R., Doi, Y., Kimura, M., Kosevich, Yu.A.: Hole and electron transfer in an 1D model of a silicate layer (2020). Preprint
4. Archilla, J.F.R., Kosevich, Yu.A., Jiménez, N., Sánchez-Morcillo, V.J., García-Raffi, L.M.: A supersonic crowdion in mica. In: J.F.R. Archilla, et al. (eds.) Quodons in mica, *Springer Ser. Mat. Sci.*, vol. 221, pp. 69–96. Springer, Cham (2015)
5. Archilla, J.F.R., Kosevich, Yu.A., Jiménez, N., Sánchez-Morcillo, V.J., García-Raffi, L.M.: Ultradiscrete kinks with supersonic speed in a layered crystal with realistic potentials. Phys. Rev. E **91**, 022912 (2015)
6. Archilla, J.F.R., Russell, F.M.: On the charge of quodons. Letters on Materials **6**, 3–8 (2016)
7. Archilla, J.F.R., Zolotaryuk, Y., Yu. A. Kosevich, Doi, Y.: Nonlinear waves in a model for silicate layers. Chaos **28**(8), 083119 (2018)
8. Bajars, J., Eilbeck, J.C., Leimkuhler, B.: Nonlinear propagating localized modes in a 2D hexagonal crystal lattice. Physica D **301-302**, 8 – 20 (2015)
9. Bajars, J., Eilbeck, J.C., Leimkuhler, B.: Numerical simulations of nonlinear modes in mica: Past, present and future. In: J.F.R. Archilla, et al. (eds.) Quodons in Mica, *Springer Ser. Mater. Sci*, vol. 221, pp. 35–67. Springer, Cham (2015)
10. Bajars, J., Eilbeck, J.C., Leimkuhler, B.: 2D mobile breather scattering in a hexagonal crystal lattice (2020). arXiv:2007.12079 [nlin.PS]
11. Cameron, J.A., Singh, B.: Nuclear data sheets for A=40. Nucl. Data Sheets **102**(2), 293–513 (2004)
12. Collins, D.R., Russell, F.M.: Computer modelling studies of solitons in layered crystals. In: R. Gruber, F. Tomasini (eds.) Proceedings of the 6th joint EPS-APS International Conference on Physics Computing, Lugano, Switzerland, 22-26 August 1994, pp. 243–246. Manno, Switzerland (1994)
13. Craig, R., Mamidzhanian, E., Wolfendale, A.W.: Ancient cosmic ray tracks in mica? Phys. Lett. B **26**, 468–470 (1968)
14. Deer, W.A., Howie, R.A., Zussman, J.: An introduction to the rock-forming minerals, 3rd edn. The Mineralogical Society, London (2013)
15. Dou, Q., Cuevas, J., Eilbeck, J.C., Russell, F.M.: Breathers and kinks in a simulated crystal experiment. Discret. Contin. Dyn. Syst. S **4**(5), 1107–1118 (2011)
16. Dubinko, V.I., Selyshchev, P.A., Archilla, J.F.R.: Reaction-rate theory with account of the crystal anharmonicity. Phys. Rev. E **83**, 041124 (2011)
17. Flach, S.: Obtaining breathers in nonlinear Hamiltonian lattices. Phys. Rev. E **51**(4), 3579–3587 (1995)
18. Korteweg, D.J., de Vries, G.: On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. Phil. Mag. (5) **39**, 422–443 (1895)
19. Krylova, K.A., Korznikova, E.A., Semenov, A.S., Bachurin, D.V., Dmitriev, S.V.: Linking tracks in mica crystals with phase transitions in a bistable lattice. Eur. Phys. J. B **93**(23), 16001 (2020)
20. MacKay, R.S., Aubry, S.: Proof of existence of breathers for time-reversible or Hamiltonian networks of weakly coupled oscillators. Nonlinearity **7**, 1623 (1994)
21. Marín, J.L., Eilbeck, J.C., Russell, F.M.: Localized moving breathers in a 2D hexagonal lattice. Phys. Lett. A **248**(2-4), 225–229 (1998)
22. Mougeot, X., Helmer, R.G.: LNE-LNHB/CEA–Table de Radionucléides, K-40 tables. `http://www.nucleide.org` (2012)
23. Russell, F.M.: The observation in mica of tracks of charged particles from neutrino interactions. Phys. Lett. **25B**, 298–300 (1967)
24. Russell, F.M.: Tracks in mica caused by electron showers. Nature **216**, 907–909 (1967)
25. Russell, F.M.: Duration of sensitive period for track recording in mica. Nature **217**, 51–52 (1968)
26. Russell, F.M.: Identification and selection criteria for charged lepton tracks in mica. Nucl. Tracks. Rad. Meas. **15**, 41–44 (1988)
27. Russell, F.M.: Positive charge transport in layered crystalline solids. Phys. Lett. A **130**, 489–491 (1988)

28. Russell, F.M.: Channelled lepton tracks and practical perspectives. Nucl. Tracks. Rad. Meas. **19**(1-4), 955–957 (1991)
29. Russell, F.M.: Decorated track recording mechanisms in muscovite mica. Nucl. Tracks. Rad. Meas. **19**, 109–113 (1991)
30. Russell, F.M.: Charge coupling to anharmonic lattice excitations in a layered crystal at 800 K (2015). arXiv:1505.03185 [cond-mat.supr-con]
31. Russell, F.M.: I saw a crystal. In: J.F.R. Archilla, et al. (eds.) Quodons in mica, *Springer Ser. Mat. Sci.*, vol. 221, pp. 475–559. Springer, Cham (2015)
32. Russell, F.M.: Tracks in mica, 50 years later. In: J.F.R. Archilla, et al. (eds.) Quodons in mica, *Springer Ser. Mat. Sci.*, vol. 221, pp. 3–33. Springer, Cham (2015)
33. Russell, F.M.: Transport properties of quodons in muscovite and prediction of hyperconductivity. In: J.F.R. Archilla, et al. (eds.) Nonlinear Systems, Vol. 2, pp. 241–260. Springer, Cham (2018)
34. Russell, F.M., Archilla, J.F.R., Frutos, F., Medina-Carrasco, S.: Infinite charge mobility in muscovite at 300 K. EPL **120**, 46001 (2017)
35. Russell, F.M., Collins, D.R.: Lattice-solitons and non-linear phenomena in track formation. Rad. Meas. **25**, 67–70 (1995)
36. Russell, F.M., Collins, D.R.: Lattice-solitons in radiation damage. Nucl. Instrum. Meth. B **105**, 30–34 (1995)
37. Russell, F.M., Eilbeck, J.C.: Evidence for moving breathers in a layered crystal insulator at 300 K. EPL **78**, 10004 (2007)
38. Russell, F.M., Eilbeck, J.C.: Persistent mobile lattice excitations in a crystalline insulator. Discret. Contin. Dyn. Syst. S **4**(5), 1267–1285 (2011)
39. Russell, F.M., Russell, A.W., Archilla, J.F.R.: Hyperconductivity in fluorphlogopite at 300 K and 1.1 T. EPL **127**(1), 16001 (2019)
40. Russell, F.M., Steeds, J.W.: Track formation by release of lattice energy. Nucl. Tracks. Rad. Meas. **22**(1-4), 65–70 (1993)
41. Schlößer, D., Kroneberger, K., Schosnig, M., Russell, F.M., Groeneveld, K.O.: Search for solitons in solids. Rad. Meas. **23**(1), 209–213 (1994)
42. Steeds, J.W., Russell, F.M., Vine, W.J.: Formation of epidote fossil positron tracks in mica. Optik **92**, 149–154 (1993)
43. Voyvodic, L., Pickup, E.: Multiple scattering of fast particles in photographic emulsions. Phys. Rev. **85**, 91–100 (1952)

Software Realization, Analysis and Experimental Investigation of Equivalent Inductance

Volodymyr Rusyn[1], Christos H. Skiadas[2], Aceng Sambas[3]

[1] Department of Radio Engineering and Information Security, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine
(E-mail: rusyn_v@ukr.net)
[2] Technical University of Crete, University Campus, 73100 Chania, Crete, Greece
(E-mail: skiadas@cmsim.net)
[3] Department of Mechanical Engineering, Universitas Muhammadiyah Tasikmalaya, Tasikmalaya, Indonesia
(E-mail: acenx.bts@gmail.com)

Abstract. Circuit realization of the inductance equivalent that contains two operational amplifiers, one capacitor, and four resistors is presented. The mathematical equation that allow convert inductor value to resistance of potentiometer is shown. Computer modeling results of the algorithm for calculate inductance was realized in the modern software LabView. Experimental results of realization of the equivalent of inductance are presented. The designed layout was applied for chaotic Chua's generator.
Keywords: Chaos, Equivalent Inductance, LabView, Chua's Generator.

1 Introduction

Chaotic theory used in many areas, such as biology [1-3], ecology [4, 5], economy [6-8], optics [9], mathematics [10, 11], memristor [12, 13], security communication systems [14, 15], etc. Many different electronics circuits generated chaotic oscillations [16-29]. One of the circuit element must be used inductor. However, there are many problems with product or buying inductor with non-standard nominal values.
In Fig. 1 shows electrical scheme that allows change inductor to operational amplifier realization.
The equation that allow convert equivalent inductance can be computed as

$$L_{eq} = \frac{R1R3R4C1}{R2}, \tag{1}$$

where R1-R4 – values of resistances, C1 – value of capacitance.

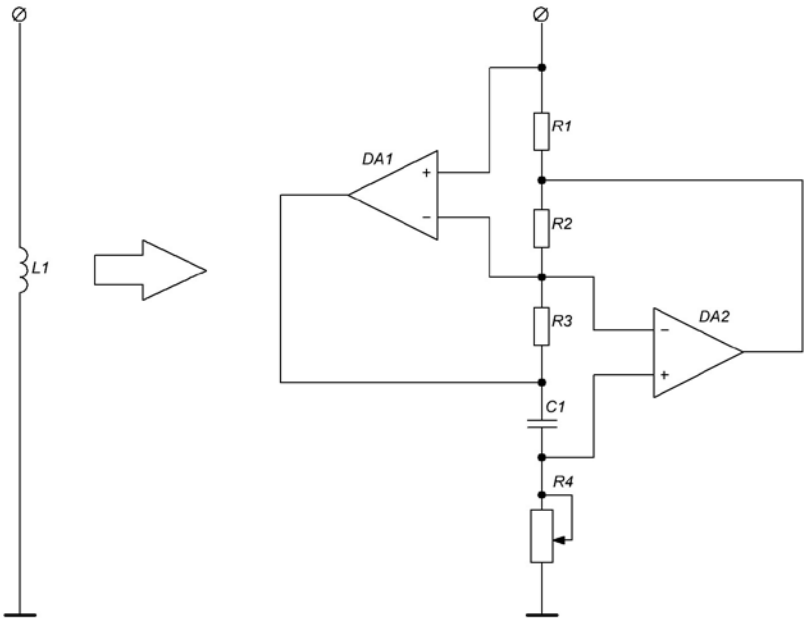


Fig. 1. Operational amplifier realization of simulated inductor

In this work, by using the Laboratory Virtual Instruments Engineering Workbench (LabVIEW), we developed an algorithm that automatically convert value of inductance to value of resistance of potentiometer R4.
The paper is organized as follows. In Sect. 2, algorithm that realize inductance equivalent, LabView software interface and practical realization are presented. In the following section, the result of experimental investigation is presented. Inductor-free simplified Chua's chaotic circuit is easily extended, and similar dynamical behaviors are exhibited through the corresponding numerical simulations and hardware experiments. The conclusionsare summarized in the last section.

2 Software and Practical Realization of theEquivalent of Inductance

Algorithm of equivalent of inductance was realized in LabView. This is a graphical programming platform that helps engineers implement all stages of development of large and small projects: from prototype creation to final testing. In this development environment, the best integration of software and hardware components with the latest computer technologies is combined today.
LabView contains all the tools for solving currentand upcoming challenges with enormous potential forinnovation, future success and effectiveness.

LabView includes powerful multi-function tools forconducting any types of measurements and development of any applications. With these tools, engineersand scientists can work in the widest range of applications and spend much less time developing. Thanks to this, LabView is a development environment for solving a wide range of research, performance enhancementsand innovations.

Fig. 2 demonstrate algorithm of equivalent of inductance that was realized in LabView

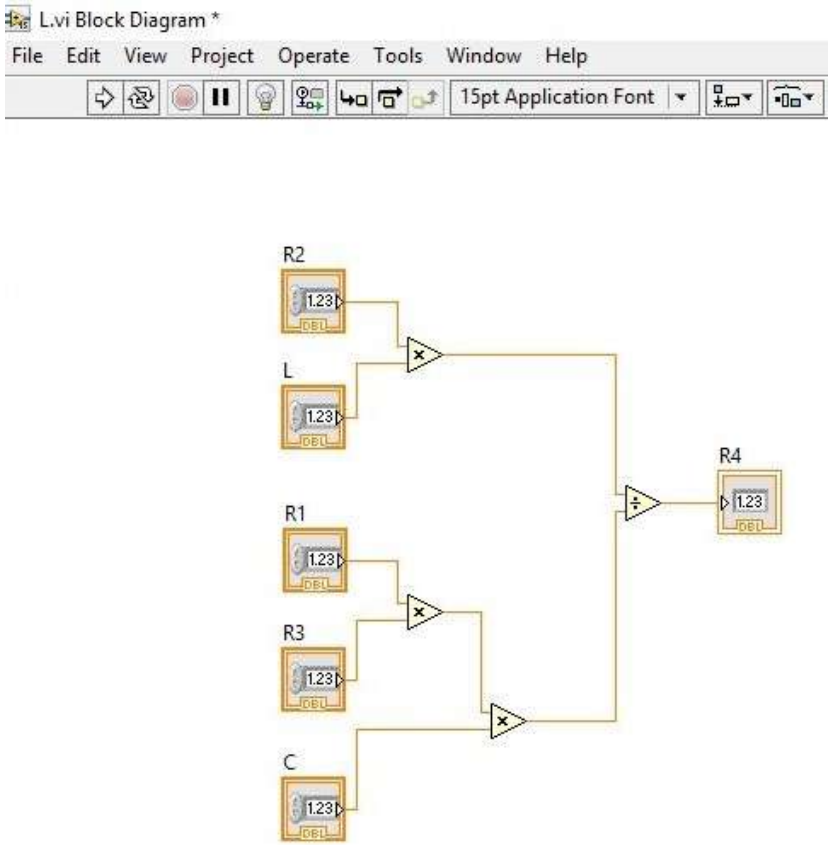


Fig. 2. Software realization of algorithm

Fig. 3 shows program interface for calculate of resistance of resistor R4.

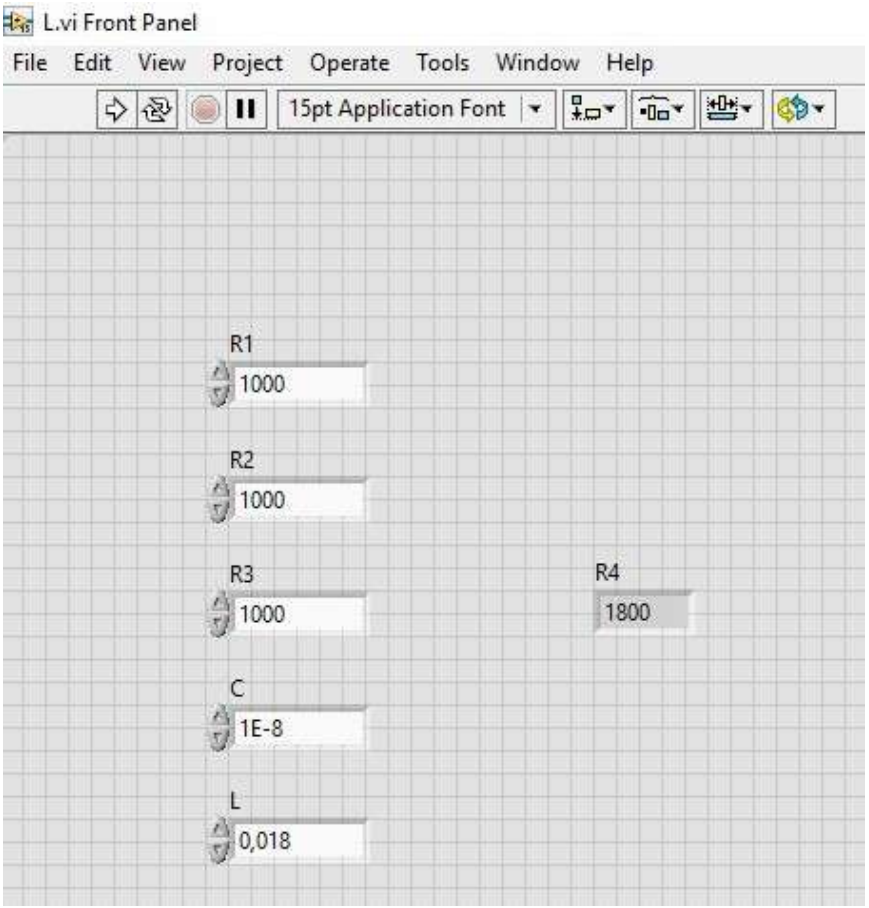


Fig. 3. Software interface for calculate of resistance of resistor R4

For example, we have next values of components:R1 = R2 = R3 = 1000 Ohm, capacitor C = 10 nF, L = 18 mH. These values we inserted in special windows of L-EQ calculator and get value for resistor R4 = 1800 Ohm. Practical realization and results are shown in Fig. 4., i.e. inductance Leq = 17.9 mH and R0 = 0.23 Ohm. Voltage source – 12 V. For power supply of the circuit was used laboratory DC power supply Hantek HT3003PB. For measurementof inductance was used LCR meter UNI-T.

Fig. 4. Practical realization of the equivalent of inductance (L = 17.9 mH)

If we changed resistance of resistor R4, we get inductance equivalent equal 20 mH. This experimental result shows in Fig. 5.

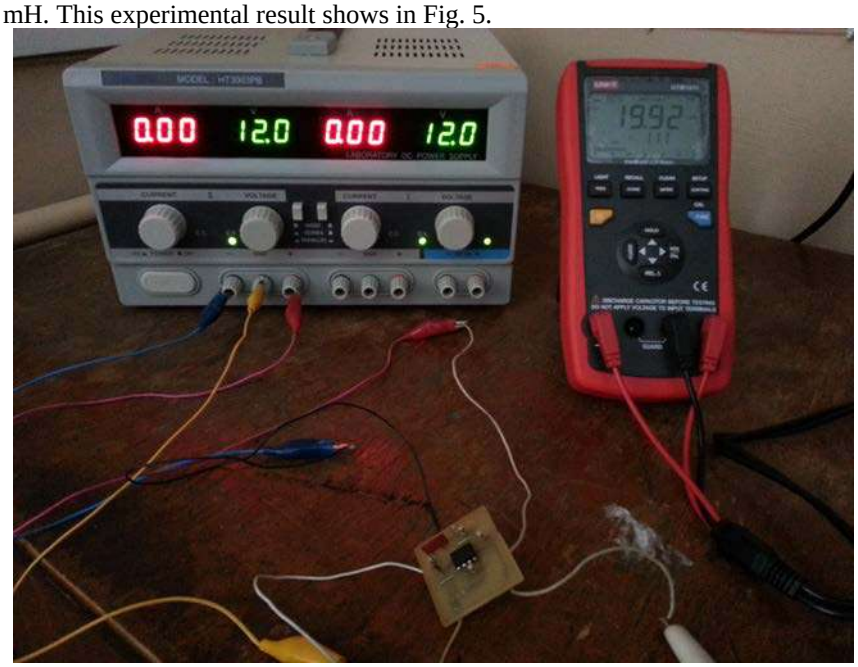


Fig. 5. Practical realization of the equivalent of inductance (L = 19.92 mH)

3 Experimental Investigation of the Equivalent of Inductance

After the advent of Chua's chaotic circuit, numerous works have been reported on different realization schemes of this circuit. We consider a realization of the double scroll chaotic Chua's attractor given by the following set of (rescaled) three coupled ODEs:

$$\begin{cases} \dfrac{dx}{dt} = \alpha\left(y - x - g(x)\right), \\ \dfrac{dy}{dt} = x - y + z, \\ \dfrac{di_L}{dt} = -\beta y, \end{cases} \tag{2}$$

where α = 10, β = 14.87, g(x) – piecewise linear function.
The circuit realization of the above is displayed in Fig. 6, with component values: capacitors C1 = 100 nF, C2 = 10 nF, DA1 – operational amplifier TL082, powered by a 12 V, GB –voltage source, inductor L1 = 18 mH, resistors R1 = 1.71 kΩ, R2 = 47 kΩ, R3 = R4 = 3.3 kΩ, R5 = 47 kΩ, R6 = R7 = 290, R8 = 1.2 kΩ, diodes VD1, VD2 – 1N4148.

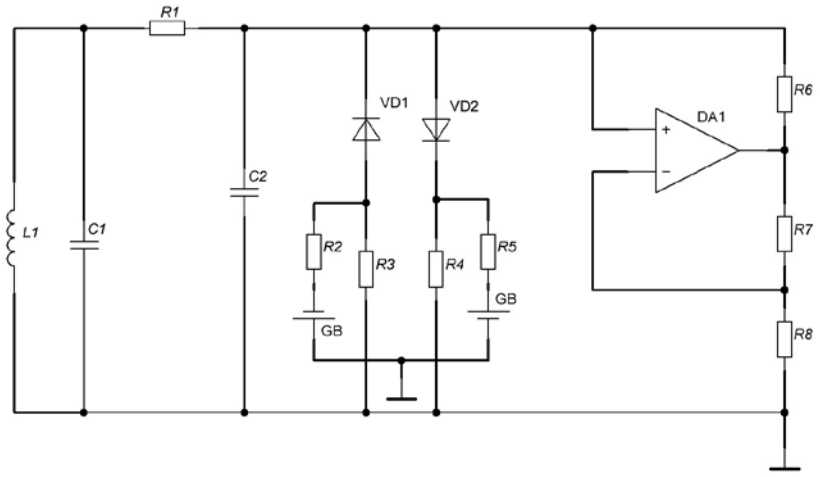


Fig. 6. The classical chaotic Chua's generator

The experimental results are captured by TektronixTDS 1002 digital oscilloscope.
Fig. 7 and Fig. 8 shows chaotic attractor and timeseries that was realized practically.

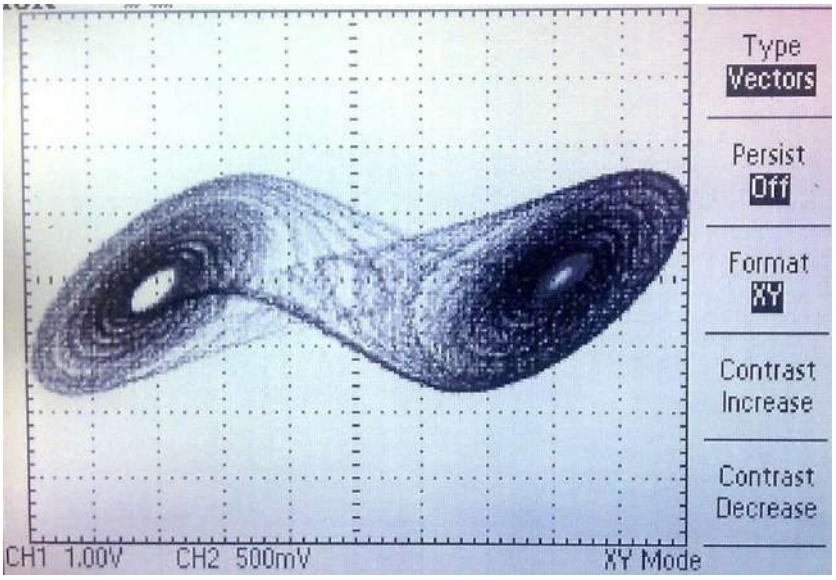


Fig. 7. Chaotic attractor

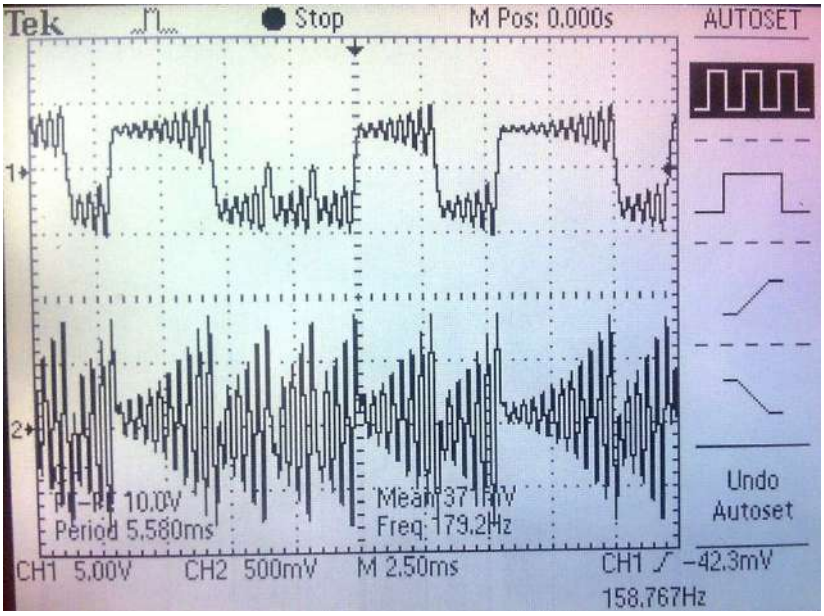


Fig. 8. The x-signal (upper) and the y-signal (lower) timeseries realized practically. Their non-periodic nature is evident.

Fig. 9 and Fig. 10 shows spectra of the classical chaotic Chua's generator.

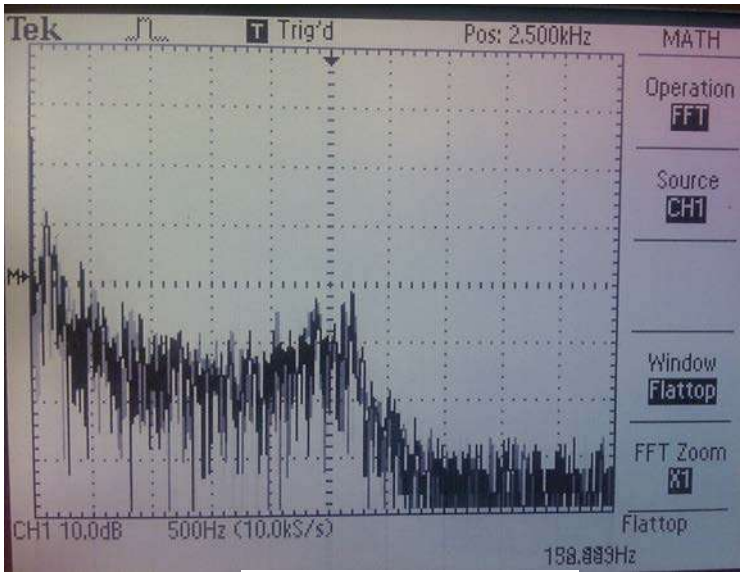


Fig. 9. The spectral distribution of V_{C1}

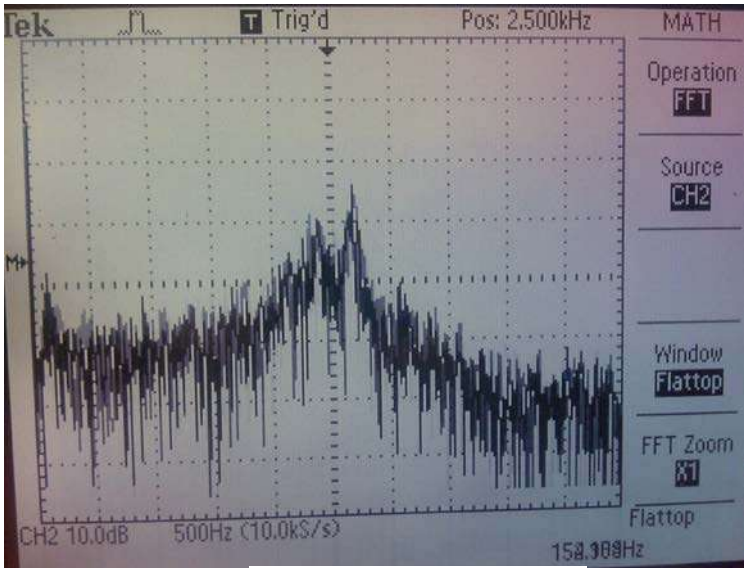


Fig. 10. The spectral distribution of V_{C2}

Conclusions

Designed L-EQ calculator allows convert of inductance values to resistance using some algorithm. Circuit realization of the inductance equivalent that contains

two operational amplifiers, one capacitor, and four resistors is presented. The mathematical equation that allow convert inductor value to resistance of potentiometer is shown. Computer modeling results of the algorithm for calculate inductance was realized in the modern software LabView. Experimental results of realization of the equivalent of inductance are presented. The designed layout was applied for chaotic Chua's generator. Chaotic attractor, timeseries and s spectral distributions are also presented.

References

1. Rhif, A., Vaidyanathan, S., Sambas, A., Mujiarto, and Subiyanto. (2019) *A Fish Biology Chaotic System and its Circuit Design*. IOP Conference Series: Journal of Physics, Vol. 1179, art. no. 012011.
2. Hajnova, V., Pribylova, L. (2017) *Two-parameter bifurcations in LPA model*. Journal of Mathematical Biology, Vol. 75, No. 5, 1235-1251.
3. Vaidyanathan, S., Feki, M., Sambas, A., Lien C. H. (2018) *A new biological snap oscillator: its modelling, analysis, simulations and circuit design*. International Journal of Simulation and Process Modelling, Vol. 13(5), 419-432.
4. Pribylova, L. (2018) *Regime shifts caused by adaptive dynamics in prey-predator models and their relationship with intraspecific*. Ecological Complexity, Vol. 36, 48-56.
5. Pribylova, L, Peniaskova, A. (2017) *Foraging facilitation among predators and its impact on the stability of predator-prey dynamics*. Ecological Complexity, Vol. 29, 30-39.
6. Rusyn V., Savko O. (2015) *Modeling of Chaotic Behavior in the Economic Model.* CHAOS 2015 – 8th Chaotic Modeling and Simulation International Conference, Proceedings 2015, 705-712.
7. Vaidyanathan S., Sambas A., Kacar S., Cavusoglu U. (2019) *A new finance chaotic system, its electronic circuit realization, passivity based synchronization and an application to voice encryption*. Nonlinear Engineering, Vol. 8, Issue 1, 193-205.
8. Lenka Pribylova (2009) *Bifurcation routes to chaos in an extended Van der Pol's equation applied to economic models.* Electronic Journal of Differential Equations, No. 53, 1-21.
9. Rusyn, V., Samila, A., Skiadas, Ch. (2020) *Computer modeling and practical realization of chaotic circuit with a light-emitting diode*. In: Fourteenth International Conference on Correlation Optics, Chernivtsi, 16-19 September 2019, pp. 113690D.
10. Skiadas, C.H. (2010) *Exact solutions of stochastic differential equations: Gompertz, generalized logistic and revised exponential*. Methodology and Computing in Applied Probability, Vol 12(2), 261-270.
11. Skiadas, C.H., Skiadas, C. (2008) Chaotic Modelling and Simulation: Analysis of Chaotic Models, Attractors and Forms. In: Taylor & Francis Group, pp. 1–345, LLC.
12. Chua, L. (1971) *Memristor – the missing circuit element.* IEEE Trans. Circuit Theory, Vol. 18, No. 5, 507-519.
13. Rusyn, V., Khrapko, S. (2019) *Memristor: Modeling and research of information properties.* Springer Proceedings in Complexity, 229-238.
14. Sambas A., Sanjaya WS. M., Mamat M., Putra Prastio R., Azar A.T. (2016) *Mathematical modelling of chaotic jerk circuit and its application in secure communication system*. Studies in Fuzziness and Soft Computing, Vol. 337, 133-153.
15. Sambas, A., Sanjaya, M. WS, Mamat, M., Diyah, H. (2013) *Design and analysis bidirectional chaotic synchronization of rossler circuit and its application for secure communication.* Applied Mathematical Sciences, Vol 7(1), 11-21.

16. Lien C.-H., Vaidyanathan S., Sambas A., Sukono, Mamat M., Sanjaya W.S.M., Subiyanto (2018) *A new two-scroll chaotic attractor with three quadratic nonlinearities, its adaptive control and circuit design*. IOP Conference Series: Materials Science and Engineering, Vol. 332, Issue 1, art. no. 012010.
17. Sambas A., Sanjaya WS. M., Mamat M., Putra Prastio R., Azar A.T. (2016) *Mathematical modelling of chaotic jerk circuit and its application in secure communication system*. Studies in Fuzziness and Soft Computing, Vol. 337, 133-153.
18. Rusyn V. B. (2014) *Modeling and Research of Chaotic Rossler System with LabView and Multisim Software Environments*. Bulletin of National Technical University of Ukraine «Kyiv Polytechnic Institute», Series Radiotechnique Radioapparatus Building, Eddition 59, 21-28.
19. Rusyn V., Stancu A., Stoleriu L. (2015) *Modeling and Control of Chaotic Multi-Scroll Jerk System in LabView*. Bulletin of National Technical University of Ukraine «Kyiv Polytechnic Institute», Series Radiotechnique Radioapparatus Building, Edition 62, 94-99.
20. Sambas A., Vaidyanathan S., Mamat M., Mada Sanjaya W.S. (2018) *A six-term novel chaotic system with hidden attractor and its circuit design*. Studies in Systems, Decision and Control, Vol. 133, 365-373.
21. Rusyn, V. B., Pavlyukevich, I., Pribylova, L. and Skiadas, H. C. (2019) *Design, modeling and research of the new non-autonomous chaotic generator*. Visnyk NTUU KPI Seriia - Radiotekhnika Radioaparatobuduvannia, 77, 13-16.
22. Rusyn V., Kushnir M., Galameiko O. (2012) *Hyperchaotic Control by Thresholding Method.* Modern Problems of Radio Engineering, Telecommunications and Computer Science - Proceedings of the 11th International Conference, TCSET'2012, art. No. 6192785, p. 67.
23. Rusyn, V., Subbotin, S. and Sambas, A. (2020) *Analysis and experimental realization of the logistic map using Arduino Pro Mini*. CEUR Workshop Proceedings, 2608, 300-310.
24. Rusyn, V., Skiadas, Ch. (2020) *Threshold Method for Control of Chaotic Oscillations*. Springer Proceedings in Complexity. Springer, pp. 217-229.
25. Rusyn V. (2017) *Modeling and Research Information Properties of Rucklidge chaotic system using LabView*. CHAOS 2017 – Proceedings: 10th Chaotic Modeling and Simulation International Conference, 739-744.
26. Rusyn, V. B., Pribylova, L. and Dimitriu, D.-G. (2018) *Control of the modified chaotic Chua's circuit using threshold method*. Visn. NTUU KPI, Radioteh. radioaparatobuduv., 75, 61-65.
27. Rusyn, V., Mohamad, M.A., Purwandari, D., Mamat, M., Titaley, J., Pinontoan, B. (2020) *Chaotic and Controlling Regimes of a New Modified Chua's Generator*. Journal of Advanced Research in Dynamical & Control Systems, Vol. 12, Issue 02, 556-561.
28. Rusyn, V., Mohamad, M., Titaley, J., Nainggolan, N., Mamat, M. (2020) *Design, computer modelling, analysis and control of the new chaotic generator*. Journal of Advanced Research in Dynamical & Control Systems, Vol. 12, Issue 02, 2306-2311.
29. Rusyn, V., Sadli, M., Mamat, M., Mujiarto, Mada Sanjaya, W.S. (2020) *Computer modelling of a new simple chaotic generator*. Journal of Physics: Conference Series, 1477, 022010.

Dynamics of a Bertrand Duopoly Game with Differentiated Goods, Heterogeneous Expectations and Relative Profit Maximization

Georges Sarafopoulos[1], Kosmas Papadopoulos[2]

[1] Department of Economics, Democritus University of Thrace, University Campus, 69100 Komotini, Greece
(E-mail: gsarafop@econ.duth.gr)
[2] Department of Economics, Democritus University of Thrace, University Campus, 69100 Komotini, Greece
(E-mail: kpapa@econ.duth.gr)

Abstract. In this article the authors investigate the dynamics of an oligopoly game in which, they consider a nonlinear Bertrand-type duopoly game with differentiated goods and heterogeneous expectations. In this study the case, where managers have a variety of attitudes toward relative performance that are indexed by their type is investigated. In this game they suppose a linear demand and cost functions. The game is modeled with a system of two difference equations. Existence and stability of equilibria of the system are studied. It is revealed that the models gives more complex, chaotic and unpredictable trajectories, as a consequence of change in the parameter k of speed of the player's adjustment, the parameter d of the horizontal product differentiation and the relative profit parameter μ. The chaotic features are justified numerically via computing Lyapunov numbers and sensitive dependence on initial conditions.

.

Keywords: Bertrand duopoly game; Relative profit maximization; Discrete dynamical system; Nash equilibrium; Stability; Bifurcation diagrams; Lyapunov numbers; Strange attractors; Chaotic Behavior.

1 Introduction

An Oligopoly is a market structure between monopoly and perfect competition, where there are only a few number of firms in the market producing homogeneous products. The dynamic of an oligopoly game is more complex because firms must consider not only the behaviors of the consumers, but also the reactions of the competitors i.e. they form expectations concerning how their rivals will act. Cournot, in 1838 has introduced the first formal theory of oligopoly. In 1883 another French mathematician Joseph Louis Francois Bertrand modified Cournot game suggesting that firms actually choose prices rather than quantities. Originally Cournot and Bertrand models were based on the premise that all players follow naive expectations, so that in every step, each player (firm) assumes the last values that were taken by the competitors without estimation of their future reactions. However, in real market conditions such an assumption is very unlikely since not all players share naive beliefs. Therefore, different approaches to firm behavior were proposed. Some authors considered

duopolies with homogeneous expectations and found a variety of complex dynamics in their games, such as appearance of strange attractors (Agiza, 1999, Agiza et al., 2002, Agliari et al., 2005, 2006, Bischi, Kopel, 2001, Kopel, 1996, Puu, 1998, Sarafopoulos, 2015, Zhang , 2009). Also models with heterogeneous agents were studied (Agiza, Elsadany , 2003, 2004, Agiza et al., 2002, Den Haan , 20013, Fanti, Gori, 2012, Tramontana, 2010, Zhang , 2007).

In the real market producers do not know the entire demand function, though it is possible that they have a perfect knowledge of technology, represented by the cost function. Hence, it is more likely that firms employ some local estimate of the demand. This issue has been previously analyzed by Baumol and Quandt, 1964, Puu 1995, Naimzada and Ricchiuti, 2008, Askar, 2013, Askar, 2014. Bounded rational players (firms) update their strategies based on discrete time periods and by using a local estimate of the marginal profit. With such local adjustment mechanism, the players are not requested to have a complete knowledge of the demand and the cost functions (Agiza, Elsadany, 2004, Naimzada, Sbragia, 2006, Zhang *et al*, 2007, Askar, 2014).

In this paper we study the dynamics of a Bertrand- type duopoly with differentiated goods where each firm behaves with heterogeneous expectations strategies. We show that the model gives more complex chaotic and unpredictable trajectories as a consequence of change in three parameters, the speed of players' adjustment, the parameter of horizontal product differentiation and the relative profit parameter. The paper is organized as follows: In Section 2, the dynamics of the duopoly game with heterogeneous expectations, linear demand and cost functions for two players are analyzed. We set first player as bounded rational and the second as a naïve player. The existence and local stability of the equilibrium points are also analyzed. In Section 3 numerical simulations are used to verify the algebraic results of Section 2 plotting the bifurcation diagrams of the game's system and to show the complex dynamics via computing Lyapunov numbers, and sensitive dependence on initial conditions.

2 The game

2.1 The construction of the game

In this study we assume that in the two companies there is a separation between ownership and management, so there is a possibility that the managers who make decisions for the company to decide at the expense of their company trying to increase the profits of the competitor. Also, we consider heterogeneous players and more specifically, we consider that the Firm 1 chooses the price of its product in a rational way, following an adjustment mechanism (bounded rational player), while the Firm 2 decides with naïve way by selecting a price that maximizes its output (naïve player). We consider a simple Bertrand-type duopoly market where firms (players) produce differentiated goods and offer

them at discrete-time periods on a common market. Price decisions are taken at discrete time periods t = 0, 1, 2,… At each period t, every firm must form an expectation of the rival's strategy in the next time period in order to determine the corresponding profit-maximizing prices for period t+1. We suppose that q_1 , q_2 are the production quantities of each firm. Also, we consider that the preferences of consumers represented by the equation:

$$U(q_1, q_2) = \alpha(q_1 + q_2) - \frac{1}{2}\left(q_1^2 + q_2^2 + 2dq_1q_2\right) \quad (1)$$

where α is a positive parameter (α > 0), which expresses the market size and $d \in (-1,1)$ is the parameter that reveals the differentiation degree of products. For example, if $d = 0$ then both products are independently and each firm participates in a monopoly. But, if $d = 1$ then one product is a substitute for the other, since the products are homogeneous. It is understood that for positive values of the parameter d the larger the value, the less diversification we have in both products. On the other hand negative values of the parameter d are described that the two products are complementary and when $d = -1$ then we have the phenomenon of full competition between the two companies. The inverse demand functions (as functions of quantities) coming from the maximizing of (1) are given by the following equations:

$$p_1(q_1, q_2) = \alpha - q_1 - dq_2 \quad \text{and} \quad p_2(q_1, q_2) = \alpha - q_2 - dq_1 \quad (2)$$

The direct demand functions (as functions of prices):

$$q_1(p_1, p_2) = \frac{\alpha(1-d) - p_1 + dp_2}{1-d^2} \quad \text{and} \quad q_2(p_1, p_2) = \frac{\alpha(1-d) - p_2 + dp_1}{1-d^2} \quad (3)$$

In this work we suppose that both players follow the same linear cost function, which is described by the following equation:

$$C_i(q_i) = c \cdot q_i \quad (4)$$

and c > 0 is the same marginal cost for two firms.

With these assumptions the profits of the firms are given by:

$$\Pi_1(p_1, p_2) = p_1q_1 - C_1(q_1) = (p_1 - c) \cdot \frac{\alpha(1-d) - p_1 + dp_2}{1-d^2} \quad (5)$$

and

$$\Pi_2(p_1, p_2) = p_1q_1 - C_2(q_2) = (p_2 - c) \cdot \frac{\alpha(1-d) - p_2 + dp_1}{1-d^2} \quad (6)$$

Then the marginal profits at the point of the strategy space are given by:

$$\frac{\partial\Pi_1}{\partial p_1} = \frac{\alpha(1-d) + c - 2p_1 + dp_2}{1-d^2} \quad , \quad \frac{\partial\Pi_1}{\partial p_2} = \frac{d(p_1 - c)}{1-d^2} \quad (7)$$

and

$$\frac{\partial\Pi_1}{\partial p_2} = \frac{\alpha(1-d) + c - 2p_2 + dp_1}{1-d^2} \quad , \quad \frac{\partial\Pi_2}{\partial p_1} = \frac{d(p_2 - c)}{1-d^2} \quad (8)$$

As it is noticed both managers care about the maximization of a utility function that contains a percentage of opponent company's profits (generalized relative profit function), which is given by:

$$U_i = (1-\mu_i)\cdot\Pi_i + \mu_i\cdot(\Pi_i - \Pi_j) = \Pi_i - \mu_i\cdot\Pi_j \tag{9}$$

where $\mu \in [0,1]$ is the percentage that the player i takes into account the opponent company's prifots. So, the marginal utility of the player i is given by the following equation:

$$\frac{\partial U_i}{\partial p_i} = \frac{\partial \Pi_i}{\partial p_i} - \mu_i\cdot\frac{\partial \Pi_j}{\partial p_i} \tag{10}$$

and the marginal utilities for each player are:

$$\frac{\partial U_1}{\partial p_1} = \frac{\alpha(1-d)+c(1+\mu d)-2p_1+d(1-\mu)p_2}{1-d^2} \tag{11}$$

and

$$\frac{\partial U_2}{\partial p_2} = \frac{\alpha(1-d)+c(1+\mu d)-2p_2+d(1-\mu)p_1}{1-d^2} \tag{12}$$

The first player is characterized as bounded rational player. According to the existing literature it means that he decides his price following a mechanism that is described by the equation:

$$\frac{p_1(t+1)-p_1(t)}{p_1(t)} = k\cdot\frac{\partial U_1}{\partial p_1}\ ,\ k>0 \tag{13}$$

Through this mechanism the player increases his level of adaptation when his marginal utility is positive or decreases his level when his marginal utility is negative, where k is the speed of adjustment of player, it is a positive parameter ($k > 0$), which gives the extend variation of price of the company 1, following a given utility signal.

The second player chooses this price that maximizes his utility function (naïve player). So, his strategy is given by the equation:

$$p_2(t+1) = \arg\max_y U_2\left(p_1(t), p_2(t)\right) \tag{14}$$

The dynamical system of the players is described by:

$$\begin{cases} p_1(t+1) = p_1(t) + k\cdot p_1(t)\cdot\dfrac{\partial U_1}{\partial p_1} \\ p_2(t+1) = \dfrac{\alpha(1-d)+c(1+\mu d)+d(1-\mu)\cdot p_1(t)}{2} \end{cases} \tag{15}$$

We will focus on the dynamics of this system to the parameter k, d and μ.

2.2 Dynamical analysis

The dynamical analysis of the discrete dynamical system involves finding equilibrium positions and studying them for stability. The ultimate goal of this algebraic study is to formulate a proposition that will be the stability condition

of the Nash Equilibrium position. Finally, these algebraic results are verified and visualized doing some numerical simulations using the program of Mathematica.

2.2.1 The equilibrium positions

The equilibriums of the dynamical system (15) are obtained as the nonnegative solutions of the algebraic system:

$$\begin{cases} p_1^* \cdot \dfrac{\partial U_1}{\partial p_1} = 0 \\ p_2^* = \dfrac{\alpha(1-d) + c(1+\mu d) + d(1-\mu)\cdot p_1^*}{2} \end{cases} \tag{16}$$

which is obtained by setting : $p_1(t+1) = p_1(t) = p_1^*$ and $p_2(t+1) = p_2(t) = p_2^*$.

- If $p_1^* = 0$ and $\dfrac{\partial U_2}{\partial p_2} = 0$ then: $p_2^* = \dfrac{\alpha(1-d) + c(1+\mu d)}{2}$ and the equilibrium position is the point:

$$E_1 = \left(0, \frac{\alpha(1-d) + c(1+\mu d)}{2}\right) \tag{17}$$

- If $\dfrac{\partial U_1}{\partial p_1} = \dfrac{\partial U_2}{\partial p_2} = 0$ then the following system is obtained:

$$\begin{cases} \alpha(1-d) + c(1+\mu d) - 2p_1^* + d(1-\mu)p_2^* = 0 \\ \alpha(1-d) + c(1+\mu d) - 2p_2^* + d(1-\mu)p_1^* = 0 \end{cases} \tag{18}$$

and the nonnegative solution of this algebraic system will give the Nash Equilibrium position $E_* = (p_1^*, p_2^*)$ where:

$$p_1^* = p_2^* = \frac{(2+d-\mu d)\cdot[\alpha(1-d) + c(1+\mu d)]}{4 - d^2(1-\mu)^2} \tag{19}$$

2.2.2 Stability of equilibrium points

To study the stability of the equilibrium positions we need the Jacobian matrix of the dynamical system Eq.(15) which is the matrix:

$$J(p_1^*, p_2^*) = \begin{bmatrix} f_{p_1} & f_{p_2} \\ g_{p_1} & g_{p_2} \end{bmatrix} \tag{20}$$

where:

$$f(p_1,p_2)=p_1+k\cdot p_1\cdot\frac{\partial U_1}{\partial p_1}$$
$$g(p_1,p_2)=\frac{\alpha(1-d)+c(1+\mu d)+d(1-\mu)\cdot p_1}{2} \tag{21}$$

and as a result the Jacobian matrix of game's discrete dynamical system Eq.(15) is the following matrix:

$$J\left(p_1^*,p_2^*\right)=\begin{bmatrix}1+k\cdot\left(\frac{\partial U_1}{\partial p_1}+p_1^*\cdot\frac{\partial^2 U_1}{\partial p_1^2}\right) & k\cdot p_1^*\cdot\frac{\partial^2 U_1}{\partial p_1\partial p_2}\\ \frac{d(1-\mu)}{2} & 0\end{bmatrix} \tag{22}$$

For the E_1 the Jacobian matrix becomes as:

$$J(E_1)=\begin{bmatrix}1+k\cdot\frac{\partial U_1}{\partial p_1} & 0\\ \frac{d\cdot(1-\mu)}{2} & 0\end{bmatrix}\begin{matrix}A=1+k\cdot\frac{\partial U_1}{\partial p_1}\\ =\\ B=\frac{d\cdot(1-\mu)}{2}\end{matrix}\begin{bmatrix}A & 0\\ B & 0\end{bmatrix} \tag{23}$$

with $Tr=A$ and $Det=0$.

From the characteristic equation of $J(E_1)$, we find the nonnegative eigenvalue:

$$r_1=Tr=1+k\cdot\frac{\partial U_1}{\partial p_1} \tag{24}$$

it's clearly seems that $|r_1|>1$ and the E_1 equilibrium is unstable.

For the E_* the Jacobian matrix becomes as:

$$J(E_*)=\begin{bmatrix}1-2k\cdot(1+c)\cdot q_1^* & -k\cdot(1-\mu)\cdot q_1^*\\ \frac{\mu-1}{2(1+c)} & 0\end{bmatrix} \tag{25}$$

with

$$Tr=1-2k\cdot(1+c)\cdot q_1^* \qquad \text{and} \qquad Det=-k\cdot\frac{(1-\mu)^2}{2(1+c)}\cdot q_1^* \tag{26}$$

To study the stability of Nash equilibrium we use three conditions that the equilibrium position is locally asymptotically stable when they are satisfied simultaneously:

$$\begin{aligned}&\text{(i)} && 1-Det>0\\ &\text{(ii)} && 1-Tr+Det>0\\ &\text{(iii)} && 1+Tr+Det>0\end{aligned} \tag{27}$$

It's easy to find that the first condition (i) is always satisfied:

$$1-\mathrm{Det}>0 \;\Leftrightarrow\; 1+k\cdot p_1^*\cdot\frac{d^2\cdot(1-\mu)^2}{2\left(1-d^2\right)}>0 \tag{28}$$

Also, the condition (ii) gives:

$$1-\mathrm{Tr}+\mathrm{Det}>0 \;\Leftrightarrow\; k\cdot p_1^*\cdot\frac{\left[4-d^2\cdot(1-\mu)^2\right]}{2\left(1-d^2\right)}>0 \tag{29}$$

and it's always satisfied because $\dfrac{\left[4-d^2\cdot(1-\mu)^2\right]}{2\left(1-d^2\right)}>0$.

Finally, the condition (iii) becomes as:

$$1+\mathrm{Tr}+\mathrm{Det}>0 \;\Leftrightarrow\; k\cdot p_1^*\cdot\frac{\left[4+d^2\cdot(1-\mu)^2\right]}{2\left(1-d^2\right)}-2<0 \tag{30}$$

Proposition:
The Nash equilibrium of the discrete dynamical system Eq.(15) is locally asymptotically stable if:

$$k\cdot p_1^*\cdot\frac{\left[4+d^2\cdot(1-\mu)^2\right]}{2\left(1-d^2\right)}-2<0$$

where

$$p_1^*=\frac{(2+d-\mu d)\cdot\left[\alpha(1-d)+c(1+\mu d)\right]}{4-d^2\cdot(1-\mu)^2}$$

3 Numerical simulations

3.1 Stability spaces

At first the 3D stability space (Fig.1) is made including the main three parameters we will focus on, the parameters k (speed of adjustment), the parameter d (product's differentiation degree) and μ (relative profit parameter). This three-dimensional space is obtained by the stability condition that is described above in Proposition, setting specific values for the other parameters $\alpha=5$ and $c=1$. Also, the two-dimensional stability region for a couple of these three parameters are presented that is resulted setting specific values of for one of the three main parameters. Specifically, the stability regions between the parameters k and μ (Fig.2) also, between the parameters d and μ (Fig.3) and between the parameters k and d (Fig.4) are presented. A useful result using the Fig.2 is that for small values of the parameter k (speed of adjustment) there is a

locally asymptotically stable Nash Equilibrium for every value of the parameter μ (relative profit parameter) into the interval [0,1]. It means that for these values of the other parameters α, c, d and k the parameter μ cannot destabilize the economy.

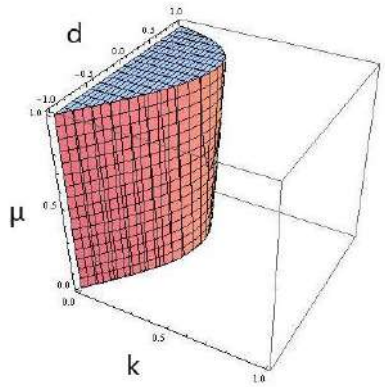


Fig. 1. Three-dimensional stability space between the parameters k, d and μ for α = 5 and c = 1.

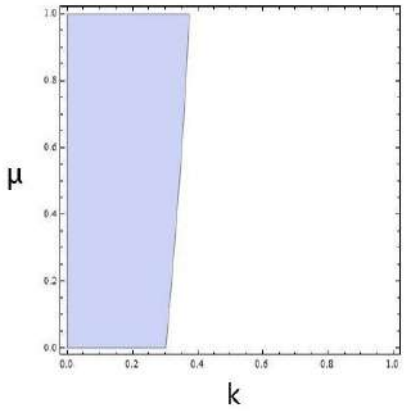


Fig. 2. Two-dimensional stability region between the parameters k (horizontal axis) and μ (vertical axis) for α = 5, c =1 and d = 0.50.

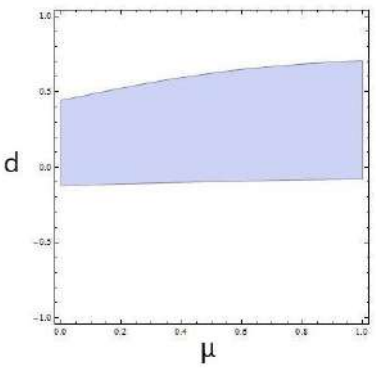


Fig. 3. Two-dimensional stability region between the parameters μ (horizontal axis) and d (vertical axis) for α = 5, c = 1 and k = 0.315.

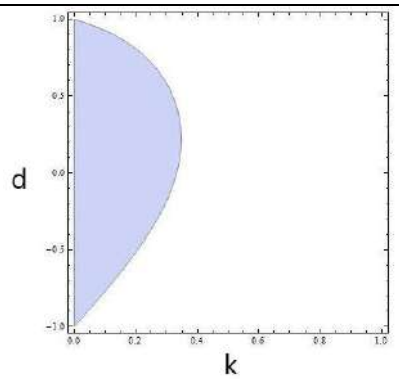


Fig. 4. Two-dimensional stability region between the parameters k (horizontal axis) and d (vertical axis) for α = 5, c = 1 and μ = 0.20.

3.2 Focusing on the parameter k

In this section some numerical simulation including bifurcation diagrams, strange attractors, Lyapunov numbers graph and Sensitive dependence on initial conditions are presented focusing on the parameter k when the other parameters are fixed taking the values: $\alpha = 5$, $c = 1$, $\mu = 0.30$ and $d = 0.50$. At first, the Nash Equilibrium for the values of these parameters becomes as:

$$p_1^* = p_2^* \approx 2.21 \Rightarrow E_*\left(p_1^*, p_2^*\right) \equiv E_*\left(2.21, 2.21\right) \tag{31}$$

and for the stability condition it means that the parameter k must take values into the interval:

$$k \in \left(0, 0.32\right) \tag{32}$$

This algebraic result is verified by the bifurcation diagrams of p_1^* (Fig.5) and p_2^* (Fig.6) with respect to the parameter k. As it seems there is a locally asymptotically stable orbit until the value of 0.32 for the parameter k and after this value doubling period bifurcations are appeared and finally, for higher values of the parameter k the system's behavior becomes chaotic and unpredictable.

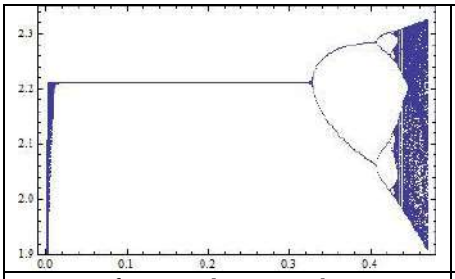


Fig. 5. Bifurcation diagram with respect to the parameter k against the variable p_1^* with 400 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $d = -0.50$ and $\mu = 0.30$.

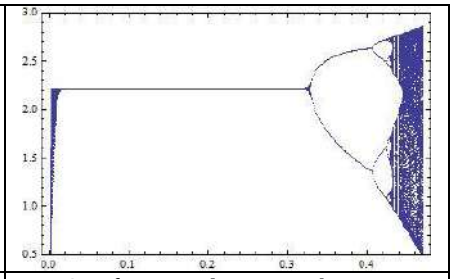


Fig. 6. Bifurcation diagram with respect to the parameter k against the variable p_2^* with 400 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $d = -0.50$ and $\mu = 0.30$.

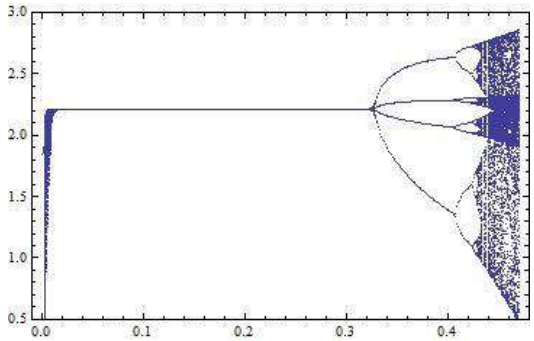


Fig.7 The two previous bifurcation diagrams of Fig.5 and Fig.6 in one.

This chaotic trajectory can create strange attractors (Fig.8) for a higher value of the parameter k like 0.47, outside the stability space. Also, computing the Lyapunov numbers (Fig.9) for this value of the parameter k and setting the same fixed values for the other parameters α, c, μ and d it seems that they are getting over the value of 1 as an evidence for the chaotic trajectory.

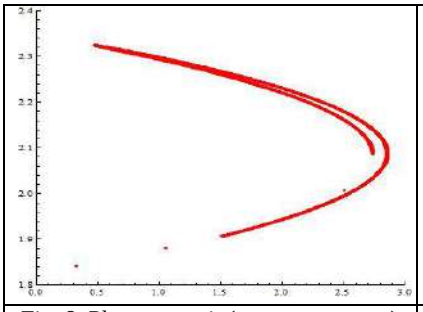


Fig. 8. Phase portrait (strange attractor) of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for α = 5, c = 1, d = -0.50, μ = 0.30 and k = 0.47.

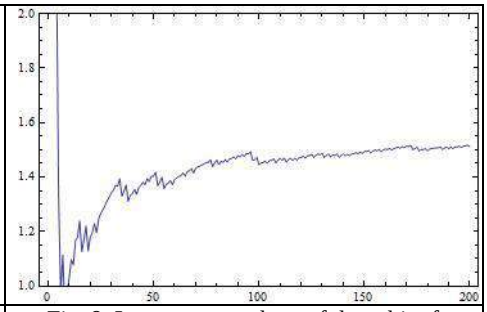


Fig. 9. Lyapunov numbers of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for α = 5, c = 1, d = -0.50, μ = 0.30 and k = 0.47.

This chaotic trajectory makes the system sensitive on initial conditions, which means that only a small change on a coordinate may change completely the system's behavior. For example, choosing two different initial conditions (0.1,0.1) (Fig.10) and (0.101,0.1) (Fig.11) with a small change at the p_1^*-coordinate and plotting the time series of system it seems that at the beginning the time series are indistinguishable, but after a number of iterations, the difference between them builds up rapidly.

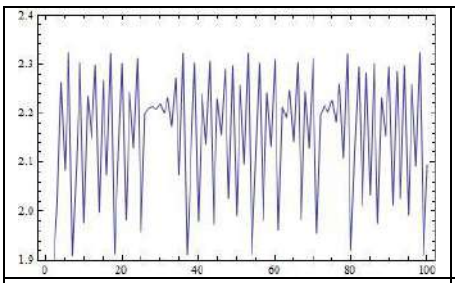


Fig. 10. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.1,0.1) of the system Eq.(15) for α = 5, c = 1, d = -0.50, μ = 0.30 and k = 0.47.

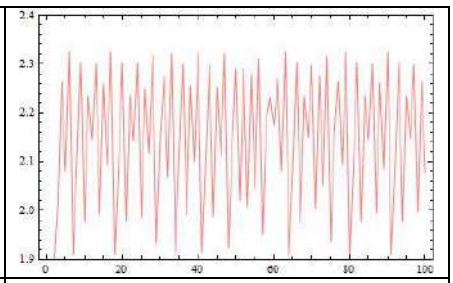


Fig. 11. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.101,0.1) of the system Eq.(15) for α = 5, c = 1, d = -0.50, μ = 0.30 and k = 0.47.

3.3 Focusing to the parameter d

Using the stability region of Fig. between the parameters μ (horizontal axis) and d (vertical axis) for α = 5, c = 1 and k = 0.315, it seems that when the parameter d takes values into a close interval there is a stable Nash equilibrium for every value of the parameter μ. For example, setting the value of 0.20 to the parameter μ, a stable Nash Equilibrium is appeared into the interval (-0.20,0.55) for the parameter d. This indication is verified by the following bifurcation diagrams of p_1^* (Fig.12) and p_2^* (Fig.13) with respect to the parameter d.

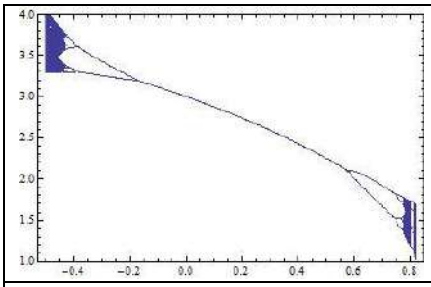


Fig. 12. Bifurcation diagram with respect to the parameter d against the variable p_1^* with 400 iterations of the map Eq.(15) for α = 5, c = 1, k = 0.30 and μ = 0.20.

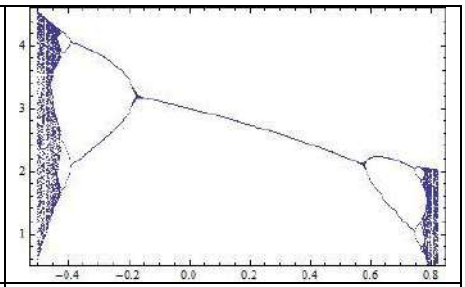


Fig. 13. Bifurcation diagram with respect to the parameter d against the variable p_2^* with 400 iterations of the map Eq.(15) for α = 5, c = 1, k = 0.30 and μ = 0.20.

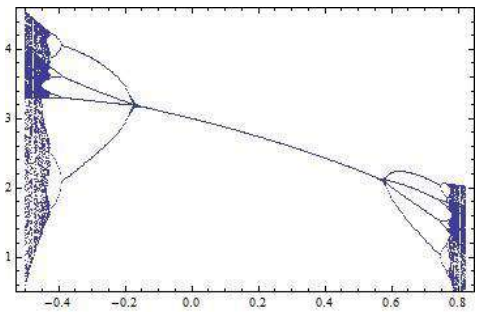


Fig.14 The two previous bifurcation diagrams of Fig.12 and Fig.13 in one.

Setting large and small values to the parameter d, strange attractors (Fig.15, Fig.19) and Lyapunov numbers (Fig.16, Fig.20) higher than the number of 1 are appeared showing the chaotic trajectories and unpredictable behavior of the system of Eq.(15) for these values of the parameter d = 0.79 and d = -0.50 outside the stability space. Finally, for these values of the parameter d the system becomes sensitive on initial conditions a result that is revealed by the sensitive dependence on initial conditions with a small change at the first coordinate and plotting the time series of system of Eq.(15) for d = 0.79 (Fig.17, Fig.18) and for d = -0.50 (Fig.21, Fig.22).

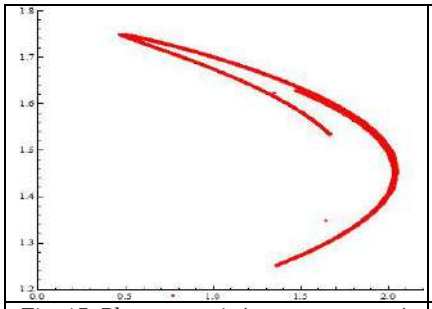


Fig. 15. Phase portrait (strange attractor) of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = 0.79$.

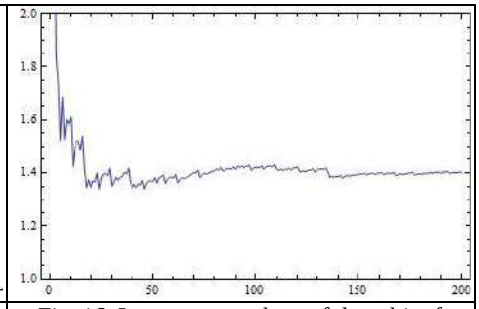


Fig. 16. Lyapunov numbers of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = 0.79$.

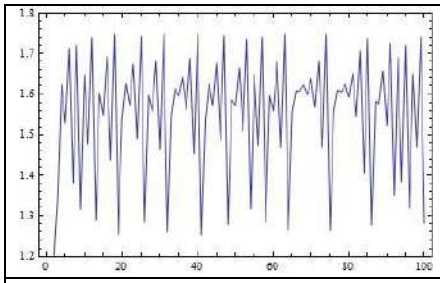


Fig. 17. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.1,0.1) of the system Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = 0.79$.

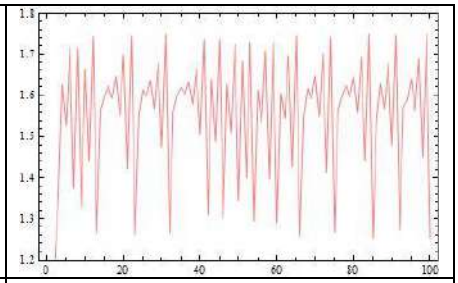


Fig. 18. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.101,0.1) of the system Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = 0.79$.

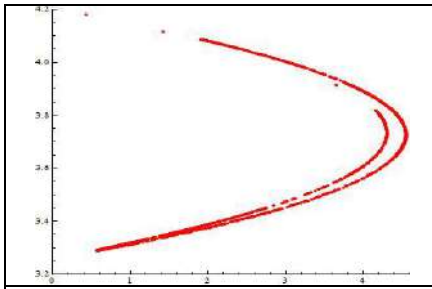


Fig. 19. Phase portrait (strange attractor) of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = -0.50$.

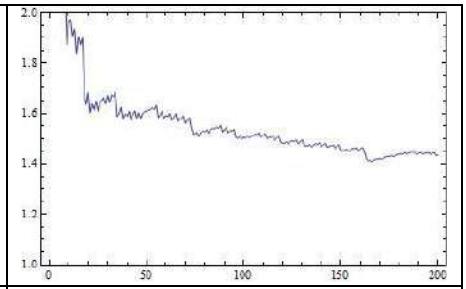


Fig. 20. Lyapunov numbers of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = -0.50$.

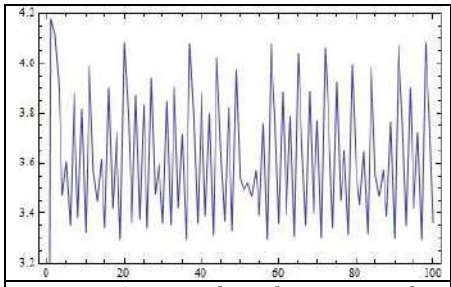


Fig. 21. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.1,0.1) of the system Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = -0.50$.

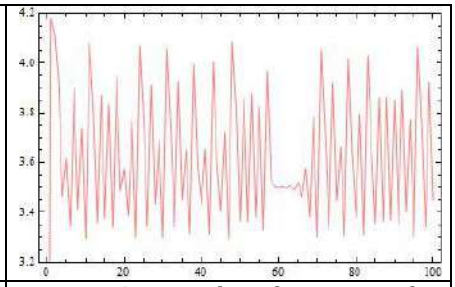


Fig. 22. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.101,0.1) of the system Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.30$, $\mu = 0.20$ and $d = -0.50$.

3.4 Focusing to the parameter μ

Using the same methods of numerical simulations focusing to the parameter μ (relative profit parameter) it is shown that small values of this parameter can destabilize the economy through doubling bifurcation diagrams (Fig.23, Fig.24, Fig.25) and strange attractors (Fig.26) and Lyapunov numbers (Fig.27) higher than the number of 1 are appeared. Also, the system of Eq.(15) becomes sensitive on small changes of initial conditions for small values of the parameter μ ($\mu = 0.05$) and it is revealed by the Fig.28 and Fig.29 of time series at first setting to the system the initial conditions of (0.1,0.1) and secondly of (0.101,0.1). As it seems, at the beginning the time series are indistinguishable, but after a number of iterations, the difference between them builds up rapidly.

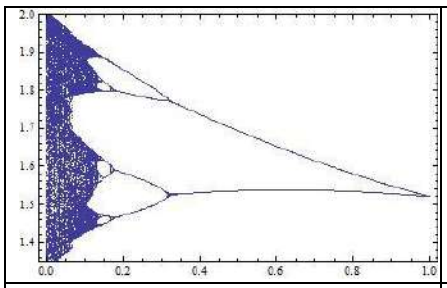


Fig. 23. Bifurcation diagram with respect to the parameter d against the variable p_1^* with 400 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.315$ and $d = 0.74$.

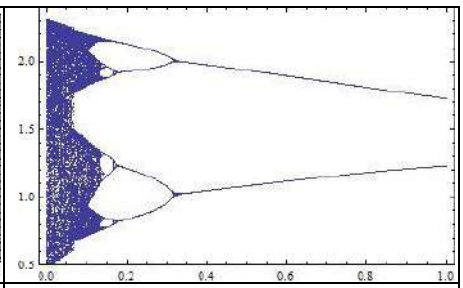


Fig. 24. Bifurcation diagram with respect to the parameter d against the variable p_2^* with 400 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.315$ and $d = 0.74$.

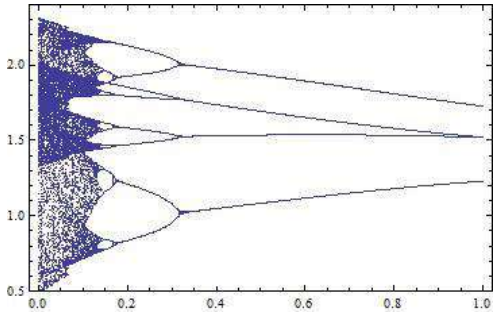


Fig.25. The two previous bifurcation diagrams of Fig.23 and Fig.24 in one.

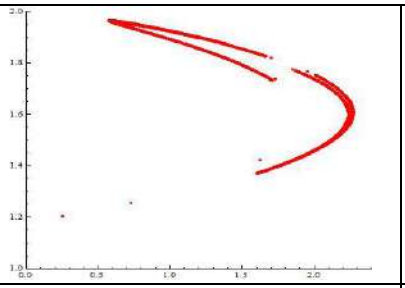


Fig. 26. Phase portrait (strange attractor) of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.315$, $d = 0.74$ and $\mu = 0.05$.

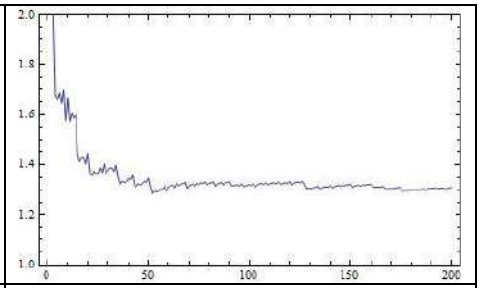


Fig. 27. Lyapunov numbers of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.315$, $d = 0.74$ and $\mu = 0.05$.

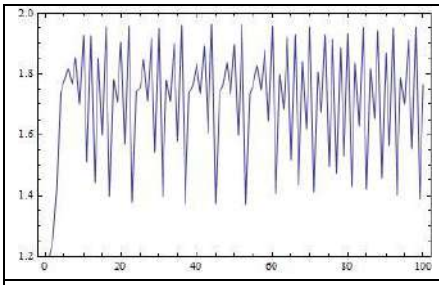


Fig. 28. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.1,0.1) of the system Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.315$, $d = 0.74$ and $\mu = 0.05$.

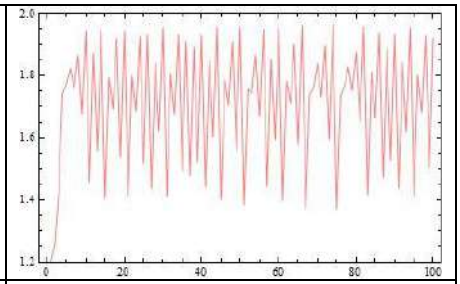


Fig. 29. Sensitive dependence on initial conditions for p_1^*-coordinate plotted against the time: the orbit of (0.101,0.1) of the system Eq.(15) for $\alpha = 5$, $c = 1$, $k = 0.315$, $d = 0.74$ and $\mu = 0.05$.

Conclusions

In this paper we analyzed the dynamics of a differentiated Bertrand duopoly with heterogeneous expectations, linear demand and cost functions. By assuming that at each time period each firm maximizes its expected relative profit under different expectations, a discrete dynamical system was obtained. Existence and stability of equilibrium of this system are studied. We showed numerically that the model gives chaotic and unpredictable trajectories. The main result is that a lower and higher degree of product differentiation relative may destabilize the Bertrand–Nash equilibrium. Also, this instability can be appeared for higher values of the speed of adjustment and lower values of the relative profit parameter. Finally, we showed also that for lower values of the speed of adjustment the equilibrium is stable for each value of the relative profit parameter.

References

1. Agiza HN, Explicit stability zones for Cournot games with 3 and 4 competitors. *Chaos Solitons Fract*. 9: 1955-66, (1998).
2. Agiza, H.N., On the analysis of stability, bifurcation, chaos and chaos control of Kopel map. *Chaos, Solitons & Fractals* 10, 1909-1916, (1999).
3. Agiza HN., On the stability, bifurcations, chaos and chaos control of Kopel map. *Chaos Solitons Fract.* 11: 1909-16, (2004).
4. Agiza HN, Elsadany AA, Chaotic dynamics in nonlinear duopoly game with heterogeneous players. *Appl. Math. Comput*. 149: 843–60.
5. Agiza HN, Elsadany AA., Nonlinear dynamics in the Cournot duopoly game with heterogeneous players. *Physica A* 320: 512–24, (2003).
6. Agiza HN, Hegazi AS, Elsadany AA., Complex dynamics and synchronization of duopoly game with bounded rationality. *Math. Comput. Simulat*. 58: 133–46, (2002).
7. Askar, S.S., On complex dynamics of monopoly market, *Economic Modelling*, 31: 586-589, (2013).
8. Askar, S. S., Complex dynamic properties of Cournot duopoly games with convex and log-concave demand function, *Operations Research Letters* 42, 85–90, (2014).
9. Baumol, W.J., Quandt, R.E., Rules of thumb and optimally imperfect decisions, *American Economic Review* 54 (2): 23–46, (1964).
10. Bischi GI, Kopel M., Equilibrium selection in a nonlinear duopoly game with adaptive expectations. *J. Econom Behav. Org*. 46: 73–100 (2001).
11. Bischi GI, Lamantia F, Sbragia L., Competition and cooperation in natural resources exploitation: an evolutionary game approach. In: Cararro C, Fragnelli V, editors. *Game practice and the environment*. Cheltenham: Edward Elgar; 187–211, (2004).
12. Bischi GI, Naimzada A., Global analysis of a dynamic duopoly game with bounded rationality. In: Filar JA, Gaitsgory V, Mizukami K, editors. *Advances in dynamic games and applications*, vol. 5. Basel: Birkhauser; 361–85, (2000).
13. Bischi, G.I., Naimzada, A.K., Sbragia, L., Oligopoly games with local monopolistic approximation, *Journal of Economic Behavior and Organization* 62 (3): 371–388, (2007).
14. Bischi, G.I., Kopel, M., Equilibrium selection in a nonlinear duopoly game with adaptive expectations. *Journal of Economic Behavior & Organization* 46, 73–100, (2001).

15. Bischi, G.I., Naimzada, A., Global analysis of a dynamic duopoly game with bounded rationality. In: Filar JA, Gaitsgory V, Mizukami K, (eds). (2000). *Advances in dynamic games and applications*, vol. 5. Basel: Birkhauser: 361-385, (1999).
16. Cournot *A. Researches into the mathematical principles of the theory of wealth.* Homewood (IL): Irwin; (1963).
17. Day, R., *Complex Economic Dynamics*. MIT Press, Cambridge, 1994.
18. Dixit, A. K., Comparative statics for oligopoly, *Internat. Econom. Rev.* 27, 107–122, (1986).
19. Dixit, A.K., A model of duopoly suggesting a theory of entry barriers. Bell Journal of Economics 10, 20–32, (1979).
20. Den Haan W.J., The importance of the number of different agents in a heterogeneous asset-pricing model. *J. Econom. Dynam. Control*, 25:721–46, (2001).
21. Elsadany A.A., Dynamics of a Cournot duopoly game with bounded rationality based on relative profit maximization. *Applied Mathematics and Computation* 294, 253–263, (2017).
22. Elaydi, S., *An Introduction to Difference Equations*, third ed., Springer-Verlag, New York, 2005.
23. Junhai Ma, Zhanbing Guo, The influence of information on the stability of a dynamic Bertrand game, *Commun Nonlinear Sci Numer Simulat* 30: 32-44, (2016).
24. Hommes, C.H., *Heterogeneous agent models in economics and finance*, in: L. Tesfatsion, K.L. Judd (Eds.), Handbook of Computational Economics, Agent-Based Computational Economics, vol. 2, Elsevier Science B.V: 1109–1186, (2006).
25. Gandolfo G., *Economic dynamics*. Berlin: Springer, 1997.
26. Gao Y., Complex dynamics in a two dimensional noninvertible map. *Chaos Solitons Fract.* 39: 1798–810, (2009).
27. Kopel M., Simple and complex adjustment dynamics in Cournot duopoly models. *Chaos Solitons Fract.* 12: 2031–48, (1996).
28. Kulenonic, M., Merino, O., *Discrete Dynamical Systems and Difference Equations with Mathematica*, Chapman & Hall/Crc., 2002.
29. Medio A, Gallo G., *Chaotic dynamics: theory and applications to economics.* Cambridge (MA): Cambridge University Press, 1995.
30. Medio A, Lines M., Introductory notes on the dynamics of linear and linearized systems. In: Lines M, editor. *Nonlinear dynamical systems in economics.* SpringerWienNewYork: CISM; 1–26, (2005).
31. Medio A, Lines M., *Nonlinear dynamics. A primer*. Cambridge (MA): Cambridge University Press, 2001.
32. Naimzada, A.K., Ricchiuti G., Complex dynamics in a monopoly with a rule of thumb, *Applied Mathematics and Computation* 203: 921–925, (2008).
33. Naimzada, A., Sbragia, L., Oligopoly games with nonlinear demand and cost functions: two boundedly rational adjustment processes, *Chaos Solitons Fractals* 29, 707–722, (2006).
34. Nakamura,Y., Price versus quantity in a mixed duopoly: The case of relative profit maximization. *Economic Modelling* 44,37-43, (2015).
35. Puu, T., The chaotic monopolist, *Chaos, Solitons & Fractals* 5 (1): 35–44, (1995).
36. Puu T., The chaotic duopolists revisited. *J Econom. Behav. Org.* 37: 385–94, (1998).
37. Puu T., Chaos in duopoly pricing. *Chaos Solitons Fract.*1:573–81, (1991).
38. Puu T., Complex oligopoly dynamics. In: Lines M, editor. *Nonlinear dynamical systems in economics.* Springer Wien NewYork: CISM; p. 165–86, (2005).
39. Sarafopoulos G., On the dynamics of a duopoly game with differentiated goods, *Procedia Economics and Finance*, 19, 146 – 153, (2015).

40. Sarafopoulos G., Complexity in a duopoly game with homogeneous players, convex, log linear demand and quadratic cost functions, *Procedia Economics and Finance*, 33, 358 – 366, (2015).
41. Sarafopoulos, G., Papadopoulos K., On a Cournot duopoly game with differentiated goods, heterogeneous expectations and a cost function including emission costs, *Scientific Bulletin - Economic Sciences* **, 1** (1), 11-22, (2017).
42. Sarafopoulos, G., Papadopoulos K., On the stability of a Cournot dynamic game under the influence of information, *Knowledge E Publishing*, **ISSN:** 2518-668X, 131-144, (2018).
43. Sarafopoulos, G., Papadopoulos K., Chaos in oligopoly models, *International Journal of Productivity Management and Assessment Technologies*, **7** (1), 50-76, (2019).
44. Sarafopoulos, G., Papadopoulos K., Complexity in Bertrand duopoly game with heterogeneous players and differentiated goods, *Economic and Financial Challenges for Eastern Europe*, Article 2 (p. 15-26), Springer, 2019.
45. Satoh A., Tanaka Y., Relative Profit Maximization and Equivalence of Cournot and Bertrand Equilibria in Asymmetric Duopoly, MPRA Paper no. 56031.
46. Sedaghat, H., Nonlinear *Difference Equations: Theory with Applications to Social Science Models*, Kluwer Academic Publishers (now Springer), 2003.
47. Singh, N., Vives, X., Price and quantity competition in a differentiated duopoly. *The RAND Journal of Economics* 15, 546–554, (1984).
48. Tramontana, F., Heterogeneous duopoly with isoelastic demand function. *Economic Modelling* 27, 350–357, (2010).
49. Westerhoff, F., *Nonlinear expectation formation, endogenous business cycles and stylized facts*, Studies in Nonlinear Dynamics and Econometrics 10 (4) (Article 4), (2016).
50. Wu, W., Chen, Z., Ip, W.H., Complex nonlinear dynamics and controlling chaos in a Cournot duopoly economic model. *Nonlinear Analysis: Real World Applications* 11, 4363–4377, (2010).
51. Zhang, J., Da, Q.,Wang, Y., Analysis of nonlinear duopoly game with heterogeneous players. *Economic Modelling* 24, 138–148, (2007).
52. Zhang J., Da Q., Wang Y., The dynamics of Bertrand model with bounded rationality, *Chaos, Solitons and Fractals* 39, 2048-2055, (2009).

On a Cournot Dynamic Game with Cost Uncertainty and Relative Profit Maximization

Georges Sarafopoulos[1], Kosmas Papadopoulos[2]

[1] Department of Economics, Democritus University of Thrace, University Campus, 69100 Komotini, Greece
(E-mail: gsarafop@econ.duth.gr)
[2] Department of Economics, Democritus University of Thrace, University Campus, 69100 Komotini, Greece
(E-mail: kpapa@econ.duth.gr)

Abstract. In this paper, a Cournot duopoly model with homogeneous goods is examined with uncertain cost function. A random linear cost function is introduced in this model for the first player. The case of homogeneous expectations is studied. The existence and uniqueness of the equilibrium are obtained. The asymptotic behavior of the equilibrium point is also investigated. Complete stability and bifurcation analysis are provided. The obtained theoretical results are verified by numerical simulations.

.

Keywords: Cournot duopoly game; Cost uncertainty; Relative profit maximization; Discrete dynamical system; Nash equilibrium; Stability; Bifurcation diagrams; Lyapunov numbers; Strange attractors; Chaotic Behavior.

1 Introduction

An Oligopoly is a market structure between monopoly and perfect competition, where there are only a few number of firms in the market producing homogeneous products. The dynamic of an oligopoly game is more complex because firms must consider not only the behaviors of the consumers, but also the reactions of the competitors i.e. they form expectations concerning how their rivals will act. Cournot, in 1838 has introduced the first formal theory of oligopoly. In 1883 another French mathematician Joseph Louis Francois Bertrand modified Cournot game suggesting that firms actually choose prices rather than quantities. Originally Cournot and Bertrand models were based on the premise that all players follow naive expectations, so that in every step, each player (firm) assumes the last values that were taken by the competitors without estimation of their future reactions. However, in real market conditions such an assumption is very unlikely since not all players share naive beliefs. Therefore, different approaches to firm behavior were proposed. Some authors considered duopolies with homogeneous expectations and found a variety of complex dynamics in their games, such as appearance of strange attractors (Agiza, 1999, Agiza et al., 2002, Agliari et al., 2005, 2006, Bischi, Kopel, 2001, Kopel, 1996, Puu, 1998, Sarafopoulos, 2015, Zhang , 2009). Also models with heterogeneous agents were studied (Agiza, Elsadany , 2003, 2004, Agiza et al., 2002, Den Haan , 20013, Fanti, Gori, 2012, Tramontana, 2010, Zhang , 2007).

In the real market producers do not know the entire demand function, though it is possible that they have a perfect knowledge of technology, represented by the cost function. Hence, it is more likely that firms employ some local estimate of the demand. This issue has been previously analyzed by Baumol and Quandt, 1964, Puu 1995, Naimzada and Ricchiuti, 2008, Askar, 2013, Askar, 2014. Bounded rational players (firms) update their strategies based on discrete time periods and by using a local estimate of the marginal profit. With such local adjustment mechanism, the players are not requested to have a complete knowledge of the demand and the cost functions (Agiza, Elsadany, 2004, Naimzada, Sbragia, 2006, Zhang *et al*, 2007, Askar, 2014).

In this paper we study the dynamics of a Cournot-type duopoly with homogeneous goods where each firm behaves with homogeneous expectations. We show that the model gives more complex chaotic and unpredictable trajectories as a consequence of change in the speed of players' adjustment. The paper is organized as follows: In Section 2, the dynamics of the duopoly game with homogeneous expectations, linear demand and cost functions and relative profit functions for two players are analyzed. A cost uncertainty is introduced into first player's utility function. We set both players as bounded rational players. The existence and local stability of the equilibrium points are also analyzed. In Section 3 numerical simulations are used to verify the algebraic results of Section 2 plotting the bifurcation diagrams of the game's system and to show the complex dynamics via computing Lyapunov numbers, and sensitive dependence on initial conditions.

2 The game

2.1 The construction of the game

In this study we assume that in the two companies there is a separation between ownership and management, so there is a possibility that the managers who make decisions for the company to decide at the expense of their company trying to increase the profits of the competitor. Also, we consider homogeneous players and more specifically, we consider that both firms choose the quantity of their productions in a rational way, following an adjustment mechanism (bounded rational players). We consider a simple Cournot-type duopoly market where firms (players) produce the same good and offer it at discrete-time periods on a common market. Production decisions are taken at discrete time periods t = 0, 1, 2,… At each period t, every firm must form an expectation of the rival's strategy in the next time period in order to determine the corresponding profit-maximizing prices for period t+1. We suppose that q_1 , q_2 are the production quantities of each firm. Also, we consider that the preferences of consumers represented by the equation:

$$U(q_1, q_2) = \alpha(q_1 + q_2) - \frac{1}{2}\left(q_1^2 + q_2^2 + 2dq_1q_2\right) \quad (1)$$

where α is a positive parameter (α > 0), which expresses the market size and $d \in (-1,1)$ is the parameter that reveals the differentiation degree of products. For example, if $d = 0$ then both products are independently and each firm participates in a monopoly. But, if $d = 1$ then one product is a substitute for the other, since the products are homogeneous. It is understood that for positive values of the parameter d the larger the value, the less diversification we have in both products. On the other hand negative values of the parameter d are described that the two products are complementary and when $d = -1$ then we have the phenomenon of full competition between the two companies. The inverse demand functions (as functions of quantities) coming from the maximizing of (1) are given by the following equations:

$$p_1(q_1, q_2) = \alpha - q_1 - q_2 \quad \text{and} \quad p_2(q_1, q_2) = \alpha - q_2 - q_1 \tag{2}$$

In this work we suppose that the first player's cost function contains an uncertainty by which the marginal cost (linear cost function) is equal to the combination between the parameters: $c_1, c_2 > 0$, which is described by the following equation:

$$C_1(q_1) = [p \cdot c_1 + (1-p) \cdot c_2] \cdot q_1 \tag{3}$$

where $p \in [0,1]$, is the positive uncertainty cost parameter.

On the other hand the second player uses a simple linear cost function that its marginal cost is equal to $c_1 > 0$ and it is described by the equation:

$$C_2(q_2) = c_1 \cdot q_2 \tag{4}$$

With these assumptions the profits of the firms are given by:

$$\Pi_1(q_1, q_2) = p_1 \cdot q_1 - C_1(q_1) = [\alpha - q_1 - q_2 - p \cdot c_1 - (1-p) \cdot c_2] \cdot q_1 \tag{5}$$

and

$$\Pi_2(q_1, q_2) = p_2 \cdot q_2 - C_2(q_2) = [\alpha - c_1 - q_1 - q_2] q_2 \tag{6}$$

Then the marginal profits at the point of the strategy space are given by:

$$\frac{\partial \Pi_1}{\partial q_1} = \alpha - p \cdot c_1 - (1-p) \cdot c_2 - 2q_1 - q_2 \quad , \qquad \frac{\partial \Pi_1}{\partial q_2} = -q_1 \tag{7}$$

and

$$\frac{\partial \Pi_2}{\partial q_2} = \alpha - c_1 - q_1 - 2q_2 \quad , \qquad \frac{\partial \Pi_2}{\partial q_1} = -q_2 \tag{8}$$

As it is noticed both managers care about the maximization of a utility function that contains a percentage of opponent company's profits (generalized relative profit function), which is given by:

$$U_i = (1-\mu_i) \cdot \Pi_i + \mu_i \cdot (\Pi_i - \Pi_j) = \Pi_i - \mu_i \cdot \Pi_j \tag{9}$$

where $\mu \in [0,1]$ is the percentage that the player i takes into account the opponent company's prifots. So, the marginal utility of the player i is given by the following equation:

$$\frac{\partial U_i}{\partial q_i} = \frac{\partial \Pi_i}{\partial q_i} - \mu_i \cdot \frac{\partial \Pi_j}{\partial q_i} \tag{10}$$

and the marginal utilities for each player are:

$$\frac{\partial U_1}{\partial q_1} = \alpha - p \cdot c_1 - (1-p) \cdot c_2 - 2q_1 - (1-\mu) q_2 \tag{11}$$

and

$$\frac{\partial U_2}{\partial q_2} = \alpha - c_1 - (1-\mu) q_1 - 2q_2 \tag{12}$$

Both players are characterized as bounded rational players. According to the existing literature it means that they decide their productions following a mechanism that is described by the equation:

$$\frac{q_i(t+1) - q_i(t)}{q_i(t)} = k \cdot \frac{\partial U_i}{\partial q_i} \ , \ k > 0 \tag{13}$$

Through this mechanism the player increases his level of adaptation when his marginal utility is positive or decreases his level when his marginal utility is negative, where k is the speed of adjustment of player, it is a positive parameter ($k > 0$), which gives the extend variation of production quantity of the each company, following a given utility signal.
The dynamical system of the players is described by:

$$\begin{cases} q_1(t+1) = q_1(t) + k \cdot q_1(t) \cdot \dfrac{\partial U_1}{\partial q_1} \\ q_2(t+1) = q_2(t) + k \cdot q_2(t) \cdot \dfrac{\partial U_2}{\partial q_2} \end{cases} \tag{14}$$

We will focus on the dynamics of this system to the parameter k.

2.2 Dynamical analysis

The dynamical analysis of the discrete dynamical system involves finding equilibrium positions and studying them for stability. The ultimate goal of this algebraic study is to formulate a proposition that will be the stability condition of the Nash Equilibrium position. Finally, these algebraic results are verified and visualized doing some numerical simulations using the program of Mathematica.

2.2.1 The equilibrium positions

The equilibriums of the dynamical system (14) are obtained as the nonnegative solutions of the algebraic system:

$$\begin{cases} q_1^* \cdot \dfrac{\partial U_1}{\partial q_1} = 0 \\ q_2^* \cdot \dfrac{\partial U_2}{\partial q_2} = 0 \end{cases} \tag{15}$$

which is obtained by setting : $q_1(t+1) = q_1(t) = q_1^*$ and $q_2(t+1) = q_2(t) = q_2^*$.

- If $q_1^* = q_2^* = 0$ then the boundary equilibrium position is the point:

$$E_0 = (0,0) \tag{16}$$

- If $q_1^* = 0$ and $\frac{\partial U_2}{\partial q_2} = 0$ then: $q_2^* = \frac{\alpha - c_1}{2}$ and the equilibrium position is the point:

$$E_1 = \left(0, \frac{\alpha - c_1}{2}\right) \tag{17}$$

- If $q_2^* = 0$ and $\frac{\partial U_1}{\partial q_1} = 0$ then: $q_1^* = \frac{\alpha - p \cdot c_1 - (1-p) \cdot c_2}{2}$ and the equilibrium position is the point:

$$E_2 = \left(\frac{\alpha - p \cdot c_1 - (1-p) \cdot c_2}{2}, 0\right) \tag{18}$$

- If $\frac{\partial U_1}{\partial q_1} = \frac{\partial U_2}{\partial q_2} = 0$ then the following system is obtained:

$$\begin{cases} \alpha - p \cdot c_1 - (1-p) \cdot c_2 - 2q_1^* - (1-\mu) \cdot q_2^* = 0 \\ \alpha - c_1 - (1-\mu) \cdot q_1^* - 2q_2^* = 0 \end{cases} \tag{19}$$

and the nonnegative solution of this algebraic system will give the Nash Equilibrium position $E_* = (q_1^*, q_2^*)$ where:

$$q_1^* = \frac{\alpha(1+\mu) + (1-\mu-2p) \cdot c_1 - 2(1-p) \cdot c_2}{4 - (1-\mu)^2} \tag{20}$$

and

$$q_2^* = \frac{\alpha(1+\mu) - (2-p+p \cdot \mu) \cdot c_1 + (1-p) \cdot (1-\mu) \cdot c_2}{4 - (1-\mu)^2} \tag{21}$$

This means that:

$$\alpha(1+\mu) + (1-\mu-2p) \cdot c_1 - 2(1-p) \cdot c_2 > 0 \tag{22}$$

and

$$\alpha(1+\mu) - (2-p+p \cdot \mu) \cdot c_1 + (1-p) \cdot (1-\mu) \cdot c_2 > 0 \tag{23}$$

2.2.2 Stability of equilibrium points

To study the stability of the equilibrium positions we need the Jacobian matrix of the dynamical system Eq.(15) which is the matrix:

$$J(q_1^*, q_2^*) = \begin{bmatrix} f_{q_1} & f_{q_2} \\ g_{q_1} & g_{q_2} \end{bmatrix} \tag{24}$$

where:

$$f(q_1,q_2)=q_1+k\cdot q_1\cdot\frac{\partial U_1}{\partial q_1}$$
$$g(q_1,q_2)=q_2+k\cdot q_2\cdot\frac{\partial U_2}{\partial q_2} \quad (25)$$

and as a result the Jacobian matrix of game's discrete dynamical system Eq.(14) is the following matrix:

$$J\left(q_1^*,q_2^*\right)=\begin{bmatrix} 1+k\cdot\left(\frac{\partial U_1}{\partial q_1}+q_1^*\cdot\frac{\partial^2 U_1}{\partial q_1^2}\right) & k\cdot q_1^*\cdot\frac{\partial^2 U_1}{\partial q_1\partial q_2} \\ k\cdot q_2^*\cdot\frac{\partial^2 U_2}{\partial q_2\partial q_1} & 1+k\cdot\left(\frac{\partial U_2}{\partial q_2}+q_2^*\cdot\frac{\partial^2 U_2}{\partial q_2^2}\right) \end{bmatrix} \quad (26)$$

For the E_0 the Jacobian matrix becomes as:

$$J(E_0)=\begin{bmatrix} 1+k\cdot\frac{\partial U_1}{\partial q_1} & 0 \\ 0 & 1+k\cdot\frac{\partial U_2}{\partial q_2} \end{bmatrix} \begin{matrix} A=1+k\cdot\frac{\partial U_1}{\partial q_1} \\ = \\ B=1+k\cdot\frac{\partial U_2}{\partial q_2} \end{matrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \quad (27)$$

with $Tr=A+B$ and $Det=A\cdot B$.

From the characteristic equation of $J(E_0)$, we find the nonnegative eigenvalues:

$$r_1=A=1+k\cdot\frac{\partial U_1}{\partial q_1} \quad \text{and} \quad r_2=B=1+k\cdot\frac{\partial U_2}{\partial q_2} \quad (28)$$

it's clearly seems that $|r_1|,|r_2|>1$ and the E_0 equilibrium is unstable.

For the E_1 the Jacobian matrix becomes as:

$$J(E_1)=\begin{bmatrix} 1+k\cdot\frac{\partial U_1}{\partial q_1} & 0 \\ -k\cdot(1-\mu)q_2^* & 1-2k\cdot q_2^* \end{bmatrix} \begin{matrix} C=1+k\cdot\frac{\partial U_1}{\partial q_1} \\ = \\ E=1-2k\cdot q_2^* \end{matrix} \begin{bmatrix} C & 0 \\ D & E \end{bmatrix} \quad (29)$$

with $Tr=C+E$ and $Det=C\cdot E$.

From the characteristic equation of $J(E_1)$, we find the nonnegative eigenvalue:

$$r_1=C=1+k\cdot\frac{\alpha(1+\mu)+(1-\mu-2p)\cdot c_1-(1-p)\cdot c_2}{2} \quad (30)$$

it's clearly seems that $|r_1|>1$, because:

$\alpha(1+\mu)+(1-\mu-2p)\cdot c_1-2(1-p)\cdot c_2>0$ Eq.(22) and the E_1 equilibrium is unstable.

For the E_2 the Jacobian matrix becomes as:

$$J(E_2)=\begin{bmatrix} 1-2k\cdot q_1^* & -k\cdot(1-\mu)q_1^* \\ 0 & 1+k\cdot\dfrac{\partial U_2}{\partial q_2} \end{bmatrix} \begin{matrix} F=1-2k\cdot q_1^* \\ = \\ H=1+k\cdot\frac{\partial U_2}{\partial q_2} \end{matrix} \begin{bmatrix} F & G \\ 0 & H \end{bmatrix} \tag{31}$$

with $Tr = F + H$ and $Det = F \cdot H$.

From the characteristic equation of $J(E_2)$, we find the nonnegative eigenvalue:

$$r_2 = H = 1 + k \cdot \frac{\alpha(1+\mu)-(2-p+p\cdot\mu)\cdot c_1+(1-p)\cdot(1-\mu)\cdot c_2}{2} \tag{32}$$

it's clearly seems that $|r_2| > 1$, because:

$\alpha(1+\mu)-(2-p+p\cdot\mu)\cdot c_1+(1-p)\cdot(1-\mu)\cdot c_2 > 0$ Eq.(23) and the E_2 equilibrium is unstable.

For the E_* the Jacobian matrix becomes as:

$$J(E_*)=\begin{bmatrix} 1+k\cdot q_1^*\cdot\dfrac{\partial^2 U_1}{\partial q_1^2} & k\cdot q_1^*\cdot\dfrac{\partial^2 U_1}{\partial q_1\partial q_2} \\ k\cdot q_2^*\cdot\dfrac{\partial^2 U_2}{\partial q_2\partial q_1} & 1+k\cdot q_2^*\cdot\dfrac{\partial^2 U_2}{\partial q_2^2} \end{bmatrix} \tag{33}$$

with

$$Tr = 2 - 2k\cdot q_1^* - 2k\cdot q_2^* \tag{34}$$

and

$$Det = 1 - 2k\cdot q_1^* - 2k\cdot q_2^* + \left[4-(1-\mu)^2\right]\cdot k^2\cdot q_1^*\cdot q_2^* \tag{35}$$

To study the stability of Nash equilibrium we use three conditions that the equilibrium position is locally asymptotically stable when they are satisfied simultaneously:

$$\begin{aligned} &\text{(i)} && 1-Det>0 \\ &\text{(ii)} && 1-Tr+Det>0 \\ &\text{(iii)} && 1+Tr+Det>0 \end{aligned} \tag{36}$$

The condition (i) gives:

$$1-Det>0 \;\Leftrightarrow\; 2k\left(q_1^*+q_2^*\right)-\left[4-(1-\mu)^2\right]\cdot k^2\cdot q_1^*\cdot q_2^*>0 \tag{37}$$

It's easy to find that the first condition (i) is always satisfied:

$$1-Tr+Det>0 \;\Leftrightarrow\; \left[4-(1-\mu)^2\right]\cdot k^2\cdot q_1^*\cdot q_2^*>0>0 \tag{38}$$

because: $\left[4-(1-\mu)^2\right]>0$.

Finally, the condition (iii) becomes as:

$$1+Tr+Det>0 \;\Leftrightarrow\; \left[4-(1-\mu)^2\right]\cdot q_1^*\cdot q_2^*\cdot k^2-4\left(q_1^*+q_2^*\right)\cdot k+4>0 \tag{39}$$

Proposition:
The Nash equilibrium of the discrete dynamical system Eq.(15) is locally asymptotically stable if:

$$2k\left(q_1^* + q_2^*\right) - \left[4 - (1-\mu)^2\right] \cdot k^2 \cdot q_1^* \cdot q_2^* > 0$$

and

$$\left[4 - (1-\mu)^2\right] \cdot q_1^* \cdot q_2^* \cdot k^2 - 4\left(q_1^* + q_2^*\right) \cdot k + 4 > 0$$

3 Numerical simulations focusing on the parameter k

From the condition (i) focusing on the parameter k we take the following inequality:

$$0 < k < \frac{2\left(q_1^* + q_2^*\right)}{\left[4 - (1-\mu)^2\right] \cdot q_1^* \cdot q_2^*} \tag{40}$$

The condition (iii) is the following:

$$\left[4 - (1-\mu)^2\right] \cdot q_1^* \cdot q_2^* \cdot k^2 - 4\left(q_1^* + q_2^*\right) \cdot k + 4 > 0$$

And its discriminant is positive:

$$\Delta = 16\left[\left(q_1^* - q_2^*\right)^2 + (1-\mu)^2\right] > 0 \tag{41}$$

so the condition (iii) is satisfied if :

$$k \in (0, k_1) \cup (k_2, +\infty) \tag{42}$$

where:

$$k_{1,2} = \frac{4\left(q_1^* + q_2^*\right) \pm \sqrt{\Delta}}{2\left[4 - (1-\mu)^2\right] \cdot q_1^* \cdot q_2^*} \tag{43}$$

are its two positive roots.
To provide some numerical evidence for the chaotic behavior of the system Eq.(14), as a consequence of change in the parameter k (the speed of adjustment), we present various numerical results here to show the chaoticity, including its bifurcations diagrams, strange attractors, Lyapunov numbers and sensitive dependence on initial conditions.
In order to study the local stability properties of the equilibrium points, it is convenient to take specific values for the other parameters: $\alpha = 5$, $c1 = 1$, $c2 = 0.5$ and $p = \mu = 0.5$. So, as a result we find that $q_1^* \approx 1.73$ and $q_2^* \approx 1.57$ and the stability condition becomes as:

$$0 < k < 0.48 \tag{44}$$

This algebraic result is verified by the bifurcation diagrams of q_1^* (Fig.1) and q_2^* (Fig.2) with respect to the parameter k. As it seems there is a locally asymptotically stable orbit until the value of 0.48 for the parameter k and after this value doubling period bifurcations are appeared and finally, for higher values of the parameter k the system's behavior becomes chaotic and unpredictable.

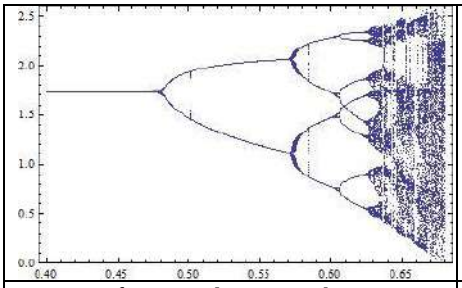


Fig. 1. Bifurcation diagram with respect to the parameter d against the variable q_1^* with 400 iterations of the map Eq.(15) for $\alpha = 5$, $c_1 = 1$, $c_2 = 0.50$, $p = 0.50$ and $\mu = 0.50$.

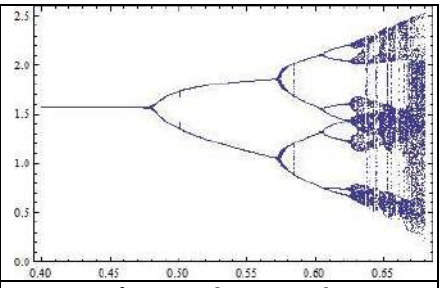


Fig. 2. Bifurcation diagram with respect to the parameter d against the variable q_2^* with 400 iterations of the map Eq.(15) for $\alpha = 5$, $c_1 = 1$, $c_2 = 0.50$, $p = 0.50$ and $\mu = 0.50$..

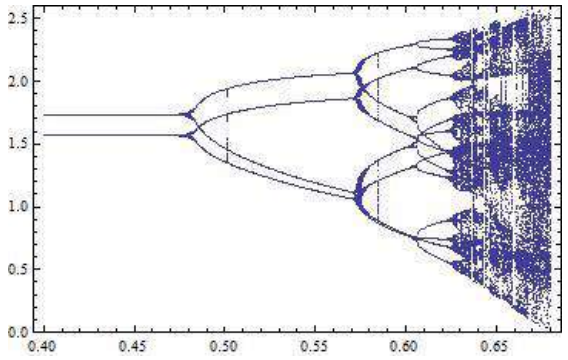


Fig.3. The two previous bifurcation diagrams of Fig.1 and Fig.2 in one.

This chaotic trajectory can create strange attractors (Fig.4) for a higher value of the parameter k like 0.675, outside the stability space. Also, computing the Lyapunov numbers (Fig.5) for this value of the parameter k and setting the same fixed values for the other parameters α, c_1, c_2, p and μ it seems that they are getting over the value of 1 as an evidence for the chaotic trajectory.

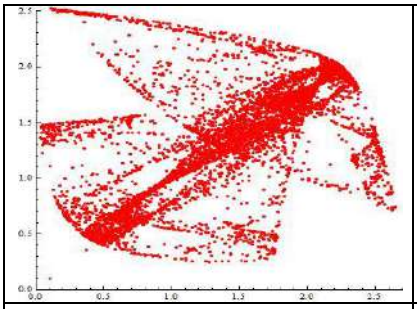


Fig. 4. Phase portrait (strange attractor) of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c_1 = 1$, $c_2 = 0.50$, $p = 0.50$, $\mu = 0.50$ and $k = 0.675$.

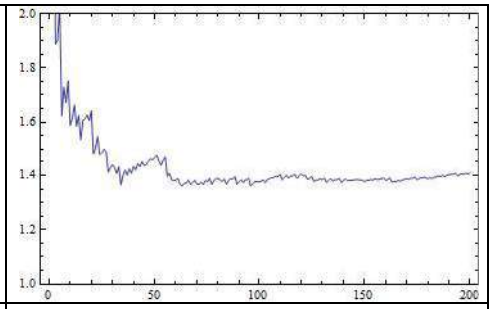


Fig. 5. Lyapunov numbers of the orbit of (0.1,0.1) with 8000 iterations of the map Eq.(15) for $\alpha = 5$, $c_1 = 1$, $c_2 = 0.50$, $p = 0.50$, $\mu = 0.50$ and $k = 0.675$.

This chaotic trajectory makes the system sensitive on initial conditions, which means that only a small change on a coordinate may change completely the system's behavior. For example, choosing two different initial conditions (0.1,0.1) (Fig.6) and (0.101,0.1) (Fig.7) with a small change at the q_1^*-coordinate and plotting the time series of system it seems that at the beginning the time series are indistinguishable, but after a number of iterations, the difference between them builds up rapidly.

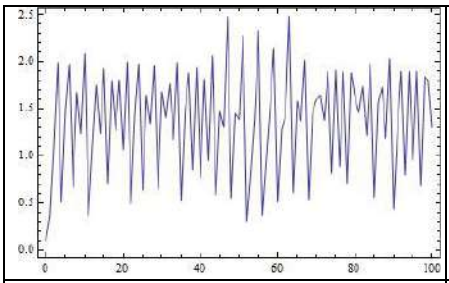


Fig. 28. Sensitive dependence on initial conditions for q_1^*-coordinate plotted against the time: the orbit of (0.1,0.1) of the system Eq.(15) for $\alpha = 5$, $c_1 = 1$, $c_2 = 0.50$, $p = 0.50$, $\mu = 0.50$ and $k = 0.675$.

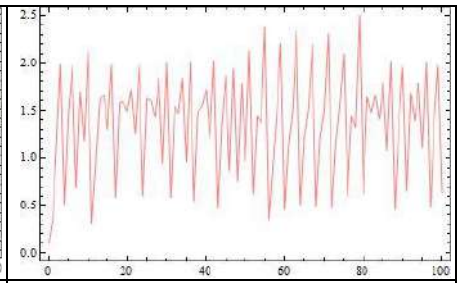


Fig. 29. Sensitive dependence on initial conditions for q_1^*-coordinate plotted against the time: the orbit of (0.101,0.1) of the system Eq.(15) for $\alpha = 5$, $c_1 = 1$, $c_2 = 0.50$, $p = 0.50$, $\mu = 0.50$ and $k = 0.675$.

Conclusions

In this paper we analyzed the dynamics of a differentiated Cournot duopoly with homogeneous expectations, linear demand and cost functions. An uncertainty of the first firm's cost function is introduced. By assuming that at each time period each firm maximizes its expected relative profit under the same expectations, a discrete dynamical system was obtained. Existence and stability of equilibrium of this system are studied. We showed numerically that the model gives chaotic and unpredictable trajectories. The main result is that higher values of the speed of adjustment may destabilize the Cournot–Nash equilibrium.

References

1. Agiza HN, Explicit stability zones for Cournot games with 3 and 4 competitors. *Chaos Solitons Fract*. 9: 1955-66, (1998).
2. Agiza, H.N., On the analysis of stability, bifurcation, chaos and chaos control of Kopel map. *Chaos, Solitons & Fractals* 10, 1909-1916, (1999).
3. Agiza HN., On the stability, bifurcations, chaos and chaos control of Kopel map. *Chaos Solitons Fract.* 11: 1909-16, (2004).
4. Agiza HN, Elsadany AA, Chaotic dynamics in nonlinear duopoly game with heterogeneous players. *Appl. Math. Comput*. 149: 843–60.
5. Agiza HN, Elsadany AA., Nonlinear dynamics in the Cournot duopoly game with heterogeneous players. *Physica A* 320: 512–24, (2003).
6. Agiza HN, Hegazi AS, Elsadany AA., Complex dynamics and synchronization of duopoly game with bounded rationality. *Math. Comput. Simulat*. 58: 133–46, (2002).
7. Askar, S.S., On complex dynamics of monopoly market, *Economic Modelling*, 31: 586-589, (2013).
8. Askar, S. S., Complex dynamic properties of Cournot duopoly games with convex and log-concave demand function, *Operations Research Letters* 42, 85–90, (2014).
9. Baumol, W.J., Quandt, R.E., Rules of thumb and optimally imperfect decisions, *American Economic Review* 54 (2): 23–46, (1964).
10. Bischi GI, Kopel M., Equilibrium selection in a nonlinear duopoly game with adaptive expectations. *J. Econom Behav. Org*. 46: 73–100 (2001).
11. Bischi GI, Lamantia F, Sbragia L., Competition and cooperation in natural resources exploitation: an evolutionary game approach. In: Cararro C, Fragnelli V, editors. *Game practice and the environment*. Cheltenham: Edward Elgar; 187–211, (2004).
12. Bischi GI, Naimzada A., Global analysis of a dynamic duopoly game with bounded rationality. In: Filar JA, Gaitsgory V, Mizukami K, editors. *Advances in dynamic games and applications*, vol. 5. Basel: Birkhauser; 361–85, (2000).
13. Bischi, G.I., Naimzada, A.K., Sbragia, L., Oligopoly games with local monopolistic approximation, *Journal of Economic Behavior and Organization* 62 (3): 371–388, (2007).
14. Bischi, G.I., Kopel, M., Equilibrium selection in a nonlinear duopoly game with adaptive expectations. *Journal of Economic Behavior & Organization* 46, 73–100, (2001).
15. Bischi, G.I., Naimzada, A., Global analysis of a dynamic duopoly game with bounded rationality. In: Filar JA, Gaitsgory V, Mizukami K, (eds). (2000). *Advances in dynamic games and applications*, vol. 5. Basel: Birkhauser: 361-385, (1999).

16. Cournot *A. Researches into the mathematical principles of the theory of wealth.* Homewood (IL): Irwin; (1963).
17. Day, R., *Complex Economic Dynamics*. MIT Press, Cambridge, 1994.
18. Dixit, A. K., Comparative statics for oligopoly, *Internat. Econom. Rev.* 27, 107–122, (1986).
19. Dixit, A.K., A model of duopoly suggesting a theory of entry barriers. Bell Journal of Economics 10, 20–32, (1979).
20. Den Haan W.J., The importance of the number of different agents in a heterogeneous asset-pricing model. *J. Econom. Dynam. Control*, 25:721–46, (2001).
21. Elsadany A.A., Dynamics of a Cournot duopoly game with bounded rationality based on relative profit maximization. *Applied Mathematics and Computation* 294, 253–263, (2017).
22. Elaydi, S., *An Introduction to Difference Equations,* third ed., Springer-Verlag, New York, 2005.
23. Junhai Ma, Zhanbing Guo, The influence of information on the stability of a dynamic Bertrand game, *Commun Nonlinear Sci Numer Simulat* 30: 32-44, (2016).
24. Hommes, C.H., *Heterogeneous agent models in economics and finance*, in: L. Tesfatsion, K.L. Judd (Eds.), Handbook of Computational Economics, Agent-Based Computational Economics, vol. 2, Elsevier Science B.V: 1109–1186, (2006).
25. Gandolfo G., *Economic dynamics*. Berlin: Springer, 1997.
26. Gao Y., Complex dynamics in a two dimensional noninvertible map. *Chaos Solitons Fract*. 39: 1798–810, (2009).
27. Kopel M., Simple and complex adjustment dynamics in Cournot duopoly models. *Chaos Solitons Fract.* 12: 2031–48, (1996).
28. Kulenonic, M., Merino, O., *Discrete Dynamical Systems and Difference Equations with Mathematica*, Chapman & Hall/Crc., 2002.
29. Medio A, Gallo G., *Chaotic dynamics: theory and applications to economics.* Cambridge (MA): Cambridge University Press, 1995.
30. Medio A, Lines M., Introductory notes on the dynamics of linear and linearized systems. In: Lines M, editor. *Nonlinear dynamical systems in economics.* SpringerWienNewYork: CISM; 1–26, (2005).
31. Medio A, Lines M., *Nonlinear dynamics. A primer.* Cambridge (MA): Cambridge University Press, 2001.
32. Naimzada, A.K., Ricchiuti G., Complex dynamics in a monopoly with a rule of thumb, *Applied Mathematics and Computation* 203: 921–925, (2008).
33. Naimzada, A., Sbragia, L., Oligopoly games with nonlinear demand and cost functions: two boundedly rational adjustment processes, *Chaos Solitons Fractals* 29, 707–722, (2006).
34. Nakamura,Y., Price versus quantity in a mixed duopoly: The case of relative profit maximization. *Economic Modelling* 44,37-43, (2015).
35. Puu, T., The chaotic monopolist, *Chaos, Solitons & Fractals* 5 (1): 35–44, (1995).
36. Puu T., The chaotic duopolists revisited. *J Econom. Behav. Org.* 37: 385–94, (1998).
37. Puu T., Chaos in duopoly pricing. *Chaos Solitons Fract.*1:573–81, (1991).
38. Puu T., Complex oligopoly dynamics. In: Lines M, editor. *Nonlinear dynamical systems in economics.* Springer Wien NewYork: CISM; p. 165–86, (2005).
39. Sarafopoulos G., On the dynamics of a duopoly game with differentiated goods, *Procedia Economics and Finance*, 19, 146 – 153, (2015).
40. Sarafopoulos G., Complexity in a duopoly game with homogeneous players, convex, log linear demand and quadratic cost functions, *Procedia Economics and Finance*, 33, 358 – 366, (2015).

41. Sarafopoulos, G., Papadopoulos K., On a Cournot duopoly game with differentiated goods, heterogeneous expectations and a cost function including emission costs, *Scientific Bulletin - Economic Sciences* **, 1** (1), 11-22, (2017).
42. Sarafopoulos, G., Papadopoulos K., On the stability of a Cournot dynamic game under the influence of information, *Knowledge E Publishing*, **ISSN:** 2518-668X, 131-144, (2018).
43. Sarafopoulos, G., Papadopoulos K., Chaos in oligopoly models, *International Journal of Productivity Management and Assessment Technologies*, **7** (1), 50-76, (2019).
44. Sarafopoulos, G., Papadopoulos K., Complexity in Bertrand duopoly game with heterogeneous players and differentiated goods, *Economic and Financial Challenges for Eastern Europe*, Article 2 (p. 15-26), Springer, 2019.
45. Satoh A., Tanaka Y., Relative Profit Maximization and Equivalence of Cournot and Bertrand Equilibria in Asymmetric Duopoly, MPRA Paper no. 56031.
46. Sedaghat, H., Nonlinear *Difference Equations: Theory with Applications to Social Science Models*, Kluwer Academic Publishers (now Springer), 2003.
47. Singh, N., Vives, X., Price and quantity competition in a differentiated duopoly. *The RAND Journal of Economics* 15, 546–554, (1984).
48. Tramontana, F., Heterogeneous duopoly with isoelastic demand function. *Economic Modelling* 27, 350–357, (2010).
49. Westerhoff, F., *Nonlinear expectation formation, endogenous business cycles and stylized facts*, Studies in Nonlinear Dynamics and Econometrics 10 (4) (Article 4), (2016).
50. Wu, W., Chen, Z., Ip, W.H., Complex nonlinear dynamics and controlling chaos in a Cournot duopoly economic model. *Nonlinear Analysis: Real World Applications* 11, 4363–4377, (2010).
51. Zhang, J., Da, Q.,Wang, Y., Analysis of nonlinear duopoly game with heterogeneous players. *Economic Modelling* 24, 138–148, (2007).
52. Zhang J., Da Q., Wang Y., The dynamics of Bertrand model with bounded rationality, *Chaos, Solitons and Fractals* 39, 2048-2055, (2009).

Driven Intrinsic Localized Modes in Soft Nonlinear Microscopic and Macroscopic Lattices

Masayuki Sato[1], Masato Sakai[1], and A. J. Sievers[2]

[1]Graduate School of Natural Science and technology, Kanazawa University, Kakumamachi, Kanazawa 920-1192, Japan
(E-mail: msato153@staff.kanazawa-u.ac.jp)
[2]Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY 14853-2501, USA
(E-mail: ajs19@cornell.edu)

Abstract. The possibility that large amplitude, localized vibrational excitations can exist in periodic physical lattices with nonlinear intersite forces was discovered over thirty years ago. The energy profiles of these intrinsic localized modes (ILMs) resemble those of localized vibrational modes at defects in a harmonic lattice. Described here are a variety of experiments on driver locked ILMs for two soft nonlinear lattices: an atomic spin array and an electrical nonlinear transmission line. CW locked ILMs in the quasi-1D antiferromagnet $(C_2H_5NH_3)_2CuCl_4$ have been found at frequencies below the antiferromagnetic resonance by employing four-wave mixing emission. A discrete step structure is observed in the emission signal as well as repeatable nonlinear ILM switching and hysteresis. These findings are compared with locked ILMs and large amplitude lattice spatial modes (LSMs) that have been measured for a driven 1-D nonlinear cyclic electric transmission line, where the nonlinear element is a saturable capacitor. Interestingly, by tuning the driver frequency away from the modal spectrum an LSM can be continuously converted into ILMs and vice versa. As a consequence, the resultant electrical energy distribution for the experimental soft nonlinear cyclic array can be either balanced or unbalanced.
Keywords: Intrinsic Localized Mode, Lattice Spatial Mode, Spin Wave, Electric Transmission Line, Soft Nonlinearity

1 Introduction

Although vibrational impurity modes in lattices have a long history[1], a major advance in the subject of excitations in discrete nonlinear lattices was the discovery that, even in the absence of impurities, some localized vibrations can be stabilized by lattice discreteness[2]. These excitations which, like impurity modes, extend only over the lattice constant scale are generically called “intrinsic localized modes” (ILMs) since they involve no disorder but come in many different shapes[3]. They are also called “lattice solitons”[4] or “discrete breathers”[5] because of similarities to some exact soliton solutions in nonlinear continuum theories. These excitations have been formalized in terms of a number of useful existence and stability criteria[6]. Both theory and a variety of experiments have been reviewed in the literature[7, 8] with experiments on microscopic antiferromagnetic lattices[9-14] and macroscopic electrical

nonlinear transmission lines (ENTL)[15-21] playing important roles. At first glance experiments on these two kinds of nonlinear systems would appear to be unrelated since the antiferromagnet is a quantum system while the ENTL is a classical one, but in certain cases both can be treated with classical equations of motion. The resulting dynamics is that both systems transmit energy over specific frequency bands. For the antiferromagnet it is spin wave energy that is transmitted, while for the ENTL it is electromagnetic. With regard to nonlinear properties of the two dynamical systems they can be designed to match as well. For the antiferromagnetic lattice both the spin exchange energy and the magnetic anisotropy energy have soft intersite and onsite nonlinearities so that the uniform antiferromagnetic resonance mode frequency decreases with increasing spin amplitude. By introducing a soft nonlinear capacitor in each cell of the ENTL, such as MOS-FETs, the low frequency uniform electromagnetic mode of this band pass filter will also decrease in frequency with increasing driving amplitude.

The purpose of this review is to examine the observed ILM findings for two driven, soft nonlinear lattices with damping, one a magnetic crystal with a microscopic lattice constant and the other an ENTL with a macroscopic one. From these very different experimental techniques on different systems come ILM findings that are remarkably similar.

2 Observation of driven antiferromagnetic ILMs

2.1 Spin wave dynamics of $(C_2H_5NH_3)_2CuCl_4$

How can an antiferromagnetic spin 1/2 system be compatible with a classical dynamical description? Below a Neél temperature of T_N=10.2K the spin 1/2 Cu^{2+} ions of $(C_2H_5NH_3)_2CuCl_4$ are oriented along the a-crystal axis, in alternate sheets of strong ferromagnetically coupled spins with weak antiferromagnetic

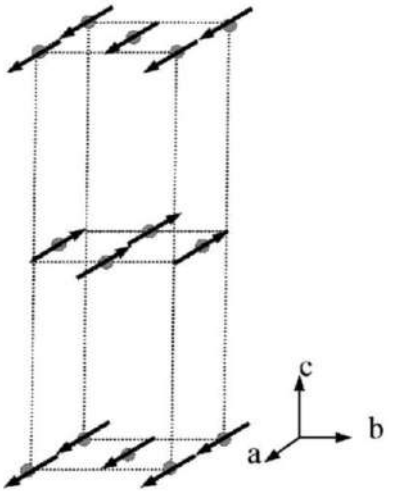


Fig. 1 Lattice and spin structure of $(C_2H_5NH_3)_2CuCl_4$. Circles denote Cu^{2+} ions and arrows indicate the spin configuration within the antiferromagnetic state. Only Cu^{2+} ions are shown for this layered, face centered, orthorhombic compound. The easy, second easy, and hard spin axes are labeled the a, b, and c crystal directions, respectively. [12]

coupling between adjacent sheets as illustrated in Fig. 1.[12] At 1.4 K the ratio of the interlayer spin exchange field to the intralayer one is 1.5×10^{-3} so the spins in a given layer are strongly aligned in the same direction, thus the low frequency spin dynamics can be approximated by a 1D two sublattice antiferromagnet with each layer represented by a macroscopic classical spin. Due to the easy axis anisotropy and the weak antiferromagnetic interaction between these classical spins the lower and upper frequency antiferromagnetic resonance (AFMR) modes are polarized along the c- and b- crystal axes, respectively.

The classical equation of motion for the normalized spin $\vec{S}_n$ at the nth sheet site, including damping and driving terms, is[9]

$$\frac{d}{dt}\vec{S}_n = -\gamma S_n \times H_n - \gamma\lambda \vec{S}_n \times \left(\vec{S}_n \times \vec{H}_n\right) \qquad (1)$$

where γ is the gyromagnetic ratio and λ the Landau damping parameter. The magnetic field acting on the nth macroscopic spin is

$$\vec{H}_n = -2J\left(\vec{S}_{n-1} + \vec{S}_{n+1}\right) - 2\vec{A}\times\vec{S}_n + h_{c0}\hat{e}_c \cos\omega t \qquad (2)$$

with J the nearest neighbor antiferromagnetic exchange constant, $\vec{A}$ the anisotropy field tensor, and h_{c0} the amplitude of the ac driving field along the *c* axis crystal direction. For this easy plane anisotropy case the two small amplitude dispersion curves are shown in Fig. 2 with the resulting uniform mode frequencies in the GHz range.

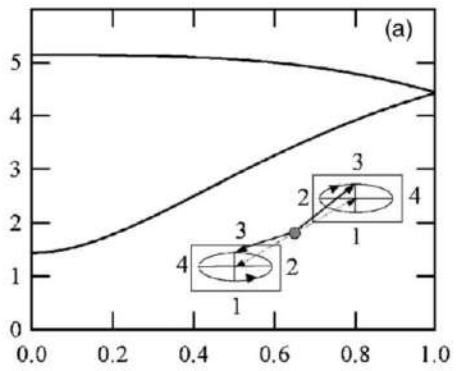


Fig. 2 Spin-wave dispersion curve for the antiferromagnet $(C_2H_5NH_3)_2CuCl_4$. Upper and lower branches along the c axis. The inset shows the uniform mode spin motion for the lower AFMR mode, which has a net ac magnetization only along the c axis. ILMs may be expected to occur below the minimum frequency of the dispersion curve. [12]

The insert shows the polarization of the lowest frequency AFMR mode with a linearly polarized transverse ac moment generated in the 1-3 direction but not in the 2-4 direction. It is the rod-shaped sample that has been studied in pulse and CW experiments.

2.2 Experimental procedure

In magnetic solids with the AFMR involving spins on the order of 10^{22} per cc the question is how to distinguish ILM dynamics from this background? The

answer is not to rely on a linear experimental technique but, since ILMs are nonlinear, to make use of a detection technique that involves a nonlinearity. Both pulse and CW four-wave techniques have been used to observe magnetic ILMs in this antiferromagnet[11-14]. The CW technique is outlined here. To illustrate the experimental approach, consider the AFMR absorption spectra shown in Fig. 3.

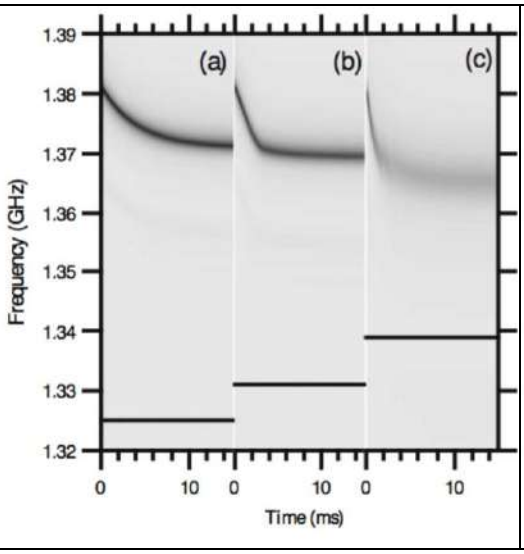


Fig. 3 Pulling the AFMR absorption frequency and capturing ILMs with a cw driver. (a) Initially the AFMR absorption line is at 1.38 GHz and the driver, 1.325 GHz, is switched on at t= 0. The frequency gap, Δf, between them continuously changes over a 15-ms time interval. (b) In this case Δf is now small enough so that the change in the AFMR frequency now occurs over 3 ms. (c) Here Δf is sufficiently small that the AFMR mode becomes unstable and ILMs are transferred to the driver frequency. [22]

A weak probe with frequency f_p is swept across the AFMR absorption line, which occurs at about 1.38 GHz at $t = 0$. Next the powerful driver with frequency f_D nearby is switched on as shown in Fig. 3(a). This driver pulls the absorption line to lower frequencies. The closer the driver frequency is to the AFMR the larger the effect as shown in Fig. 3(b). Figure 3(c) illustrates the point where the AFMR becomes unstable. So far all we see is the linear absorption spectrum. To access the nonlinearity associated with the ILM a four-wave mixing technique was employed. The resulting power emitted by the ILM alone, $P_{ILM}^{(3)}$, is detected at the spectrum analyzer frequency $f_{sp} = 2f_D - f_p$. Further analysis shows that among other elements[12]

$$\sqrt{P_{ILM}^{(3)}} \mu n_{ILM} P_D \qquad (3)$$

where n_{ILM} is the number of ILMs emitting and P_D is the driving power. Because n_{ILM} has integer values this relation will identify steps in the square root of the emitted power as the ILMs appear and disappear.

Figure 4 illustrates the turn on of the emission when the AFMR becomes instable. Both the absorption spectrum and the emission spectrum are superimposed on the same figure. Here the breakup of the absorption pattern

occurs at around 9 ms while simultaneously the emission grows rapidly in strength, as ILMs become locked at the driver frequency[14].

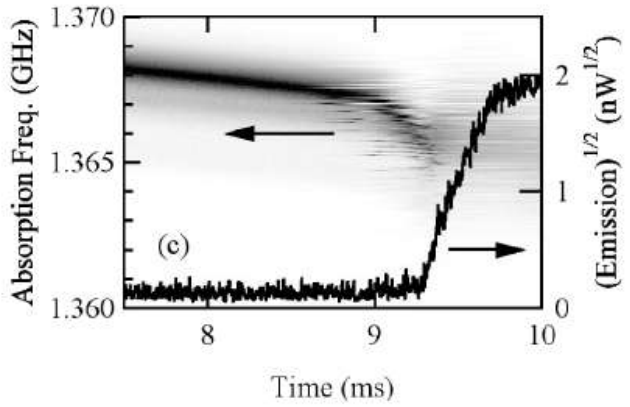


Figure 4 illustrates the turn on of the emission when the AFMR becomes instable. Both the absorption spectrum and the emission spectrum are superimposed on the same figure. Here the breakup of the absorption pattern occurs at around 9 ms while simultaneously the emission grows rapidly in strength, as ILMs become locked at the driver frequency[14].

2.3 Observing steady state magnetic ILMs

A time sweep of the emission spectrum does not have sufficient resolution to pick out individual ILMs so a more refined technique is required. With ILMs locked to the driver it is also possible to change the frequency gap, Δf, between the AFMR and the driver, by changing the AFMR frequency. Described here is the technique of tuning by changing the sample temperature. Since the anisotropy energy depends on the sublattice magnetization it depends on the temperature so both positive and negative manipulation of the frequency gap Δf is possible by simply sweeping the sample temperature.

By monitoring the four-wave emission as a function of sample temperature variation one finds the data shown in Fig. 5(a). It maps out two hysteresis loops, each for a single ILM[14]. Slowly increasing the sample temperature decreases the sublattice magnetization, decreases Δf and generates a locked ILM as shown by the dotted curve (arrow pointing to the right). Increasing the sample temperature produces the solid curve, arrow to the left, ending with the destruction of the ILM. In Fig. 5(a) the results for three different driver power levels are displayed with the lowest temperature emission steps up and down for a 50 mW driver. Figure 5(b) presents an expanded view of the lowest temperature sweep described in (a). Superimposed on that trace is the amplitude pattern for the switching of a nonlinear oscillator model to qualify the hysteretic switching that that has been observed. Interestingly it is found that the step heights do not change significantly as the power of the locking driver is changed.

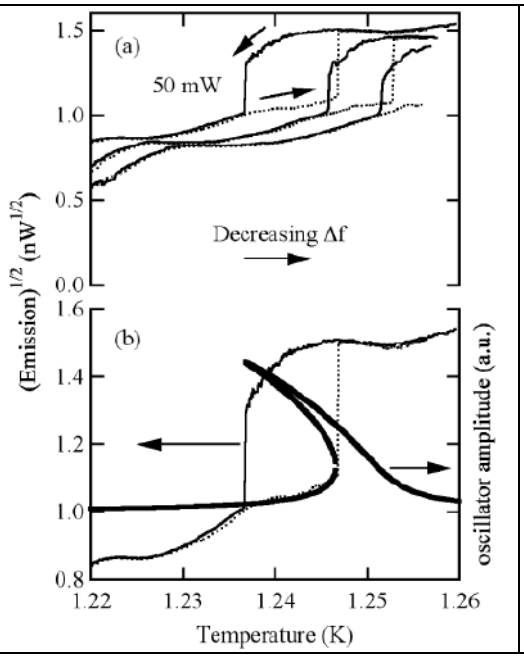


Fig. 5 Observation of locked ILM switching by tuning sample temperature. (a) Increasing the temperature decreases the frequency gap, Δf (dotted lines for three different driving powers). Solid lines are for decreasing temperatures. (b) Comparison of the 50 mW data with those expected for a driven nonlinear oscillator (thick solid line). [14]

3 Electric lattice with soft, saturable, nonlinearity

3.1 Experimental setup

Figure 6(a) shows the components within a unit cell of a 32 element electric nonlinear transmission line (ENTL). The nonlinear capacitor consists of two anti-paralleled N-channel MOS-FETs (IRFU-120). When the gate is negatively biased, holes in a P-semiconductor (for the N-channel FET) are accumulated below the gate between an oxide and semiconductor, and the boundary works as a conductive sheet. The capacitance between the gate and source electrodes is large with its value limited by the thickness of the oxide. When the gate is positively biased, the semiconductor is inverted to N-type and the boundary layer forms a conductive sheet. The capacitance between the drain and the gate is again large with its value limited by the thickness of the oxide. The two anti-paralleled capacitances together, identified as $C(V)$, are plotted as $1/C$ versus the applied DC bias in Fig. 1(b). Since the capacitance increases with absolute voltage, the resonance frequency decreases with increasing amplitude (soft nonlinearity); however, the decrement saturates abruptly when the voltage becomes large[23, 24]. One might consider a drawback of the MOS-FET to be its saturation property; however, we suggest that such behavior gives rise to an electrical analogue of a well studied mechanical nonlinear system containing piecewise linear springs in which the restoring force has a change in slope.

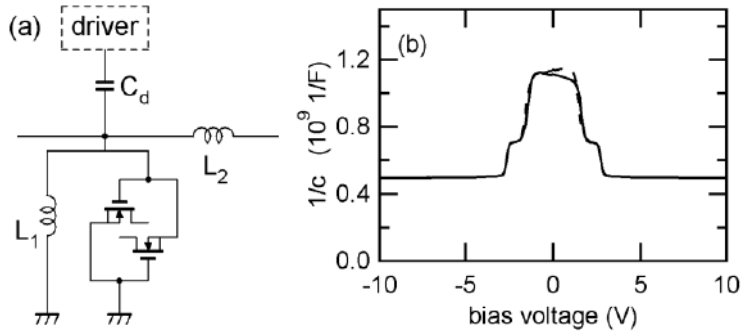


Fig. 6(a) Design of the nonlinear element for ENTL. Two MOS-FETs (IRFU-120) equal a nonlinear capacitor together with coil L_1(200 mH) to form a resonant circuit. L_2(200 mH) provides coupling to the next element. The driver for each element is via coupling capacitor C_d(70pF). (b) Inverse versus capacitance of the element illustrating the nonlinearity. Small AC voltage applied together with the DC bias voltage.

For this cyclic ENTL with an onsite capacitive nonlinearity $C(V_n)$ the dynamical equation of motion for the lattice voltage V_n as a function of time becomes

$$\left(C(V_n)+C_d\right)\frac{d^2V_n}{dt^2}+\left.\frac{dC(V)}{dV}\right|_{V_n}\left(\frac{dV_n}{dt}\right)^2+\frac{V_n}{L_1}+\frac{L_1w_0}{Q}\frac{dV_n}{dt}+\frac{1}{L_2}\left(2V_n-V_{n+1}-V_{n-1}\right)=C_d\frac{d^2}{dt^2}V_d \quad (4)$$

where Q is the quality factor and $C(V_n)$ is the differential capacitance $C(V)=dq/dV$ and q is the stored charge. The other parameters are defined in Fig. 6(a).

Figure 7 shows the linear band frequencies as a function of wave number. Because of the soft nonlinearity, an ILM can be generated below the bottom of the extended wave band. This shape is similar to that of the lower spin wave dispersion curve shown in Fig. 2.

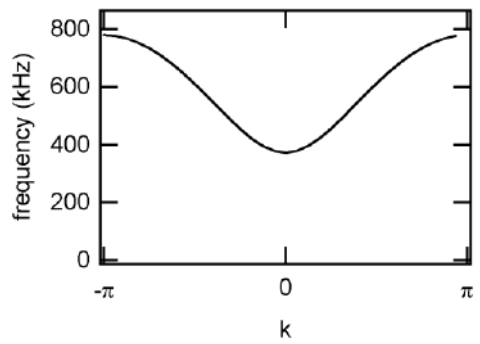


Fig. 7 Linear dispersion curve for the ENTL lattice. The circuit contains the element in Fig. 6(a). ILMs are expected below the minimum frequency of the dispersion curve.

3.2 Experimental results

Figure 8 shows the driver frequency dependence of two kinds of nonlinear excitations: large amplitude lattice spatial modes[25] (LSMs) and ILMs for the soft nonlinear lattice containing the nonlinear element of Fig. 6(a). At each driver frequency, a snap shot of the spatial voltage pattern at the maximum voltage versus time was captured. Those snapshots were ordered as a function of frequency resulting in Figs. 8(a)-(d). Arrows indicate sweep direction of the

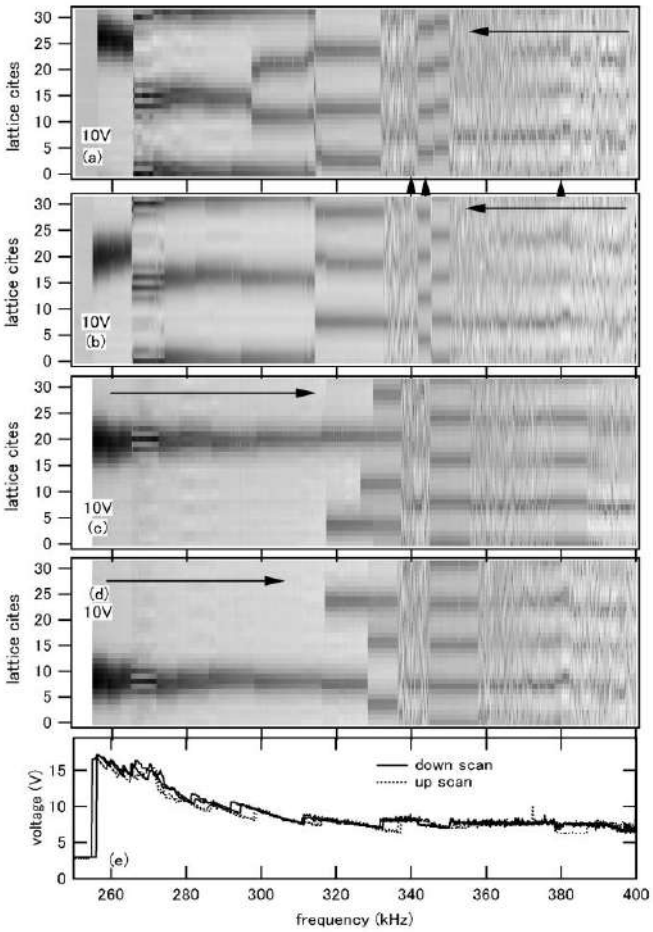


Fig. 8 Measured frequency dependence of driver locked LSMs and ILMs for the soft saturable 32 element nonlinear lattice. Four panel sweeps illustrate conversions. Driver frequency decreases in panels (a) and (b), and increases in panels (c) and (d). Snap shots at a moment when voltage is the maximum are mapped as a function of the driver frequency. Panel (e) displays maximum voltage as a function of frequency. Single excitation profiles such as in panel (a) 265-255 kHz, and (c) 255-317 kHz are due to ILMs. Other strip patterns at higher frequencies are due to LSMs.

driver frequency. Panel (e) demonstrates how the amplitude increases with decreasing driving frequency. At a frequency larger than 350 kHz in panels (a) and (b), a noisy pattern is observed. As the frequency is lowered, a 4 peak-LSM appears, followed by a noisy pattern, then a 3 peak-LSM, a 2 peak-LSM, and finally a single ILM is observed. LSM patterns are made from a few cyclic component waves, so that their peaks are mostly equidistance.

In addition, the LSM widths are observed to increase with decreasing frequency. This width dependence is causes by the saturable nonlinearity. Because of the lattice discreteness, the widening accompanies symmetry change of the peak from cite-center to bond-center or vise-versa. For example, the transition from cite-centered to bond-centered shapes is observed at 297.5 kHz. Panel 8(b) shows a sequential measurement. Similar patterns to those in 8(a) are observed.

Panels 8(c) and 8(d) display scan-up measurement results starting from the single ILM state. Hysteresis is observed for ILM and LSM generation. Panels (c)-(d) indicate that a single ILM is stable at frequencies around 255-317 kHz, while an LSM is stable at frequencies larger than 317 kHz. At a frequency below 317 kHz, there may be a stable two ILM state. Two peaks in panels (a)-(b) at these frequencies may be an LSM seeded ILM array.

The noisy pattern and 4-peaked LSM are very similar between Figs. 8(a)-(d). This signature indicates that the wave position is strongly influenced by lattice irregularity, because the wave samples the entire lattice. On the other hand, the single ILM position shifts, because the ILM feel only nearby impurities.

3.3 Analysis

To illustrate how the LSM is generated the multi-channel oscilloscope voltage data has been transformed in (time, space) domain into the (frequency, wavenumber) domain. Figure 9(a, b, c) show the results for driver frequencies indicated by the horizontal arrows in the panels and also identified by the three vertical arrow heads at bottom of Fig. 8(a). The 2D-FT log amplitude is displayed (gray-scale), where darker indicates larger amplitude. The solid curves (red) in frame (a)-(b) represent the superimposed linear dispersion curve. Excitation point is at k=0 (uniform mode) in this wavenumber-frequency space. Because of the low Q of the system it is not necessary that this frequency and the driver frequency be coincident. In addition, the nonlinearity and uniformly excited lattice ensures that the rest of the dispersion curve is shifted down as well. All other displayed intensities come from this one driving source. In panels (a) and (c), a faint gray intensity curve is apparent across the (k,w) range (dashed red curve in (a) is a guide to eye). This intensity display is from the shifted nonlinear dispersion curves, excited by the noisy pattern displayed in Fig. 8(a). The two dark spots that appear on the dashed curve at the driver frequency in Fig. 9(a) are from secondary excitation waves generated by a four-wave mixing process $2(k=0,W) \Rightarrow (k_+,W)+(k_-,W)$. This determines the number of LSM peaks in real space. For example, the spot locations are at $k=\pm 4\times(2p/32)$ in panel

(a), producing a 4 peak-LSM by mixing with the k=0 uniform excitation, where $(2p/32)$ is the k-space unit for the N=32 lattice. All other spots are excited by successive nonlinear wave generation processes.

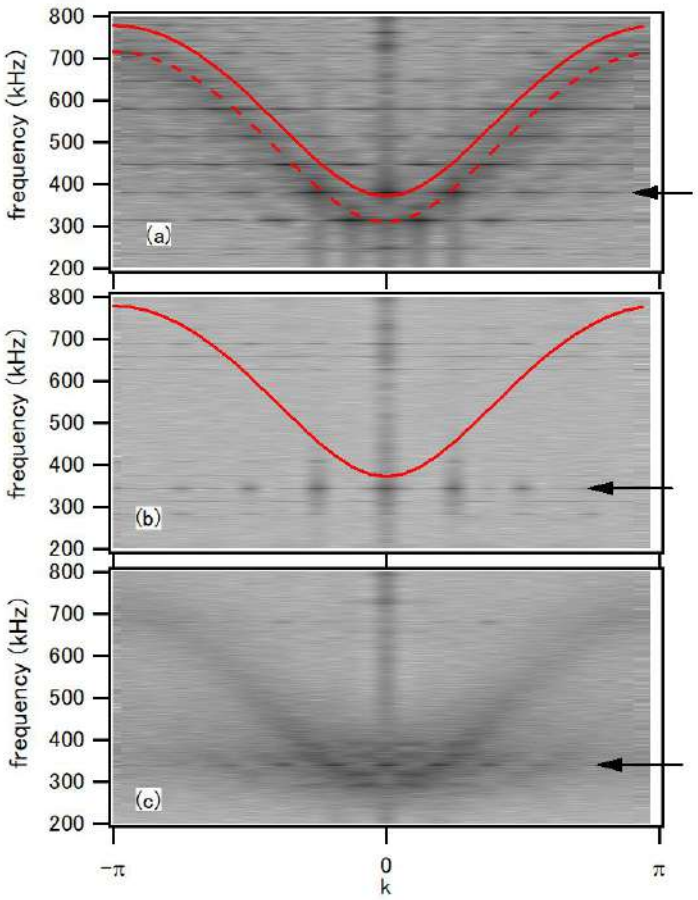


Fig. 9 2D-Fourier transform maps of voltage records for the ENTL system. 32 channel oscilloscope data from Fig. 8(a) are transformed into the frequency and wavenumber domain by time- and spatial-Fourier transform. Three panels are identified by driver frequencies: (a) 380 kHz, (b) 343.8 kHz and (c) 340 kHz. (See horizontal arrows.) Same driver frequencies are also identified at the bottom of Fig. 8(a). Log amplitude is displayed to emphasis small signals. Solid curves (red) in panels (a)-(b) are the superimposed linear dispersion curve. In panel (a) and (c), intensity from shifted dispersion curve is faintly observed. (Dashed red in (a) is a guide to the eye.) Because of random excitation in Fig. 8(a) at those frequencies, normal modes on the cyclic dispersion curve are excited and observed in this 2D-FT figure. In panel (b), the shifted nonlinear dispersion curve is not seen, since there are no such random noise vibrations. The wave number spots are produced by LSM formation at the driver frequency. Such wavenumber spots are seen also in panels (a) and (c).

In panel 9(b), the intensity spots are only at the driver frequency. In real space, a clean 4 peak-LSM is found. Here it is hard to identify a nonlinearly shifted dispersion curve in the absence of the noisy pattern. The secondary spot positions are the same as those in panel (a). In panel (c), the secondary spots are shifted to $k = \pm 3^{\circ}\left(2\rho/32\right)$. These are for 3-peaked LSM. A clean 3-peaked LSM is observed at the slightly lower frequency 330kHz.

The number of LSM peaks decreases with decreasing driver frequency, because the shift of the nonlinear dispersion curve relative to the driver frequency becomes smaller, and two secondary spot spacing in k-space becomes smaller. With a large number of lattice elements the transition from an LSM to an ILM array could happen at a still lower frequency, as the interaction decreases between neighboring peaks. In this case, many ILMs may remain. In Fig. 8, because of the saturation of the nonlinearity the peak widths become wider as the frequency decreases. To maintain the driver excitation, the frequency difference between the LSM (or ILMs) and the driver must be compensated by the nonlinearity.

The number of LSM peaks is roughly lattice size divided by wavelength determined by the secondary excitation wave. The maximum number of LSM peaks depends on the largest wavenumber of the four-wave mixing process, that is, the larger the wavenumber, the more peaks. The rule of thumb is the larger the spectral band width, the smaller the secondary wave number, if the frequency shift remains the same.

4. Discussion and conclusions

Both soft nonlinear lattices described here, the antiferromagetic and the electric transmission line, are essentially simple 1D systems, one with a microscopic lattice constant and the other with a macroscopic one. Although in each case the experimental techniques used to observe and explore the properties of ILMs have been quite different fundamentally similar results have been observed. In both systems the experiments consisted of frequency locking an ILM to a driver. For the magnetic system, because these excitations are strongly nonlinear, four-wave mixing emission spectroscopy is an ideal way to enhance the ILM signal over that produced by the more numerous plane wave spin excitations. The result is countable ILMs for an atomic spaced system. Experimentally, it is easiest to keep track of the difference between the AFMR and driver frequency as the important ILM parameter, since one had the flexibility to vary either the AFMR frequency, via the sample temperature, or to vary directly the driver frequency. Precision measurements allow the properties of single ILMs to be studied and their hysteresis curves to be measured. Surprisingly the step heights do not change significantly with locking driver power.

More extensive information about the general dynamic properties of a soft nonlinear 1D lattice has been obtained by studying a cyclic ENTL where each lattice element is monitored. Such a driven nonlinear transmission line, with periodic boundary conditions and elements that contain a saturable nonlinear capacitor, has been used to generate ILMs below the modal spectrum.

An ILM switching hysteresis signature, similar to that observed for the antiferromagnet, is a natural feature. An additional finding is the observation of LSMs within the spectrum. The most dramatic feature is that by simply changing the driver frequency the spectrum can evolve continuously from an LSM pattern distributed around the ring, with a successive decrease in the number of LSM-peaks, to multiple ILMs localized on a few sites, and finally to a single ILM. A four-wave mixing process plays a key role in determining the resulting LSM signature. As a consequence, the resultant AC energy distribution for the experimental soft nonlinear cyclic array can either be balanced or unbalanced.

Acknowledgements

MS is supported by JSPS Kakenhi Grant No. JP16K13716. AJS is supported by a Podell Endowment Award for Research and Scholarship.

References

1. I. M. Lifshitz, Some problems of the dynamic theory of non-ideal crystal lattices, Nuovo Cim. Suppl. **3**, 716 (1956).
2. A. S. Dolgov, The localization of vibrations in a nonlinear crystal structure, Sov. Phys. Solid State **28**, 907 (1986).
3. A. J. Sievers, S. Takeno, Intrinsic localized modes in anharmonic crystals, Phys. Rev. Lett. **61**, 970 (1988).
4. D. N. Christodoulides, R. I. Joseph, Discrete self-focusing in nonlinear arrays of coupled waveguides, Opt. Lett. **13**, 794 (1988).
5. S. Flach, C. R. Willis, Discrete breathers, Phys. Repts. **295**, 182 (1998).
6. S. Aubry, Breathers in Nonlinear Lattices: Existence, Linear Stability and Quantization, Physica D **103**, 201 (1997).
7. D. K. Campbell, S. Flach, Y. S. Kivshar, Localizing energy through nonlinearity and discreteness, Physics Today **57**, 43 (2004).
8. S. Flach, A. Gorbach, Discrete breathers-Advances in theory and applications, Phys. Repts. **467**, 1 (2008).
9. R. Lai, A. J. Sievers, Nonlinear nanoscale localization of magnetic excitations in atomic lattices, Phys. Repts. **314**, 147 (1999).
10. U. T. Schwarz, L. Q. English, A. J. Sievers, Experimental generation and observation of intrinsic localized spin wave modes in an antiferromagnet, Phys. Rev. Lett. **83**, 223 (1999).
11. M. Sato, A. J. Sievers, Direct observation of the discrete character of intrinsic localized modes in an antiferromagnet, Nature **432**, 486 (2004).
12. M. Sato, A. J. Sievers, Counting discrete emission steps from intrinsic localized modes in a quasi-1D antiferromagnetic lattice, Phys. Rev. B **71**, 214306 (2005).

13. L. Q. English, M. Sato, A. J. Sievers, Nanoscale intrinsic localized modes in an antiferromagnet, J. Appl. Phys. **89**, 6707 (2001).
14. J. P. Wrubel, M. Sato, A. J. Sievers, Controlled switching of intrinsic localized modes in a one-dimensional lattice, Phys. Rev. Lett. **95**, 264101 (2005).
15. R. Hirota, K. Suzuki, Theoretical and experimental studies of lattice solitons in nonlinear lumped networks, Proc. IEEE **61**, 1483 (1973).
16. K. Fukushima, M. Wadati, Y. Narahara, Envelope solitons in a new nonlinear transmission line, J. Phys. Soc. Jpn. **49**, 1953 (1980).
17. P. Marquie, J. M. Bilbault, M. Remoissenet, Observation of nonlinear localized modes in an electrical lattice, Phys. Rev. E **51**, 6127 (1995).
18. L. Q. English, F. Palmero, A. J. Sievers, P. G. Kevrekidis, D. H. Barnak, Traveling and stationary intrinsic localized modes and their spatial control in electrical lattices, Phys. Rev. E **81**, 046605 (2010).
19. L. Q. English et al., Nonlinear localized modes in two-dimensional electrical lattices, Phys. Rev. E **88**, 022912 (2013).
20. F. Palmero, L. Q. English, X.-L. Chen, e. al., Experimental and numerical observation of dark and bright breathers in the band gap of a diatomic electrical lattice, Phys. Rev. E **99**, 032206 (2019).
21. M. Remoissenet, Waves Called Solitons, (Springer-Verlag, Berlin, ed. Third, 1999), pp. 76.
22. M. Sato and A. J. Sievers, Experimental and numerical exploration of intrinsic localized modes in an atomic lattice, J. Biol. Phys. **35**, 57(2009)
23. W. Shi, S. Shige, Y. Soga, M. Sato, A. J. Sievers, Intrinsic localized modes in a nonlinear electrical lattice with saturable nonlinearity, EPL **103**, 30006 (2013).
24. S. Shige, K. Miyasaka, W. Shi, Y. Soga, M. Sato A. J. Sievers, Experimentally observed evolution between dynamic patterns and intrinsic localized modes in a driven nonlinear electrical cyclic lattice, EPL **121**, 30003(2018); Erratum: EPL **126**, 29901 (2019)
25. V. M. Burlakov, Interference of mode instabilities and pattern formation in anharmonic lattices, Phys. Rev. Lett. **80**, 3988 (1996).

Tilted Nadaraya-Watson Regression Estimator

Mohammad T. Shakeri[1], Farzaneh Boroumand[2,1], Hassan Doosti[2,3], Nino Kordzakhia[2,4], and Mehdi Salehi[5]

[1] Department of Biostatistics, Health School, Mashhad University of Medical Sciences, Mashhad, Iran
(E-mail: shakerimt@mums.ac.ir)
[2] Department of Mathematics and Statistics, Faculty of Science and Engineering, Macquarie University, NSW, Australia
(E-mail: Farzaneh.Boroumand@mq.edu.au)
[3] (E-mail: Hassan.Doosti@mq.edu.au)
[4] (E-mail: Nino.Kordzakhia@mq.edu.au)
[5] Department of Mathematics and Statistics, University of Neyshabur, Neyshabur, Iran
(E-mail: salehi.sms@neyshabur.ac.ir)

Abstract. In nonparametric statistics the tilting techniques are sustainably used for adjusting an empirical distribution by replacing uniform distribution of weights by general multinomial distribution. In this paper a tilting approach has been used for minimizing "the distance" to an infinite order (IO) regression estimator, a comparator regression function estimator. We also provide the simulation study results illustrating the tilted version of the Nadaraya-Watson (N-W) estimator performs better than its classical analog (the N-W estimator) in terms of Median Integrated Squared Error (MISE). In addition, the performance of the tilted N-W regression function estimator has been examined using the Italy's COVID-19 deaths data.
Keywords: Nadaraya-Watson estimator, Tilted Nadaraya-Watson estimator, Infinite order estimator, Kernel estimator, Trapezoidal kernel, Cross-validation function, MISE, ISE..

1 Introduction

Let the regression model be

$$Y_i = r(X_i) + \epsilon_i, \;\; 1 \leq i \leq n,$$

where $(Y_1, X_1), (Y_2, X_2), \ldots, (Y_n, X_n)$, are the data pairs, the design variable $X \sim f_X$, X and ϵ are independent, ϵ_i's are independent and identically distributed (iid) errors with zero mean $E(\epsilon) = 0$ and variance $E(\epsilon^2) = \sigma^2$. The regression function $r(x)$ is unknown. The kernel, local polynomial regression and orthogonal series methods are commonly used for estimating an unknown regression function. The kernel and local polynomial regression function estimators are also known as smoothers, see Wasserman [1], Fan and Gijbels [2], Hall and Racine [18] and the references within.

In this paper, we propose a tilting method which represents an optimised modification of Nadaraya-Watson estimator. An estimator $\hat{r}_n$ of r is a linear smoother if

$$l(x) = (l_1(x), ..., l_n(x))^T$$

and

$$\hat{r}_n = l(x)^T Y = \sum_{i=1}^{n} l_i(x_i) Y_i,$$

where l_i are weights and for all x, $\sum_{i=1}^{n} l_i(x) = 1$. The $l_i(x)$ for Nadaraya-Watson estimator is defined as

$$l_i(x) = \frac{K(\frac{X_i - x}{h})}{\sum_{j=1}^{n} K(\frac{X_j - x}{h})}, h > 0. \tag{1}$$

In (1), K is the weighting function that assigns the values to the design points X_i according to proximity to x. The Nadaraya-Watson (N-W) estimator, which is the kernel estimator, depends on the bandwidth parameter $h > 0$, so-called smoothing parameter. As the bandwidth increases the kernel estimator tends to a flat function, due to this property the N-W estimator is often referred to as the locally constant estimator.

In the tilting approach, the empirical distribution is being adjusted by replacing the equal weights $1/n$ by p_i, where $p_i \geq 0$ and $\sum_i^n p_i = 1$, [4]. The tilting method is sustainably used for an unknown density function estimation. Grenander [5] proposed tilting-based method by imposing some restrictions on density estimators. The empirical likelihood-based methods and distance measure approach are used for estimation of tilting parameters within regression function estimators. The empirical likelihood-based method is a semi-parametric method that allows to find a parameter through estimating equations. Owen [6] was the first who proposed the empirical likelihood method as an alternative to likelihood ratio tests. Chen [7], Zhang [8], Mülle *et al.* [9] used the empirical likelihood-based methods for estimating tilting parameters. In the distance measure approach, the tilted estimator is obtained by minimizing some distance, subject to constraints. Hall and Presnell [10], Carroll *et al.* [11], Doosti and Hall [12], Doosti *et al.* [13] used various distance measures for estimating density functions. Namely, in Doosti *et al.* [13] have introduced a cross-validation function tailored to this estimation problem. They had shown that the proposed density function estimator performs better than the conventional kernel-based estimators.

The aim of this study is to introduce the tilted nonparametric N-W regression function estimator which is obtained by minimizing the distance to a comparator estimator. In this paper, an infinite order flat-top kernel estimator is selected as the comparator estimator. The infinite order flat-top kernel estimator, also known as the trapezoidal kernel, can be defined through Fourier transform which is flat near the origin and is infinitely differentiable.

2 Notation and preliminary results

Definition 1. Let λ be the Fourier transform of kernel K, and we select g (g is not unique) to make $\lambda(s)$, $\lambda^2(s)$, ans $s\lambda(s)$ integrable. For a fixed constant $c > 0$

$$\lambda(s) = \begin{cases} 1, & |\ s\ | \leq c \\ g(|\ s\ |), & |\ s\ | > c \end{cases}.$$

The flat-top kernel is

$$K(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \lambda(s) e^{-isx} ds. \tag{2}$$

The infinite order regression estimator was introduced by McMurry and Politis [14] in the form of linear smoother

$$\check{r}_n = \sum\nolimits_{i=1}^{n} \check{l}_i(x_i) Y_i,$$

$$\check{l}_i(x) = \frac{K_{IO}(\frac{X_i - x}{h})}{\sum_{j=1}^{n} K_{IO}(\frac{X_j - x}{h})}.$$

K_{IO} refers to any kernel which fulfils the definition 1. The idea behind using an infinite order kernel for estimating r_n is that these type of kernels reduce the bias asymptotically at the rate $O(h^k)$, [14]. The trapezoidal kernel satisfies definition 1

$$K_T(x) = \frac{2(cos(x/2) - cos(x))}{\pi x^2}.$$

The Fourier transform of the trapezoidal kernel K_T is

$$\lambda(s) = \begin{cases} 1 & \mid s \mid \leq 1/2, \\ 2(1 - \mid s \mid) & 1/2 < \mid s \mid \leq 1, \\ 0 & \mid s \mid > 1. \end{cases}$$

We denote $\hat{r}_n(.|\theta)$, the tilted estimator, where $\theta = (h, p)$ is the vector of unknown parameters. Our objective is to estimate θ by minimizing the distance measure between $\hat{r}_n(.|\theta)$ and $\check{r}$. For preserving the convergence rate of $\check{r}_n$, in this paper, we use the L_2-metric as the distance between these two estimators, [12].

3 Numerical study

This section contains the results of the numerical study carried out for analysing the performance of the tilted N-W estimators. The exponential regression function, $r(x) = x + 4exp(-2x^2)/\sqrt{2\pi}$, has been used with the uniform design density and normally distributed error terms for generation of 500 data sets. The relative performance of the tilted N-W estimator has been assessed with respect to varying sample sizes and standard deviation levels. The assessment has been made by comparing the Median Integrated Squared Error (MISE) and ISE, the latter belonging to [-2,2]. The cross-validation function method has been employed for an optimal bandwidth selection for Nadaraya-Watson estimator, [1]. For an infinite order estimator, the bandwidths were selected using the rule of thumb introduced by Politis in [14] which is available form R-package 'iosmooth'. In fact, the bandwidth for tilted N-W estimator has been estimated within the proposed numerical procedure.

In table 1, we provides the MISEs for the simulated data. Evidently, for fixed sample size, as the variance increases the tilted N-W estimators perform better than others. For larger sample sizes, in contrast, the N-W estimator, outperforms the tilted N-W estimator. However, for smaller sample sizes and the moderate standard deviation levels, the tilted N-W estimator remains, at some extent, superior to the conventional estimators.

n	σ	IO	NW	NW p4	NW p10
60	0.3	0.1559	**0.0663**	0.1308	0.1529
	0.5	0.1980	**0.1398**	0.1724	0.1953
	0.7	0.2515	**0.2152**	0.2316	0.2492
	1	0.3588	0.3697	**0.3418**	0.3650
	1.5	0.6530	0.6281	**0.6197**	0.6520
	2	1.0524	0.9892	**0.9871**	1.0597
100	0.3	0.1195	**0.0442**	0.1034	0.1191
	0.5	0.1432	**0.0914**	0.1253	0.1426
	0.7	0.1781	**0.1443**	0.1607	0.1766
	1	0.2490	0.2324	**0.2305**	0.2469
	1.5	0.4165	0.4366	**0.4041**	0.4144
	2	0.6487	0.6780	**0.6107**	0.6371
200	0.3	0.0991	**0.0232**	0.0891	0.0997
	0.5	0.1089	**0.0470**	0.0993	0.1086
	0.7	0.1256	**0.0822**	0.1172	0.1253
	1	0.1577	**0.1299**	0.1533	0.1589
	1.5	0.2401	0.2542	**0.2386**	0.2416
	2	0.3568	0.3878	**0.3464**	0.3534
1000	0.3	0.0801	**0.0058**	0.0776	0.0800
	0.5	0.0823	**0.0125**	0.0790	0.0825
	0.7	0.0853	**0.0207**	0.0801	0.0845
	1	0.0922	**0.0356**	0.0830	0.0917
	1.5	0.1080	**0.0716**	0.0972	0.1074
	2	0.1294	0.1286	**0.1209**	0.1308

Table 1. MISE for the Infinite Order (IO) estimator with the trapezoidal kernel, N-W estimator and tilted N-W estimator with 4 (p4) and 10 (p10) weighting nodes.

4 Real data

In this section, the tilted N-W estimator along with two other kernel-based estimators are being used for a curve fitting to real data. We shall apply the tilted N-W estimator approach to Italy's COVID-19 daily deaths data from 23 February 2020 to 6 May 2020 downloaded from https://www.ecdc.europa.eu.

The interest in COVID-19 death rate modelling among scientists is growing rapidly since the outbreak of the pandemic began [15], [16], [17]. Along with the tilted N-W estimator, we applied the N-W, and IO estimators. The tilted N-W estimator performed the best in terms of the Mean Square Errors (MSE).

Table 2 provides the MSE for each estimator. In terms of minimising the MSE, the tilted N-W estimator ranked first, followed by N-W and IO estimators resulting in the relative improvement of 3% and 4%, respectively. Slightly, improved performance of the tilted N-W estimator is attributed to the lower MSE components at the edges versus other kernel-based regression function estimators which are generally known for so-called "edge effect", [19].

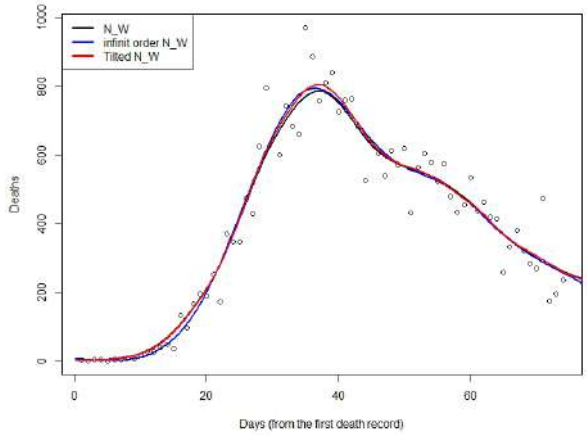


Fig. 1. Fitting N-W, IO, and tilted N-W regression curves to the logarithm of the COVID-19 deaths data.

	infinite order estimator	Nadaraya-Watson estimator	tilted Nadaraya-Watson estimator
MSE	3819.803	3778.425	3667.398

Table 2. MSEs for Nadaraya-Watson, infinite order, and tilted Nadaraya-Watson estimators.

Acknowledgement

This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (NCI), which is supported by the Australian Government. This research forms part of the corresponding author's PhD thesis approved by the ethics committee of Mashhad University of Medical Sciences with the project code 971017.

References

1. L. Wasserman. All of nonparametric statistics, *Springer Science and Business Media*, 2006.

2. J. Fan and I. Gijbels. Local polynomial modelling and its applications: monographson, *statistics and applied probability*, 66. Vol. 66. CRC Press, 1996.
3. P. G. Hall and J. S. Racine. Infinite order cross-validated local polynomial regression, *Journal of Econometrics*, 185.2 (2015), pp. 510–525.
4. P. G. Hall and Q. Yao. Data tilting for time series, *Journal of the Royal StatisticalSociety: Series B (Statistical Methodology)*, 65.2 (2003), pp. 425–442.
5. U. Grenander. On the theory of mortality measurement: part ii, *In:Scandinavian Actuarial Journal*, 1956.2 (1956), pp. 125–153.
6. A. B. Owen. Empirical likelihood ratio confidence intervals for a single functional, *Biometrika*,75.2 (1988), pp. 237–249.
7. S. X. Chen. Empirical likelihood-based kernel density estimation, *Australian Journal of Statistics*, 39.1 (1997), pp. 47–56.
8. B. Zhang. A note on kernel density estimation with auxiliary information, *Communications in Statistics-Theory and Methods*, 27.1 (1998), pp. 1–11.
9. U. U. Müller, A. Schick, and W. Wefelmeyer. Weighted residual-based densityestimators for nonlinear autoregressive models, *Statistica Sinica*, (2005), pp. 177–195.
10. P. G. Hall and B. Presnell. Intentionally biased bootstrap methods, *Journal of theRoyal Statistical Society: Series B (Statistical Methodology)*, 61.1 (1999), pp. 143–158.
11. R. J. Carroll, A. Delaigle, and P. Hall. Testing and estimating shape-constrainednonparametric density and regression in the presence of measurement error, *Journal of the American Statistical Association*, 106.493 (2011), pp. 191–202.
12. H. Doosti and P. G. Hall. Making a non-parametric density estimator more attractive,and more accurate, by data perturbation, *Journal of the Royal Statistical Society: SeriesB (Statistical Methodology)*,78.2 (2016), pp. 445–462.
13. H. Doosti, P. G. Hall, and J. Mateu. Nonparametric tilted density function estimation: A cross-validation criterion, *Journal of Statistical Planning and Inference*, 197(2018), pp. 51–68.
14. T. L. McMurry and D. N. Politis. Nonparametric regression with infinite orderflat-top kernels, *Journal of Nonparametric Statistics*, 16.3-4 (2004), pp. 549–562.
15. R. Verity . Estimates of the severity of coronavirus disease 2019: a model-basedanalysis, *The Lancet infectious diseases*, (2020).
16. C. Anastassopoulou . Data-based analysis, modelling and forecasting of the COVID-19 outbreak, *PloS one*,15.3 (2020), e0230405.
17. J. M. Nichols. Modelling daily death patterns in the COVID - 19 Pandemic using Linear Regression, Fast Fourier Transforms, ODE Solver and Generalized Poissonian Distributions in FORTRAN, *Working paper Texas AM University (TAMU)*,(15 July 2020).
18. P. G. Hall and J. S. Racine, Infinite order cross-validated local polynomial regression, *Journal of Econometrics*, 185.2 (2015), pp. 510-525.
19. P. G. Hall and T. E. Wehrly. A geometrical method for removing edge effects from kernel-type nonparametric regression estimators, *Journal of the American Statistical Association*, 86.415 (1991), pp. 665–672.

Homoclinic Orbits and Solitary Waves within the Non-dissipative Lorenz Model and KdV Equation

B.-W. Shen*

Department of Mathematics and Statistics
San Diego State University
5500 Campanile Drive
San Diego, CA, 92182, USA
*Email: bshen@mail.sdsu.edu
Web: http://bwshen.sdsu.edu

Abstract. Recent studies using the classical Lorenz model and generalized Lorenz models present abundant features of both chaotic and oscillatory solutions that may change our view on the nature of weather as well as climate. In this study, the mathematical universality of solutions in different physical systems is presented. Specifically, the main goal is to reveal mathematical similarities for solutions of homoclinic orbits and solitary waves within a three-dimensional non-dissipative Lorenz model (3D-NLM), the Korteweg-de Vries (KdV) equation, and the Nonlinear Schrodinger (NLS) equation. A homoclinic orbit for the X, Y, and Z state variables of the 3D-NLM connects the unstable and stable manifolds of a saddle point. The X and Z solutions for the homoclinic orbit can be expressed in terms of a hyperbolic secant function ($sech$) and a hyperbolic secant squared function ($sech^2$), respectively. Interestingly, these two solutions have the same mathematical form as solitary solutions for the KdV and NLS equations, respectively. After introducing new independent variables, the same second order ordinary differential equation (ODE) and solutions for the Z component and the KdV equation were obtained. Additionally, the ODE for the X component has the same form as the NLS for the solitary wave envelope. Finally, how a logistic equation, also known as the Lorenz error growth model, and an improved error growth model can be derived by simplifying the 3D-NLM is also discussed. Future work will compare the solutions of the 3D-NLM and KdV equation to understand the different physical role of nonlinearity in their solutions, and the solutions of the error growth model and the 3D-NLM, as well as other Lorenz models, in order to propose an improved error growth model to better represent linear and nonlinear error growth for both oscillatory and non-oscillatory solutions.

Keywords: Homoclinic Orbits, Solitary Waves, Lorenz Model, KdV Equation, Logistic Equation, Error Growth Model, Homoclinic Bifurcation.

13^{th} *CHAOS Conference Proceedings, 9-12 June 2020*

1 Introduction

After the pioneering studies by Prof. Lorenz (Lorenz, 1963a, 1972; Gleick, 1987; Anthes, 2011), the chaotic nature of weather with a finite predictability has been well recognized. Our current understanding of chaotic nature is indeed largely derived from the analysis of the chaotic and unstable solutions of Lorenz models (e.g., Lorenz, 1963a and 1969a). On the other hand, with a goal of revealing the true nature of weather, a recent comprehensive literature review on Lorenz models and high-dimensional Lorenz models (e.g., Curry, 1978; Curry et al., 1984; Faghih-Naini and Shen, 2018; Howard and Krishnamurti, 1986; Hermiz et al., 1995; Thiffeault and Horton, 1996; Musielak et al., 2005; Reyes and Shen, 2019; Roy and Musielak, 2007; Moon et al., 2017; Felicio and Rech, 2018; Shen 2014-2019a) suggests a need for taking other types of solutions into consideration. For example, in addition to three types of solutions that include steady-state, chaotic, and limit cycle solutions (e.g., Sparrow, 1982; Pedlosky and Frenzen, 1980; Ghil et al. 2010), Shen and colleagues discussed two kinds of attractor existence using newly developed high-dimensional Lorenz models (e.g., Reyes and Shen, 2019; Shen, 2019a, b; Shen et al., 2019). The first kind of attractor coexistence consists of chaotic and steady-state solutions, while the second kind of attractor coexistence includes limit cycle and steady-state solutions. While everything centers around periodic solutions prior to Lorenz studies and while great attention has been paid to non-periodic solutions following Lorenz studies, a refined view on the dual nature of chaos and order (i.e., non-periodic and periodic solutions) is suggested for weather (Shen 2019a,b; Shen et al., 2020).

Prior to Lorenz's chaotic solutions in the 1960s, the occurrence of a nonlinear periodic solution (i.e., limit cycle) in the atmosphere was indeed illustrated using so-called dishpan experiments in the laboratory in the 1950s (Fultz et al. 1959; Hide 1953). Based on the experiment, three types of solutions were discovered, including: (1) steady state solutions, (2) irregular chaotic solutions, and (3) vacillation (Lorenz 1963b; Ghil and Childress 1987; Lorenz, 1993). Using a model with 14 ODEs (e.g., Lorenz, 1963b) that is different from the well-known 1963 Lorenz model, Prof. Lorenz discussed symmetric vacillation (SV), unsymmetric vacillation (UV), and their transition. The vacillations are closed orbits within the phase space. Although Lorenz (1963b) documented the transition between the two types of vacillations (Figure 5 of Lorenz, 1963b), he did not discuss the condition under which the transition occurred. Later, Ghil and Childress (1987) suggested a hypothetical role for the homoclinic orbit in connection with the transition from the UV to SV regime (e.g., Figure 5.6 of Ghil and Childress, 1987). In the 1970s, studies suggested that amplitude vacillation may be viewed as a limit cycle solution (e.g., Pedloksy 1972; Smith 1975; Smith and Reilly 1977). As a result, improving our understanding of homoclinic orbits and nonlinear oscillatory solutions has the potential to address the transition between UV and SV.

Although the Lorenz 1963 model with 3 ODEs originally came from the Saltzman model with 7 ODEs (e.g., Saltzman, 1962; Lorenz, 1993), the Lorenz model with simplicity has been used as a powerful tool for revealing various

types of solutions, as discussed above. Additionally, the Lorenz 1963 model also produces other types of solutions under certain conditions. For example, the three-dimensional non-dissipative Lorenz model (3D-NLM), that represents the original three-dimensional Lorenz model (3DLM) with sufficiently large heating parameters, was studied in order to reveal two types of oscillatory solutions with small and large cycles, and homoclinic orbits (Shen, 2018). The two types of oscillatory solutions are not isolated and, thus, are not the same as the limit cycle within the dissipative system. A homoclinic orbit that connects stable and unstable manifolds at the saddle point may be viewed as the limiting case of periodic solutions. The homoclinic orbit is also viewed as a separatrix that separates periodic solutions with small and large cycles. Recently, related approaches have been applied for deriving non-dissipative 5D and 7D Lorenz models in order to examine quasi-periodic solutions that include additional incommensurate frequencies (Shen and Faghih-Naini, 2017; Faghih-Naini and Shen, 2018). To obtain better predictability within numerical models, detecting non-chaotic processes that include periodic, quasi-periodic, and steady-state solutions based on observations and model simulations is recommended (e.g., Shen et al., 2020).

In a recent study (e.g., Shen, 2019a), the 3D-NLM was compared to the original Lorenz model and simplified Lorenz models, including the so-called Lorenz geometric model (Guckenheimer and William, 1979; Smale, 1998; Tucker, 2002; Hirsch et al. 2013) and the (nonlinear) limiting equation (e.g., Sparrow, 1982) in order to reveal the role of a saddle point and nonlinear terms (referred to as the nonlinear feedback loop in Shen, 2014) in creating chaotic responses. Within the 3D-NLM, a homoclinic orbit that is associated with the appearance of periodic orbits is not orbitally stable, and its presence tends to introduce irregular behavior associated with chaos (Guckenheimer and Holmes, 1983). Such a feature is presented in Appendix A. Therefore, improving our understanding of the dynamics of homoclinic orbits may be helpful for determining the conditions under which large predictability may be present.

Researchers often analyze the homoclinic orbit within a 2D or higher dimensional space and focus on its association with the saddle point, as well as homoclinic points where stable and unstable manifolds intersect. On the other hand, individual components for the homoclinic orbit within the 3D-NLM also possess interesting dynamics. For example, analytical solutions for the time varying X and Z components can be expressed in terms of the hyperbolic secant function $sech$ and the hyperbolic secant squared function $sech^2$, respectively. As a result, they appear as a solitary pattern in temporal space, providing motivation for comparing the mathematical solutions of homoclinic orbits and so-called solitary waves, showing the mathematical universality of solutions in different physical systems.

In addition to chaotic behavior, a major focus of nonlinear dynamics has been on nonlinear wave systems that have analytical solutions. One such nonlinear system is the Korteweg-de Vries (KdV) equation (e.g., Boussinesq, 1877; Koreweb and deVries, 1895). The KdV equation with quadratic nonlinearity governs the dynamics of weakly dispersive, weakly nonlinear water waves (Miles, 1981). The system possesses a family of periodic solutions that can

be written in terms of the Jacobi elliptic function (e.g., *cn*; Whitham, 1974; Baines, 1995). Waves with *cn* functions are referred to as cnoidal waves. One limit of the *cn* function is a hyperbolic secant function. Such a limiting solution does not represent a periodic wave but moves as a single isolated 'hump' with a unchanging pattern (Lighthill, 1978). The solution, that appears as a result of the balance between nonlinearity and dispersion, is called a solitary wave or soliton (e.g., Zabusky and Kruskal, 1965; Zabusky and Porter, 2010; Balmforth, 1995; Boyd, 2015).

The first observed solitary wave was documented one hundred and seventy years ago (e.g., Russell, 1844). In the atmosphere, Lin and Goff (1988) documented observations of a long-lived mesoscale wave that persisted for approximately 9 hours and its wavelength was approximately 185 km. These authors suggested that the observed wave may be viewed as a solitary wave, while Rottman and Einaudi (1993) pointed out that a wave ducting mechanism associated with a critical level at approximately 8 km may also be important in trapping wave energy (e.g., Rottman and Grimshaw, 2002).

To evaluate the performance of weather forecasting models, a logistic equation has been used for describing the evolution of the root mean square (rms) average forecast error for ensemble runs (Lorenz 1969b, 1982, 1996; Nicolis 1992; Kalnay, 2003; Zhang et al. 2019). The logistic equation is also known as the Lorenz error growth model, and its solution with an initial positive value is a sigmoid function that is a monotonically increasing function with non-negative growth rates. Specifically, given an initial error with a small, positive value, errors that are described by the sigmoid function grow at an initial larger growth rate, then at a nonlinear smaller growth rate, and eventually approach a constant defined as a saturated error that has a zero growth rate. Earlier efforts were made to show how a logistic equation with a linear and a quadratic term was derived (e.g., Nicolis 1992). Here, how the 3D-NLM can be simplified to yield the Lorenz error growth model and an improved error growth model is briefly discussed.

The paper is organized as follows. In Section 2, a review of the 3D-NLM and its solutions for homoclinic orbits is provided. Equations and solutions of the 3D-NLM, NLS, and KdV are then compared in order to reveal mathematical similarities for solutions of the homoclinic orbits and solitary waves. A derivation of the Lorenz error growth model and the improved error growth model from the 3D-NLM is also presented. Concluding remarks are provided at the end. Appendix A discusses the sensitive dependence of solutions on initial conditions that are close to a homoclinic orbit and homoclinic bifurcation using the 3D-NLM with additional positive and negative dissipative terms. A brief review on the so-called Lorenz error growth model is provided in Appendix B.

2 The 3D-NLM, KdV, and Logistic Equations

In this section, a review and additional analysis for the 3D-NLM, as well as a comparison to the NLS and KdV equations, are provided. How to simplify the 3D-NLM to yield a logistic equation, which is known as a Lorenz error growth model, is then discussed.

2.1 The 3D-NLM and its Homoclinic Orbits

The 3D-NLM and its solutions for homoclinic orbits are presented below. As discussed in Shen (2018), as well as in Appendix A, the equations of the 3D-NLM and its two conservation laws can be written, as follows:

$$\frac{\mathrm{d}^2X}{\mathrm{d}\tau^2} + \left(\frac{X^2}{2} - (\sigma r + \frac{C_1}{C_o})\right) X = 0, \tag{1}$$

$$\overline{\mathrm{KE}} + \overline{\mathrm{PE}} = C_o\left(\frac{X^2}{2} - \sigma Z\right) = C_1, \tag{2}$$

$$\overline{\mathrm{KE}} + \overline{\mathrm{APE}} = \frac{C_o}{2}\left(X^2 - \frac{\sigma}{r}(Y^2+Z^2)\right) = C_2. \tag{3}$$

Here, X, Y, and Z are the state variables. The two time-independent parameters are the Prandtl number (σ) and the normalized Rayleigh number (r), also referred to as a heating parameter. $\overline{\mathrm{KE}}$, $\overline{\mathrm{PE}}$, and $\overline{\mathrm{APE}}$ represent the domain-averaged kinetic energy, potential energy, and available potential energy, respectively (e.g., Blender and Lucarini, 2013; Shen, 2014). C_o, C_1, and C_2 are constants, indicating energy conservation (e.g., Shen, 2014). Eqs. (2)-(3) are related to the two Nambu Hamiltonians (Nambu, 1973; Nevir and Blender, 1994; Roupas, 2012). See details in Shen (2018).

For the homoclinic orbit that begins and ends at the saddle point (i.e., the origin), both C_1 and C_2 are zero and remain zero forever. Thus, Eqs. (1-3) lead to:

$$\frac{\mathrm{d}^2X}{\mathrm{d}\tau^2} + \left(\frac{X^2}{2} - \sigma r\right) X = 0, \tag{4}$$

$$\frac{X^2}{2} - \sigma Z = 0, \tag{5}$$

$$X^2 - \frac{\sigma}{r}(Y^2+Z^2) = 0. \tag{6}$$

To facilitate discussions, two non-trivial turning points, X_t, are defined when the second derivative of X becomes zero (i.e., $X^2/2 - \sigma r = 0$ in Eq. (4)), yielding $X_t = \pm\sqrt{2\sigma r}$. Equations (5-6) lead to:

$$Y^2 = \frac{1}{\sigma^2}\left(\sigma r X^2 - \frac{X^4}{4}\right). \tag{7}$$

Since $Y^2 \geq 0$, we have $|X| \leq X_{max} = 2\sqrt{\sigma r}$. As a result, when $Y = 0$, $|X| = X_{max}$, and $Z = 2r$ from Eq. (5). The special point $(X, Y, Z) = (2\sqrt{\sigma r}, 0, 2r)$ is used as the initial condition in order to obtain an analytical solution of the homoclinic orbit. Without a loss of generality, unless stated otherwise, the positive X component for the intervals $[0, X_t]$ and $[X_t, X_{max}]$ are mainly analyzed. By plugging $Y = \sigma^{-1}dX/d\tau$ (i.e., Eq. 3 of Shen 2018) into Eq. (7), we have:

$$\left(\frac{dX}{d\tau}\right)^2 - \sigma r X^2 + \frac{X^4}{4} = 0, \tag{8a}$$

yielding:

$$\frac{dX}{d\tau} = \pm\sqrt{\sigma r X^2 - \frac{X^4}{4}}. \tag{8b}$$

The term X^4 is associated with a nonlinear feedback loop that consists of two nonlinear terms ($-XZ$ and XY). Without this term, the system in Eq. (4) becomes linear. As discussed in Shen (2018), the "contracting component" of the homoclinic orbital solution for $\tau \in [0, \infty)$, with an initial condition of $(X, Y, Z) = (2\sqrt{\sigma r}, 0, 2r)$, can be written as follows:

$$X(\tau) = \frac{4\sqrt{\sigma r}}{e^{\sqrt{\sigma r}\tau} + e^{-\sqrt{\sigma r}\tau}} = 2\sqrt{\sigma r} sech(\sqrt{\sigma r}\tau), \tag{9a}$$

$$Y(\tau) = -4r \frac{e^{\sqrt{\sigma r}\tau} - e^{-\sqrt{\sigma r}\tau}}{\left(e^{\sqrt{\sigma r}\tau} + e^{-\sqrt{\sigma r}\tau}\right)^2} = -(2r) tanh(\sqrt{\sigma r}\tau) sech(\sqrt{\sigma r}\tau), \tag{9b}$$

$$Z(\tau) = \frac{X^2(\tau)}{2\sigma} = (2r) sech^2(\sqrt{\sigma r}\tau). \tag{9c}$$

As a result of the property for the hyperbolic secant function, the derivatives of the X component, as well as the Z component, are non-positive (e.g., $dX/d\tau \leq 0$) for $\tau \in [0, \infty)$. Eqs. (9a) and (9c) are referred to as the monotonically decreasing component of the homoclinic orbit. Since Y is not a monotonic function for $\tau \in [0, \infty)$, the term "contracting component" is used to describe the solution in the $X-Y$ phase space. Such a component begins at $(X, Y, Z) = (2\sqrt{\sigma r}, 0, 2r)$ for $\tau = 0$ and approaches the origin for $\tau \to \infty$. The time τ_t, when the second derivative of X is zero (i.e., at $X = X_t = \sqrt{2\sigma r}$), can be determined by $sech(\sqrt{\sigma r}\tau_t) = \sqrt{2}/2$ in Eq. (9a), yielding $\tau_t = 0.881/\sqrt{\sigma r}$. Therefore, for the decreasing component of X within $\tau \in [0, \infty)$, $d^2X/d\tau^2$ is negative in association with stronger nonlinearity for $\tau \in [0, \tau_t)$ and is positive for association with weaker nonlinearity for $\tau \in (\tau_t, \infty)$.

Note that the solution of Z at τ_t ($Z(\tau_t)$) is equal to r, which is half of its maximum ($Z_{max} = 2r$). The interval $[-\tau_t, \tau_t]$ is then used to define a "pulse width" (or "temporal pulse width") within which the Z component is greater than or equal to half of the maximum. As a result, the pulse width is determined by the Rayleigh parameter (r) and the Prandtl parameter (σ), while its amplitude is solely determined by the heating parameter r. In other words, the heating parameter determines both pulse width and amplitude. Below, this feature is compared to the nonlinear solitary wave solutions of the KdV equation.

Since the system is invariant under $\tau \to -\tau$ and $Y \to -Y$ (e.g., Strogatz, 2015), the solution $(X(\tau), -Y(\tau), Z(\tau))$ in backward time $\tau \in (-\infty, 0]$ represents the expanding component of the homoclinic solution. The solution begins at the origin for $\tau \to -\infty$ and then moves to the point $(X, Y, Z) = (2\sqrt{\sigma r}, 0, 2r)$ for $\tau = 0$. The corresponding X and Z components are referred to as the monotonically increasing components. Similarly, for the increasing component within $\tau \in [(-\infty, 0]$, solutions within $\tau \in (-\tau_t, 0)$ have a stronger nonlinearity than those within $\tau \in (-\infty, -\tau_t)$.

Figure 1 displays the solution of the homoclinic orbit for $\tau \in [-1, 1]$, with $\sigma = 10$ and $r = 10$. Such a choice leads to $X(\tau = -1) = 20 sech(-10) = 0.0018$, which is close to 0, and allows us to effectively analyze solutions within a finite interval instead of an infinite interval $\tau \in (-\infty, \infty)$. In panels (a)-(c), blue lines display the contracting components for X, Y, and Z for $\tau \in [0, 1]$, as

described by Eqs. (9a)-(9c). Red lines show the expanding components for $\tau \in [-1, 0]$, as discussed above. Panel (d) plots the homoclinic orbit within the X-Y space. The blue curve for the monotonically decreasing component begins at $(X, Y, Z) = (2\sqrt{\sigma r}, 0, 2r)$ and moves forward in time towards the origin. The red curve represents the monotonically increasing component moving from the origin at $\tau = -1$ to $(X, Y, Z) = (2\sqrt{\sigma r}, 0, 2r)$ at $\tau = 0$. Although the homoclinic orbit connects the stable and unstable manifolds of the saddle point, the trajectory "takes forever" to reach the saddle point (as $t \to \infty$) and is, therefore, not a periodic solution. Alternatively, the homoclinic orbit may be viewed as a special periodic solution with a period of infinity. In Figure 1, vertical lines are plotted at $\tau = \pm\tau_t$ or $X = X_t$, where $d^2X/d\tau^2 = 0$. At $\tau = \tau_t$, $Z(\tau_t) = Z_{max}/2$. As a result, a pulse width is determined by $[-\tau_t, \tau_t]$.

2.2 A Comparison of the 3D-NLM with the NLS and KdV Equations

Previously, Shen (2018) showed that the 3D-NLM represents a special form of the Duffing equation whose solutions are Jacobi elliptic functions (e.g., cn). Below, it is additionally shown that the equation for $(dX/d\tau)^2$ in Eq. (8a) shares the same mathematical form as the NLS equation for the amplitude of a traveling wave ($h(x, \tau)$), defined as follows:

$$\left(\frac{dh}{dx}\right)^2 + \delta h^2 + \frac{\gamma}{2}h^4 = 0. \tag{10}$$

Here, the lower case x represents the distance in physical space. Parameters δ and γ are negative and positive, respectively (e.g., Haberman, 2013). As documented, the solution to Eq. (10) is a hyperbolic secant function, which is also the solution of X, as shown in Eq. (9a). Table 1 lists the above equations.

Below, an ODE for the Z of the 3D-NLM is first derived in order to perform a comparison with the KdV equation. By differentiating both sides in Eq. 5 and making a square, we obtain:

$$\left(\frac{dZ}{d\tau}\right)^2 = \frac{1}{\sigma^2}X^2(\frac{dX}{d\tau})^2. \tag{11}$$

By using Eq. (8a) to replace $\frac{dX}{d\tau}$ in Eq. (11), we have:

$$\left(\frac{dZ}{d\tau}\right)^2 = \frac{1}{\sigma^2}\left(\sigma r X^4 - \frac{X^6}{4}\right). \tag{12}$$

Applying Eq. (5) to replace the above X^4 and X^6 leads to:

$$\left(\frac{dZ}{d\tau}\right)^2 = 4\sigma r Z^2 - 2\sigma Z^3. \tag{13}$$

By differentiating Eq. (13) and dividing it by $dZ/d\tau$, we can obtain an equation for the second derivative of Z, as follows:

$$\frac{d^2Z}{d\tau^2} + 3\sigma Z^2 - 4\sigma r Z = 0. \tag{14}$$

For a comparison below, the independent variable is rescaled by introducing $\zeta = \sqrt{\sigma}\tau$ with a different time scale, turning Eq. (14) into:

$$\frac{d^2Z}{d\zeta^2} + 3Z^2 - 4rZ = 0. \tag{15}$$

Next, the KdV equation for comparison with the 3D-NLM is presented. Well documented is the fact that the KdV contains a family of periodic solutions within a three parameter space (e.g., Lighthill, 1978; Whitham, 1974). These three parameters include a wave's amplitude, wavelength, and water depth. Two non-dimensional parameters can be defined by the ratio of amplitude to water depth and the ratio of water depth to wavelength. The periodic solutions can be written in terms of the Jacobi elliptic function (e.g., *cn*; Whitham, 1974; Baines, 1995), referred to as cnoidal waves. One limit of the *cn* function is a hyperbolic secant function, which is the solitary wave solution with one non-dimensional parameter. The solitary wave solution is discussed below. The KdV equation is written as follows:

$$\frac{\partial u}{\partial t} + 6u\frac{\partial u}{\partial x} + \frac{\partial u^3}{\partial x^3} = 0. \tag{16}$$

By introducing a traveling-wave coordinate $\xi = (x - ct)$ and assuming $f(\xi) = u(x,t)$ (e.g., Sprott 2010), we can turn Eq. (16) into:

$$\frac{d^2f}{d\xi^2} + 3f^2 - cf = A, \tag{17}$$

where c is a parameter and A is an integration constant that is determined as $A = 0$ when solutions satisfy:

$$\lim_{\xi\to\pm\infty} f(\xi) = 0 \lim_{\xi\to\pm\infty} \frac{df(\xi)}{d\xi} = 0. \tag{18}$$

As a result, Eq. (15) of the 3D-NLM and Eq. (17) of the KdV equation are idential when $c = 4r$. A solitary wave solution of the KdV in Eq. (17) is a hyperbolic secant squared function, as follows:

$$u(x,t) = \frac{c}{2} sech^2\left(\frac{\sqrt{c}(x-ct)}{2}\right). \tag{19}$$

The solution has one dimensionless parameter c that is a ratio between wave amplitude and water depth (Baines, 1995). By replacing (c) and $(x - ct)$ by $(4r)$ and (ζ) (that is equal to $\sqrt{\sigma}\tau$), respectively, the above is identical to Eq. (9c).

As summarized in Table 1, the above discussions indicate that the governing equations for the Z and X components of the homoclinic orbit within the 3D-NLM have the same mathematical form as the KdV and NLS equations, respectively. Their corresponding solutions also have the same function form.

2.3 A Logistic Equation from a Simplified 3D-NLM

The so-called Lorenz error growth model that is indeed a logistic equation has been used in the error analysis of numerical solutions for decades. Although derivations for a one-variable, two-parameter model in the form of a stochastically driven logistic equation have been provided, (e.g., Nicolis, 1992), a consis-

tent approach for improving the error growth model is still needed (e.g., Zhang et al., 2019). Another goal of this study is to propose an error growth model that can represent the evolution of errors for both chaotic and non-chaotic solutions. Here, the Lorenz error growth model whose solution is a sigmoid function is first derived through simplification of the 3D-NLM. The sigmoid function with an initial small positive value is also a monodically increasing function with non-negative growth rates and is viewed as an approximated solution, while the monotonically increasing components of the homoclinic orbit is treated as a true error solution for verifications of the approximated solution. Simply speaking, a comparison of the sigmoid function and the increasing component of the homoclinic orbit is made to understand the condition under which the derived error growth model may represent the original system.

Applying a Taylor series expansion, Eq. (8a) can be approximated as follows:

$$\frac{dX}{d\tau} \approx \pm\sqrt{\sigma r}X\left(1-\frac{X^2}{8\sigma r}\right), \tag{20a}$$

which can be written:

$$\frac{dE}{d\tau} \approx \pm 2\sqrt{\sigma r}E\left(1-\frac{E}{8\sigma r}\right), \tag{20b}$$

after introducing $E = X^2$. Eq. (20b) is a typical logistic equation that has been used as an error growth model (Lorenz, 1969b, 1982). A saturated value is defined when $dE/d\tau = 0$, leading to $E = 8\sigma r$ and $X = 2\sqrt{2\sigma r}$. Here, the saturated value of X is greater than the maximum $X_{max} = 2\sqrt{\sigma r}$ for the homoclinic orbit. The solution of Eq. (20a) is a sigmoid function, written as follows:

$$E = \frac{Be^{2\sqrt{\sigma r}\tau}}{1+\frac{B}{8\sigma r}e^{2\sqrt{\sigma r}\tau}}, \tag{21a}$$

yielding:

$$X = \pm\sqrt{E}. \tag{21b}$$

The integration constant, B, is determined by $E(\tau=-1) = X^2(\tau=-1)$ using Eq. (9a). Figure 2 shows that the solution of Eq. (21b) captures the evolution of the increasing component of X reasonably well at linear and weakly nonlinear stages. The weakly nonlinear stage is defined in intervals where X is close to X_t but smaller than X_{max}. However, the sigmoid function cannot represent the solution at fully nonlinear stages near $X = X_{max}$ and overestimates values in this interval. By comparison, for $\tau \approx 0$, a *sech* function that represents the homoclinic orbit can be approximated with a *cosine* function, as shown in Figure 2. The result additionally indicates that the non-monotonicity near X_{max} for the homoclinic orbit cannot be represented by a monotonic sigmoid function. Note that the term "fully nonlinear stage" should not be confused with the term "weakly nonlinear system" for the KdV equation. The latter simply indicates limited nonlinearity in the partial differential equation of the KdV system.

The above analysis suggests that a self-consistent, improved error growth model can be obtained when higher order terms are included in the Taylor series

expansion. For example, an improved, three-parameter error growth model is written as follows:

$$\frac{dX}{d\tau} \approx \pm\sqrt{\sigma r}X\left(1 - \frac{X^2}{8\sigma r} - \frac{X^2}{36\sigma^2 r^2}\right), \tag{22a}$$

which can be written:

$$\frac{dE}{d\tau} \approx \pm 2\sqrt{\sigma r}E\left(\alpha - \beta E - \gamma E^2\right), \tag{22b}$$

where $E = X^2$, $\alpha = 1$, $\beta = 1/(8\sigma r)$, and $\gamma = 1/(36\sigma^2 r^2)$. Compared to the original logistic equation (e.g., Eq. 20b), the above equation includes an additional parameter γ. The specific value of $\gamma = 1/(36\sigma^2 r^2)$ was selected to produce a smooth transition towards the saturated value at the nonlinear stage, as compared to the true solution of the homoclinic orbit (Figure 3).

3 Concluding Remarks

In this study, the mathematical similarities for the solutions of homoclinic orbits and solitons within the 3D-NLM, NLS, and KdV equations were presented. The simplification of the 3D-NLM into the so-called error growth model and an improved error growth model was also discussed. As shown, the X and Z components of the homoclinic orbit, which are a hyperbolic secant function ($sech$) and a hyperbolic secant squared function ($sech^2$), respectively, have the same mathematical form as solutions for the solitary wave envelope of the NLS equation and the solitary wave of the KdV equation, respectively. Specifically, the same second order ODE for the Z component and the KdV, and the same solitary pattern solutions for both systems were obtained. Table 1 provides a summary.

A comparison of the 3D-NLM and KdV equation suggests that the heating parameter (r) determines the amplitude and pulse width of the Z component for the homoclinic orbit and that the phase speed (c) plays the same mathematical role in the solitary wave of the KdV equation. Within the 3D-NLM, the Prandtl parameter (σ) only impacts the pulse width of the Z component, and both the pulse width and amplitude of the X component. Since homoclinic orbits and solitary waves represent the limiting case for oscillatory solutions within their systems, their existence may indicate the potential presence of periodic solutions (e.g., cnoidal waves), as discussed in Appendix A. Periodic solutions should have a higher predictability than chaotic solutions. On the other hand, homoclinic bifurcation, as discussed in Appendix A, and tangling may also appear in association with the presence of homoclinic orbits. Such features are less predictable, in particular, in higher dimensional dissipative or non-dissipative Lorenz models. To understand predictability in real world systems, all of these features will be the subject of a future study.

Additionally, a future study will compare the aforementioned solutions to improve our understanding of the physical role of nonlinearity in the solutions. For example, the homoclinic orbit and oscillatory orbits of the 3D-NLM appear as a balance of nonlinearity and heating. While heating acts as a forcing, nonlinearity plays a role as a restoring force. By comparison, a balance between

nonlinearity and dispersion is found within the solitary waves of the KdV. The steepening effect of nonlinearity prevents energy dispersion, yielding an unchanged shape for the solution. Although both homoclinic and solitary wave solutions share the mathematical similarities, the physical role of nonlinearity is not necessarily the same in the 3D-NLM and the KdV equation.

How the 3D-NLM can be simplified into a Lorenz error growth model with a sigmoid function as a solution was also illustrated. Related derivations were applied to propose an improved error growth model whose solution was "verified" against the analytical solution of the homoclinic orbit. A comparison of the sigmoid function and the monotonically increasing component of the homoclinic orbit suggests that the former can reasonably represent the latter at linear and weakly non-linear stages. In comparison, solutions of the improved error growth model may provide a smoother transition towards the fully nonlinear stage with better representation for the saturated value. However, the solutions of the original and improved error growth models cannot represent the homoclinic orbit at full nonlinear stages, including the transition from a monotonically increasing component to a monotonically decreasing component. Such a transition can be captured using a cosine function that is a good approximation to the *sech* function at fully nonlinear stages (i.e., $X \approx X_{max}$ or $\tau \approx 0$). To understand the error dynamics of oscillatory solutions, which is currently being examined for a future publication, features of the transition should be included in error growth models.

Acknowledgements

We thank anonymous reviewers, M. Dunster, M. Ghil, B. Hunt, and X. Zeng for valuable comments and discussions, and Ms. S. Faghih-Naini for her help in verifying derivations.

Appendix A: Homoclinic Bifurcation

Here, I first discuss the sensitive dependence of solutions on initial conditions that are close to the homoclinic orbit. I then illustrate how the three- and higher- dimensional non-dissipative Lorenz models (NLMs) may be used to illustrate homoclinic bifurcation. The 3D NLM (3D-NLM) with an additional dissipative term (ϵY) is written as:

$$\frac{\mathrm{d}X}{\mathrm{d}\tau} = \sigma Y, \tag{A1}$$

$$\frac{\mathrm{d}Y}{\mathrm{d}\tau} = -XZ + rX - \epsilon Y, \tag{A2}$$

$$\frac{\mathrm{d}Z}{\mathrm{d}\tau} = XY. \tag{A3}$$

The above with $\epsilon = 0$ represents the 3D-NLM. Equations (A1)-(A3) can be transformed into the following equation for X:

$$\frac{\mathrm{d}^2X}{\mathrm{d}\tau^2} + \epsilon\frac{dX}{d\tau} + \left(\frac{X^2}{2} - (\sigma r + \frac{C_1}{C_o})\right) X = 0. \tag{A4}$$

Here, it should be noted that Equation (A4) is mathematically different from the so-called Burgers-KdV equation for traveling solitary wave solutions (e.g., Eq. 9 of Feng and Meng, 2007). If we simply consider the case with $C_1 = 0$ that may represent an orbit beginning or ending at the saddle point, Eq. (A4) illustrates homoclinic bifurcation with $\epsilon < 0$, $\epsilon = 0$, and $\epsilon > 0$ (e.g., Jordan and Smith, 2007), as discussed below.

Equations (A1)-(A3) were solved using the numerical methods that were discussed in Shen (2014, 2019). For the numerical solutions, $\sigma = 10.0$ and $r = 100.0$ are used. Since the 3D-NLM (i.e., Equations A1-A3 with $\epsilon = 0$) is conservative, it cannot produce the classical chaotic solutions that appear in the original dissipative 3DLM. On the other hand, as a result of the appearance of the homoclinic orbit, two nearby orbits with starting points near the homoclinic orbit may significantly diverge, showing the dependence on initial conditions. Such can be shown using 10 orbits with starting points at $(X, Y, Z) = (2\sqrt{\sigma r}, \delta, 2r)$, here δ is a small number. Selected frames at $\tau = 0.5, 0.9$, and 1.18 are provided in Figure A1, while the corresponding animation can be found at `https://goo.gl/EUiFX1`. All of the initial conditions are shown on the top left of the panels. These solutions display sensitive dependence on initial conditions but are not the classical chaotic orbits (which should be dense in the phase space). As illustrated below, they may become more complicated when dissipations are added.

The top panels in Figure A2 display an oscillatory solution with a large cycle, homoclinic orbits, and an oscillatory solution with a small cycle, respectively, within the 3D-NLM. The first and third types of solutions have initial conditions of $(X, Y, Z) = (0, 5, 0)$ and $(X, Y, Z) = (5, 0, 0)$, respectively. Homoclinic orbits were plotted using the analytical solutions provided in Eqs. (9a) and (9b). Within Eqs. (A1)-(A3) (i.e., the 3D-NLM with one dissipative term), the bottom panels reveal the so-called homoclinic bifurcation with negative ($\epsilon = -0.1$), zero ($\epsilon = 0$), and positive ($\epsilon = 0.1$) dissipative terms, respectively. The approach with the inclusion of the dissipative term (ϵY) will be applied to study the homoclinic bifurcation, as well as the dynamics of homoclinic orbits using 5D- or higher-dimensional NLMs (e.g., Faghih-Naini and Shen, 2018; Shen and Faghih-Naini, 2017) in a future study.

Appendix B: A Brief Review of the Lorenz Error Growth Model

Here, I provide a brief review on the error growth model to discuss how a nonlinear quadratic term may be included and how additional terms may be introduced to improve the error growth model.

The so-called error growth model, in the form of the logistic equation with two parameters, was first proposed in Lorenz (1969b). When Lorenz (1982) presented the model, he stated that the inclusion of the nonlinear quadratic term in the error growth model is reasonable but not readily verifiable. Since then, while the error growth model has been used to analyze the errors in real world numerical models, continual efforts have been made to provide justification for the inclusion of the nonlinear quadratic term. For example, Nicolis

(1992) provided derivations for a stochastically driven logistic equation with two parameters.

Recently, Zhang et al. (2019) proposed a revised logistic equation with two parameters that can better describe the errors of numerical simulations obtained from the state-of-the-art weather models. However, such a revised error growth model is empirical, lacking justifications for both mathematical and physical consistency between the empirical model and the data. To mitigate uncertainties, below, I first discuss major features of the solutions within the logistic model (e.g., "error saturation") in order to verify the solutions of an improved error growth model, as proposed in the main text.

Since the 1960s, a logistic equation (e.g., Eq. 20b) has been used for describing the evolution of root mean square (rms) average forecast error for ensemble runs (e.g., Lorenz 1969b, 1996; Nicolis 1992; Kalnay 2003; Zhang et al. 2019). Given initial conditions with small positive values, the solution of Equation (20b), that may represent the rms averaged error for ensemble runs, possesses the following features:

(1) it grows linearly in association with the linear term;

(2) its growth is suppressed by the nonlinear term when the error becomes relatively large; and

(3) it becomes saturated and remains at a constant value of $8\sigma r$ with a zero-growth rate; $8\sigma r$ is conventionally defined as saturated error.

The above features indicate three different stages: linear growth, nonlinear growth, and saturated stages, which appear in many nonlinear simulations with different nonlinear (dissipative) models, including: the ECMWF Integrated Forecast System (IFS) and the U.S. Finite-Volume Global Forecast System (FVGFS) (e.g., Zhang et al. 2019), the NCAR Whole Atmosphere Community Climate Model, (e.g., Liu et al. 2009), the Lorenz (1963) model (e.g., Ding and Li 2007), and the Lorenz (1984) model (Nicolis 1992), etc. Here, it should be noted that while error saturation may appear in association with the chaotic solutions of a dissipative system, such a feature may also be found in systems that display sensitivity to initial conditions. For example, as shown in Figure A1, the 3D-NLM also displays error saturation for the ensemble runs that begin close to the homolinic orbit. A system with computational chaos is another example (e.g., Lorenz, 1989).

The aforementioned features additionally illustrate a transition from an unstable point with initial positive growth rates to a stable point with zero growth rates, displaying the so-called heteroclinic dynamics (e.g., Balmforth, 1995). The features may also be found in the ensemble runs of the 3D-NLM, as discussed in Appendix A. As a result, a future task is to understand why a single orbit displays chaotic or homoclinic dynamics but rms error of ensemble runs reveals heteroclinic dynamics. In the main text, I discuss how an improved error growth model can be obtained by including higher order terms in the Taylor series expansion.

References

1. Anthes, R.: Turning the tables on chaos: is the atmosphere more predictable than we assume?, UCAR Magazine, spring/summer, available at: https://news.ucar.edu/4505/turning-tables-chaos-atmosphere-more-predictable-we-assume (last access: 14 February 2020), 2011.
2. Baines, P. G.: Topographic ects in stratified ows. Cambridge: Cambridge University Press. 482pp, 1995.
3. Balmforth, N.J.: Solitary Waves and Homoclinic Orbits. Annual Review of Fluid Mechanics 1995 27:1, 335-373.
4. Blender, R. and Lucarini, V.: Nambu representation of an extended Lorenz model with viscous heating, Physica D, 243, 86–91, 2013.
5. Boussinesq, J.: Essai sur la theorie des eaux courantes (`http://gallica.bnf.fr/ark:/12148/bpt6` k56673076), Memoires presentes par divers savants ‘ lAcad. des Sci. Inst. Nat. France, XXIII, pp. 1-680, 1987.
6. Boyd, J.P. DYNAMICAL METEOROLOGY: Solitary Waves. Encyclopedia of Atmospheric Sciences (Second Edition) 2015, Pages 417-422, 2015.
7. Curry, J. H.: Generalized Lorenz systems, Commun. Math. Phys., 60, 193–204, 1978.
8. Curry, J. H., Herring, J. R., Loncaric, J., and Orszag, S. A.: Order and disorder in two- and three-dimensional Benard convection, J. Fluid Mech., 147, 1–38, 1984.
9. Ding, R. O., and J. P. Li: Nonlinear finite-time Lyapunov exponent and predictability. Phys. Lett., 354A, 396400, 2007.
10. Faghih-Naini, S. and B.-W. Shen: Quasi-periodic in the Five-dimensional Non-dissipative Lorenz Model: The Role of theExtended Nonlinear Feedback Loop. International Journal of Bifurcation and Chaos, Vol. 28, No. 6 (2018) 1850072 (20 pages).DOI: 10.1142/S0218127418500724, 2018.
11. Feng, Z.-S. and Q.-G Meng,: Burgers-Korteweg-de Vries equation and its traveling solitary waves. Science in China Series A: Mathematics. Mar., 2007, Vol. 50, No. 3, 412422, 2007.
12. Felicio, C.C. and Rech P.C., On the dynamics of five- and six-dimensional Lorenz models, Journal of Physics Communications, vol. 2, 025028, 2018.
13. Fultz, D., R. R. Long, G. V. Owens, W. Bohan, R. Kaylor, and J. Weil, 1959: "Studies of Thermal Convection in a Rotating Cylinder with Some Implications for Large-Scale Atmospheric Motion." Meteorol. Monographs (American Meteorological Society) 21, no. 4.
14. Ghil, M., P. Read and L. Smith, 2010: Geophysical flows as dynamical systems: the influence of Hides experiments. in Astronomy & Geophysics, vol. 51, no. 4, pp. 428-435, Aug. 2010. doi: 10.1111/j.1468-4004.2010.51428.x
15. Ghil, M. and S. Childress, 1987: Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics. Springer, New York, 485pp.
16. Gleick, J.: Chaos: Making a New Science, Penguin, New York, 360 pp., 1987.
17. Guckenheimer, J., and Williams, R. F.: Structural stability of Lorenz attractors. Publ. Math. IHES. 50, 59, 1979.
18. Guckenheimer, J. and P. Holmes: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, New York, 459pp, 1983.
19. Haberman, R.: Applied Partial Differential Equations, with Fourier Series and Boundary Valule Problems. 5th Ed. Pearson Education, Inc. 756 pp, 2013.
20. Hermiz, K. B., Guzdar, P. N., and Finn, J. M.: Improved low-order model for shear flow driven by Rayleigh–Benard convection, Phys. Rev. E, 51, 325–331, 1995.

21. Hide, R., 1953: Some Experiments on Thermal Convection in a Rotating Liquid. Quart. f. Roy. Meteorol. Soc. 79 (1953): 161.
22. Hirsch, M., S. Smale, and R. L. Devaney: Differential Equations, Dynamical Systems, and an Introduction to Chaos. 3rd edition. Academic Press, 432 pp, 2013.
23. Howard, L. N. and Krishnamurti, R. K.: Large-scale flow in turbulent convection: a mathematical model, J. Fluid Mech, 170, 385–410, 1986.
24. Jordan D. W. and P. Smith: Nonlinear Ordinary Differential Equations. An Introduction for Scientists and Engineers. 4th edition. Oxford University Press, New York, 560pp, 2007.
25. Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge, New York. 369 pp, 2003.
26. Korteweg, D. J. and de Vries, G.: On the change of long waves advancing in a rectangular canal, and on a new type of long stationary waves, Phil. Mag. (Series 5) 39, pp. 422-433, 1895.
27. Li, J., and R. Ding: Temporal-spatial distribution of atmospheric predictability limit by local dynamical analogs. Mon. Wea. Rev., 139, 32653283, 2011:
28. Lighthill, J.: Waves in Fluids. Cambridge University Press, 504 pp, 1978.
29. Lin, Y.-L., and R. C. Goff: A study of a mesoscale solitary wave in the atmosphere originating near a region of deep convection. J. Atmos. Sci., 45, 194-205, 1988.
30. Liu H.-L., F. Sassi, and R. R. Garcia: Error Growth in a Whole Atmosphere Climate Model. J. Atmos. Sci., 66, 173-186, 2009.
31. Lorenz,E. N.: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130–141, 1963a.
32. Lorenz, E. N.: The mechanics of vacillation. J. Atmos. Sci., 20, 448-464, 1963b.
33. Lorenz, E. N.: The predictability of a flow which possesses many scales of motion. Tellus, 21, 289307, 1969a.
34. Lorenz, E. N.: Atmospheric predictability as revealed by naturally occurring analogues. J. Atmos. Sci., 26, 636-646, 1969b.
35. Lorenz, E. N.: Predictability: Does the flap of a butterflys wings in Brazil set off a tornado in Texas? Proc. 139th Meeting of AAAS Section on Environmental Sciences, New Approaches to Global Weather: GARP, Cambridge, MA, AAAS, 5 pp, 1972.
36. Lorenz, E.N.,: Atmospheric predictability experiments with a large numerical model. Tellus, 34, 505-513, 1982.
37. Lorenz, E.N.: Computational chaos: a prelude to computational instability. Physica, 35D, 299-317, 1989.
38. Lorenz, E. N.: The Essence of Chaos. University of Washington Press, Seattle, 227pp, 1993.
39. Lorenz, E. N.: "Predictability A problem partly solved" (PDF).Seminar on Predictability, Vol. I, ECMWF, 1996.
40. Miles, J. W.: The Korteweg-de Vries equation: a historical essay. Journal of Fluid Mechanics 106, 131-147, DOI: https://doi.org/10.1017/S0022112081001559, 1981.
41. Moon, S., B.-S. Han, J. Park, J. M. Seo and J.-J. Baik, Periodicity and Chaos of High-Order Lorenz Systems. International Journal of Bifurcation and Chaos, Vol. 27, No. 11 (2017) 1750176 (11 pages) DOI: 10.1142/S0218127417501760, 2017.
42. Musielak, Z. E., Musielak, D. E. and Kennamer, K. S.: The onset of chaos in nonlinear dynamical systems determined with a new fractal technique, Fractals, 13, 19–31, 2005.
43. Nambu, Y.: Generalized Hamiltonian dynamics, Phys. Rev. D, 7, 2403, doi:10.1103/PhysRevD.7.2405, 1973.

44. Nevir, P. and Blender, R.: Hamiltonian and Nambu representation of the non-dissipative Lorenz equations, Beitraege zur Physik der Atmosphaere, 67, 133–140, 1994.
45. Nicolis, C., 1992: Probabilistic aspects of error growth in atmospheric dynamics. Q. J. R. Meteorol. SOC. (1992), 118, pp. 553-568.
46. Pedlosky, 1972: Limit cycles and unstable baroclinic waves. J. Atmos. Sci., 29, 53-63.
47. Pedlosky, J. and C. Frenzen, 1980: Chaotic and Periodic Behavior of Finite-Amplitude Baroclinic Waves. J. Atmos. Sci. 37, 1177-1196.
48. Reyes, T. and B.-W. Shen: A Recurrence Analysis of Chaotic and Non-Chaotic Solutions within a Generalized Nine-Dimensional Lorenz Model. Chaos, Solitons & Fractals. 125 (2019), 1-12. https://doi.org/10.1016/j.chaos.2019.05.003, 2019.
49. Rottman, J. W., and R. Grimshaw: Atmospheric internal solitary waves, in Environmental Stratified Flows, Environ. Fluid Mech., edited by R. Grimshaw, chap. 3, pp. 6188, Springer, New York.
50. Rottman, J. W., and F. Einaudi: Solitary waves in the atmosphere. J. Atmos. Sci., 50, 2116-2136, 1993.
51. Roupas, Z.: Phase space geometry and chaotic attractors in dissipative nambu mechanics, J. Phys. A-Math. Theor., 45, 195101, doi:10.1088/1751-8113/45/19/195101, 2012.
52. Roy, D. and Musielak, Z. E.: Generalized Lorenz models and their routes to chaos. I. energy-conserving vertical mode truncations, Chaos Soliton. Fract., 32, 1038–1052, 2007a.
53. Russell, J S (1844). Report on Waves. in Report of the 14th Meeting of the British Association for the Advancement of Science : 311390.
54. Saltzman, B.: Finite amplitude free convection as an initial value problem, J. Atmos. Sci., 19, 329–341, 1962.
55. Shen, B.-W.: Nonlinear feedback in a five-dimensional Lorenz model, J. Atmos. Sci., **71**, 1701-1723. doi: http://dx.doi.org/10.1175/JAS-D-13-0223.1, 2014.
56. Shen, B.-W.: Nonlinear Feedback in a Six-dimensional Lorenz Model. Impact of an additional heating term. Nonlin. Processes Geophys., 22, 749-764, doi:10.5194/npg-22-749-2015, 2015.
57. Shen, B.-W.: Hierarchical scale dependence associated with the extension of the nonlinear feedback loop in a seven-dimensional Lorenz model. Nonlin. Processes Geophys., 23, 189-203, doi:10.5194/npg-23-189-2016, 2016.
58. Shen, B.-W.: On an extension of the nonlinear feedback loop in a nine-dimensional Lorenz model. Chaotic Modeling and Simulation (CMSIM), 2: 147157, 2017.
59. Shen, B.-W.: On periodic solutions in the non-dissipative Lorenz model: the role of the nonlinear feedback loop. Tellus A: 2018, 70, 1471912, https://doi.org/10.1080/16000870.2018.1471912, 2018.
60. Shen, B.-W.: Aggregated Negative Feedback in a Generalized Lorenz Model. International Journal of Bifurcation and Chaos, Vol. 29, No. 3 (2019) 1950037 (20 pages). https://doi.org/10.1142/S0218127419500378, 2019a.
61. Shen, B.-W.: On the Predictability of 30-day Global Mesoscale Simulations of Multiple African Easterly Waves during Summer 2006: A View with a Generalized Lorenz Model. Geosciences 2019, 9(7), 281; https://doi.org/10.3390/geosciences9070281, 2019b.
62. Shen, B.-W., T. Reyes, and S. Faghih-Naini: Coexistence of Chaotic and Non-Chaotic Orbits in a New Nine-Dimensional Lorenz Model. In: Skiadas C., Lubashevsky I. (eds) 11th Chaotic Modeling and Simulation International Conference. CHAOS 2018. Springer Proceedings in Complexity. Springer, Cham. `https://doi.org/10.1007/978-3-030-15297-0_22`, 2019.

63. Shen, B.-W., R. A. Pielke Sr., X. Zeng, J.-J. Baik, T.A.L. Reyes, S. Faghih-Naini, R. Atlas, and J. Cui: Is Weather Chaotic? Coexistence of Chaos and Order within a Generalized Lorenz Model. 100th AMS meeting. Boston, MA, 12-16 January 2020. (a manuscript also submtited to BAMS, which is available from ResearchGate http://doi.org/10.13140/RG.2.2.21811.07204), 2020.
64. Shen, B.-W. and S. Faghih-Naini: On recurrent solutions in high-dimensional non-dissipative Lorenz models. The 10th Chaos Modeling and Simulation International Conference (CHAOS2017), Barcelona, Spain, 30 May - 2 June, 2017.
65. Smale, S.: Mathematical Problems for the Next Century. The Mathematical Intelligencer 20, no. 2, pages 715, 1998.
66. Smith, R.K.:A note on a theory of vacillating baroclinic waves. J. Atmos. Sci., 32, 2027, 1975.
67. Smith, R. K. and J. M. Reilly: On a theory of amplitude vacillation in baroclinic waves: Some numerical solutions. J. Atmos. Sci., 34, 1256-1260, 1977.
68. Sparrow, C.: The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors. Springer, New York. Appl. Math. Sci., 269 pp, 1982:
69. Sprott, J. C.: Elegant Chaos. World Scientific, Singapore. 285 pp, 2010.
70. Strogatz, S. H.: Nonlinear Dynamics and Chaos. With Applications to Physics, Biology, Chemistry, and Engineering. Westpress view, Boulder, CO, 513 pp, 2015.
71. Thiffeault, J.-L. and Horton, W.: Energy-conserving truncations for convection with shear flow, Phys. Fluids, 8, 1715–1719, 1996.
72. Tucker, W.: A Rigorous ODE Solver and Smale's 14th Problem. Found. Comput. Math. 2, 53-117, 2002.
73. Whitham, G. B.: Linear and Nonlinear Waves. John Wiley & Sons, Inc., New York. 636 pp.
74. Zabusky, N. J. and M. A. Porter: Soliton. Scholarpedia, 5(8):2068. doi:10.4249/scholarpedia.2068, 2010.
75. Zabusky, N. J. and Kruskal, M. D.: Interaction of Solitons in a Collisionless Plasma and the Recurrence of Initial States. Phys. Rev. Let. 15, 240-243, 1965.
76. Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, K. Emanuel, What Is the Predictability Limit of Midlatitude Weather? J. of Atmos. 76, 1077-1091, 2019.

Table 1: A comparison of the 3D-NLM, Duffing, NLS, KdV, and Logistic equations. *cn*, *sech*, and *Sig* represent the Jacobi elliptic, hyperbolic secant, and sigmod functions, respectively. See details in the main text.

3D-NLM	Other Models	Solutions
The Equation for X'' $\frac{\mathrm{d}^2X}{\mathrm{d}\tau^2} - (\sigma r + \frac{C_1}{C_o})X + \frac{X^3}{2} = 0$	The Duffing Equation $\frac{d^2X}{d\tau^2} + \delta\frac{dX}{d\tau} + \alpha X + \beta X^3 = \gamma cos(\omega\tau)$ $\delta = 0$, $\gamma = 0$, $\alpha = -(\sigma r + \frac{C_1}{C_o})$, $\beta = 1/2$	cn
The Equation for $(X')^2$ $\left(\frac{dX}{d\tau}\right)^2 - \sigma r X^2 + \frac{X^4}{4} = 0$	The Nonlinear Schrodinger Equation $\left(\frac{dh}{dx}\right)^2 + \delta h^2 + \frac{\gamma}{2}h^4 = 0$ $\delta < 0$, $\gamma > 0$	$sech$
The Equation for Z'' $\frac{d^2Z}{d\zeta^2} + 3Z^2 - 4rZ = 0$ $\zeta = \sqrt{\sigma}\tau$	The Korteweg-de Vries Equation $\frac{d^2f}{d\xi^2} + 3f^2 - cf = 0$ $c = 4r$	$sech^2$
The Equation for X' $\frac{dX}{d\tau} = \pm\sqrt{\sigma r}X\sqrt{1 - \frac{X^2}{4\sigma r}}$	The Logistic Equation $\frac{dE}{d\tau} \approx \pm 2\sqrt{\sigma r}E\left(1 - \frac{E}{8\sigma r}\right)$, $E = X^2$	$\approx Sig$

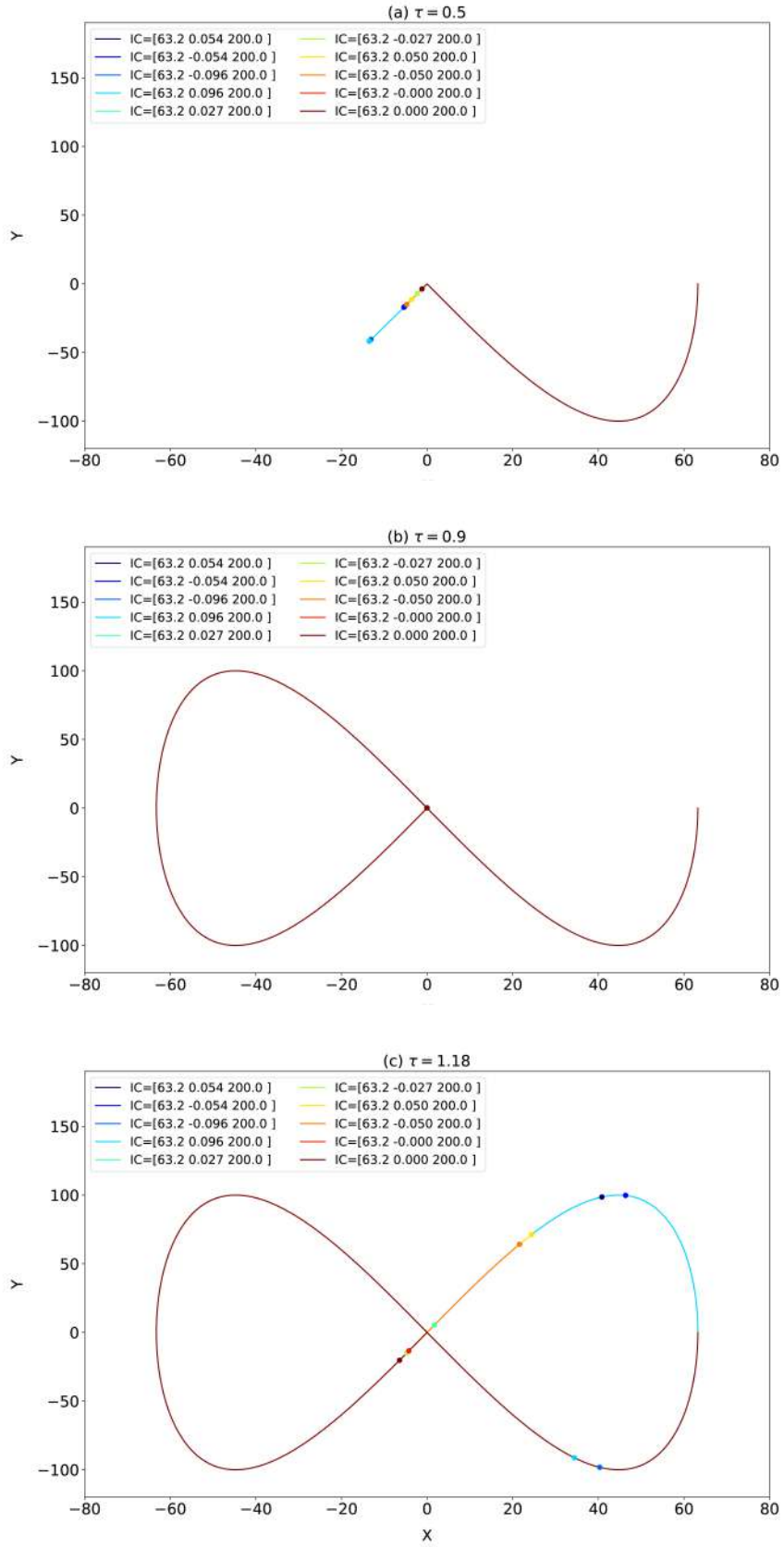


Figure A1: Sensitivity to initial conditions within the 3D-NLM. Ten orbits begin at $(X, Y, Z) = (2\sqrt{\sigma r}, \delta, 2r)$, here δ is a small number. Panels (a)-(c) display selected frames at $\tau = 0.5, 0.9$, and 1.18, respectively. The corresponding animation can be found at `https://goo.gl/EUiFX1`.

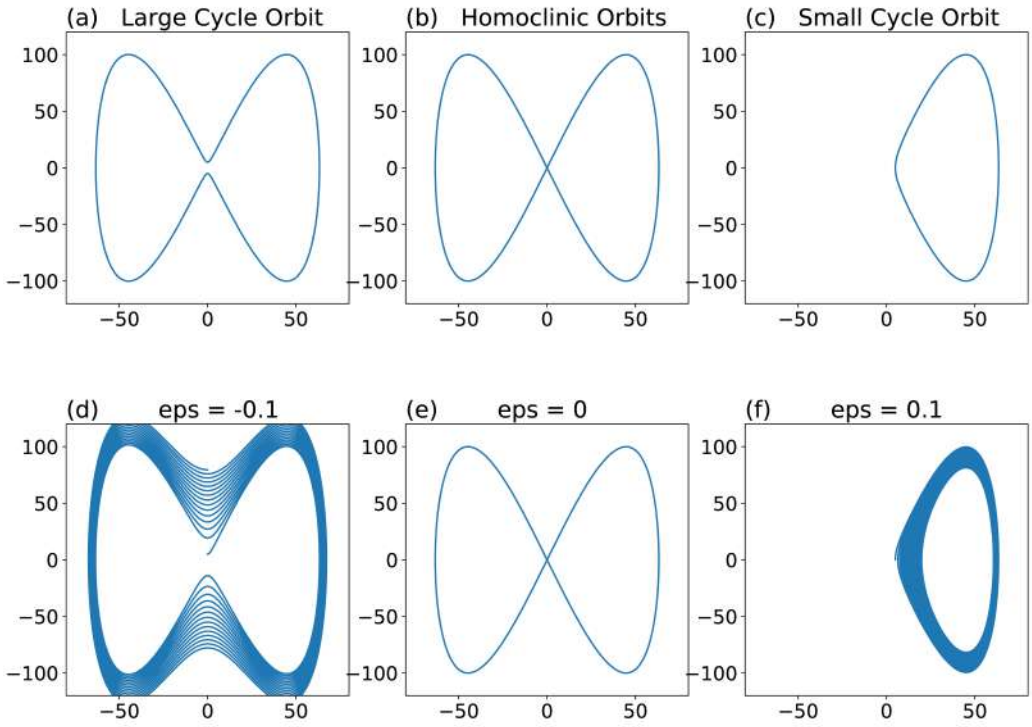


Figure A2: Top panels display two types of oscillatory solutions (a, c) and homoclinic orbits (b) within the 3D-NLM. Bottom panels (d)-(f) illustrate homoclinic bifurcation using Eqs. (A1)-(A3) with negative, zero, and positive dissipations, respectively. Panels (a, d) display numerical solutions with the same initial condition of $(X, Y, Z) = (0, 5, 0)$, while panels (c, f) use the same initial condition of $(X, Y, Z) = (5, 0, 0)$. Panels (b, e) display analytical solutions depicted in Eq. (9).

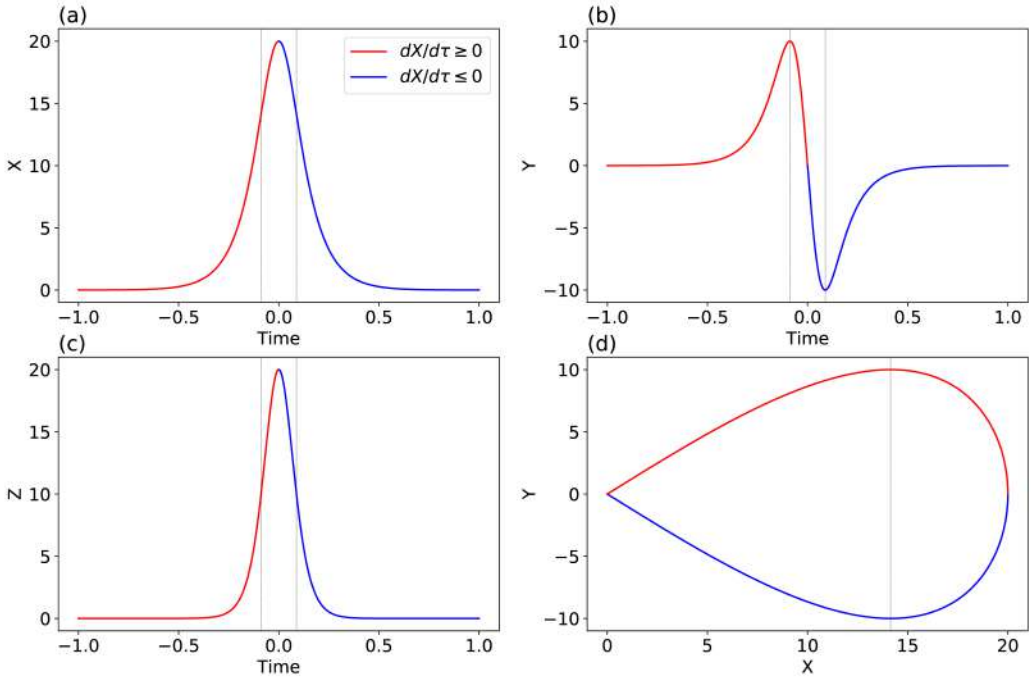


Figure 1: Solutions of the homoclinic orbit for X (a), Y (b), and Z (c) within the 3D-NLM for $\sigma = r = 10$. The homoclinic orbit in the $X - Y$ space is shown in panel (d). Blue and red lines represent the monotonically decreasing and increasing components of X (as well as Z) for the homoclinic orbit, respectively. Thin vertical lines are plotted at $\tau = \pm\tau_t = \pm 0.0881$ or $X = X_t = 14.14$, where $d^2X/d\tau^2 = 0$. Pulse width is determined by $[-\tau_t, \tau_t]$.

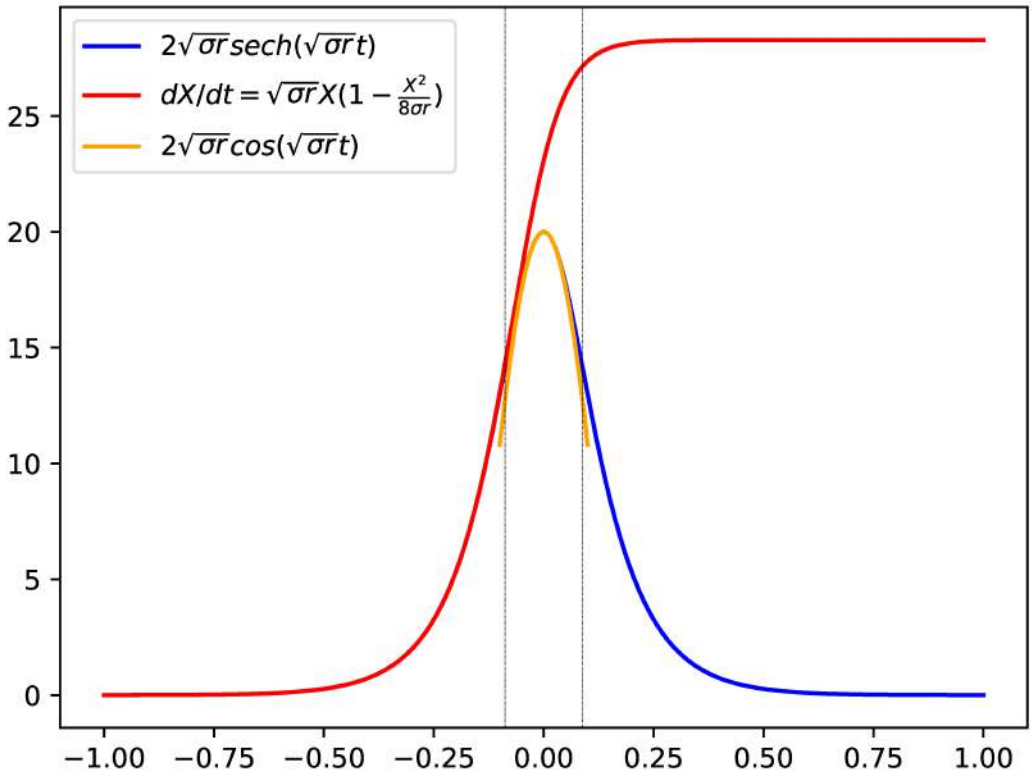


Figure 2: A comparison of the homoclinic orbit (in blue), a sigmoid function (in red), and a cosine function (in orange), showing the non-negative growth rates and the non-monotonic features of the homoclinic orbit. Two vertical lines indicates the time (τ_t) when $X = X_t = \sqrt{2\sigma r}$. Note that $d^2X/d\tau^2 < 0$ appears in association with stronger nonlinearity for $X \in (X_t, X_{max})$, i.e., $\tau \in (-\tau_t, \tau_t)$.

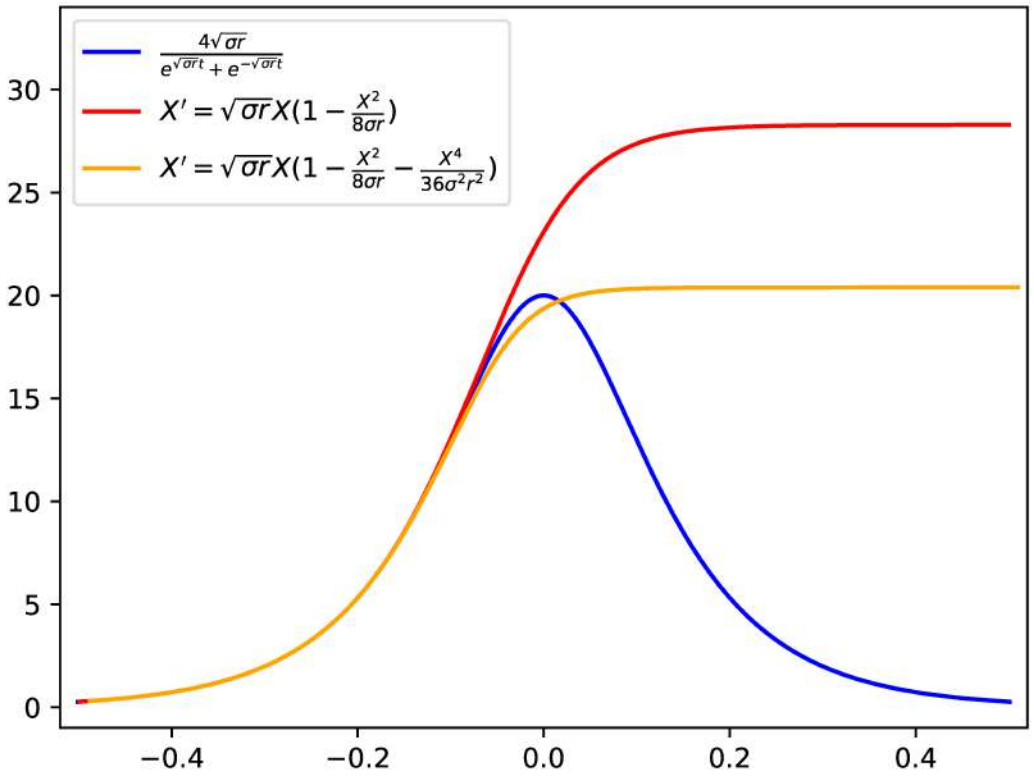


Figure 3: A comparison of the homoclinic orbit (in blue) with approximated solutions, a sigmoid function of the logistic equation (in red), and a solution of the three-parameter error growth model (in orange).

Is Weather Chaotic?
Coexisting Chaotic and Non-Chaotic Attractors within Lorenz Models

B.-W. Shen[1,*], R. A. Pielke Sr.[2], X. Zeng[3], J.-J. Baik[4],
S. Faghih-Naini[1,5], J. Cui[1,6], R. Atlas[7], and T. A. L. Reyes[1]

[1] San Diego State University
E-mail: bshen@mail.sdsu.edu
[2] CIRES and ATOC, University of Colorado at Boulder
[3] The University of Arizona
[4] Seoul National University
[5] University of Bayreuth and Friedrich-Alexander University Erlangen-Nuremberg
[6] North Carolina State University
[7] AOML, National Oceanic and Atmospheric Administration

Abstract.

The pioneering study of Lorenz in 1963 and a follow-up presentation in 1972 changed our view on the predictability of weather by revealing the so-called butterfly effect, also known as chaos. Over 50 years since Lorenz's 1963 study, the statement of ``weather is chaotic'' has been well accepted. Such a view turns our attention from regularity associated with Laplace's view of determinism to irregularity associated with chaos. Here, a refined statement is suggested based on recent advances in high-dimensional Lorenz models and real-world global models. In this study, we provide a report to: (1) Illustrate two kinds of attractor coexistence within Lorenz models (i.e., with the same model parameters but with different initial conditions). Each kind contains two of three attractors including point, chaotic, and periodic attractors corresponding to steady-state, chaotic, and limit cycle solutions, respectively. (2) Suggest that the entirety of weather possesses the dual nature of chaos and order associated with chaotic and non-chaotic processes, respectively. Specific weather systems may appear chaotic or non-chaotic within their finite lifetime. While chaotic systems contain a finite predictability, non-chaotic systems (e.g., dissipative processes) could have better predictability (e.g., up to their lifetime). The refined view on the dual nature of weather is neither too optimistic nor pessimistic as compared to the Laplacian view of deterministic unlimited predictability and the Lorenz view of deterministic chaos with finite predictability.

13th *CHAOS Conference Proceedings, 9 - 12 June 2020*

1 Introduction

Is weather chaotic? A view that weather is chaotic was proposed and is recognized based on the pioneering work of Lorenz (1963a) who first introduced the concept of deterministic chaos. Defined as aperiodic solutions that display sensitive dependence on initial conditions (ICs), chaos is also known as the butterfly effect. In a follow-up conference presentation in 1972 (Lorenz 1972), the concept of sensitivity to ICs was further discussed by addressing whether a butterfly's flap may lead to a chain of responses that remotely generates a tornado. Since then, the butterfly effect has come to be known as a metaphor for indicating the huge impact of a tiny perturbation on the formation of a tornado. The original Lorenz 1963 study and the 1972 presentation, as well as his 1969 study (Lorenz 1969a), laid the foundation for chaos theory that is viewed as the third scientific achievement in the 20th century, after relativity and quantum mechanics, inspiring numerous studies in multiple fields, including earth science, mathematics, philosophy, physics, etc. (Gleick 1987).

While periodic solutions were a main focus until the Lorenz (1963) study, non-periodic solutions have increasingly received attention over the past fifty years. Lorenz's discovery has led to the statement of "weather is chaotic" and to a paradigm shift in the view of finite predictability from the Laplacian view of unlimited deterministic predictability. The idea of finite predictability for chaotic weather has prompted a search for the upper limit of predictability that was determined as two weeks based on the analyses of unstable solutions from simplified models and data (e.g., Lorenz 1969a). With the above being said, our current view on the chaotic nature of weather and a predictability limit of two weeks are based on the understanding of chaotic (as well as unstable) solutions obtained from elegant but simple models. To facilitate discussions, we define two kinds of predictability, including (1) intrinsic predictability that is dependent only on flow itself and (2) practical predictability that is limited by the imperfect initial conditions and/or (mathematical) formulas (Lorenz 1963b; Shen 2014).

Chaotic solutions are just one type of solution that occurs over finite intervals of time-independent parameters within the Lorenz model. To reveal the true nature of weather, we should take into consideration other types of solutions within original Lorenz models and newly developed generalized Lorenz models (Guckenheimer and Williams 1979; Sparrow 1982; Pielke and Zeng 1994; Smale 1998; Tucker 2002; Musielak et al. 2005; Roy and Musielak 2007; Yang and Chen 2008; Sprott et al. 2013; Moon et al. 2017, 2019; Felicio and Rech 2018; Macek 2018; Faghih-Naini and Shen 2018; Reyes and Shen 2019; Shen 2014-2018, 2019a). For example, in addition to chaotic solutions, two types of non-chaotic solutions indeed appear over different intervals of parameters within the Lorenz model (Sparrow 1982). Furthermore, recent studies using a generalized high-dimensional Lorenz model (e.g., Shen 2019a; Shen et al. 2019; Reyes and Shen 2019) showed that chaotic and non-chaotic solutions may coexist within the

same model parameters but for different ICs (e.g., Sprott et al. 2005; Sprott and Xiong 2015). Thus, it is important to understand whether or not and how other types of solutions and their coexistence may help illustrate a more comprehensive view on the nature of weather, and improve our understanding of predictability associated with different types of solutions. Specifically, we may ask whether the statement of ``weather is chaotic'' that exclusively considers chaotic solutions is scientifically precise.

To address the above, here, we first provide a review of major solutions using the Lorenz model (LM), including three types of solutions or three attractors in Section 2. In this study, a specific type of solution is referred to as an ``attractor'', defined as the smallest attracting point set that cannot be decomposed into two or more subsets with distinct regions of attraction (e.g., Sprott et al. 2013). We then summarize our recent findings for two kinds of attractor coexistence (i.e., with the same model parameters but with different initial conditions) using a newly developed, generalized, high-dimensional LM (GLM) (e.g., Shen 2019a) in Section 3. Section 4 is presented in order to support the findings for two kinds of attractor coexistence using the original LM with different parameters. Based on an analysis of the LM and the GLM, we suggest a refined view on the dual nature of weather in Section 5. Additional support for this view is also presented by the review of prior studies. Concluding remarks are provided in Section 6.

2. The Lorenz 1963 Model

In his 1963 study, Prof. Lorenz presented an elegant system of three ordinary differential equations (ODEs) derived from the governing equations for the Rayleigh-Benard convection (e.g., Saltzman 1962; Lorenz 1963a). The system describes the time evolution of three variables, X, Y, and Z, as follows:

$$\frac{dX}{d\tau} = \sigma Y - \sigma X, \tag{1}$$

$$\frac{dY}{d\tau} = -XZ + rX - Y, \tag{2}$$

$$\frac{dZ}{d\tau} = XY - bZ. \tag{3}$$

Here, τ is the dimensionless time. Three time-independent parameters include the Prandtl number (σ), the normalized Rayleigh number (r), also called the heating parameter, and a function of the ratio between the vertical and horizontal scales of the convection (b). $(\mathrm{X}, \mathrm{Y}, \mathrm{Z})$ represent the amplitudes of the three Fourier modes for dynamic and thermodynamic variables. The system contains three types of physical processes, including buoyancy/heating terms (represented by σY and rX), dissipative terms (represented by $-\sigma$X, $-$Y, and $-$bZ), and nonlinear processes (indicated by $-$XZ and XY). With the exception of the heating parameter (r), the following parameters are kept constant: $\sigma = 10$ and $\mathrm{b} = 8/3$. Control and parallel runs are performed in order to reveal the difference (or divergence) of two solutions. The only difference between control and parallel runs is that a

parallel run includes tiny perturbations ($\epsilon = 10^{-10}$) or finite perturbations ($\epsilon = -0.9$) in initial conditions.

Using the state variables X, Y, and Z as coordinates, a phase space can be constructed for an analysis of solutions. An orbit or a trajectory is defined as the time varying components of solutions within the phase space. The dimension[1] of the phase space is equal to the number of time-dependent variables or to the number of ODEs. Thus, equations (1)-(3) with three variables are referred to as a three-dimensional Lorenz model (3DLM). High-dimensional LMs contain more than three variables.

Lorenz's Chaotic and Non-Chaotic Attractors

Depending on the competitive or collective impact of nonlinear processes and linear buoyancy/heating and dissipative processes, various types of solutions (i.e., different attractors) appear within the Lorenz model. Historically, the dependence of their appearance on the strength of heating measured by the parameter (r) has been a focus. Steady-state, chaotic, and nonlinear oscillatory solutions have been shown to occur under conditions of weak, moderate, and strong heating, respectively (e.g., Sparrow 1982; Drazin 1992)[2]. In Fig. 1., the three different types of solutions are shown using r = 20, 28, and 350, respectively. The top panels display solutions for control runs within the *X*-*Y* space, while bottom panels display the time evolution of the Y components for both control and parallel runs. For a steady-state solution, its orbit eventually approaches a single point, that is, a non-trivial equilibrium point within the *X*-*Y* space (Fig. 1a), appearing as a point attractor; and its amplitude remains constant over time after arriving at the equilibrium point. Mathematically, equilibrium points, also called critical points, are defined as solutions of the time-independent nonlinear system (e.g., no time derivatives in Eqs. (1)-(3), Guckenheimer and Holmes 1983)[3]. When the heating parameter exceeds the critical value of $r_c = 24.74$, the 3DLM with r = 28 displays the so-called chaotic solution or a chaotic attractor with irregular oscillations. The solution's boundary within the *X*-*Y* space appears as a

[1] The term "dimension" is conventionally used for a system of ODEs (e.g., Hirsch et al. 2013; Thompson and Stewart 2002). In this study, the 5DLM, 7DLM, and 9DLM are referred to as high-dimensional or high-order Lorenz models (e.g., Moon et al. 2017).

[2] Similar findings for the dependence of various solutions (i.e., chaotic and limit cycle solutions) on the strength of heating were also reported using a two-layer, quasi-geostrophic model that describes the finite-amplitude evolution of a single baroclinic wave by Pedlosky and Frenzen (1980).

[3] In our 5D-, 7D-, and 9D LMs, we can obtain closed form solutions of trivial and non-trivial equilibrium points and use them to verify the numerical solutions of equilibrium points.

tilted "8" pattern. Interestingly, when heating becomes larger (e.g., r = 350), the system produces a nonlinear periodic solution known as a limit cycle solution or a periodic attractor, as shown in Figs. 1c and 1f. Additional details on the characteristics of nonlinear oscillatory solutions may be found in earlier studies (e.g., Shimizu 1979; Sparrow 1982; Strogatz 2015) and/or recent studies (e.g Reyes and Shen 2019; Shen 2019a, b). Below, the impact of a tiny initial perturbation on three attractors, including a point attractor, a chaotic attractor, and a periodic attractor, is further discussed.

Parallel runs with a tiny initial perturbation ($\in= 10^{-10}$) are compared to control runs in order to reveal the difference of initial, nearby trajectories within the phase space. For steady-state and nonlinear oscillatory solutions, control and parallel runs produce almost identical results, only appearing in red, for example, in Figs. 1d and 1f. The runs indicate insignificant impacts by a tiny initial perturbation. In other words, steady-state and nonlinear oscillatory solutions are insensitive to a tiny change in ICs. In comparison, within the chaotic regime, two solution orbits whose starting points are very close to each other display very different time evolutions, as clearly shown in blue and red in Fig. 1e. The phenomenon is called the sensitive dependence of solutions on ICs and only appears within a chaotic solution.

Boundedness and Divergence of Chaotic Trajectories

Within the chaotic regime, a sensitive dependence of solutions on ICs is referred to as the butterfly effect (BE, e.g., Lorenz 1993, 2008). As shown in Fig. 2a, the term ``butterfly'' was partly used due to its geometric pattern within the *Y*-*Z* space (e.g., Lorenz 1993). A butterfly pattern with a finite size and varying curvatures within the phase space also qualitatively suggests an important feature of solution boundedness. Therefore, BE means that a tiny change in an IC can produce a very different time evolution of a solution for three variables (X, Y, Z). However, the separation (or divergence) of two orbits should be bounded by the size of a butterfly pattern.

The average separation rate (i.e., an average rate of divergence) of nearby trajectories has been quantitatively measured using the Lyapunov exponent (LE, Wolf et al. 1985; Zeng et al. 1991, 1993). A positive LE suggests an exponential rate in the averaged separation of two infinitesimally nearby trajectories over an infinite period of time (e.g., Eqs. (25)-(26) of Shen 2014). Chaotic solutions within the 3DLM, as well as high-dimensional LMs, have a positive LE. Since the LE is defined as a long-term averaged separation, researchers often misinterpret the divergence of two nearby, but finitely separated, chaotic trajectories within the 3DLM as continuing over time and lasting forever. The misunderstanding also makes people believe that an unconstrained solution is due to the divergent nature of chaos (e.g., Hilborn 2000). In fact, in addition to a positive LE, solution boundedness is another major feature of a chaotic system.

Due to solution boundedness, a trajectory should recurve within the phase space (e.g., Hilborn 2000). Therefore, time-varying (local) growth rates along a chaotic orbit are observed (e.g., Zeng et al. 1993) and may become negative, as indicated by a negative finite time LE (e.g., Nese 1989; Eckhardt and Yao 1993; Ding and Li 2007; Bailey 2011). In other words, the infinite-time limit in the definition of an LE does not imply a monotonically increasing separation between two nearby trajectories over a long period of time. Two initial nearby trajectories can quickly separate and reach the bound of their separation.

3. The Generalized Lorenz Model

The 3DLM produces three different attractors and each attractor exclusively appears within the phase space, depending on the interval of system parameters. The 3DLM with a single-type solution suggests that either chaos or order exclusively exists. By comparison, two different solutions may coexist and dominate system dynamics in a separate region (i.e., a different subspace) within the phase space within the same model, and with the same parameters, but with different ICs. Attractor coexistence has mainly been studied using conservative Hamiltonian systems (e.g., Hilborn 2000), but can also be found in the forced dissipative 3DLM (e.g., Yorke and Yorke 1979; Drazin 1992; Ott 2002). Below, we first discuss two kinds of attractor coexistence using the GLM, and then apply the GLM to understand the whether the 3DLM can also possess two kinds of attractor coexistence.

Based on our recent studies (e.g., Shen, 2014-2019a), we successfully developed a GLM that: (1) is derived based on partial differential equations for the Rayleigh-Benard convection[4]; (2) allows a large number of modes, say M modes, where M is an odd number greater than three; and (3) produces aggregated negative feedback[5] that is accumulated from the feedback of various smaller-scale processes, yielding a larger effective dissipation in higher dimensional LMs (Shen 2019a; Shen et al. 2019). *As a result of aggregated negative feedback, a higher-dimensional LM requires a larger critical value for the Rayleigh parameter (r_c) for the onset of chaos.* For example, the r_c for the 5DLM, 7DLM, and 9DLM are 42.9, 116.9, and 679.8, respectively, as compared to a r_c of 24.74 for the 3DLM (Shen 2019a). Fig. 2 displays chaotic solutions obtained from the 3D, 5D, 7D, and 9D LMs with different heating parameters. Therefore, an intial tiny perturbation with the same strength may play a different role within the GLM with a different value of M. Such a feature shows a dependence on the number

[4] By comparison, chaotic models in Lorenz (1996/2006, 2005) were not derived from physics-based partial differential equations.

[5] Negative feedback can be found within the so-called Lorenz-Stenflo system that extends the 3DLM with one additional ODE containing one additional mode that takes rotation into consideration (e.g., Xavier and Rech 2010; Park et al. 2015, 2016).

of selected modes. Namely, it depends on the degree of spatial complexity associated with a various number of modes of the GLM.

Two Kinds of Attractor Coexistence

The GLM with M = 5 or M = 7 (i.e., 5DLM or 7DLM) also produces three different types of solutions, including a steady-state, chaotic, and limit cycle/torus[6]. More importantly, the GLM with M = 9 (i.e., 9DLM) displays two kinds of attractor coexistence, each with two different attractors. For the first kind of coexistence, both chaotic and steady-state solutions occur concurrently using the same model and the same parameters. The only difference is their ICs. Such a coexistence shares properties similar to that of the 3DLM but appears over a wider range of the Rayleigh parameter (e.g., $679.8 < r < 1{,}058$), as compared to the small interval (e.g., $24.06 < r < 24.74$) for the 3DLM.

In addition to the first kind of attractor coexistence, the 9DLM is able to produce the second kind of attractor coexistence, consisting of nonlinear, periodic (i.e., limit cycle) orbits, and steady-state solutions at large Rayleigh parameters (e.g., r = 1,600). The new kind of coexistence was recently documented in Shen (2019a), Shen et al. (2019), and Reyes and Shen (2019). Additionally, coexisting two periodic solutions were documented using the 9DLM with r = 1120 (e.g., Shen 2019a). As a result, when system parameters change at a large time scale (e.g., at climate time scales), different kinds of attractor coexistence may alternatively or concurrently appear, leading to complexities that better resemble real weather and climate.

Two Kinds of IC Dependence and Final State Sensitivity

Depending on system parameters, ICs and the dimension of the model (say the value of M within the GLM), a modeling system may contain one or more attractors[7] within the phase space. Since different attractors coexist, we expect different kinds of solution dependence on ICs, as illustrated using the 9DLM with r = 680 that produces the coexistence of chaotic and non-chaotic orbits. Control runs apply three sets of ICs at different locations within the phase space: close to the non-trivial equilibrium point, near the origin (i.e., a saddle point), and at point (100, 100, 100, 100, 100, 100, 100, 100, 100). For parallel runs, a finite-amplitude perturbation ($\in = -0.9$) is added into the ICs. In Fig. 3, solutions of the control runs are shown in blue, while results of parallel runs are displayed in green, red, or orange. Top panels display the time evolution of the Y components, while bottom panels present solutions within the *X*-*Y* space. The model with r = 680

[6] A torus is defined as a composite motion with two (or more) oscillatory frequencies whose ratio is irrational (e.g., Faghih-Naini and Shen 2018).
[7] The coexistence of chaotic and quasi-periodic orbits has been recently documented in a modified Lorenz system by Saiki et al. (2017).

produces the coexistence of steady-state and chaotic orbits, displaying a dependence on ICs. For the first case (Figs. 3a and 3d) with the IC that is close to the non-trivial equilibrium point, the orbit moves toward the equilibrium point, producing steady-state solutions. Since the orbit spirals into the non-trivial equilibrium point within the *X*-*Y* space, it is also called a spiral sink solution. For the second case (Figs. 3b and 3e) where an IC is close to a saddle point at the origin but away from the non-trivial equilibrium point, solutions still approach the same non-trivial equilibrium point as a steady-state solution, while initially displaying a different time evolution as compared to the first case. On the other hand, for the third case (Figs. 3c and 3f), the model produces a chaotic solution, different from the steady-state solution. A comparison between control and parallel runs suggests that an initial perturbation only has a short-term impact on the initial transient evolution of steady-state solutions but can lead to a very different evolution for chaotic solutions[8].

When coexisting chaotic and regular attractors from 256 different initial conditions are plotted within the X-Y phase space, Figure 4 clearly shows that chaotic and non-chaotic orbits occupy two different regions (or two different subspaces). Additional details on the spatial distribution of 256 initial conditions may be found in Shen et al. (2019). As a result of the different regions of attraction for coexisting attractors, final state sensitivity may appear (e.g., Grebogi et al. 1983) when ICs begin near the boundary of two different attractors. Such a sensitivity creates a different challenge for prediction.

Finite and Deterministic Predictability

The rate of a growing initial error with time has been used to determine predictability, suggesting a finite predictability in chaotic (or unstable) systems. Such a growth rate is proportional to the divergence of two nearby trajectories measured using a Lyapunov exponent. Within chaotic regimes of the 3DLM, as well as within the GLM that contains one positive LE and solution boundedness, time-varying divergence and convergence of nearby trajectories yields time-varying growth rates and, thus, time-varying predictability. Estimated predictability over a short period should display a dependence on various initial states[9]. By comparison, when non-chaotic (i.e., steady-state or nonlinear periodic) solutions appear as a single type of solution or coexist with another type of solution, their predictability should be deterministic (unlimited). Stated conservatively, the non-chaotic solution should remain predictable until it is changed by time varying parameters that represent heating or dissipations. As a

[8] Such a dependence on initial conditions, close to (or away from) the non-trivial equilibrium point, can be shown by the following YouTube video for a double pendulum (between 1:00-1:20): https://www.youtube.com/watch?v=LfgA2Auyo1A.

[9] As a result, we agree with Prof. Arakawa that the predictability limit is not necessarily a fixed value (Lewis 2005).

result, very different intrinsic predictability may appear and depend on ICs within a system that possesses the coexistence of chaotic and non-chaotic attractors.

4. Attractor Coexistence within the 3DLM

Within chaotic solutions of the 3DLM that has no stable equilibrium points, a tiny perturbation can always lead to a very different time evolution. Stated alternatively, within the chaotic regime, the system, in the absence of energy sinks for steady-state solutions, does not have a mechanism for completely removing the impact of a tiny perturbation on state variables. By comparison, within the GLM with M = 9, or higher, that possesses coexisting chaotic and steady-state solutions, a tiny initial perturbation may play a very different role. A tiny perturbation may have no long-term impact when it appears to be associated with a steady-state solution that approaches one of stable equilibrium points, suggesting that the perturbation eventually dissipates. On the other hand, a tiny perturbation may lead to a large impact on the time evolution of the chaotic solution. As a result, the 9DLM with a dual role for a tiny initial perturbation over a wide range of the heating parameter, as well as other features such as hierarchical scale dependence, is more realistic than the classical 3DLM with typical parameters. On the other hand, we may ask whether the 3DLM with different parameters may also produce two kinds of attractor coexistence, providing additional support to the findings of the GLM.

Next, we first discuss the coexistence of the 3DLM with typical parameters that include $\sigma = 10$. We then address the question of whether $\sigma = 10$ is a magic choice. As simply shown in the animation, https://goo.gl/scqRBo, the 3DLM with the same parameters, including r = 24.4, $\sigma = 10$, and b = 8/3, but with different ICs, produces two types of solutions that include chaotic or steady-state solutions, yielding the first kind of attractor coexistence. However, such a coexistence only appears over a very small range of r, giving the length of an interval less than 0.7 (i.e., $24.06 < r < r_c = 24.74$), and, thus, its characteristics and potential role in revealing the nature of weather has not been well appreciated.

For the past 50 years, although various types of solutions for Lorenz (1963) have been documented, chaotic solutions have been the main focus. As discussed in the main text, since chaotic solutions appear over a finite range of parameters, their applicability in revealing the nature of weather depends on the realism of not only the models employed but also model parameter values. In his book in 1993, Lorenz humbly expressed that it may not have been possible for him to discover the butterfly-pattern solution if a realistic value of $\sigma = 1$ was used, as shown below:

``

I was lucky in more ways than one. An essential constant of the model is the Prandtl number -- the ratio of the viscosity of the fluid to the thermal conductivity. Barry had chosen the value 10.0 as having the order of magnitude of the Prandtl

number of water. As a meteorologist, he might well have chosen to model convection in air instead of water, in which case he would probably have used the value 1.0. With this value the solutions of the three equations would have been periodic, and I probably would never have seen any reason for extracting them from the original seven.
''

Therefore, one may wonder how fortunate Prof. Lorenz was and whether a realistic value of $\sigma = 1$ may have influenced our view on the nature of weather. We make an attempt of addressing the question by analyzing a GLM with M = 9 and examining a 3DLM with $\sigma = 1$. As discussed in Shen (2019a), the GLM with M = 9 has stable, non-trivial equilibrium points for all $r > 1$ when $\sigma = 10$ and b = 8/3. To have stable, non-trivial equilibrium points for $\sigma = 1$ within the 3DLM, we chose b = 2/5. Such a choice leads to two kinds of attractor coexistence, a unique feature first identified within the 9DLM (Shen 2019a). With $\sigma = 1$ in the 3DLM, the first kind of coexistence includes chaotic and steady-state solutions at a moderate heating parameter (e.g., r = 170, as shown in Fig. 5). Table 1 lists initial conditions for the results provided in Fig. 5. Thus, chaotic solutions may still appear within the 3DLM for a realistic value of $\sigma = 1$, but they coexist with steady-state solutions. The appearance of chaotic solutions depends not only on the range of the heating parameter but also on the ICs. Additionally, the second kind of coexistence that consists of a limit cycle and a steady-state solution appears at a large heating parameter (e.g., r = 250, not shown).

Both traditional and new model configurations with $(\sigma, b) = (10, 8/3)$ and $(1, 2/5)$, respectively, can produce chaotic solutions. For the traditional configuration that has been well applied in numerous studies since Lorenz (1963a), all three equilibrium points are unstable when $r > 24.74$. The stability of the three equilibrium points for $\sigma = 10$, as well as for $\sigma = 1$, is illustrated in Fig. 6. The non-existence of stable equilibrium points within the chaotic regime makes it easier to obtain chaotic solutions. However, no tiny, initial perturbation can completely lose its impact within the chaotic regime. We may interpret this as a finding that a tiny, initial perturbation cannot completely dissipate (before leading to a large impact). By comparison, for the new configuration of $\sigma = 1$, while the origin is still a saddle point, the two, non-trivial equilibrium points are stable (Fig. 6b). The existence of stable equilibrium points enables the coexistence of chaotic and steady-state solutions, the latter of which has no long-term memory regarding a tiny, initial perturbation.

As a result of coexistence for $\sigma = 1$ within the 3DLM, a proper choice of initial conditions is required in order to simulate a chaotic solution. Without knowing this, Prof. Lorenz thought it may have been impossible to obtain a "strange'' solution if $\sigma = 1$ was first used in the Saltzman (1962) model, giving no motivation for him to work on the 3DLM. In other words, the value of $\sigma = 10$ used in the original study (e.g., Saltzman 1962) was indeed a "fortunate" choice so that an unexpected irregularly oscillatory solution could be revealed, inspiring

Prof. Lorenz to develop the 3DLM in order to discover interesting chaotic features. However, on the other hand, we now understand that such a configuration can only depict a partial picture for the nature of weather. Based on our results and analysis, a realistic system should include physical processes for (some of) the tiny disturbances in order to completely dissipate. Since it produces the coexistence of chaotic and steady-state solutions and since the steady-state solution has no long-term memory of tiny perturbations, the 3DLM with the new configuration of $\sigma = 1$ satisfies the objective. Such a system, which is similar to the 9DLM that produces two kinds of coexisting attractors, provides a more realistic view on the true nature of weather than the original 3DLM with a typical configuration. The above results support the idea that two kinds of attractor coexistence should be taken into consideration to reveal the nature of weather.

5. A Refined View on the Nature of Weather

Within the forced dissipative 3DLM, chaotic solutions appear within a finite range of parameters (e.g., heating parameter), bounded on one side by stable, steady-state solutions and on the other side by nonlinear periodic solutions. Since climate and weather involve open systems (e.g., McGuffie and Henderson-Sellers 2014), an assumption of constant parameters within numerical simulations using the 3DLM, as well as high-dimensional LMs, is not realistic (e.g., Daron and Stainforth 2015). Time varying parameters that lead to different attractors should be used in models for realistic climate or weather (Shen et al. 2020, in preparation). For example, when a moderate heating becomes weaker (or stronger), a steady-state solution (or a limit cycle) may appear. Since regular and chaotic solutions may alternatively appear, chaotic solutions alone may not be able to represent the entirety of weather.

Additionally, our results show that chaotic and non-chaotic solutions may coexist and two kinds of attractor coexistence may alternatively appear within the 9DLM using time varying parameters. The analysis suggests a need to refine our view of weather by taking the dual nature associated with attractor coexistence into consideration. To this end, we suggest, contrary to the traditional view that weather is chaotic, that weather is, in fact, a superset that consists of both chaotic and non-chaotic processes, including both order and chaos.

Vacillation, Coexisting Two LCs, and Coexisting Two Time-scale Orbits

The (potential) occurrence of a regular nonlinear periodic solution (i.e., limit cycle) in the atmosphere was first illustrated by laboratory experiments using dishpans. Based on experiments by David Fultz (Fultz et al. 1959) and Raymond Hide (Hide 1953), Lorenz (1993) suggested three types of solutions, including:

(1) steady state solutions, (2) irregular chaotic solutions, and (3) vacillation. "Amplitude vacillation" is defined as a solution whose amplitude grows and periodically decays in a regular cycle (Lorenz 1963c; Ghil and Childress 1987; Ghil et al. 2010). Studies by Pedlosky (1972), Smith (1975), and Smith and Reilly (1977) found that *amplitude vacillation can be viewed as a limit cycle solution*. By conducting a study for observational characteristics of low-frequency variability, Ghil and Robertson (2002) suggested that *40-day, intra-seasonal oscillations* may arise from a bifurcation off the blocking flow and *may be represented by a limit cycle with a period of 40 days.*

As discussed earlier, we showed that the 3DLM with a realistic value of $\sigma = 1$ also generates two kinds of attractor coexistence. Additionally, the coexistence of two stable limit cycle solutions was documented using the Lorenz 1984 model (Lorenz 1984, 1990; Masoller et al. 1992; Pielke and Zeng 1994; Veen 2002a, b; Wang et al. 2014) that also contains three types of solutions, including steady state, periodic solutions, and chaotic solutions. Using a seasonally varying forcing term with a time scale of 12 months, Lorenz (1990) showed that chaos appears during winter (within a specific range of parameters) and two coexisting LCs during summer (within a different range of parameters). Such numerical results also support the view of the dual nature of chaos and order that alternatively appear. The above results suggest that once summer begins and has been observed, a better predictability for a limit cycle solution may be expected during each cycle of the solution in summer, as compared to that in winter. More recently, Lucarini and Bodai (2019) applied a multistable system with coexisting attractors to reveal the bistability of the climate system with both positive and negative feedback (e.g., Garashchuk et al. 2019; Lucarinii and Bodai 2019).

Coexisting solutions at two time scales, that are not the same as the coexisting attractors discussed above, have also been documented within the scientific literature. Related studies additionally support the refined view on the nature of weather. For example, co-existence of fast and slow manifolds has been discussed by Lorenz (1986, 1992), Lorenz and Krishnamurthy (1987), and Curry et al. (1995). Both types of solutions in Lorenz (1986) are non-chaotic. By comparison, fast and slow "variables" that are chaotic may also coexist within coupled systems (e.g., Pena and Kalnay 2004; Mitchell and Gottwald 2012). In fact, an analysis using a singular perturbation method (Bender and Orszag 1978) indicates that the GLM also possesses the coexistence of slow and fast variables that correspond to large and very small spatial modes (e.g., Eq. (2) and Eq. (4) of Shen 2019a in a high-dimension phase space). A current trend is to include time-varying parameters to increase the complexities of low order systems (e.g., Lucarini 2019). It can be shown that a higher dimensional Lorenz model (e.g., 7DLM) can be viewed as a lower-dimensional Lorenz model (e.g., 5DLM) with a period forcing, suggesting that the complexities of spatial mode-mode interaction may lead to the temporal complexities.

Error Saturations and Computational Chaos

In real-world weather models, the appearance of (fully) chaotic solutions may be indicated by error saturations, defined as follows. A logistic equation has been used to describe the evolution of root mean square (rms) average forecast error for ensemble runs (Lorenz 1969b, 1996; Nicolis 1992; Kalnay 2003; Zhang et al. 2019). Given an initial condition with a small value, the solution of the logistic equation has time varying, non-negative growth rates (e.g., growing at an initial larger growth rate, then at a nonlinear smaller growth rate, and eventually approaching a constant defined as a saturated error that has a zero growth rate). The occurrence of error saturation at a fully nonlinear stage indicates a comparable number of members with positive and negative error growth rates at a given time. Such a result is consistent with the features of a positive LE and solution boundedness associated with a specific chaotic solution.

The error growth model with non-negative growth rates may describe the statistical behavior of the system within which the majority of small errors tends to grow. *By comparison, the error growth model cannot accurately represent the initial, transient evolution of the rms averaged forecast error associated with large ensemble members with periodic or decaying components whose growth rates are small.* For periodic solutions such as vacillation (Lorenz 1969b), an ensemble averaged error may grow (or decay) with time when a large (or small) ensemble number of growing errors and a small (or large) ensemble number of decaying errors are averaged. As a result, when oscillatory waves were simulated, their rms errors may oscillate with time rather than become saturated. For example, oscillatory rms errors appeared after 40-day simulations in Figure 5 of Liu et al. (2009) who performed global simulations using the Community Atmosphere Model (Collins et. al. 2004). An additional example can be found in 30-day simulations of multiple African Easterly Waves (AEWs) using a global mesoscale model that produced oscillatory correlation coefficients (Shen 2019b).

On the other hand, it should be noted that error saturations may appear in association with computational chaos that is a numerical artifact. For example, Lorenz (1989) presented several cases in order to show that while differential equations of a model may possess nonlinear limit cycle solutions, the corresponding discrete version of the model with large time steps produces a sensitive dependence of solutions on the initial condition, referred to as computational chaos. As a result, the appearance of error saturations (as well as positive LE) that appear within numerical models does not necessarily represent the chaotic nature of weather. *Due to the appearance of computational chaos, an estimate of a practical predictability limit using saturation errors should be interpreted with caution, as it does not necessarily represent an intrinsic predictability limit for real weather.*

5. Concluding Remarks

The statement of "weather is chaotic" has been introduced to indicate the chaotic nature of weather with a finite intrinsic predictability. The statement has also been cited to embrace a practical predictability limit of two weeks. The finite intrinsic and practical predictability are indeed largely derived from the chaotic and unstable solutions of Lorenz models (e.g., Lorenz 1963a, 1969a). In other words, the current view of "weather is chaotic" does not take into consideration other types of solutions within original Lorenz models and new types of solutions within newly developed generalized Lorenz models.

In this study, we first applied the aforementioned models in order to reveal three types of solutions and two kinds of attractor coexistence, indicating different intrinsic predictability for different solutions. We then suggested a refined view on the dual nature of chaos and order in weather. In contrast to the current view that focuses on chaotic solutions with a predictability limit (of two weeks), our refined view suggests that coexisting chaotic and non-chaotic systems can have different intrinsic predictability. *The refined view may unify the theoretical understanding of different predictability within Lorenz models with recent numerical simulations of advanced global models that can simulate large-scale tropical waves beyond two weeks (e.g., Shen 2019b; Judt 2020).*

The refined view with a duality of chaos and order is fundamentally different from the Laplacian view of deterministic predictability and the Lorenz view of deterministic chaos. The refined view that is not too optimistic nor too pessimistic suggests both potential and challenges. For non-chaotic processes with steady-state or nonlinear periodic solutions, their intrinsic predictability is deterministic (e.g., up to the lifetime of a dissipative solution or the time scale of the forcing) and their practical predictability can be continuously increased by improving the accuracy of the model and the initial conditions. For limit cycle solutions that may be associated with computational chaos, accurate simulations with better predictability, as compared to chaotic solutions, can be obtained by increasing temporal resolutions and/or removing redundant dissipations. To reveal longer predictability or better estimates on predictability in model and observation data, we will focus on developing schemes for the detection of chaotic and non-chaotic solutions (e.g., Sprott and Xiong 2015; Reyes and Shen 2019) and examining the roles of butterfly effects in multiscale simulations using high-resolution global models.

In addition to the chaotic nature of weather with a finite predictability, another major influential impact of the 3DLM is that the sensitive dependence on initial condition, referred to as the butterfly effect of the first kind, has been inaccurately metaphorized to indicate the ability of a butterfly flap in creating a tornado, referred to as the butterfly effect of the second kind (Shen 2014). To understand their roles in reality and numerical models, the two different kinds of butterfly

effects are being analyzed based on a comprehensive review of historical literature and recent understanding of chaos dynamics.

Acknowledgments:
We thank reviewers of the manuscript, the editor, and Drs. R. Anthes, B. Bailey, J. Buchmann, D. Durran, M. Ghil, B. Mapes, Z. Musielak, T. Krishnamurti (Deceased), C.-D. Lin, J. Rosenfeld, R. Rotunno, I. A. Santos, C.-L. Shie, S. Vannitsem, and F. Zhang (Deceased) for valuable comments and discussions. We appreciate the eigenvalue analysis provided by Mr. N. Ferrante. We are grateful for support from the College of Science at San Diego State

References

Bailey, B., 2011: Quantifying the predictability of noisy space-time dynamical processes. Statistics and its Interface, 4, 535-549.

Bender, C. M. and Orszag, S. A. 1978. Advanced Mathematical Methods for Scientists and Engineers. McGraw-Hill, New York, pp. 593.

Collin, W. D., and Coauthors, 2004: Description of the NCAR Community Atmosphere Model (CAM3). NCAR Tech. Note NCAR/TN-464+STR, 226pp.

Curry, J.H., S.E. Haupt, and M.N. Limber, 1995: Truncated Systems, Initialization, and the Slow Manifold, Tellus, 47A, 145161.

Daron, J. D. and D. A. Stainforth, 2015: On quantifying the climate of the nonautonomous Lorenz-63 model. Chaos 25, 043103 (2015); https://doi.org/10.1063/1.4916789.

Ding, R. O., and J. P. Li, 2007: Nonlinear finite-time Lyapunov exponent and predictability. Phys. Lett., 354A, 396–400.

Drazin, P.G., 1992: Nonlinear Systems. Cambridge. 333pp.

Eckhardt, B., and D. Yao, 1993: Local Lyapunov exponents in chaotic systems. Physica D, 65, 100–108.

Faghih-Naini, S. and B.-W. Shen, 2018: Quasi-periodic in the Five-dimensional Non-dissipative Lorenz Model: The Role of the Extended Nonlinear Feedback Loop. International Journal of Bifurcation and Chaos, Vol. 28, No. 6 (2018) 1850072 (20 pages). DOI: 10.1142/S0218127418500724.

Felicio, C. C. and Rech, P. C., 2018: On the dynamics of five- and six-dimensional Lorenz models. J. Phys. Commun. 2, 025028.

Franz, M. and M. Zhang, 1995: Suppression and creation of chaos in a periodically forced Lorenz system. Physical Review E. Vol 52, 4, 3558-3565.

Fultz, D., R. R. Long, G. V. Owens, W. Bohan, R. Kaylor, and J. Weil, 1959: "Studies of Thermal Convection in a Rotating Cylinder with Some Implications for Large-Scale Atmospheric Motion." Meteorol. Monographs (American Meteorological Society) 21, no. 4.

Garashchuk, I, R., D. I. Sinelshchikov, A. O. Kazakov, and N. A. Kudryashov, 2019: Hyperchaos and multistability in the model of two interacting microbubble contrast agents editors-pick. Chaos 29, 063131 (2019); https://doi.org/10.1063/1.5098329

Ghil, M., P. Read and L. Smith, 2010: Geophysical flows as dynamical systems: the influence of Hide's experiments. in Astronomy & Geophysics, vol. 51, no. 4, pp. 428-435, Aug. 2010. doi: 10.1111/j.1468-4004.2010.51428.x

Ghil, M. and A. W. Robertson, 2002: "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? PNAS February 19, 2002 99 (suppl 1) 2493-2500; https://doi.org/10.1073/pnas.012580899
Ghil, M. and S. Childress, 1987: Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics. Springer, New York, 485pp.
Gleick, J., 1987: Chaos: Making a New Science, Penguin, New York, 360 pp.
Grebogi, C., S. W. McDonald, E. Ott and J. A. Yorke, 1983: Final state sensitivity: An obstruction to predictability. Physics Letters A. Volume 99, Issue 9, 26 December 1983, Pages 415-418.
Guckenheimer, J., and Williams, R. F., 1979: Structural stability of Lorenz attractors. Publ. Math. IHES. 50, 59.
Guckenheimer, J., and P. Holmes, 1983: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, New York, 459pp.
Hide, R., 1953: Some Experiments on Thermal Convection in a Rotating Liquid. Quart. f. Roy. Meteorol. Soc. 79 (1953): 161.
Hilborn, R. C., 2000: Chaos and Nonlinear Dynamics. An Introduction for Scientists and Engineers. 2nd ed. Oxford University Press, New York, pp. 650.
Hirsch, M., S. Smale, and R. L. Devaney, 2013: Differential Equations, Dynamical Systems, and an Introduction to Chaos. 3rd edition. Academic Press, 432 pp.
Judt, F., 2020: Atmospheric Predictability of the Tropics, Middle Latitudes, and Polar Regions Explored through Global Storm-Resolving Simulations. Journal of The Atmospheric Sciences, 77, 257-276. https://doi.org/10.1175/JAS-D-19-0116.1
Kalnay, E., 2003: Atmospheric Modeling, Data Assimilation and Predictability. Cambridge, New York. 369 pp.
Lewis, 2005: Roots of ensemble forecasting. Monthly weather review, 133, 7, 1865-1885.
Liu, H.-L., F. Sassi, and R. R. Garcia, 2009: Error Growth in a Whole Atmosphere Climate Model. J. Atmos. Sci., 66, 173-186.
Lorenz, E., 1963a: Deterministic nonperiodic flow, J. Atmos. Sci., 20, 130-141.
Lorenz, E., 1963b: The predictability of hydrodynamic flow. Trans. N.Y. Acad. Sci., Ser. II, 25, No. 4, 409-432.
Lorenz, 1963c: The mechanics of vacillation. J. Atmos. Sci., 20, 448-464.
Lorenz, E. N., 1969a: The predictability of a flow which possesses many scales of motion. Tellus, 21, 289–307.
Lorenz, 1969b: Atmospheric predictability as revealed by naturally occurring analogues. J. Atmos. Sci., 26, 636-646.
Lorenz, E. N., 1972: Predictability: Does the flap of a butterfly's wings in Brazil set off a tornado in Texas? Proc. 139th Meeting of AAAS Section on Environmental Sciences, New Approaches to Global Weather: GARP, Cambridge, MA, AAAS, 5 pp. [Available online at http://eaps4.mit.edu/research/Lorenz/Butterfly_1972.pdf.]
Lorenz, E.N., 1984: Irregularity: a fundamental property of the atmosphere. Crafoord Prize Lecture, presented at the Royal Swedish Academy of Sciences, Stockholm, September 28, 1983. Tellus, 36A, 98-110. (LM84)
Lorenz, E.N., 1986: On the existence of a slow manifold, J. Atmos. Sci. 43, 154471557.
Lorenz, E.N., 1989: Computational chaos: a prelude to computational instability. Physica, 35D, 299-317.
Lorenz, E.N., 1990: Can chaos and intransitivity lead to interannual variability? Tellus, 42A, 378-389.
Lorenz, E.N., 1992: The slow manifold. What is it? J. Atmos. Sci., 49, 24492451.
Lorenz, E. N., 1993: The Essence of Chaos. University of Washington Press, Seattle, 227pp.

Lorenz, E. N., 1996: "Predictability – A problem partly solved" (PDF). Seminar on Predictability, Vol. I, ECMWF. (also, published as Lorenz (2006))
Lorenz, E. N., 2005: Designing Chaotic Models. J. Atmos. Sci. 62, 1574-1587.
Lorenz, E. N., 2006: "Predictability – A problem partly solved" . In T. Palmer & R. Hagedorn (Eds.), Predictability of Weather and Climate (pp. 40-58). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511617652.004
Lorenz, E. N., 2008: The butterfly effect. Premio Felice Pietro Chisesi e Caterina Tomassoni award lecture, University of Rome, Rome, April, 2008.
Lorenz, E.N. and V. Krishnamurthy, 1987: On the nonexistence of a slow manifold, J. Atmos. Sci., 44, 29402950.
Lucarini, V. 2019 : Stochastic resonance for nonequilibrium systems. Physical Review E. 100, 062124. DOI: 10.1103/PhysRevE.100.062124.
Lucarini, V. and T. Bodai, 2019 : Transitions across Melancholia States in a Climate Model: Reconciling the Deterministic and Stochastic Points of View. Physical Review Letter 122, 158701 (2019).
Macek, W. M., 2018: Nonlinear Dynamics and Complexity in the Generalized Lorenz System. Nonlinear Dynamics. 94, 2957-2968. https://doi.org/10.1007/s11071-018-4536-z.
Masoller, C., A.C. Sicardi Schifino, L. Romanelli, 1992: Regular and chaotic behavior in the new Lorenz system. Physics Letters A 167 (1992) 185-190.
McGuffie, K. and A. Henderson-Sellers, 2014: The Climate Modeling Primer. 4th Ed. John Wiley & Sons, Ltd. 439 pp.
Mitchell L. and G. A. Gottwald, 2012: Data Assimilation in Slow–Fast Systems Using Homogenized Climate Models. J. Atmos. Sci., 69, 1359-1377. DOI: 10.1175/JAS-D-11-0145.1
Moon, S., B.-S. Han, J. Park, J. M. Seo and J.-J. Baik, 2017: Periodicity and Chaos of High-Order Lorenz Systems International Journal of Bifurcation and Chaos, Vol. 27, No. 11 (2017) 1750176 (11 pages) DOI: 10.1142/S0218127417501760.
Moon, S., J. M. Seo1, B.-S. Ha, J. Park, and J.-J. Baik, 2019: A physically extended Lorenz system. Chaos 29, 063129 (2019); https://doi.org/10.1063/1.5095466
Musielak, Z. E., Musielak, D. E. and Kennamer, K. S., 2005: The onset of chaos in nonlinear dynamical systems determined with a new fractal technique, Fractals, 13, 19-31.
Nese, J. M., 1989: Quantifying Local Predictability in Phase Space. Physica D., 35, 237-250.
Nicolis, C., 1992: Probabilistic aspects of error growth in atmospheric dynamics. Q. J. R. Meteorol. SOC. (1992), 118, pp. 553-568.
Ott, E., 2002: Chaos in Dynamical Systems. 2nd Edition. Cambridge University Press. 478pp.
Park, J., H. Lee, Y.-L. Jeon, and J.-J. Baik, 2015: Periodicity of the Lorenz-Stenflo equations. Physica Scripta, 90, 065201.
Park, J., B.-S. Han, H. Lee, Y.-L. Jeon, and J.-J. Baik, 2016: Stability and periodicity of high-order Lorenz-Stenflo equations. Physica Scripta, 91, 065202.
Pedlosky, 1972: Limit cycles and unstable baroclinic waves. J. Atmos. Sci., 29, 53-63.
Pedlosky, J. and C. Frenzen, 1980: Chaotic and Periodic Behavior of Finite-Amplitude Baroclinic Waves. J. Atmos. Sci. 37, 1177-1196.
Peña, M. and Kalnay, E., 2004: Separating fast and slow modes in coupled chaotic systems, Nonlin. Processes Geophys., 11, 319-327, https://doi.org/10.5194/npg-11-319-2004, 2004.

Pielke, R., 2008: The real butterfly effect. [Available online at http://pielkeclimatesci.wordpress.com/2008/04/29/the-realbutterfly-effect/.]
Pielke, R.A. and X. Zeng, 1994: Long-term variability of climate. J. Atmos. Sci., 51, 155-159.
Reyes, T. and B.-W. Shen, 2019: A Recurrence Analysis of Chaotic and Non-Chaotic Solutions within a Generalized Nine-Dimensional Lorenz Model. Chaos, Solitons & Fractals. 125 (2019), 1-12. https://doi.org/10.1016/j.chaos.2019.05.003
Roy, D. and Z. E. Musielak, 2007: Generalized Lorenz models and their routes to chaos. I. energy-conserving vertical mode truncations, Chaos Soliton. Fract., 32, 1038-1052.
Saiki, Y., E. Sander, and J. A. Yorke, 2017: Generalized Lorenz equations on a three-sphere. The European Physical Journal Special Topics, 226, 1751-1764.
Saltzman, B., 1962: Finite amplitude free convection as an initial value problem, J. Atmos. Sci., 19, 329-341.
Shen, B.-W., 2019a: Aggregated Negative Feedback in a Generalized Lorenz Model. International Journal of Bifurcation and Chaos, Vol. 29, No. 3 (2019) 1950037 (20 pages). https://doi.org/10.1142/S0218127419500378
Shen, B.-W., 2019b: On the Predictability of 30-Day Global Mesoscale Simulations of African Easterly Waves during Summer 2006: A View with the Generalized Lorenz Model. *Geosciences* 2019, 9, 281. https://doi.org/10.3390/geosciences9070281
Shen, B.-W., 2018: On periodic solutions in the non-dissipative Lorenz model: the role of the nonlinear feedback loop. Tellus A: 2018, 70, 1471912, https://doi.org/10.1080/16000870.2018.1471912.
Shen, B.-W., 2017: On an extension of the nonlinear feedback loop in a nine-dimensional Lorenz model. Chaotic Modeling and Simulation (CMSIM), 2: 147–157, 2017.
Shen, B.-W., 2016: Hierarchical scale dependence associated with the extension of the nonlinear feedback loop in a seven-dimensional Lorenz model. Nonlin. Processes Geophys., 23, 189-203, doi:10.5194/npg-23-189-2016, 2016.
Shen, B.-W., 2015: Nonlinear Feedback in a Six-dimensional Lorenz Model. Impact of an additional heating term. Nonlin. Processes Geophys., 22, 749-764, doi:10.5194/npg-22-749-2015, 2015.
Shen, B.-W., 2014: Nonlinear Feedback in a Five-dimensional Lorenz Model. J. of Atmos. Sci., 71, 1701–1723. doi:http://dx.doi.org/10.1175/JAS-D-13-0223.1
Shen, B.-W., T. A. L Reyes and S. Faghih-Naini, 2019: Coexistence of Chaotic and Non-Chaotic Orbits in a New Nine-Dimensional Lorenz Model. In: Skiadas C., Lubashevsky I. (eds) 11th Chaotic Modeling and Simulation International Conference. CHAOS 2018. Springer Proceedings in Complexity. Springer, Cham. https://doi.org/10.1007/978-3-030-15297-0_22
Shimizu, T., 1979: Analytical Form of the Simplest Limit Cycle in the Lorenz Model: Physica, 97A, 383-398.
Smale, S., 1998: Mathematical Problems for the Next Century. The Mathematical Intelligencer 20, no. 2, pages 7–15.
Smith, R.K., 1975: A note on a theory of vacillating baroclinic waves. J. Atmos. Sci., 32, 2027.
Smith, R. K. and J. M. Reilly, 1977: On a theory of amplitude vacillation in baroclinic waves: Some numerical solutions. J. Atmos. Sci., 34, 1256-1260.
Sparrow, C., 1982: The Lorenz Equations: Bifurcations, Chaos, and Strange Attractors. Springer, New York. Appl. Math. Sci., 269 pp.
Sprott, J. C. and A. Xiong, 2015: Classifying and Quantifying Basins of Attraction. Chaos 25, 083101 (2015). http://dx.doi.org/10.1063/1.4927643

Sprott, J. C., X. Wang and G. Chen, 2013: Coexistence of Point, periodic and Strange attractors. International Journal of Bifurcation and Chaos, Vol. 23, No. 5. DOI:10.1142/S0218127413500934.
Sprott, J.C., J.A. Vano, J.C. Wildenberg, M.B. Anderson, J.K. Noel, 2005: Coexistence and Chaos in Complex Ecologies. Physics Letters A 335 (2005) 207–212.
Strogatz, S. H., 2015: Nonlinear Dynamics and Chaos. With Applications to Physics, Biology, Chemistry, and Engineering. Westpress view, Boulder, CO, 513 pp.
Thompson, J. M. T. and H. B. Stewart, 2002: Nonlinear Dynamics and Chaos. 2nd edition. John Wiley & Sons, LTD, 437 pp.
Tucker, W., 2002: A Rigorous ODE Solver and Smale's 14th Problem. Found. Comput. Math. 2, 53-117.
Veen, L. van, 2002a: Time scale interaction in low-order climate models. Utrecht University Repository. (PhD Dissertation).
Veen, L., 2002b: Baroclinic flow and the Lorenz-84 mode. Int. J. Bifurcation Chaos, 13, 2117 (2003). DOI: http://dx.doi.org/10.1142/S0218127403007904
Wang, H., Y. Yu, and G. Wen, 2014: Dynamical Analysis of the Lorenz-84 Atmospheric Circulation Model. Journal of Applied Mathematics Volume 2014, Article ID 296279, 15 pages. http://dx.doi.org/10.1155/2014/296279.
Wolf, A., J. B. Swift, H. L. Swinney, and J. A. Vastano, 1985: Determining Lyapunov Exponents from a Time Series. Physica, 16D, 285-317.
Xavier, J. C., and P. C. Rech, 2010: Regular and chaotic dynamics of the Lorenz-Stenflo system. International Journal of Bifurcation and Chaos, 20 145-152.
Yang, Q. and G. Chen, 2008: A chaotic system with one saddle and two stable node-foci. International Journal of Bifurcation and Chaos, Vol. 18, No. 5, 05.2008, p. 1393-1414.
Yorke, J. and E. Yorke, 1979: Metastable chaos: The transition to sustained chaotic behavior in the Lorenz model. J. Stat. Phys. 21, 263–277.
Zeng, X., R. Eykholt, and R.A. Pielke, 1991: Estimating the Lyapunov-exponent spectrum from short time series of low precision. Phys. Rev. Lett., 66, 3229-3232.
Zeng, X., R. A. Pielke Sr., and R. Eykholt, 1993: Chaos Theory and Its Applications to the Atmosphere. Bulletin of the Atmospheric Meteorological Society. Vol. 74, No. 4, 631-644.
Zhang, F., Y. Q. Sun, L. Magnusson, R. Buizza, S.-J. Lin, J.-H. Chen, K. Emanuel, 2019: What Is the Predictability Limit of Midlatitude Weather? J. of Atmos. 76, 1077-1091.

Table 1: Initial conditions (ICs) for revealing the coexistence of two attractors for $\sigma = 1$, $b = 0.4$, and $r = 170$ within the 3DLM. $X_c = Y_c = \sqrt{b(r-1)}$ and $Z_c = (r-1)$. The six rows provide the ICs for Fig. A1.

X	Y	Z
X_c	$Y_c + 1$	Z_c
-X_c	$-Y_c + 1$	Z_c
0	1	0
-76.72346293	37.62433028	-146.96230812
-27.75526885	167.67883615	3.66782724
136.44623635	99.45689394	-19.76741851

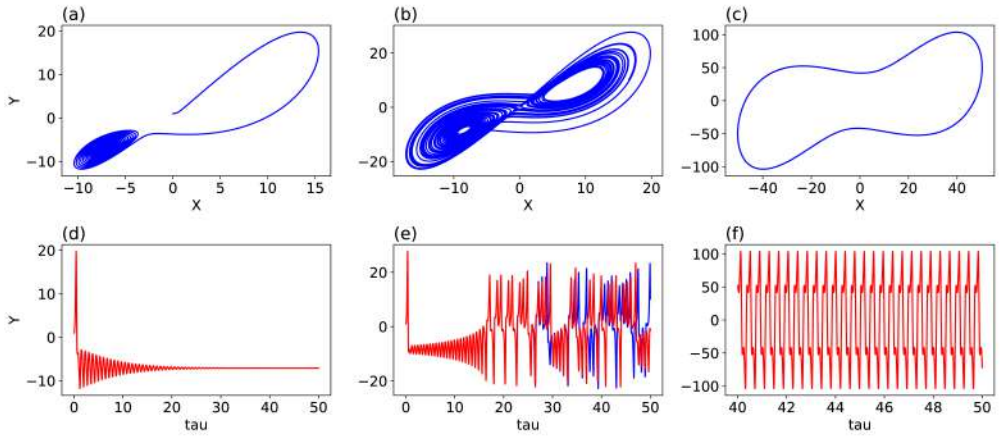


Figure 1: Three types of solutions within the 3DLM. Left, middle, and right panels display steady-state, chaotic, and limit cycle solutions at small, moderate, and large heating parameters (i.e., $r = 20, 28$, and 350), respectively. The solutions are categorized into a point attractor, a chaotic attractor, and a periodic attractor, respectively. Top panels show orbits within the $X - Y$ space and bottom panels depict the time evolution of Y. Blue lines provide solutions from control runs. To display results from parallel runs, red lines are added in the bottom panels. Sensitive dependence on initial conditions is shown in panel (e) with two visible lines. Two panels, (b) and (e), are reproduced from Shen (2019b).

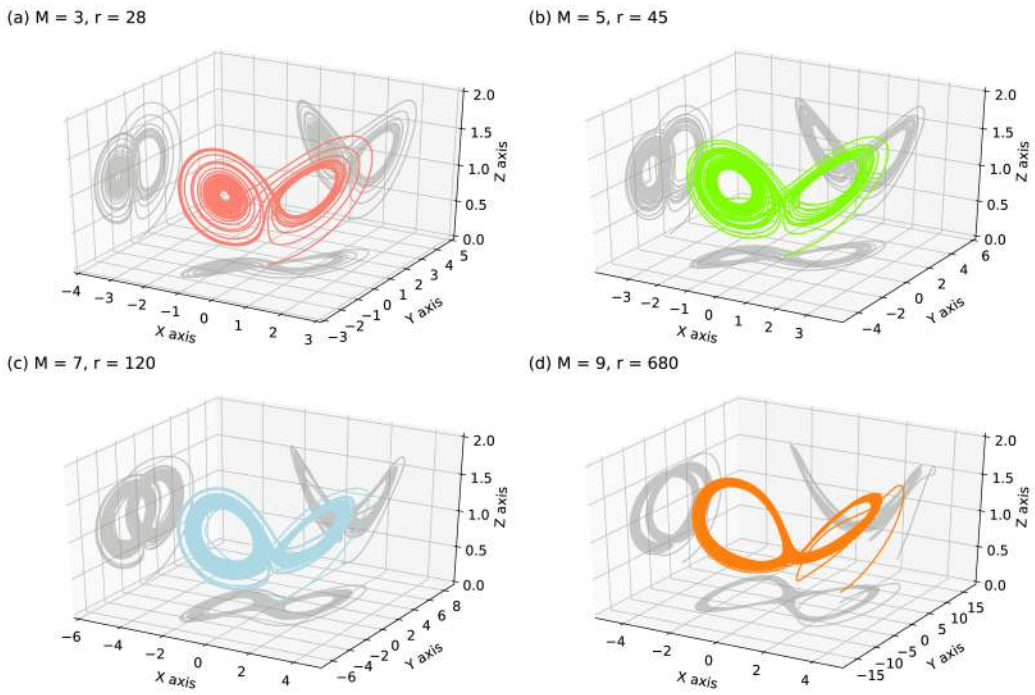


Figure 2: Chaotic solutions in the $X - Y - Z$ phase space within the 3D, 5D, 7D, and 9D Lorenz models (LMs). Panels (a)-(c) use the same initial conditions with Y = 1 and the remaining as zero, while panel (d) uses an IC with 100 for all variables. Variables (X, Y, Z) are normalized by $2\sqrt{r-1}$, $2\sqrt{r-1}$, and $(r-1)$, respectively. A larger heating parameter is required for the onset of chaos in a higher-dimensional LM. Also see detailed analysis of solutions in Shen (2016) and Shen (2019a).

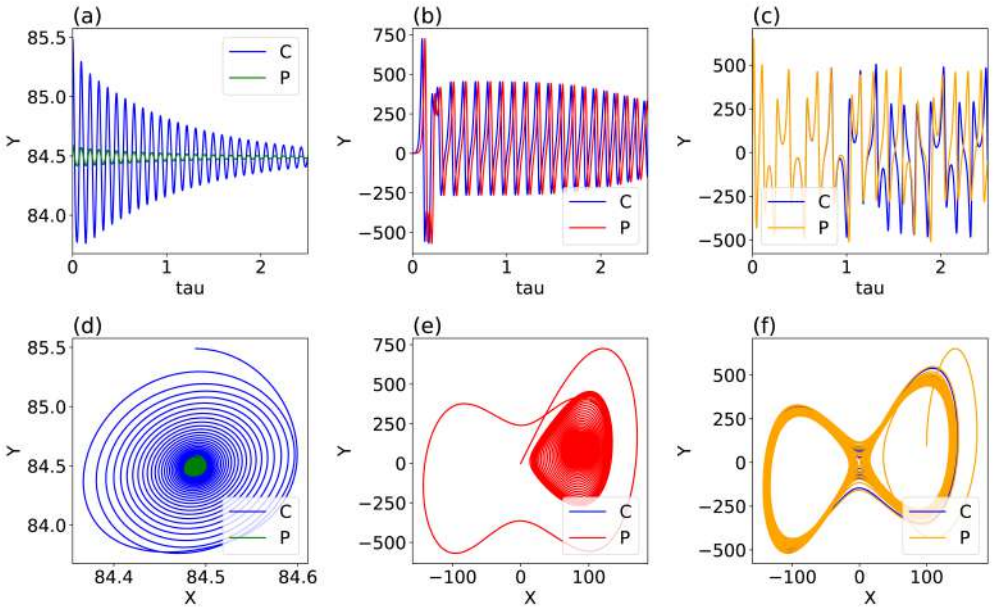


Figure 3: Solutions of the GLM with $M = 9$ and $r = 680$. Initial conditions for the three cases are placed near the non-trivial critical point (a,d), the origin (i.e., trivial critical point) (b,e), and at $(100, 100, 100, 100, 100, 100, 100, 100, 100)$ (c,f). Top panels show the time evolution of Y for $t \in [0, 2.5]$, while bottom panels display the corresponding solutions $t \in [0, 10]$ within the $X - Y$ space. Control and parallel runs are denoted by 'C' and 'P', respectively. A finite-amplitude perturbation ($\epsilon = -0.9$) is added into the parallel runs. Two panels, (c) and (f), are reproduced from Shen (2019a).

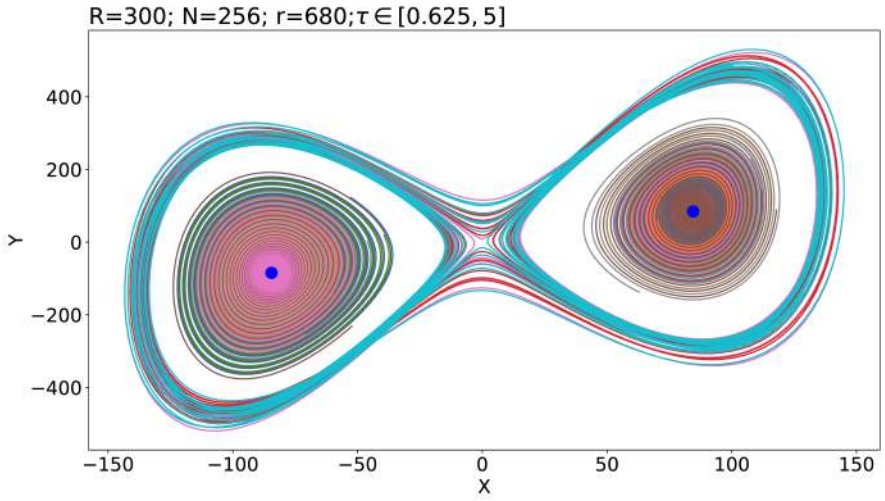


Figure 4: Coexistence of chaotic and non-chaotic orbits starting with 256 different initial conditions(ICs) for $\tau \in [0.625, 5]$. Chaotic orbits recurrently return back to the saddle point at the origin. Non-chaotic orbits eventually approach one of two stable critical points as shown in large blue dots. Chaotic and non-chaotic orbits occupy different regions of attraction within the phase space.

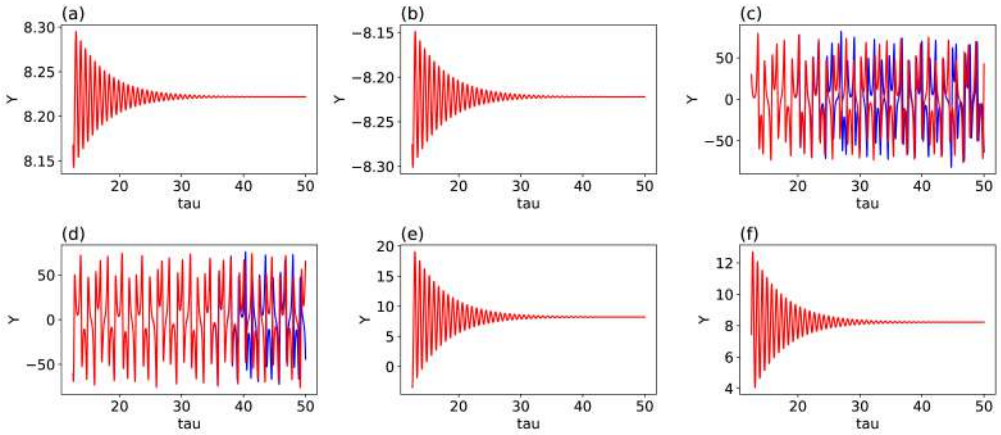


Figure 5: A co-existence of chaotic (c, d) and non-chaotic (a, b, e, f) solutions using the same parameters for $\sigma = 1$, $b = 0.4$, and $r = 170$ within the 3DLM. Blue and red lines display solutions from the control and parallel runs, respectively. Initial conditions for the results in six panels are listed in Table 1.

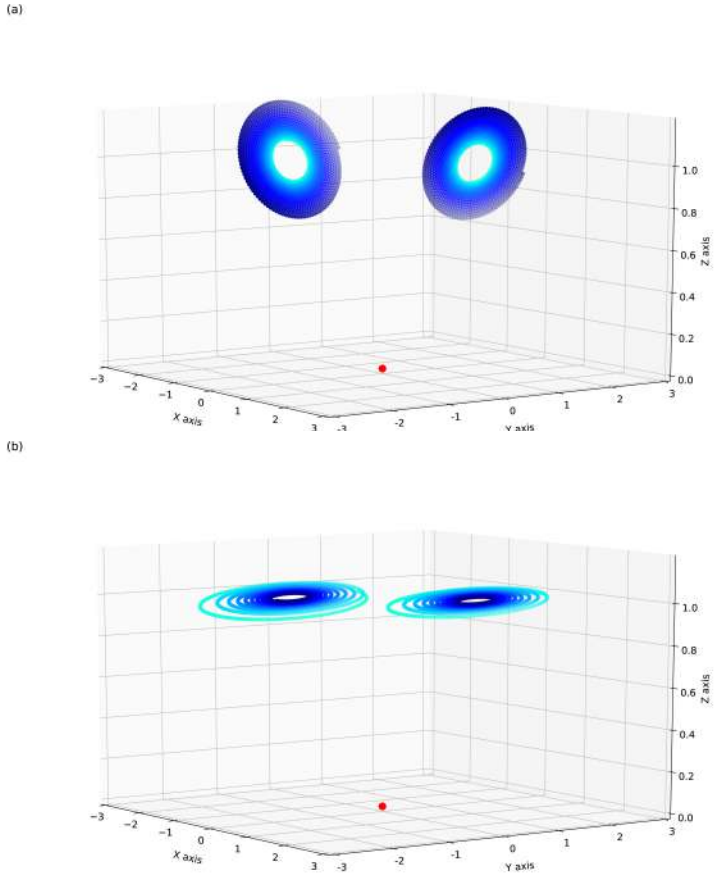


Figure 6: Local behavior near the two non-trivial critical points for the 3DLM with $\sigma = 10$ (a) and $\sigma = 1$ (b). Lighter blue dots indicate the locations of orbits at earlier times. A red dot indicates the origin, which is a saddle point. Orbits in panel (a) spiral away from the non-trivial critical points while orbits in panel (b) spiral toward the non-trivial critical points.