

# Modeling of Chaotic Processes by Means of Antisymmetric Neural ODEs

Vasiliy Ye. Belozyorov

*Department of Applied Mathematics, Dnipro National University, Gagarin's avenue, 72, 49010, Dnipro, Ukraine, belozyve2017@gmail.com*

Danylo V. Dantsev

*Chaotic Dynamics Incorporated, 201-45793, Luckakuck Way, Chilliwack BC V2R 5S3, Canada, danylo@chaodyna.com*

Correspondence should be addressed to Vasiliy Belozyorov; belozyve2017@gmail.com

**Abstract.** The main goal of this work is to construct an algorithm for modeling chaotic processes using special neural ODEs with antisymmetric matrices (antisymmetric neural ODEs) and power activation functions (PAFs). The central part of this algorithm is to design a neural ODEs architecture that would guarantee the generation of a stable limit cycle for a known time series. Then, one neuron is added to each equation of the created system until the approximating properties of this system satisfy the well-known Kolmogorov theorem on the approximation of a continuous function of many variables. In addition, as a result of such an addition of neurons, the cascade of bifurcations that allows generating a chaotic attractor from stable limit cycles is launched. We also consider the possibility of generating a homoclinic orbit whose bifurcations lead to the appearance of a chaotic attractor of another type. In conclusion, the conditions under which the found attractor adequately simulates the chaotic process are discussed. Examples are given.

**Key words:** system of ordinary autonomous differential equations, neural network, antisymmetric matrix, power activation function, Lyapunov stability, limit cycle, homoclinic orbit, strange attractor, search algorithm

## 1 Introduction

Today recurrent neural networks (RNN) are one of the main tools for modeling various dynamic processes. It should be noted that the quality of modeling with the help of RNN depends on the activation functions used in the network (see [1] – [7]).

As for activation functions, the good results of modeling various processes were obtained precisely with the help of those neural networks in which the well-known rectified linear units (ReLU) were used [4, 6, 8]. Naturally, any generalizations of ReLU deserve attention. Therefore, below we will consider some of them.

We will not focus on the advantages or disadvantages of one or another activation function, but will focus only on those properties that are essentially used in this work.

In recent years, an interesting idea has appeared to interpret a system of ordinary differential equations in the form of a suitable neural network (residual network) [9] – [12]. Precisely this interpretation is implemented in the present work: a system of differential equations (a system of so-called neural ODEs) is considered as a continuous analogue of some RNN [13] – [17]. It should be noted that in [17] the neural network was considered as a linear control system closed by nonlinear feedback. In this case, the ReLU activation functions played the role of the functions constituting the feedback. The task of modeling was not to bring the trajectories of the model and the real process closer together, but to find the algebraic invariants that determine the behavior of the model built for a known time series. If the corresponding invariants for different lengths of this time series turn out to be equal, then we can talk about the adequacy of the model and the real process.

Below we will focus on two areas of research, which can be formulated in the following questions.

1) If a neural network models a certain dynamic process, then how to guarantee the stability or boundedness of solutions of the system of differential equations describing a continuous analog of the aforementioned neural network?

2) In the theory of bifurcations, the following result is well known: in any determinate system, chaotic processes arise as a result of bifurcations of limit cycles or homoclinic orbits [18] – [21]. Therefore, how to design the architecture of neural ODEs system so that the

resulting architecture would generate a limit cycle? (It is now known that most types of chaos in systems of differential equations begin with bifurcations of limit cycles [18] – [20].)

The answer to the first question will be successful if the activation functions are chosen so that Lyapunov analysis can be done for the resulting system of neural ODEs [22] – [24]. Piecewise continuous functions, each part of which is a power function, can be proposed as such functions. The use of power activation functions (PAF) in neural networks is a generalization of the rectified linear units. In the present time ReLU are standard functions to increase the depth of learning. Therefore, power activation functions are an obvious generalization of ReLU.

Note that for systems of neural ODEs with PAF, the answer to the first question has already been partially obtained in article [25]. (In the present paper, the results proved in [25] will be generalized.) As for the second question, the main part of the article will be devoted to the answer to this question.

It should be said that in a large part of the article neural ODEs with chaotic modes are discussed. Chaos constitutes the basic form of collective neural activity for all processes and functions of perception. It acts as a controlled noise source to ensure uninterrupted access to previous memorized images and memorizing new ones. Chaos allows the system to be always active, ridding it of the need to wake up or enter a stable state every time the input changes [26]. Many researchers agree that the best from the point of view of storing and processing information is the regime of ordered chaos [27]. On the one hand, this mode has all the advantages of chaos, on the other hand, this mode can be controlled. The set of states through which the trajectory of a chaotic system passes is called a chaotic attractor. Therefore, the conditions for the existence of chaotic attractors in systems of neural ODEs are the subject of research in this paper.

The final sections of the article are devoted to the reconstruction of neural ODE systems. For this purpose, several algorithms have been developed for determining the parameters of ODE systems for known time series. The essence of these algorithms lies in the fact that they use the special structure of neural ODEs (antisymmetric neural ODEs) with which it is possible to generate a limit cycle. After that, by choosing the weight coefficients, we obtain such bifurcations of the indicated cycle, which lead to the simulation of a real chaotic process.

## 2 Mathematical preliminaries

We now recall several well-known results from the theory of approximation of real functions of  $n$  variables [28] – [30].

Let  $\mathbb{X}$  be an arbitrary set in the linear space  $\mathbb{R}^n$ . By  $\mathbf{C}(\mathbb{X})$  denote a set of continuous real functions of  $n$  variables with domain of definition  $\mathbb{X}$ .

**Definition 1.** A set of real functions  $\mathbb{F} \subset \mathbf{C}(\mathbb{X})$  is called separating points of the set  $\mathbb{X} \subset \mathbb{R}^n$  if for any different  $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{X}$ , there exists a function  $f \in \mathbb{F}$  such that  $f(\mathbf{x}_1) \neq f(\mathbf{x}_2)$ .

**Definition 2.** A collection of functions  $\mathbb{F} \subset \mathbf{C}(\mathbb{X})$  is called closed with respect to a function of one variable  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  if  $\phi(f) \in \mathbb{F}$  for any  $f \in \mathbb{F}$ .

**Theorem 1** ([28]). Let  $\mathbb{X} \subset \mathbb{R}^n$  be a compact space and  $\mathbf{C}(\mathbb{X})$  be the algebra of continuous real functions on  $\mathbb{X}$ . Let also the set  $\mathbb{F} \subset \mathbf{C}(\mathbb{X})$  containing the constant 1 be the linear subspace closed with respect to the nonlinear continuous function  $\phi$  and separating points of the set  $\mathbb{X}$ . Then  $\mathbb{F}$  is dense in  $\mathbf{C}(\mathbb{X})$ .

Theorem 1 can be interpreted as a statement about the universal approximation possibilities of arbitrary nonlinearity: using linear operations and a single nonlinear element  $\phi$ , one can construct an algorithm that builds an analytical model of any continuous function with any desired accuracy.

From an applied point of view, Theorem 1 can be presented as follows.

Let  $(\mathbf{u}, \mathbf{v}) \equiv ((u_1, \dots, u_n), (v_1, \dots, v_n))$  be a scalar product of vectors  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ .

Let  $F(x_1, \dots, x_n)$  be a real continuous function defined on a closed bounded domain  $\mathbb{D} \subset \mathbb{R}^n$ . Let also  $\epsilon > 0$  be any arbitrarily small number, which means the accuracy of the approximation.

**Theorem 2** ([28] – [30]). Let  $\psi$  be a continuous nonlinear real function of one variable. Then there exist an integer  $m > 0$ , a set of vectors  $\mathbf{a}_j \in \mathbb{R}^n$ , and sets of real numbers  $\xi_j$  and  $b_j$ ;  $j = 1, \dots, m$ , such that the function

$$H(\mathbf{x}) \equiv H(x_1, \dots, x_n) = \sum_{j=1}^m \xi_j \psi(\mathbf{a}_j, \mathbf{x}) + b_j \quad (1)$$

approximates the given function  $F(x_1, \dots, x_n)$  with the error  $\epsilon$  in the entire domain  $\mathbb{D}$ .

Thus, for any vector  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{D}$ , we have  $|F(x_1, \dots, x_n) - H(x_1, \dots, x_n)| < \epsilon$ .

In terms of neural networks, this theorem can be formulated as follows. Any continuous function of several variables can be realized with any accuracy using a two-layer neural network with a sufficient number of neurons and one nonlinear activation function in the hidden layer.

### 3 Periodic solutions of neural ODEs

We assume that we know the dimension  $n$  of the real phase space in which the considered dynamic process  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))^T \in \mathbb{R}^n$  takes place [17, 25, 26]. We will also assume that the functions  $x_1(t), \dots, x_n(t)$  are continuous and differentiable with respect to time  $t$  on the interval  $[0, \infty)$ . Our goal will be to model this process with a suitable system of neural ODEs. This system will be based on new concepts, which are demonstrated below.

Introduce the following power functions [25]:

$$g(u, \alpha \vee \beta) = \begin{cases} -(-u)^\beta & \text{if } (u < 0 \text{ and } \beta > 0); 0 \text{ if } (u < 0 \text{ and } \beta = 0) \\ u^\alpha & \text{if } (u \geq 0 \text{ and } \alpha > 0); 0 \text{ if } (u \geq 0 \text{ and } \alpha = 0) \end{cases} \quad (2)$$

and

$$g(u, \alpha \vee \beta) = \begin{cases} (-u)^\beta & \text{if } (u < 0 \text{ and } \beta > 0); 0 \text{ if } (u < 0 \text{ and } \beta = 0) \\ u^\alpha & \text{if } (u \geq 0 \text{ and } \alpha > 0); 0 \text{ if } (u \geq 0 \text{ and } \alpha = 0). \end{cases} \quad (3)$$

**Definition 3**[25]. *Representation (2) ((3)) is called an odd (even) activation function.*

It is obvious that the linear combination

$$\mathcal{H}(u, \mathbf{s}) = s_1 g(u, \alpha_1 \vee \beta_1) + \dots + s_k g(u, \alpha_k \vee \beta_k), \quad (4)$$

where  $\mathbf{s} = (s_1, \dots, s_k)$  and  $s_1 \geq 0, \dots, s_k \geq 0$ , of the odd (even) functions of form (2) (form (3)) again is the odd (even) function (see [25]).

Using these concepts, we can refine the form of function  $\psi(u)$  in Theorem 2.

1) If the activation function  $\psi(u) := \psi(u, \alpha \vee \beta)$  is odd, then in representation (1) we can consider  $\xi_j = 1; j = 1, \dots, m$ .

Indeed, let  $\xi_j < 0$ . Then

$$\xi_j \psi(u, \alpha \vee \beta) = \left\{ \begin{array}{l} ((-\xi_j)^{1/\beta} (-u))^\beta \text{ if } (u < 0) \\ -((-\xi_j)^{1/\alpha} u)^\alpha \text{ if } (u \geq 0) \end{array} \right\} = \psi(-(-\xi_j)^{(1/\alpha) \vee (1/\beta)} u, \alpha \vee \beta).$$

Similarly, let  $\xi_j > 0$ . Then

$$\xi_j \psi(u, \alpha \vee \beta) = \left\{ \begin{array}{l} -(\xi_j^{1/\beta} u)^\beta \text{ if } (u < 0) \\ (\xi_j^{1/\alpha} u)^\alpha \text{ if } (u \geq 0) \end{array} \right\} = \psi(\xi_j^{(1/\alpha) \vee (1/\beta)} u, \alpha \vee \beta).$$

Thus, we have

$$\xi_j \psi(\mathbf{a}_j, \mathbf{x}) + b_j = \left\{ \begin{array}{l} \psi[\xi_j^{(1/\alpha) \vee (1/\beta)}(\mathbf{a}_j, \mathbf{x}) + \xi_j^{(1/\alpha) \vee (1/\beta)} b_j] \text{ if } (\xi_j > 0) \\ \psi[-(-\xi_j)^{(1/\alpha) \vee (1/\beta)}(\mathbf{a}_j, \mathbf{x}) - (-\xi_j)^{(1/\alpha) \vee (1/\beta)} b_j] \text{ if } (\xi_j < 0) \end{array} \right\}.$$

If we now introduce redesignations

$$\psi[\xi_j^{(1/\alpha) \vee (1/\beta)}(\mathbf{a}_j, \mathbf{x}) + \xi_j^{(1/\alpha) \vee (1/\beta)} b_j] \equiv \psi[(\bar{\mathbf{a}}_j, \mathbf{x}) + \bar{b}_j],$$

then we get formula (1) in which  $\xi_j = 1; j = 1, \dots, m$ .

2) Let us turn to Definition 1 in which we will consider  $\mathbb{X} = \mathbb{R}$  and  $\mathbb{F}$  the set of all odd functions. Since for any  $f \in \mathbb{F}$  and any  $a \in \mathbb{X}$  equation  $f(u) = a$  has a single root, then it is clear that the set  $\mathbb{F}$  separates the points of the set  $\mathbb{X}$ .



$$\frac{h_n(x_n, \gamma_n)}{\gamma_n + 1}, (x_1, \dots, x_n) \Big) + 0.5 \left( (x_1, \dots, x_n), \left( \frac{h_1(x_1, \gamma_1)}{\gamma_1 + 1}, \dots, \frac{h_n(x_n, \gamma_n)}{\gamma_n + 1} \right) \right), \quad (8)$$

where  $\forall i \in \{1, \dots, n\}$

$$\gamma_i = \begin{cases} \beta_i, & \text{if } x_i < 0, \\ \alpha_i, & \text{if } x_i \geq 0. \end{cases}$$

(a1) The case of strictly odd function  $h_i(x_i, \alpha_i \vee \beta_i)$ ;  $i = 1, \dots, n$  [25].

Since fractions  $\beta_i + 1$  and  $\alpha_i + 1$  have an even numerator and odd denominator, then the function  $V(x_1, \dots, x_n)$  will be positive definite.

Further, from the definition of function  $V(x_1, \dots, x_n)$  and system (7) it follows that

$$\begin{aligned} \dot{V}_i(x_1(t), \dots, x_n(t)) &= \left( \frac{h_1(x_1, \gamma_1 + 1)}{\gamma_1 + 1} + \dots + \frac{h_n(x_n, \gamma_n + 1)}{\gamma_n + 1} \right)'_t \\ &= h_1(x_1(t), \alpha_1 \vee \beta_1) \cdot \dot{x}_1(t) + \dots + h_n(x_n(t), \alpha_n \vee \beta_n) \cdot \dot{x}_n(t) \\ &= \left( h_1(x_1(t), \alpha_1 \vee \beta_1), h_2(x_2(t), \alpha_2 \vee \beta_2), \dots, h_n(x_n(t), \alpha_n \vee \beta_n) \right) S \begin{pmatrix} h_1(x_1(t), \alpha_1 \vee \beta_1) \\ h_2(x_2(t), \alpha_2 \vee \beta_2) \\ \vdots \\ h_n(x_n(t), \alpha_n \vee \beta_n) \end{pmatrix} \\ &\quad + (\mathbf{h}(\mathbf{x}), \mathbf{f}(\mathbf{x})), \end{aligned} \quad (9)$$

where  $S := 0.5(A + A^T - 2rI)$ .

Introduce the norm of matrix  $Q = \{q_{ij}\} \in \mathbb{R}^n$  by the following formula:

$$\|Q\| = \sum_{1 \leq i, j \leq n} |q_{ij}|.$$

Similarly, we define the norm of vector  $\mathbf{u} = (u_1, \dots, u_n)^T$ :

$$\|\mathbf{u}\| = \sum_{1 \leq i \leq n} |u_i|.$$

Now we estimate the derivative  $\dot{V}_i(x_1(t), \dots, x_n(t))$  of function  $V(x_1(t), \dots, x_n(t))$ , taking into account the fact that matrix  $S$  is negative definite:

$$\begin{aligned} \dot{V}_i(\mathbf{x}(t)) &\leq \lambda_{\max}(S) \cdot (h_1^2(x_1(t), \alpha_1 \vee \beta_1) + h_2^2(x_2(t), \alpha_2 \vee \beta_2) + \dots + h_n^2(x_n(t), \alpha_n \vee \beta_n)) \\ &\quad + W(x_1(t), \dots, x_n(t))), \end{aligned}$$

where  $\lambda_{\max}(S)$  denotes the maximal eigenvalue of symmetric matrix  $S$  and  $W(x_1(t), \dots, x_n(t)) := \|\mathbf{h}(\mathbf{x})\| \cdot \|\mathbf{f}(\mathbf{x})\|$  is a positive definite function.

The last inequality can be rewritten as follows:

$$\frac{d}{dt} V(x_1(t), \dots, x_n(t)) \leq W(x_1(t), \dots, x_n(t)) - b \cdot (h_1(x_1(t), 2\gamma_1) + \dots + h_n(x_n(t), 2\gamma_n)), \quad (10)$$

where  $b = -\lambda_{\max}(S) > 0$ .

The solution of inequality (10) can be found by the formula

$$\begin{aligned} &V(x_1(t), \dots, x_n(t)) \\ &\leq V_0 \cdot \exp \int_0^t \left[ \frac{W(x_1(\tau), \dots, x_n(\tau)) - b \cdot (h_1(x_1(\tau), 2\gamma_1) + \dots + h_n(x_n(\tau), 2\gamma_n))}{V(x_1(\tau), \dots, x_n(\tau))} \right] d\tau, \end{aligned} \quad (11)$$

where the constant  $V_0 = V(x_1(0), \dots, x_n(0)) > 0$ .

Note that the functions  $W(x_1, \dots, x_n)$ ,  $V(x_1, \dots, x_n)$ , and  $H(x_1, \dots, x_n) = h_1(x_1, 2\gamma_1) + \dots + h_n(x_n, 2\gamma_n)$  are positive definite power functions. In addition,  $\deg W(x_1, \dots, x_n) = q \cdot \max(\gamma_{ij}, \delta_{ij}) < 2q$ , and  $\deg H(x_1, \dots, x_n) = 2q$ ;  $j = 1, \dots, k_i$ ;  $i = 1, \dots, n$ .

In this case, on the one hand, there exists a moment  $T_0 > 0$  such that if  $t > T_0$ , then  $W(x_1, \dots, x_n) - bH(x_1, \dots, x_n) < 0$ , and

$$\lim_{t \rightarrow \infty} \frac{W(x_1(t), \dots, x_n(t)) - bH(x_1(t), \dots, x_n(t))}{V(x_1(t), \dots, x_n(t))} < 0. \quad (12)$$

Thus, we have  $V(x_1(t), \dots, x_n(t)) \rightarrow 0$  at  $t \rightarrow \infty$ . But on the other hand this fact means that if the function  $V(x_1(t), \dots, x_n(t))$  is small enough, then there exists the moment  $T_1 > T_0 > 0$  such that if  $t > T_1$ , then  $W(x_1(t), \dots, x_n(t)) - bH(x_1(t), \dots, x_n(t)) > 0$  and the positive function  $V(x_1(t), \dots, x_n(t))$  increases, and so on.

Denote by  $\mathbb{H} = \{h_1, \dots, h_n\} \subset \mathbb{R}^n$  the set of all points satisfying condition  $W(h_1, \dots, h_n) - bH(h_1, \dots, h_n) \leq 0$ . Since function  $H(x_1, \dots, x_n)$  is positive definite, then  $\mathbb{H}$  is the compact positively invariant set with respect to (7). Therefore, if solution  $\mathbf{x}(t)$  of system (7) belongs to  $\mathbb{H}$ , then it is bounded. It means that the solution  $V(x_1, \dots, x_n)$  of equation (9) also should be bounded.

Now we assume that there exists a moment  $T_u$  such that for  $t = T_u$   $\mathbf{x}(T_u) \notin \mathbb{H}$ . Then inequality (12) holds.

Denote by  $\mathbb{S}$  the ball centered at the origin whose surface passes through point  $T_u \in \mathbb{S}$ . In this case  $\mathbb{H} \subset \mathbb{S}$ . Therefore, by virtue of (12) and according to LaSalle's Theorem [25], there is a moment  $T_s > T_u$  such that  $\mathbf{x}(T_s) \in \mathbb{H}$ . Again we get that solution  $\mathbf{x}(t)$  of system (7) starting at  $\mathbb{S}$  belongs to  $\mathbb{S}$ . In addition,  $\mathbf{x}(t)$  is attracted to the boundary of  $\mathbb{H}$  as  $t \rightarrow +\infty$ . Thus, it is bounded.

Now we use Comparison Principle [25]. Then it remain to compare the solution  $V(x_1, \dots, x_n)$  of equation (9) and a similar solution of inequality (10). From here it follows the boundedness of solution  $\mathbf{x}(t)$  of system (7) for any initial condition  $\mathbf{x}_0 \in \mathbb{R}^n$ . This completes the proof of case (a1) for strictly odd functions.

(a2) The case of odd function  $h_i(x_i, \alpha_i \vee \beta_i)$ ;  $i = 1, \dots, n$  [25].

Now we can apply Theorem 4.1 [25] to equation (7). Then all the ideas that were used in the proof of Theorem 4.1 [25] can be directly carried over to the proof of Theorem 3. Since

$$\min(\deg h_1(x_1, \alpha_1 \vee \beta_1), \dots, \deg h_n(x_n, \alpha_n \vee \beta_n)) > \max(\deg f_1(x_1, \dots, x_n), \dots, \deg f_n(x_1, \dots, x_n)),$$

it remains to verify only one condition of Theorem 4.1 [25]: the symmetric matrix  $A + A^T - 2rI$  must be negative definite. The last condition can always be achieved by choosing the sufficiently large parameter  $r \geq 0$ . It is clear that the same statement will also hold for system (6). This remark completes the proof.  $\square$

**Comment.** In the general case the function  $V(x_1, \dots, x_n)$  is not the Lyapunov function for system (7). It is guaranteed to be the Lyapunov function if  $\mathbf{f}(\mathbf{x}) \equiv 0$ .

Let us compose from functions (4) the following vector- function:

$$\mathcal{H}(\mathbf{x}, \mathbf{s}_1, \dots, \mathbf{s}_n) = (\mathcal{H}_1(x_1, \mathbf{s}_1), \dots, \mathcal{H}_n(x_n, \mathbf{s}_n))^T,$$

where  $\mathbf{s}_j = (s_{j1}, \dots, s_{jk_j}), j = 1, \dots, n$ .

Let

$$q = \min(\deg \mathcal{H}_1(x_1, \mathbf{s}_1), \dots, \deg \mathcal{H}_n(x_n, \mathbf{s}_n)).$$

**Corollary of Theorem 3.** *Let the vector  $\mathbf{h}(\mathbf{x})$  in system (7) be replaced by the vector  $\mathcal{H}(\mathbf{x}, \mathbf{s}_1, \dots, \mathbf{s}_n)$ . Then, under the conditions of Theorem 3, any solution of system  $\dot{\mathbf{x}}(t) = (A - rI)\mathcal{H}(\mathbf{x}, \mathbf{s}_1, \dots, \mathbf{s}_n) + \mathbf{f}(\mathbf{x})$  is bounded.*

*Proof* repeats the proof of Theorem 3 if in this theorem we replace the function  $V(x_1, \dots, x_n)$  (8) by the function

$$V(x_1, \dots, x_n) = \int \mathcal{H}_1(x_1, \mathbf{s}_1) dx_1 + \dots + \int \mathcal{H}_n(x_n, \mathbf{s}_n) dx_n,$$

where all the indicated integrals are integrals of power functions.  $\square$

Now, we will assume that in system (7)  $r = 0$  and  $\mathbf{f}(\mathbf{x}) = 0$ . Then we get the following system

$$\dot{\mathbf{x}}(t) = A\mathbf{h}(\mathbf{x}). \quad (13)$$

**Theorem 4.** *Let  $\alpha_i > 0$  and  $\beta_i > 0$ , and the power function  $h_i(u_i, \alpha_i \vee \beta_i)$  be odd;  $i = 1, \dots, n$ . Let also  $A \in \mathbb{R}^{n \times n}$  be an antisymmetric matrix such that  $\text{rank } A = m \leq n$ . Let  $\mathbb{V} \subset \mathbb{R}^n$  be a  $(n - m)$ -dimensional subspace of  $\mathbb{R}^n$  such that  $A\mathbb{V} = \mathbf{0}$  (if  $m = n$ , then  $\mathbb{V} = \mathbf{0}$ ; for odd  $n > 1$ , we always have  $m < n$  and  $\mathbb{V} \neq \mathbf{0}$ ). Then any solution  $\mathbf{x}(\mathbf{x}_0, t)$  of system (13) starting from point  $\mathbf{x}_0 = (x_{10}, \dots, x_{n0})^T \notin \mathbb{V}$  is periodic.*

*Proof.* We will again use the Lyapunov function  $V(x_1, \dots, x_n)$  of form (8), which was introduced in the proof of Theorem 3.

Let  $\mathbb{S} = \{\mathbf{x} \in \mathbb{R}^n | V(\mathbf{x}) \leq V(\mathbf{x}_0)\}$ . Then the set  $\mathbb{S} \subset \mathbb{R}^n$  is compact positively invariant with respect to (13). In addition, from here it follows that  $V'_t(\mathbf{x}) = 0.5\mathbf{h}^T(\mathbf{x})(A + A^T)\mathbf{h}(\mathbf{x}) = 0$ . This means that the solution  $\mathbf{x}(\mathbf{x}_0, t)$  of system (13) starting from a point  $\mathbf{x}_0 \in \mathbb{R}^n$  forms a closed curve on the boundary of the set  $\mathbb{S} \subset \mathbb{R}^n$  and therefore the solution  $\mathbf{x}(\mathbf{x}_0, t)$  is periodic [31].  $\square$

Now consider the following system of ordinary differential equations:

$$\dot{\mathbf{x}}(t) = \mathbf{h}(A\mathbf{x} + \mathbf{b}) \iff \begin{cases} \dot{x}_1(t) = h_1(A_1 \cdot \mathbf{x} + b_1) \\ \dots \dots \dots \dots \dots \dots \dots \\ \dot{x}_n(t) = h_n(A_n \cdot \mathbf{x} + b_n) \end{cases} \quad (14)$$

with the known vector of initial values  $\mathbf{x}_0$ . Here  $A_i$  is the  $i$ -th row of the matrix  $A$ ;  $i = 1, \dots, n$ .

**Theorem 5.** *Let, under the conditions of Theorem 4,  $m = n - 1$  and  $n$  be odd. If  $\alpha_i = \beta_i$  and  $b_i = 0$ ,  $i = 1, \dots, n$ , then any solution of system (14) is periodic; if at least for one  $k \in \{1, \dots, n\}$ , we have  $\alpha_k \neq \beta_k$  or  $b_k \neq 0$ , then any solution of system (14) is a winding of an infinite cylinder with generatrix  $\mathbb{V}$ . In addition, there exist numbers  $T > 0$  and  $\lambda = \lambda(T) \in \mathbb{R}$  that do not depend on  $\mathbf{x}_0$ , such that for any solution  $\mathbf{x}(t)$  of system (14), we have*

$$\mathbf{x}(t + T) - \mathbf{x}(t) = \lambda\mathbb{V},$$

where  $\lambda = 0$  if and only if  $\alpha_i = \beta_i$  and  $b_i = 0$ ,  $i = 1, \dots, n$ .

*Proof.* (b1) With the help of suitable changes of variables  $\mathbf{y} = A\mathbf{x} + \mathbf{b}$ , we transform system (14) into the following system (generally speaking, not equivalent to system (14)):

$$\frac{d}{dt}(A\mathbf{x}(t) + \mathbf{b}) = A \cdot \begin{pmatrix} h_1(A_1\mathbf{x}(t) + b_1) \\ \dots \dots \dots \dots \dots \dots \dots \\ h_n(A_n\mathbf{x}(t) + b_n) \end{pmatrix}, \quad (15)$$

where  $\mathbf{b} = (b_1, \dots, b_n)^T$ . Then in the new variables  $\mathbf{y}$  we get the system  $\dot{\mathbf{y}}(t) = A\mathbf{h}(\mathbf{y})$ . (This is system (13).)

It is clear that in this case all the conditions of Theorem 4 are fulfilled and we obtain a periodic solution  $\mathbf{p}(t)$  of the system  $\dot{\mathbf{y}}(t) = A\mathbf{h}(\mathbf{y})$ . In addition, taking into account the new variables  $\mathbf{y}$ , system (14) can be written in the form  $\dot{\mathbf{x}}(t) = \mathbf{h}(\mathbf{y})$ . From here it follows that

$$\mathbf{x}(t) = \int \mathbf{h}(\mathbf{y}(t))dt. \quad (16)$$

It is important to note that the solutions  $\mathbf{q}(t)$  of system (14) are not (generally speaking) periodic. (It is found from the solution of the equation  $\mathbf{p} = A\mathbf{q} + \mathbf{b}$ , where for odd  $n$  matrix  $A$  is singular.)

It is known that the indefinite integral of a continuous periodic function of period  $T$  is the sum of a periodic function of the same period  $T$  and some linear function. Thus, from (16) and (14) it follows that  $\mathbf{q}(t) = \mathbf{q}_T(t) + t\lambda\mathbb{V} \in \mathbb{R}^n$  and

$$\frac{d}{dt}(\mathbf{q}_T(t) + t\lambda) = \mathbf{h}(A(\mathbf{q}_T(t) + t\lambda\mathbb{V}) + \mathbf{b}).$$

Here  $\mathbf{q}_T(t) = (q_{T1}(t), \dots, q_{Tn}(t))^T \in \mathbb{R}^n$  is a periodic vector function.

Let  $\mathbb{W} = A(\mathbb{R}^n)$  be a linear subspace in  $\mathbb{R}^n$ . Denote by  $A|_{\mathbb{W}}$  the restriction of  $A$  to  $\mathbb{W}$ . Then, we have  $\mathbf{q}_T(t) + t\lambda\mathbb{V} = (A|_{\mathbb{W}})^{-1}(\mathbf{p}(t) - \mathbf{b})$ , where  $t\lambda\mathbb{V}$  is a straight line in  $\mathbb{R}^n$  passing through the origin. Thus, the set  $\{\mathbf{q}_T(t) + t\lambda\mathbb{V}\} \subset \mathbb{R}^n$  is a curve wound on the cylinder with generatrix  $\mathbb{V}$ . (The projection of the periodic curve  $\mathbf{q}_T(t) + t\lambda\mathbb{V}$ ,  $t_0 \leq t \leq t_0 + T$  onto any  $(n - 1)$ -dimensional hyperplane  $\mathbb{P} \subset \mathbb{R}^n$  such that  $\mathbb{P} \perp \mathbb{V}$  is a directrix of cylinder.)

(b2) Let  $\mathbb{V} = (v_1, \dots, v_n)^T \in \mathbb{R}^n$ . Consider the situation  $\alpha_1 = \beta_1$  and  $b_1 = 0$ . Then at  $A_1 \cdot \mathbf{x} \geq 0$  the first equation of system (14) will not change, and at  $A_1 \cdot \mathbf{x} < 0$  this equation will be represented as  $\dot{x}_1(t) = -h_1(A_1 \cdot \mathbf{x})$ .

Thus, if  $A_1 \cdot \mathbf{x} \geq 0$ , then  $x_1(t) = q_{T1}(t) + t\lambda \cdot v_1$ ; if  $A_1 \cdot \mathbf{x} < 0$ , then  $x_1(t) = q_{T1}(t) - t\lambda \cdot v_1$ . Since the initial conditions are the same for both equations, then, in accordance with the well-known Cauchy theorem on the existence and uniqueness of solution, there must be  $\lambda = 0$ . Repeating the same reasoning for each of the equations of system (14) for  $\alpha_i = \beta_i$  and  $b_i = 0$ , we finally obtain  $\lambda = 0$ ;  $i = 2, \dots, n$ . The last statement completes the proof of the whole Theorem 5.  $\square$

The next figure demonstrates the statements of Theorem 5 (see Fig. 1):

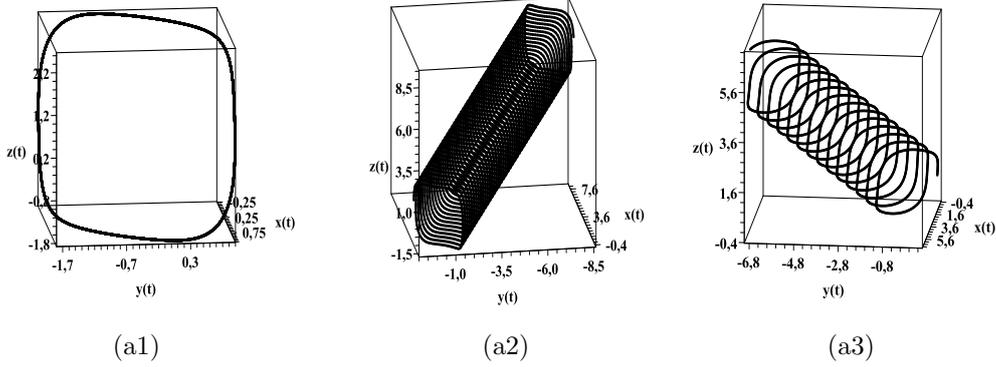


Fig. 1. Trajectories of the system  $\dot{x} = (-y - z + b_1)^{r_1}$ ,  $\dot{y} = (x - z + b_2)^{r_2}$ ,  $\dot{z} = (x + y + b_3)^{r_3}$ :  
(a1)  $r_1 = (1.5 \vee 1.5)$ ,  $r_2 = (2.5 \vee 2.5)$ ,  $r_3 = (2 \vee 2)$ , and  $b_1 = b_2 = b_3 = 0$ ;  
(a2)  $r_1 = (1.5 \vee 1.5)$ ,  $r_2 = (2.5 \vee 2.5)$ ,  $r_3 = (2 \vee 2)$ , and  $b_1 = 1, b_2 = b_3 = 0$ ;  
(a3)  $r_1 = (1.5 \vee 2.5)$ ,  $r_2 = (2.5 \vee 2.5)$ ,  $r_3 = (2 \vee 2)$ , and  $b_1 = b_2 = b_3 = 0$ .

**Definition 4.** System (14) in which matrix  $A$  is antisymmetric, power activation functions  $h_i(u_i, \alpha_i \vee \beta_i)$ ,  $i = 1, \dots, n$ , are odd is called the unperturbed system of antisymmetric neural ODEs.

In addition, the following system

$$\dot{\mathbf{y}}(t) = T^{-1} \cdot \begin{bmatrix} h_1(A_1 \cdot T \cdot \mathbf{y} + b_1) \\ \vdots \\ h_n(A_n \cdot T \cdot \mathbf{y} + b_n) \end{bmatrix} \quad (17)$$

obtained from system (14) of antisymmetric neural ODEs with the help of the linear invertible transformation  $T \in \mathbb{R}^{n \times n}$  ( $\mathbf{x} = T \cdot \mathbf{y}$ ) will also be called the unperturbed system of antisymmetric neural ODEs.

**Definition 5.** Let the matrix  $A$  of system (6) be antisymmetric and  $r \geq 0$ . Then, under the conditions of Theorems 3, system (6) is called a perturbed system of antisymmetric neural ODEs.

## 4 On conditions for the appearance of chaos in system (7)

It is known that the following question often arises when modeling chaotic processes: can the created model generate chaotic behavior? The same question can be extended to model (7) (or (6)).

Let us denote by the symbol

$$\deg \mathbf{h}(\mathbf{x}) = (\deg h_1(x_1, \alpha_1 \vee \beta_1), \dots, \deg h_n(x_n, \alpha_n \vee \beta_n)) \in \mathbb{N}^n,$$

where  $\mathbb{N}$  is the set of natural numbers. Then the inequality  $\deg \mathbf{h}(\mathbf{x}) > \deg \mathbf{v}(\mathbf{x}) = (\deg v_1(x_1, \gamma_1 \vee \delta_1), \dots, \deg v_n(x_n, \gamma_n \vee \delta_n))$  means that  $\alpha_i > \gamma_i \geq 1$  and  $\beta_i > \delta_i \geq 1$ ;  $i = 1, \dots, n$ .

Consider the following simplified version of system (7):

$$\dot{\mathbf{x}}(t) = B\mathbf{x} + \mathbf{v}(\mathbf{x}) + (A - rI)\mathbf{h}(\mathbf{x}). \quad (18)$$

Here matrices  $A, B \in \mathbb{R}^{n \times n}$ , and the matrix  $A$  is antisymmetric; the power vector functions  $\mathbf{v}(\mathbf{x}), \mathbf{h}(\mathbf{x}) \in \mathbb{R}^n$  such that  $\deg \mathbf{h}(\mathbf{x}) > \deg \mathbf{v}(\mathbf{x}) > (1, \dots, 1)$ . (If  $\mathbf{v}(\mathbf{x}) \equiv \mathbf{0}$ , then  $\deg \mathbf{h}(\mathbf{x}) > (1, \dots, 1)$ .)

Let us construct for system (18) a positive definite function  $V(x_1, \dots, x_n)$  (8). Then, we have  $\dot{V}_t(x_1, \dots, x_n) = 0.5 \cdot (\mathbf{h}^T(\mathbf{x})(B\mathbf{x} + \mathbf{v}(\mathbf{x})) + (\mathbf{x}^T B^T + \mathbf{v}^T(\mathbf{x}))\mathbf{h}(\mathbf{x})) - r \cdot \mathbf{h}^T(\mathbf{x})\mathbf{h}(\mathbf{x})$ .

We denote by  $\mathbb{W}$  the set of all points from  $\mathbb{R}^n$  such that  $\dot{V}_t(x_1, \dots, x_n) \leq 0$ . Let also  $\mathbb{L} \subset \mathbb{W}$  be the set of all points in  $\mathbb{W}$  such that  $\dot{V}_t(x_1, \dots, x_n) = 0$ . We also denote by  $\mathbb{X} \subset \mathbb{W}$  an open set in  $\mathbb{W}$  such that  $\forall \mathbf{x} = (x_1, \dots, x_n)^T \in \mathbb{X} \quad \dot{V}_t(x_1, \dots, x_n) < 0$ . (Thus,  $\mathbb{X} \cup \mathbb{L} = \mathbb{W}$ .)

**Theorem 6.** Let the point  $\mathbf{0}$  be a unique equilibrium point for system (18) in  $\mathbb{W}$ . Suppose also that:

- 1) point  $\mathbf{0}$  is a saddle point;
- 2) there exists a value of parameter  $r > 0$  such that for an arbitrary vector of initial values  $(x_{10}, \dots, x_{n0})^T \in \mathbb{X}$  the solutions  $x_1(x_{10}, \dots, x_{n0}, t), \dots, x_n(x_{10}, \dots, x_{n0}, t)$  of system (18) satisfy to equality  $\liminf_{t \rightarrow \infty} V(x_1(x_{10}, \dots, x_{n0}, t), \dots, x_n(x_{10}, \dots, x_{n0}, t)) = 0$ .

Then under the conditions of Theorem 3 in system (18) there exists a chaotic dynamics.

*Proof.* According to Theorem 3, there exists the value  $r = r_0$  such that all solutions of system (18) are bounded. Therefore, for  $r = r_0$  the set  $\mathbb{W}$  is a compact positively invariant set with respect to system (18).

According to LaSalla's theorem every solution of system (18) starting in  $\mathbb{W}$  approaches to the largest invariant set  $\mathbb{M} \subset \mathbb{L}$  as  $t \rightarrow \infty$  (see [25, 31]). In our case, by assumption  $A^T + A = 0$  and condition 1) of Theorem 6, the role of the set  $\mathbb{M}$  can be played either by a limit cycle or by a homoclinic trajectory connected at  $\mathbf{0}$  (see Theorem 4). If both conditions 1) and 2) of Theorem 6 are satisfied, then there exists a sequence of values  $r_{01} > r_{02} > \dots > r_{0k} > \dots$  of parameter  $r$  such that  $\lim_{k \rightarrow \infty} r_{0k} = r_c > 0$  and the set  $\mathbb{M}(r_c)$  at the critical value  $r_c$  is the homoclinic orbit. (Indeed, let  $\mathbb{N}_s$  and  $\mathbb{N}_u$  be stable and unstable manifolds of the point  $\mathbf{0}$  [21, 34]. Let's denote by  $\mathbf{x}_0 = (x_{01}, \dots, x_{0n})^T \in \mathbb{N}_u$  the starting point. Since at the point  $\mathbf{x}_0$  we have  $\dot{V}_t(x_{01}, \dots, x_{0n}) \geq 0$ , then the solution  $\mathbf{x}(\mathbf{x}_0, t)$  of system (18) should be attracted to a certain limit cycle in  $\mathbb{L}$ . According to condition 2) of Theorem 6, this limit cycle for some value of parameter  $r$  will pass arbitrarily close to the origin (to the manifold  $\mathbb{N}_s$ ). In other words, near point  $\mathbf{0}$  on trajectory  $\mathbf{x}(\mathbf{x}_0, t)$  there will be point  $\mathbf{x}_1 = (x_{11}, \dots, x_{1n})^T$  such that  $\dot{V}_t(x_{11}, \dots, x_{1n}) \leq 0$ . Therefore, there must be the value  $r_c$  of parameter  $r$  for which  $\mathbb{N}_s \cap \mathbb{N}_u \neq \emptyset$ . This means the existence of homoclinic orbit.)

The last statement enables us to construct a discrete mapping for system (18). Theorems 4 and 5 allow us to assert that a limit cycle can exist in perturbed system (18) for some values of parameter  $r$ . Let  $T_0$  be the period of this cycle. In this case, the continuous relation (11) for the positive definite function  $V(t) \equiv V(x_1(t), \dots, x_n(t))$  can be rewritten as

$$V(x_1(t_{k+1}), \dots, x_n(t_{k+1})) \leq V(x_1(t_k), \dots, x_n(t_k)) \\ \times \exp \int_{t_k}^{t_{k+1}} \left[ \frac{W(x_1(\tau), \dots, x_n(\tau)) - b \cdot (h_1(x_1(\tau), 2\gamma_1) + \dots + h_n(x_n(\tau), 2\gamma_n))}{V(x_1(\tau), \dots, x_n(\tau))} \right] d\tau, t_{k+1} - t_k = T_0;$$

where  $k = 0, 1, \dots$

A discrete analogue of this relation according to the technique described in [32] – [34] can be represented in the following form:

$$V_{k+1} = V_k \exp(p + \phi(V_k) - rV_k^q); k = 0, 1, 2, \dots \quad (19)$$

Here  $p > 0$ ;  $\phi(u)$  is a linear combination of power functions  $u^{q_i}$  of one variable  $u > 0$  (with  $q_i > 0$ ) and  $0 < \deg \phi(u) = \max q_i < q$ ,  $i = 1, \dots, l$ .

In [32] – [34] it is shown that for some  $r = r_c$  mapping (19) generates a chaotic dynamics. Therefore, system (18) at  $r = r_c$  will also exhibit chaotic behavior.  $\square$

**Comment 1.** Theorems 4 and 5 guarantee the existence of periodic trajectories for unperturbed systems. This means that for small perturbations, periodic motions (limit cycles) will appear in system (18). Namely the existence of limit cycle in system (18) allows starting a cascade of bifurcations leading to the appearance of a chaotic attractor.

**Comment 2.** Note that the condition  $A^T + A = 0$  cannot be omitted. Indeed, if this condition is satisfied, then  $\forall r > 0$   $A^T + A - 2rI = -2rI < 0$ . Otherwise (at  $A^T + A \neq 0$ ), the inequality  $A^T + A - 2rI < 0$  does not hold  $\forall r > 0$ . Therefore, if  $A^T + A \neq 0$  condition 2) of Theorem 6 generally speaking cannot be achieved.

Nevertheless, the results obtained for system (18) can be strengthened if matrix  $A$  is replaced by a similar non-antisymmetric matrix  $A_n = H^{-1}AH$  ( $A_n^T + A_n \neq 0$ ):

$$\dot{\mathbf{x}}(t) = B\mathbf{x} + \mathbf{v}(\mathbf{x}) + (A_n - rI)\mathbf{h}(\mathbf{x}). \quad (20)$$

In order to check condition 2) of Theorem 6 we can proceed in the following way.

1. In system (18) set parameter  $r$  large enough. 2. Decrease the parameter  $r$  until the limit cycle appears in system (18). 3. Continue slowly decreasing the parameter  $r$  until limit cycles of any multiplicity appear in system (18). (This means that a non-negative function  $V(x_1(t), \dots, x_n(t))$  will have any number of minimums.) If parameter  $r$  is close enough to value  $r_c$  (which is unknown), then the minimums of function  $V(x_1(t), \dots, x_n(t))$  will take arbitrarily small (not zero!) values. This will mean that  $r \approx r_c$  (see Fig. 1).

#### 4.1 Homoclinic chaos

A large number of chaotic attractors arise when so-called homoclinic orbits exist in a dynamical system (see [18] – [21], [32] – [37], and references to papers on chaotic dynamics cited elsewhere).

**Definition 6** [35]. A bounded trajectory  $\mathbf{x}(t, \mathbf{x}_0) \in \mathbb{R}^n$  of system (7) is called a homoclinic orbit if the trajectory converges to the same equilibrium point as  $t \rightarrow \pm\infty$ .

Let  $A = (a_1, \dots, a_n) \in \mathbb{R}^{n \times n}$  be the antisymmetric matrix composed of columns  $a_i \in \mathbb{R}^n; i = 1, \dots, n$ .

Let us  $\forall s \in \{1, \dots, n\}$  denote by symbol  $A_s \in \mathbb{R}^{n \times n}$  ( $A^s \in \mathbb{R}^{n \times n}$ ) the matrix obtained from matrix  $A$  by replacing the column  $a_s$  (row  $-a_s^T$ ) with the column  $-a_s$  (row  $a_s^T$ ).

**Definition 7.** Matrix  $A_s$  (or  $A^s$ ) will be called partially antisymmetric.

Let  $p < n$  be a positive integer. If columns  $a_{s_1}, \dots, a_{s_p}, 1 \leq s_1 < \dots < s_p \leq n$  of the matrix  $A$  are replaced by columns  $-a_{s_1}, \dots, -a_{s_p}$ , then the resulting matrix will also be denoted as  $A_{s_1 \dots s_p}$  and called partially antisymmetric. A similar designation  $A^{s_1 \dots s_p}$  is retained for the rows.

Let  $\mathbf{h}(\mathbf{x}) = (h_1(x_1, \alpha_1 \vee \beta_1), \dots, h_n(x_n, \alpha_n \vee \beta_n))^T$  and  $\mathbf{f}(\mathbf{x}) = (f_1(x_1, \gamma_1 \vee \delta_1), \dots, f_n(x_n, \gamma_n \vee \delta_n))^T$  be two vectors of power functions.

**Theorem 7.** Let  $\psi$  and  $\phi \neq 0$  be arbitrary real numbers and let  $\mathbf{h}(\mathbf{x})$  be odd function. Consider the following system of ordinary differential equations

$$\dot{\mathbf{x}}(t) = \psi A_{s_1 \dots s_p} \mathbf{f}(\mathbf{x}) + \phi A \mathbf{h}(\mathbf{x}), \quad (21)$$

in which it is assumed that  $\deg \mathbf{h}(\mathbf{x}) > \deg \mathbf{f}(\mathbf{x})$ . Then any solution  $\mathbf{x}(t, \mathbf{x}_0)$  of system (21) is periodic. In addition, if the function  $\mathbf{f}(\mathbf{x})$  is odd, then there exists a vector of initial conditions  $\mathbf{x}_0^*$  such that trajectory  $\mathbf{x}(t, \mathbf{x}_0^*)$  is the homoclinic orbit connected at equilibrium point  $\mathbf{0}$ .

*Proof.* (a1) Let  $\psi = 0$ . Then the assertion of Theorem 7 follows from Theorem 4.

Let  $p = 1$  and  $s_1 = s$ . Introduce the vector  $\mathbf{f}_s(\mathbf{x}) = (f_1(x_1), \dots, -f_s(x_s), \dots, f_n(x_n))^T$ , where  $s \in \{1, \dots, n\}$ .

(a2) Let  $\psi \cdot \phi \neq 0$ .

Consider two scalar products:

$$\mathbf{h}^T(\mathbf{x}) \cdot \dot{\mathbf{x}} = \psi \mathbf{h}^T(\mathbf{x}) \cdot A_s \mathbf{f}(\mathbf{x}) + \phi \mathbf{h}^T(\mathbf{x}) \cdot A \mathbf{h}(\mathbf{x}) = \psi \mathbf{h}^T(\mathbf{x}) A \mathbf{f}_s(\mathbf{x})$$

and

$$\mathbf{f}_s^T(\mathbf{x}) \cdot \dot{\mathbf{x}} = \psi \mathbf{f}_s^T(\mathbf{x}) \cdot A_s \mathbf{f}(\mathbf{x}) + \phi \mathbf{f}_s^T(\mathbf{x}) \cdot A \mathbf{h}(\mathbf{x}) = \psi \mathbf{f}_s^T(\mathbf{x}) \cdot A_s^s \mathbf{f}(\mathbf{x}) + \phi \mathbf{f}_s^T(\mathbf{x}) \cdot A \mathbf{h}(\mathbf{x}) = \phi \mathbf{h}^T(\mathbf{x}) A^T \mathbf{f}_s(\mathbf{x}).$$
 (Here we have used the obvious equality:  $A_s^s + (A_s^s)^T = 0$ .)

From here it follows that  $\phi \mathbf{h}^T(\mathbf{x}) \cdot \dot{\mathbf{x}} + \psi \mathbf{f}_s^T(\mathbf{x}) \cdot \dot{\mathbf{x}} = 0$ . Therefore, if we define the derivative  $\dot{V}(\mathbf{x})$  of some function  $V(\mathbf{x})$  by formula  $\dot{V}(\mathbf{x}) \equiv \phi \mathbf{h}^T(\mathbf{x}) \cdot \dot{\mathbf{x}} + \psi \mathbf{f}_s^T(\mathbf{x}) \cdot \dot{\mathbf{x}}$ , then we will have

$$\begin{aligned} V(\mathbf{x}) \equiv & \phi \frac{h_1(x_1, \gamma_1 + 1)}{\gamma_1 + 1} + \dots + \phi \frac{h_n(x_n, \gamma_n + 1)}{\gamma_n + 1} + \psi \frac{f_1(x_1, \xi_1 + 1)}{\xi_1 + 1} + \dots \\ & + \psi \frac{f_{s-1}(x_{s-1}, \xi_{s-1} + 1)}{\xi_{s-1} + 1} - \psi \frac{f_s(x_s, \xi_s + 1)}{\xi_s + 1} + \dots + \psi \frac{f_n(x_n, \xi_n + 1)}{\xi_n + 1} = C(\mathbf{x}_0) = const, \end{aligned}$$

where

$$\forall i \in \{1, \dots, n\} \quad \gamma_i = \begin{cases} \beta_i, & \text{if } x_i < 0, \\ \alpha_i, & \text{if } x_i \geq 0, \end{cases} \quad \xi_i = \begin{cases} \delta_i, & \text{if } x_i < 0, \\ \gamma_i, & \text{if } x_i \geq 0. \end{cases}$$

Without loss of generality, we can assume that  $\phi > 0$ . Then from the conditions of Theorem 7 it follows that for a sufficiently large norm  $\|\mathbf{x}\|$  of the vector  $\mathbf{x}$ , we have  $V(\mathbf{x}) > 0$ . This means that if  $V(\mathbf{x}_0) = C(\mathbf{x}_0) > 0$ , then the function  $V(\mathbf{x}) = C(\mathbf{x}_0) > 0$  is bounded. Consequently, any solution of system (21) is closed and therefore periodic.

Now let the vector of initial conditions  $\mathbf{x}_0^*$  be such that  $V(\mathbf{x}_0^*) = 0$ . Therefore, if the function  $\mathbf{f}(\mathbf{x})$  is odd, then the function  $V(\mathbf{x})$  is even. From here it follows that  $V(\mathbf{x}_0^*) = V(-\mathbf{x}_0^*) = 0$  and in addition,  $V(\mathbf{0}) = 0$ . Consequently, there is a trajectory of the system (21), which for

$t = 0$  leaves some point  $\mathbf{x}_0 \neq \mathbf{0}$  arbitrarily close to  $\mathbf{x}_0^* \approx \mathbf{0}$  and at  $t \rightarrow \pm\infty$  approaches to the point  $\mathbf{0}$ .

Now, in system (21), we will change the sign of time:  $t \rightarrow -\tau$ . Then we will have

$$\dot{\mathbf{x}}(\tau) = -\psi A_{s_1 \dots s_p} \mathbf{f}(\mathbf{x}) - \phi A \mathbf{h}(\mathbf{x}). \quad (22)$$

Further, we apply to the study of solutions of equation (22) the same technique as for the study of solutions of equation (21). As a result, instead of analyzing of the equation  $V(\mathbf{x}(t)) = C(\mathbf{x}_{t=0}) > 0$ , we come to an analysis of the equation  $-V(\mathbf{x}(\tau)) = -C(\mathbf{x}_{\tau=0}) < 0$ . Thus, based on the structure of function  $V(\mathbf{x}(t))$ , we have  $V(\mathbf{x}(t)) = V(\mathbf{x}(\tau)) = V(\mathbf{x}(-t)) = V(-\mathbf{x}(t)) = C(\mathbf{x}_0) > 0$ . This means the existence of a homoclinic orbit.

(a3) Now let  $p > 1$ . Introduce the vector

$$\mathbf{f}_{s_1 \dots s_p}(\mathbf{x}) = (f_1(x_1), \dots, f_{s_1-1}(x_{s_1-1}), -f_{s_1}(x_{s_1}), \dots, -f_{s_p}(x_{s_p}), f_{s_{p+1}}(x_{s_{p+1}}), \dots, f_n(x_n))^T.$$

Then the proof of case  $p > 1$  completely repeats the proof of case  $p = 1$  if we take into account the obvious equality:  $A_{s_1 \dots s_p}^{s_1 \dots s_p} + (A_{s_1 \dots s_p}^{s_1 \dots s_p})^T = 0$ .  $\square$

Since  $\phi \neq 0$ , we can assume that  $\phi = 1$ . Consider the following special case of system (18):

$$\dot{\mathbf{x}}(t) = B\mathbf{x} + \psi A_{s_1 \dots s_p} \mathbf{f}(\mathbf{x}) + (A - rI)\mathbf{h}(\mathbf{x}), r \geq 0. \quad (23)$$

(Here the matrix  $A$  is antisymmetric and the matrix  $A_{s_1 \dots s_p}$  is partially antisymmetric.)

Thus, if conditions of Theorems 6 and 7 are satisfied for system (23), then we will get some chaotic behavior of this system. (Note that if the vectors  $\mathbf{h}(\mathbf{x})$  and  $\mathbf{f}(\mathbf{x})$  in system (21) is replaced by the vectors  $\mathcal{H}(\mathbf{x}, \mathbf{s}_1, \dots, \mathbf{s}_n)$  and  $\mathcal{F}(\mathbf{x}, \mathbf{s}_1, \dots, \mathbf{s}_n)$ , then by virtue of Corollary of Theorem 3, the statement of Theorem 7 is preserved.)

**Definition 8.** *Chaos arising in system (23) will be called homoclinic.*

1. Suppose that in system (21) we have  $n = 2$ ,  $\psi = 1$ , and  $\mathbf{f}(\mathbf{x}) = (x, y)^T$ ,  $\mathbf{h}(\mathbf{x}) = (x^3, y^3)^T$ ,

$$A_s = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Then, depending on parameter  $\phi$ , we obtain the following phase portraits (see Fig. 2):

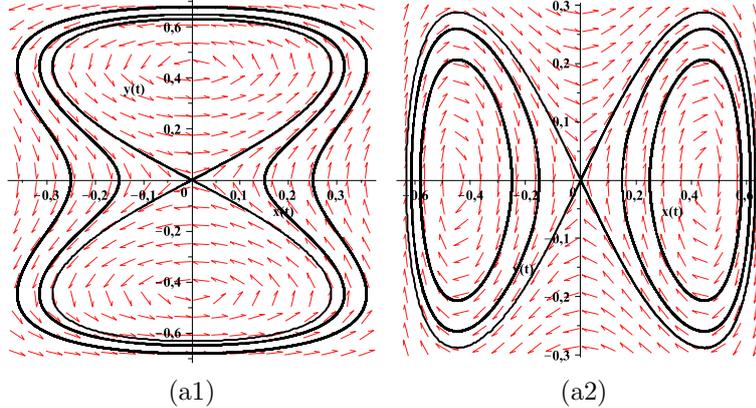


Fig. 2. Two types of homoclinic orbits in system (21) at  $\phi = 5$  (a1) and  $\phi = -5$  (a2). The homoclinic orbit resulting from the transformation (at  $\|\mathbf{x}_0\| \rightarrow 0$ ): (a1) one periodic trajectory and its self-intersection at point  $\mathbf{0}$ ; (a2) two different periodic trajectories lying to the left and right of the vertical axis and their merging at point  $\mathbf{0}$ . The point  $\mathbf{0}$  is saddle; the eigenvalues of the Jacobi matrix at the point  $\mathbf{0}$  are  $\pm 1$ .

2. Now suppose that in system (23) we have  $n = 3$ ,  $r = 0.01$ , and  $\mathbf{f}(\mathbf{x}) = (x^{1.2}, y^{1.2}, z^{1.2})^T$ ,  $\mathbf{h}(\mathbf{x}) = (x^2, y^2, z^2)^T$ ,

$$B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -0.25 & -0.46 \\ 0 & 0.46 & 0.26 \end{pmatrix}, A_s = \begin{pmatrix} 0 & -0.8 & -0.675 \\ 0.8 & 0 & 0.7 \\ -0.675 & 0.7 & 0 \end{pmatrix},$$

$$A = \begin{pmatrix} -0.01 & -0.8 & 0.675 \\ 0.8 & -0.01 & -0.7 \\ -0.675 & 0.7 & -0.01 \end{pmatrix}.$$

Then, depending on parameter  $\psi$ , we obtain the following phase portraits (see Fig.3):

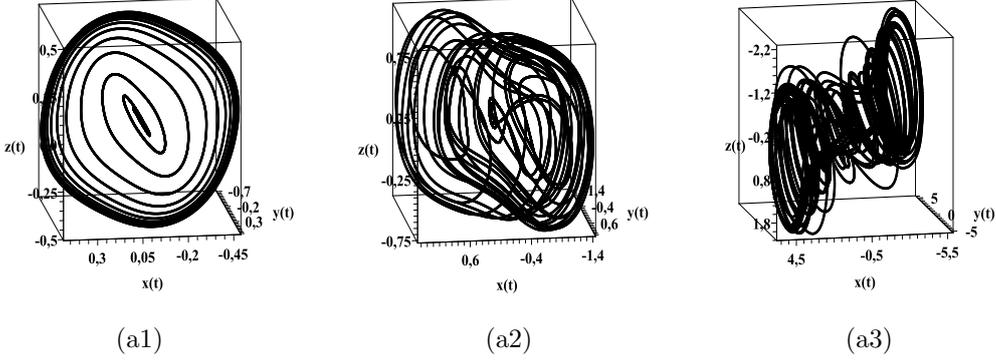


Fig. 3. The birth of homoclinic chaos in system (23), depending on the change in parameter  $\psi$ : (a1)  $\psi = -0.5$ ; (a2)  $\psi = -1$ ; (a3)  $\psi = -3$ .

The result of Theorem 7 can be generalized in the following way.  
Let  $P \in \mathbb{R}^{n \times n}$  be a real matrix. We introduce the vector

$$\mathbf{h}(P\mathbf{x}) = \left( h_1 \left( \sum_{j=1}^n p_{1j} x_j \right), \dots, h_n \left( \sum_{j=1}^n p_{nj} x_j \right) \right)^T$$

(see (14)).

Introduce also the following  $l = 2^n$  matrices:

$$G_1 = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, G_2 = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, \dots$$

$$G_{l-1} = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}, G_l = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 \end{pmatrix}.$$

Then, Equation (23) can be rewritten as

$$\dot{\mathbf{x}}(t) = B\mathbf{x} + \psi A G_i \mathbf{f}(\mathbf{x}) + (A - rI)\mathbf{h}(\mathbf{x}); i \in \{1, \dots, l\}. \quad (24)$$

Equation (24) can be used in two directions.

3. Simulation of systems containing a limit cycle.

In this case,  $\psi = 0$  and instead of (24), you can use the system

$$\dot{\mathbf{x}}(t) = \mathbf{c} + B\mathbf{x} + (A - rI)\mathbf{h}(\mathbf{x}), \mathbf{c} \in \mathbb{R}^n, r > 0$$

or the system

$$\dot{\mathbf{x}}(t) = \mathbf{c} + B\mathbf{x} + \mathbf{h}((A - rI)\mathbf{x}), \mathbf{c} \in \mathbb{R}^n, r > 0. \quad (25)$$

(Here, the shift vector  $\mathbf{c}$  was added to the right side of system (24).)

4. Simulation of systems containing a homoclinic orbit (for example, Lorenz-like systems [37]).

Clearly, if  $\psi \neq 0$ , then among the solutions of equation (24) there exists a homoclinic orbit.

In addition, the homoclinic orbit can be obtained in the following way.

Let  $\psi = -1$ ,  $\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{x})$ ,  $\mathbf{c}, \mathbf{c}_0 \in \mathbb{R}^n$ ,  $r \geq 0$ , and instead of (24), you can use the system

$$\dot{\mathbf{x}}(t) = \mathbf{c} + B\mathbf{x} + \mathbf{h}((A - rD_0)G_i\mathbf{x} + \mathbf{c}_0) - \mathbf{h}((A - rD_0)G_j\mathbf{x} + \mathbf{c}_0), \quad (26)$$

where  $D_0 = \text{diag}(d_{11}, \dots, d_{nn}) \in \mathbb{R}^{n \times n}$ ;  $i, j \in \{1, \dots, l\}$ . (The case  $D_0 = I$  is not excluded.)

Then, for certain values of the parameters, the homoclinic orbit will exist among the solutions of system (26). (Note that in system (26) the components  $h_i(v_i)$ ,  $i = 1, \dots, n$ , of vector  $\mathbf{h}(\mathbf{v})$  can be either even or odd activation functions.)

For example, for the Lorenz system

$$\dot{x} = \sigma(y - x), \dot{y} = ax - y - xz, \dot{z} = -bz + xy, \quad (27)$$

we have:  $\mathbf{h}(\mathbf{x}) = \mathbf{h}(x, y, z) = (h_1(x), h_2(y), h_3(z))^T = (x^2, y^2, z^2)^T$ ,  $\mathbf{c} = \mathbf{c}_0 = 0, r = 0$ ,

$$B = \begin{pmatrix} -\sigma & \sigma & 0 \\ a & -1 & 0 \\ 0 & 0 & -b \end{pmatrix}, A = AG_1 = 0.5 \begin{pmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{pmatrix}, AG_2 = 0.5 \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix},$$

$$\mathbf{h}(AG_1\mathbf{x}) - \mathbf{h}(AG_2\mathbf{x}) = \begin{pmatrix} 0.25(y+z)^2 \\ 0.25(-x+z)^2 \\ 0.25(-x-y)^2 \end{pmatrix} - \begin{pmatrix} 0.25(y+z)^2 \\ 0.25(x+z)^2 \\ 0.25(x-y)^2 \end{pmatrix} = \begin{pmatrix} 0 \\ -xz \\ xy \end{pmatrix}.$$

The use of representation (26) to approximate derivatives can be motivated by the following considerations.

Let  $h(u)$  be the even activation function (3). Then the function  $\phi(u) = h(au + c) - h(-au + c)$ ;  $a, c \in \mathbb{R}$ , is the extended odd activation function [25].

Indeed, without loss of generality, we can assume that  $a \geq 0$  and  $c \geq 0$ . Let  $u^*$  be the root of function  $\phi(u)$  (it is easy to check that this root is unique).

It is clear that if  $u \geq u^*$  ( $u < u^*$ ), then  $\phi(u) \geq 0$  ( $\phi(u) < 0$ ). Thus, if there are numbers  $k, i, j$  such that  $h_k((A - rI)G_i\mathbf{x}) - h_k((A - rI)G_j\mathbf{x}) \equiv 0$ , then the  $k$ -th equation of system (26) is a composition of only odd activation functions (linear and  $\phi(u_k)$ , where  $u_k$  is a function of  $\mathbf{x}$ );  $k \in \{1, \dots, n\}; i \neq j$ . Since any of these functions separates points (see Definition 1), the  $k$ -th equation of system (26) satisfies the approximation Theorem 2. (Let  $\mathbf{c}_0 \neq 0$ . Obviously, if numbers  $i, j \in \{1, \dots, l\}$ , and  $i \neq j$  are such that  $G_i + G_j = 0$ , then all nonzero components of vector  $\mathbf{h}((A - rI)G_i\mathbf{x} + \mathbf{c}_0) - \mathbf{h}((A - rI)G_j\mathbf{x} + \mathbf{c}_0)$  are odd activation functions.)

## 4.2 Examples of attractors of 3D power systems

**1. Chaotic attractor.** Consider a specific system (18) for  $n = 3$ . Assume that in this system

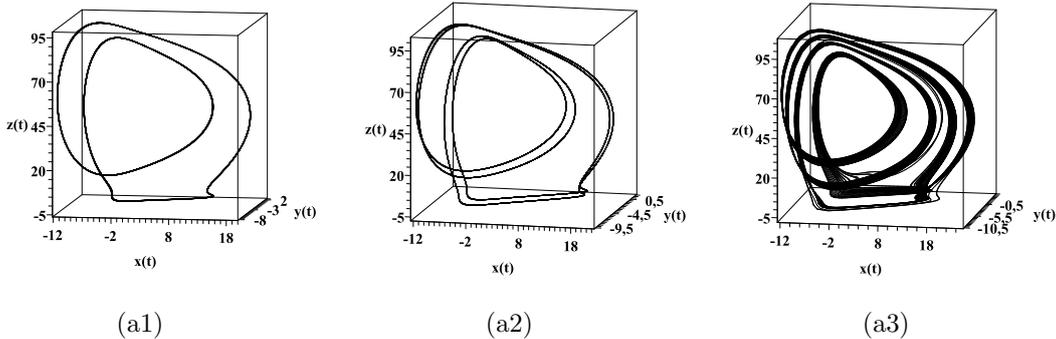
$$\mathbf{v}(\mathbf{x}) \equiv \mathbf{0}, B = \begin{pmatrix} 0.1 & -0.2 & 4.8 \\ -2.9 & -1.4 & 2.0 \\ -0.1 & -1.7 & 1.9 \end{pmatrix}, A - rI = \begin{pmatrix} -r & -0.1 & -0.7 \\ 0.1 & -r & -0.2 \\ 0.7 & 0.2 & -r \end{pmatrix}.$$

The eigenvalues of the matrix  $B$  are  $\lambda_{1,2} = -1.1645 \pm 2.4811i, \lambda_3 = 2.9291$ . Thus, the origin is the saddle focus.

We introduce the following vectors:  $\mathbf{x} = (x, y, z)^T$ ,  $\mathbf{h}(\mathbf{x}) = (h_1(x), h_2(y), h_3(z))^T$ , where

$$\begin{aligned} h_1(x) &= -(-x)^{3.0} \text{ if } (x < 0) \text{ and } x^{1.8} \text{ if } (x \geq 0) \\ h_2(y) &= -(-y)^{3.0} \text{ if } (y < 0) \text{ and } y^{1.8} \text{ if } (y \geq 0) \\ h_3(z) &= -(-z)^{3.0} \text{ if } (z < 0) \text{ and } z^{1.5} \text{ if } (z \geq 0). \end{aligned}$$

The following Figure 4 shows the transition of system (18) from regular regime to chaotic behavior:



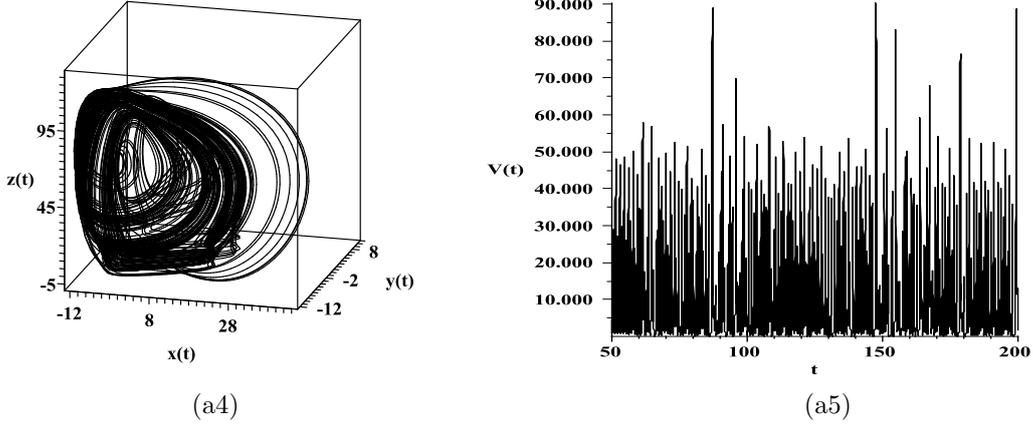


Fig. 4. A cascade of bifurcations of the limit cycle of system (18) for different values of  $r$  (transition to chaos): (a1)  $r = 0.024$ , (a2)  $r = 0.019$ , (a3)  $r = 0.0183$ , (a4)  $r = r_c \approx 0.0142$  (it is chaos); the verification of condition 2) of Theorem 6 is shown in the graph (a5).

Thus, we have shown that systems (6) can simulate a chaotic processes.

**2. Strange attractor.** Theorems 3 and 6 are valid if  $r > 0$ . In case  $r = 0$  the statements of these theorems have not been proved. In this regard, consider 3D system (20) in which  $r = 0$  (compare with system (17)):

$$\mathbf{v}(\mathbf{x}) \equiv \mathbf{0}, B = \begin{pmatrix} 0 & 0.01 & 0 \\ -0.01 & 0 & 1 \\ 1 & -2 & -0.05 \end{pmatrix}, A_n = H^{-1}AH = \begin{pmatrix} 0 & 0 & 0 \\ 0.01 & 0 & -0.1 \\ 0 & 0.005 & 0 \end{pmatrix}.$$

Figure 5 presents a new type of attractors that can be generated by system (20) with odd activation functions  $f_1(x) = x, f_2(y) = y, f_3(z) = z$  (here  $\alpha = \beta = 1$ ) and  $h_1(x) = x^{\gamma \vee \delta}, h_2(y) = y^{\gamma \vee \delta}, h_3(z) = z^{\gamma \vee \delta}$ :

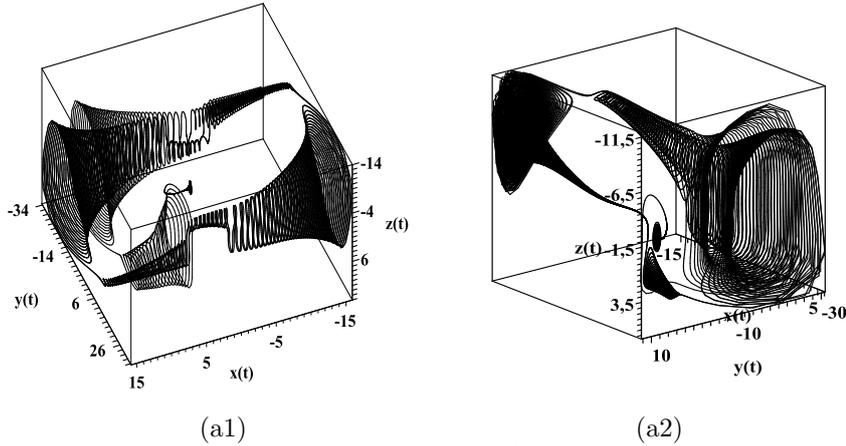


Fig. 5. New strange attractors in system (20): (a1)  $\gamma = 3.5714, \delta = 3$  and (a2)  $\gamma = 4.1428, \delta = 3$

Note that matrix  $A_n = H^{-1}AH$  is similar to antisymmetric matrix

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -\sqrt{5}/100 \\ 0 & \sqrt{5}/100 & 0 \end{pmatrix}.$$

Thus, system (20) allows simulating more complex processes than system (18).



where  $0 < 1 + 3n \leq m \leq N$ . (The number of training elements  $m$  in the time series should be more than the number of unknown coefficients  $1 + 3n$  in any of the equations of the system.)

1.1. Introduce positive numbers  $step, \gamma, \delta$ , and integers  $M_a > 0, M_b > 0$ . (Let, for example, be:  $step := 0.1, \gamma := 1.5, \delta := 1.5$ , and  $M_a := 20, M_b := 20$ .)

1.2. Construct the matrix of numerical derivatives

$$DER := \frac{1}{\Delta t} \begin{pmatrix} g_{11} - g_{10} & g_{21} - g_{20} & \cdots & g_{n1} - g_{n0} \\ g_{12} - g_{11} & g_{22} - g_{21} & \cdots & g_{n2} - g_{n1} \\ \vdots & \vdots & \vdots & \vdots \\ g_{1,m} - g_{1,m-1} & g_{2,m} - g_{2,m-1} & \cdots & g_{n,m} - g_{n,m-1} \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

1.3. Construct the matrix

$$U := \begin{pmatrix} u(g_{11}) & u(g_{21}) & \cdots & u(g_{n1}) \\ u(g_{12}) & u(g_{22}) & \cdots & u(g_{n2}) \\ \vdots & \vdots & \vdots & \vdots \\ u(g_{1,m}) & u(g_{2,m}) & \cdots & u(g_{n,m}) \end{pmatrix} \in \mathbb{R}^{m \times n}.$$

2. Put  $k_a := 1, k_b := 1$ . (Start a double cycle for integer variables  $k_a$  and  $k_b$ .)

2.1. Fix the number  $\alpha := \gamma + k_a \cdot step$ .

2.2. Fix the number  $\beta := \delta + k_b \cdot step$ .

2.3. Construct a perturbation matrix

$$H := \begin{pmatrix} h(g_{11}) & h(g_{21}) & \cdots & h(g_{n1}) \\ h(g_{12}) & h(g_{22}) & \cdots & h(g_{n2}) \\ \vdots & \vdots & \vdots & \vdots \\ h(g_{1,m}) & h(g_{2,m}) & \cdots & h(g_{n,m}) \end{pmatrix} \in \mathbb{R}^{m \times n},$$

and the Jacobi matrix for system (28):

$$W := \left( \begin{array}{cccc|cc} 1 & g_{11} & \cdots & g_{n1} & & \\ 1 & g_{12} & \cdots & g_{n2} & & \\ \vdots & \vdots & \vdots & \vdots & & \\ 1 & g_{1,m} & \cdots & g_{n,m} & U & H \end{array} \right) \in \mathbb{R}^{m \times (3n+1)}.$$

2.4. Check the condition  $\text{rank } W = 3n + 1$ . If  $\text{rank } W < 3n + 1$  then fix some number  $\mu \in (0, 1]$  else  $\mu := 0$ .

2.5. Find the matrix  $Y$  using the least squares method:

$$Y^T = (W^T W + \mu I)^{-1} W^T \cdot DER \in \mathbb{R}^{(3n+1) \times n}.$$

(Here  $I \in \mathbb{R}^{(3n+1) \times (3n+1)}$  is the identity matrix.)

3. Compute the matrix of errors

$$ERR := (e_{ij}) = DER - W \cdot Y^T \in \mathbb{R}^{m \times n}$$

and the total computational error at iterative steps  $k_a$  and  $k_b$ :

$$E_{k_a, k_b} := \sum_{i=1}^m \sum_{j=1}^n e_{ij}^2.$$

4. If  $k_b < M_b$  then  $k_b := k_b + 1$  and go to item 2.2.

4.1. If  $k_a = M_a$  then go to item 5 else  $k_b := 1, k_a := k_a + 1$  and go to item 2.1.

5. Compute the integer numbers  $l_a \in \{1, \dots, M_a\}$  and  $l_b \in \{1, \dots, M_b\}$  such that

$$E_{l_a, l_b} = \min(E_{1,1}, \dots, E_{1,M_b}, E_{2,1}, \dots, E_{2,M_b}, \dots, E_{M_a,1}, \dots, E_{M_a, M_b}).$$

(End of the double cycle for integer variables  $k_a$  and  $k_b$ .)

6. Print the matrix  $Y \equiv Y_{l_a, l_b} \in \mathbb{R}^{n \times (3n+1)}$ , the numbers  $\alpha > 1, \beta > 1$  and stop the algorithm.

7. Solve system (28). If the solutions of system (28) diverge, then to increase the values  $\gamma$  and  $\delta$  by a small number  $\Delta > 0$  ( $\gamma := \gamma + \Delta > 0, \delta := \delta + \Delta$ ) and go to item 2. If for several values  $\gamma$  and  $\delta$  the solutions of system (28) still diverge, then stop the algorithm.

**Comment 1.** If there are no matrices  $B$  or  $D$  in system (28), then there are no fixed degrees of activation functions. Indeed, let  $B = 0$ , then Algorithm 1 works directly. If  $D = 0$  holds, then the following redesignations of variables  $B \rightarrow D, \gamma \rightarrow \alpha$ , and  $\delta \rightarrow \beta$  should be introduced into Algorithm 1.

**Comment 2.** If the matrix  $D + D^T$  is negative definite, then the resulting model of neural ODE will generate bounded solutions (see [25]).

## 5.2 Algorithm 2: antisymmetric matrices are used in the architecture of neural ODEs

It is known that in most real dynamical systems, chaotic phenomena arise as a result of the development of certain periodic processes. In this regard, in order to model the chaotic behavior of such systems, it is desirable to include a mechanism generating a limit cycle in the architecture of the neural network. In turn, the cascade of bifurcations of the limit cycle will lead to the appearance of chaos in the modeled system. The following algorithm is designed to simulate limit cycles in real dynamic processes.

As an object of modeling, we will choose some generalization of system (25). Consider the following matrix

$$S = \begin{pmatrix} S_1 \\ S_2 \\ \vdots \\ S_{n-1} \\ S_n \end{pmatrix} = \begin{pmatrix} -d_{11} & d_{12} & \dots & d_{1,n-1} & d_{1n} \\ -d_{12} & -d_{22} & \dots & d_{2,n-1} & d_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -d_{1,n-1} & -d_{2,n-1} & \dots & -d_{n-1,n-1} & d_{n-1,n} \\ -d_{1n} & -d_{2n} & \dots & d_{n-1,n} & -d_{nn} \end{pmatrix} \in \mathbb{R}^{n \times n} \quad (29)$$

such that  $(S + D_0) + (S + D_0)^T = 0$ . (Thus,  $(S + D_0)$  is antisymmetric.)

Now, we introduce the following system

$$\begin{cases} \dot{x}_1(t) = f_1(c_{10} + \sum_{i=1}^n c_{1i}x_i) + h_1(S_1 \cdot \mathbf{x} + b_1), \\ \dot{x}_2(t) = f_2(c_{20} + \sum_{i=1}^n c_{2i}x_i) + h_2(S_2 \cdot \mathbf{x} + b_2), \\ \dots \\ \dot{x}_n(t) = f_n(c_{n0} + \sum_{i=1}^n c_{ni}x_i) + h_n(S_n \cdot \mathbf{x} + b_n). \end{cases} \quad (30)$$

Here  $f_i(u), h_i(v)$  are power functions, the functions  $h_i(v) = \text{piecewise}(v < 0, -(-v)^{\gamma_i}, v^{\gamma_i})$  are odd and  $\gamma_i > 1$ ;  $i = 1, \dots, n$ . (System (25) follows from system (30) if we put  $\deg f_i(u) = 1$ ,  $D_0 = I$ , and  $b_i = 0$  in the last system.)

It is known that the basis of search algorithms is the Jacobi matrix  $Jac$ . Let's form this matrix for system (30) using representation (29). In this case, the vector of parameters  $\mathbf{z}$  will be constructed as follows :

$$\mathbf{z} = (c_{10}, \dots, c_{1n}, \dots, c_{n0}, \dots, c_{nn}, d_{11}, \dots, d_{1n}, d_{22}, \dots, d_{2n}, \dots, d_{nn}, b_1, \dots, b_n)^T \in \mathbb{R}^{C_n^2 + n(n+2)}.$$

Let  $m = 1$ . We introduce the following auxiliary matrices that will be needed to construct the Jacobi matrix for system (30):

$$J_1 = \text{Diag}_1 \cdot \begin{pmatrix} (1, x_1, \dots, x_n) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & (1, x_1, \dots, x_n) \end{pmatrix} \in \mathbb{R}^{n \times (n+1)n},$$

$$J_2 = \text{Diag}_2 \cdot (-\mathbf{x}, P_1, P_2, \dots, P_{n-1}, I_n) \in \mathbb{R}^{n \times (C_n^2 + n + 1)},$$

where

$$\text{Diag}_1 = \begin{pmatrix} f'_1(c_{10} + \sum_{i=1}^n c_{1i}x_i) & 0 & \dots & 0 \\ 0 & f'_2(c_{20} + \sum_{i=1}^n c_{2i}x_i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & f'_n(c_{n0} + \sum_{i=1}^n c_{ni}x_i) \end{pmatrix} \in \mathbb{R}^{n \times n},$$

$$Diag_2 = \begin{pmatrix} h'_1(S_1 \cdot \mathbf{x} + b_1) & 0 & \dots & 0 \\ 0 & h'_2(S_2 \cdot \mathbf{x} + b_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h'_n(S_n \cdot \mathbf{x} + b_n) \end{pmatrix} \in \mathbb{R}^{n \times n},$$

and

$$P_i = \begin{pmatrix} \overline{\overline{O_{(i-1) \times (n-i)}}} \\ x_{i+1}, \dots, x_n \\ \overline{\overline{-x_i I_{n-i}}} \end{pmatrix} \in \mathbb{R}^{n \times (n-i)};$$

$O_{(i-1) \times (n-i)} \in \mathbb{R}^{(i-1) \times (n-i)}$  is the zero matrix;  $I_{n-i} \in \mathbb{R}^{(n-i) \times (n-i)}$  and  $I_n \in \mathbb{R}^{n \times n}$  are the identity matrices;  $i = 1, \dots, n-1$ .

Now, we briefly describe the algorithm for adjusting the weights for system (30).

1. Introducing objects that do not change in the iterative process.

1.1. Fix a learning selections (they are formed from time series):

$$\begin{aligned} g_{10}, g_{11}, g_{12}, \dots, g_{1m} &\equiv \mathbf{g}_1^T \\ g_{20}, g_{21}, g_{22}, \dots, g_{2m} &\equiv \mathbf{g}_2^T \\ \dots &\dots \\ g_{n0}, g_{n1}, g_{n2}, \dots, g_{nm} &\equiv \mathbf{g}_n^T. \end{aligned}$$

1.2. Fix triplets of real numbers  $\alpha_i \geq 1$ ,  $\beta_i \geq 1$  and  $\gamma_i > \alpha_i, \gamma_i > \beta_i$ ;  $i = 1, \dots, n$ .

1.3. Introduce the matrix

$$G = \begin{pmatrix} 1 & g_{11} & g_{21} & \dots & g_{n1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & g_{1m} & g_{2m} & \dots & g_{nm} \end{pmatrix} = (\mathbf{1}, \mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n) \in \mathbb{R}^{m \times (n+1)}.$$

1.4. Construct the columns of numerical derivatives:

$$D_i = \frac{1}{\Delta t} \begin{pmatrix} g_{i1} - g_{i0} \\ g_{i2} - g_{i1} \\ \vdots \\ g_{i,m} - g_{i,m-1} \end{pmatrix} \in \mathbb{R}^m; i = 1, \dots, n.$$

2.  $k = 0$ . (The beginning of a global cycle.) Introduce the nonzero vector of initial approximations:

$$Y_k := (c_{10}, \dots, c_{1n}, \dots, c_{n0}, \dots, c_{nn}, d_{11}, \dots, d_{1n}, d_{22}, \dots, d_{2n}, \dots, d_{nn}, b_1, \dots, b_n)^T \in \mathbb{R}^{C_n^2 + n(n+2)};$$

$$(\text{for example}) Y_0 := (1, \dots, 1, d_{11} = 0, \dots, d_{nn} = 0, 0, \dots, 0) \in \mathbb{R}^{C_n^2 + n(n+2)},$$

the previous error  $T_{k-1} = T_{-1} := 10^{10}$ , and the number  $\epsilon := 0.001$ .

2.1. Construct the matrices

$$F_i = \begin{pmatrix} f_{11}^{(i)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & f_{mm}^{(i)} \end{pmatrix} \in \mathbb{R}^{m \times m},$$

where  $f_{jj}^{(i)} = \text{piecewise}(c_{i0} + \sum_{s=1}^n c_{is} g_{sj} < 0, (-1)^{k_i} \beta_i (-c_{i0} - \sum_{s=1}^n c_{is} g_{sj})^{\beta_i - 1}, \alpha_i (c_{i0} + \sum_{s=1}^n c_{is} g_{sj})^{\alpha_i - 1})$ ,

and

$$H_i = \begin{pmatrix} h_{11}^{(i)} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & h_{mm}^{(i)} \end{pmatrix} \in \mathbb{R}^{m \times m},$$

where  $h_{jj}^{(i)} = \text{piecewise}(\sum_{l=1}^n S_{il}g_{lj} + b_i < 0, \gamma_i \cdot (-\sum_{l=1}^n S_{il}g_{lj} - b_i)^{\gamma_i-1}, \gamma_i \cdot (\sum_{l=1}^n S_{il}g_{lj} + b_i)^{\gamma_i-1})$ ;  
 $k_i = 0(k_i = 1)$  if  $f_i(\dots)$  is odd (even);  $i = 1, \dots, n$ ;  $j = 0, 1, \dots, m-1$ .

2.2. Now we will use auxiliary matrices  $J_1$  and  $J_2$  obtained for  $m = 1$  in order to construct the Jacobi matrix of system (30) in case  $m > 1$ . To do this, we need to replace scalar variables  $1, x_1, \dots, x_n$ , respectively, by vectors  $\mathbf{1}, \mathbf{g}_1, \dots, \mathbf{g}_n \in \mathbb{R}^m$ . Then, we have

$$Jac = (J_1(\mathbf{1}, \mathbf{g}_1, \dots, \mathbf{g}_n) | J_2(\mathbf{1}, \mathbf{g}_1, \dots, \mathbf{g}_n)) = W_k$$

$$= \left( \begin{array}{ccccc|cccc} F_1 \cdot G & 0 & 0 & \dots & 0 & -H_1 \cdot \mathbf{g}_1 & H_1 \cdot \mathbf{g}_2 & H_1 \cdot \mathbf{g}_3 & \dots & H_1 \cdot \mathbf{g}_n \\ 0 & F_2 \cdot G & 0 & \dots & 0 & 0 & -H_2 \cdot \mathbf{g}_1 & 0 & \dots & 0 \\ 0 & 0 & F_3 \cdot G & \dots & 0 & 0 & 0 & -H_3 \cdot \mathbf{g}_1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & F_n \cdot G & 0 & 0 & 0 & \dots & -H_n \cdot \mathbf{g}_1 \end{array} \right)$$

$$\left( \begin{array}{ccc|ccc} 0 & 0 & \dots & 0 & H_1 \cdot \mathbf{1} & 0 & 0 & \dots & 0 \\ -H_2 \cdot \mathbf{g}_2 & H_2 \cdot \mathbf{g}_3 & \dots & 0 & 0 & H_2 \cdot \mathbf{1} & 0 & \dots & 0 \\ 0 & -H_3 \cdot \mathbf{g}_2 & \dots & 0 & 0 & 0 & H_3 \cdot \mathbf{1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -H_n \cdot \mathbf{g}_2 & -H_n \cdot \mathbf{g}_n & 0 & 0 & \dots & H_n \cdot \mathbf{1} \end{array} \right).$$

Here  $W_k$  is  $(nm \times (C_n^2 + n(n+2)))$ -matrix and  $\mathbf{1} = (1, 1, \dots, 1)^T \in \mathbb{R}^m$ .

2.3. Check condition  $\text{rank } W_k = C_n^2 + n(n+2)$ . If  $\text{rank } W_k < C_n^2 + n(n+2)$  then fix some number  $\mu \in (0, 1]$  else  $\mu := 0$ .

2.4. Calculation of the right side of system (30):

$$R_{ij} = \text{piecewise}(c_{i0} + \sum_{s=1}^n c_{is}g_{sj} < 0, (-1)^{k_i}(-c_{i0} - \sum_{s=1}^n c_{is}g_{sj})^{\beta_i}, (c_{i0} + \sum_{s=1}^n c_{is}g_{sj})^{\alpha_i})$$

$$+ \text{piecewise}(\sum_{l=1}^n S_{il}g_{lj} + b_i < 0, -(-\sum_{l=1}^n S_{il}g_{lj} - b_i)^{\gamma_i}, (\sum_{l=1}^n S_{il}g_{lj} + b_i)^{\gamma_i});$$

where  $k_i = 1(k_i = 0)$  if  $f_i(\dots)$  is odd (even);  $i = 1, \dots, n$ ;  $j = 0, \dots, m-1$ .

2.5. Form vectors:

$$R_i = \begin{pmatrix} R_{i0} \\ R_{i1} \\ \vdots \\ R_{i,m-1} \end{pmatrix} \in \mathbb{R}^m; i = 1, \dots, n.$$

3. Compute the vector of errors

$$E_k = (e_1, \dots, e_m, e_{m+1}, \dots, e_{2m}, e_{2m+1}, \dots, e_{3m}, \dots, e_{n-1,m}, \dots, e_{nm})^T := \begin{pmatrix} D_1 - R_1 \\ D_2 - R_2 \\ \vdots \\ D_n - R_n \end{pmatrix} \in \mathbb{R}^{nm}$$

and the total computation error at iteration step  $k$ :

$$T_k := \sum_{j=1}^{nm} e_j^2.$$

4. If

$$T_{k-1} > T_k + \epsilon,$$

then compute the vector

$$Y_{k+1}^T := Y_k^T + (W_k^T \cdot W_k + \mu I)^{-1} \cdot W_k^T \cdot E_k \in \mathbb{R}^{C_n^2 + n(n+2)}$$

else go to item 6. (Here  $I \in \mathbb{R}^{(C_n^2 + n(n+2)) \times (C_n^2 + n(n+2))}$  is the identity matrix.)

5. Put  $k := k + 1$ ,

$$(c_{10}, \dots, c_{1n}, \dots, c_{n0}, \dots, c_{nn}, d_{11}, \dots, d_{1n}, d_{22}, \dots, d_{2n}, \dots, d_{nn}, b_1, \dots, b_n)^T := Y_k$$

and go to item 2.1.

6. Print a graph of error changes  $T_0, T_1, \dots, T_{k-1}$ . (End of the cycle of  $k$  for finding the local minimum.)

7. Print the vector  $Y_{k-1} \in \mathbb{R}^{C_n^2 + n(n+2)}$  and stop the algorithm.

8. Solve system (30). (If the solutions of system (30) diverge, then in the initial conditions replace the value  $d_{11} = 0$  with a sufficiently small value  $|d_{11}|$  and go to item 2.)

**Comment 1.** A chaotic dynamic process modeling should begin with building a limit cycle. As Theorems 4 and 5 show, for a periodic trajectory to appear in system (30), it is sufficient that  $b_1 = \dots = b_n = 0$ . Therefore, Algorithm 2 can be started by putting  $b_1 = \dots = b_n = 0$ .

**Comment 2.** If the constructed model does not satisfy the necessary requirements, then it is possible to increase the number of neurons in the equation in accordance with the architecture of system (6).

Let  $n = 3$ . In order to simulate a chaotic process, it is necessary to introduce perturbations into the right-hand sides of system (30). These perturbations are generated by adding one neuron each, two neurons each, three neurons each (nonlinearity), etc. to the right parts of system (30). As a result, we can get the following system:

$$\begin{cases} \dot{x}_1(t) = f_1(a_{10} + \sum_{j=1}^3 a_{1j}x_j) + f_1(b_{10} + \sum_{j=1}^3 b_{1j}x_j) + f_1(c_{10} + \sum_{j=1}^3 c_{1j}x_j) \\ \quad + h_1(-d_{11}x_1 + d_{12}x_2 + d_{13}x_3), \\ \dot{x}_2(t) = f_2(a_{20} + \sum_{j=1}^3 a_{2j}x_j) + f_2(b_{20} + \sum_{j=1}^3 b_{2j}x_j) + f_2(c_{20} + \sum_{j=1}^3 c_{2j}x_j) \\ \quad + h_2(-d_{12}x_1 - d_{22}x_2 + d_{23}x_3), \\ \dot{x}_3(t) = f_3(a_{30} + \sum_{j=1}^3 a_{3j}x_j) + f_3(b_{30} + \sum_{j=1}^3 b_{3j}x_j) + f_3(c_{30} + \sum_{j=1}^3 c_{3j}x_j) \\ \quad + h_3(-d_{13}x_1 - d_{23}x_2 - d_{33}x_3). \end{cases} \quad (31)$$

Here

$f_i(u_{ij}) = \text{piecewise}(u_{ij} < 0, (-1)^{k_i}(-u_{ij})^{\beta_i}, u_{ij}^{\alpha_i})$ ;  $h_i(v_i) = \text{piecewise}(v_i < 0, -(-v_i)^{\gamma_i}, v_i^{\gamma_i})$  are real piecewise continuous power functions of one variable;  $k_i = 1$  ( $k_i = 0$ ) if  $f_i(\dots)$  is odd (even);  $i, j = 1, \dots, 3$ ;

$a_{ij}, b_{ij}, c_{ij} \in \mathbb{R}$ ;  $\alpha_i \geq 1, \beta_i \geq 1, \gamma_i > \alpha_i \geq 1, \gamma_i > \beta_i \geq 1$ ;  $i = 1, \dots, 3$ ;  $j = 0, \dots, 3$ ;

$d_{11}, d_{12}, d_{13}, d_{22}, d_{23}, d_{33} \in \mathbb{R}$ .

The addition of neurons can be stopped when a suitable approximation accuracy will be obtained.

## 6 Modeling

### 6.1 Applications of Algorithm 1

To test the operability of Algorithm 1, the following system [36] will be used:

$$\begin{cases} \dot{x}(t) = y(t), \\ \dot{y}(t) = -x(t) + z(t), \\ \dot{z}(t) = 2 - 0.8x^2(t) + z^2(t). \end{cases} \quad (32)$$

For system (32), we solve the Cauchy problem with initial data  $x_0 = 0, y_0 = 2.3, z_0 = 0$ . From the obtained continuous  $x(t), y(t), z(t)$ , we form time series with step  $\Delta t = 1$ . To denote these series, we use the notations  $x_1(t) = x(t), x_2(t) = y(t), x_3(t) = z(t)$ ;  $t = 0, 1, 2, \dots, 300$ . Thus, we have  $m = 300$ .

We assume that  $B = 0$ . Then the use of Algorithm 1 with odd activation functions leads to such coefficients of system (28):

$$c_0 = \begin{pmatrix} 0.00020 \\ -0.01967 \\ 1.96685 \end{pmatrix}, C = \begin{pmatrix} 0.00955 & 0.99992 & -0.00951 \\ -0.95541 & 0.00848 & 0.95059 \\ -4.44895, & 0.15144 & 4.93128 \end{pmatrix},$$

$$D = \begin{pmatrix} 0.00000 & 0.00000 & 0.00000 \\ -0.03484 & 0.00131 & 0.04126 \\ 3.48409 & -0.13084 & -4.12689 \end{pmatrix}; \alpha = 0.66666, \beta = 1.55555. \quad (33)$$

Now, in order to simulate processes in system (32), we assume  $D = 0$  and again use Algorithm 1, but with even activation functions. Then we get such coefficients of system (28):

$$c_0 = \begin{pmatrix} 0.00021 \\ -0.02067 \\ 2.06700 \end{pmatrix}, C = \begin{pmatrix} 0.01000 & 0.99990 & -0.01000 \\ -0.99964 & 0.01018 & 0.99956 \\ -0.02591 & -0.01766 & 0.03368 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.00000 & 0.00000 & 0.00000 \\ 0.00735 & 0.00000 & 0.00926 \\ -0.73552 & -0.00296 & 0.92592 \end{pmatrix}; \gamma = 2.05263, \delta = 2.05263 \quad (34)$$

and

$$c_0 = \begin{pmatrix} 0.00006 \\ -0.00619 \\ 0.61919 \end{pmatrix}, C = \begin{pmatrix} 0.00989 & 0.99993 & -0.00021 \\ -0.98893 & 0.00679 & 0.98823 \\ -1.09730 & 0.32107 & 1.16668 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.00000 & 0.00000 & 0.00000 \\ 0.02060 & -0.00296 & -0.02438 \\ -2.06034 & 0.29609 & 2.43823 \end{pmatrix}; \gamma = 0.66666, \delta = 1.55555. \quad (35)$$

Phase portraits of system (32) and (33) are presented in Fig.6:

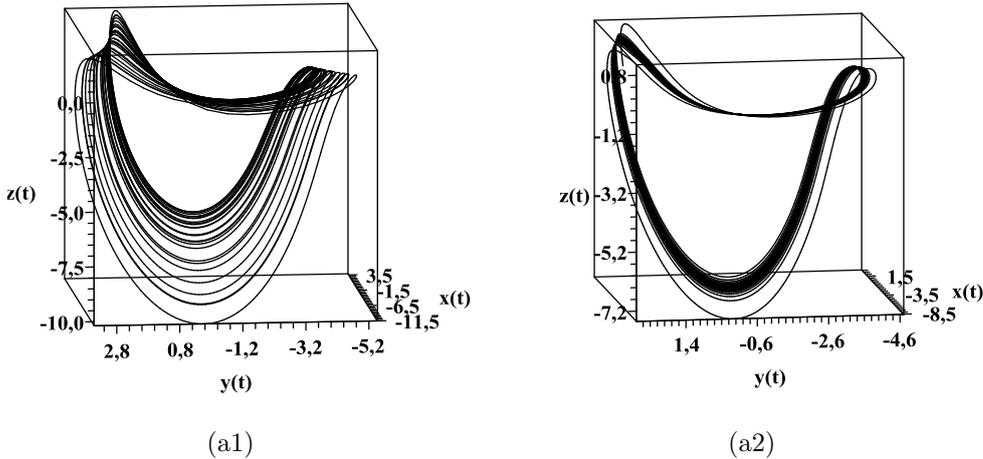


Fig. 6. Phase portraits of system (32)(a1) and (33)(a2)

Phase portraits of system (34) and (35) are presented in Fig.7:

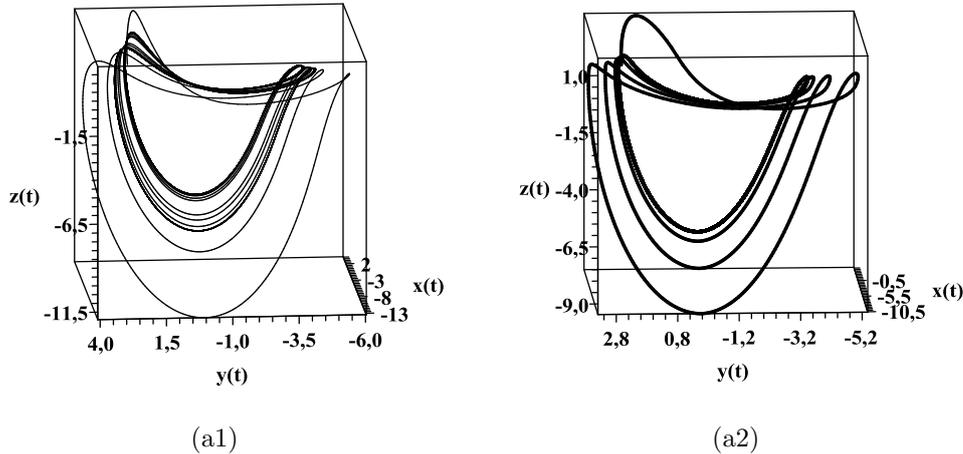


Fig. 7. Phase portraits of system (34)(a1) and (35)(a2)

Thus, we can assert that the use of even activation functions leads to more accurate modeling (especially, system (34)) than use of odd activation functions.

## 6.2 Applications of Algorithm 2

Consider the following special case of system (26):

$$\begin{cases} \dot{x}_1(t) = a_{10} + \sum_{j=1}^3 a_{1j}x_j + h(-d_{11}x_1 + d_{12}x_2 + d_{13}x_3) + \nu h(-l_1d_{11}x_1 + l_2d_{12}x_2 + l_3d_{13}x_3) \\ \dot{x}_2(t) = a_{20} + \sum_{j=1}^3 a_{2j}x_j + h(-d_{12}x_1 - d_{22}x_2 + d_{23}x_3) + \nu h(-l_1d_{12}x_1 - l_2d_{22}x_2 + l_3d_{23}x_3) \\ \dot{x}_3(t) = a_{30} + \sum_{j=1}^3 a_{3j}x_j + h(-d_{13}x_1 - d_{23}x_2 - d_{33}x_3) + \nu h(-l_1d_{13}x_1 - l_2d_{23}x_2 - l_3d_{33}x_3). \end{cases} \quad (36)$$

Here if  $\nu = 0$ , then there is a limit cycle; if  $\nu = -1$ , then there is a homoclinic orbit.

In addition, if  $\nu = 0$ , then  $h(v_i) = \text{piecewise}(v_i < 0, -(-v_i)^\gamma, v_i^\gamma)$  is an odd function; if  $\nu = -1$ , then  $h(v_i) = \text{piecewise}(v_i < 0, (-v_i)^\gamma, v_i^\gamma)$  is an even function. Here  $h(v_i)$  are real piecewise continuous power functions of one variable;  $\gamma > 1$ ;  $i = 1, \dots, 3$ ;  $d_{11}, d_{12}, d_{13}, d_{23} \in \mathbb{R}$ , where  $d_{11} \geq 0$ .

A triple of integers  $(l_1, l_2, l_3)$  accepts only one set of  $2^n = 8$  ( $n = 3$ ) possible:

$$(-1, -1, -1); (-1, -1, 1); (-1, 1, -1); (1, -1, -1); (1, 1, -1); (1, -1, 1); (-1, 1, 1); (1, 1, 1).$$

The triple  $(l_1, l_2, l_3)$  indicates one of the possible types of homoclinic orbits that can be realized in system (36).

In this subsection, Algorithm 2 will only be used to simulate systems with possible homoclinic orbits ( $\nu = -1$ ). Therefore, instead of the nonlinear part  $\mathbf{h}(S\mathbf{x} + \mathbf{b})$  in the system (30), the nonlinear part  $\mathbf{h}(SG_1\mathbf{x} + \mathbf{b}) - \mathbf{h}(SG_2\mathbf{x} + \mathbf{b})$  should be used.

In order to check the performance of Algorithm 2, several well-known systems were used. From the obtained continuous  $x(t), y(t), z(t)$ , we form time series with step  $\Delta t = 1$ , where  $t = 0, 1, 2, \dots, 200$ . Thus, we have  $m = 200$ .

1. Lorenz system (27) at  $x_0 = 1, y_0 = 1, z_0 = 8$ . For the Lorentz approximation, we need to take  $(l_1, l_2, l_3) = (-1, 1, 1)$ .

1.1. For the classical Lorenz attractor, we have for system (27)  $\sigma = 10, a = 27, b = 2.7$  and  $h(v_i) = \text{piecewise}(v_i < 0, (-v_i)^2, v_i^2)$ ;  $i = 1, \dots, 3$ . A graph of this system is shown in Figure 8 (a1).

1.2. We form time series for system (27) and assign  $h(v_i) = \text{piecewise}(v_i < 0, (-v_i)^{1.5}, v_i^{1.5})$ ;  $i = 1, \dots, 3$ . Then Algorithm 2 applied to system (36) leads to the following results:

$$A_0 = \begin{pmatrix} 0.71269 \\ 0.69584 \\ 3.92098 \end{pmatrix}, A = \begin{pmatrix} -10.01641 & 9.94437 & -0.00899 \\ 54.83057 & -2.42695 & -0.02138 \\ -0.44238 & 0.27646 & -3.32041 \end{pmatrix},$$

$d_{11} = 0.00607, d_{12} = 2.63824, d_{13} = 1.08829, d_{22} = d_{11}, d_{23} = 1.68198, d_{33} = d_{11}$ . The graph of this system is shown in Figure 8 (a2).

1.3. By analogy with (31), add the following terms to the corresponding right-hand sides of each of equations (36):

$$h_1(a_{10} + \sum_{j=1}^3 a_{1j}x_j), h_1(a_{20} + \sum_{j=1}^3 a_{2j}x_j), h_1(a_{30} + \sum_{j=1}^3 a_{3j}x_j), \quad (37)$$

where  $h_1(v_i) = \text{piecewise}(v_i < 0, -(-v_i)^{1.1}, v_i^{1.1})$ ;  $i = 1, \dots, 3$ . In this case, Algorithm 2 leads to the following results:

$$A_0 = \begin{pmatrix} -0.09276 \\ -0.37044 \\ -6.40899 \end{pmatrix}, A = \begin{pmatrix} -4.25589 & 4.15270 & 0.00343 \\ 27.12403 & -2.14111 & 0.00926 \\ 0.00761 & -0.00245 & -1.53054 \end{pmatrix},$$

$d_{11} = 0.00081, d_{12} = 2.75407, d_{13} = 0.83019, d_{22} = d_{11}, d_{23} = 2.63333, d_{33} = d_{11}$ . The graph of this system is shown in Figure 8 (a3).

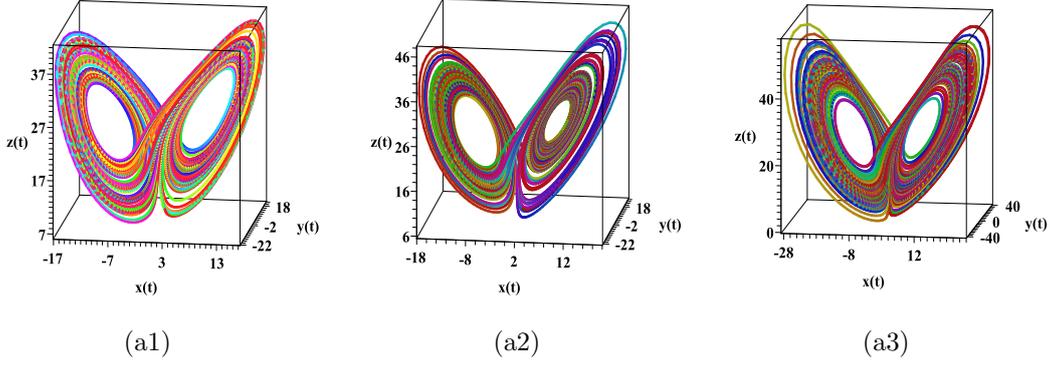


Fig. 8. Phase portraits of system (27)(a1), (36)(a2), and [(36)+(37)](a3)

Thus, in the case 1.3 (see Fig. 8(a3)), there was no improvement in the quality of approximation.

2. Consider the following system without equilibrium points [38]:

$$\begin{cases} \dot{x}(t) = 10y(t), \\ \dot{y}(t) = -x(t) + 3z(t) + x(t)z(t), \\ \dot{z}(t) = 1 + x(t) - z(t) - x(t)y(t) + 0.25x(t)z(t), \end{cases} \quad (38)$$

where  $x_0 = y_0 = z_0 = 1$ . The graph of this system is shown in Figure 9 (a1).

2.1. From the obtained continuous  $x(t), y(t), z(t)$  of system (38), we form time series with step  $\Delta t = 1$ , where  $t = 0, 1, 2, \dots, 5000$ . Thus, we have  $m = 5000$ . We assign  $h(v_i) = \text{piecewise}(v_i < 0, (-v_i)^2, v_i^2)$ ;  $i = 1, \dots, 3$ . Then Algorithm 2 applied to system (36) leads to the following results:

$$A_0 = \begin{pmatrix} 0.00017 \\ -0.05053 \\ 2.22924 \end{pmatrix}, A = \begin{pmatrix} -0.01871 & 9.98680 & 0.02484 \\ -0.99742 & -0.36465 & 3.00322 \\ 0.99120 & 0.01164 & -0.52040 \end{pmatrix},$$

$d_{11} = -0.10007, d_{12} = -0.00573, d_{13} = -0.00573, d_{23} = 43.63002, d_{22} = -0.11820, d_{33} = 10.81835$ . The graph of this system is shown in Figure 9 (a2).

2.2. Now we will assign  $h(v_i) = \text{piecewise}(v_i < 0, (-v_i)^{1.7}, v_i^{1.7})$ ;  $i = 1, \dots, 3$ . In this case, Algorithm 2 leads to the following results:

$$A_0 = \begin{pmatrix} 0.05826 \\ 1.49581 \\ 13.51462 \end{pmatrix}, A = \begin{pmatrix} -0.22723 & 9.75617 & 0.24837 \\ -1.93969 & -3.59190 & 2.94801 \\ 0.92691 & 0.18019 & -6.16120 \end{pmatrix},$$

$d_{11} = 0.01850, d_{12} = 1.30116, d_{13} = 1.30705, d_{23} = -0.71111, d_{22} = 0.01463, d_{33} = -0.16564$ . The graph of this system is shown in Figure 9 (a3).

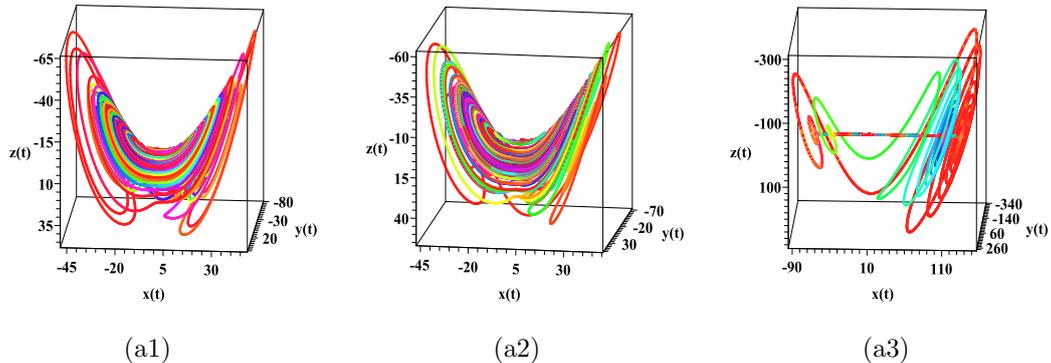


Fig. 9. Phase portraits of system (38)(a1), system (36) at  $\deg h(v) = 2$  (a2), and system (36) at  $\deg h(v) = 1.7$  (a3)

Thus, the quality of the approximation deteriorates if  $\deg h(v) < 2$ .

## 7 Conclusion

As can be seen from the above examples, Algorithms 1 and 2 correctly indicate the tendency of the behavior of dynamic processes described by certain time series. Now, based on the obtained models, we can describe them more precisely: for example, if approximate the derivatives by finite differences of the second or third order. Various combinations of odd and even power functions can also be used in the simulation. (Corollary of Theorem 2 guarantees the achievement of the required approximation accuracy only for odd activation functions. Nevertheless, the introduction of even activation functions sometimes makes it possible to achieve the required accuracy with a smaller number of terms.) Moreover, adding new power-law nonlinearities to the model equations also improves the quality of the approximation.

The main results of this article are as follows:

1. Due to the concept of odd activation function introduced in [25], a simplified version of the classical approximation Theorem 2 (this is Corollary of Theorem 2) was obtained. As a result, representation (1), where generally speaking  $\xi_j \neq 1$ , was replaced by representation (5) with  $\xi_j = 1$ ;  $j = 1, \dots, m$ .

2. Conditions for the existence of periodic solutions in neural systems of ordinary differential equations with power nonlinearities were found.

3. Theorems 3 and 6 represent a constructive approach to solving the problem of the existence of chaos in dynamical systems with power-law nonlinearities. (This approach is more general than the approaches indicated in [32] – [34].)

4. Algorithms for approximating time series using antisymmetric neural ODE, the architecture of which allows the possibility of modeling limit cycles or homoclinic orbits, have been developed (Algorithm 2 and its modifications associated with modeling systems (26) and (36)).

5. The developed methods for adjusting the coefficients of antisymmetric neural ODEs for solving the approximation problem make it possible to use the obtained coefficients as initial weights for deep learning of recurrent neural networks for which the mentioned neural ODEs were designed.

Indeed, let us replace the reconstructed neural ODE  $\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}) \in \mathbb{R}^n$  with its difference analogue  $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{f}(\mathbf{x}_k) \cdot \Delta t \in \mathbb{R}^n$ , where  $k = 0, 1, \dots, K < \infty$ , and  $K$  is the number of layers. Then, by adjusting the value of step  $\Delta t > 0$ , we can get RNN ( $\Delta t \rightarrow 0$ ), which adequately simulates the process under study for some  $(\Delta t)^*$  and  $K^*$ .

### Conflict of Interests

The authors declare that they have no conflict of interests.

### Data Availability Statement

All data generated or analysed during this study are included in this published article (and its supplementary information files).

## References

- [1] Haykin, S.: Neural Networks. A Comprehensive Foundation, Second Edition. Pearson Education, Prentice Hall (2005).
- [2] Izhikevich, E. M.: Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting. The MIT Press Cambridge, Massachusetts, London, England (2007).
- [3] Cessac, B.: A view of neural networks as dynamical systems. – arXiv:0901.2203v2[nlin.AO], 1 – 62 (2009).
- [4] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. The MIT Press Cambridge, Massachusetts, London, England (2017).
- [5] Krotov, D., Hopfield, J.: Large associative memory problem in neurobiology and machine learning. – arXiv:2008.06996v1[q-bio. NC], 1 – 8 (2020).
- [6] Li, Q.: Dynamical Systems and Machine Learning. Summer School, Peking University (2020).
- [7] Yang, C- L., Suh, C. S.: A general framework for dynamic complex networks. Journal of Vibration Testing and System Dynamics **5** (1), 87 – 111 (2021).

- [8] Sonoda, S., Murata, N.: Neural network with unbounded activation functions is universal approximator. *Applied and Computational Harmonic Analysis* **43**, 233 – 268 (2017).
- [9] Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. – arXiv:1806.07366v5[cs.LG], 1 – 18 (2019).
- [10] Chen, R.T.Q., Duvenaud, D.: Neural networks with cheap differential operators. – arXiv:1912.03579v1[cs.LG], 1 – 11 (2019).
- [11] Kelly, J., Bettencourt, J., Johnson, M., J., Duvenaud, D.: Learning differential equations that are easy to solve. – arXiv:2007.04504v2[cs.LG], 1 – 18 (2020).
- [12] Jia, J., Benson, A.R.: Neural jump stochastic differential equations. – arXiv:1905.10403v3[cs.LG], 1 – 14 (2020).
- [13] Chang, B., Chen, M., Haber, E., Chi, E.D. : Antisymmetric RNN: A dynamical system view on recurrent neural networks. In conference ICLR, May 6 - 9, New Orleans, Louisiana, USA, 1 – 15 (2019).
- [14] Lorraine, J., Vicol, P., Duvenaud, D.: Optimizing millions of hyperparameters by implicit differentiation. – arXiv:1911.02590v1[cs.LG], 1 – 18 (2019).
- [15] Massaroli, S., Poli, M., Park, J., Yamashita, A., Asama, H.: Dissecting neural ODEs. – arXiv:2002.08071v3[cs.LG], 1 – 23 (2020).
- [16] Queiruga, A.F., Erichson, N.B., Taylor, D., Mahoney, M.W.: Continuous-in-depth neural networks. – arXiv:2008.02389v1[cs.LG], 1 – 29 (2020).
- [17] Belozyorov, V. Ye., Dantsev, D. V., Volkova, S. A.: On equivalence of linear control systems and its usage to verification of the adequacy of different models for a real dynamic process. *Journal of optimization, differential equations and their applications (JODEA)* **28**(1), 43 – 97 (2020).
- [18] Magnitskii, N.A.: Universal theory of dynamical chaos in nonlinear dissipative systems of differential equations. *Commun. Nonlinear Sci. Numer. Simul.* **13**, 416 – 433 (2008).
- [19] Magnitskii, N.A.: Universality of transition to chaos in all kinds of nonlinear differential equations. In: *Nonlinearity, Bifurcation and Chaos – Theory and Applications*, Chapter 6. Intech, 133–174 (2012).
- [20] Magnitskii, N. A.: Bifurcation theory of dynamical chaos. *Chaos Theory*, Chapter 11. Intech, 197 – 215 (2018).
- [21] Belozyorov, V.Ye.: Universal approach to the problem of emergence of chaos in autonomous dynamical systems. *Nonlinear Dynamics* **95** (1), 579 – 595 (2019).
- [22] Tatar, N.-E.: Hopfield neural networks with unbounded monotone activation functions. Hindawi Publishing Corporation. *Advances in Artificial Neural Systems* **2012**, ID 571358-1 – 5 (2012).
- [23] Richards, S.M., Berkenkamp, F., Krause, A.: The Lyapunov neural network: adaptive stability certification for safe learning of dynamical systems. – arXiv:1808.00924v2[cs.SY], 1 – 11 (2018).
- [24] Haber, E., Ruthottom, L.: Stable architectures for deep neural networks. – arXiv:1705.03341v3[cs.LG], 1 – 23 (2019).
- [25] Belozyorov, V. Ye., Dantsev, D. V.: Stability of neural ordinary differential equations with power nonlinearities. *Journal of optimization, differential equations and their applications (JODEA)* **28** (2), 21 – 46 (2020).
- [26] Belozyorov, V.Ye., Kosariev, Ye. M., Pulin, M. M., Sychenko, V. G., Zaytsev, V. G.: A new mathematical model of dynamic process in direct current traction power supply system. *Journal of optimization, differential equations and their applications (JODEA)* **27** (1), 21 – 55 (2019).

- [27] Wang, Z., Liang, J., Liu, Y.: Mathematical problems for complex networks. Hindawi Publishing Corporation Mathematical Problems in Engineering **2012**, ID 934680-1 – 5 (2012).
- [28] Gorban, A. N., Dunin-Barkovsky, V. L., Kirdin, A. N., Mirkes, E.M., Novokhodko, A. Yu., Rossiev, D. A., Terekhov, S. A., Senashova, M. Yu., Tsaregorodtsevm, V. G.: Neuroinformatics. Nauka, Novosibirsk (1998) (in rus).
- [29] Zhang, H., Gao, X., Unterman, J., Arodz, T. : Approximation capabilities of neural ordinary differential equations. – arXiv: 1907.12998v1[cs.LG], 1 – 11 (2019).
- [30] Li, Q., Lin, T., Shen, Z.: Deep learning via dynamical systems: an approximation perspective. – arXiv:1912.10382v1[cs.LG], 1 – 30 (2019).
- [31] Khalil, H. K.: Nonlinear Systems, 2nd edition. Prentice Hall, Upper Saddle River (1996).
- [32] Belozyorov, V. Ye.: A novel search method of chaotic autonomous quadratic dynamical systems without equilibrium points. Nonlinear Dyn. **86**, 835 – 860 (2016).
- [33] Belozyorov, V. Ye., Volkova, S. A.: Role of logistic and Ricker’s maps in appearance of chaos in autonomous quadratic dynamical systems. Nonlinear Dyn. **83**, 719 – 729 (2016).
- [34] Belozyorov, V. Ye.: On novel conditions of chaotic attractors existence in autonomous polynomial dynamical systems. Nonlinear Dyn. **91**, 2435 – 2452 (2018).
- [35] Belozyorov, V.Ye.: On existence of homoclinic orbits for some types of autonomous quadratic systems of differential equations. Applied Mathematics and Computation, **217**(9), 4582 – 4595 (2011).
- [36] Jafary, S., Sprott, J. C., Mohammad Reza Hashemi Golpayegani, S. : Elementary quadratic chaotic flows no equilibria. Physics Letters A **377**, 699 – 702 (2013).
- [37] Wang, X., Chen, G.: A gallery Lorenz-like and Chen-like attractors. Int. J. Bifurc. Chaos **23**(4), ID 1330011-1 – 20 (2013).
- [38] Dantsev, D.: A novel type of chaotic attractor for quadratic systems without equilibriums. Int. J. Bifurc. Chaos **28**(1), ID 1850001-1 – 7 (2018).