Applied nonlinear dynamics Habilitation thesis

Lenka Přibylová

October 6, 2022

2

Department of Mathematics and Statistics,
Faculty of Science,
Masaryk University,
61137 Kotlářská 2, Brno, Czech Republic

email: pribylova@math.muni.cz

Commentary:

The habilitation thesis deals with applications of nonlinear dynamics and related phenomena in various fields of science. It presents both the basic types of bifurcations of equilibria and cycles that are common in application due to their wide range of existence in nature and advanced types of local and non-local bifurcations of equilibria, cycles and tori that are more complex and may lead even to an unpredictable dynamics. Typical nonlinear phenomena are the disappearance of equilibria and cycle, bistability in the broader sense of more attractors coexistence, hysteresis phenomena, cycle and tori births, and breaks that correspond to system destabilization mechanisms that may lead to abrupt regime shifts analogously to hysteresis. Moreover, we discuss the synchronization phenomenon as well as chaotic dynamics and routes to chaos.

The attached or linked published articles show applications of bifurcation theory on specific models to explain dynamical phenomena and their consequences. These original articles have my full or significant authorship. For the paper [22] published with co-author Robert Mařík as a corresponding author, I made the bifurcation analysis related to the generalized Hopf bifurcation. For other articles where I am not the corresponding and lead author, these are articles published in collaboration with my students where I am a co-author and supervisor. The last published paper [11] in Nonlinear Dynamics appeared online in September 2022 and includes results made in collaboration with André Botha from the University of South Africa, a specialist on Josephson junctions. The idea of embedding near the Hopf-Hopf bifurcation is mine. We wrote the text equally.

The approach to the description of nonlinear phenomena using bifurcation theory is presented in five chapters, which end with a summary of the author's contributions to published scientific papers. Chapter 6, Collection of author's published articles concerning nonlinear phenomena, describes and includes all papers that are relevant to the habilitation thesis. The Appendix is only supplementary part of the habilitation thesis[1].

In all papers relevant to the habilitation thesis, nonlinear phenomena related to equilibria bifurcations are included. The analysis of local bifurcations of equilibria usually precedes a more advanced analysis of the dynamical system. Nevertheless, in some cases, it is very substantial itself, either in terms of hysteresis and bistability or loss of stability. The

---

[1]I include it because I consider this application of dynamical systems modelling relevant in the last three years, given my position as the head of the Analysis Group of the National Institute for Pandemic Control and given the number of my publications and other scientific activities during the period of the covid-19 pandemic.

hysteresis phenomenon studied in [27] is related to the mechanism of a trait-dependent cycle in an eco-evolutionary slow-fast system. Another interesting analysis of a biochemical model of hematopoietic stem cells provided in [29] connected a pitchfork bifurcation and its symmetry breaking with the differentiation of erythroid and myeloid cell types from the priming stage through a biochemical switch.

In the field of applied nonlinear phenomena in population biology I made some significant results in predator-prey models with predator inference [28] or foraging facilitation [30]. Both papers study complex dynamics of predator-prey systems with a generalized family of encounter rates that includes a big range of standard functional responses such as Beddington-DeAngelis, Hassel-Varley, or Cosner types. We found a threshold for predator extinction, some kind of inner Allee effect, and confirmed destabilization analogous to the paradox of enrichment due to the increase in predator cooperation. The paper [22] deals with a predator-prey model with an age-structured population of the predator with cannibalism and generalized Hopf bifurcation with respect to the food availability of the adult predator is found. The paper [16] introduced the predator-prey model with the Allee effect in prey and in which individual predators can use Hawk and Dove tactics in fighting over caught prey, which caused a division into two submodels. We analyzed also another age-structured model, the discrete LPA model, with very complex behavior in [13] in the non-chaotic part of the parameter space.

My colleague Veronika Eclerová and I proposed and published in [14] a useful method based on Gröbner basis to partly automate the analytical computation of typical bifurcation manifolds in full parameter space. The approach was implemented in Maple by Veronika Eclerová as a part of her dissertation thesis under my supervision.

Paper [24] studies the route to chaos in a macroeconomic model with foreign capital investment. Even for common economic parameters and for common values of the capital-output ratio, a big range of parameters allows chaotic unpredictable regimes. Another macroeconomic application was a generalization of the neoclassical one-sector Solow-Swan model and its modification of the Kaldor-Pasinetti type with a non-constant labor growth rate published in [25]. This model reveals more complex and even chaotic dynamics.

A useful insight into bifurcations in the complex domain was published in Springer Proceedings in Complexity [26]. This seems to be a perspective approach when applied to bifurcations of limit cycles since it generalizes period doubling to the m-fold bifurcation associated with Arnold tongues and synchronization.

My recent scientific work focuses on a bifurcation theory approach in systems that can exhibit synchronizations. The paper [11] introduces an embedding approach that can be successfully used for harmonically driven systems. Artificial parameters give birth to a torus where the synchronizations and route to chaos may be studied by means of local bifurcation theory, Arnold tongues and m-fold cusps of the m-fold branches of cycles in their neighborhood. This approach is perspective in many other areas. My team already presented some results in neuronal modeling and synchronization of coupled neurons [39], [33]. We started to study seasonality effects on population dynamics of predator-prey models and epidemiologic models with a new team member Deeptajyoti Sen and Veronika Eclerová. Deeptajyoti Sen received MSCA grant from the EU under my supervision. We have just submitted a new paper with our recent interesting results on the effect of seasonal changes in the Allee effect, which are likely due to climate change, on population dynamics. The most promising part of applied bifurcation theory on synchronization is modeling coupled neural networks and neuronal mass networks with possible new approaches to study focal epilepsy. Our team started to analyze EEG data with very and ultra high-frequency oscillations provided by prof. Brázdil's research group. Our collaboration could lead to interesting results in the future, we are submitting together a proposal for the GAMU Interdisciplinary Project this year.

Contents

Acknowledgements

My deep gratitude belongs to my family, especially my husband Zdeněk, my mother, and my deceased father, who supported me all my life in my scientific carrier.

I would like to express deep gratitude to my former doctoral supervisor doc. RNDr. Josef Kalas, CSc. since opening the scientific world for me. I am grateful to prof. RNDr. Zdeněk Pospíšil, Dr. for his kind help, tremendous insight, broad knowledge, and experience in applied dynamical systems and modeling.

My sincere thanks go to my dear friend, former student, but over time, my most indispensable colleague RNDr. Veronika Eclerová, Ph.D. Her enthusiasm is infectious, her work ethic and speed encourage competitiveness, her skills complement mine perfectly, and in my opinion, we have become a very well-functioning team.

Similarly, my many thanks go to the Head of our department prof. RNDr. Jan Slovák, DrSc., who supported my work and research. Great merit goes to all my team, to my postdoctoral fellow Deeptajyoti Sen, Ph.D., M.Sc., to all my students, doctoral, master, and even bachelor students who have done a lot of hard work with me (or without me) on the research and projects.

There are many more outstanding people in our Department but I found great colleagues that work mainly in the Czech Academy of Science. The most prominent collaborator is doc. Ing. Luděk Berec, Dr. My cooperation with him resulted not only in high-quality publications but most importantly in a strong bond of friendship that has developed between us, based on our common experience during the pandemic. He was always completely reliable in both scientific and human judgment.

Chapter 1

Introduction

The dynamically changing world around us can be described reasonably well by the interrelationships between quantifiable variables in many cases. In general, a dynamical system can be defined as a trinity $\{T, \mathbf{X}, \boldsymbol{\varphi}^t\}$, where T is a number set (at least a monoid) that represents the time (it can be discrete or continuous or more generally be represented by a topological monoid), $\mathbf{X}$ is a metric space called a state space or a phase space, and $\boldsymbol{\varphi}^t$ is a family of evolution operators depending on a time parameter $t \in T$ defined as $\boldsymbol{\varphi}^t : \mathbf{X} \to \mathbf{X}$ that map the initial state $\mathbf{x}_0 \in \mathbf{X}$ to a state $\mathbf{x}_t = \boldsymbol{\varphi}^t \mathbf{x}_0 \in \mathbf{X}$. Such a definition of a dynamic system covers almost anything that we can quantify. Imagine a cooling cup of coffee on a table described by Newton's law of cooling using a differential equation; motions of celestial bodies defined by laws of gravity; protein production in a cell described by kinetic chemical equations; a number of infected subjects during the COVID-19 epidemic described by a system of difference equations or stock market movements described by stochastic differential equations. These all are examples of such dynamical systems.

For the simplicity of the model of the observed dynamical process, we often use a deterministic system, where $\boldsymbol{\varphi}^0 = \mathbf{id}$. This property says that the system does not change its state spontaneously. Such a model is not always suitable, but it is in many cases. If we work with physical quantities such as the model of cooling coffee or the motion of celestial bodies, the laws of physics are relentlessly deterministic until a randomly flying comet comes across or a random movement of a hand awkwardly overturns the cup of coffee. However, we can often omit the randomness even with completely randomly behaving systems. For example, in a multi-particle system such as molecules involved in a chemical reaction or people meeting during an epidemic, the system can be modeled deterministically - using aggregation.

However, in my scientific work, as well as in the following text, I focus on

even more specific deterministic dynamical systems, so-called autonomous systems. For autonomous systems, the rules of evolution do not change over time. The notation of a deterministic system as autonomous is rather technical. We can transform time or embed the system in a suitable hyperspace.

An autonomous dynamical system is a deterministic dynamical system $\{T, \mathbf{X}, \boldsymbol{\varphi}^t\}$ that satisfies $\boldsymbol{\varphi}^{t+s} = \boldsymbol{\varphi}^t \circ \boldsymbol{\varphi}^s$, i.e., $\forall \mathbf{x} \in \mathbf{X} : \boldsymbol{\varphi}^{t+s}\mathbf{x} = \boldsymbol{\varphi}^t(\boldsymbol{\varphi}^s\mathbf{x})$ whenever both sides of the equation are defined.

A typical example of an autonomous system is a continuously time-dependent state $\mathbf{x} = \mathbf{x}(t) \in \mathbb{R}^m$, for $t \in \mathbb{R}$, that is changing according to a system of ordinary differential equations

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\varepsilon}) \tag{1.1}$$

or an iterative process

$$\mathbf{x}_{n+1} = \mathbf{f}(\mathbf{x}_n, \boldsymbol{\varepsilon}), \quad n \in \mathbb{N}, \text{ or } n \in \mathbb{Z}, \tag{1.2}$$

where $\mathbf{f} : \mathbb{R}^{m+p} \to \mathbb{R}^m$ is a smooth enough function and $\boldsymbol{\varepsilon} \in \mathbb{R}^p$ are given parameters. Thus, these are dynamical systems where the time dependence is not explicit, only through changing state variables.

It would seem that such restrictions on dynamical systems would be too big to describe and explain the complexity of the world around us. But the opposite is true. Not only that a large number of dynamical systems in various fields can be described this way, but the dynamics of such systems themselves are so diverse that typical dynamical phenomena can be modeled very well. Even such systems can describe phenomena with unpredictable dynamics, i.e., chaotic.

Chapter 2

Bifurcations of equilibria and related phenomena

The simplest dynamics allowed by equations (1.1) or (1.2), respectively, is near equilibrium dynamics. An equilibrium is given implicitly in the state-parameter space $\mathbb{R}^m \times \mathbb{R}^p$ since it is given by the manifold

$$\mathbf{f}(\mathbf{x}, \varepsilon) = \mathbf{0}, \text{ or } \mathbf{f}(\mathbf{x}, \varepsilon) - \mathbf{x} = \mathbf{0}, \text{ respectively.}$$

These zero loci in state variable $\mathbf{x}$ are called equilibrium manifolds. At this place, it is worth noting the difference in the definition of a manifold as a topological space that is locally Euclidean and an algebraic variety. The term manifold is used in dynamical systems theory, although the term variety may be more appropriate. However, the Czech language does not distinguish between the names. An affine algebraic variety is defined as a set of solutions of a system of polynomial equations over complex numbers[1]. A variety can also contain singular points, where tangent space cannot be defined (in contrast, a manifold is regular at each point). The equilibrium manifold of the system (1.1) or (1.2), respectively, can also contain singular points in the state-parameter space. However, we most often use terms as branches of equilibria or branches of the equilibrium manifold in such a case.

At this point, we finally come to the concept of bifurcation and gradually to the promised applications.

The implicit equilibrium can be locally defined in the neighborhood of a fixed parameter value ε_0 thanks to the implicit function theorem as a function $\varepsilon \mapsto \boldsymbol{\beta}(\varepsilon)$ satisfying $\boldsymbol{\beta}(\varepsilon_0) = \mathbf{x}_0$ and $\mathbf{f}(\boldsymbol{\beta}(\varepsilon), \varepsilon) \equiv \mathbf{0}$ or $\mathbf{f}(\boldsymbol{\beta}(\varepsilon), \varepsilon) \equiv$

---

[1]In applications, we are of course interested in its real part, basic dynamic phenomena are described using systems of polynomial equations.

$\beta(\varepsilon)$, respectively. Simply put, $\beta(\varepsilon)$ corresponds to the equilibrium for given parameter ε.

However, neither the existence nor the uniqueness of the equilibrium $\beta(\varepsilon)$ is guaranteed if Jacobi matrix $D\mathbf{f}(\mathbf{x}_0, \varepsilon_0)$ has a zero eigenvalue in the continuous case or equal to 1 in the discrete case. A small parametric perturbation may cause the disappearance of the equilibrium point (in any neighborhood of ε_0), or create a new branch of equilibrium solutions (hence the name, branching = bifurcation) and of course, the equilibrium may change its stability during the transition through ε_0. In other words, the perturbation may cause a general qualitative change in system behavior.

By a local bifurcation of the system (1.1) or (1.2), respectively, near equilibrium $\mathbf{x}_0 = \beta(\varepsilon_0)$ with the critical value of the parameter $\varepsilon = \varepsilon_0$, we understand the qualitative change of dynamics in the neighborhood of the critical value ε_0, that is, the phase portraits near equilibrium $\mathbf{x}_0$ while crossing the bifurcation parameter ε_0 are not locally topologically equivalent, i.e., no homeomorphism maps the phase portraits near the equilibrium to each other. A local change in dynamics (bifurcation) can occur when an eigenvalue of the Jacobi matrix $D\mathbf{f}(\mathbf{x}_0, \varepsilon_0)$ has zero real part in the continuous case or modulus equal to 1 in the discrete case.

Obviously, local bifurcations will play an important role in applications because they cause changes in dynamics with a slight change in the parameter value. It can be a change in the rate of a chemical reaction that makes a cell start to produce a specific protein; it can be a reduction in the average number of contacts between people that causes the outbreak of an epidemic; it can be a slight percentage increase in catches of fish that causes their extinction; or an increase in the average ocean temperature that causes a change in the dynamics of sea currents and the Earth's ecosystem.

2.1   Fold of the equilibrium manifold

The simplest local bifurcation is the fold of an equilibrium manifold, or fold bifurcation, sometimes called the saddle-node bifurcation[2]. This is the case when the Jacobi matrix $D\mathbf{f}(\mathbf{x}_0, \varepsilon_0)$ has exactly one eigenvalue $\lambda = 0$ and no other eigenvalue with zero real part in the continuous case or $\lambda = 1$ and no other eigenvalue with modulus equal to 1 in the discrete case. In applications, other eigenvalues have typically negative real parts in the continuous case or moduli less than 1 in the discrete case, since in such

---

[2]This name is based on the frequent situation where the saddle and node merge in the fold, which is not always necessary. The name tangent bifurcation is sometimes used as well as a limit point.

a case, the bifurcation guarantees a change of stability – from a stable equilibrium to an unstable one. The system can then be reduced to the so-called central manifold, which always exists locally in the neighborhood of the equilibrium and is an invariant set to which close trajectories are attracted. To give the reader an idea of the generality of the results of the bifurcation theory, I present here the elementary local bifurcation of an equilibrium in a continuous case with respect to one parameter. It is possible to work analogously with the system of iterations (maps) (1.2).

Theorem 2.1.1. *Suppose a one-parametric one-dimensional equation*

$$\dot{x} = f(x, \varepsilon), \quad x \in \mathbb{R}, \ \varepsilon \in \mathbb{R}, \tag{2.1}$$

where f is a smooth function, has equilibrium $x = x_0$ for $\varepsilon = \varepsilon_0$ and $\lambda = f_x(x_0, \varepsilon_0) = 0$. Suppose that conditions

$$f_{xx}(x_0, \varepsilon_0) \neq 0 \qquad \text{non-degeneracy condition,}$$
$$f_\varepsilon(x_0, \varepsilon_0) \neq 0 \qquad \text{transversality condition.}$$

are satisfied. Then the system (2.1) *is locally topologically equivalent near equilibrium x_0 and parameter ε_0 with one of four fold bifurcation normal forms*

$$\dot{x} = \pm\alpha \pm x^2$$

with parameter α near zero.

An example of a fold bifurcation diagram for the case (choice of signs) $\dot{x} = \alpha - x^2$ is in Fig. 2.1. The critical point of the fold bifurcation is most often labeled LP (limit point) in the bifurcation diagrams, and for this normal form, it is the vertex of the parabola $\alpha = x^2$, i.e., the origin of the parameter-state space. Despite its simplicity, this bifurcation is related to many nonlinear phenomena and can be applied in many situations.

2.2   Switches and hysteresis

A typical example of an abrupt change caused by the fold bifurcation is a biochemical switch. A very simplified kinetic equation of genetic protein production in a cell has the form

$$\dot{g} = k_1 \frac{g^2}{1+g^2} - k_2 g,$$

where g quantifies the amount of protein and k_1, k_2 are positive parameters related to the rates of chemical reactions. The first term $k_1 \frac{g^2}{1+g^2}$ corresponds
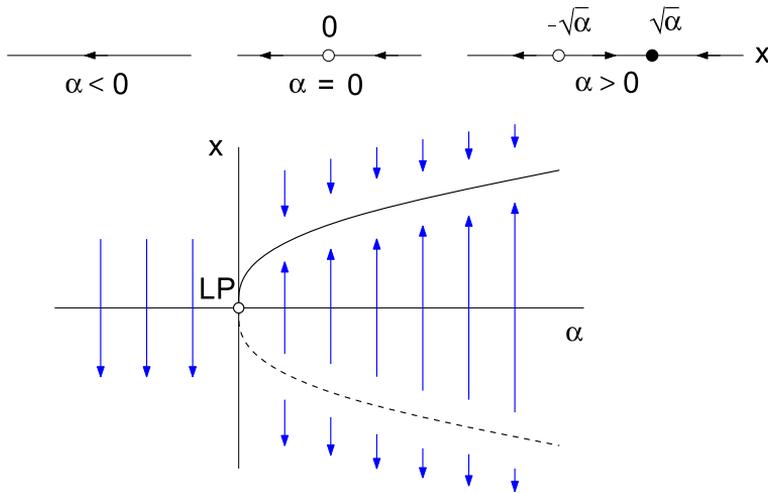
Figure 2.1: Bifurcation diagram for the fold bifurcation normal form.

to gene expression through transcription and translation. The protein is formed by an autocatalytic reaction, which is not mentioned here and is described by the so-called Hill function. The second term $k_2 g$ corresponds to the degradation of the protein.

The new variable $x(t) = g(\frac{t}{k_1})$ and parameter $a = \frac{k_2}{k_1}$ transforms equation to much more simpler form

$$\dot{x} = \frac{x^2}{1+x^2} - ax. \tag{2.2}$$

Any equilibrium satisfies

$$x\left(\frac{x}{1+x^2} - a\right) = 0.$$

We get two branches of equilibrium manifolds (depending on parameter a): the trivial equilibrium and the non-trivial manifold

$$a = \frac{x}{1+x^2}.$$

We are only interested in non-negative protein concentration values and the positive ratio of reaction rates (parameter a). But the origin is important since it is the intersection of two branches of equilibria (so-called transcritical bifurcation). The point $[a^*, x^*] = [\frac{1}{2}, 1]$ is the limit point, or fold of the non-trivial equilibrium manifold (shown in Fig. 2.2). At this point, fold bifurcation occurs, with the upper part of this equilibrium branch stable and the lower unstable.
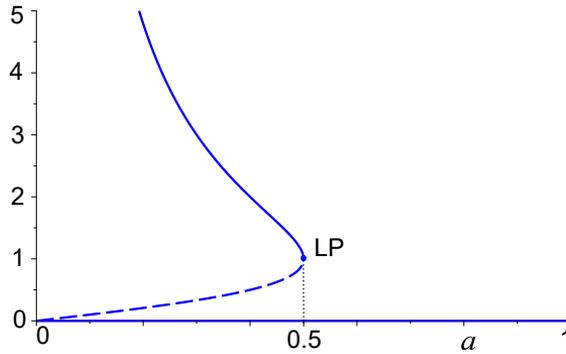
Figure 2.2: Rovnovážné variety rovnice (2.2) v závislosti na parametru a.

If the parameter $a = \frac{k_2}{k_1}$ exceeds the critical value $\frac{1}{2}$, the gene stops being produced and a biochemical switch to the zero branch of the equilibrium occurs.

Very often, two such switches occur in dynamical systems. This creates a phenomenon called hysteresis. A system with hysteresis typically exhibits a kind of delay or prevention of return to the original state. Hysteresis is well known in ferromagnetic materials (sketch in Fig. 2.3), which, after exposure to a magnetic field, exhibit magnetic properties for some time, after which the internal magnetic field disappears.

However, this phenomenon also appears in applications in other fields – biology, medicine, economics, etc. In fact, it is the fold of the equilibria manifold depending on two parameters. In bifurcation theory, the critical point of this equilibrium variety is called the two-parameter local bifurcation of the cusp equilibrium.

In this case, the system can be reduced near the equilibrium to a central manifold and written in the form of the normal form of the two-parametric cusp bifurcation

$$\dot{y} = \varepsilon_1 + \varepsilon_2 y \pm y^3. \tag{2.3}$$

Equilibria lie on the variety $M : \varepsilon_1 + \varepsilon_2 y \pm y^3 = 0$, which is shown in Fig. 2.4. The first derivative according to y vanishes (the fold bifurcation condition) on the curve satisfying $\varepsilon_2 \pm 3y^2 = 0$, which has two branches and the cusp that connects them. Hence the name of this bifurcation.

The branches T_1, T_2 correspond to the extinctions of the pair equilibrium points in the fold of the manifold M, i.e. they are the bifurcation boundaries of the fold bifurcation (LP). If we exclude y from these two equations, we get the projection of the folds into the parameter plane $(\varepsilon_1, \varepsilon_2)$, which is a curve of a typical V-shape with a point of return at the origin.
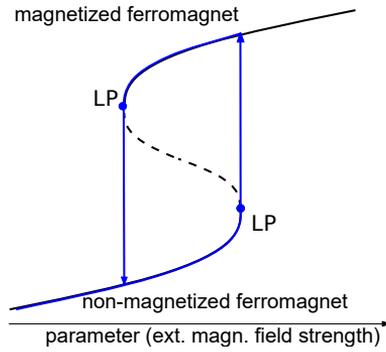
Figure 2.3: Sketch of hysteresis in a ferromagnetic material as a function of the magnitude of the external magnetic field.
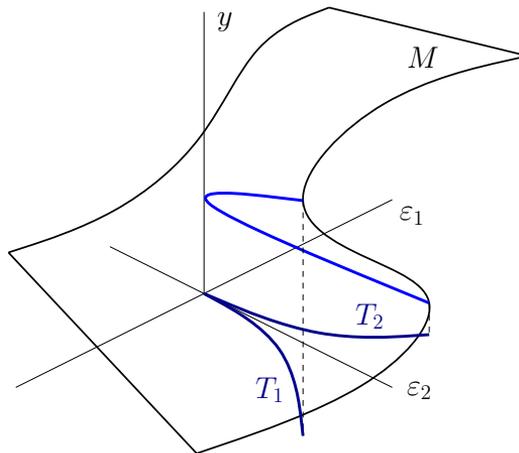

Figure 2.4: Equilibrium manifold of the equation (2.3) depending on two parameters ε_1 and ε_2.
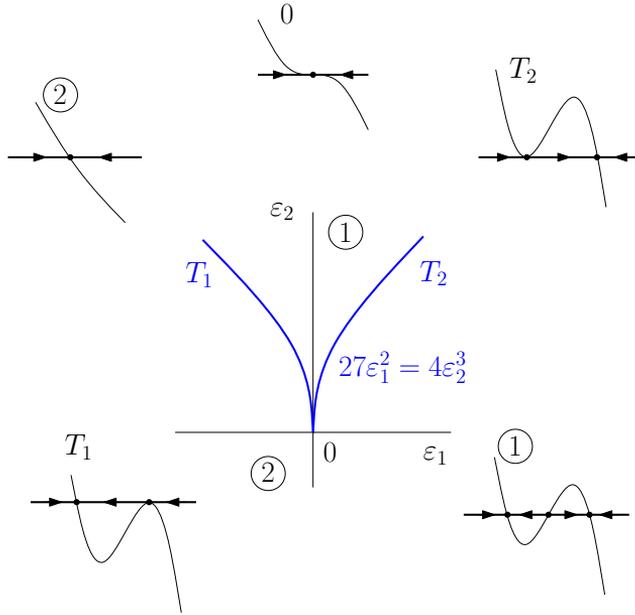
Figure 2.5: Bifurcation diagram of the cusp bifurcation normal form $\dot{y} = \varepsilon_1 + \varepsilon_2 y - y^3$.

For the minus sign, it is shown in Fig. 2.5 and is of the form

$$27\varepsilon_1^2 - 4\varepsilon_2^3 = 0.$$

The areas marked 1 and 2 are structurally stable[3], in which the system has three or a single stable equilibrium, respectively. The manifolds T_1 and T_2 correspond to a one-parametric fold bifurcation, they have codimension 1 in the 2-dimensional parameter space. Their intersection is the cusp point, which has dimension 0, i.e. codimension 2 in the 2-dimensional parameter space. The cusp bifurcation point is therefore called a local two-parameter bifurcation or a codimension-2 bifurcation point. This typical phenomenon does not depend on the form of the equations and arises generically in the situation when two branches of the fold bifurcation touch. There are, of course, exact conditions (the so-called non-degeneracy and transversality conditions similar to the 2.1.1 theorem) under which the general system undergoes a cusp bifurcation, but the geometric idea of the two-folding of the equilibrium manifold depicted in Fig. 2.4 is much more useful for our purposes.

---

[3]Structural stability means the preservation of the existence of three equilibria or a single stable equilibrium, respectively, even with a small change of parameters

Hysteresis thus appears in a situation where sigmoid-type nonlinearity is present. Such systems are very common in nature. Examples are population models where the nature of the sigmoid is determined by, for example, predation functions or other environment-dependent functions. The appetite of a predator is limited by some upper bound, and other functions affecting the prey population are often also sigmoidal in nature, precisely because of the lower and upper bounds, which are determined by the capacity of the environment or the size of other interacting populations. In biochemistry it is similar. Especially, Hill functions that are common in biochemistry are sigmoidal in nature. They are also encountered in the above simple model of a biochemical switch. Such functions describe the kinetics in chemical reactions that typically arise for autocatalytic reactions in living cells, while chemical reactions outside living organisms, on the other hand, are typically sigmoid-free so that they lead to equilibrium states without giving rise to more complex phenomena.

Fold bifurcation and hysteresis are not the only phenomena related to the singularity of the equilibrium manifold. Other generic one-parameter bifurcations are the transcritical bifurcation (generic intersection of two branches of equilibrium manifolds) and the pitchfork bifurcation, which exhibits some symmetry and its applications appear, for example, in evolutionary branching models. In fact, in the pitchfork bifurcation, the intersection of two branches of equilibria occurs just at the fold of one of the branches. In applications, this makes it possible to model precisely the phenomena of branching, where one stable equilibrium disappears (and becomes unstable) and at the same time two other stable branches arise in its vicinity, see Fig. 2.6.

2.3 Author's contribution to the field

In fact, the basic types of bifurcations of equilibria are included in all my scientific papers and contributions. Equilibria disappearance, hysteresis, and pitchfork branching are ubiquitous and present as basic phenomena in all systems I ever studied in my carrier. They are usually continued to more advanced analysis of at least two-parametric bifurcations and that's why I will choose more interesting phenomena described in my works later. Here, I would like to mention hysteresis phenomena first of all. I studied and explained in [27] a hysteresis-related mechanism of trait-dependent cycle in the eco-evolutionary slow-fast system and hysteresis caused by a double fold of the equilibrium manifold in a population model with respect to the predator encounter rate change, more precisely a change in predator interference strength [28]. Another partial result of my paper [25]
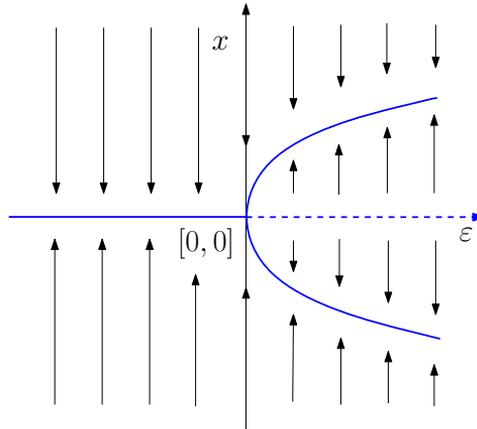
Figure 2.6: Bifurcation diagram of the pitchfork bifurcation normal form $\dot{x} = \varepsilon x - x^3$.

that dealt mainly with the emergence of chaotic dynamics was bistability in a continuous Solow-Swan model. And an interesting example was a pitchfork bifurcation and its symmetry breaking (sometimes called an imperfect bifurcation) in a biochemical model of hematopoietic stem cells studied in [29] that can cause a switch due to the pitchfork bifurcation. One branch of equilibria corresponds to the differentiation of erythroid and megakaryocytic cells and the other branch corresponds to the differentiation of myeloid and lymphoid cells. Such a model explains the priming stage of hematopoietic stem cells as well as the biochemical switch since the basins of attraction of the two equilibrium branches are separated by the unstable branch merging from the pitchfork bifurcation, together with continuation of this phenomenon into the imperfect non-symmetric part of the parameter space that is probable in nature. Moreover, the disorder in hematopoietic differentiation could be caused by a shift out of this near pitchfork region.

Chapter 3

Bifurcations of cycles and related phenomena

One can certainly expect cycles, i.e. oscillations, to be another important dynamic with interesting applications. Pendulums, electromagnetic waves, alternating currents, heart muscle activity or the motion of planets are all examples of cyclic dynamics. It is also very interesting that one of the fundamental nonlinear phenomena, the origin of oscillations, is associated with the local bifurcation mentioned in the previous text.

3.1 Birth of cycles by Hopf bifurcation

For simplicity, we will deal only with the continuous system (1.1) with a smooth enough right side. If the two complex eigenvalues of the Jacobi matrix $D\mathbf{f}(\mathbf{x}_0, \boldsymbol{\varepsilon}_0)$ for a given equilibrium $\mathbf{x}_0$ change so that their real parts cross the imaginary axis when the parameter changes, the equilibrium manifold is given uniquely with respect to the parameter $\boldsymbol{\varepsilon}$, but the stability changes. Generically, a so-called Hopf bifurcation occurs, when a stable focus turns into an unstable focus. Trajectories close to equilibrium thus exhibit damped or gradual oscillations. It can be shown that in the generic case (under certain conditions of non-degeneracy and transversality) this bifurcation is necessarily associated with the formation of a limit cycle. Either a stable cycle can arise near an unstable focus (we speak of a supercritical Hopf bifurcation) or an unstable one near a stable focus (we speak of a subcritical bifurcation). As in the case of fold bifurcation, the decision on the type of bifurcation is given by the sign of a certain term in the normal form. The description of the cycle is simpler in complex or polar normal form. Here we give the complex normal form of the supercritical
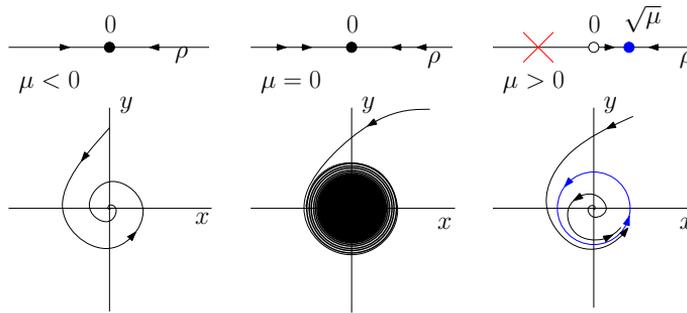
Figure 3.1: Supercritical Hopf bifurcation.

Hopf bifurcation

$$\dot{z} \ = \ (\mu + i)z \ - z \ |z|^2,$$

where the Euler form of the complex number $z = \rho e^{i\varphi}$ gives the polar form

$$\dot{\rho} \ = \ \rho(\mu - \rho^2), \qquad (3.1)$$
$$\dot{\varphi} \ = \ 1.$$

The equation (3.1) is the pitchfork bifurcation normal form. So for $\mu \leq 0$ the origin is the only stable equilibrium. For $\mu > 0$, there is another equilibrium point $\rho = \sqrt{\mu}$ (we can omit the negative value, it is meaningless in this representation, it is a distance). The origin is unstable in this case for $\mu > 0$, the equilibrium $\rho = \sqrt{\mu}$ is stable. This point corresponds to the stable limit cycle near the origin (see Figure 3.1). From the complex normal form, it can be seen that the complex conjugated eigenvalues are $\mu \pm i$. For simplicity, the parameter ε is replaced here directly by the real part of the eigenvalue μ which is the origin of the qualitative change in the dynamics.

In terms of applications, these two types of Hopf bifurcation are very different. If we imagine a supercritical bifurcation on some real phenomenon, we find that it is not a very significant change. In the small neighborhood of the critical parameter value, although the equilibrium loses stability, it is replaced by small oscillations (the amplitude of the oscillations increases with the square root of the parameter). An example of this is the emergence of a tone when playing the flute or when the tea kettle whistles. At low air velocity, no tone is produced, but when a certain threshold is exceeded, it is very quiet. A strong tone (large amplitude of the same frequency) is produced by a strong full breath into the instrument. The whistle of a kettle boiling water also starts gently and gradually gains in strength. Conversely, subcritical bifurcation is usually associated with catastrophic phenomena

such as the aeroelastic phenomenon of the wings of an airplane fluttering, which may break due to the wind gust that exceeds a certain threshold [34]. Similarly, there are known cases of bridge collapses [10] ("the most famous" is probably the collapse of the Tacoma Bridge), the occurrence of engine vibrations when turbines are shutting down or during aircraft landing [3]. In this case, it is the disappearance of stable equilibrium (surrounded by an unstable cycle) and the rapid emergence of strong oscillations with a large amplitude.

In real applications, this phenomenon is similar to hysteresis and is similarly related to the two-parameter bifurcation, when another stable cycle is present near the unstable cycle. It is also called bistability, i.e. the simultaneous appearance of two stable attractors. Thus, a subcritical Hopf bifurcation can be non-catastrophic. In neurons, this is the most typical way for lasting oscillations (bursting) to arise. In fact, when the threshold is crossed, a cycle with progressively increasing amplitude does not arise as in the supercritical bifurcation, but the system jumps directly to the stable branch of the other cycle with a large amplitude. As an example, consider the simplest FitzHugh–Nagumo model of a neuron.

$$\dot{V} = V - \frac{1}{3}V^3 - W + i,$$
$$\dot{W} = a\left(bV - cW + d\right),$$

where V is the membrane potential, W is the return-related variable, i is the delivered current, and a, b, c, d are the parameters. Note that the change in voltage V across the axon membrane described by the first equation is a sigmoid curve. The other parameters and the state variable W are based on the description of the kinetics of chemical reactions at the axon membrane (transfer signal is mediated by changes in the concentrations of ions K^+, Na^+, Cl^- and protein anions), $a, b, c > 0$. The second equation is a restoring equation, has a slower response (hence the parameter a, which has a small value, is retained) and allows for the generation of a pulse that subsequently terminates.

Figures 3.2 and 3.3 demonstrate why the neuron responds by sending an oscillating signal only for certain values of the supplied current i. The equilibrium plotted with respect to the parameter i as the red curve $[V(i), W(i)]$ passes through two subcritical Hopf bifurcations. However, the pink cycles are unstable and the system does not stabilize on them, so we never observe them. We only observe significant oscillations corresponding to the blue cycles or equilibria. Until the current is increased to a sufficient value (threshold), the neuron does not respond; if the current is too large, it does not either.
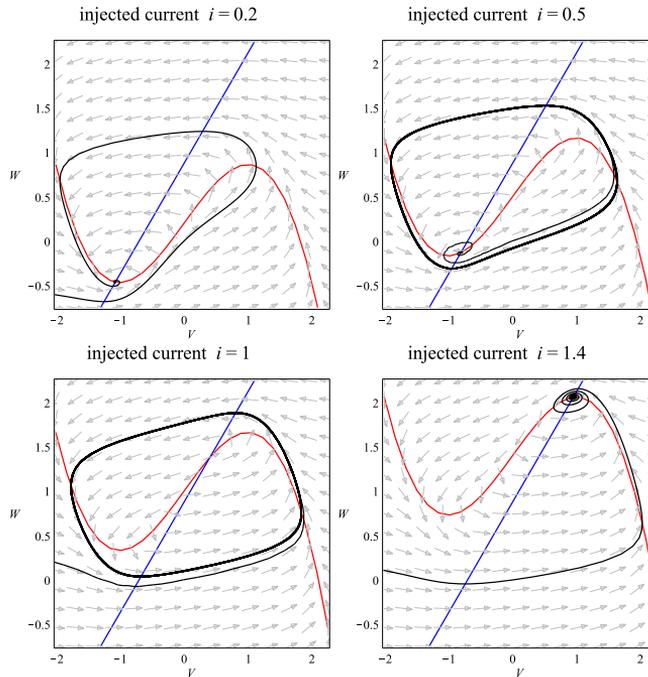
Figure 3.2: Phase portraits for $a = 0.2, b = 1, c = 0.8, d = 0.7$. From conference poster CMSB 2018, computed by Veronika Eclerová (Hajnová) in Matcont.
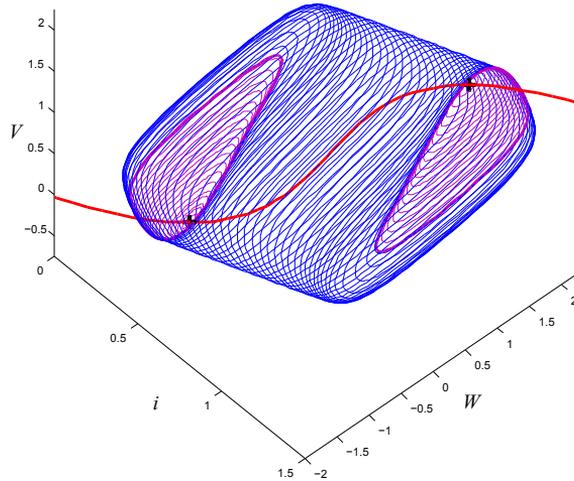
Figure 3.3: Cycles occurring near equilibrium manifold $a = 0.2, b = 1, c = 0.8, d = 0.7$. From conference poster CMSB 2018, computed by Veronika Eclerová (Hajnová) in Matcont.

3.2   Extinction of oscillations on the saddle

Limit cycles are commonly observed in predator-prey population models. Systems of this type often have a saddle point in the admissible non-negative region of the state space. Significant trajectories that approach limit equilibrium near the saddle point for $t \to \pm\infty$ are also important in terms of applications, regardless of whether the system is continuous or discrete. These trajectories can separate parts of the state space and thus separate regions in which the trajectories have completely different dynamical properties. Saddle trajectories also got their name - separatrices. Thus a limit cycle can (and often must) split on a saddle separatrix loop. Thus, for example, in predator-prey models, large oscillations can suddenly cease and lead to a significant change in population size, even to the extinction of a species; a multitude of factors ranging from the mode of predation to evolutionary pressure can play a role in such a change in dynamics, and bifurcation theory is able to disentangle some of these links. Some applications in papers of which I am an author or co-author related to this phenomena [28], [30], [13], [27] are be briefly described in subsection 3.4. The study of such global bifurcations is more difficult and uses mostly continuation numerical methods. However, due to their close connection with local bifurcations of higher codimension, there are also methods that use

an analytical approach, e.g., by using Gröbner bases [14]. This approach makes it possible to partially automate the finding of important bifurcation manifolds for systems (1.1) and (1.2) with function **f** in the form of a polynomial or rational function, and to allow a wider range of people, not just specialized mathematicians, to perform at least partial bifurcation analysis of dynamical systems. Indeed, it turns out that just such systems appear in descriptions of biochemical reactions in cells or in the aforementioned models of population biology or neuroscience, and understanding the dependence of changes in dynamics on changes in the parameters of the dynamical system is crucial.

3.3   Fold of the cycle manifold

You have probably noticed that in the example of the subcritical Hopf bifurcation in the neuron model, a fold of the manifold that belonged to limit cycles was present. In fact, one can move from the study of cycles in a clever way to the study of equilibria – fixed points of maps. Indeed, if we make a transversal section through a limit cycle (in the space of state variables), we can look at the trajectories in its neighborhood as iterations of the representation given by the intersection of the trajectory with this section. This idea is brilliant, and its author Henri Poincaré was equally brilliant. He was many decades ahead of his time in the study of dynamical systems.

We can apply the local bifurcation theory for fixed points (equilibria) of the Poincaré map[1] to cycles. Thus, the appearance or disappearance of a pair of fixed points (usually one stable and one unstable) of the Poincaré map by a generic fold bifurcation gives rise or disappearance to the corresponding cycles exactly as we see in Fig. 3.3.

3.4   Author's contribution to the field

Much of my scientific work has been in the field of applied nonlinear phenomena in population biology and specifically predator-prey models. Described nonlinear phenomena are related to various bifurcations, some of these bifurcations are advanced and not mentioned in the above text. I studied predator-prey models with predator inference [28] or foraging facilitation [30] that is with a specific type of the encounter rate of predator that is dependent on the predator density. Predator inference is the

---

[1]Poincaré map is that ingenious first return map defined on the transversal section of a cycle described above

case when the encounter rate function λ of the predator is decreasing with predator density and on the other hand, foraging facilitation describes the situation when the predators cooperate and λ increases with the predator density. Both cases are common in nature. The mostly used Holling type II functional response is a neutral case. Our generalized family of encounter rates included, except the Rosenzweig-MacArthur predator-prey model [31] with paradox of enrichment, also many other commonly used functional responses as Beddington-DeAngelis [1], [9], Hassel-Varley [15] or Cosner types [4].

In the case of predator interference more complex dynamics appear – from cusp, bistability of equilibria, and hysteresis phenomenon that can abruptly switch the population levels in both predator and prey populations, to both types of Hopf bifurcation. Supercritical Hopf bifurcation is analogous to the paradox of enrichment in the Rosenzweig-MacArthur predator-prey model [31] when the increase of availability in prey population destabilizes the predator population while crossing the Hopf bifurcation threshold. It can be similarly observed with respect to the strength of the predator interference. In the case of the Rosenzweig-MacArthur model [31], the predator nullcline is vertical, intersects the prey nullcline (at most) once, and a supercritical Hopf bifurcation with the appearance of stable limit cycles occurs once the predator nullcline crosses the vertex of the prey nullcline from right to left. When the Holling type II functional response is replaced by the Beddington–DeAngelis one [1, 9] the predator nullcline stays straight but leans to the right. As a consequence, the paradox of enrichment occurs for larger carrying capacities when compared to the Holling type II functional response, if at all, which has given rise to the notion that predator interference is a major stabilizing factor of predator-prey dynamics [9, 36].Finally, for sufficiently strong predator interference functional responses from the family, we consider the cause of the predator nullcline to bend to the right, and if sufficiently bent, it may intersect the prey nullcline at three points, thus giving rise to hysteresis and fold and cusp bifurcations. Fold and cusp bifurcations have not been observed in the previous two cases. Moreover, stable limit cycles that may appear on the lower branch of the hysteresis curve due to the Hopf bifurcation may break up on its middle branch due to a homoclinic bifurcation and stabilize the dynamics on a higher level, see Figure 3.4. The existence of generalized Hopf bifurcation and subcritical Hopf bifurcation gives birth to an abrupt change in stability, a switch onto a large amplitude stable cycle that encounters also a neighborhood of zero, and a small exogenous perturbation may lead to a transition to extinction. Interestingly, the most complex behavior appears at the middle values of the interference strength.
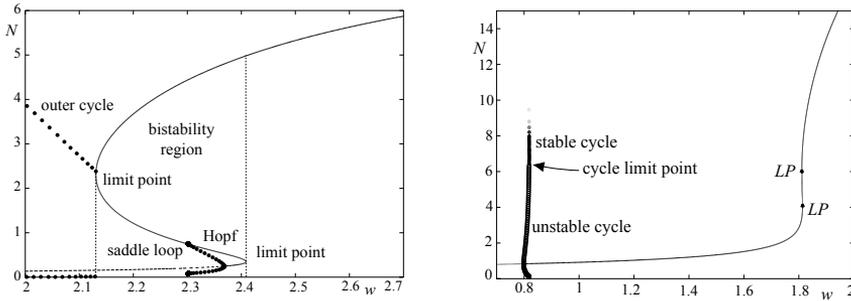
Figure 3.4: A homoclinic cycle break (left), a subcritical Hopf bifurcation and a large amplitude outer cycle (right). From [28].

Bistability is possible also for the foraging facility case [30]. We studied two types of encounter rates - unbounded and bounded. We found a threshold for predator extinction, some kind of inner Allee effect, destabilization analogous to the Rosenzweig-MacArthur paradox of enrichment due to the increase in predator cooperation and typical extinction phenomenon on the homoclinic bifurcation of the saddle loop.

Similar results can be found in our papers [22] and [16]. The first paper analyzed a system proposed in [21] as a predator-prey model with an age-structured population of the predator with cannibalism. Both supercritical and subcritical Hopf bifurcations occur with respect to the food availability of the adult predator. The second paper introduced the predator-prey model with the Allee effect in prey and in which individual predators can use Hawk and Dove tactics in fighting over caught prey, which caused a division into two submodels. The bifurcation analysis revealed the existence of saddle-node bifurcation, two supercritical Hopf bifurcations, and two Bogdanov-Takens bifurcations in the aggregated model. Moreover, we have shown the existence of fold of limit cycles, which is in neither submodel alone.

We analyzed also another age-structured model and its complex behavior in [13]. It was a discrete population model for a flour beetle Tribolium castaneum called LPA model introduced and partly studied in [5, 7, 8]. We studied the non-chaotic part of the parameter space with very complex behavior, non-local bifurcations related to multistability, Neimark-Sacker bifurcation, symmetry of 1:2 resonance bifurcation, Chenciner bifurcation, and the cusp of the invariant loop. We showed that slight perturbations (variability, white noise) in initial conditions lead to similar unpredictability in this non-chaotic part of the parametric space as the "usual" deterministic
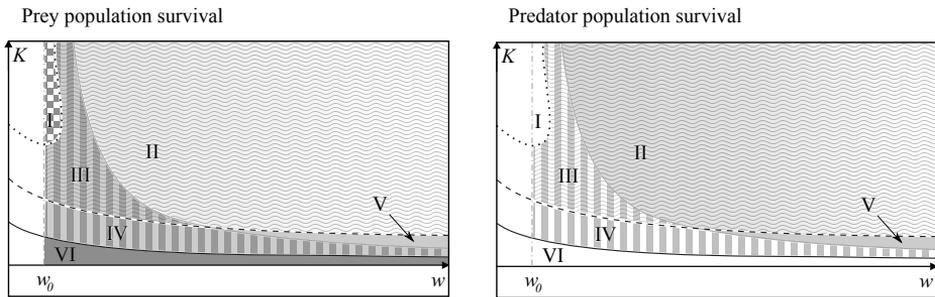
Figure 3.5: Survival of the prey and predator populations in case of foraging facilitation. Extinction regions are shown in white, stable survival in grey. Vertical stripes correspond to the existence of two stable equilibria. White-grey stripes show regions where the population may go extinct (when the initial conditions are out of the basin of attraction of the stable coexistence equilibrium); while the light and dark grey stripes correspond to two levels of population equilibria. Waves correspond to oscillations and the checkerboard pattern corresponds to areas where the trajectories pass very close to the extinction point and survival of the population may be a matter of chance. From [30].
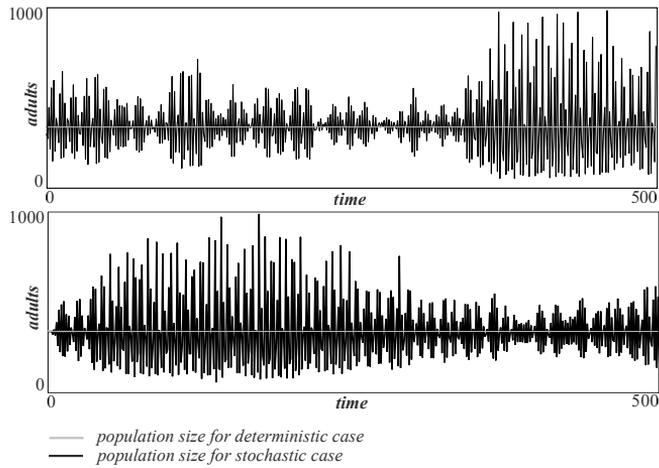
Figure 3.6: Two time series simulating the number of adults for the same initial parameters and the same initial conditions near Chenciner bifurcation. From [13].

chaos itself.

Over time, my colleague Veronika Eclerová and I proposed a useful method to partly automate the analytical computation of typical bifurcation manifolds in full parameter space using Gröbner basis method for some types of dynamical systems [14] and presented it on International Conference on Computational Methods in Systems Biology 2018. The approach was implemented in Maple by Veronika Eclerová on Maplesoft Application Center[2]
`https://www.maplesoft.com/Applications/Detail.aspx?id=154567` and therefore it can be used in applied research as a supporting autonomous computation even by non-experts in bifurcation theory.

---

[2]This is a part of her successfully defended dissertation thesis under my supervision.

Chapter 4

Chaotic dynamics

Chaos theory is another important area with applications. The emergence of chaos is closely linked to discrete dynamics. First of all, the Poincaré map of a limit cycle can go through various types of discrete bifurcations. The folding corresponds to the eigenvalue equal to 1 of the Poincaré map (we often speak of a cycle or Floquet multiplier), but changes in dynamics also occur in other cases when some eigenvalue crosses the unit circle. If the eigenvalue crosses -1, a so-called flip bifurcation of the Poincaré map occurs. A flip bifurcation of a map is a local bifurcation of a fixed point where the eigenvalue of the map is equal to one after precisely two iterations. Therefore, a cycle of the Poincaré map of period 2 is born from a cycle of period 1, in other words, the fixed point of the map breaks into an oscillating pair of points. If we study continuous oscillations using Poincaré maps, then this emergence of a 2 cycle of the Poincaré map implies that the limit cycle of a continuous dynamical system also undergoes a period doubling, and from a topological point of view the doubled cycle is the boundary of the Möbius strip. However, the generic flip bifurcation does not end with a single doubling. If, as the parameter changes, an eigenvalue crosses -1 from the inside the unit circle to the outside, then the multiplier of the doubled cycle will also decrease in value from 1 and may again cross -1 to form a 4-cycle, then an 8-cycle, and in this way a period doubling will gradually occur – the so-called Feigenbaum cascade of bifurcations, which is a very common way to produce chaotic unpredictable behavior in parameter-dependent dynamical systems. Chaotic dynamics is very sensitive to initial conditions. In fact, a small change in the initial conditions leads to an exponentially increasing change in the distances of the original and perturbed trajectories, and this phenomenon has come to be known as the butterfly effect. In one of his lectures, the creator of the famous model that brought the new concept of instability of the Earth's

atmosphere, mathematician and meteorologist Edward Lorenz, used this metaphor to explain the phenomenon in weather dynamics: the flapping of a butterfly's wings in Brazil can cause a tornado in Texas. It should be added that this very phenomenon makes both long-term weather forecasts and many other predictions impossible. The degree of divergence of nearby trajectories can, however, be measured by the so-called maximum Lyapunov exponent, and thus obtain at least some predictability horizon m, i.e., a time interval over which we can accept the prediction error. The emergence of deterministic chaotic dynamics by cascading period doubling is typical and can be found in many applications.

A famous and historically very important experiment was that of Albert Libchaber. In 1977, he created a stainless steel cylinder into which he inserted liquid helium and heated the bottom surface of the cylinder. Afterward, he verified experimentally that the turbulent flow of liquid helium that arises after the breakdown of the convective flow produces this cascade of period doubling. In 1982, he published a similar experiment with mercury[1] [20], where he even measured an estimate of the Feigenbaum theoretically derived constant, which does not depend on the precise form of the system and is general for period-doubling cascades.

Over the last 5 decades, the discovery of the ubiquity of chaos has led to a new perspective on many areas. For example, see the articles on chaotic dynamics in neuroscience. It seems that in the brain, chaos is desirable (!) and, conversely, stable periodic dynamics is an undesirable state – an epileptic seizure [32]. The excitable cells in the heart muscle work synchronously periodically and the chaotic dynamics lead to fibrillation of the heart [35]. The chaotic dynamics in the Universe and the Solar System had been touched upon in advance by Poincaré himself, although he did not and could not yet suspect its extent. Today, the chaotic rotation of Saturn's moon Hyperion and the axis of rotation of Mars is described and explained, NASA used its knowledge of chaotic dynamics to send the ISEE-3/ICE spacecraft on a nearly fuel-free journey to a comet back in 1985, Saturn's rings are being studied for their fractal structure of chaotic attractor [19], even the maximum Lyapunov exponent for the Solar System is calculated [18]. Based on the Lyapunov exponent, the predictability of the system can then be estimated as 36 days for the rotation of Hyperion and 5 million years for the stability of the Solar System.

If you want to play at home, you can try the tap experiment. It's not so easy, a lever faucet is not suitable, but a good old (ideally even dripping) tap is sufficient laboratory equipment. If the faucet is almost

---

[1]A magnetic field is induced in the rotating mercury convective currents, which was measurable by damping the electrical oscillators using the principle of frequency analysis.

closed, and very slightly leaks, the water drips. Plink, silence, plink, silence, plink, silence. It can be quite a disturbing periodic sound. The same plink, the same silence. Plink and... that's a cycle (of length 1). If you're dexterous enough to loosen the tap just a little, it will drip differently. Plink, plink, silence, plink, plink, silence. Then maybe you can set up a four-drop cycle... Because it will quickly start dripping aperiodically. You're seeing and hearing a chaotic attractor. If you don't believe me, look in [38].

4.1 Routes to chaos

Feigenbaum's period-doubling cascade is not the only route to chaos, although it is certainly the most studied. Chaotic dynamics can occur in other ways as well. We can mention some: Shilnikov route to spiral chaos related to Shilnikov bifurcation, Ruelle–Takens–Newhouse route to the quasiperiodic onset of chaos that is born from the break of a quasiperiodic torus, crisis route to chaos near collision of a chaotic attractor with an unstable invariant set, etc.

One important chaotic phenomenon (of a boundary crisis type) is intermittency. Intermittency is a special type of dynamics where almost periodic behavior is randomly interrupted. In applications, we most often encounter intermittency appearing near the fold bifurcation of some stable cycle whose Poincaré section is an odd-period cycle, most often a 3-cycle. In this case, we are at the boundary of the stable window and the dynamics exhibit so-called intermittency of type I. The transient behavior is dominated by 'ghosts' of the merged cycles. There are type II and III intermittencies associated with complex unit eigenvalues and -1.
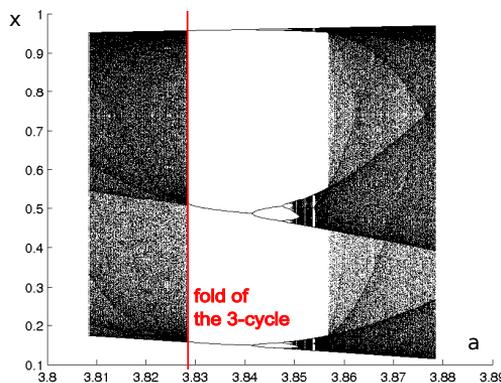
Stable 3-cycle of the logistic map $x \to ax(1-x)$ for $a = 1 + \sqrt{8} + 0.0001$:
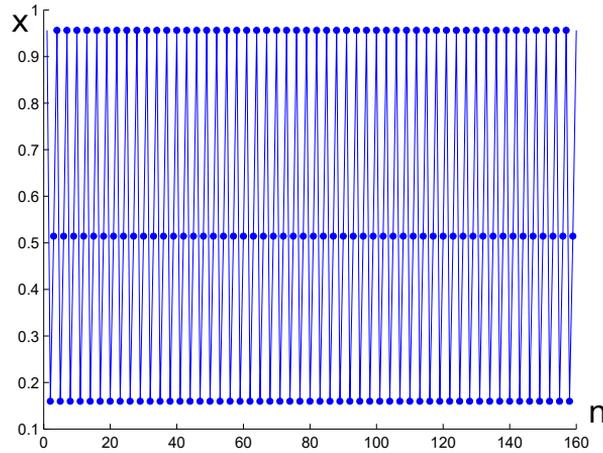
Intermittency in the neighbourhood of an unstable 3-cycle of the logistic map $x \to ax(1-x)$ for $a = 1 + \sqrt{8} - 0.0001$:
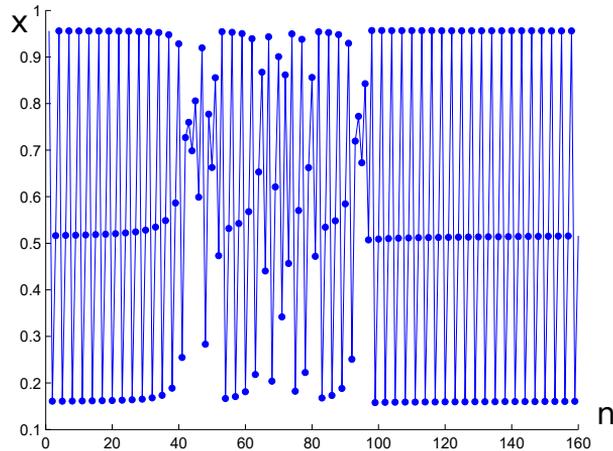
Intermittency appears in models of mechanical systems as irregular behavior that very often needs to be removed. For example, in robotics, stabilization possibilities of biped robot walking [40] are analyzed using models that exhibit both period-doubling and intermittency routes to chaos [12].

4.2   Author's contribution to the field

Some of my papers are related to chaotic dynamics and bifurcations that lead to it. One of them is a system that can be interpreted as a macroeco-

nomic model with foreign capital investment introduced in [37] or an idealized economic model in [2]. In [24], I studied the route to the chaos in this system through the period doubling of a cycle together with other bifurcations as generalized Hopf bifurcation, Bogdanov-Takens, and of course period doubling of a cycle. Moreover, I was able to detect the unstable three-cycle in the chaotic area that (according to theory, Sharkovski or Yorke theorems) implies the existence of full chaos. I also measured chaos using the maximal Lyapunov exponents and by a method of 0-1 test for chaos to confirm the results of bifurcation analysis.

The bifurcation analysis thus revealed flaws in the conclusion of the original paper [37] that claims that if the capital inflow/savings ratio is less than double the ratio of capitalized profit then the system is in a stable state. Even for common economic parameters and for common values of capital-output ratio $v \ll 1$, the first Lyapunov coefficient is positive for a quite big range of parameters, there can exist trajectories corresponding to unstable trade cycles that may even change into non-periodic bounded chaotic unpredictable regime.

Another macroeconomic application was a generalization of the neo-classical one-sector Solow-Swan model and its modification of the Kaldor-Pasinetti type with a non-constant labor growth rate. Bistability due to a fold bifurcation in the continuous-time model may lead to a switch in capital-labor ratio published in [25]. Discrete-time models reveal more complex and even chaotic dynamics for Beverton-Holt and logistic type modeling of the labor dynamics. We found that instabilities and complex dynamics may be driven by the capital-labor ratio dynamics, labor dynamics, or both together. Usually saving rate is crucial for the stabilization of the capital dynamics. Complex behavior may arise in the case that shareholders saving rate exceeds the workers saving rate, moreover, it does not depend on how high or low the rates are, so for the stabilization policy we should differentiate them. There may be also cases of instabilities (cycles or chaos) such that saving rates cannot influence the dynamics at all, since it is based on the behavior of the labor.

Another part of my research is theoretical. Most of the work is in progress till now, but some results were already presented and published in Springer Proceedings in Complexity [26]. It shows that the common bifurcations in real systems can be viewed more generally in the complex domain. This insight is particularly useful when applied to bifurcations of limit cycles, not only for a narrow group of period doubling, but in general for the m-fold associated with Arnold tongues and synchronization. Supplementary worksheets were published on Maplesoft Application Center `https://www.maplesoft.com/Applications/Author.aspx?id=357540`.

Chapter 5

Synchronization

Synchronization is generally understood as the adaptation of the dynamical behavior of objects based on some interaction between them. Christiaan Huygens is considered to be the first scientist to observe and describe this phenomenon. In a nautical experiment to determine longitude, he found that a pair of pendulum clocks suspended from a common wooden beam synchronized their swings. And it was Christiaan Huygens who contributed to accurate time-measuring by using the pendulum clock, which he patented.

The simplest type of synchronization is the synchronization of two oscillators. Driven oscillators are being used for a very long time and all of us have tried such a synchronized system on a swing, swinging a child. This is a case of an oscillator driven by an external periodic force. The external periodic force (driver or master) affects the frequency of another oscillator (slave). Except swinging a child on a swing, let's mention circadian rhythms of living organisms regulated by the rotation of the Earth, the beat of the heart rate controlled by a pacemaker, or various electrotechnical equipment to mention but a few. Huygens's clocks are a case of mutual coupling, that is a pair of coupled oscillators that influence each other. Another example of synchronization is the synchronized dynamics in a network of coupled neurons in the heart or brain. It is known that also stimulation of a group of non-interacting nearly identical oscillators by external noise can lead to their synchronization [23].

From the perspective of bifurcation theory, synchronization or partial synchronization of a group of oscillators can be understood as a type of bifurcation on a multidimensional torus that breaks it into a lower dimension, even into a limit cycle. Limit cycles and tori in continuous systems are so frequent that we can say that such systems are almost ubiquitous. The applications are significant in physics (astronomy, cosmology, elec-

tromagnetism, optics), geophysics, evolutionary and population biology, hydrology, climatology, neuroscience (neuronal excitations, bursting, neuronal synchronization), biochemistry (autocatalytic reactions in cells, feedback systems, gene expression, cell cycles, circadian rhythms), engineering (aeroelastic phenomena, structural engineering), computer science and numerical methods, economics and epidemiology (endogenous cycles), etc.

The simplest bifurcation that allows the birth of a torus from a cycle is the Neimarck-Sacker bifurcation of a limit cycle. It is originally a discrete bifurcation of a fixed point associated with the existence of a unit complex conjugated pair of eigenvalues, in this case, the fixed point of the Poincaré map of the limit cycle. Similarly to the Hopf bifurcation, it gives birth to a new invariant set around the equilibrium. Dynamics and bifurcations on and near a torus are much more complicated but can be understood using the Poincaré section, maps on a circle, transformation into a complex domain, etc. That's why it is important to put these bifurcations and related phenomena in the context of discrete dynamics. This is because the Poincaré section of an attractor of a continuous system is an attractor of a related discrete system, and so the changes in discrete dynamics can represent typical changes of continuous attractors. In the case of the emergence of chaotic trajectories, they are even used to distinguish the typical routes to chaos.

5.1   Author's contribution to the field

Systems, where a bifurcation theory approach is useful for understanding synchronizations, are externally driven oscillators in the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, A\sin\omega t, \boldsymbol{\varepsilon}), \qquad (5.1)$$

where similarly to (1.1) $\mathbf{f} : \mathbb{R}^{m+p+1} \to \mathbb{R}^m$ is a smooth enough function and $\boldsymbol{\varepsilon} \in \mathbb{R}^p$ are given parameters and the term $A\sin\omega t$ is a harmonic external force. If the values of amplitude A and the frequency ω are suitable, the external oscillation forces the oscillator with some other natural frequency[1] to oscillate synchronously - or in some rational ratio. The explanation of this phenomenon is possible by fold bifurcation of cycles on the torus.

The system (5.1) can be embedded into the autonomous system

---

[1]We assume that the system without external driving force exhibits stable oscillatory dynamics itself since we want to explain the origin of synchronized oscillations.

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, u_2, \boldsymbol{\varepsilon}),$$
$$\dot{u}_1 = \omega\big(\mu u_1 - u_2 - u_1(u_1^2 + u_2^2)\big), \tag{5.2}$$
$$\dot{u}_2 = \omega\big(u_1 + \mu u_2 - u_2(u_1^2 + u_2^2)\big),$$

where the added 2-dimensional driving system for $\mathbf{u} = (u_1, u_2)^T$ is a normal form of a supercritical Hopf bifurcation with a stable equilibrium at the origin for negative μ and with a stable limit cycle $S = \{(u_1, u_2) : u_1^2 + u_2^2 = \mu\}$ for positive μ. So for $\mu = A^2$, the dynamics of (5.2) on the invariant manifold $\mathbb{R}^m \times S$ are identical (topologically equivalent) to the dynamics of the system (5.1). Moreover, the asymptotic stability of the cycle S implies good numerical properties for the continuation of bifurcation manifolds of the system. In the driving system, the artificial parameter μ gives birth to a stable limit cycle at zero, but in the system (5.2), the manifold $\mu = 0$ gives birth to a torus since the natural oscillations exist in the original system. The torus birth is also called the Neimark-Sacker bifurcation since the related Poincaré map undergoes this bifurcation. Synchronization of the original oscillations and the driving oscillations is born here and is present in stable regions of so-called Arnold tongues. Boundaries of these regions are fold bifurcation of the limit cycles manifolds that live on the invariant torus. Stable and unstable cycles merge on the invariant torus at the limit point manifolds. The stable cycle is an attractor of the system and it has a special property – the ratio of frequency ω of the driving system and the frequency of the driven system (5.1) is exactly rational. The reason is deeper and is related to the discrete dynamics on a circle. The Arnold tongues are born at so-called resonances at the Neimark-Sacker manifold that are cusps of the fold branches in their neighborhood.

A similar approach can be used for systems that arise in Josephson junctions modeling[2]. The model equations in the form of an autonomous system are given by

$$\dot{V} = I - \sin\varphi - \beta V + A\sin\theta,$$
$$\dot{\varphi} = V, \tag{5.3}$$
$$\dot{\theta} = \omega,$$

where φ is the phase difference between the macroscopic wave functions of the two superconducting layers (S-layers) and V is the voltage across the insulating barrier (I-layer) of the S-I-S layers forming the junction.

---

[2]Josephson junctions generate stable voltages that depend only on an applied frequency and fundamental constants, so it is used as the most accurate method to generate or measure voltage. By international agreement in 1990, it is the basis for voltage standards around the world. Josephson junctions are widely used in science and engineering.

More generally we can write such a system as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{A}\sin\mathbf{y}, \boldsymbol{\varepsilon}),$$
$$\dot{\mathbf{y}} = \mathbf{g}(\mathbf{x}, \boldsymbol{\varepsilon}),$$

(5.4)

where $\mathbf{f}$ and $\mathbf{g}$ are smooth enough functions and $\boldsymbol{\varepsilon}$ are given parameters and part of the system is driven by harmonic and dynamically coupled oscillations. You can see that such a system is a generalization of (5.1) too. Analogously to the previous method used for externally driven oscillators (that can be also viewed as a special case of this type), we can embed the system into an autonomous system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{v}, \boldsymbol{\varepsilon}),$$
$$\dot{\mathbf{u}} = \left(\boldsymbol{\mu}\mathbf{u} - \mathbf{v} - \mathbf{u}(\mathbf{u}^2 + \mathbf{v}^2)\right) \cdot \mathbf{g}(\mathbf{x}, \boldsymbol{\varepsilon}),$$
$$\dot{\mathbf{v}} = \left(\mathbf{u} + \boldsymbol{\mu}\mathbf{v} - \mathbf{v}(\mathbf{u}^2 + \mathbf{v}^2)\right) \cdot \mathbf{g}(\mathbf{x}, \boldsymbol{\varepsilon}),$$

(5.5)

where the dot product $\cdot$ is understood as element-by-element multiplication (using the chain rule).

Let's say $\mathbf{y}, \mathbf{u}, \mathbf{v} \in \mathbb{R}^k$. One can easily check that the k-dimensional torus $\{[\mathbf{x}, \mathbf{u}, \mathbf{v}] \mid u_i^2 + v_i^2 = \mu_i \text{ for } i \in \{1, \dots, k\}\}$ is an invariant set of system (5.5) for $\mu_i = A_i^2$. We can set

$$u_i = A_i \cos y_i, \quad v_i = A_i \sin y_i \quad \text{for } i \in \{1, \dots, k\}$$

or equivalently $\mathbf{u} = \mathbf{A}\cos\mathbf{y}$, $\mathbf{v} = \mathbf{A}\sin\mathbf{y}$ for any solution $(\mathbf{x}, \mathbf{y})$ of system (5.4). At the moment

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{v}, \boldsymbol{\varepsilon}), \ \dot{\mathbf{u}} = -\mathbf{v}\mathbf{g}(\mathbf{x}, \boldsymbol{\varepsilon}), \ \dot{\mathbf{v}} = \mathbf{u}\mathbf{g}(\mathbf{x}, \boldsymbol{\varepsilon})$$

is the exact dynamics on the embedded invariant k-dimensional torus and all solutions of the system (5.4) are mapped to it.

Generalization of the system (5.4) into this form and reformulation of the problem in terms of bifurcation theory gives a new possibility to study such systems semi-analytically with the support of numerical continuation software, and avoid problems with unbounded phase variables and stiffness of the system.

The Josephson junction model (5.3) can be embedded into the following system

$$\dot{V} = I - y - \beta V + w,$$
$$\dot{x} = \left(px - y - x\left(x^2 + y^2\right)\right)V,$$
$$\dot{y} = \left(x + py - y\left(x^2 + y^2\right)\right)V,$$
$$\dot{u} = \left(au - w - u\left(u^2 + w^2\right)\right)\omega,$$
$$\dot{w} = \left(u + aw - w\left(u^2 + w^2\right)\right)\omega.$$

(5.6)

We explained the Shapiro steps origin in Josephson junctions, the route to chaos inside Arnold tongues, and the bistability phenomenon in the paper [11] using this bifurcation theory approach to the generalization (5.6). The method allows analysis and stable numerical continuations in the system which is stiff in the original form.
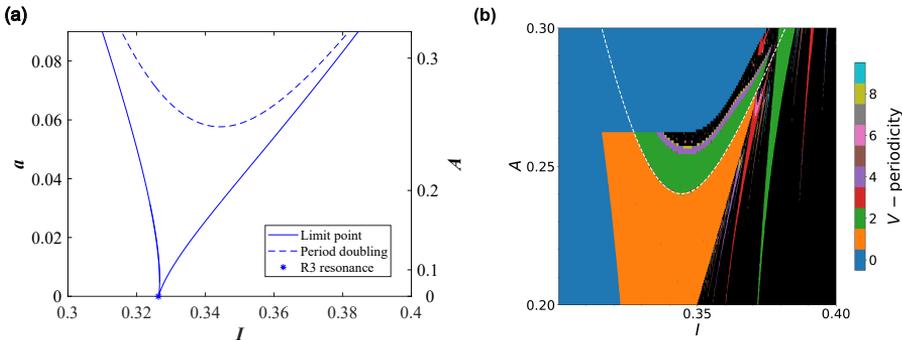

Figure 5.1: (a) Bifurcation manifolds of the system (5.6) in the (I, a) parameter space at $p = 1$, $\beta = 0.2$, and $\omega = 0.5$ (near strong 1:3 resonance). The figure shows the manifold of period-doubling bifurcation of (1,3)-cycle, which is part of the inner structure of the (1,3)-cycle Arnold tongue. (b) Periodicity in V of the system (5.3) relative to the drive cycle for parameters $\beta = 0.2$, and $\omega = 0.5$, note reparameterization $a = A^2$. The dashed white line reparameterizes the period doubling curve in (a), the black regions indicate chaos, and zero periodicity (blue) indicates the superconducting (zero average voltage) state. From [11].

Figure 5.1(b) depicts periods of V relative to the driving in the system (5.3) computed for each point of the parametric space as a period of a solution of (5.3) for fixed initial conditions. The continuation of the Arnold tongue that belongs to a fold bifurcation of a 3-cycle on an invariant torus near 1:3 resonance point is depicted at 5.1(a). You can see it nicely corresponds to 5.1(b) with reparametrization $a = A^2$ in the lower part of the Figure while the upper blue part disturbs the Arnold tongue. The blue area in Figure 5.1 belongs to a synchronized state with zero average voltage (superconducting state). This phenomenon can be explained by proximity to the Hopf-Hopf bifurcation in system (5.5). There is bistability present in the system (5.3) or (5.6), respectively. The invariant torus born from the resonance cusp point is not the only invariant stable set in the system, there is a stable torus and a stable cycle. Their basins of attraction determine the initial conditions in the state space that are attracted to one of the two attractors (so the blue area in Figure 5.1(b) inside the Arnold

tongue according to Figure 5.1(a) were computed for initial conditions in the cycle basin of attraction and not the torus basin). It is the Hopf-Hopf bifurcation [17] that give birth to bistable dynamics for some parameters – a limit cycle near zero (synchronized stable cycle colored red at Figure that belongs to zero average voltage) and a torus with higher frequency dynamics (inside the Arnold tongues, colored green at Figure). Moreover, a rigorously derived bifurcation diagram near Hopf-Hopf bifurcation (see [17]) implies the existence of an additional phenomenon, called cycle blow-up, when the torus breaks into a heteroclinic trajectory tending to a limit cycle.

To sum up the important nonlinear phenomena present in the system, there are infinitely many Arnold tongues (in the parameter space) with period doubling structures inside and routes to chaos on each torus belonging to a given Arnold tongue and there is also a bistability present and that divides the state space initial conditions into two sets with completely different dynamical behaviors.

This approach seems to be very perspective in many other areas from applications in physics to ecological or epidemiological models with seasonally varying parameters. Embedding helps to study birth and break of the torus, synchronization onset due to m-folding of cycles and also to visualize the dynamics on the torus or near the torus.

Bibliography

[1] John R Beddington. Mutual interference between parasites or predators and its effect on searching efficiency. *The Journal of Animal Ecology*, pages 331–340, 1975.

[2] Safieddine Bouali. Feedback loop in extended van der pol's equation applied to an economic model of cycles. *International Journal of Bifurcation and Chaos*, 9(04):745–756, 1999.

[3] Athanasios Chasalevris. Stability and Hopf bifurcations in rotor-bearing-foundation systems of turbines and generators. *Tribology International*, 145:106154, 2020.

[4] Chris Cosner, Donald L DeAngelis, Jerald S Ault, and Donald B Olson. Effects of spatial grouping on the functional response of predators. *Theoretical population biology*, 56(1):65–75, 1999.

[5] Robert F Costantino, RA Desharnais, Jim M Cushing, and Brian Dennis. Chaotic dynamics in an insect population. *Science*, 275(5298):389–391, 1997.

[6] Estee Y Cramer, Evan L Ray, Velma K Lopez, Johannes Bracher, Andrea Brennen, Alvaro J Castro Rivadeneira, Aaron Gerding, Tilmann Gneiting, Katie H House, Yuxin Huang, et al. Evaluation of individual and ensemble probabilistic forecasts of covid-19 mortality in the us. *Medrxiv*, 2021.

[7] Jim Michael Cushing. *An introduction to structured population dynamics*. SIAM, 1998.

[8] Jim Michael Cushing, Robert F Costantino, Brian Dennis, Robert Desharnais, and Shandelle Marie Henson. *Chaos in ecology: experimental nonlinear dynamics*, volume 1. Elsevier, 2003.

[9] Donald Lee DeAngelis, RA Goldstein, and Robert V O'Neill. A model for tropic interaction. *Ecology*, 56(4):881–892, 1975.

[10] Simona Di Nino and Angelo Luongo. Nonlinear aeroelastic in-plane be-
     havior of suspension bridges under steady wind flow. *Applied Sciences*,
     10(5):1689, 2020.

[11] Veronika Eclerová, Lenka Přibylová, and André Botha. Embedding
     nonlinear systems with two or more harmonic phase terms near the
     Hopf-Hopf bifurcation. *Nonlinear Dynamics*, 2022.

[12] Hassene Gritli, Safya Belghith, and Nahla Khraeif. Intermittency and
     interior crisis as route to chaos in dynamic walking of two biped robots.
     International Journal of Bifurcation and Chaos, 22(03):1250056, 2012.

[13] Veronika Hajnová and Lenka Přibylová. Two-parameter bifurcations in
     LPA model. *Journal of mathematical biology*, 75(5):1235–1251, 2017.

[14] Veronika Hajnová and Lenka Přibylová. Bifurcation manifolds in
     predator-prey models computed by Gröbner basis method. *Mathe-
     matical biosciences*, 312:1–7, 2019.

[15] MP Hassell, JH Lawton, and JR Beddington. The components of
     arthropod predation: I. the prey death-rate. *The Journal of Animal
     Ecology*, pages 135–164, 1976.

[16] Jitka Kühnová and Lenka Přibylová. A predator-prey model with
     allee effect and fast strategy evolution dynamics of predators using
     hawk and dove tactics. *Tatra Mountains Mathematical Publications*,
     50(1):13–24, 2011.

[17] Yuri A Kuznetsov. *Elements of Applied Bifurcation Theory*, volume
     112 of *Applied Mathematical Sciences*. Springer, New York, 2nd edi-
     tion, 1998.

[18] Jacques Laskar. Large scale chaos and marginal stability in the solar
     system. *Celestial Mechanics and Dynamical Astronomy*, 64(1):115–
     162, 1996.

[19] Jun Li and Martin Ostoja-Starzewski. Edges of Saturn's rings are
     fractal. *SpringerPlus*, 4(1):1–8, 2015.

[20] A Libchaber, C Laroche, and Stephan Fauve. Period doubling cascade
     in mercury, a quantitative measurement. *Journal de Physique Lettres*,
     43(7):211–216, 1982.

[21] Kjartan G Magnússon. Destabilizing effect of cannibalism on a struc-
     tured predator–prey system. *Mathematical biosciences*, 155(1):61–75,
     1999.

[22] Robert Marik and Lenka Pribylova. An age-structured model of cannibalism. *Electronic Journal of Differential Equations (EJDE)[electronic only]*, 2006:Paper–No, 2006.

[23] Alexander B Neiman and David F Russell. Synchronization of noise-induced bursts in noncoupled sensory neurons. *Physical review letters*, 88(13):138103, 2002.

[24] Lenka Pribylova. Bifurcation routes to chaos in an extended van der pol's equation applied to economic models. *Electronic Journal of Differential Equations (EJDE)[electronic only]*, 2009:Paper–No, 2009.

[25] Lenka Pribylová. The Solow-Swan model generalization with non-constant labor growth rate. *Bullet Czech Econo Soc*, 18(28):20–32, 2011.

[26] Lenka Přibylová. Bifurcations of one-dimensional one-parametric maps revisited. In *Chaotic Modeling and Simulation International Conference*, pages 215–227. Springer, 2018.

[27] Lenka Přibylová. Regime shifts caused by adaptive dynamics in prey-predator models and their relationship with intraspecific competition. *Ecological Complexity*, 36:48–56, 2018.

[28] Lenka Přibylová and Luděk Berec. Predator interference and stability of predator–prey dynamics. *Journal of mathematical biology*, 71(2):301–323, 2015.

[29] Lenka Přibylová and Barbora Losová. Symmetry breaking for GATA-1/PU. 1 model. In *International Conference on Computational Methods in Systems Biology*, pages 360–363. Springer, 2019.

[30] Lenka Přibylová and Anna Peniašková. Foraging facilitation among predators and its impact on the stability of predator–prey dynamics. *Ecological Complexity*, 29:30–39, 2017.

[31] Michael L Rosenzweig. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. *Science*, 171(3969):385–387, 1971.

[32] Steven J Schiff, Kristin Jerger, Duc H Duong, Taeun Chang, Mark L Spano, and William L Ditto. Controlling chaos in the brain. *Nature*, 370(6491):615–620, 1994.

[33] Jan Ševčík and Lenka Přibylová. Forced van der Pol Oscillator – synchronization from the bifurcation theory point of view. In *Chaotic Modeling and Simulation International Conference*, pages 411–420. Springer, 2022.

[34] Himanshu Shukla and Mayuresh J Patil. Nonlinear state feedback control design to eliminate subcritical limit cycle oscillations in aeroelastic systems. *Nonlinear Dynamics*, 88(3):1599–1614, 2017.

[35] Michael Small, Dejin Yu, and Robert G Harrison. Observation of a period doubling bifurcation during onset of human ventricular fibrillation. *International Journal of Bifurcation and Chaos*, 13(03):743–754, 2003.

[36] Georg AK van Voorn, Dirk Stiefs, Thilo Gross, Bob W Kooi, Ulrike Feudel, and Sebastiaan ALM Kooijman. Stabilization due to predator interference: comparison of different analysis approaches. *Mathematical Biosciences & Engineering*, 5(3):567, 2008.

[37] Miloslav Vošvrda. Bifurcation routes and economic stability. *Bulletin of the Czech Econometric Society*, 14:43–60, 2001.

[38] HN Núñez Yépez, AL Salas Brito, CA Vargas, and LA Vicente. Chaos in a dripping faucet. *European Journal of Physics*, 10(2):99, 1989.

[39] Jakub Záthurecký and Lenka Přibylová. Coupled FitzHugh-Nagumo type neurons driven by external voltage stimulation. In *Chaotic Modeling and Simulation International Conference*, pages 537–550. Springer, 2022.

[40] Wafa Znegui, Hassène Gritli, and Safya Belghith. Stabilization of the passive walking dynamics of the compass-gait biped robot by developing the analytical expression of the controlled poincaré map. *Nonlinear Dynamics*, 101(2):1061–1091, 2020.

Chapter 6

Collection of author's published articles concerning nonlinear phenomena

Articles in international journals (Q1-4 journal SCIMAGO ranking is in applied mathematics, analysis, mathematics or modeling and simulation field):

Applications in population biology:

Q2 Robert Mařík and Lenka Přibylová. An age-structured model of cannibalism. *Electronic Journal of Differential Equations*, 106, 11 p., **2006**.

Q3 Jitka Kühnová and Lenka Přibylová. A predator-prey model with Allee effect and fast strategy evolution dynamics of predators using hawk and dove tactics. *Tatra Mountains Mathematical Publications*, 50(1): 13–24, **2011**.

Q1 Lenka Přibylová and Luděk Berec. Predator interference and stability of predator-prey dynamics. *Journal of mathematical biology*, 71(2):301–323, **2015**.
Publicly shared by Springer Nature: `https://rdcu.be/cWLGW`

Q2 Lenka Přibylová and Anna Peniašková. Foraging facilitation among predators and its impact on the stability of predator-prey dynamics. *Ecological Complexity*, 29:30–39, **2017**.

Q1 Veronika Hajnová and Lenka Přibylová. Two-parameter bifurcations in LPA model. *Journal of mathematical biology*, 75(5):1235–1251, **2017**.
Publicly shared by Springer Nature: `https://rdcu.be/cWPyX`

Q2 Lenka Přibylová. Regime shifts caused by adaptive dynamics in prey-predator models and their relationship with intraspecfic competition. *Ecol. Complex.*, 36:48–56, **2018**.

Q2 Veronika Hajnová and Lenka Přibylová. Bifurcation manifolds in predator-prey models computed by Gröbner basis method. *Mathematical biosciences*, 312:1–7, **2019**.

Applications in physics:

Q1 Veronika Eclerová, Lenka Přibylová and André E. Botha. Embedding nonlinear systems with two or more harmonic phase terms near the Hopf-Hopf bifurcation *Nonlinear Dynamics*, **2022**.
Publicly shared by Springer Nature: `https://rdcu.be/cWPDc`

Applications in economics:

Q3 Lenka Baráková. Asymptotic behavior and existence of a limit cycle of cubic autonomous systems. *Demonstratio Mathematica*, 34(3): 559–576, **2001**.

Q2 Lenka Baráková and Evgenii P. Volokitin. Bifurcation diagram of a cubic three-parameter autonomous system. *Electronic Journal of Differential Equations*, 83, 16 p., **2005**.

Q2 Lenka Přibylová. Bifurcation routes to chaos in an extended Van der Pol's equation applied to economic models. *Electronic Journal of Diff. Equations*, 53, 21 p., **2009**.

– Lenka Přibylová. The Solow-Swan model generalization with non-constant labor growth rate. *Bullet Czech Econo Soc*, 18(28): 20–32, **2011**.

Conference articles in proceedings:

- Lenka Přibylová. Bifurcations of one-dimensional one-parametric maps revisited. In: *Chaotic Modeling and Simulation International Conference*, Springer, Cham, 215–227, **2018**.

- Lenka Přibylová and Barbora Losová. Symmetry Breaking for GATA-1/PU. 1 Model. In: *International Conference on Computational Methods in Systems Biology*, Springer, Cham, 360–363, **2019**.

- Jan Ševčík and Lenka Přibylová. Forced van der Pol oscillator – synchronization from the bifurcation theory point of view . In: *Chaotic Modeling and Simulation International Conference CHAOS 2021*, Springer, Cham, 411–420, **2022**

- Jakub Záthurecký and Lenka Přibylová. Coupled Fitzhugh-Nagumo Type Neurons Driven by External Voltage Stimulation. In: *Chaotic Modeling and Simulation International Conference CHAOS 2021*, Springer, Cham, 537–550, **2022**

- Veronika Eclerová, Lenka Přibylová, and André E. Botha. Transformation of master-slave systems with harmonic terms for improved stability in numerical continuation. In: *The 15th CHAOS 2022 International Conference Proceedings*, accepted, to appear **2022**

AN AGE-STRUCTURED MODEL OF CANNIBALISM

ROBERT MAŘÍK, LENKA PŘIBYLOVÁ

ABSTRACT. We investigate the predator-prey model with cannibalism in the predator population, suggested by Magnusson [5] in 1995. We explore the model by a theory of bifurcations, based mainly on the results of Bautin. Among others, we show that the limit cycle appearing in the model due to the Andronov-Hopf bifurcation may be stable or unstable.

1. INTRODUCTION

Magnusson [5] introduced the predator-prey system with age structure and cannibalism in the predator population in the form

$$
\begin{aligned}
\dot{X} &= AY - \mu_a X + \gamma SXY + VCXZ, \\
\dot{Y} &= \lambda X - AY - \mu_j Y - SXY, \\
\dot{Z} &= LZ - QZ^2 - VZX,
\end{aligned}
\tag{1.1}
$$

where X is the population of adult predators, Y the population of juvenile predators, Z the population of prey, T is the time and the dot denotes derivative with respect to T.

In model (1.1) the Lotka–Volterra type of interspecific interaction is considered. The parameters μ_a and μ_j describe the natural death rate of the adult and juvenile predators, respectively. The constant λ is the birth rate of predators and A is the rate at which juvenile predators mature into adults. The term VXZ describes the rate at which adult predators kill the prey and the constant $C \in (0,1)$ is an efficiency of conversion of sources obtained by killing the prey to the increase of the fitness of population of adult predators. In a similar way, the term SXY is the rate at which adult predators kill juvenile predators and the corresponding increase of fitness of adult predators is proportional to this term by the constant $\gamma \in (0,1)$.

Remark that the population of prey is subjected to the logistic growth in the absence of predators. Magnusson in [5] used $Q = 0$ and proved that for suitable values of parameter S an Andronov-Hopf bifurcation occurs and the increase of the cannibalism rate can destabilize the system. This result is in the same paper extended to small values of Q, i.e. for the logistic growth with large carrying capacity, when the competition in the prey population is not significant.

Later Kaewmanee and Tang [3] reexamined model (1.1) (with general Q) and obtained similar results.

As pointed out already in [2], it is usually much easier to show that the Andronov-Hopf bifurcation appears in a particular model than to distinguish whether the periodic solution is stable or whether it is unstable and does not describe a state which may really appear in nature. In both papers [3], [5], authors prove that the Andronov-Hopf bifurcation takes place, but the stability of the limit cycle is not analyzed. All authors provide numerical results showing the stable limit cycle and in this sense some of the results are rough.

The aim of this paper is to continue in the study of dynamical system (1.1) and to obtain deeper results concerning its stability, topological properties and types of bifurcations (especially stability and uniqueness of limit cycles using Lyapunov coefficients and theory developed in [1, 4]). Among others, we focus our attention to more parameters, not only to the parameter S. In order to make the computations manageable we consider $Q = 0$ (as in [5]), i.e., we suppose no intraspecific competition in the prey population.

The paper is organized as follows. The next chapter recalls some facts concerning the theory of bifurcations. The Magnusson's mathematical model is studied in Section 3. Section 4 contains some numeric results and a discussion about possible phase portraits for various values of parameters in the model. Among others, we show that an unstable limit cycle exists for suitable values of parameters.

2. Preliminaries

2.1. **Mathematical models.** Many problems of natural science can be solved using mathematical models. Some of deterministic models can be described by dynamical systems of autonomous differential equations. These systems contain studied variables and coefficients given by internal attributes of the problem or by the external conditions. Values of this type (constant according to time) are parameters, since the studied variables and behavior of the whole dynamical system depend on them. General mathematical model can be represented by the following system

$$\dot{x} = F(x, \alpha), \tag{2.1}$$

where $x \in \mathbb{R}^n$ are studied phase variables, $\alpha \in \mathbb{R}^m$ are parameters and F is sufficiently smooth.

2.2. **Bifurcation diagrams and normal forms.** Changing parameters α, the phase portrait of the system (2.1) changes. There are two possibilities of the change: the phase portrait stays topologically equivalent to the previous one or not. Generally the space of parameters can be divided into structurally stable domains with topologically equivalent phase portraits. This division together with related phase portraits is called a bifurcation diagram of the system (2.1). Structurally unstable boundaries correspond with some particular type of bifurcation that can be found by transforming the system (2.1) to its normal form. Some typical systems (already in the normal form) describing particular types of bifurcations were studied as model systems for many types of bifurcations.

2.3. **Andronov-Hopf bifurcation.** Now let us turn our attention to the bifurcation which will be proved in the age-structured model of cannibalism, the Andronov-Hopf bifurcation. First of all, recall some basic facts about this bifurcation. This

bifurcation is related to an existence of a pair of purely imaginary eigenvalues. By crossing the imaginary axis and changing the real part of eigenvalues from negative to positive value a stable focus changes to an unstable focus. This change is accompanied by a birth of a cycle. The model system for Andronov-Hopf bifurcation

$$\dot{x}_1 = \mu x_1 - x_2 - x_1(x_1^2 + x_2^2)$$
$$\dot{x}_2 = x_1 + \mu x_2 - x_2(x_1^2 + x_2^2)$$

(2.2)

possesses a unique stationary point $x_1 = 0 = x_2$ for every real μ. The eigenvalues at this stationary points are $\lambda = \mu \pm i$. Introducing complex-valued variable $z = x_1 + ix_2$ the system (2.2) takes the form $\dot{z} = (\mu + i)z - z|z|^2$ which becomes in polar coordinates to a system

$$\dot{\rho} = \rho(\mu - \rho^2)$$
$$\dot{\phi} = 1.$$

(2.3)

From here it follows that the stationary point is a stable focus for $\mu < 0$ and an unstable focus for $\mu > 0$. If $\mu > 0$ then a solution $\rho = \sqrt{\mu}$ presents a limit cycle around the origin.
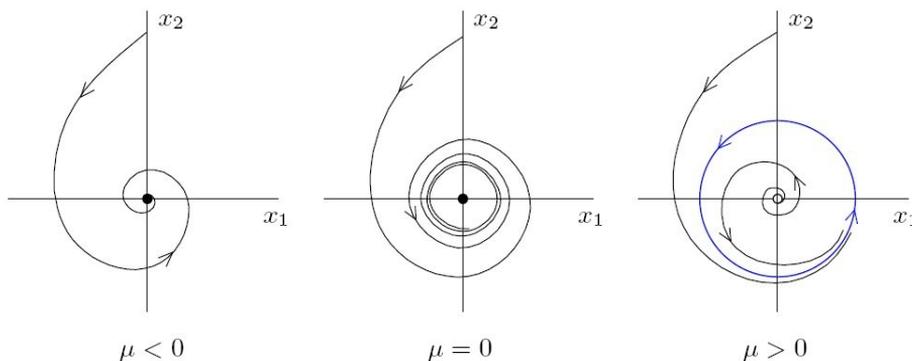

FIGURE 1. Hopf-Andronov bifurcation.

An important problem is whether the limit cycle stable or unstable. One of the concepts which enables us to decide which of these two possibilities really arises is the concept of Lyapunov coefficient l_1, shortly introduced in the following paragraphs.

Lemma 2.1. *Consider general one-parameter system*

$$\dot{x} = f(x, \mu), \quad x \in \mathbb{R}^2, \quad \mu \in \mathbb{R}$$

(2.4)

which possesses a stationary point $x = 0$ for every sufficiently small $|\mu|$ and let

$$\lambda_{1,2}(\mu) = \psi(\mu) \pm i\omega(\mu),$$

where $\psi(0) = 0$, $\omega(0) = \omega_0 > 0$, be the eigenvalues of this stationary point.

There exists an invertible transformation depending on the parameter μ which converts system (2.4) into the complex form

$$\dot{z} = \Big(\psi(\mu) + i\omega(\mu)\Big)z + c_1(\mu)z|z|^2 + O(|z|^4)$$

Remark 2.2 (stability of limit cycle in Andronov-Hopf bifurcation)**.** The number

$$l_1(\mu) = \frac{\operatorname{Re} c_1(\mu)}{\omega(\mu)} - \psi(\mu)\frac{\operatorname{Im} c_1(\mu)}{\omega^2(\mu)}$$

is called the *first Lyapunov coefficient*. The sign of $l_1(0)$ determines stability or instability of the limit cycle. A stable limit cycle appears near the equilibrium in the case $l_1(0) < 0$ for μ close to zero. This process is referred as "supercritical bifurcation", since an equilibrium at stationary point is replaced by small oscillations. On the contrary, if $l_1(0) > 0$, then "subcritical bifurcation" takes place, since the stable equilibrium enclosed by an *unstable* limit cycle changes to an unstable focus.

Bautin [1] introduced the terms "subcritical" and "supercritical" bifurcation and derived formulas for the first Lyapunov coefficient for systems of two and three equations with analytical coefficients. These formulas are too long to repeat them on this place. For some special systems these formulas take simpler forms than presented in [1].

Another approach which enables to distinguish between the stable and unstable limit cycle is presented in [2]. In this book the first Lyapunov coefficient is replaced by the quantity μ_2 which is the first nonvanishing term of an infinite series, see [2, Chapter 1] for more details.

Finally, recall that the period of the limit cycle approaches $\frac{2\pi}{\omega(0)}$ as the parameter approaches the critical value.

2.4. **Shoshitaishvili theorem.**

Lemma 2.3. *Let for the system* (2.1) *be $\alpha \in \mathbb{R}^1$ and $x = 0$ be a stationary point for $\alpha = 0$ with $n_0 \neq 0$ purely imaginary eigenvalues. Then there exists a local invariant manifold $W^c(\alpha)$ in the neighbourhood of zero. The manifold is attractive, if all other eigenvalues have positive real parts.*

The system (2.1) *can be restricted to $W^c(\alpha)$ for small $|\alpha|$ by a local projection of $W^c(\alpha)$ to T^c (generalized eigenspace corresponding to the union of purely imaginary eigenvalues). On this so-called central manifold, the system* (2.1) *can be represented in the new coordinates by a system*

$$\dot{u} = \Phi(u, \alpha), \quad u \in \mathbb{R}^{n_0}.$$

Theorem 2.4 (Shoshitaishvili theorem)**.** *The system* (2.1) *is locally topologically equivalent to a suspension of the central manifold, which can be represented by a system*

$$\dot{u} = \Phi(u, \alpha),$$
$$\dot{y} = -y,$$
$$\dot{z} = z,$$

where $u \in \mathbb{R}^{n_0}$, $y \in \mathbb{R}^{n_-}$ and $z \in \mathbb{R}^{n_+}$ ($n_\pm$ is the number of eigenvalues with positive or negative real parts).

For the proof of the above theorem, see [6].

3. Mathematical model

Let us return our attention to the system (1.1). First of all, we non-dimensionalise the system and obtain a smaller amount of parameters. Remark that we are interested especially in the interspecific parameters V and S and from this reason we perform the transformation which preserves these parameters.

Lemma 3.1. *Transformation $X = \frac{A}{\gamma}x$, $Y = \frac{A}{\gamma^2}y$, $Z = \frac{A}{C\gamma}z$, $T = \frac{\gamma}{A}t$ transforms the dynamical system (1.1) with $Q = 0$ into*

$$\dot{x} = y - mx + Sxy + Vxz,$$
$$\dot{y} = nx - oy - Sxy, \tag{3.1}$$
$$\dot{z} = pz - Vxz,$$

where the dot, $\cdot$ represents $\frac{d}{dt}$, $m = \frac{\mu_a \gamma}{A} > 0$, $n = \frac{\lambda \gamma^2}{A} > 0$, $o = \gamma\left(1 + \frac{\mu_j}{A}\right) > 0$, $p = \frac{L\gamma}{A} > 0$.

For $o \neq 1$ and $m \neq n$, System (3.1) possesses three stationary points $S_0 = [x_0, y_0, z_0]$, $S_1 = [x_1, y_1, z_1]$ and $S_2 = [x_2, y_2, z_2]$, where

$$x_0 = \frac{p}{V}, \quad y_0 = \frac{np}{oV + pS}, \quad z_0 = \frac{(mo - n)V + p(m - n)S}{V(oV + pS)}, \tag{3.2}$$

$$x_1 = \frac{om - n}{S(n - m)}, \quad y_1 = \frac{1}{S} \cdot \frac{om - n}{o - 1}, \quad z_1 = 0, \tag{3.3}$$

$$x_2 = y_2 = z_2 = 0. \tag{3.4}$$

The statements of this lemma follow immediately.

From a practical point of view, the most interesting stationary point is the point S_0.

The absence of prey in the stationary point S_1 is explained in [3] and [5] as the existence of an alternative food for the predator. This alternative food is not incorporated in our model. However, the existence of an alternative food is not necessary for real interpretation of the system. Particularly, there are lakes in nature with predator fishes only. In such ecosystems, the juvenile predators are the major food for adults.

In fact, the set $z = 0$ is an invariant set of the system. In this set system (3.1) becomes

$$\dot{x} = y - mx + Sxy,$$
$$\dot{y} = nx - oy - Sxy \tag{3.5}$$

and describes the states with no prey. However, the main aim is to study the steady states in which all populations are present. Therefore we focus our attention to the stationary point S_0.

3.1. **Stationary point S_0.** We will focus our attention to the stationary point S_0. This point is in the first octant if

$$(mo - n)V + p(m - n)S > 0. \tag{3.6}$$

This condition covers three mutually different cases

(1) $mo - n > 0$, $m - n > 0$ and V, S are arbitrary positive numbers
(2) $mo - n > 0$, $m - n < 0$ and $V > \frac{S(n-m)p}{mo-n}$
(3) $mo - n < 0$, $m - n > 0$ and $V < \frac{S(n-m)p}{mo-n}$

The transformation

$$x' = x - x_0,$$
$$y' = y - y_0,$$
$$z' = z - z_0$$

moves the stationary point S_0 into the origin in new coordinates. In the new variables $[x', y', z']$ system (3.1) reads as

$$\dot{x}' = y' - mx' + S(x'y' + x'y_0 + y'x_0) + V(x'z' + x'z_0 + x_0z'),$$
$$\dot{y}' = nx' - oy' - S(x'y' + x'y_0 + x_0y'), \tag{3.7}$$
$$\dot{z}' = pz' - V(x'z' + x'z_0 + x_0z'),$$

where dot is the derivative with respect to t. The Jacobi matrix of system (3.7) evaluated at the origin is

$$
J(0,0,0) = \begin{pmatrix} -m + Sy_0 + Vz_0 & 1 + Sx_0 & Vx_0 \\ n - Sx_0 & -o - Sx_0 & 0 \\ -Vz_0 & 0 & p - Vx_0 \end{pmatrix}
$$
$$
= \begin{pmatrix} -\frac{y_0}{x_0} & 1 + Sx_0 & Vx_0 \\ o\frac{y_0}{x_0} & -n\frac{x_0}{y_0} & 0 \\ -Vz_0 & 0 & 0 \end{pmatrix}
\tag{3.8}
$$

and the characteristic polynomial of this matrix is

$$
|J(0,0,0) - \nu I| = \begin{vmatrix} -\frac{y_0}{x_0} - \nu & 1 + Sx_0 & Vx_0 \\ o\frac{y_0}{x_0} & -n\frac{x_0}{y_0} - \nu & 0 \\ -Vz_0 & 0 & -\nu \end{vmatrix} = -(\nu^3 + A\nu^2 + B\nu + C), \tag{3.9}
$$

where

$$A = \frac{y_0}{x_0} + n\frac{x_0}{y_0},$$
$$B = n - o\frac{y_0}{x_0}(1 + Sx_0) + V^2 x_0 z_0,$$
$$C = V^2 n \frac{x_0^2 z_0}{y_0}.$$

Consider the most important case, when the point S_0 is in the interior of the first octant. In this case clearly $A > 0$ and $C > 0$. Let us examine the expression

$AB - C$. An evaluation shows that

$$
\begin{aligned}
AB - C &= \Big(\frac{y_0}{x_0} + n\frac{x_0}{y_0}\Big)\Big(n - o\frac{y_0}{x_0}(1 + Sx_0) + V^2 x_0 z_0\Big) - V^2 n\frac{x_0^2 z_0}{y_0} \\
&= \frac{y_0}{x_0}\Big(n - o\frac{y_0}{x_0}(1 + Sx_0) + V^2 x_0 z_0\Big) + n^2\frac{x_0}{y_0} - no(1 + Sx_0) \\
&= \Big[1 + n\Big(\frac{x_0}{y_0}\Big)^2\Big]\Big[n - o\frac{y_0}{x_0} - Soy_0\Big]\frac{y_0}{x_0} + V^2 y_0 z_0 \\
&= \Big[1 + n\Big(\frac{x_0}{y_0}\Big)^2\Big]\Big[Sy_0 - Soy_0\Big]\frac{y_0}{x_0} + V^2 y_0 z_0 \\
&= \Big[1 + n\Big(\frac{x_0}{y_0}\Big)^2\Big]\frac{y_0^2}{x_0}S(1 - o) + V^2 y_0 z_0 \\
&= \frac{1}{x_0}\Big[(y_0^2 + nx_0^2)S(1 - o) + Vpy_0 z_0\Big].
\end{aligned}
\tag{3.10}
$$

Hence if $o < 1$, then $AB - C$ is positive. In this case all real parts of eigenvalues are negative and the stationary point S_0 is stable.

The only possibility which allows the point S_0 become to be unstable is for $o > 1$. Under this condition the point S_0 lies in the first octant if and only if either

$$mo > m > n,$$

or

$$mo > n > m \quad \text{and} \quad V > Sp\frac{n - m}{mo - n}.$$

3.2. **Case $o > 1$.** We have to write the expression $AB - C$ in terms of parameters of system (3.1) to obtain correct conclusions concerning the influence of parameters on stability. (Remember that all x_0, y_0 and z_0 depend on the parameters and hence (3.10) cannot be used to obtain correct conclusions. This step seems to be omitted in [3].) A direct computation shows that the expression $AB - C$ can be written in the form of the product of a positive factor and the three-degree polynomial $P(\cdot)$ in variable $\frac{V}{S}$ as follows:

$$AB - C = \frac{pnS^3}{V(oV + pS)^2}P\Big(\frac{V}{S}\Big), \tag{3.11}$$

where

$$P(\tau) = (mo-n)\tau^3 + \big[p(m-n)+(1-o)(n+o^2))\big]\tau^2 + 2op(1-o)\tau + (1-o)p^2. \tag{3.12}$$

According to the Descarte's rule of signs, the polynomial $P(\tau)$ possesses a unique positive zero. An evaluation shows that $P\big(p\frac{n-m}{mo-n}\big)$ is negative. Really

$$
\begin{aligned}
P\Big(p\frac{n - m}{mo - n}\Big) &= p^3\frac{(n - m)^3}{(mo - n)^2} + \Big[p(m - n) + (1 - o)(n + o^2)\Big]\frac{(n - m)^2}{(mo - n)^2}p^2 \\
&\quad + 2op(1 - o)p\frac{n - m}{mo - n} + (1 - o)p^2 \\
&= (1 - o)p^2\Big[(n + o^2)\frac{(n - m)^2}{(mo - n)^2} + 2o\frac{n - m}{mo - n} + 1\Big] \\
&= (1 - o)p^2\Big[n\frac{(n - m)^2}{(mo - n)^2} + \Big(o\frac{n - m}{mo - n} + 1\Big)^2\Big].
\end{aligned}
\tag{3.13}
$$

We summarize the above computations in the following theorem.

Theorem 3.2. *Let either $mo > m > n$ or $mo > n > m$ hold. Denote by τ_k the unique positive zero of the polynomial* (3.12).

  (1) *If $\frac{V}{S} > \tau_k$, then all real parts of eigenvalues of Jacobi matrix at S_0 are negative and S_0 is a stable singular point in the first octant.*
  (2) *If either*

$$mo > m > n \quad and \quad \frac{V}{S} < \tau_k$$

  or

$$mo > n > m \quad and \quad p\frac{n-m}{mo-n} < \frac{V}{S} < \tau_k,$$

  then at least one of the eigenvalues of Jacobi matrix at S_0 has a positive real part and S_0 is an unstable singular point in the first octant.
  (3) *If $mo > n > m$ and $\frac{V}{S} < p\frac{n-m}{mo-n}$, then S_0 is a singular point outside the first octant.*

Proof. From the assumptions it follows that $o > 1$. From equality (3.13) it follows that $\tau_k > p\frac{n-m}{mo-n}$. Now Theorem follows from (3.11), (3.13) and from the well known Ruth–Hurwitz criterion. □

Remark 3.3. Note that if $o < 1$, then $AB - C$ is always positive and the point S_0 is stable. No bifurcation occurs if $o < 1$.

If $n > mo > m$, then $z_0 < 0$ for all positive values of V and S and the point z_0 lies outside the first octant.

If $mo > n > m$, then $\frac{V}{S} > p\frac{n-m}{mo-n}$ is a necessary and sufficient condition for $\min(x_0, y_0, z_0) > 0$.

4. Andronov-Hopf bifurcation of the system (3.1)

In this section we consider system (3.1) with positive parameters m, n, p, S, V and $o > 1$ (according to the previous remark). We make a natural assumption $\min(x_0, y_0, z_0) > 0$.

Corollary 4.1. *Let τ_k be the zero of the polynomial* (3.12) *for arbitrary fixed positive parameters m, n, o, p, S. Then $V_k = \tau_k S$ is the critical value of Andronov-Hopf bifurcation of the system* (3.1) *for the stationary point S_0.*

Proof. Follows immediately from the Theorem (3.2), since the characteristic polynomial (3.9) has two purely imaginary eigenvalues, if $AB - C = 0$ (notice that $A > 0$ and $C > 0$), that is while $P(\tau) = 0$. □

Remark 4.2. Both the supercritical and the subcritical bifurcation types occur in the system (3.1). The type of Andronov-Hopf bifurcation is determined by the first Lyapunov coefficient l_1, as Remark 2.2 shows. This coefficient can be calculated numerically for given critical parameters.

For example, let $m = 7$, $n = 3$, $o = 3$, $p = 5$ and $S = 1$. The polynomial (3.12) has a unique zero $\tau_k \doteq 2.251$ and $V_k = \tau_k$ is a critical value of Andronov-Hopf bifurcation. Following [1, Chapter III], we transfer the stationary point S_0 to the origin and find that $l_1 \doteq -1.683\pi$. In this case the supercritical bifurcation takes place and the existence of a stable limit cycle in the neighborhood of an unstable focus for $V < V_k$ near V_k is proved. Figure 2 shows parts of two ω-limit trajectories (with initial conditions $[x(0) = 10, y(0) = 1, z(0) = 14]$ and $[x(0) = 1, y(0) = 1.5, z(0) = 2]$) converging to the central manifold for $V = 1.9$. The limit cycle lies

on the central manifold in the free space between them. This free space arose due to stoping calculating the trajectories.
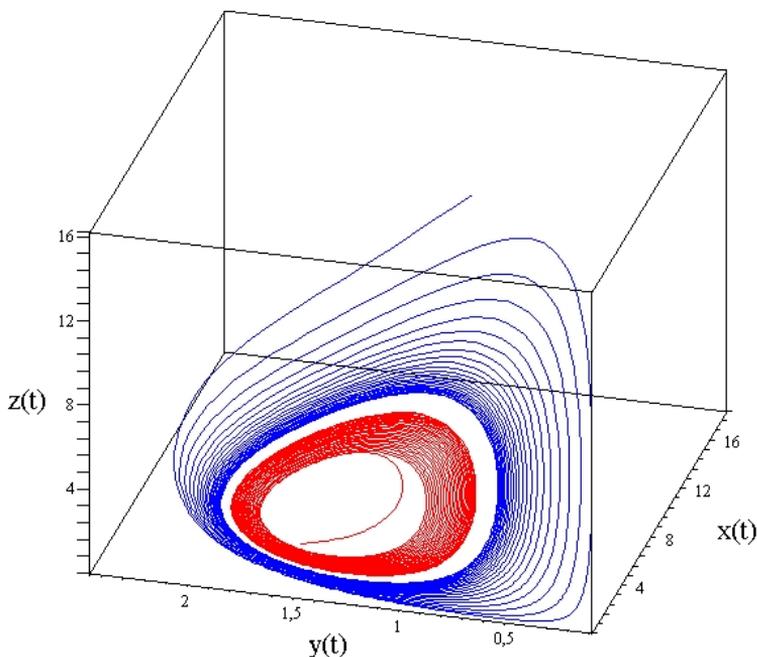

FIGURE 2. A stable limit cycle.

On the other hand, let $m = 4$, $n = 3$, $o = 3$, $p = 5$ and $S = 1$. The polynomial (3.12) has a unique zero $\tau_k \doteq 4.079$ and $V_k = \tau_k$ is a critical value of Andronov-Hopf bifurcation. Calculating the first Lyapunov coefficient for V_k we get $l_1 \doteq 0.317\pi$, so the subcritical bifurcation takes place. In this case, there exists an unstable limit cycle in the neighborhood of a stable focus for $V > V_k$ near V_k. Figure 3 shows parts of two trajectories for $V = 5.5$. One trajectory (with initial conditions $[x(0) = 2, y(0) = 1, z(0) = 2]$) is converging to the stable focus S_0 on the central manifold and the other is a stable separatrix (with initial conditions $[x(0) = 2, y(0) = 2, z(0) = 0]$, $z = 0$ is an invariant set) of the saddle point in the origin. These trajectories are very close to each other and they are converging to the central manifold, so we can suppose that the unstable limit cycle is somewhere in between on the central manifold (unlike the planar phase space, in the 3d phase space it is very hard to find the unstable limit cycle by drawing trajectories).

In view of the above computations, we can see that another type of bifurcation must take place here. Changing parameter m from 7 to 4 and counting the critical value V_k we got two qualitatively different phase portraits. You can verify that for $m \doteq 4.254$, the critical value $V_k \doteq 3.792$ and $l_1 = 0$. The figure 4 shows the trajectories near the center (initial conditions $[x(0) = 2, y(0) = 2, z(0) = 2]$, $[x(0) =$
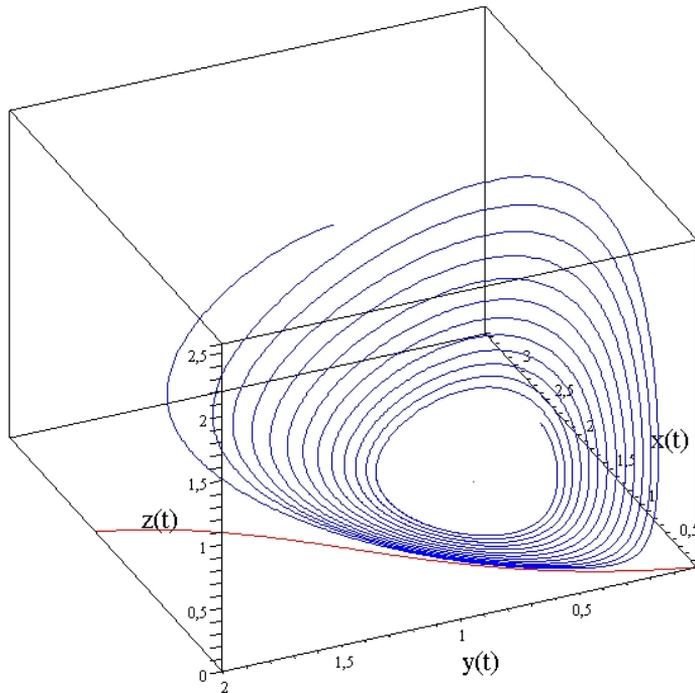
FIGURE 3. An unstable limit cycle.

$2, y(0) = -2, z(0) = 2]$ $[x(0) = 2, y(0) = -2, z(0) = 1]$, $[x(0) = 2, y(0) = 1, z(0) = 2]$, $[x(0) = 2, y(0) = 2, z(0) = 0]$).

   This is so called generalized Hopf (Bautin) bifurcation and more limit cycles can arise in the neighbourhood of the stationary point S_0.

Conclusion. We investigated the system introduced by Magnusson [5] as a model of cannibalism. We derived the conditions under which the Andronov-Hopf bifurcation takes place. These conditions are derived in terms of the parameters of the original model and among others, we provide an equation which describes the critical value for which the bifurcation occurs. Using theory of Bautin we proved that both subcritical and supercritical bifurcations may take place in this model and hence the limits cycle enclosing the stationary point need not to be stable. This phenomenon was not mentioned in any of previous works on this system.

References

[1] N. N. Bautin: *Povedenie dinamicheskih sistem vblizi granic oblasti ustoychivosti*, OGIZ GOS-TEXIZDAT, Leningrad-Moskva, 1949
[2] B. D. Hassard, N. D. Kazarinoff, Y.–H. Wan: *Theory and applications of Hopf bifurcation*, Cam. Univ. Press (1981), russian transl. Moskva, Mir (1985).
[3] C. Kaewmanee, I. M. Tang: Cannibalism in an age-structured predator-prey system, Ecol. Mod. 167 (2003), 213–220.

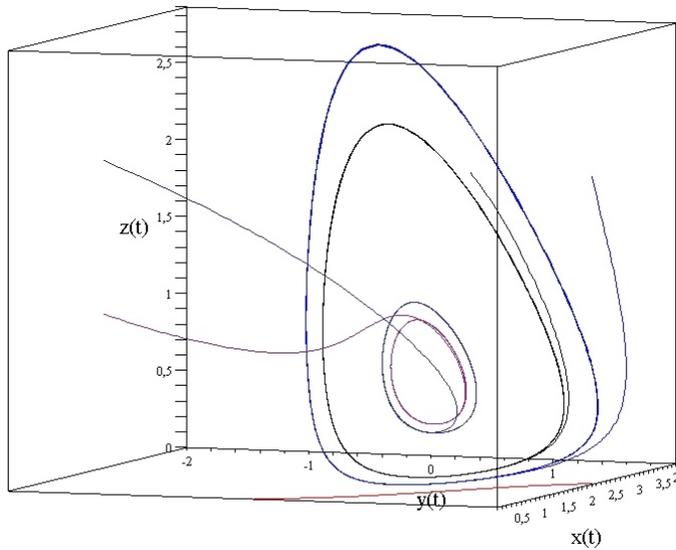FIGURE 4. Phase portrait of the center.

[4] Lyapunov: *Povedenie dinamicheskih sistem vblizi granic oblasti ustoychivosti*, OGIZ GOS-TEXIZDAT, Leningrad-Moskva, 1949

[5] K. G. Magnusson: Destabilizing efect on cannibalism on a structures predator-prey system, Math. Biosci. 155 (1999), 61–75.

[6] A. N. Shoshitaishvili; *Bifurkacii topologicheskovo tipa vektornovo polya vblizi osoboy tochki*, Trudy seminara im. I.G. Petrovskovo, No. 1 (1975) 279–309.

ROBERT MAŘÍK

DEPT. OF MATHEMATICS, MENDEL UNIVERSITY, ZEMĚDĚLSKÁ 3, 613 00 BRNO, CZECH REPUBLIC
  E-mail address: marik@mendelu.cz

LENKA PŘIBYLOVÁ

DEPT. OF APPLIED MATHEMATICS, MASARYK UNIVERSITY, JANÁČKOVO NÁM. 2A, 662 95 BRNO, CZECH REPUBLIC
  E-mail address: pribylova@math.muni.cz

VERSITA

A PREDATOR-PREY MODEL WITH ALLEE EFFECT AND FAST STRATEGY EVOLUTION DYNAMICS OF PREDATORS USING HAWK AND DOVE TACTICS

Jitka Kühnová — Lenka Přibylová

ABSTRACT. In this work we present the predator-prey model with Allee effect and Hawk and Dove tactics in fighting over caught prey implemented as fast strategy evolution dynamics. We extend the work of Auger, Parra, Morand and Snchez (2002) using the prey population embodying Allee effect and analogously to this work we get two connected submodels with polymorphic and monomorphic predator population. We get much richer dynamics, in each submodel we find local bifurcations (saddle-node, supercritical Hopf caused by Allee effect and Bogdanov-Takens) and a global bifurcation of limit cycles caused by the strategy evolution that is not possible in any of the submodels that can lead to a bluesky extinction of both populations.

1. The model

In our model, we expect the prey is limited by a carrying capacity and a threshold of survivance, which means that prey becomes extinct when its density is lower than θ. It is typical for a lot of populations — only few could grow up from just one or two individuals. This so called Allee effect is not considered in [1].

The predator population is divided into two types related to the predator behaviour. Whenever two predators meet after catching the prey, they choose their behaviour strategy. They can initiate aggressive behaviour and fight over the prey (we expect they are equally likely to be injured) or they retreat. The first predators are so called Hawks, the second are Doves. From the game theory it is known as the Hawk and Dove game. Whenever two Hawk predators meet, they both initiate aggressive behaviour, the conflict results and one of them gets the prey (gain $G > 0$). The cost of the conflict (C — given effort and got injuries)

JITKA KÜHNOVÁ — LENKA PŘIBYLOVÁ

reduces individual fitness of the other by some positive constant value C. When a Hawk meets a Dove, the Dove immediately retreats and the Hawk obtains the resource and when two Doves meet the resource is shared equally between them. The fitness payoffs for the Hawk-Dove game can be summarized according to the following matrix

$$A = \left(\begin{array}{cc} \frac{G-C}{2} & G \\ 0 & \frac{G}{2} \end{array} \right). \tag{1}$$

Let the function p in time t be the function of the number of predators which is the sum of Hawk $p_H(t)$ and Dove $p_D(t)$ predators. Let $x(t)$ and $y(t)$ be the proportions of predators with Hawks and Doves tactics

$$x(t) = \frac{p_H(t)}{p(t)}, \quad y(t) = \frac{p_D(t)}{p(t)} = 1 - x(t),$$

respectively.

Now we need to create a model in fast time scale by using replicators equations

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = x(\Delta_H - \Delta),$$

$$\frac{\mathrm{d}y}{\mathrm{d}\tau} = y(\Delta_D - \Delta),$$

where Δ_H is the gain of an individual always using the Hawk strategy, Δ_D is the gain of an individual always using the Dove strategy and Δ is the average gain of an individual playing the two tactics. Using the fact that $x + y = 1$ in any time t and after some algebra we get single equation:

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \frac{x}{2}(1 - x)(G - Cx). \tag{2}$$

The stationary points are 0, 1 a $G/C = x^*$. When $G < C$ we denote $x^* = \frac{G}{C}$. This stationary point is asymptotically stable, population of the predators is polymorphic with proportion G/C Hawks and $1 - G/C$ Doves. When $G > C$ we denote $x^* = 1$, and this stationary point is asymptotically stable, population of the predators is monomorphic (there are only Hawks).

For model in slow time scale we need the equation for population of the prey. Because of Allee effect in prey population, intraspecific competition and constant harvesting by predator we get

$$\frac{\mathrm{d}n}{\mathrm{d}t} = rn \left(\frac{n}{\theta} - 1 \right) \left(1 - \frac{n}{K} \right) - anp,$$

where θ is the threshold of survivance, r is the growth rate of prey population, K is the carrying capacity and a is a predation force parameter.

2

A PREDATOR-PREY MODEL WITH ALLEE EFFECT

For equations of predator populations we use assumptions from [1] and we get

$$\frac{\mathrm{d}p_H}{\mathrm{d}t} = -bp_H + \left(\alpha \left(\frac{G-C}{2} \right) \frac{p_H}{p} + \alpha G \frac{p_D}{p} \right) p_H,$$

$$\frac{\mathrm{d}p_D}{\mathrm{d}t} = -bp_D + \alpha \frac{G}{2} \frac{p_D}{p} p_D,$$

where b is the mortality rate of the predator (without prey predator becomes extinct), α is a conversion coefficient of gain and cost into biomass of predators. All coefficients are positive.

We assume that the model in fast time scale is established in the stationary point x^* and by using the theory of aggregated model [1] we get:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = rn \left(\frac{n}{\theta} - 1 \right) \left(1 - \frac{n}{K} \right) - anp,$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -bp + \frac{\alpha G}{2}p - \frac{\alpha C}{2}(x^*)^2 p.$$

If the gain depends on the prey density as $G(n) = an$, we obtain two different aggregated models:

Model I, $n < \frac{C}{a}$,

$$\frac{\mathrm{d}n}{\mathrm{d}t} = rn \left(\frac{n}{\theta} - 1 \right) \left(1 - \frac{n}{K} \right) - anp,$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -bp + \frac{\alpha a}{2}np - \frac{\alpha a^2}{2C}n^2 p.$$

Model II, $n > \frac{C}{a}$,

$$\frac{\mathrm{d}n}{\mathrm{d}t} = rn \left(\frac{n}{\theta} - 1 \right) \left(1 - \frac{n}{K} \right) - anp,$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -bp + \frac{\alpha a}{2}np - \frac{\alpha C}{2}p.$$

Now denote:

$$P(n) = r \left(\frac{n}{\theta} - 1 \right) \left(1 - \frac{n}{K} \right),$$

$$Q(n) = -b + \frac{\alpha a}{2}n - \frac{\alpha a^2}{2C}n^2 \qquad \text{for } n < \frac{C}{a},$$

$$= -b + \frac{\alpha a}{2}n - \frac{\alpha C}{2} \qquad \text{for } n > \frac{C}{a}.$$

We get:

$$\frac{\mathrm{d}n}{\mathrm{d}t} = n[P(n) - an],$$

$$\frac{\mathrm{d}p}{\mathrm{d}t} = Q(n)p.$$

(3)

JITKA KÜHNOVÁ — LENKA PŘIBYLOVÁ

2. Analysis of the model

Nullclines of the system:

- nullclines for both models: $n = 0, p = 0, p = \frac{1}{a}P(x)$.

- nullclines for the first model: $n_{1,2}^* = \frac{C}{2a}\left(1 \mp \sqrt{1 - \frac{8b}{\alpha C}}\right)$.

- nullcline for the second model: $n_3^* = \frac{C}{a} + \frac{2b}{\alpha a}$.

For n_1^*, n_2^*, n_3^* it is always true:

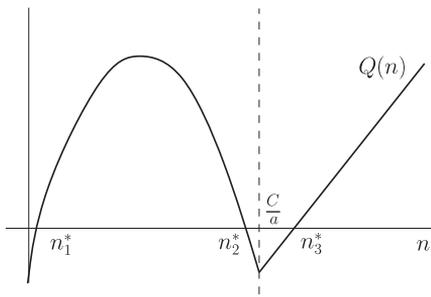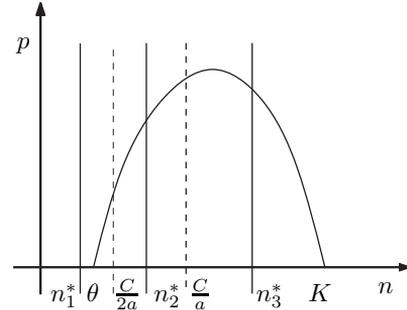$$0 < n_1^* \leq \frac{C}{2a} \leq n_2^* < \frac{C}{a} < n_3^* . \tag{4}$$

FIGURE 1
Function $Q(n)$.

FIGURE 2
Phase portrait of the combined model.

It is clear that, there are six stationary points:

$$(0,0), (\theta,0), (K,0), (n_1^*,p_1^*), (n_2^*,p_2^*) \quad \text{and} \quad (n_3^*,p_3^*).$$

The last three points exist only when

$$n_1^*, n_2^*, n_3^* \in (\theta, K).$$

For finding out the type and stability of all stationary points it is necessary to examine Jacobian matrix for all of them. The general Jacobian matrix

$$\boldsymbol{J}(n^*, p^*) = \begin{pmatrix} P(n^*) - ap^* + n^*P'(n^*) & -an^* \\ p^*Q'(n^*) & Q(n^*) \end{pmatrix}$$

1. $(0,0)$: always **stable node or focus**. Because of that, in every situation there are trajectories that end at beginning. That means both the prey and the predator become extinct.

4

A PREDATOR-PREY MODEL WITH ALLEE EFFECT

2. $(\theta, 0)$**: saddle** for $\theta \in (0, n_1^*) \cup (n_2^*, n_3^*)$, **unstable node or focus**, otherwise.

3. $(K, 0)$**: saddle** for $K \in (n_1^*, n_2^*) \cup (n_3^*, \infty)$, **stable node or focus**, otherwise.

4. (n_1^*, p_1^*)**:** for $n_1^* < \frac{K+\theta}{2}$ **unstable**, for $n_1^* > \frac{K+\theta}{2}$ **stable node or focus**.

5. (n_2^*, p_2^*)**:** always **saddle**.

6. (n_3^*, p_3^*)**:** for $n_3^* < \frac{K+\theta}{2}$ **unstable**, for $n_3^* > \frac{K+\theta}{2}$ **stable node or focus**.

At the first sight there could be five different situations:

1. only (0,0) is locally stable stationary point,

2. (0,0) and $(K, 0)$ are locally stable stationary points,

3. (0,0) and (n_1^*, p_1^*) are locally stable stationary points,

4. (0,0), $(K, 0)$ and (n_1^*, p_1^*) are locally stable stationary points,

5. (0,0) and (n_3^*, p_3^*) are locally stable stationary points.

The first two situations are not interesting - neither population survive, there is nothing to examine; the prey could not satisfy predator, predator becomes extinct and prey is stabilized on its carrying capacity. The third and the fifth situations have stable stationary point of coexistence of both populations, the fourth has either coexistence point or extinction of predator and prey stabilized on its carrying capacity.

From the analysis one could see that stationary points

$$(n_1^*, p_1^*) \quad \text{and} \quad (n_3^*, p_3^*)$$

could be stable or unstable focuses which implies there could be two Hopf bifurcations.

To examine stability of the limit cycle it is necessary to translate the system to the normal form of the Hopf bifurcation. First, we shift the system to the origin, second, we make linear transformation to the normal form and consequently we find the first Lyapunov coefficient whose sign determines the stability of the limit cycle.

Generally, our system (3) has the stationary point (n^*, p^*). Let us transform the variables:

$$n = \xi_1 + n^*,$$
$$p = \xi_2 + p^*.$$

We get the new system

$$\boldsymbol{\xi}' = \boldsymbol{f}(\boldsymbol{\xi}) \tag{5}$$

that could be rewritten as

$$\boldsymbol{\xi}' = \boldsymbol{A}_0 \cdot \boldsymbol{\xi} + \boldsymbol{F}(\boldsymbol{\xi}), \qquad \boldsymbol{\xi} = (\xi_1, \xi_2)^T,$$

JITKA KÜHNOVÁ — LENKA PŘIBYLOVÁ

where $\boldsymbol{A}_0$ is Jacobian matrix of the system (5) in the origin, denote $\det \boldsymbol{A}_0 = \omega^2$.

Now we change variables with matrix $\boldsymbol{T}$ of real eigenvectors of $\boldsymbol{A}_0$:

$$\boldsymbol{\xi} = \boldsymbol{T} \cdot \boldsymbol{\mu},$$
$$\boldsymbol{\mu}' = \boldsymbol{T}^{-1} \boldsymbol{A}_0 \, \boldsymbol{T} \boldsymbol{\mu} + \boldsymbol{T}^{-1} \boldsymbol{F}(\boldsymbol{T}\boldsymbol{\mu}),$$
$$\boldsymbol{P}(\boldsymbol{\mu}) = \boldsymbol{T}^{-1} \boldsymbol{F}(\boldsymbol{T}\boldsymbol{\mu}),$$
$$\boldsymbol{\mu}' = \boldsymbol{J} \cdot \boldsymbol{\mu} + \boldsymbol{P}(\boldsymbol{\mu}).$$

$\boldsymbol{J}$ is in the real Jordan form, in detail

$$\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}' = \begin{pmatrix} 0 & -\omega \\ \omega & 0 \end{pmatrix} \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} P(\mu_1, \mu_2) \\ R(\mu_1, \mu_2) \end{pmatrix}.$$

The stability of the limit cycle is determined by the first Lyapunov coefficient

$$\boxed{l_1(0) = \tfrac{1}{8\omega} L_1 + \tfrac{1}{8\omega^2} L_2}$$

with

$$L_1 = P_{111} + P_{122} + R_{112} + R_{222},$$
$$L_2 = P_{12}(P_{11} + P_{22}) - R_{12}(R_{11} + R_{22}) - P_{11}R_{11} + P_{22}R_{22},$$

where the lower indices mean partial derivatives of $\boldsymbol{P}$ respective to components evaluated at $\boldsymbol{\mu} = \boldsymbol{0}$.

We have the eigenvectors of $\boldsymbol{A}_0$:

$$\begin{pmatrix} 1 \\ -\frac{i\omega}{an_i^*} \end{pmatrix}, \quad \begin{pmatrix} 1 \\ \frac{i\omega}{an_i^*} \end{pmatrix},$$

and matrix $\boldsymbol{T}$

$$\begin{pmatrix} 1 & 0 \\ 0 & \frac{\omega}{an_i^*} \end{pmatrix}, \qquad i = 1, 3.$$

For (n_1^*, p_1^*) we have the functions $P(\mu_1, \mu_2), R(\mu_1, \mu_2)$:

$$P(\mu_1, \mu_2) = -\frac{r\mu_1^3}{\theta K} + \frac{r(K + \theta - 3n_1^*)}{\theta K}\mu_1^2 - \frac{\omega}{n_1^*}\mu_1\mu_2 + \frac{r}{K\theta}(K + \theta - 2n_1^*)n_1^*\mu_1,$$

$$R(\mu_1, \mu_2) = \frac{\alpha a}{2C}\left[(C - 2an_1^*)\mu_1\mu_2 - a\left(\frac{an_1^*p_1^*}{\omega} + \mu_2\right)\mu_1^2\right].$$

6

After some algebra we get the first Lyapunov coefficient

$$l_1(0) = -\frac{r\alpha a^2 n_1^* p_1^*}{4CK\theta\omega^3}(C - a n_1^*),$$

and because the stationary point (n_1^*, p_1^*) is from the first model, where

$$n < \frac{C}{a}, \quad \text{then} \quad l_1(0) < 0$$

and the limit cycle which appears in this Hopf bifurcation is stable and Hopf bifurcation is supercritical.

For (n_3^*, p_3^*), we have the functions $P(\mu_1, \mu_2), R(\mu_1, \mu_2)$:

$$P(\mu_1, \mu_2) = -\frac{r}{K\theta}\mu_1^3 - \frac{rn_3^*}{K\theta}\mu_1^2 - \frac{\omega}{n_3^*}\mu_1\mu_2,$$

$$R(\mu_1, \mu_2) = \frac{\alpha a}{2}\mu_1\mu_2.$$

It is easy to see that the first Lyapunov coefficient is

$$l_1(0) = -\frac{r}{2K\theta\omega} < 0.$$

The limit cycle which appears in this Hopf bifurcation is stable and Hopf bifurcation is supercritical.

Hopf bifurcation corresponding to the stationary point (n_1^*, p_1^*) comes up for

$$n_1^* = \frac{K + \theta}{2}$$

and according to (4) it must hold

$$K + \theta < \frac{C}{a}.$$

That means the second model has no influence in contrast with Hopf bifurcation corresponding to the stationary point (n_3^*, p_3^*), where because of the influence of the first model it appears the global bifurcation with appearance of the new unstable limit cycle (see Figure 3).
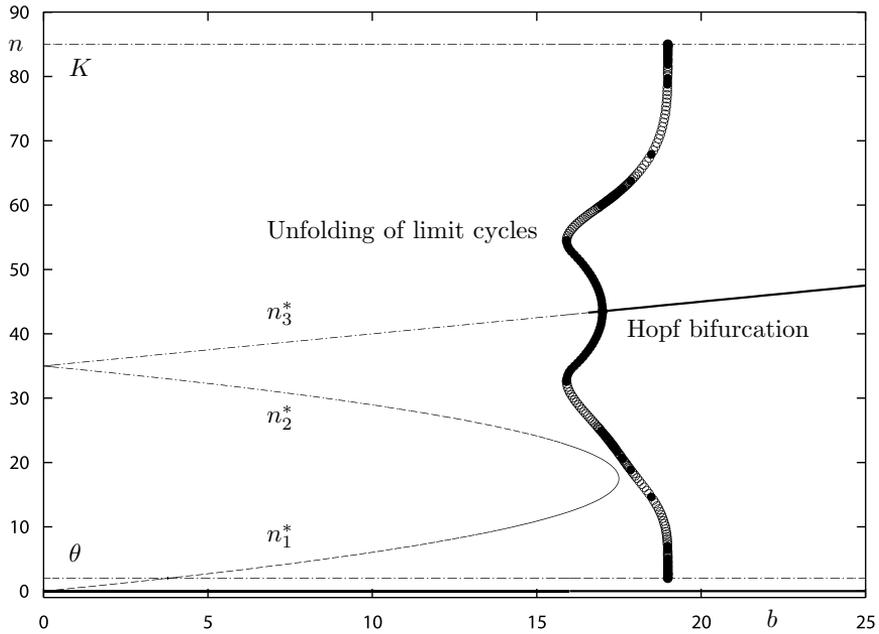
JITKA KÜHNOVÁ — LENKA PŘIBYLOVÁ

FIGURE 3. Stable cycle branch is presented by filled circles, unstable by unfilled circles. Unstable cycle branch splits on $(\theta, 0)$ and $(K, 0)$.

When

$$n_1^* = \frac{C}{2a},$$

we get saddle-node bifurcation and in combination with Hopf bifurcation we have Bogdanov-Takens bifurcation. By using analogous normalizing process like in computing of the first Lyapunov coefficient with assuming the condition for n_1^* we transform the system into the normal form of the Bogdanov-Tankens bifurcation.

We get the eigenvectors of $\boldsymbol{A}_0$

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ -\frac{1}{an_1^*} \end{pmatrix},$$

matrix $\boldsymbol{T}$

$$\begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{an_1^*} \end{pmatrix},$$

8

A PREDATOR-PREY MODEL WITH ALLEE EFFECT

and functions $P(\mu_1, \mu_2), R(\mu_1, \mu_2)$:

$$P(\mu_1, \mu_2) = -\frac{r{\mu_1}^3}{\theta K} + \frac{r(K + \theta - 3n_1^*)}{\theta K}\mu_1^2 + \frac{\mu_1\mu_2}{n_1^*} + \frac{r}{K\theta}(K + \theta - 2n_1^*)n_1^*\mu_1,$$

$$R(\mu_1, \mu_2) = -\frac{\alpha a^2 n_1^*}{2C}\left[p_1^*(C - 2an_1^*) - (C - 2an_1^*)\frac{\mu_2}{an_1^*} - a\left(p_1^* - \frac{\mu_2}{an_1^*}\right)\mu_1\right]\mu_1.$$

The genericity condition for Bogdanov-Tankens bifurcation according to [2] is

$$s = \operatorname{sgn}\big(b_{20}(a_{20} + b_{11})\big) \neq 0,$$

where

$$P(\mu_1, \mu_2) = a_{20}\mu_1^2 + P_1(\mu_1, \mu_2),$$
$$R(\mu_1, \mu_2) = b_{20}\mu_1^2 + b_{11}\mu_1\mu_2 + R_1(\mu_1, \mu_2).$$

Since

$$s = \operatorname{sgn}\left(\frac{\alpha a^3}{2C}n_1^*p_1^*\left(-\frac{rn_1^*}{K\theta} + \frac{\alpha a}{2C}(C - 2an_1^*)\right)\right),$$

s is nonzero variable when $\theta \neq K$, which is satisfied since $\theta < K$ in a common situation.

Another Bogdanov-Takens bifurcation is possible when

$$\lim_{b\to 0} n_2^* = \lim_{b\to 0} n_3^* = \frac{C}{a},$$

it is caused by combinations of two models. This implies possibility of splitting of the unstable limit cycle on the saddle point (n_2^*, p_2^*) see Figure 4.

Near the Bogdanov-Takens bifurcation there is a unique smooth curve corresponding to a saddle homoclinic bifurcation, where the limit cycle splits on the separatrix loop of the saddle. In our case, see the bifurcation diagram Figure 3, the supercritical Hopf bifurcation curve of the stable limit cycle has a limit point, where the stable cycle turns to be unstable one. It is a typical fold bifurcation of the limit cycles. In a small parameter area there coexist two nearby cycles - stable and unstable one. The unstable one then splits on the separatrix homoclinic loop of one of the saddles (see Figure 3 and 4).
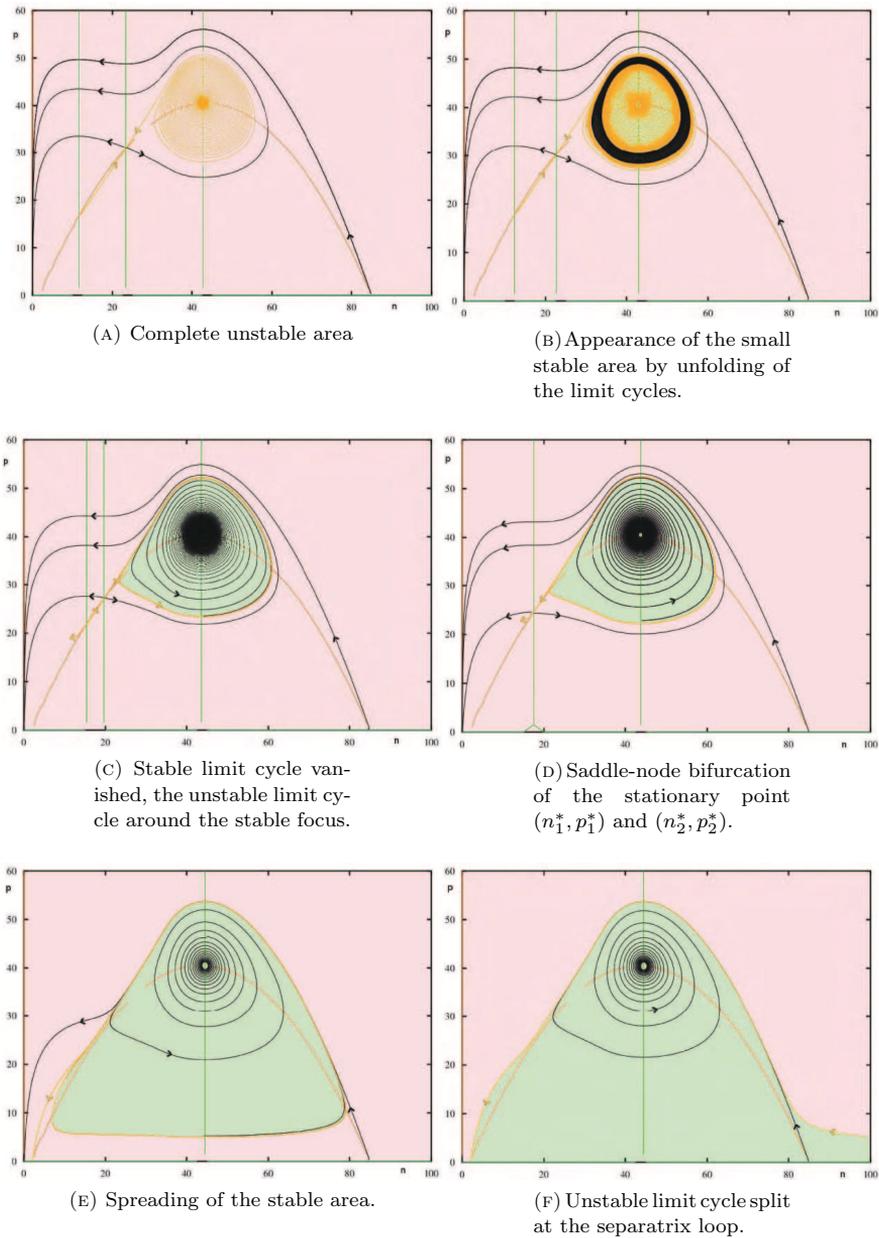
JITKA KÜHNOVÁ — LENKA PŘIBYLOVÁ

FIGURE 4. Stable cycle branch is presented by filled circles, unstable with unfilled circles. Unstable cycle branch splits on (n_2^*, p_2^*).

Let us look closer to this in some phase portraits (Figure 5). As it can be seen in Figure 5(a) for a small value of parameter b there is a huge unstable area, where both populations become extinct. As the parameter b increases there is unfolding of the stable cycle that arises from the supercritical Hopf bifurcation. This causes appearance of a small stable area "from nothing". When the parameter b crosses the critical value of the local supercritical Hopf bifurcation at the maximum of the parabola nullcline, the stable limit cycle vanishes and the branch of the unstable cycle around the stable focus continues until it splits at the separatrix loop of one of the saddle points. The green area is the basin of attraction of the stable limit cycle or the stable focus, respectively. It is a large area that maintains the populations. On the other hand, decreasing of the value of parameter b (mortality rate of the predator) causes disappearance of the stable green area and there is a global bifurcation—a bluesky catastrophe, extinction of both populations "from nothing".

A PREDATOR-PREY MODEL WITH ALLEE EFFECT

(A) Complete unstable area

(B) Appearance of the small stable area by unfolding of the limit cycles.

(C) Stable limit cycle vanished, the unstable limit cycle around the stable focus.

(D) Saddle-node bifurcation of the stationary point (n_1^*, p_1^*) and (n_2^*, p_2^*).

(E) Spreading of the stable area.

(F) Unstable limit cycle split at the separatrix loop.

FIGURE 5. Phase portraits.

3. Conclusion

In this paper, we have introduced the predator-prey model with Allee effect in prey and in which individual predators can use Hawk and Dove tactics in fighting over caught prey, that caused division into two submodels. We have shown, by bifurcation analysis and respective phase portraits, existence of saddle-node bifurcation, two supercritical Hopf bifurcations and two Bogdanov-Takens bifurcations in this aggregated model. Moreover, we have shown existence of global bifurcation of unfolding of two limit cycles, which is in neither submodel alone.

REFERENCES

[1] AUGER, P.—PARRA, R. B—MORAND, S.—SNCHEZ, E.: *A predator-prey model with predators using hawk and dove tactics*, Mathematical Biosciences **177&178** (2002), 185–200.
[2] KUZNETSOV, Y. A.: *Elements of Applied Bifurcation Theory*. Springer-Verlag, Berlin, New York, Inc. 1998.

Jitka Kühnová
Faculty of Science
University of Hradec Králové
Rokitanského 62
CZ–500-03 Hradec Králové
CZECH REPUBLIC
E-mail: jitka.kuhnova@seznam.cz

Lenka Přibylová
Department of Mathematics and Statistics
Faculty of Science
Masaryk University
Kotlářská 2
CZ–611-37 Brno
CZECH REPUBLIC
E-mail: pribylova@math.muni.cz

12

Contents lists available at ScienceDirect

Ecological Complexity

Original Research Article

Foraging facilitation among predators and its impact on the stability of predator–prey dynamics

Lenka Přibylová *, Anna Peniašková

Department of Mathematics and Statistics, Section of Applied Mathematics, Faculty of Science, Masaryk University, Kotlářská 2, 611 37 Brno, Czech Republic

ABSTRACT

Predator foraging facilitation may strongly influence the dynamics of a predator–prey system. This behavioral pattern is well-observed in real life interactions, but less is known about its possible impacts on the predator–prey dynamics. In this paper we analyze a modified Rosenzweig–MacArthur model, where a predator-dependent family of functions describing predator foraging facilitation is introduced into the Holling type II functional response. As the general assumption of foraging facilitation is that higher predator densities give rise to an increased foraging efficiency, we model predator facilitation with an increasing encounter rate function. Using the tools of bifurcation analysis we describe all the nonlinear phenomena that occur in the system provoked by foraging facilitation, these include the fold, Hopf, transcritial, homoclinic and Bogdanov–Takens bifurcation. We show that foraging facilitation can stabilize the coexistence in the predator–prey system for specific rates, but in most of the cases it can have fatal consequences for the predators themselves.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Behavioral interaction among the entities of predator or prey populations, both positive and negative, may change and modify the basic findings of population biology and theory of ecology and may also have a substantial effect on the individual's foraging success (Skalski and Gilliam, 2001). The importance of the inclusion of predator behavior into the predator–prey interaction models has been recognized (e.g. Abrams and Ginzburg, 2000; Arditi and Akcakaya, 1990; Lima, 2002) but little is known about the nonlinear phenomena that may appear in predator–prey systems with interaction due to bifurcations. Empirical studies and observations have shown the presence of interacting behavior in animal populations; lions or baboons are very often subjects of such observations (Heinsohn and Packer, 1995). Interacting predator–prey systems especially with cooperative or foraging behavior are often modeled as a social dilemma in game theory

with several strategies of behavior to adopt (Iwasa, 1982; Packer and Ruttan, 1988; Brown et al., 1999; Doebeli and Hauert, 2005). The motives and evolution even of human cooperative behavior is not thoroughly known currently; Richerson and Boyd (2001) argue that deeper understanding of human behavior may lead to a better explanation of other species.

Predator functional response, that is, the per capita feeding rate of predators upon their prey, is a basic concept in the predator–prey theory from its beginning. There is a variety of functional response types, almost every textbook (e.g. Begon et al., 1990; Krebs, 2001) refers to the traditional classification of the three density-independent types (Holling, 1959): Holling type I (linear), Holling type II (concave increase) and Holling type III (sigmoid increase) functional responses. There is a wide range of different functional relations of feeding rate with respect to the predator population density (Skalski and Gilliam, 2001). The most illustrative choices are an increasing or a decreasing feeding rate. One may interpret the decreasing feeding rate as a predator population where the entities interfere. Predator interference is a collective term that includes a number of specific mechanisms from stealing subdued prey to mechanisms connected with territorial behavior

* Corresponding author.
 E-mail address: pribylova@math.muni.cz (L. Přibylová).

of the predator or the prey. There is a wide range of literature available on interference, such as (Arnqvist et al., 2006; Ruxton et al., 1992; Ruxton, 1995).

Increasing feeding rate, on the other hand, represents foraging facilitation among the entities of the predator population. Although foraging facilitation is the opposite process to predator interference, contemporary predator–prey theory often overlooks this aspect despite the fact that an inclusion of the facilitation in ecology theory may influence the results and the learnings gained so far in predator–prey theory (Giraldeau and Caraco, 2000; Bruno et al., 2003). Foraging facilitation can manifest in different ways, it can be an ability to locate or capture the prey in a bigger group (Cosner et al., 1999), intraspecific cooperation (Courchamp and Macdonald, 2001), protecting mechanisms against other species (Krause and Ruxton, 2002), or accessibility of public information on the availability of food (Bijleveld et al., 2015).

Recent studies devoted to density dependent functional responses and their effects on predator–prey dynamics (e.g. Nilsson et al., 2006; Fryxell et al., 2007; Boukal et al., 2008; Berec, 2010; Pribylova and Berec, 2014) show that density dependent functional responses bring about new stabilizing and destabilizing mechanisms of predator–prey dynamics. The aim of the present paper is to build on the findings of Berec (2010) and Pribylova and Berec (2014) and analyze these mechanisms for a family of functional responses representing positive interaction among the entities of the predator population. While Berec (2010) gives a brief overview on the number and the character of the coexistence equilibria, we present a complete and thorough analysis of the system with foraging facilitation. In Berec (2010), Berec calls for further, more detailed studies on the models with predator facilitation and interference, with special attention on the limit cycle behavior. A step towards the understanding of the phenomena that may occur in systems with interacting predators was the analysis of interfering predators in Pribylova and Berec (2014). However, the picture is not complete without the predator facilitation case, which is the question we address in the recent paper. While the notation is similar to that in Pribylova and Berec (2014), the approach of our analysis is different and so are our findings. We have considered the family of functional responses suggested in Berec (2010) with unbounded above encounter rate and a more realistic family of functional responses with bounded above encounter rate. The latter corresponds with the findings of several empirical studies (Bijleveld et al., 2015) and studies on game theory models (Packer and Ruttan, 1988). While Berec (2010) considers only the number and character (i.e. stability) of the coexistence equilibria, our analysis goes deeper and examines the bifurcations that may occur. We concentrate on the conditions of appearance of multi-stable regions.

Observations of cooperative behavior, mainly cooperative hunting, have been published several times (Packer and Ruttan, 1988; Creel and Creel, 1995; Heinsohn and Packer, 1995). A general rule of cooperative hunting strategy is that the benefits of group hunting for each hunter have to outweigh the benefits of solitary hunting. Packer and Ruttan (1988) showed that cooperative hunting can be the evolutionarily stable strategy for species hunting a single small prey, a small group size of the hunters hunting a single large prey, or for a large group of hunters hunting multiple large preys. Species with cooperative hunting strategy have an increasing above bounded hunting success function with respect to the hunter group size. Packer's findings validate the consideration of an increasing above bounded encounter rate at least for several specific predator–prey populations (such as lions and hyenas).

The analyzed model is a modification of the classic Rosenzweig–MacArthur predator–prey model (Rosenzweig, 1971) with Holling type II functional response of predators. The Rosenzweig–MacArthur model demonstrates the paradox of enrichment where stable oscillations bifurcate out of a stable equilibrium once the environmental carrying capacity of the prey exceeds a critical value (Kot, 2001). We show that our model keeps this very typical behavior after the encounter rate functions for foraging facilitation are introduced. Another destabilization may appear due to the homoclinic bifurcation that causes splitting of the stable cycle, thus ending the oscillations and consequently causing the extinction of the predators. This phenomenon influences the prey density as well.

2. Model

2.1. Generic model with predator foraging facilitation

Let us consider the following modified Rosenzweig–MacArthur predator–prey model with the predator-dependent functional response

$$
\begin{aligned}
\frac{dN}{dt} &= rN\left(1 - \frac{N}{K}\right) - Pf(N, P),\\
\frac{dP}{dt} &= ePf(N, P) - mP.
\end{aligned}
\tag{1}
$$

In this model, N and P are prey and predator densities, respectively, r is the intrinsic per capita growth rate of the prey, K is the environmental carrying capacity of the prey, m is the per capita predator mortality rate, and e is the efficiency with which the consumed prey is transformed into new predators. The density of the prey increases logistically in the absence of the predator, while the predator dies out exponentially in the absence of the prey. The predator functional response $f(N, P)$ is a generalized Holling type II functional response with the predator encounter rate λ being a smooth and increasing function of the predator density P

$$
f(N, P) = \frac{\lambda(P)N}{1 + h\lambda(P)N},
\tag{2}
$$

where $\lambda(0) \geq 0$, $\lambda'(P) > 0$, and h denotes the predator handling time of one prey. The general assumption of foraging facilitation is that higher predator densities give rise to an increased foraging efficiency and hence increase encounter rate for any member of the foraging party.

Combining the generic model (1) with the functional response (2), our primary model is

$$
\begin{aligned}
\frac{dN}{dt} &= rN\left(1 - \frac{N}{K}\right) - \frac{\lambda(P)N}{1 + h\lambda(P)N}P,\\
\frac{dP}{dt} &= e\frac{\lambda(P)N}{1 + h\lambda(P)N}P - mP.
\end{aligned}
\tag{3}
$$

2.2. Specific models

Berec (2010) considers a specific encounter rate function

$$
\lambda(P) = \lambda_0(b + P)^w,
\tag{4}
$$

where $b \geq 0$, $\lambda_0 > 0$. For a positive w, $\lambda(P)$ is increasing. Encounter rate (4) is an increasing unbounded above function, concave for $w \in (0, 1)$ and convex for $w > 1$. Note that for $w < 0$, encounter rate (4) models a negative interaction. Varying w in (4) produces a family of functional responses, whereas varying b or λ_0 does not qualitatively change the functional response. Substituting $b = 0$ and $\omega = -w$ into (4) leads to the Hassel–Varley functional response

$$
f(N, P) = \frac{\lambda_0(N/P^\omega)}{1 + h\lambda_0(N/P^\omega)}.
\tag{5}
$$

Setting $b = 0$ and $w = 1$ one obtains the form used by Cosner et al. (1999)

$$f(N,P) = \frac{\lambda_0 NP}{1 + h\lambda_0 NP}. \tag{6}$$

Encounter rate function (4) covers the Beddington–DeAngelis functional response if $b > 0$ and $w = -1$ (Beddington, 1975; DeAngelis et al., 1975). For $w = 0$ the function simplifies to a constant $\lambda(P) = \lambda_0$ and the functional response $f(N, P)$ is of the Holling type II.

The generalized model (3) for $w \in \mathbb{R}$ has been studied partly by Berec (2010). Bifurcation analysis has been presented for the interference case ($w < 0$) in Pribylova and Berec (2014) showing that sufficiently strong predator interference may bring stabilizing mechanisms. In the model with foraging facilitation we complete the analysis that was left to be done. We focus to the cases with bi-stable regions and cyclic behavior. The emergence and additional disappearance stand in our special interest, because the effect of limit cycles on the persistence of the populations can be fatal. Our findings about the role and effect of the Bogdanov–Takens bifurcation along with the generic topological structure of the parametric space of the carrying capacity and the rate of facilitation are novel and unique. Apart from the general analysis of the model (3) we present our findings for two specific types of encounter rate functions, function (4) and a function first introduced in this paper to the best of our knowledge

$$\lambda(P) = L - \frac{\lambda_0}{(b + P)^w}, \tag{7}$$

where $b \geq 0$, $\lambda_0 > 0$, $w > 0$, and $\lambda(0) \geq 0$. Encounter rate (7) is increasing, concave and bounded above. It represents a more realistic assumption corresponding with the findings of several empirical and theoretical studies on cooperative behavior, such as Packer and Ruttan (1988) and Bijleveld et al. (2015). For a large population density of predators, (7) saturates at L and the functional response $f(N, P)$ asymptotically approaches that of the Holling type II. In both specifically chosen encounter rate forms (4) and (7), one may interpret the parameter w as a measure of the facilitation rate among the predator population. For $w = 0$ there is no facilitation and the encounter rate is not dependent on predator population density. With increasing w the encounter rate for the same predator population density increases. The additional parameter L, of course, make the analysis more complicated than the case of the form (4) and that's why the manifolds derivations for the form (7) are not as much analogous to those published in Pribylova and Berec (2014) as for the form (4).

In the following, we focus on the coexistence equilibria of the primary model (3) and on the characterization of the local bifurcations that may occur. These include limit point equilibrium disappearance, appearance of a limit cycle due to the Hopf bifurcation, and disappearance of a stable limit cycle due to the homoclinic bifurcation. Another non-local phenomenon connected to the transcritical bifurcation near the Hopf bifurcation causes persistence of a huge stable limit cycle or a huge homoclinic saddle separatrix orbit which leads to the extinction of the predator population even for increasing facilitation.

3. Model analysis

This part is based on bifurcation theory. We derive the existence conditions of the listed non-linear phenomena and describe their consequences on the dynamics of the predator–prey populations. The analysis of the model goes along the lines of Pribylova and Berec (2014) but additional information have to be considered, especially in the case of model (3) with the above bounded

encounter rate function. We have decided to present a thorough, step-by-step bifurcation analysis and empower the reader to follow the idea behind the steps. Readers familiar with Pribylova and Berec (2014) may go through this part very quickly except for the nullclines description in the first part and the transcritical bifurcation manifold derivation. Qualitatively the results are summarized and biological consequences are presented in Section 4.

3.1. Multiple coexistence equilibria

The invariant nullclines $N = 0$ and $P = 0$ of the primary model (3) correspond to the exponential extinction of the predators with absent prey and to the logistic dynamics of the prey with absent predator, respectively. The stability properties of the two boundary equilibria $[0, 0]$ and $[K, 0]$ related to the logistic dynamics of the prey with absent predator were studied by Berec (2010). He found that the extinction equilibrium $[0, 0]$ is always a saddle point and the prey-only equilibrium $[K, 0]$ changes its stability due to the transcritical bifurcation.

The non-zero nullclines of system (3) are

$$
\begin{aligned}
F_1(N,P) &= r\left(1 - \frac{N}{K}\right) - \frac{\lambda(P)}{1 + h\lambda(P)N}P = 0 \quad \text{for } N' = 0, \\
F_2(N,P) &= e\frac{\lambda(P)N}{1 + h\lambda(P)N} - m = 0 \quad \text{for } P' = 0.
\end{aligned}
$$

One may easily obtain that the slope of the $F_2(N, P) = 0$ nullcline is negative $((dP/dN) = -(\lambda(P)/N\lambda'(P)) < 0)$ and the nullcline is a decreasing function $P = f_2(N)$ in the first quadrant. The nullcline $F_1(N, P) = 0$ intersects the coordinate axes in $[K, 0]$ and $[0, P_0]$, where $P_0 > 0$ denotes the only solution of $\frac{r}{P} = \lambda(P)$. Since in the first quadrant

$$\frac{dF_1}{dP}(N,P) = -\frac{\lambda(P) + hN\lambda^2(P) + P\lambda'(P)}{(1 + h\lambda(P)N)^2} < 0,$$

the nullcline is a continuous function $P = f_1(N)$. The described behavior is a generic feature of the nullclines, therefore it does not depend on the specific form of the encounter rate functions. Fig. 1 shows possible intersections of the nullclines.

A coexistence equilibrium $E^* = [N^*, P^*]$ of model (3) has to satisfy

$$N^* = K\frac{C_2}{\lambda(P^*)}, \tag{8}$$

$$P^* = C_1\frac{\lambda(P^*) - C_2}{\lambda^2(P^*)}, \tag{9}$$

where $C_1 = (er/(e - hm))$ and $C_2 = (m/(K(e - hm)))$. If $e - hm \leq 0$, no coexistence equilibrium exists. Therefore, $e - hm > 0$ is a valid assumption implying also $C_1 > 0$ and $C_2 > 0$, while

$$\lambda(P^*) > C_2. \tag{10}$$

Eq. (9) defines the implicit function

$$F(P^*) := P^*\lambda^2(P^*) - C_1(\lambda(P^*) - C_2) = 0 \tag{11}$$

of the non-zero predator equilibrium P^* (having one-to-one correspondence to the non-zero prey equilibrium N^* via (8)). In a limit point related to merging and disappearance of two equilibria on a fold of the equilibrium manifold, the value of P^* satisfies $F'(P^*) = 0$, which can be equivalently expressed as

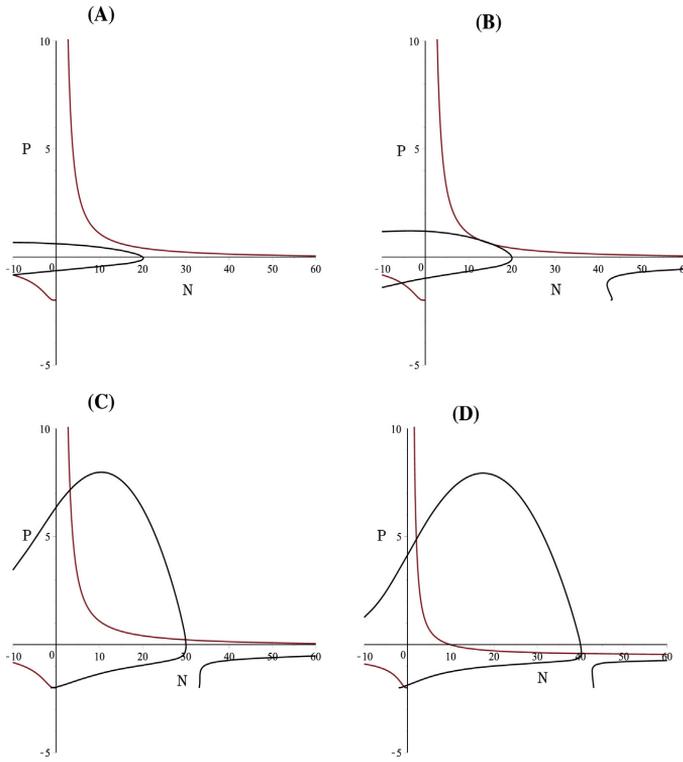$$\lambda^3(P^*) - C_1(2C_2 - \lambda(P^*))\lambda'(P^*) = 0. \tag{12}$$

Fig. 1. Nullclines of the system (3) with encounter rate function (7) for $L = 5$, $\lambda_0 = 5.1783$, $e = 1$, $m = 1$, $h = 0.25$, $b = 2$. Parameters of the subplots are: (A) $w = 0.055$, $K = 20$, $r = 0.05$, (B) $w = 0.055$, $K = 20$, $r = 0.17$, (C) $w = 0.055$, $K = 30$, $r = 2.5$, (D) $w = 0.09$, $K = 40$, $r = 2.5$.

Analogously to the results published in Přibylova and Berec (2014) it can be proven that condition (12) is both necessary and sufficient for vanishing of the Jacobian of model (3) at the coexistence equilibrium $[N^*, P^*]$. It is obvious that for P^* to be a limit point, it has to satisfy $\lambda(P^*) < 2C_2$ since $\lambda'(P^*) > 0$ and Eq. (12) holds.

Crossing the fold bifurcation manifold, the coexistence equilibria disappear leading to the extinction of the predator population while the prey survives at the carrying capacity K as shown in Fig. 2. The phase portraits of system (3) with encounter rate (7) are presented for varying prey growth rate r with other parameters fixed. A decrease in growth rate r of the prey causes shrinking of the basin of attraction of the stable coexistence equilibrium (the grey area) and brings about the extinction of the predators at the fold bifurcation critical value. All trajectories starting in the white area end in E_K. The parameter values $K = 8$, $w = 0.06$, $e = 1$, $m = 1$, $h = 0.25$, $L = 5$, $\lambda_0 = 5.1783$, $b = 2$ satisfy equation (12) for a growth rate of $r = 1.2095$. Below this rate the predator population can never survive. The merging equilibria E_1^* and E_2^* are marked as equilibrium $E^* \doteq [5.7997, 1.9294]$. For a relatively high prey birth rate r, two stable states are present in the system: coexistence and the extinction of the predators (as the prey population does not represent a sufficient supply for the predators). With decreasing birth rate, the prey population stabilizes on the level of the carrying capacity K while the predator population goes extinct. The disappearance of multiple stable coexistence states caused by decreasing prey birth rate is influenced by the predator foraging facilitation rate. The expression representing the fold bifurcation manifold for system (3) with the encounter rate functions (4) and (7) may be obtained as

described below. Denoting $\Lambda = \lambda(P^*)$ in (9) and using the expressions (14a) or (14b), the fold bifurcation manifolds can be written in polynomial form as

$$
\begin{cases}
b\Lambda^2 + \Lambda C_1(1+w) - C_1 C_2(2w+1) = 0 & \text{for encounter rate (4),} \quad (13a) \\
b\Lambda^3 + \Lambda^2 C_1(1-w) + \Lambda C_1(w(L+2C_2) - C_2) - 2C_1 C_2 Lw = 0 & \text{for encounter rate (7).} \quad (13b)
\end{cases}
$$

The encounter rate differentials are

$$
\lambda'(P) =
\begin{cases}
\dfrac{\lambda(P)w}{b+P} & \text{for encounter rate (4),} \quad (14a) \\[2ex]
\dfrac{(L-\lambda(P))w}{b+P} & \text{for encounter rate (7).} \quad (14b)
\end{cases}
$$

Note that the coefficients of the polynomials (13a) and (13b) are independent of λ_0. Any positive root Λ^* of these polynomials fulfilling the condition of the fold bifurcation (12) has to satisfy

$$
E^* = [N^*, P^*] = \left[K\frac{C_2}{\Lambda^*}, C_1\frac{\Lambda^* - C_2}{\Lambda^{*2}} \right]
$$

for a critical bifurcation value of the parameter $\lambda_0 = \lambda_0^*$. The parameter values for the two studied encounter rates are

$$
\lambda_0^* =
\begin{cases}
\dfrac{\Lambda^*}{(b+P^*)^w} & \text{for encounter rate (4),} \quad (15a) \\[2ex]
(L-\Lambda^*)(b+P^*)^w & \text{for encounter rate (7).} \quad (15b)
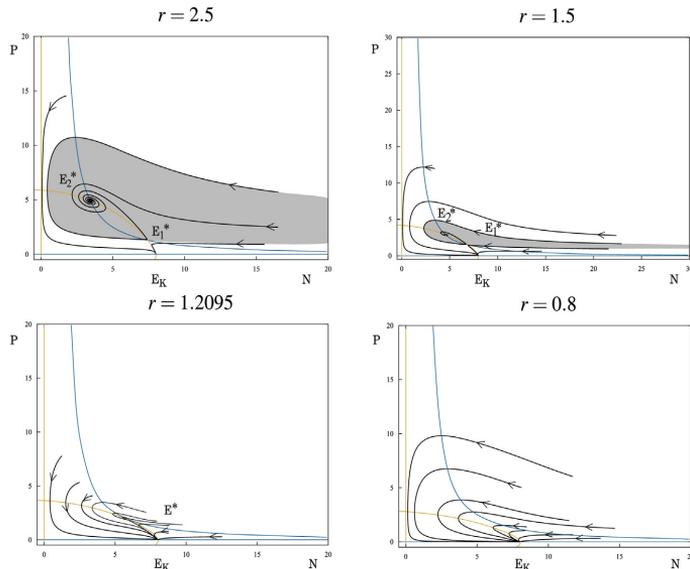\end{cases}
$$

Fig. 2. Phase portrait of system (3) with the encounter rate function (7) for parameters $K = 8$, $w = 0.06$, $e = 1$, $m = 1$, $h = 0.25$, $L = 5$, $\lambda_0 = 5.1783$, $b = 2$, and decreasing r.

3.2. Coexistence equilibrium Hopf bifurcation

The coexistence equilibrium can undergo Hopf bifurcation meaning that a local qualitative change of the phase space becomes visible. A periodic solution emerges or dies out from the equilibrium as the bifurcation parameter crosses a critical value where the real parts of both of the eigenvalues transversally cross zero and two purely imaginary eigenvalues of the Jacobian evaluated at the coexistence equilibrium appear. Assuming that for $w = 0$ the encounter rate simplifies to a constant and therefore system (3) simplifies to the classic Rosenzweig–MacArthur model exhibiting the well-analyzed paradox of enrichment caused by the Hopf bifurcation. We expect the Hopf bifurcation to appear in our system with foraging facilitation as well.

Necessary conditions of the Hopf bifurcation at a coexistence equilibrium $[N^*, P^*]$ are

$$\lambda'(P^*) = \frac{(eC_2 - mh(\lambda(P^*) - C_2))\lambda^2(P^*)}{em(\lambda(P^*) - C_2)}, \qquad (16)$$

$$\lambda^3(P^*) - C_1(2C_2 - \lambda(P^*))\lambda'(P^*) > 0, \qquad (17)$$

where (16) is the necessary and sufficient condition for vanishing of the trace of the Jacobian of model (3) at a coexistence equilibrium. The positivity of the Jacobian at the equilibrium is guaranteed by (17).

Substituting $\lambda'(P^*)$ from (16) into (17), the condition (17) can be rewritten as

$$\lambda(P^*)(\lambda(P^*) - C_2) > \frac{rh}{e - hm}(\lambda(P^*) - 2C_2)\left(\lambda(P^*) - C_2\frac{e + hm}{hm}\right). \qquad (18)$$

A graphical solution of inequality (18) is presented in Fig. 3 where the convex parabolas represent the left-hand and the right-hand sides of expression (18) and $\lambda_H(P^*)$ is the unique point that belongs to their intersection in the interval $(C_2, 2C_2)$. Since we focus to the case of predator–prey coexistence where the condition (10) holds, it is obvious that the condition (18) is valid for all $\lambda(P^*)$ between the roots of the right-hand side of expression (18) and for

$(rh/(e - hm)) \leq 1$ always (that is for low enough r). It is worth to mention, that the Hopf bifurcation necessary condition and the fold bifurcation condition both can be fulfilled for $\lambda(P^*) \leq 2C_2$ (the grey hatched area in Fig. 3), therefore one may expect to detect the Bogdanov–Takens bifurcation at this area. The coexistence equilibrium with $\lambda(P^*) \in (C_2, 2C_2)$ is prone to undergo some very significant qualitative changes that are important especially for the biological interpretation the model.

The analysis of the Hopf bifurcation manifold uses analogous steps to those used in the case of the fold bifurcation manifold. Substituting (4) and (7) into expression (16), we obtain a polynomial equation whose coefficients do not depend on the parameter λ_0. For $\Lambda = \lambda(P^*)$ the polynomial takes the form

$$A\Lambda^3 + B\Lambda^2 + C\Lambda + D = 0 \text{ for encounter rate (4), where} \qquad (19)$$

$$\begin{aligned} A &= K^2hb(e - hm)^3 > 0, \\ B &= K(e - hm)^2(hKer - b(e + hm) + wKe(e - hm)), \\ C &= -Ke(e - hm)(wm(e - hm) + r(2hm + e)) < 0, \\ D &= emr(e + hm) > 0 \end{aligned} \qquad (20)$$
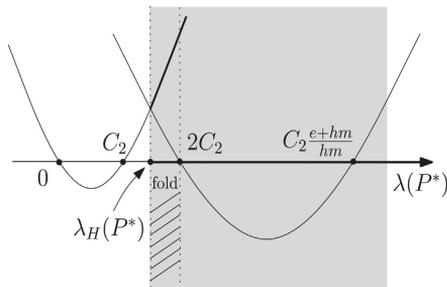

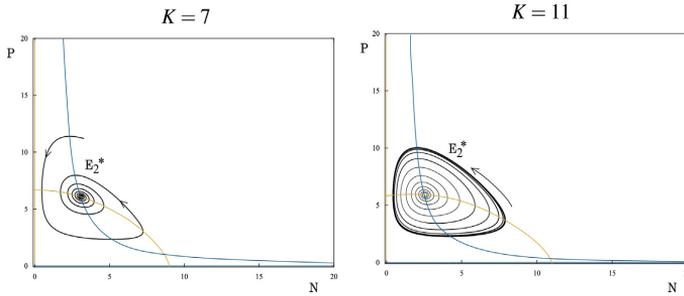**Fig. 3.** Graphic solution of inequality (18).

Fig. 4. Phase portrait of system (3) with encounter rate function (7) for $K = 7$ (left) and $K = 11$ (right). All other parameters are fixed at $w = 0.07$, $e = 1$, $m = 1$, $h = 0.25$, $L = 5$, $\lambda_0 = 5.1783$, $b = 2$, $r = 3$.

or

$$A\Lambda^3 + B\Lambda^2 + C\Lambda + D = 0 \quad \text{for encounter rate (7), where} \tag{21}$$

$$\begin{aligned}
A &= K^2 hb(e - hm)^3 > 0, \\
B &= K(e - hm)^2(hKer - b(e + hm) - wKe(e - hm)), \\
C &= Ke(e - hm)(wm(e - hm) - r(2hm + e) + wLK(e - mh)^2), \\
D &= em(r(e + hm) - KLw(e - hm)^2).
\end{aligned} \tag{22}$$

There is a correspondence between the roots Λ^* of (19) or (21) and the critical value of λ_0 described by the expressions of the equilibrium point (8) and (9). Because of the biological interpretation of the model, only cases where $\Lambda^* > 0$ and $\lambda_0 > 0$ are of our interest. The effect of the Hopf bifurcation on the model with predator encounter rate (7) is shown in Fig. 4 for $K = 7$ (left) and $K = 11$ (right). All other parameters are fixed at $w = 0.07$, $e = 1$, $m = 1$, $h = 0.25$, $L = 5$, $\lambda_0 = 5.1783$, $b = 2$, $r = 3$. We have chosen K as the bifurcation parameter of the Hopf bifurcation but any other parameter (except λ_0) is a valid choice. K has been chosen in order to respect traditions as the paradox of enrichment was described with respect to K (Rosenzweig, 1971). A periodic solution appears for $w > w_{HB}$. The critical value w_{HB} separates the region of locally stable focus coexistence equilibrium and the region of unstable focus coexistence equilibrium with a stable limit cycle. For weak foraging facilitation $w < w_{HB}$ the populations can coexist. Higher foraging facilitation leads to non-stability and oscillations in densities of the populations.

3.3. Coexistence equilibrium and predator extinction equilibrium transcritical bifurcation

Due to the transcritical bifurcation, a coexistence equilibrium merges with the prey-only equilibrium E_K for $\lambda(P^*) = C_2$ on the manifold defined by $\lambda(0) = C_2$, that is for

$$K = \begin{cases} \dfrac{m}{e - mh} \dfrac{1}{\lambda_0 b^w} & \text{for encounter rate (4),} \tag{23a} \\[3mm] \dfrac{m}{e - mh} \dfrac{1}{L - \frac{\lambda_0}{b^w}} & \text{for encounter rate (7).} \tag{23b} \end{cases}$$

This phenomenon has a huge effect especially on the population of the predator. The prey-only equilibrium E_K remains the only stable equilibrium after K crosses the critical bifurcation value, this leads to predator extinction.

3.4. Coexistence equilibrium Bogdanov–Takens bifurcation

Bogdanov–Takens bifurcation is connected to a non-linear phenomenon of merging equilibria with a nearby limit cycle. There are various regimes in the neighborhood of the Bogdanov–Takens

bifurcation manifold and the dynamic transition between these regions may be stabilizing, destabilizing or even fatal – depending on which of the one-parameter manifold is crossed or on the direction of the shift. A generic Bogdanov–Takens manifold is a two parametric manifold (with codimension 2) that connects the one-parametric fold and Hopf bifurcation manifolds (with codimension 1) and another one-parametric manifold of a non-local bifurcation called homoclinic, since it belongs to a homoclinic trajectory. Typical regime shifts near the Bogdanov–Takens point can be deduced from the bifurcation diagram 5: coexistence equilibria disappearance on the fold bifurcation manifold, destabilizing of the stable coexistence equilibrium to oscillations or split of a stable cycle on the separatrix loop of homoclinic bifurcation. All of these phenomena have significant effect on population persistence and stability. For deeper insight see Kuznetsov (1998).

A necessary condition for the Bogdanov–Takens bifurcation to occur in system (3) is a concurrent fold and Hopf bifurcation. This condition can be expressed as a quadratic equation combining conditions (12) and (16)
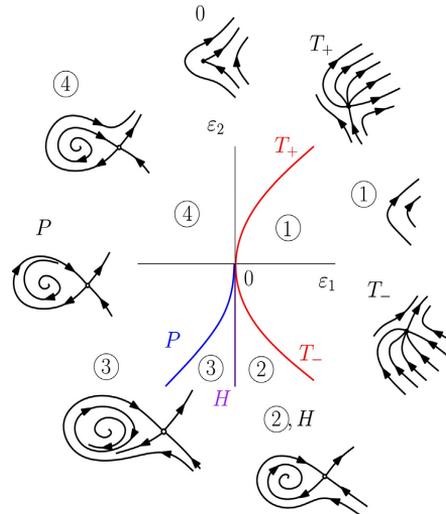
$$A\Lambda^2 + B\Lambda + C = 0, \tag{24}$$

Fig. 5. Bifurcation diagram near Bogdanov–Takens bifurcation point [0, 0] for its normal form $y_1 = y_2$, $y_2 = \varepsilon_1 + \varepsilon_2 y_1 + y_1^2 - y_1 y_2$. T^+ is the fold bifurcation curve with one zero and one positive eigenvalue, T^- is the fold bifurcation curve with one zero and one negative eigenvalue, H is the Hopf bifurcation curve with two purely imaginary eigenvalues, and P is the homoclinic bifurcation curve. See also Kuznetsov (1998).

where the coefficients are

$$A = K^2(e-mh)^2(h(m+r)-e),$$
$$B = K(e-mh)(m(e-mh)-r(e+3mh)),$$
$$C = 2mr(e+mh).$$

(25)

The specific form of the encounter rate functions do not influence the solution of this equation. The necessary condition of the Bogdanov–Takens bifurcation is identical for the cases with functions (4), (7), and even for predator interference. The parameter values of (4) and (7) have to be chosen to meet several
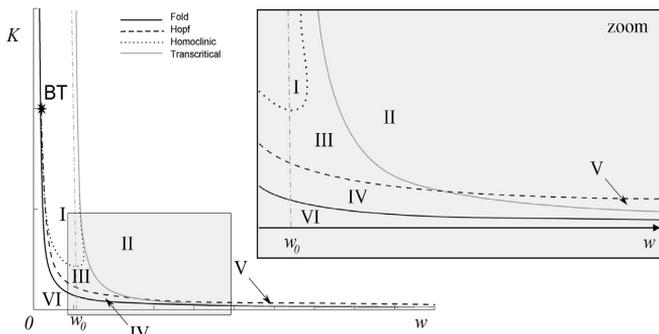

Fig. 6. Bifurcation diagram in parameter space w versus K with the fold, Hopf, homoclinic, and transcritical bifurcation curves and the Bogdanov–Takens point for parameter values $e = 1$, $m = 1$, $h = 0.25$, $r = 3$, $L = 5$, $\lambda_0 = 5.1783$, and $b = 2$. A close-up of the grey rectangular area is also shown.
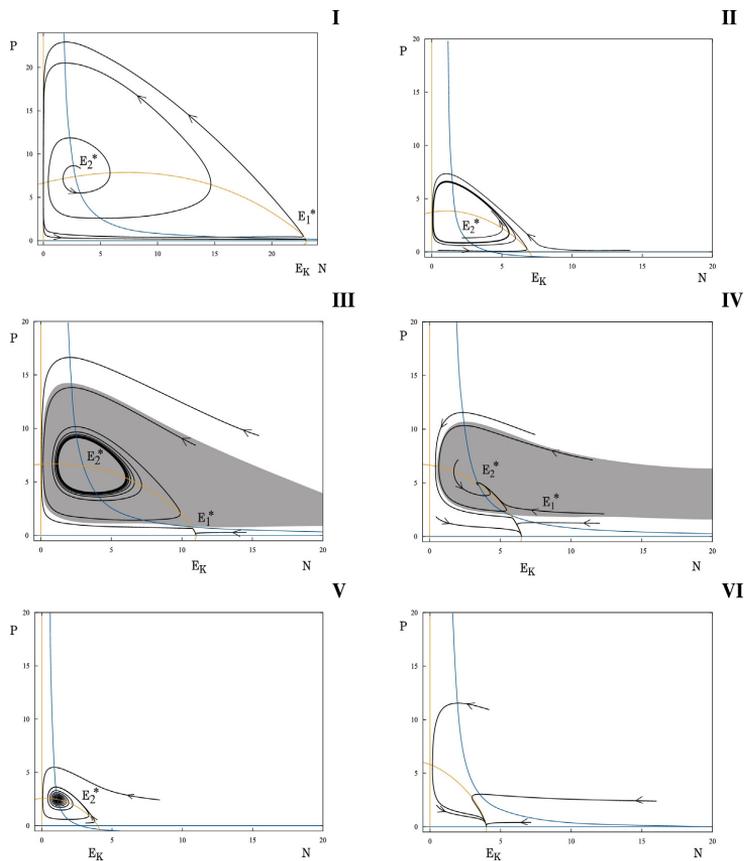
Fig. 7. Dynamics of model (3) with the bounded encounter rate function (7) for $e = 1$, $m = 1$, $h = 0.25$, $r = 3$, $L = 5$, $\lambda_0 = 5.1783$, $b = 2$. I: unstable focus E_2^* and stable node E_K for $w = 0.06$, $K = 23$. II: two unstable equilibria E_K and E_2^* and a stable limit cycle for $w = 0.15$, $K = 7$. III: unstable focus E_2^* and a stable limit cycle with its basin of attraction for $w = 0.06$, $K = 11$. IV: one stable node E_K and two coexistence equilibria: stable equilibrium E_2^* with its basin of attraction (grey) and saddle point E_1^* for $w = 0.06$, $K = 6.5$. V: one stable coexistence equilibrium E_2^* for $w = 0.2$, $K = 4$. VI: one stable node E_K for $w = 0.07$, $K = 4$.

conditions. In case of the unbounded encounter rate (4), the existence of a positive root of (24) implies the existence of an appropriate λ_0. In case of (7), the condition $\lambda(P^*) > 0$ has to be fulfilled. The effect of the Bogdanov–Takens bifurcation on the evolution of the population densities is described in Section 4. The numerical continuation has been carried out for the bounded above encounter rate function (7). Even though the Bogdanov–Takens bifurcation point has no relevant biological interpretation for this specific continuation, the one-parameter bifurcation manifolds continued from this point divide the parametric space into domains with significant features. These domains correspond to qualitatively different dynamical regimes that are biologically highly important. Regime shifts near the above mentioned bifurcation manifolds are predictable and continuable. The topological features of the manifolds in the neighborhood of the Bogdanov–Takens manifold does not depend on the specific form of the encounter rate function. That implies that the pure existence of the Bogdanov–Takens bifurcation guarantees wild changes of the behavior of the model nearby the bifurcation manifold for biologically significant parameters. These changes represent a generic feature.

4. Results of the bifurcation analysis

In Section 3 we have described the nonlinear phenomena we have found in system (3) with foraging facilitation. In this section we provide complete description of the dynamical changes that may occur in system (3) with the bounded above encounter rate function (7). The bifurcation diagram in Fig. 6 shows the equivalence partitioning of the parameter space w versus K into separate regions with topologically different phase space properties. The corresponding phase portraits are shown in Fig. 7. The bifurcation diagram in Fig. 6 expresses a generic property of system (3) with foraging facilitation. Different parameter values and different encounter rate functions generate quantitatively slightly different results but the qualitative topological features are equivalent. Qualitative equivalence is guaranteed by the fact that one may continue all bifurcation manifolds from the Bogdanov–Takens bifurcation manifold. Our results show the genericity of the described phenomena in the sense, that if the Bogdanov–Takens bifurcation is present in the predator–prey model with predator facilitation, all of the consequent phenomena are detectable as well. The bifurcation analysis was created by the MatCont continuation software.

The survival diagram depicted in Fig. 8 summarizes the obtained information on the predator and prey populations in scenarios that may take place for varying values of parameters w and K. There are six different scenarios that may occur – the corresponding phase portraits are shown in Fig. 7.

For low carrying capacity of the prey K (area VI in Fig. 6), no matter how high or low the predator foraging facilitation rate w is, only the prey survives. This finding corresponds with Courchamp et al. (1999) and with empirical studies such as Creel and Creel (1995) and Creel (1997)). For higher values of K, after passing the fold bifurcation manifold (into area IV), the system has two stable states. The prey-only equilibrium E_K is still stable and a stable coexistence equilibrium appears. The prey density at this coexistence equilibrium is lower than K. How the populations evolve in time depends on the initial conditions. For slightly larger values of K and larger w the basin of attraction of the coexistence equilibrium gets bigger, which means that the predator population is prone to survive. There is a minimum predator population density to prevent extinction, which can be explained as a type of an inner Allee effect. Such a type of the Allee effect mechanism has been mentioned already in Courchamp et al. (2008) and Berec (2010). Crossing the transcritical bifurcation manifold (into area V), the coexistence

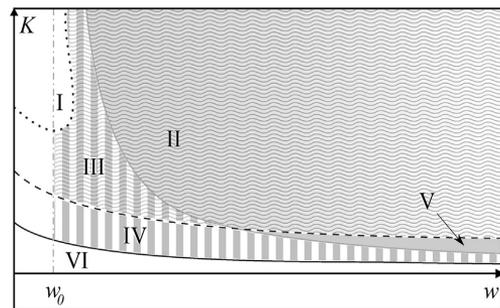Prey population survival
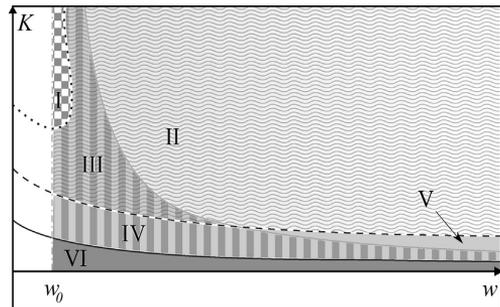
Predator population survival

Fig. 8. Survival of the prey and predator populations according to the dynamics of system (3) with encounter rate (7) for various facilitation rates w and environmental carrying capacities K. All other parameters are fixed at $e = 1$, $m = 1$, $h = 0.25$, $r = 3$, $L = 5$, $\lambda_0 = 5.1783$, and $b = 2$. Extinction regions are shown in white, stable survival in grey. Vertical stripes correspond to two stable equilibria. White-grey stripes show regions where the population may go extinct (when the initial conditions are out of the basin of attraction of the stable coexistence equilibrium); while the light and dark grey stripes correspond to two levels of population equilibria. Waves correspond to oscillations and the checkerboard pattern corresponds to areas where the trajectories pass very close to extinction point and survival of the population may be a matter of chance.

equilibrium becomes globally stable for higher facilitation w. For parameter values from this area the positive interaction among the predator population stabilizes the densities in a coexistence equilibrium.

The Hopf bifurcation curve separates the stable co-equilibrium area (area IV and V) from the area with a stable limit cycle (area III and II). The higher the carrying capacity K, the bigger the amplitude of the limit cycle. This phenomenon corresponds with the findings about the paradox of enrichment – increasing the carrying capacity evokes periodic behavior of the population densities, destabilizes the system, and may lead even to extinction. The population densities oscillate for parameter values from area III, but the populations survive. However, for high enough K, the limit cycle gets bigger and the population densities can reach extremely low levels. Due to unpredictable events this may lead to the extinction of the predator or to the extinction of both populations. The prey-only equilibrium is still stable; some initial conditions lead to the survival of the prey and the extinction of the predator. Crossing the transcritical bifurcation curve (from area III into area II) changes stability of the prey only equilibrium. For higher foraging facilitation rates w the population densities oscillate. Area III
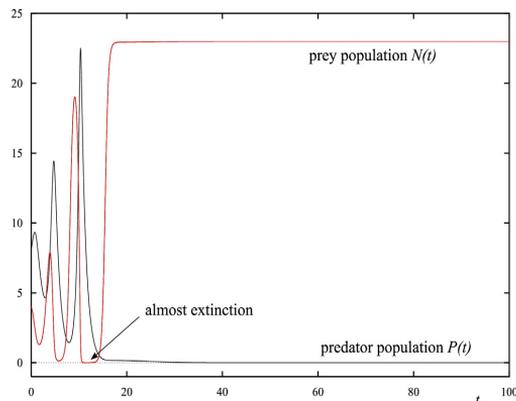
Fig. 9. Dynamics of the prey and predator populations of system (3) with encounter rate (7) for parameters $K = 23$, $w = 0.06$, $e = 1$, $m = 1$, $h = 0.25$, $r = 3$, $L = 5$, $\lambda_0 = 5.1783$, and $b = 2$ and for initial conditions $N(0) = 4$ and $P(0) = 8$. Notice the closeness to extinction point. The survival of the prey population may be a matter of chance.

and area I are separated by the homoclinic bifurcation curve, on which the stable limit cycle splits and the prey-only equilibrium becomes the only stable state. The predator population becomes overpopulated first followed by extinction while the prey population can survive at the carrying capacity. However, the trajectories pass so close to the trivial equilibrium that a small perturbation would lead to the extinction of both populations (see Fig. 9). Added Allee effect in the prey population may stabilize the trivial equilibrium. This phenomenon could motivate further studies. The provided bifurcation analysis has shown that although for specific facilitation rates the positive interaction of the predators may be beneficial, in most of the cases it can have fatal consequences for the predators themselves.

5. Summary

Several papers have emphasized the influence of interactions, both intra- and interspecific, on predator–prey populations (e.g. Abrams and Ginzburg, 2000; Arditi and Akcakaya, 1990; Lima, 2002). Observations and empirical studies have given information about interacting behavior in real populations; negative (interference) and positive interaction (facilitation) can be observed on different trophic levels (Packer and Ruttan, 1988; McCarthy and Ginzburg, 1995; Creel and Creel, 1995; Bertness et al., 1999; Bijleveld et al., 2015). Intraspecific predator interference has received quite some attention in the literature (e.g. Arnqvist et al., 2006; Ruxton et al., 1992; Ruxton, 1995), less is known about predator facilitation. Facilitation or cooperation is often modeled as a social dilemma in game theory. Packer and Ruttan (1988) claimed, that cooperation is not always the evolutionarily stable strategy for the predator. In the present paper, by means of bifurcation analysis of a predator–prey model we have proved that predator facilitation is not always beneficial for the predator population.

We have introduced an increasing encounter rate function based on our assumption that the predator encounter rate is higher for higher predator densities. In this paper we study the coexistence equilibria and the bifurcations they may undergo for specific parameter values; the fold, Hopf, transcritical, and Bogdanov–Takens bifurcations are all present in our model. We have derived the necessary and sufficient conditions of these phenomena as well as the formulae describing the bifurcation

manifolds for a system with a general encounter rate function and also two specifically chosen encounter rate functions. The unbounded encounter rate function was proposed by Berec (2010), while the bounded above encounter rate function introduced in this paper corresponds with game theory models (Packer and Ruttan, 1988). This function has never been studied before to the best of our knowledge. The accomplished bifurcation analysis of the system shows that predator facilitation or predator cooperation is not always beneficial for the predator. We have found that foraging facilitation is fatal for predators in case of low prey densities. The density threshold is given by the fold bifurcation manifold (12). This result is in agreement with previously published works (e.g. Courchamp et al., 1999, 2008; Berec, 2010); these state that predator foraging facilitation is a mechanism that invokes the Allee effect on the side of the predator. In case of high prey carrying capacities, low predator foraging facilitation is dangerous for both populations. The threshold is given by the homoclinic bifurcation manifold that can be continued only numerically from any point of the Bogdanov–Takens bifurcation manifold given by (25). For some middle level of carrying capacity, a very high rate of predator foraging facilitation is the safest and most stable. This stable area is bounded by the transcritical (23b) and the Hopf (16) bifurcation manifolds satisfying (18).

Our most important finding that predator facilitation is not beneficial in most cases may explain why predator cooperation in nature is only observed in specific scenarios. Cooperating only in certain conditions seems to be an evolutionary advantage in many species. In the future, we intend to continue our research by including the Allee effect in our model and by studying additional interference dominant at high population densities.

References

Abrams, P.A., Ginzburg, L.R., 2000. The nature of predation: prey dependent, ratio dependent or neither? Trends Ecol. Evol. 15, 337–341.
Arditi, R., Akcakaya, H.R., 1990. Underestimation of mutual interference of predators. Oecologia 83, 358–361.
Arnqvist, G., Jones, T.M., Elgar, M.A., 2006. Sex-role reversed nuptial feeding reduces male kleptoparasitism of females in Zeus bugs (*Heteroptera Veliidae*). Biol. Lett. 2, 491–493.
Beddington, J.R., 1975. Mutual interference between parasites or predators and its effect on searching efficiency. J. Anim. Ecol. 44, 331–340.
Begon, M., Harper, J.L., Townsend, C.R., 1990. Ecology: Individuals, Populations and Communities, 2nd ed. Blackwell Scientific Publications, Oxford.
Berec, L., 2010. Impacts of foraging facilitation among predators on predator–prey dynamics. Bull. Math. Biol. 72, 94–121.
Bertness, M.D., Leonard, G.H., Levine, J.M., Bruno, J.F., 1999. Climate-driven interactions among rocky intertidal organisms caught between a rock and a hot place. Oecologia 120, 446–450.
Bijleveld, A.I., van Gils, J.A., Jouta, J., Piersma, T., 2015. Benefits of foraging in small groups: an experimental study on public information use in red knots *Calidris canutus*. Behav. Process. 117, 74–81.
Boukal, D.S., Berec, L., Krivan, V., 2008. Does sex-selective predation stabilize or destabilize predator–prey dynamics? PLoS ONE 3 (7), e2687.
Brown, J.S., Laundre, J.W., Gurung, M., 1999. The ecology of fear: optimal foraging, game theory, and trophic interactions. J. Mammal. 80, 385–399.
Bruno, J.F., Stachowitz, J.J., Bertness, M.D., 2003. Inclusion of facilitation into ecological theory. Trends Ecol. Evol. 18, 119–125.
Cosner, C., DeAngelis, D., Ault, J.S., Olson, D.B., 1999. Effects of spatial grouping on the functional response of predators. Theor. Popul. Biol. 56, 65–75.
Courchamp, F., Berec, L., Gascoigne, J., 2008. Allee Effects in Ecology and Conservation. Oxford University Press, Oxford.
Courchamp, F., Clutton-Brock, T., Grenfell, B., 1999. Inverse density dependence and the Allee effect. Trends Ecol. Evol. 14, 405–410.
Courchamp, F., Macdonald, D.W., 2001. Crucial importance of pack size in the African wild dog *Lycaon pictus*. Anim. Conserv. 4, 169–174.
Creel, S., Creel, N.M., 1995. Communal hunting and pack size in African wild dogs, *Lycaon pictus*. Anim. Behav. 50, 1325–1339.

Creel, S., 1997. Cooperative hunting and group size: assumptions and currencies. Anim. Behav. 54, 1319–1324.

DeAngelis, D.L., Goldstein, R.A., O'Neill, R.V., 1975. A model for tropic interaction. Ecology 56, 881–892.

Doebeli, M., Hauert, C., 2005. Models of cooperation based on the prisoner's dilemma and snowdrift game. Ecol. Lett. 8, 748–766.

Fryxell, J.M., Mosser, A., Sinclair, A.R.E., Packer, C., 2007. Group formation stabilizes predator–prey dynamics. Nature 449, 1041–1044.

Giraldeau, L.-A., Caraco, T., 2000. Social Foraging Theory. Princeton University Press, Princeton, NJ.

Heinsohn, R., Packer, C., 1995. Complex cooperative strategies in group-territorial African lions. Science 269, 1260–1262.

Holling, C.S., 1959. Some characteristics of simple types of predation and parasitism. Can. Entomol. 91, 385–398.

Iwasa, Y., 1982. Vertical migration of zooplankton: a game between predator and prey. Am. Nat. 120, 171–180.

Kot, M., 2001. Elements of Mathematical Ecology. Cambridge University Press, Cambridge.

Krause, J., Ruxton, G.D., 2002. Living in Groups. Oxford University Press, Oxford.

Krebs, C.J., 2001. Ecology, 5th ed. Benjamin Cummings, San Francisco.

Kuznetsov, Y.A., 1998. Elements of Applied Bifurcation Theory, Applied Mathematical Sciences, 2nd ed., vol. 112. Springer-Verlag, Berlin, Heidelberg, New York.

Lima, S.L., 2002. Putting predators back into behavioral predator–prey interactions. Trends Ecol. Evol. 17, 70–75.

McCarthy, M.A., Ginzburg, L.R., 1995. Predator interference across trophic chains. Ecology 76, 1310–1319.

Nilsson, A.P., Lundberg, P., Bronmark, C., Persson, A., Turesson, H., 2006. Behavioral interference and facilitation in the foraging cycle shape the functional response. Behav. Ecol. 18, 354–357.

Packer, C., Ruttan, L., 1988. The evolution OD cooperative hunting. Am. Nat. 132, 159–198.

Pribylova, L., Berec, L., 2014. Predator interference and stability of predator–prey dynamic. J. Math. Biol. 69, 1–23.

Richerson, P.J., Boyd, R., 2001. The Evolution of Subjective Commitment to Groups: A Tribal Instincts Hypothesis. Evolution and the Capacity for Commitment. Russell Sage, New York, pp. 186–220.

Rosenzweig, M.L., 1971. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. Science 171, 385–387.

Ruxton, G.D., 1995. Short term refuge use and stability of predator–prey models. Theor. Popul. Biol. 47, 1–17.

Ruxton, G.D., Gurney, W.S.C., Roos, A.M.D., 1992. Interference and generation cycles. Theor. Popul. Biol. 42, 235–253.

Skalski, G.T., Gilliam, J.F., 2001. Functional responses with predator interference: viable alternatives to the Holling type II model. Ecology 82, 3083–3092.

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Original Research Article

Regime shifts caused by adaptive dynamics in prey–predator models and their relationship with intraspecific competition

Lenka Přibylová

Department of Mathematics and Statistics, Section of Applied Mathematics, Faculty of Science, Masaryk University, Kotlářská 2, Brno 611 37, Czech Republic

ARTICLE INFO

ABSTRACT

The paper concerns with regime shifts between multiple attractors in ecological predator–prey models and hysteresis phenomena caused by evolution. We present a survey of eco-evolutionary models with an adaptive trait affecting the prey defence or activity that influence predator functional response and give overview of typical consequences of the trait evolution to the predator–prey dynamics together with important references to related adaptive dynamics research. The selection and mutation process is modelled by a resident-mutant model (possible mutant invasion into a monomorphic resident population). Model derivations are given in detail for all of the common functional responses (Holling's type I, II, III and generalized). Different types of adaptive trait value dependences with respect to transient dynamics are distinguished according to the effect to the eco-system: we prove that if the prey adaptive trait evolution influence only the functional response of the predator, stable dynamics and irreversible abrupt regime changes are typical, whereas reversible regime shifts or more complex dynamics caused by adaptivity of the prey trait occur for trait adaptations that bring an advantage against predator together with intraspecific competition asymmetry. We confirm possibility of hysteresis eco-evolutionary cycle, persistent oscillations between different attractors of the ecological subsystem driven by adaptive trait dynamics.

1. Introduction

Population dynamics is an important part of biology. Population growth and ecological interactions are still studied throughout last decades enormously although the basic principles and ideas are known for more than a century. The fundamental works that concern the laws of exponential and logistic population growth of one-species populations or the basic interaction models have been modified, generalized and reformulated many times. Google Scholar returns almost 5 million links for the search phrase "population growth model"! A considerable part of the research on population dynamics addresses the question of the determinants of population growth and of the interactions among populations. Apart from this general qualitative approach more specific models are used, especially for the purposes of management and control, see e.g. coral-algae growth models (Mumby et al., 2007)), specific food-chain models (Kuznetsov et al., 2001) and so on. New effective technologies allow us to study individual-based models, see e.g. Grimm et al. (2003) and to accomplish computer simulation models as in e.g. Boit et al. (2012).

A typical system that describes a dynamical population model (when spatial distribution is omitted) is a system of parameter dependent ordinary differential equations

$$\dot{n} = \phi(n, a), \tag{1}$$

where n is a population density vector of species in the ecosystem and a is a vector of parameters as birth rates of particular species, their carrying capacities and so on. There is a plenty of studies of such systems that analyse dependence of the long-term behaviour on parameters (see e.g. Boukal et al., 2007; Kar, 2006; Mohammed et al., 2018; Mumby et al., 2007; Rinaldi et al., 1993; Scheffer et al., 1997 and many others). These works give insight to the principles involved, give possibilities to manage and control the systems, but also show that rapid changes and unexpected behaviour can happen. As a good and well-known example may serve the spruce budworm model introduced in Ludwig et al. (1978) that explains hysteresis loops in the dynamics of the population density of the budworm or variety of prey–predator models with stable limit cycles (Abrams and Walters, 1996; Boukal et al., 2007; Rosenzweig and MacArthur, 1963; Steele and Henderson,

1992 and others). Besides the multiple equilibria and the limit cycles, more complex behaviour can be observed in some models: multiple cycles and other attractors are studied in e.g. González-Olivares and Rojas-Palma (2011) and the presence of chaotic dynamics in ecological models is shown in some ecological models (see e.g. Hastings and Powell, 1991; Huisman and Weissing, 2002; Kuznetsov et al., 2001).

Evolution as a change in the inherited traits of biological populations give the species possibility to adapt to their environments by means of natural selection. During the last two decades techniques based on game theory were developed and they are referred to as adaptive dynamics techniques (see e.g. Dercole and Rinaldi, 2008; Dieckmann and Law, 1996; Geritz et al., 1997; Metz et al., 1995). They link population dynamics to evolutionary dynamics for understanding the long-term ecological and evolutionary consequences of small mutations in the traits expressing the phenotype. Although the adaptive dynamics theory is still in development (for new results see e.g. Della Rossa et al., 2015; Dercole, 2005; Dercole, 2016; Dercole et al., 2016; Geritz et al., 2016), already obtained results have often been surprising and have given a very promising insight into the complexity of nature. There are also possibilities to use the adaptive dynamics techniques in other fields as economics, since for example innovations of products may be viewed as adaptive traits and the market equilibria existence and stability depend on the innovation dynamics (see Dercole et al., 2008).

In this paper an ecosystem (1) is considered in this eco-evolutionary context. The species are affected by adaptive change in the trait that creates feedback between the evolutionary and population dynamics. Generally, the adaptive changes in traits are described by so called canonical equation (Dieckmann and Law, 1996), a system of equations

$$\dot{x} = \varepsilon g(n, a, x), \tag{2}$$

where x is generally a vector of traits of the resident species. Hereafter, only one main trait is considered, so the equation is scalar. The evolution process does not have the same timescale as the population dynamics, so the parameter $\varepsilon > 0$ scales the rate of evolutionary change. Usually $\varepsilon \ll 1$ and Eq. (2) is slow with regard to the ecological timescale of Eq. (1). Function g is explained in more detail in the next section, for now it is enough to mention that it covers the feedback of the population dynamic process and that it is generally nonlinear. This feature is responsible for bifurcations.

Bifurcations connected to adaptive trait can be responsible for regime shifts in ecosystem (1). Trait value x may affect the vector parameter a and so it can be (at least locally) considered as a function $a = a(x)$. The trait value influences the population dynamics throughout this dependence. If ecosystem (1) reaches a stable equilibrium for a specific trait value x_r of the monomorphic resident population, the system stays at the stable equilibrium until that trait value belongs to an evolutionarily stable strategy. This condition may be violated and the feedback between population and adaptive dynamics can cause transitions from one regime to another (Dercole, 2003), abrupt and unexpected changes of population densities (Dercole, 2005; Parvinen, 2005), diversification of the species to polymorphic populations (Dercole et al., 2016; Gallien et al., 2018; Landi et al., 2013 or Hui et al., 2017) and so on.

In this study we focus on monomorphic populations with slow adaptive change of the trait that is strongly connected to some parameters of the population dynamic model (1). Of course that it is not sufficient to answer even basic questions about the persistence of the population or its stabilization at some equilibrium or attractor by analysis of the eco-system alone if mutations are taken into account (for example in Dercole et al., 2003; Landi et al., 2013 Rosenzweig-MacArthur model was transformed into a resident-mutant model by adding a third equation for the mutant population and it was shown that the evolutionary model is much richer than the resident population model). The aim of this survey is not only to present different types of ecological regime transitions between multiple ecological attractors caused by

adaptive dynamics, but also to distinguish typical cases when ecosystem stays stable and when the evolution causes reversible or irreversible transients. We prove that for one specialized adaptive trait changes (for example if the trait adapts specially to prevent the prey against predators and does not significantly affects other parameters of the ecosystem), only the stable case or an irreversible regime shift is possible, whereas reversible regime shifts or more complex dynamics caused by adaptivity of the prey trait occur for trait changes that bring in advantage against predator together with intraspecific competition asymmetry.

The analysis of a population dynamic model is, of course, a first step to understand the long-term behaviour of an ecosystem, especially in cases of more complex dynamics with multiple or chaotic attractors. Quasi-equilibria and quasi-attractors of the fast subsystem are essential, but the transitions between them depend on the adaptive dynamics (the slow system) strongly. Generally, the trait values are not the parameters of the population dynamic model, but the parameters are functions of the trait values. Adaptive dynamics works according to evolutionary game theory principles and that explains why additional slow nonlinear equations give feedback to the population dynamics. The slow dynamics shift the parameters of the fast population dynamic subsystem that tends to stabilize on quasi-attractors, but different types of regime transitions may happen. These transitions are not the same as the transitions caused by parameter shifts at the population dynamic subsystem (usually described by bifurcation diagrams of population models). As a prototype example of an evolving community, a prey–predator community with multiple attractors at the population dynamic subsystem is taken. The adaptive dynamics involve one trait of the prey that affects the parameters of the functional response of the predator. Such a model serves as a good demonstrative example, because the modelled situation is easy to imagine and also to explain. One may imagine the mutant prey phenotype change as getting stronger through a genetic mutation and selection process and therefore the new resident prey with adapted trait is able to defend itself more effectively from a predator. It is clear from this example, that the handling time, the encounter rate parameters or other parameters in the functional response may be influenced and varied by the trait value. This parametric change may significantly affect population dynamics near local bifurcation of the resident equilibrium.

We present cases with qualitatively preserved eco-dynamics and also cases with expected irreversible regime transitions. We also present an example of a system, when adaptive dynamics induce hysteresis and periodic regime shifts. The presented results are in agreement with Dercole et al. (2002) and Muratori and Rinaldi (1991). More complex dynamics may occur (see Dercole et al., 2010; Wilsenach et al., 2017 or Dercole and Rinaldi, 2010).

2. Adaptive dynamics and slow-fast eco-evolutionary system

We assume the resident population model as the system (1) of the form

$$\dot{n} = \phi(n, a),$$

where $n = (n_1, \cdots, n_N)$ is a population density vector for an arbitrary number N of species (the species are characterized by an index $i = 1, \cdots, N$) and $a = (a_1, \cdots, a_m)$ is a vector of parameters. The change in the population size or density is described by function $\phi: \mathbb{R}^{N+m} \to \mathbb{R}^N$, that is a smooth enough function dependent on parameters such as the birth and the death rates of the species, carrying capacities and various parameters of functional responses. The parameters $a = a(x)$ of the system (1) are functions of the resident trait value $x = x_r$. We assume that system (1) dynamics is fast, so the resident population settles in a dynamical equilibrium $n^* = n^*(x)$.

Mutations are assumed to be sufficiently rare, but once a mutant has entered the population, it may grow or go extinct according to the invasion exponent (fitness) of mutants in the resident population. The

invasion exponent can be interpreted as the initial exponential growth rate of mutants (Diekmann et al., 2004; Metz et al., 1992 or Hui et al., 2016). The mutant population trait value is labelled as x_m. The invasion exponent is a function $W(x_m, x_r, n^*)$ of the mutant and resident trait values, the population densities at the ecosystem and parameter a. Parameter a is assumed to be a function of the trait value, that is $a(x_r)$ for the resident population and $a(x_m)$ for the mutant population. If the mutant population invades with a positive invasion exponent, it becomes a new resident (Dercole and Geritz, 2016). An evolutionary change of an adaptive trait caused by a slightly different growing mutant population occurs on a much longer timescale than that of individuals' births, interactions and deaths, which determine the ecological dynamics of population. The approach of adaptive dynamics theory provides a deterministic approximation of the underlying stochastic processes of mutation and selection and the final result is that the trait values x vary in accordance with canonical equation

$$\dot{x} = \varepsilon\, n^*\, s\,(n^*, x), \tag{3}$$

where

$$s(n^*, x_r) = \left.\frac{\partial W(x_m, x_r, n^*)}{\partial x_m}\right|_{x_m = x_r} \tag{4}$$

is the selection gradient in the resident equilibrium n^* and x denotes the adaptive trait value of the dynamically resident population, indexed x_m and $x = x_r$ distinguish between mutant and resident trait values, non-negative $\varepsilon \ll 1$ scales the rate of evolutionary change. For more details see e.g. Dercole and Rinaldi (2008) or Dieckmann and Law (1996).

Together this yields a slow-fast eco-evolutionary system

$$\dot{x} = \varepsilon\, n\, s(n, x), \tag{5}$$

$$\dot{n} = f(n, x) \tag{6}$$

with positive $\varepsilon \ll 1$.

A typical property of a slow-fast system is that nullclines or attractors of the fast system are reached much faster than attractors of the whole system. Consequently the fast subsystem (6) that corresponds to the population dynamics is highly important to study. The long-term behaviour of the adaptive eco-evolutionary system is determined by the stable attractors of the fast subsystem (6). On the other hand, the slow dynamics of the adaptive trait may cause these attractors to lose their stability in the long evolutionary timescale, so the partial analysis of the fast subsystem is not sufficient to answer basic questions about the extinction or persistence of the populations, stabilizing of population densities, periods of oscillations and so on. It is obvious that the answers to these questions are influenced by many factors, even if we assume very simple eco-evolutionary models. For example, various types of functional responses of predator in the predator–prey model may pose various attractors of the fast subsystem, but also various dependences on parameters and adaptive trait values and various selection gradients. It depends on the situation, which model is more appropriate.

In the next sections, we would like to present typical eco-evolutionary regime shifts connected to local bifurcations of equilibria with aim to answer the questions that arose in the text above. Firstly, what types of the adaptive trait value dependences do not influence the population dynamics massively. Secondly, whether or when abrupt regime changes caused by adaptive dynamics are possible, expectable or predictable. And thirdly, we will show more complex dynamics, than the irreversible regime shifts: persistent oscillations between different attractors of the fast ecological subsystem driven by adaptive trait dynamics as a hysteresis eco-evolutionary cycle.

All these typical dynamics will be demonstrated on well-known and standard examples of population models: a one-species model with Holling's type I, II and III functional response and generalized functional responses. What can be useful to the readers is the fact that resident-mutant models and the corresponding eco-evolutionary systems

for all these functional responses are derived properly. We will assume that the prey trait adapts to increase the protection against predators, so it influences the predation rate and the handling time. We may interpret the increasing protection as a more efficient defence, or getting stronger, bigger, or as many other mechanisms. This adaptive trait value may also cause trait-dependent intraspecific competition between the resident and mutant prey and this case is also studied and found important to distinguish.

3. Resident-mutant model in the prey population

Let us consider a resident prey population model

$$\dot{n} = r n \left(1 - \frac{n}{K}\right) - N \varphi(n), \tag{7}$$

where r is a specific growth rate, K is the carrying capacity of the environment, N is predator population density and $\varphi(n)$ is the functional response of the predator, that is a number of prey killed by predator per predator per a unit of time. Function φ depends on other parameters that will be specified later. Functional responses used in standard population models are usually the Holling's type responses or their generalized types and that is the reason why we focus on Holling's type functional responses in this paper. The population of the prey is modelled by a logistic function in the absence of the predator, but it can be changed. The Allee effect or any other appropriate growth function can be implemented easily.

The derivation of the resident-mutant model follows the method described in Dercole and Rinaldi (2008). The derivation of the model with Holling's type I or II functional responses is relatively simple, but that is not the case with the Holling's type III and the generalized functional responses. That's why we present a proper derivation of the generalized Holling's type III functional response for resident n_r and mutant population n_m here. For detailed explanation of Holling's type functional responses derivations see Dawes and Souza (2013). To keep the resident-mutant models derivations simple, it is assumed that only the encounter rate depends on the adaptive trait of the prey, but the handling time is not. This case with non-adaptive handling time is also used for simulations later. The resident-mutant model for the case of simultaneous encounter rate and handling time adaptivity is analogous and its functional responses (15) are presented without derivation.

Let Δn_r and Δn_m be the number of resident and mutant prey consumed by one predator, Δt_t is the total time for predator to encounter a prey and Δt_s is time available for searching. In case of the type I response, Δn_r and Δn_m depend linearly on prey densities. The linear dependence coefficient is the probability of a given predator encountering prey in a fixed time interval $\Delta t_t = \Delta t_s$ in some fixed area, predators spend no time to handle the prey or this is omitted. This probability generally depends on the adaptive trait. In cases of Holling's type II, III and generalized functional responses, the searching time is the total time decreased by the handling time of the encountered prey, that is

$$\Delta t_s = \Delta t_t - h \cdot (\Delta n_r + \Delta n_m), \tag{8}$$

where h is the handling time of one prey (assumed here as non-dependent on the adaptive trait). Now

$$\Delta n_r = \Lambda(n_r, n_m, x_r) \cdot n_r \cdot \Delta t_s, \quad \text{and} \quad \Delta n_m = \Lambda(n_r, n_m, x_m) \cdot n_m \cdot \Delta t_s, \tag{9}$$

where the function Λ is the encounter rate that depends on the adaptive trait of the prey and possibly on the prey population. It is assumed that the predator encounters the prey with some efficiency rate that is assumed to be a constant λ for Holling's type II functional response or linearly growing with the prey density λn in the case of Holling's type III functional response. This is motivated by the assumption that learning behaviour occurs in the predator population with a consequent increase in the discovery rate as more encounters with prey occur, see Dawes and Souza (2013). Generalized Holling's type III functional response is derived for an encounter rate λn^{k-1} for $k > 1$ that is increasing

with n, while $k = 1$ gives Holling's type II functional response (and Holling's type I functional response is the case of $h = 0$). Herein, it will be assumed that the efficiency rate λ depends on the adaptive trait, $k \geq 1$ and $h > 0$ and the learning in the predator population is affected by the whole population $n = n_r + n_m$. The derivation is different, if the learning process is diversified.

Putting (8) and (9) together with the encounter rate derivation gives

$$\Delta n_r = \lambda_r(n_r + n_m)^{k-1} \cdot n_r \cdot (\Delta t_t - h \cdot (\Delta n_r + \Delta n_m)), \tag{10}$$

$$\Delta n_m = \lambda_m(n_r + n_m)^{k-1} \cdot n_m \cdot (\Delta t_t - h \cdot (\Delta n_r + \Delta n_m)), \tag{11}$$

where λ_r and λ_m stay for the efficiency rates on the prey resident and mutant populations. Consequently after division by Δt_t, the functional responses $\varphi(n_r) = \frac{\Delta n_r}{\Delta t_t}$ for the resident and $\varphi(n_m) = \frac{\Delta n_m}{\Delta t_t}$ for the mutant population satisfy a set of equations

$$\begin{aligned}\varphi(n_r) &= \lambda_r(n_r + n_m)^{k-1} \cdot n_r \cdot (1 - h \cdot (\varphi(n_r) + \varphi(n_m))), \\ \varphi(n_m) &= \lambda_m(n_r + n_m)^{k-1} \cdot n_m \cdot (1 - h \cdot (\varphi(n_r) + \varphi(n_m))),\end{aligned} \tag{12}$$

for $\varphi(n_r)$ and $\varphi(n_m)$ that gives solutions

$$\varphi(n_r) = \frac{\lambda_r n^{k-1} n_r}{1 + h\overline{\lambda} n^k}, \tag{13}$$

$$\varphi(n_m) = \frac{\lambda_m n^{k-1} n_m}{1 + h\overline{\lambda} n^k} \tag{14}$$

with $n = n_r + n_m$ and the mean efficiency value $\overline{\lambda} = \lambda_r \frac{n_r}{n} + \lambda_m \frac{n_m}{n}$. An analogous derivation with the handling time codependence on the adaptive trait leads to functional responses

$$\begin{aligned}\varphi(n_r) &= \frac{\lambda_r n^{k-1} n_r}{1 + \overline{h\lambda} n^k}, \\ \varphi(n_m) &= \frac{\lambda_m n^{k-1} n_m}{1 + \overline{h\lambda} n^k},\end{aligned} \tag{15}$$

where $\overline{h\lambda} = h_r \lambda_r \frac{n_r}{n} + h_m \lambda_m \frac{n_m}{n}$ is the mean value of the handling time times efficiency rate.

The resident-mutant model with no trait-dependent intraspecific competition in the prey population is

$$\begin{aligned}\dot{n}_r &= r n_r \left(1 - \frac{n_r + n_m}{K}\right) - N \frac{\lambda_r n^{k-1} n_r}{1 + h\overline{\lambda} n^k}, \\ \dot{n}_m &= r n_m \left(1 - \frac{n_r + n_m}{K}\right) - N \frac{\lambda_m n^{k-1} n_m}{1 + h\overline{\lambda} n^k}.\end{aligned} \tag{16}$$

The resident-mutant model with a trait-dependent intraspecific competition in the prey population is

$$\begin{aligned}\dot{n}_r &= r n_r \left(1 - \frac{n_r}{K} - \alpha(x_r, x_m)\frac{n_m}{K}\right) - N \frac{\lambda_r n^{k-1} n_r}{1 + h\overline{\lambda} n^k}, \\ \dot{n}_m &= r n_m \left(1 - \alpha(x_m, x_r)\frac{n_r}{K} - \frac{n_m}{K}\right) - N \frac{\lambda_m n^{k-1} n_m}{1 + h\overline{\lambda} n^k}.\end{aligned} \tag{17}$$

where $\alpha(x_r, x_m)$ is called the competition function (see MacArthur, 1970 or Gatto, 1990). The competition is said to be symmetric if $\alpha(x_r, x_m) = \alpha(x_m, x_r)$ for all possible trait values x_r and x_m and asymmetric otherwise. A natural assumption on the competition function is $\alpha(x, x) = 1$ for any x. Intraspecific competition is trait-independent in the case of $\alpha(x_r, x_m) \equiv 1$ for any possible trait values x_r and x_m.

4. Eco-evolutionary model for prey population with trait-independent intraspecific competition

In this section we use principles of the adaptive dynamics theory to model (16). The basic concept is the invasion exponent that is the expected growth rate of an initially rare mutant in the environment set by the resident. The invasion exponent can be interpreted as the fitness in evolutionary game theory. The selection gradient is the slope of the invasion exponent at trait value $x_m = x_r$, that is in the moment when

the mutation appears in the resident monomorphic population. The sign of the selection gradient determines an increase or a decrease in trait values of the invasion successful mutants that become a new resident population according to (5). Let's remind that the resident fast system starts at the equilibrium $n = n^*$.

Let us follow the analysis of the resident-mutant model with trait-independent intraspecific competition. The invasion exponent in the prey population (16) is

$$W(x_m, x_r, n) = \frac{\dot{n}_m}{n_m}\Big|_{n_m \to 0} = r\left(1 - \frac{n}{K}\right) - N\lambda_m \frac{n^{k-1}}{1 + h\lambda_r n^k},$$

where $n = n_r$ is the resident population and it is assumed that the efficiency rate $\lambda_m = \lambda(x_m)$ is a locally defined smooth function at least near $x_m = x_r$ (the trait value is a measurable real quantity such as the prey escape speed). The selection gradient (4) is the slope of the invasion exponent of the mutant population in the monomorphic resident population

$$s(n, x_r) = \frac{\partial W(x_m, x_r, n)}{\partial x_m}\Big|_{x_m = x_r} = -N\lambda'(x_r)\frac{n^{k-1}}{1 + h\lambda(x_r)n^k}, \tag{18}$$

where $\lambda'(x) = \frac{d\lambda(x)}{dx}$. Consequently, the slow-fast eco-evolutionary system (5) and (6) has the form

$$\dot{x} = -\varepsilon n^*(x)\lambda'(x)\frac{n^{k-1}N}{1 + h\lambda(x)n^k}, \tag{19}$$

$$\dot{n} = rn\left(1 - \frac{n}{K}\right) - \frac{\lambda(x)n^k N}{1 + h\lambda(x)n^k}, \tag{20}$$

where $h = 0$ and $k = 1$ for Holling's type I functional response, $h > 0$ and $k = 1$ for Holling's type II functional response, $h > 0$ and $k = 2$ for Holling's type III functional response and $h > 0$ and $k > 1$ for generalized Holling's type functional response.

Let us assume that $x = x_r$ and the fast equation (20) has a stable equilibrium $n^*(x_r)$. The slow evolution equation (19) is in equilibrium for $\lambda'(x_r) = 0$ and it is clear that this equilibrium is stable if $\lambda(x_r)$ is a local minimum. It models the most effective prey mutant strategy against the predator. An evolution process that minimizes the predator encounter efficiency is advantageous and the minimum is the evolutionary stable strategy of the prey. If the evolution equation (19) is out of the equilibrium, the state of the eco-evolutionary system will copy the nullcline in the direction of $-\lambda'(x)$. That explains why the quasi-equilibria have to be studied.[1] It is very helpful that the ecological model equilibria are already well-studied in population biology models.

Let us focus on population models that exhibit multiple equilibria (for example using Holling's type III functional response for $k = 2$). What kind of impact will the evolutionary process have on the prey trait and equilibrium density? We assume that there exists a local minimum $\lambda(x_{\min})$ for a trait value $x_{\min}$.[2] The slow equation (19) has a vertical nullcline $x = x_{\min}$ (see Fig. 1), while $\dot{x} < 0$ on the right side of the nullcline and $\dot{x} > 0$ on the left side of the nullcline. The fast equation (20) nullcline that corresponds to quasi-equilibria is non-linear with possibly multiple stable equilibria on $x = x_{\min}$. Fig. 1 presents the example of model with an S-shaped quasi-equilibrium manifold (e.g. for Holling's type III functional response), where stable quasi-equilibria tend to the stable equilibrium on the upper branch of the S-shaped nullcline or on the lower branch of the S-shaped nullcline. Obviously,

---

[1] This is known very well for slow-fast systems in neuroscience or biochemistry, where the notion of quasi-equilibrium is commonly used.

[2] This seems to be a natural assumption, since the adaptive trait change has got some limits. For example, an increase of the handling time or a decrease of the encounter rate by getting the prey more strong or agile take costs in return, the prey is getting tired quickly and can be caught more easily after some time. Of course, some adaptive traits can have infimum at infinity, but this case can be treated in analogous way since also different trait quantification may be taken, for example infinity can be shifted to zero on the x-axes by its reciprocal value.
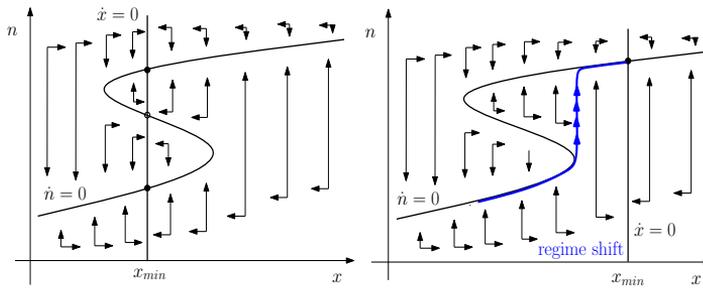
Fig. 1. A model with multiple quasi-equilibria with trait-independent intraspecific competition. Population density n versus trait value x on axes with nullclines and phase-diagram present stable (left figure) and transient (right figure) eco-evolutionary dynamics.

the S-shape of the nullcline can be distorted, mirrored or doubled due to reparametrization (the hysteresis is usually displayed in parameter versus population axes), the representation of the parameter λ as a function of the trait x versus n axes, but multiple quasi-equilibria generally appear. Fig. 1 is just a draft of this phenomenon. Specific simulations are in Figs. 2 and 3. Slight and slow changes of the trait value due to evolutionary dynamics may cause abrupt changes in the population if the quasi-equilibrium nullcline folding crosses the path of the adaptive dynamics of mutating trait value. In the left part of the Fig. 1 the minimum $x_{\min}$ lies in the multiple equilibria area and the adaptive dynamics keeps the quasi-equilibria on the same branch with no ecological regime shift. On the other hand, in the right part of the Fig. 1 the minimum $x_{\min}$ lies out of the multiple equilibria area and the adaptive dynamics can cause abrupt change while the quasi-equilibria branch disappears due to fold bifurcation. This is a regime shift that is not reversible afterwards.

Simulated eco-evolutionary dynamics with impossible (Fig. 2) and possible (Fig. 3) regime shifts are presented for the model (19) and (20) with $k = 2$ and $\lambda(x) = \lambda_0 + \lambda_1(x - x_{\min})^2$. The encounter rate is taken as the most simple function with a minimum in $x_{\min}$, but any other function with a minimum in $x_{\min}$ can be used (or its Taylor expansion near $x_{\min}$). Since the trajectories start from a quasi-equilibrium, n^* can be replaced by n. Parameters used in simulations are $r = 0.95$, $K = 10$, $N = 1$, $\varepsilon = 0.001$, $\lambda_1 = 1$, $x_{\min} = 2$ and $h = 0.4$. Parameter λ_0 is used to set the level of the encounter rate and it shifts the S-shaped λ versus n equilibrium manifold of a population model according to $x_{\min}$. In case $\lambda_0 = 1$, $x_{\min}$ is situated in the multiequilibria area of the $\dot{n} = 0$ null-cline. More precisely $\lambda(x_{\min})$ lies between the two fold bifurcation points of the λ versus n population equilibrium manifold. There is no possibility to shift the regime to a different branch for $\lambda_0 = 1$ (Fig. 2) since no fold is left in the direction of the evolution. On the contrary, abrupt irreversible regime shift is simulated for $\lambda_0 = 0.5$ (Fig. 2) when the fold bifurcation point is present.

These simulations are examples of a general phenomenon for systems with adaptive traits that cause no intraspecific competition. The presented results can be generalized. A resident-mutant model with trait-independent intraspecific competition has the form

$$\dot{n}_r = \rho_r n_r - N\varphi_r \tag{21}$$

$$\dot{n}_m = \rho_m n_m - N\varphi_m, \tag{22}$$

where the growth functions ρ_r and ρ_m does not depend on the mutant trait value, whereas the predator (or other) functional responses φ_r and φ_m do, so the selection gradient has to be in the form of a multiplication and the x-nullcline will be a vertical line. That implies the evolution dynamics cannot affect the globally stable equilibrium of the fast population dynamics if intraspecific competition is trait-independent, that is in the case that the mutant trait value affects only the predator functional response or more generally, in case that the adaptive trait is

getting highly specialized to a specific functional response. In the case of multiple equilibria it can cause abrupt changes from one regime to another and these changes are irreversible and stable in the eco-evolutionary context. More precisely ecosystems with quasi-equilibria manifolds that depend on an adaptive parameter and exhibit folds (typically near cusp bifurcation points) will stay on the same slowly changing branch tending to an equilibrium on it or the adaptive parameter will irreversibly cross the bifurcation point of the folded manifold and jump to a different branch (level) with a new stable equilibrium. The theory of bifurcations provides good tools to find the fold bifurcation points of the population dynamic models even for big systems with a lot of parameters and variables and the position of the adaptive parameter minimum together with the critical fold bifurcation manifold of quasi-equilibria give a strong tool to expect or even predict these dynamic transitions. Let us mention here that also new software tools are being developed to handle forecasting bifurcation-induced regime shifts by computing of the so called distance to bifurcation (see for example Dobson, 1993 or Tamba, 2015 with developed MATLAB Toolkit).

5. Eco-evolutionary model for prey population with trait-dependent intraspecific competition

It is very common that the trait value influences the population growth function too. Imagine a bird that has increased its escape speed. This mutant is possibly quicker in feeding too. Imagine a bug that has increased its shard hardness, possibly it took some costs that decrease its breeding or it has become heavier and that makes flying or feeding more difficult. This trait-dependent intraspecific competition should be taken into account in a lot of cases.

The slow-fast eco-evolutionary system (5) and (6) with trait-dependent intraspecific competition has the form

$$\dot{x} = -\varepsilon n^*(x) n \left(\frac{r}{K} \frac{\partial \alpha(x_m, x)}{\partial x_m} \bigg|_{x_m = x} + \lambda'(x) \frac{n^{k-2}N}{1 + h\lambda(x)n^k} \right), \tag{23}$$

$$\dot{n} = rn\left(1 - \frac{n}{K}\right) - \frac{\lambda(x)n^k N}{1 + h\lambda(x)n^k}, \tag{24}$$

where α is the intraspecific competition function from the resident-mutant model (17) and x_m stays for mutant and $x = x_r$ for resident trait values. The Taylor series representation of the function α in variable x_m at $x_m = x_r = x$ has the form

$$\alpha(x_m, x) = 1 + \alpha_1(x_m - x) + \alpha_2(x_m - x)^2 + \dots$$

and the term with $\alpha_1 = \frac{\partial \alpha(x_m, x)}{\partial x_m} \bigg|_{x_m = x}$ vanishes for symmetric trait-dependent intraspecific competition, since it has to be an even function of $(x_m - x)$. Consequently, the model and the results stay the same for the case of symmetric competition as in the case of trait-independent
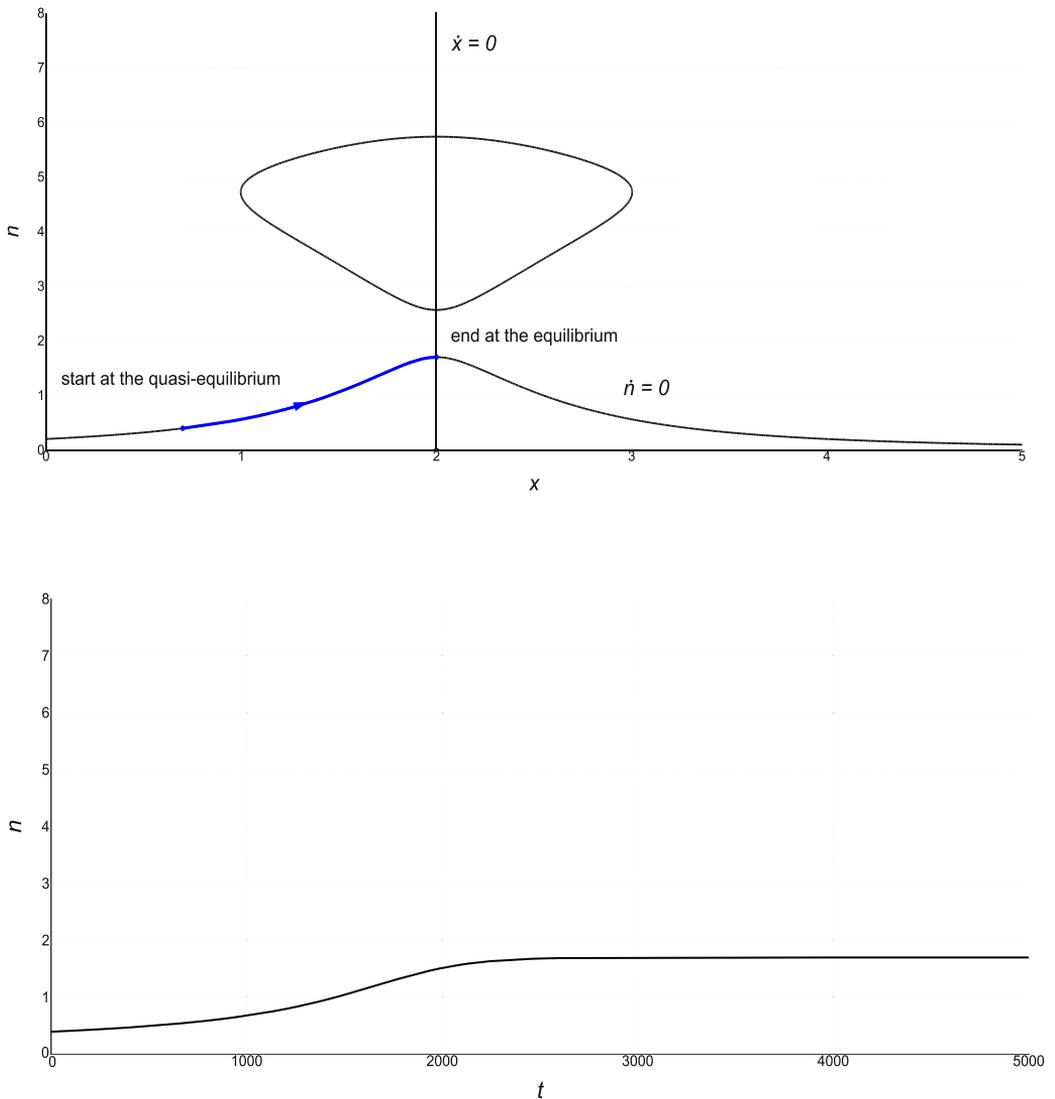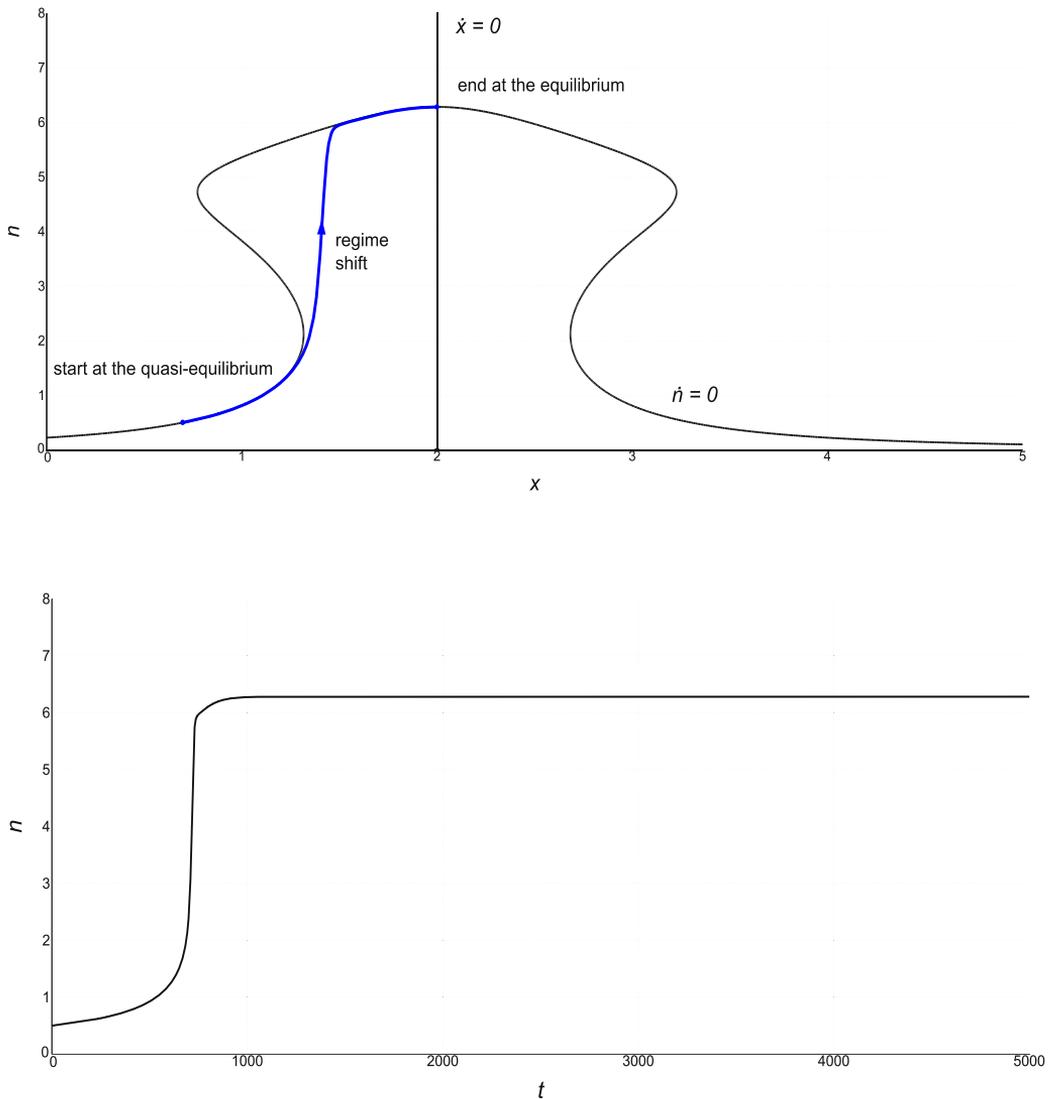
Fig. 2. Simulation of a model with multiple quasi-equilibria and trait-independent intraspecific competition and symmetric $\lambda(x)$ at level $\lambda_0 = 1$.

intraspecific competition. This is natural, since the trait value change for symmetric intraspecific competition gives no advantage over the resident population to the mutant population except of the advantage mediated by the predator population. But in case of an asymmetric competition, the trait value change gives advantage or disadvantage to the mutant population (for example in the carrying capacity or the birth rate change with regard to the resident population) and in that case $\left.\frac{\partial \alpha(x_m,x)}{\partial x_m}\right|_{x_m=x} \neq 0$ generally. In this case the eco-evolutionary dynamics can change in a more complex way, since the evolutionary nullcline need not stay vertical.

5.1. Hysteresis caused by the trait change

Let us consider the slow-fast eco-evolutionary system (23) and (24) with $\lambda = \frac{\lambda_0}{x+1}$ and with an asymmetric intraspecific competition, where $\left.\frac{\partial \alpha(x_m,x)}{\partial x_m}\right|_{x_m=x} = \alpha_1 > 0$. This specific form of λ was used since it is a monotonic function for $\lambda_0 \neq 0$ and $x > 0$ and it can represent dependence of the encounter rate on such a simple quantified trait value as the strength or the speed of the prey. The simulation in Fig. 4 shows that such a system exhibits a hysteresis cycle, continual dynamical transitions from one branch of the S-shaped quasi-equilibria manifold to another. Parameters used for the simulation were $r = 0.375$, $K = 3.3$, $N = 1$, $\lambda_0 = 10$, $h = 3$, $\alpha_1 = 0.3$ and $\varepsilon = 0.01$. But which specific parameter values are used is not so important, since the same phenomenon

Fig. 3. Simulation of a model with multiple quasi-equilibria and trait-independent intraspecific competition and symmetric $\lambda(x)$ at level $\lambda_0 = 0.5$.

happens any time the unstable branch of the eco-quasi-equilibria nullcline crosses the evolutionary nullcline. The stable eco-evolutionary limit cycle existence can be proved by methods of bifurcation theory.

6. Discussion

The impact of the evolutionary process on the ecosystem in case of population models that exhibit multiple equilibria depends on the influence of the trait to intraspecific competition. The evolutionary process may induce stability, regime shifts, hysteresis, cycles or more complex dynamics caused by the trait change. More complex eco-evolutionary dynamics are impossible in the case of trait-independent or symmetric intraspecific competition, where the eco-system stays qualitatively unaffected and only near fold points of equilibria manifold

irreversible regime shifts can happen (abrupt changes in population density or even extinction is possible, see e.g. Dercole, 2005). In the case of asymmetric intraspecific competition more complex dynamics can occur, for example a sequence of regime shifts. The asymmetry in intraspecific competition as a consequence of a trait mutation induced by evolutionary advantage of the mutant relative to predator is the driving force of complexity. Intraspecific competition as well as interspecific competition create new qualitative features of the ecosystem. The more adaptive traits are considered, the more complex dynamics can occur. These results are in agreement with recent research presented in Dercole et al. (2010) and Dercole and Rinaldi (2010), where authors introduced first example of evolutionary chaos. Their tritrophic food chain system with three adaptive traits and asymmetric intraspecific competition evince a period-doubling bifurcation cascade
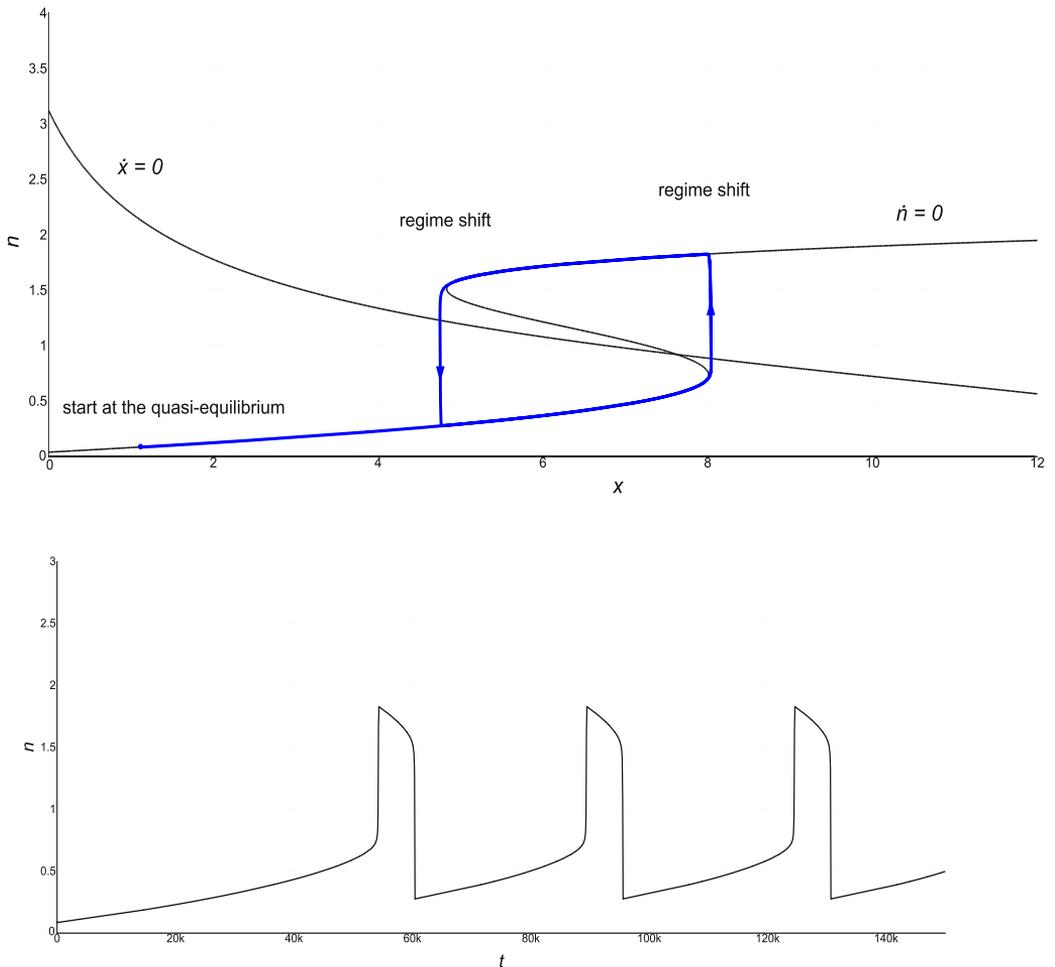
Fig. 4. Simulation of a model with multiple quasi-equilibria and asymmetric intraspecific competition.

with respect to a changing parameter (mutational rate), the Hopf bifurcation brings in a stable cycle that goes through a cascade of flip bifurcations (of a cycle) which is a typical route to chaotic dynamics.

Methods used in this study are based on the assumption that the ecosystem is settled at an equilibrium attractor, but models of ecosystems show also emergence of limit cycles and various other attractors (also chaotic). The slow evolutionary equation (3) is derived with the assumption that the resident population settles to a point attractor. Out of equilibria eco-system state variables should be replaced by their temporal mean value at the resident attractor. For detailed explanation see Dieckmann and Law (1996). This is still an open problem and usually nothing more than numerical simulations can be done except some special cases where explicit calculation of the attractor is possible as for slow-fast prey–predator dynamics presented by Dercole et al. (2006). On this place I have to mention that AD canonical equations in the chaos evincing model from Dercole and Rinaldi (2010) are considered without this temporal mean value of resident population (since ecological and evolutionary timescales are completely separated, so evolution is infinitely slower than ecology). On the other hand, there are examples of ecosystems with evolutionary dynamics on the same

timescale, see e.g. Thompson (1998). It would be also interesting to find out whether emergence of chaotic behaviour is sensitive to the time-scale variance.

7. Conclusion

Adaptive dynamics was surely influenced by the brave, fascinating and impressive ideas of Prigogine's self-organization through non-equilibrium dynamics in Nicolis and Prigogine (1977) and by Kauffman's origins of order with autonomous agents interlocked in a highly complex biosystem (see Kauffman, 1993). While modelling the ecosystem we usually assume given relationships between species (with changeable parameters of course), since in the short time scale it is possible. It is - usually. But that relationships may change much more due to adaptivity. The resident-mutant relationship can bring in the intraspecific competition as a new feature that has to be included into the model. And the model can get more complex, the dynamics is changing with an increase in the number of variables, parameters and relations. The eco-evolutionary dynamics are non-linear with all their features, bifurcation manifolds, basins of attraction and attractors, so it

can produce order, but also cycles, bifurcations and chaotic behaviour as well (see recent research Dercole et al., 2010; Dercole and Rinaldi, 2010; Wilsenach et al., 2017). There is that important word - can. Adaptive dynamics for specialized adaptive traits can be considered as the source of order or stability (but also the source of possible danger of abrupt irreversible change in population density, even a species extinction). More complex systems with less specialized adaptive traits that influence more entities of the system produce more complex dynamics and the selection and mutation process brings no order, only the slow-fast dynamics makes it less visible. As the complexity of entities increases, the power of selection is limited and complexity catastrophes occur, since selection cannot climb to the fitness landscape peaks (cited from Kauffman, 1993). This is exactly what bifurcation and catastrophe theories claim. The slow-fast system tends to quasi-attractors of the fast system, but as the whole it can be very complex and unpredictable as well, the order may be seeming, just caused by the slow-fast feature primarily. Indeed, laboratory experiments confirm that prey evolution can substantially alter predator–prey dynamics on similar timescales, see Yoshida et al. (2003). Focusing on multiple timescales should be very useful in future research. These contemplations correspond with the ideas about the importance of transient dynamics indicated by many nowadays biologists, see for example Hastings (2004). What is clear is that the ecosystems cannot be studied separately from evolution, but on the other hand, to understand the quasi-attractors and transient dynamics, ecosystems have to be studied separately also. It appears that our models are not good enough in a deeper way. An increase in number of the variables, relations and parameters used in the model violates the types of possible dynamics and the evolving ecosystems are getting more complex as well as their inner relations. The relevance of ecosystem models and modelling process itself crucially depends on information whether and how the evolution dynamics influence the ecosystem dynamics.

Acknowledgement

References

Abrams, P.A., Walters, C.J., 1996. Invulnerable prey and the paradox of enrichment. Ecology 77 (4), 1125–1133.

Boit, A., Martinez, N.D., Williams, R.J., Gaedke, U., 2012. Mechanistic theory and modelling of complex food-web dynamics in lake constance. Ecol. Lett. 15 (6), 594–602.

Boukal, D.S., Sabelis, M.W., Berec, L., 2007. How predator functional responses and allee effects in prey affect the paradox of enrichment and population collapses. Theor. Popul. Biol. 72 (1), 136–147.

Dawes, J., Souza, M., 2013. A derivation of holling's type I, II and III functional responses in predator–prey systems. J. Theor. Biol. 327, 11–22.

Della Rossa, F., Dercole, F., Landi, P., 2015. The branching bifurcation of adaptive dynamics. Int. J. Bifurcation Chaos 25 (07), 1540001.

Dercole, F., 2003. Remarks on branching-extinction evolutionary cycles. J. Math. Biol. 47 (6), 569–580.

Dercole, F., 2005. Border collision bifurcations in the evolution of mutualistic interactions. Int. J. Bifurcation Chaos 15 (07), 2179–2190.

Dercole, F., 2016. The ecology of asexual pairwise interactions: the generalized law of mass action. Theor. Ecol. 9 (3), 299–321.

Dercole, F., Della Rossa, F., Landi, P., 2016. The transition from evolutionary stability to branching: a catastrophic evolutionary shift. Sci. Rep. 6, 26310.

Dercole, F., Dieckmann, U., Obersteiner, M., Rinaldi, S., 2008. Adaptive dynamics and technological change. Technovation 28 (6), 335–348.

Dercole, F., Ferriere, R., Gragnani, A., Rinaldi, S., 2006. Coevolution of slow–fast populations: evolutionary sliding, evolutionary pseudo-equilibria and complex red queen dynamics. Proc. R. Soc. Lond. B Biol. Sci. 273 (1589), 983–990.

Dercole, F., Ferrière, R., Rinaldi, S., 2002. Ecological bistability and evolutionary reversals under asymmetrical competition. Evolution 56 (6), 1081–1090.

Dercole, F., Ferriere, R., Rinaldi, S., 2010. Chaotic red queen coevolution in three-species food chains. Proc. R. Soc. Lond. B Biol. Sci. rspb20100209.

Dercole, F., Geritz, S.A., 2016. Unfolding the resident–invader dynamics of similar strategies. J. Theor. Biol. 394, 231–254.

Dercole, F., Irisson, J.-O., Rinaldi, S., 2003. Bifurcation analysis of a prey-predator coevolution model. SIAM J. Appl. Math. 63 (4), 1378–1391.

Dercole, F., Rinaldi, S., 2008. Analysis of Evolutionary Processes: The Adaptive Dynamics Approach and its Applications. Princeton University Press.

Dercole, F., Rinaldi, S., 2010. Evolutionary dynamics can be chaotic: a first example. Int. J. Bifurcation Chaos 20 (11), 3473–3485.

Dieckmann, U., Law, R., 1996. The dynamical theory of coevolution: a derivation from stochastic ecological processes. J. Math. Biol. 34 (5–6), 579–612.

Diekmann, O., et al., 2004. A Beginner's Guide to Adaptive Dynamics. Banach Center Publications 63, 47–86.

Dobson, I., 1993. Computing a closest bifurcation instability in multidimensional parameter space. J.Nonlinear Sci. 3 (1), 307–327.

Gallien, L., Landi, P., Hui, C., Richardson, D.M., 2018. Emergence of weak-intransitive competition through adaptive diversification and eco-evolutionary feedbacks. J. Ecol. 106 (3), 877–889.

Gatto, M., 1990. A general minimum principle for competing populations: some ecological and evolutionary consequences. Theor. Popul. Biol. 37 (3), 369–388.

Geritz, S.A., Metz, J.A., Kisdi, É., Meszéna, G., 1997. Dynamics of adaptation and evolutionary branching. Phys. Rev. Lett. 78 (10), 2024.

Geritz, S.A., Metz, J.A., Rueffler, C., 2016. Mutual invadability near evolutionarily singular strategies for multivariate traits, with special reference to the strongly convergence stable case. J. Math. Biol. 72 (4), 1081–1099.

González-Olivares, E., Rojas-Palma, A., 2011. Multiple limit cycles in a gause type predator–prey model with holling type iii functional response and allee effect on prey. Bull. Math. Biol. 73 (6), 1378–1397.

Grimm, V., Dorndorf, N., Frey-Roos, F., Wissel, C., Wyszomirski, T., Arnold, W., 2003. Modelling the role of social behavior in the persistence of the alpine marmot marmota marmota. Oikos 102 (1), 124–136.

Hastings, A., 2004. Transients: the key to long-term ecological understanding? Trends Ecol. Evol. 19 (1), 39–45.

Hastings, A., Powell, T., 1991. Chaos in a three-species food chain. Ecology 72 (3), 896–903.

Hui, C., Minoarivelo, H.O., Landi, P., 2017. Modelling coevolution in ecological networks with adaptive dynamics. Math. Methods Appl. Sci. Special issue paper.

Hui, C., Richardson, D.M., Landi, P., Minoarivelo, H.O., Garnas, J., Roy, H.E., 2016. Defining invasiveness and invasibility in ecological networks. Biol. Invasions 18 (4), 971–983.

Huisman, J., Weissing, F.J., 2002. Oscillations and chaos generated by competition for interactively essential resources. Ecol. Res. 17 (2), 175–181.

Kar, T.K., 2006. Modelling and analysis of a harvested prey–predator system incorporating a prey refuge. J. Comput. Appl. Math. 185 (1), 19–33.

Kauffman, S.A., 1993. The Origins of Order: Self Organization and Selection in Evolution. Oxford University Press, USA.

Kuznetsov, Y.A., De Feo, O., Rinaldi, S., 2001. Belyakov homoclinic bifurcations in a tritrophic food chain model. SIAM J. Appl. Math. 62 (2), 462–487.

Landi, P., Dercole, F., Rinaldi, S., 2013. Branching scenarios in eco-evolutionary prey-predator models. SIAM J. Appl. Math. 73 (4), 1634–1658.

Ludwig, D., Jones, D.D., Holling, C.S., 1978. Qualitative analysis of insect outbreak systems: the spruce budworm and forest. J. Anim. Ecol. 315–332.

MacArthur, R., 1970. Species packing and competitive equilibrium for many species. Theor. Popul. Biol. 1 (1), 1–11.

Metz, J.A., Geritz, S.A., Meszéna, G., Jacobs, F.J., Van Heerwaarden, J.S., 1995. Adaptive Dynamics: A Geometrical Study of the Consequences of Nearly Faithful Reproduction.

Metz, J.A., Nisbet, R.M., Geritz, S.A., 1992. How should we define fitness for general ecological scenarios? Trends Ecol. Evol. 7 (6), 198–202.

Mohammed, M., Landi, P., Minoarivelo, H., Hui, C., 2018. Frugivory and seed dispersal: extended bi-stable persistence and reduced clustering of plants. Ecol. Modell. 380, 31–39.

Mumby, P.J., Hastings, A., Edwards, H.J., 2007. Thresholds and the resilience of caribbean coral reefs. Nature 450 (7166), 98–101.

Muratori, S., Rinaldi, S., 1991. A separation condition for the existence of limit cycles in slow-fast systems. Appl. Math. Model. 15 (6), 312–318.

Nicolis, G., Prigogine, I., 1977. Self-organization in Nonequilibrium Systems. Wiley, New York.

Parvinen, K., 2005. Evolutionary suicide. Acta Biotheor. 53 (3), 241–264.

Rinaldi, S., Muratori, S., Kuznetsov, Y., 1993. Multiple attractors, catastrophes and chaos in seasonally perturbed predator-prey communities. Bull. Math. Biol. 55 (1), 15–35.

Rosenzweig, M.L., MacArthur, R.H., 1963. Graphical representation and stability conditions of predator-prey interactions. Am. Nat. 209–223.

Scheffer, M., Rinaldi, S., Gragnani, A., Mur, L.R., van Nes, E.H., 1997. On the dominance of filamentous cyanobacteria in shallow, turbid lakes. Ecology 78 (1), 272–282.

Steele, J.H., Henderson, E.W., 1992. The role of predation in plankton models. J. Plankton Res. 14 (1), 157–172.

Tamba, T.A., 2015. Forecasting Regime Shifts in Nonlinear Dynamical Processes. University of Notre Dame Ph.D. thesis.

Thompson, J.N., 1998. Rapid evolution as an ecological process. Trends Ecol. Evol. 13 (8), 329–332.

Wilsenach, J., Landi, P., Hui, C., 2017. Evolutionary fields can explain patterns of high-dimensional complexity in ecology. Phys. Rev. E 95 (4), 042401.

Yoshida, T., Jones, L.E., Ellner, S.P., Fussmann, G.F., Hairston, N.G., 2003. Rapid evolution drives ecological dynamics in a predator–prey system. Nature 424 (6946), 303–306.

Bifurcation manifolds in predator–prey models computed by Gröbner basis method

Veronika Hajnová[*], Lenka Přibylová

Department of Mathematics and Statistics, Section of Applied Mathematics, Faculty of Science, Masaryk University, Kotlářská 2, Brno, 611 37, Czechia

A B S T R A C T

Many natural processes studied in population biology, systems biology, biochemistry, chemistry or physics are modeled by dynamical systems with polynomial or rational right-hand sides in state and parameter variables. The problem of finding bifurcation manifolds of such discrete or continuous dynamical systems leads to a problem of finding solutions to a system of non-linear algebraic equations. This approach often fails since it is not possible to express equilibria explicitly. Here we describe an algebraic procedure based on the Gröbner basis computation that finds bifurcation manifolds without computing equilibria. Our method provides formulas for bifurcation manifolds in commonly studied cases in applied research – for the fold, transcritical, cusp, Hopf and Bogdanov–Takens bifurcations. The method returns bifurcation manifolds as implicitly defined functions or parametric functions in full parameter space. The approach can be implemented in any computer algebra system; therefore it can be used in applied research as a supporting autonomous computation even by non-experts in bifurcation theory. This paper demonstrates our new approach on the recently published Rosenzweig–MacArthur predator–prey model generalizations in order to highlight the simplicity of our method compared to the published analysis.

1. Introduction

The aim of this paper is to describe a method of finding bifurcation manifolds using the Gröbner basis computation. The approach is presented on the Rosenzweig–MacArthur model generalizations. The original model was published in 1963 in [21] as a generalization of Lotka–Voltera model. Dynamics of two interacting populations is described by the following system of differential equations:

$$
\begin{aligned}
\frac{\mathrm{d}N}{\mathrm{d}t} &= rN\left(1 - \frac{N}{K}\right) - g(N, P)P := f_1(N, P),\\
\frac{\mathrm{d}P}{\mathrm{d}t} &= eg(N, P)P - mP := f_2(N, P),
\end{aligned}
\tag{1}
$$

where the state variables N, P denote prey population and predator population density and t is time. A common notation is used - parameter r denotes the growth rate of the prey, K is carrying capacity of the prey, m is the death rate of the predator and e is the conversion efficiency of the predator. We assume that all parameters are positive. Function $g(N, P)$ is a specific functional response. Models with number of different types of functional responses $g(N, P)$ are studied in current research, see for example [3,19,20,22] or others.

All parameters of a dynamical system such as (1) can fluctuate over time. External factors typically cause these fluctuations. Bifurcation theory

provides a background to describe the consequences of those fluctuations. More precisely, bifurcation theory brings in tools to analyze qualitative changes in dynamics of a given family of differential equations. The topological structure of the state–parameter space is described by a bifurcation diagram, usually presented in parameter space as a partition of the parameter space to structurally stable domains with topologically equivalent dynamics, while their boundaries are given by the bifurcation manifolds. That is why the computation of bifurcation manifolds is crucial in applied research. Let us recall that biochemical switches and bistability are related to fold or cusp bifurcations, see e.g. [7,11,17,23]. Chemical oscillations and current research of network phenomena such as synchronization or chimera states are related to Hopf bifurcation, see e.g. [1,18,24]. Control methods of harvesting or preserving populations, sudden population extinction explanation are related to fold or Bogdanov–Takens bifurcations, see [8,9,13] or [19]. However, the analysis itself is often very complicated and usually requires numerical continuation methods and software, for example, Matcont [14] or Auto [25]. In case of a differential system with polynomial or rational right-hand sides, where the problem of finding bifurcation manifolds leads to a problem of solving a system of polynomial equations, see [12], we can avoid numerical continuation and derive the results analytically. We present our approach that allows us to analyse the system algorithmically and compute

bifurcation manifolds as implicit or parametric functions in full parameter space using known algorithms to compute the Gröbner basis of the polynomial system[1]. There are several known algorithms implemented in current mathematical software, such as the Buchberger algorithm, see [6].

System (1) has polynomial or rational right-hand sides if function g (N, P) is rational or polynomial. This approach was already used in [15,16], but to the best of our knowledge, it is not a commonly used procedure despite its apparent advantages.

To obtain the system of algebraic equations described in previous paragraph, let us consider a general dynamical system with continuous time in the form

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{x}' = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\varepsilon}). \tag{2}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $n \in \mathbb{N}$ is a vector of state variables, $\boldsymbol{\varepsilon} \in \mathbb{R}^l$, $l \in \mathbb{N}$ is a vector of parameters, and t is time. The Jacobi matrix of system (2) is denoted by $J(\boldsymbol{x}, \boldsymbol{\varepsilon})$. If (x^*, ε^*) is a fold or a transcritical bifurcation point, then (x^*, ε^*) is a solution of the system

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\varepsilon}) = \boldsymbol{0},$$
$$\det(J(\boldsymbol{x}, \boldsymbol{\varepsilon})) = 0. \tag{3}$$

A proof can be found in [12]. Similarly, if (x^*, ε^*) is a Hopf bifurcation point, then (x^*, ε^*) is a solution of the system

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\varepsilon}) = \boldsymbol{0},$$
$$\det(2J(\boldsymbol{x}, \boldsymbol{\varepsilon}) \odot I_n) = 0, \tag{4}$$

where I_n stands for an n dimensional identity matrix and $\odot$ denotes the bialternate matrix product, for its definition see [12], page 486. In special case $n = 2$ we obtain commonly known conditions

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\varepsilon}) = \boldsymbol{0},$$
$$\text{trace}(J(\boldsymbol{x}, \boldsymbol{\varepsilon})) = 0.$$

Previous statements described necessary conditions for existence of the Hopf and the fold one-parameter bifurcations. Those conditions are not sufficient to classify the bifurcation type. Generally solutions of systems of algebraic equations (3) or (4) consist of fold or Hopf bifurcation points, multi-parameter bifurcation points or degenerated bifurcation points.

Although it might seem that the use of the Gröbner basis method is only a different approach for computation of bifurcation points, it is not the case. Opposed to the commonly used analysis of a dynamic system, this method enables the computation of the bifurcation manifolds without the need to calculate equilibria explicitly. Therefore, we can also present some new results for the Rosenzweig–MacArthur model generalizations. In our paper we applied the method to models derived and studied in [3,19,20] or [22]. For Gröbner basis computations we used Maple package Groebner, see [27].

2. Model analysis and results

2.1. The Rosenzweig–MacArthur model with predator interference

In this section we focus on the Rosenzweig–MacArthur model in form (1). The functional response $g(N, P)$ is considered in following form:

$$g(N, P) = \frac{\Lambda(P)N}{\Lambda(P)hN + 1},$$
$$\Lambda(P) = \frac{\Lambda_0}{(b + P)^2} \tag{5}$$

---

[1] The Gröbner basis elimination method can be likened to the Gauss elimination method - it is, in fact, its generalization. The original set of polynomials is transformed into another set with the same properties, in our case the Gröbner basis generates the vanishing ideal of the original polynomial set. So similarly to Gauss elimination, the original problem is transformed into an equivalent problem, but the algorithm provides a variable elimination.

where parameter $h > 0$ is handling time and function $\Lambda(P)$ is an attack rate with $\Lambda_0 > 0$, $b > 0$. Functional response $g(N, P)$ is a generalization of Holling type II functional response, see [10]. Function $\Lambda(P)$ is a decreasing function, therefore function $g(N, P)$ is decreasing in P for any fixed N, and it represents interference among predators. The Rosenzweig–MacArthur model with this particular functional response was studied in [3,19].

Note that it is ineffective to try to find an equilibrium of system (1) with functional response (5) analytically. Using the standard methods for finding equilibria of system (1) with functional response (5) we end up with a polynomial equation, where the leading term is generally fourth-degree.

Assuming a lexicographic order $N > P > r > K > h > e > m > b > \Lambda_0$, it is possible to find Gröbner basis for the set of polynomials

$$S_1 = \left\{ \text{numerator}(f_1(N, P)), \text{numerator}(f_2(N, P)), \text{numerator}\left(\det\begin{pmatrix} \frac{df_1(P,N)}{dN} & \frac{df_1(P,N)}{dP} \\ \frac{df_2(P,N)}{dN} & \frac{df_2(P,N)}{dP} \end{pmatrix}\right) \right\}$$

$$\tag{6}$$

and by factorization of the first polynomial in the basis we obtain polynomials

$$\begin{aligned}
p_1(N, P, r, K, h, e, m, b, \Lambda_0) =\ & 16\,r^3K^2e^3b + 4\,r^2K^2he^2m\Lambda_0 - 4\,r^2K^2e^3\Lambda_0 \\
& + 128\,r^2Ke^2mb^2 + 144\,rKhem^2b\Lambda_0 - 144\,rKe^2mb\Lambda_0 \\
& + 256\,rem^2b^3 + 27\,Kh^2m^3\Lambda_0^2 - 54\,Khem^2\Lambda_0^2 \\
& + 27\,Ke^2m\Lambda_0^2, \\
p_2(N, P, r, K, h, e, m, b, \Lambda_0) =\ & -Khm\Lambda_0 + Ke\Lambda_0 - mb^2, \\
p_3(N, P, r, K, h, e, m, b, \Lambda_0) =\ & -hm + e.
\end{aligned}$$

To eliminate state variables N, P, it is crucial to place them first in assumed lexicographic order. The overall computation time of computations depends also on ordering within the set of state variables N, P and the set of parameters $r, K, h, e, m, b, \Lambda_0$, see [4].

The polynomials vanish at fold or transcritical bifurcation points and so we get implicit description of the corresponding bifurcation manifolds. The case where $e = hm$ is not biologically relevant. The transcritical bifurcation in the intersection of coexistence equilibrium branch and prey equilibrium $[K, 0]$ without predator branch, see [19], is described by $p_2 = 0$. On the other hand $p_1 = 0$ implicitly defines the fold manifold that is derived in full parameter space, which is a new result not mentioned in [19]. Similarly assuming the same lexicographic order $N > P > r > K > h > e > m > b > \Lambda_0$, it is possible to find the Gröbner basis for the set of polynomial

$$S_2 = \left\{ \text{numerator}(f_1(N, P)), \text{numerator}(f_2(N, P)), \text{numerator}\left(\text{trace}\begin{pmatrix} \frac{df_1(P,N)}{dN} & \frac{df_1(P,N)}{dP} \\ \frac{df_2(P,N)}{dN} & \frac{df_2(P,N)}{dP} \end{pmatrix}\right) \right\}$$

$$\tag{7}$$

to obtain Hopf bifurcation manifold (new result). Factorization of the first polynomial in the basis implies that polynomial

$$\begin{aligned}
q(N, P, r, K, h, e, m, b, \Lambda_0) =\ & r^3K^2h^2e^4l^2 - 4\,r^2K^2h^3e^3m^2l^2 + 6\,r^2K^2h^2e^4ml^2 \\
& - 4\,r^2K^2he^5mbl - 2\,r^2K^2he^5\Lambda_0^2 + 4\,r^2K^2e^6b\Lambda_0 \\
& + 2\,r^2Kh^3e^2m^2b\Lambda_0^2 + 4\,r^2Kh^2e^3mb\Lambda_0 + 2\,r^2Khe^4b\Lambda_0^2 \\
& + 4\,rK^2h^4e^2m^4\Lambda_0^2 - 16\,rK^2h^3e^3m^3\Lambda_0^2 + 16\,rK^2h^2e^4m^3b\Lambda_0 \\
& + 20\,rK^2h^2e^4m^2\Lambda_0^2 - 32\,rK^2he^5m^2b\Lambda_0 - 8\,rK^2he^5m\Lambda_0^2 \\
& + 16\,rK^2e^6mb\Lambda_0 + rKh^5m^4\Lambda_0^3 + 2\,rKh^4em^3\Lambda_0^3 \\
& + 4\,rKh^3e^2m^4b^2\Lambda_0 - 2\,rKh^3e^3m^3b\Lambda_0^2 - 12\,rKh^2e^3m^3b^2\Lambda_0 \\
& - 2\,rKh^2e^3m^2b\Lambda_0^2 - 2\,rKh^2e^3m\Lambda_0^3 - 4\,rKhe^4m^2b^2\Lambda_0 \\
& + 2\,rKhe^4mb\Lambda_0^2 - rKhe^4\Lambda_0^3 + 12\,rKe^5mb^2\Lambda_0 \\
& \quad + 2\,rKe^5b\Lambda_0^2 \\
& + rh^4m^4b^2\Lambda_0^2 + 4\,rh^3em^3b^2\Lambda_0^2 + 6\,rh^2e^2m^2b^2\Lambda_0^2 \\
& + 4\,rhe^3mb^2\Lambda_0^2 + re^4b^2\Lambda_0^2 + 8\,K^2h^4e^2m^5\Lambda_0 \\
& - 16\,K^2h^3e^3m^3b\Lambda_0 - 24\,K^2h^3e^3m^4\Lambda_0^2 + 48\,K^2h^2e^4m^4b\Lambda_0 \\
& + 24\,K^2h^2e^4m^3\Lambda_0^2 - 48\,K^2he^5m^3b\Lambda_0 - 8\,K^2he^5m^2\Lambda_0^2 \\
& + 16\,K^2e^6m^2b\Lambda_0 + 8\,Kh^3e^2m^5b^2\Lambda_0 - 16\,Kh^2e^3m^5b^3 \\
& - 8\,Kh^2e^3m^4b^2\Lambda_0 + 32\,Khe^4m^4b^3 - 8\,Khe^4m^3b^2\Lambda_0 \\
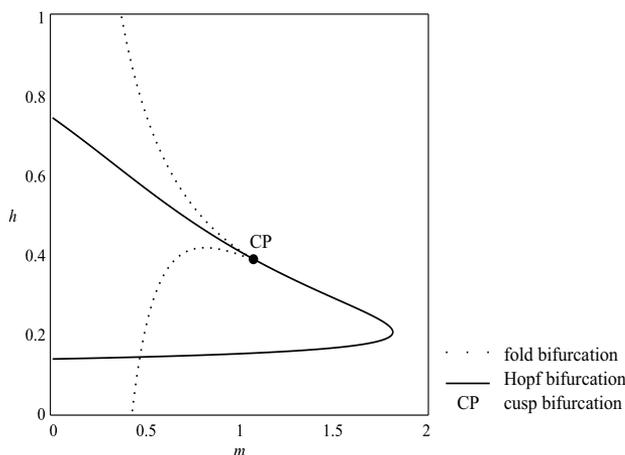& - 16\,Ke^5m^3b^3 + 8\,Ke^5m^2b^2\Lambda_0
\end{aligned}$$

Fig. 1. The fold bifurcation manifold, implicitly defined by equation $p_1 = 0$, and the Hopf bifurcation manifold, implicitly defined by equation $q = 0$, in the generalized Rosenzweig–MacArthur model depending on m (a death rate of a predator) and h (handling time) with $K = 5$, $\Lambda_0 = 28$, $r = 3.5$, $e = 1$, $b = 1.3$. Note that the solid bifurcation manifold consists of both neutral saddle points and Hopf bifurcation points.

vanishes at Hopf bifurcation points. Figs. 1 and 2 depict the fold bifurcation manifold implicitly defined by $p_1 = 0$ and the Hopf bifurcation manifold implicitly defined by $q = 0$. It is crucial to point out that the vanishing ideal of (7) does not account for the scenario, where the Jacobi matrix determinant is positive. Therefore the expressed manifold consists of Hopf bifurcation points as well as neutral saddle points. Additional analysis of the manifold subset that satisfies $\det J > 0$ is necessary to avoid spurious results. This could be done numerically, analytically or using Gröbner basis with an additional parameter $\delta = \det J$ and an additional polynomial $\delta - \det J$. The Hopf bifurcation manifold in parametric space border is the contour $\delta = 0$, so the border is usually the Bogdanov–Takens bifurcation manifold (see the next example below).

A cusp bifurcation, denoted by CP in Fig. 1, is a two-parameter bifurcation of equilibria. For the normal form of this bifurcation see [12]. A cusp bifurcation manifold is a co-dimension two manifold, submanifold of the fold bifurcation manifold. At cusp points, two branches of fold bifurcation manifolds of co-dimension 1 meet tangentially in a typical V-shape, see Fig. 1. Let us assume that we are interested in parameter m and h analysis and we seek hysteresis or transient

dynamics. Because generally, as well as in this example, cusp bifurcation points can be found as points where the implicitly defined function has no derivative, according to the implicit function theorem, cusp points are solutions of a set of polynomial equations $\frac{\partial p_1}{\partial h} = 0$ and $\frac{\partial p_1}{\partial m} = 0$. Eliminating h and m respectively using the Gröbner basis method again, we get the cusp bifurcation points explicitly as

$$m = \frac{1}{12}\frac{Ker}{b},$$

$$h = \frac{4}{3}\frac{(9\Lambda_0 - 32br)b}{Kr\Lambda_0}.$$

The bifurcation analysis provided in this paper is not complete, for more see [19]. Dynamical model in form (1) with functional response (8) is in fact a special case of model studied in [19]. In [19] function $\Lambda(P)$ is considered in form $\Lambda(P) = \frac{\Lambda_0}{(b + P)^w}$. In our paper parameter w is set to critical value 2. The Hopf curve passes directly through the cusp bifurcation point (CP), which is a degenerate situation, see Fig. 1.

Such computations, made for example in Maple, may serve as quick support to more in-depth analysis or simulations. Following code shows all necessary computations in Maple.

```
>  lambda := (P) -> 1/(b+P)^2;

>  g := (N, P) -> lambda(P)*N/(1+lambda(P)*h*N);

>  f1 := (N, P)-> r*N*(1-N/K)-g(N, P)*P;

>  f2 := (N, P) -> e*g(N, P)*P-m*P;

>  Jacobi := VectorCalculus[Jacobian]([f1(N, P), f2(N, P)], [N, P]);

>  with(LinearAlgebra):

>  p1 := numer(f1(N, P))/N;

>  p2 := numer(f2(N, P))/P;

>  p3 := numer(Determinant(Jacobi));

>  p4 := numer(Trace(Jacobi));

>  gb1 := Groebner[Basis]([p1, p2, p3], plex(N, P, r, K, h, e, m, b, l));

>  factor(gb1[1]);

>  gb2 := Groebner[Basis]([p1, p2, p4], plex(N, P, r, K, h, e, m, b, l));
```

V. Hajnová and L. Přibylová

2.2. The Rosenzweig–MacArthur model with a functional response increasing in both predator and prey densities

This section is again devoted to the Rosenzweig–MacArthur model in form (1). The functional response $g(N, P)$ is considered in the following form:

$$g(N, P) = \frac{\Lambda(P)N}{\Lambda(P)hN + 1},$$
$$\Lambda(P) = Ce_0 P, \tag{8}$$

where parameter $h > 0$ is the handling time and function $\Lambda(P)$ is an attack rate with $\Lambda_0 > 0$. $C > 0$ is the amount of prey captured by a predator per encounter and e_0 is the total encounter coefficient between the predator and the prey. Function $g(N, P)$ is increasing in P for any fixed N, and it represents cooperation of predators in foraging. The Rosenzweig–MacArthur model with this particular functional response was studied in [22] or in a general form that included the special case of foraging facilitation among predators in [20].

In the remainder of the section we use the same scaling and notation as [22] to simplify connections between results in [22] and those obtained in our paper. The rescaling is

$$rt = T, \quad \frac{N}{K} = x, \quad hCe_0\,KP = y, \quad \frac{1}{Ce_0(hK)^2 r} = \alpha, \quad \frac{e}{rh} = \beta, \quad \frac{m}{r} = \gamma$$

and the system simplifies to

$$\frac{dx}{dT} = x(1-x) - \frac{\alpha y^2 x}{yx + 1} := f_3,$$
$$\frac{dy}{dT} = \frac{\beta y^2 x}{yx + 1} - \gamma y := f_4. \tag{9}$$

Using the same technique as in the previous section, assuming a lexicographic order $x \succ y \succ \gamma \succ \beta \succ \alpha$ we obtain fold and Hopf bifurcation manifolds:

fold: $\quad 27\,\alpha\,\gamma^2 - 4\beta^2 + 4\beta\,\gamma = 0 \tag{10}$

$$S_2 = \left\{ \text{numerator}(f_3(x, y)),\ \text{numerator}(f_4(x, y)),\ \text{numerator}\left(\det\begin{pmatrix} \dfrac{df_3(x, y)}{dN} & \dfrac{df_3(x, y)}{dP} \\[2ex] \dfrac{df_4(x, y)}{dN} & \dfrac{df_4(x, y)}{dP} \end{pmatrix}\right), \right.$$
$$\left. \text{numerator}\left(\text{trace}\begin{pmatrix} \dfrac{df_3(x, y)}{dN} & \dfrac{df_3(x, y)}{dP} \\[2ex] \dfrac{df_4(x, y)}{dN} & \dfrac{df_4(x, y)}{dP} \end{pmatrix}\right) \right\}, \tag{12}$$

to obtain polynomials that define to Bogdanov–Takens bifurcation manifold. There are four polynomials in the Gröbner basis that have state variables x, y eliminated:

$$p_1 = 81\,\alpha\,\beta^2\gamma - 243\,\alpha^2\gamma - 108\,\alpha\,\beta\,\gamma - 9\,\alpha\gamma - 4\beta\gamma$$
$$p_2 = 81\,\alpha^2\gamma^2 - 54\,\alpha^2\gamma - 6\,\alpha\,\beta\,\gamma + 6\,\alpha\gamma^2 - 2\,\alpha\gamma$$
$$p_3 = 27\,\alpha\gamma^2 - 6\,\beta^2\gamma + 6\,\beta\,\gamma^2 + 2\beta\gamma$$
$$p_4 = 81\,\alpha\,\gamma^3 - 54\,\alpha\gamma^2 - 4\,\beta\gamma.$$

Assuming all parameters are non-zero we obtain following simplification:

$$\alpha = \frac{4}{27}\frac{3\gamma - 1}{(-2 + 3\gamma)^2}$$
$$\beta = \frac{(3\gamma - 1)\gamma}{-2 + 3\gamma}. \tag{13}$$

Fig. 3 shows the fold bifurcation manifold and the Hopf bifurcation manifold, described implicitly by (10), (11) and the Bogdanov–Takens bifurcation manifold in their intersection, described explicitly by (13). These results correspond to those published in [22]. Notice that there is no need to find equilibria of a cubic equation, and further results are received. For example, critical $\alpha_{\beta\gamma}$ is derived in [22] as a function of β and γ. Not only is our method straightforward, it also provides full-parameter expression of the Bogdanov–Takens manifold where α and β are parameterized by γ variable, so it explains the statements in [22] very clearly and simply.

Following code shows all necessary computations in Maple. Note that variable γ is denoted by c because symbol γ is protected in Maple.

```
>   f3 := (x, y) -> x*(1-x)-alpha*x*y^2/(1+x*y);

>   f4 := (x, y) -> beta*x*y^2/(1+x*y)-c*y;

>   Jacobi := VectorCalculus[Jacobian]([f3(x, y), f4(x, y)], [x, y]);

>   p1 := numer(f3(x, y))/x;

>   p2 := numer(f4(x, y))/y;

>   p3 := numer(LinearAlgebra[Determinant](Jacobi));

>   p4 := numer(LinearAlgebra[Trace](Jacobi));

>   gb1 := Groebner[Basis]([p1, p2, p3], plex(x, y, c, beta, alpha));

>   gb2 := Groebner[Basis]([p1, p2, p4], plex(x, y, c, beta, alpha));

>   gb3 := Groebner[Basis]([p1,p2,p3,p4], plex(x, y, c, beta, alpha));
```

Hopf: $\quad \beta^5\gamma - 4\beta^4\gamma^2 + 6\beta^3\gamma^3 - 4\beta^2\gamma^4 + \beta\,\gamma^5 - \beta^5 + 5\beta^4\gamma - 9\beta^3\gamma^2$
$$\qquad + 7\beta^2\gamma^3 - 2\beta\,\gamma^4 + \beta^3\alpha + 3\alpha\,\beta^2\gamma +$$
$$\qquad 3\,\alpha\,\beta\,\gamma^2 + \alpha\,\gamma^3 - 2\beta^4 + 5\gamma\beta^3 - 4\gamma^2\beta^2 + \gamma^3\beta - \beta^3 + \beta^2\gamma = 0 \tag{11}$$

Using the same lexicographic order $x \succ y \succ \gamma \succ \beta \succ \alpha$, it is possible to find Gröbner basis for set of polynomials

3. Conclusion

Our paper focuses on the use of Gröbner basis to find bifurcation manifolds in the generalized Rosenzweig–MacArthur model. This approach allows us to find implicit formulae for bifurcation manifolds in whole parameter space for systems of differential equations without computing equilibria analytically. An analogous procedure can be used for difference equations.

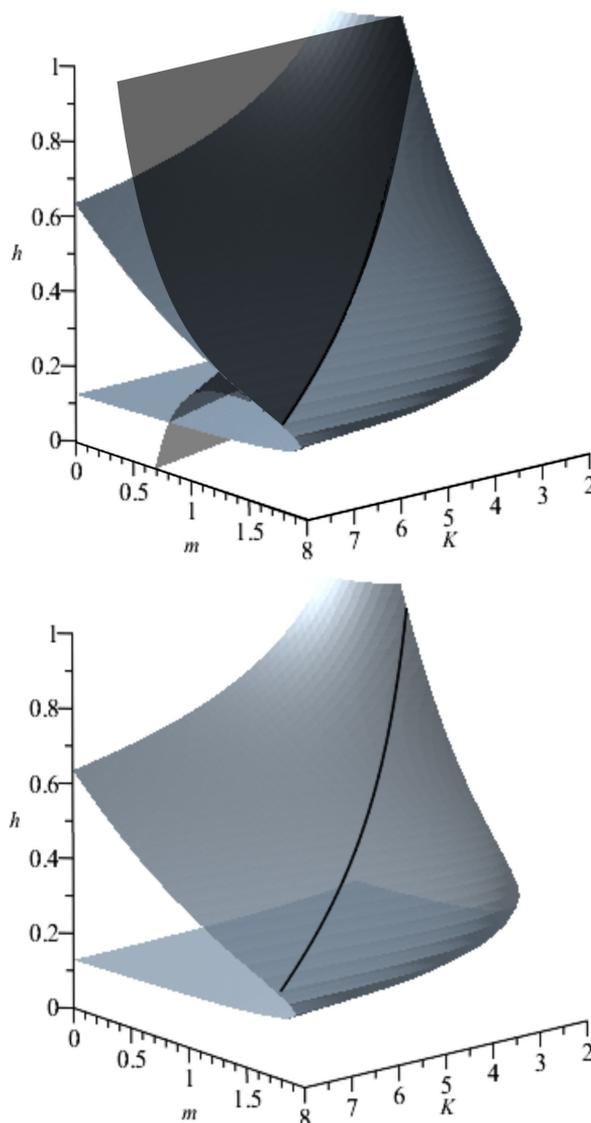The problem of finding bifurcation manifolds of a system of

Fig. 2. The top figure: the fold bifurcation manifold (black), implicitly defined by $p_1 = 0$, and the Hopf bifurcation manifold (gray), in the generalized Rosenzweig–MacArthur model depending on m (a death rate of a predator), h (handling time) and K (capacity of environment), with $\Lambda_0 = 28$, $r = 3.5$, $e = 1$, $b = 1.3$. The bottom figure: the cusp bifurcation manifold (black) and the Hopf bifurcation manifold (gray) for the same setting. Note that the depicted bifurcation manifold consists of both neutral saddle points and Hopf bifurcation points.

differential (or difference) equations generally leads to a system of non-linear equations. Gröbner basis method can be used for a system of algebraic equations, and this requirement is fulfilled for systems of differential or difference equations with polynomial or rational right-hand sides.

Several issues related to our proposed method need to be considered. Firstly, it is not possible to easily distinguish computed bifurcation manifold parts that belong to different equilibria or different domains of topological equivalence, since we get manifold expression in parameter space without the direct equilibria correspondence.

Secondly, it is not possible to assume additional properties of state variables such as positivity, so these questions have to be studied afterward. Lastly, in many practically relevant cases it is not possible to obtain Gröbner basis in reasonable time due to the computational complexity of the algorithm. Despite the mentioned disadvantages, the method has a considerable advantage because it can be applied to numerous cases and bifurcation manifolds can be computed automatically by a computer, not only by experts in non-linear dynamics. The main advantage from the expert point of view is that full-parameter space representation of bifurcation manifolds is provided. That can provide
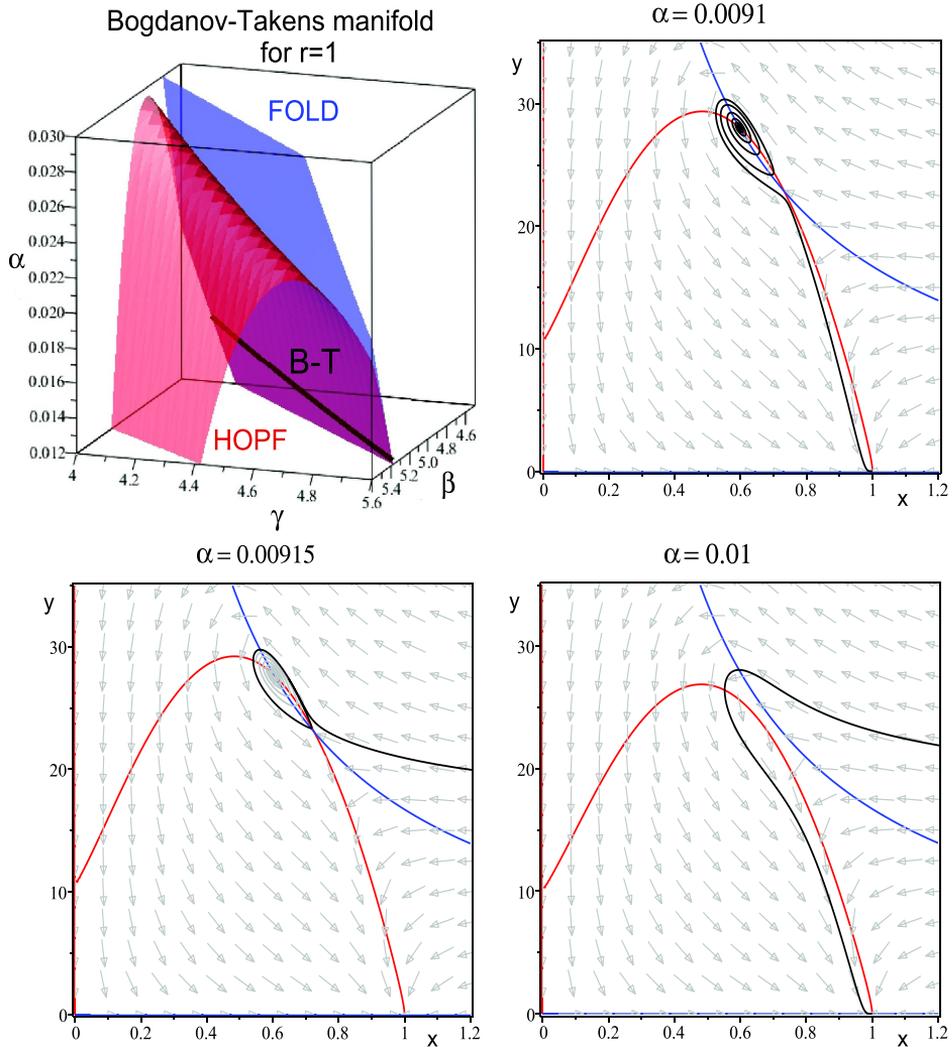
Fig. 3. Bogdanov–Takens, Hopf, fold bifurcation manifolds and phase portraits for $\gamma = 4.6$ and $\beta = 4.875$ and various α. In figures of phase portraits the red line is x-nullcline, the blue line is y-nullclines, black a gray lines are arbitrary trajectories, gray arrows represent the direction field. By increasing value of the parameter α, the invariant sets in displayed area changes from saddles and stable foci to unstable foci with stable limit cycles through the homoclinic loop to its split and presence of no invariant set at all. Note that in the figure displaying bifurcation manifolds (top left), the red bifurcation manifold consists of Hopf bifurcation points as well as neutral saddles (for values of the parameter α smaller than critical Bogdanov–Takens points).

new results as we presented here for the generalized Rosenzweig–MacArthur models.

As for future research, we believe it could be interesting to investigate how the proposed techniques can be combined with formal methods, [2], or with results published in [5].

In our paper, we used software Maple for Gröbner basis computation. There are other software options; for example, it is implemented in Matlab or Mathematica, see [26,28].

Acknowledgments

References

[1] D.M. Abrams, et al., Solvable model for chimera states of coupled oscillators, Phys. Rev. Lett. 101 (8) (2008) 084103.

[2] N. Beneš, L. Brim, M. Demka, S. Pastava, D. Šafránek, A model checking approach to discrete bifurcation analysis, FM 2016: Formal Methods, Springer International Publishing, 2016, pp. 55–63.

[3] L. Berec, Impacts of foraging facilitation among predators on predator-prey dynamics, Bull. Math. Biol. 72 (1) (2010) 94–121, https://doi.org/10.1007/s11538-009-9439-1.

[4] W. Boege, R. Gebauer, H. Kredel, Some examples for solving systems of algebraic equations by calculating Groebner bases, J. Symb. Comput. 2 (1) (1986) 83–98, https://doi.org/10.1016/S0747-7171(86)80014-1.

[5] C. Chen, W. Wu, Revealing bistability in neurological disorder models by solving parametric polynomial systems geometrically, Artificial Intelligence and Symbolic Computation, Springer International Publishing, 2018, pp. 170–180.

[6] D.A. Cox, J.B. Little, D. O'Shea, Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, Third ed., Springer, New York, 2007.

[7] O. Díaz-Hernández, M. Santillán, Bistable behavior of the lac operon in e. coli when induced with a mixture of lactose and TMG, Front. Physiol. 1 (2010) 22. 1–12

[8] E. González-Olivares, et al., Dynamical complexities in the Leslie-Gower predator-prey model as consequences of the allee effect on prey, Appl. Math. Model. 35 (1) (2011) 366–381.

[9] F.M. Hilker, M. Langlais, H. Malchow, The allee effect and infectious diseases: extinction, multistability, and the (dis-) appearance of oscillations, Am. Nat. 173 (1) (2008) 72–88.

[10] C.S. Holling, The components of predation as revealed by a study of small-mammal predation of the european pine sawfly, Can. Entomol. 91 (05) (1959) 293–320, https://doi.org/10.4039/Ent91293-5.

[11] D. Jost, Bifurcation in epigenetics: implications in development, proliferation, and diseases, Phys. Rev. E 89 (1) (2014) 010701.

[12] Y.A. Kuznetsov, Elements of Applied Bifurcation Theory, Second ed., Springer, New York, 1998.

[13] J. Lyu, R. Parshad, et al., The effects of invasive epibionts on Crab-Mussel communities: a theoretical approach to understand mussel population decline, SIAM Conference on Mathematics of Planet Earth, (2018). Preprint: https://arxiv.org/abs/1810.04256

[14] A. Dhooge, W. Govaerts, Y.A. Kuznetsov, H.G.E. Meijer, B. Sautois, MATCONT and CL MATCONT: Continuation Toolboxes in Matlab, (2006). Manual, Utrecht, Belgium

[15] W. Niu, D. Wang, Algebraic analysis of bifurcation and limit cycles for biological systems, Algebraic Biology: Third International Conference, AB 2008, Castle of Hagenberg, Austria, July 31–August 2, 2008 Proceedings, no. 5147, (2008), pp. 156–171, https://doi.org/10.1007/978-3-540-85101-1_12.

[16] W. Niu, D. Wang, Algebraic approaches to stability analysis of biological systems, Math. Comput. Sci. (1) (2008) 3, https://doi.org/10.1007/s11786-007-0039-x.

[17] B. Novák, J.J. Tyson, Design principles of biochemical oscillators, Nat. Rev. Mol. Cell Biol. 9 (12) (2008) 981.

[18] M.J. Panaggio, et al., Chimera states in networks of phase oscillators: the case of two small populations, Phys. Rev. E 93 (1) (2016) 012218.

[19] L. Přibylová, L. Berec, Predator interference and stability of predator-prey dynamics, J. Math. Biol. 71 (2) (2015) 301–323, https://doi.org/10.1007/s00285-014-0820-9.

[20] L. Přibylová, A. Peniašková, Foraging facilitation among predators and its impact on the stability of predator–prey dynamics, Ecol. Complexity 29 (2017) 30–39, https://doi.org/10.1016/j.ecocom.2016.11.006.

[21] M.L. Rosenzweig, R.H. MacArthur, Graphical representation and stability conditions of predator-prey interactions, Am. Nat. 97 (895) (1963) 209–223, https://doi.org/10.1086/282272.

[22] K. Ryu, W. Ko, M. Haque, Bifurcation analysis in a predator-prey system with a functional response increasing in both predator and prey densities, Nonlinear Dyn. (2018) 1–18.

[23] J.J. Tyson, B. Novak, Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions, J. Theor. Biol. 210.2 (2001) 249–263.

[24] M. Xiao, C. Jinde, Genetic oscillation deduced from hopf bifurcation in a genetic regulatory network with delays, Math. Biosci. 215 (1) (2008) 55–63.

[25] AUTO: Software for Continuation and Bifurcation Problems in Ordinary Differential Equations, Computational Mathematics and Visualization Laboratory (CMVL), Concordia University in Montreal, Canada, 1996. [online]. [Accessed 10 October 2018]. Available at: http://indy.cs.concordia.ca/auto/

[26] gbasis: Reduced Groebner Basis, MATLAB Documentation, MathWorks, 2018. [online]. [Accessed 10 October 2018]. Available at: https://www.mathworks.com/help/symbolic/gbasis.html

[27] Groebner Basis: Compute a Groebner Basis, Maplesoft, a division of Waterloo Maple Inc., Waterloo, 2016. Maplesoft [online]. [Accessed 10 October 2018]. Available at: https://www.maplesoft.com/support/help/Maple/view.aspx?path=Groebner

[28] GroebnerBasis, Wolfram Language & System Documentation Center, 2007. [online]. Wolfram Research. [Accessed 10 October 2018]. Available at:https://reference.wolfram.com/language/ref/GroebnerBasis.html

Asymptotic Behaviour and Existence of a Limit Cycle of Cubic Autonomous Systems

By Lenka Baráková

Abstract. In this paper a 2-dimensional real autonomous system with polynomial right-hand sides of a concrete type is studied. Hopf bifurcation is analysed and existence of a limit cycle is proved. A new formula to determine stability or unstability of this limit cycle is introduced. A positively invariant set, which is globally attractive, is found. Consequently, existence of a stable limit cycle around an unstable critical point is proved and also a sufficient condition for non-existence of a closed trajectory in the phase space is given. Global characteristics of the system are studied. An application in economics to the dynamic version of the neo-keynesian macroeconomic IS-LM model is presented.

1. Introduction

There is a lot of papers dealing with asymptotic properties and existence of a limit cycles of 2-dimensional polynomial autonomous differential systems. A great deal of scientific effort was devoted to quadratic systems, see e.g. [4], [5], [12], [13]. In the present paper we shall consider the real dynamical autonomous system in the plane

$$
(1.1) \qquad
\begin{aligned}
\dot{x}_1 &= \alpha(a_0 + a_1 x_1 + b x_2 + a_2 x_1^2 + a_3 x_1^3), \\
\dot{x}_2 &= \beta(c_0 + c x_1 + d x_2),
\end{aligned}
$$

where $x_1 \in IR$, $x_2 \in IR$, $\alpha, \beta > 0$ are real parameters and $a_0, a_1, b, a_2, a_3, c_0, c, d$ are real coefficients satisfying following assumptions:

$$
(1.2) \qquad b < 0, \quad a_3 < 0, \quad c > 0, \quad d < 0
$$

and the quadratic equation

$$
(1.3) \qquad a_1 + 2a_2 x + 3a_3 x^2 = 0 \quad \text{has two real roots.}
$$

2

2.   Null curves of the system (1.1)

1. x_1 - null curve $x_2 = \varphi(x_1)$ of the system (1.1) is defined for all $x_1 \in IR$ and satisfies:

$$\varphi(x_1) = -\frac{a_0 + a_1 x_1 + a_2 x_1^2 + a_3 x_1^3}{b}.$$

It follows that

$$\frac{d\varphi}{dx_1} = -\frac{a_1 + 2a_2 x_1 + 3a_3 x_1^2}{b}.$$

That yields

$$\frac{d\varphi}{dx_1} < 0 \qquad \text{for} \qquad x_1 \in (-\infty, M) \cup (N, \infty),$$

$$\frac{d\varphi}{dx_1} > 0 \qquad \text{for} \qquad x_1 \in (M, N),$$

where $M = \frac{-a_2 + \sqrt{a_2^2 - 3a_3 a_1}}{3a_3}, \quad N = \frac{-a_2 - \sqrt{a_2^2 - 3a_3 a_1}}{3a_3}.$

2. x_2 - null curve $x_2 = \psi(x_1)$ of the system (1.1) is defined for all $x_1 \in IR$ and satisfies:

$$\psi(x_1) = -\frac{c_0 + c x_1}{d}.$$

It follows that

$$\frac{d\psi}{dx_1} = -\frac{c}{d} = const. > 0.$$

Critical points of (1.1) are the intersections of these null curves. Obviously, the conditions (1.2) assure that there exists at least one critical point of (1.1).

3.   Linearization of the system (1.1)

Let

$$\begin{aligned}
A &= \begin{pmatrix} \alpha(a_1 + 2a_2 x_1 + 3a_3 x_1^2) & \alpha b \\ \beta c & \beta d \end{pmatrix}, \\
\operatorname{tr} A &= \alpha(a_1 + 2a_2 x_1 + 3a_3 x_1^2) + \beta d, \\
\det A &= \alpha\beta\big[d(a_1 + 2a_2 x_1 + 3a_3 x_1^2) - cb\big], \\
p_A(\lambda) &= \det(A - \lambda I) = \lambda^2 - \lambda \operatorname{tr} A + \det A.
\end{aligned}$$

Assume that A is a regular matrix. The matrix A is a matrix of a linearized system (1.1). The matrix A, its trace $\operatorname{tr} A$, its determinant $\det A$ and its characteristic polynom $p_A(\lambda)$ are functions of x_1. The system (1.1) has at least one critical point. Let (x_1^*, x_2^*) be any critical point of the system (1.1). In the following part we will discuss the type

Baráková, Asymptotic Behaviour and Existence of a Limit Cycle   3

of the critical point (x_1^*, x_2^*) and we will write $\operatorname{tr} A$, $\det A$ instead of $\operatorname{tr} A(x_1^*, x_2^*)$ and $\det A(x_1^*, x_2^*)$ to make the computation clearer.

Denoting $\lambda_1, 2$ roots of the characteristic polynom, we obtain

$$\lambda_{1,2} = \frac{\operatorname{tr} A}{2} \pm \frac{\sqrt{D}}{2},$$

where $D = (\operatorname{tr} A)^2 - 4 \det A$.

1) The point (x_1^*, x_2^*) is a positive attractor in the case that $x_1^* \in (-\infty, M) \cup \langle N, \infty)$. $\frac{d\varphi}{dx_1}(x_1^*) < 0$ implies $\operatorname{tr} A < 0$. Two cases are possible:

(i) If D is negative, $(\operatorname{tr} A)^2 < 4 \det A$, then the critical point (x_1^*, x_2^*) is a stable focus.

(ii) If D is non-negative, $(\operatorname{tr} A)^2 \geq 4 \det A$, then the critical point (x_1^*, x_2^*) is a stable node.

2) In the case that $x_1^* \in (M, N)$, three types can occur at the point (x_1^*, x_2^*).

Positive attractor:

(3.1) $\operatorname{tr} A < 0,$ that is $\alpha(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}) + \beta d < 0,$

(3.2) $\det A > 0,$ that is $\alpha\beta\big[d(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}) - cb\big] > 0.$

Similarly to the previous case 1) we obtain either a stable node or a stable focus.

Negative attractor:

(3.3) $\operatorname{tr} A > 0,$ that is $\alpha(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}) + \beta d > 0,$

        $\det A > 0,$ that is $\alpha\beta\big[d(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}) - cb\big] > 0.$

This can be either an unstable node or an unstable focus.

Saddle point:

The case of a saddle point occurs if and only if

(3.4) $\det A < 0.$

It is seen that it coresponds to the inequality

$$\frac{d\varphi}{dx_1}(x_1^*) = -\frac{a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}}{b} > -\frac{c}{d} = \frac{d\psi}{dx_1}(x_1^*).$$

4.   Hopf Bifurcation in the system (1.1)

Here we will show that Hopf bifurcation can occur for $\alpha = \tilde{\alpha}$ in the case that $x_1^* \in (M, N)$ holds at the critical point $P = (x_1^*, x_2^*)$. According to the Hopf theorem on existence of a limit cycle (see [1], chap. 26, page 406), these two conditions have to be fulfilled at the critical point P:

4

(i) $\operatorname{tr} A = 0$ and $\det A > 0$ — the eigenvalues are purely imaginary:

$$\lambda_{1,2} = \frac{\pm\sqrt{(\operatorname{tr} A)^2 - 4\det A}}{2} = \pm\sqrt{-\det A},$$

(ii) $\operatorname{tr} A$ satisfies

$$\frac{d\operatorname{tr} A}{d\alpha}\,|_{\alpha=\tilde{\alpha}} > 0.$$

Suppose that coefficients of the right-hand side of (1.1) are fixed and let α and β be parameters. With respect to the previous results, changes of stability in the phase space caused by changes of parameters can occur only if $x_1^* \in (M, N)$ (note that $\det A > 0$ in condition (i) excludes a saddle point). Furthermore, stability or unstability of the critical point depends on the proportion of parameters α and β, so we can fix $\beta = \tilde{\beta} > 0$ arbitrary. Since

$$\alpha > 0, \quad \tilde{\beta} > 0, \quad b < 0, \quad d < 0, \quad \frac{d\varphi}{dx_1}(x_1^*) > 0,$$

there exists $\alpha = \tilde{\alpha} > 0$ such that

$$\operatorname{tr} A = \tilde{\alpha}(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}) + \tilde{\beta}d = 0.$$

If the value of parameter α is moved, the value of $\operatorname{tr} A$ changes. $0 < \alpha < \tilde{\alpha}$ implies $\operatorname{tr} A < 0$ and the critical point P is stable; $\alpha > \tilde{\alpha}$ implies $\operatorname{tr} A > 0$ and the critical point P is unstable. The condition (i) is fulfiled for $\alpha = \tilde{\alpha}$. Since

$$\frac{d\varphi}{dx_1}(x_1^*) = -\frac{a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}}{b} > 0,$$

which implies

$$\frac{d\operatorname{tr} A}{d\alpha}\,|_{\alpha=\tilde{\alpha}} = \frac{d[\alpha(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}) + \tilde{\beta}d]}{d\alpha} = a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2} > 0,$$

the condition (ii) is also fulfilled.

The critical value of α, where Hopf bifurcation occurs is

$$\tilde{\alpha} = -\frac{\tilde{\beta}d}{a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}}.$$

According to Hopf theorem, if for $\alpha = \tilde{\alpha}$ the critical point P is not a centre, then for values of α sufficiently close to $\tilde{\alpha}$ there exists a periodic solution in some neighbourhood of P.

We proved existence of a limit cycle, but that is not enough to describe the asymptotic behaviour of the solutions of (1.1). To find the type of bifurcation, we are going to use Hopf theorem on stability of the limit cycle (see [10], chap. 4, page 344). Here we will cite a simplified version to make clear the further calculation.

Hopf theorem (on stability)

Consider a general planar analytic system

$$\dot{x} = ax + by + p(x, y),$$
$$\dot{y} = cx + dy + q(x, y),$$

with parameter $a + d$ (the trace of the linearized matrix), where

$$p(x, y) = \sum_{i+j \geq 2} a_{ij} x^i y^j, \quad q(x, y) = \sum_{i+j \geq 2} b_{ij} x^i y^j$$

are analytic functions. The origin is a critical point.

If $a + d = 0$, $\Delta = ad - bc > 0$, then the type of Hopf bifurcation with the parameter $a + d$ is given by the sign of the Lyapunov number σ:

$$
\begin{aligned}
\sigma = \ & -\frac{3\pi}{2b\Delta^{3/2}} \big\{ ac(a_{11}^2 + a_{11}b_{02} + a_{02}b_{11}) + ab(b_{11}^2 + a_{20}b_{11} + a_{11}b_{02}) + \\
& + c^2(a_{11}a_{02} + 2a_{02}b_{02}) - 2ac(b_{02}^2 - a_{20}a_{02}) - 2ab(a_{20}^2 - b_{02}b_{20}) - \\
& - b^2(2a_{20}b_{20} + b_{11}b_{20}) + (bc - 2a^2)(b_{11}b_{02} - a_{11}a_{20}) - \\
& - (a^2 + bc)\big[3(cb_{03} - ba_{30}) + 2a(a_{21} + b_{12}) + (ca_{12} - bb_{21})\big] \big\}.
\end{aligned}
$$

If $\sigma < 0$, then for $a + d > 0$ sufficiently close to 0, there exists a stable limit cycle, supercritical Hopf bifurcation occurs.

If $\sigma > 0$, then for $a + d < 0$ sufficiently close to 0, there exists an unstable limit cycle, subcritical Hopf bifurcation occurs.

P r o o f. See [1] or [10], older theory can be found in [2] and [3].

So the sign of σ determines the asymptotic behaviour of those solutions of (1.1), whose ω-limit set is the limit cycle guaranteed by Hopf bifurcation.

To find the formula for σ, we have to translate the system (1.1) to the origin. After the transformation

$$y_1 = x_1 - x_1^*, \quad y_2 = x_2 - x_2^*$$

we obtain the system (1.1) in the form:

$$
\begin{aligned}
(4.1) \quad \dot{y}_1 &= \alpha[(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2})y_1 + by_2 + (a_2 + 3a_3 x_1^*)y_1^2 + a_3 y_1^3], \\
\dot{y}_2 &= \beta(cy_1 + dy_2).
\end{aligned}
$$

After some arrangements, we obtain the following formula for the system (4.1) with $\alpha = \tilde{\alpha}$, $\beta = \tilde{\beta}$:

$$(4.2) \qquad \sigma = \frac{3\pi}{2\Delta^{3/2}} \tilde{\alpha}^3 M_1^2 \left[2\frac{M_2^2}{M_1} - 3a_3 \left(1 - \frac{b}{M_1}\frac{c}{d}\right) \right],$$

where $M_1 = a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2} > 0$, $M_2 = a_2 + 3a_3 x_1^*$, $\Delta = \tilde{\alpha}\tilde{\beta}[(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2})d - cb]$. It is seen that the sign of σ does not depend on parameters α and β. The sign of σ can be both plus or minus as it is shown in the following two examples:

Example 1. Let

$$
\begin{aligned}
a_0 &= 40, \quad a_1 = -9.64, \quad a_2 = 1, \quad a_3 = -0.033, \\
b &= -0.8, \quad c_0 = -3, \quad c = 0.7, \quad d = -0.3
\end{aligned}
$$

6

and fix

$$\beta = 1.$$

The system (1.1) with this coefficients satisfies assumptions (1.2) and (1.3) and the only critical point is (x_1^*, x_2^*):

$$x_1^* \doteq 9.9526, \quad x_2^* \doteq 13.2228$$

The critical value of the parameter α, where Hopf bifurcation occurs is $\tilde{\alpha} \doteq 0.6538$. According to the formula (4.2), we get

$$\sigma \doteq -1.2612 < 0.$$

A stable limit cycle occurs for those values of α, which are sufficiently close to $\tilde{\alpha}$ and such that $\operatorname{tr} A > 0$.
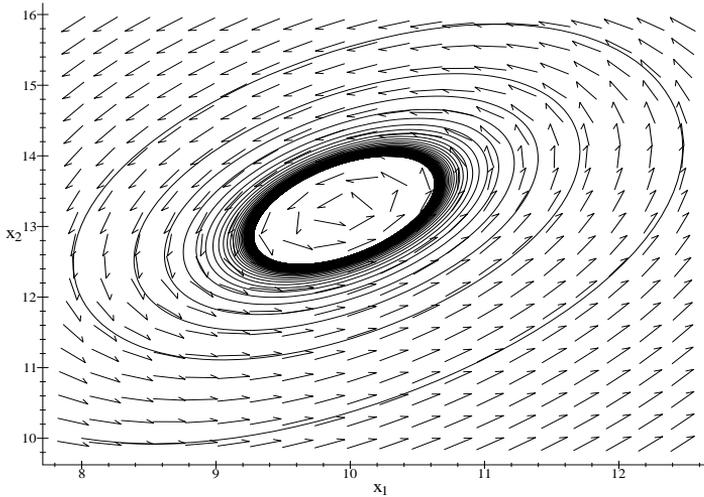
Fig. 1 $\alpha = 0.67 > \tilde{\alpha}$: a stable limit cycle around an unstable focus

Example 2. Let

$$a_0 = 1, \quad a_1 = 2, \quad a_2 = 3, \quad a_3 = -0.1,$$
$$b = -1, \quad c_0 = 2, \quad c = 100, \quad d = -2$$

and fix

$$\beta = 1.$$

The system (1.1) with this coefficients satisfies assumptions (1.2) and (1.3) and the only critical point is (x_1^*, x_2^*).

$$x_1^* = 0, \quad x_2^* = 1.$$

The critical value of the parameter α, where Hopf bifurcation occurs is $\tilde{\alpha} = 1$. According to the formula (4.2), we get

$$\sigma \doteq 0.0180 > 0.$$

An unstable limit cycle occurs for those values of α, which are sufficiently close to $\tilde{\alpha}$ and such that tr $A < 0$.
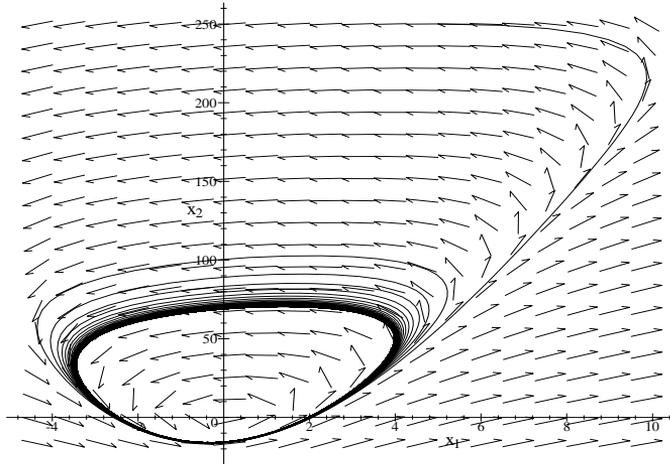
Fig. 2 $\alpha = 0.93 < \tilde{\alpha}$: an unstable limit cycle around a stable focus.

In the second example, Theorem 5.3 of the next section guarantees existence of a stable limit cycle, which contains the unstable cycle.

5.   Existence of a globally attractive set

In this section a positively invariant set, which is globally attractive will be found using a suitable Lyapunov-like function. This result is usefull partly to characterization of the global behaviour of solutions of (1.1) and partly for the application of the Poincaré-Bendixson theorem. These give more precise statements about the phase portrait, which are presented in the section 6.

Theorem 5.1. *If (4.1) is the previous translated system belonging to the system (1.1) and satisfying assumptions (1.2) and (1.3), then*

$$\left\{ [y_1, y_2] \in I\!R^2 \mid \frac{\beta c}{2} y_1^2 - \frac{\alpha b}{2} y_2^2 \leq R \right\}$$

8

is a positively invariant set of the system (4.1) for each $R \geq -\frac{\varrho}{2\beta d}$, where

$$(5.1) \quad \varrho = \max_{y_1 \in I\!R} \left(\alpha\beta c y_1^2 (a_3 y_1^2 + (a_2 + 3a_3 x_1^*) y_1 + a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2} - \frac{\beta}{\alpha} d) \right).$$

Proof. Here, to make the further calculations clearer, we will write the system (4.1) with simplified coefficients:

$$\begin{aligned}
\dot{y}_1 &= Py_1 - Qy_2 + Sy_1^2 - Ty_1^3, \\
\dot{y}_2 &= Uy_1 - Wy_2,
\end{aligned}$$

where we denote

$$\begin{aligned}
P &= \alpha(a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}), \\
Q &= -\alpha b > 0, \\
S &= \alpha(a_2 + 3a_3 x_1^*), \\
T &= -\alpha a_3 > 0, \\
U &= \beta c > 0, \\
W &= -\beta d > 0.
\end{aligned}$$

Let $g(y_1) = -Uy_1^2(Ty_1^2 - Sy_1 - P - W)$. Notice that function $g(y_1)$ is bounded from above on $I\!R$, since g is continuous and

$$\lim_{y_1 \to \pm\infty} g(y_1) = -\infty.$$

This together with $g(0) = 0$ yields

$$\varrho = \max_{y_1 \in I\!R} g(y_1) \geq 0.$$

Let

$$R \geq \frac{\varrho}{2W} \geq 0$$

be arbitrary.

Put $y = (y_1, y_2)$ and consider the function $V(y) = \frac{U}{2}y_1^2 + \frac{Q}{2}y_2^2 = \frac{\beta c}{2}y_1^2 - \frac{\alpha b}{2}y_2^2$.

If $R = 0$, then the set given by the inequality $V(y) \leq R$ is the origin, which is a critical point of (4.1) and the statement is proved in this case.

If $R > 0$, then $V(y) = R^*$ is an ellipse for any $R^* \geq R$ and the set $V(y) \leq R^*$ contains the origin and the ellipse $V(y) = R$. Let $y = y(t) = (y_1(t), y_2(t))$ be any non-trivial solution of (4.1). If $V(y(t)) = R^*$, the solution $y = y(t)$ intersects the ellipse $V(y) = R^*$ at the time t. Then the following relations hold for any t such that $V(y(t)) = R^*$:

$$\begin{aligned}
V'(y(t)) &= Uy_1\dot{y}_1 + Qy_2\dot{y}_2 = \\
&= U(Py_1^2 - Qy_1y_2 + Sy_1^3 - Ty_1^4) + Q(Uy_1y_2 - Wy_2^2) = \\
&= -W(Uy_1^2 + Qy_2^2) + U(P + W)y_1^2 + USy_1^3 - UTy_1^4 = \\
&= -2WR^* + Uy_1^2(-Ty_1^2 + Sy_1 + P + W) \leq \\
&\leq -2WR^* + \varrho \leq 0.
\end{aligned}$$

Clearly the last inequality is strict for $R^* > R$. This implies that trajectories of (4.1) intersect the ellipses $V(y) = R^*$, $R^* > R$, in the direction from their exteriors to their interiors. Consequently, if $y(t^*)$ is an element of $V(y) \leq R$ for some t^*, i.e. if $V(y(t^*)) \leq R$, then the trajectory corresponding to the solution $y(t)$ cannot leave the set $V(y) \leq R$ for $t \geq t^*$. Hence $V(y) \leq R$ is a positively invariant set and the proof is completed.

Theorem 5.2. *Let (x_1^*, x_2^*) be a critical point of the system (1.1) with assumptions (1.2) and (1.3) and let ϱ be defined by (5.1). Then the set*

$$E = \left\{ [x_1, x_2] \in IR^2 \mid \frac{\beta c}{2}(x_1 - x_1^*)^2 - \frac{\alpha b}{2}(x_2 - x_2^*)^2 \leq -\frac{\varrho}{2\beta d} \right\}$$

is globally attractive.

Proof. In view of Theorem 5.1, any set

$$\left\{ [x_1, x_2] \in IR^2 \mid \frac{\beta c}{2}(x_1 - x_1^*)^2 - \frac{\alpha b}{2}(x_2 - x_2^*)^2 \leq R \right\},$$

where $R \geq -\frac{\varrho}{2\beta d}$, is positively invariant. Thus every solution $x(t)$ of (1.1) is bounded and therefore defined for all $t \geq t_0$, where t_0 is an initial value of t for the solution $x(t)$. Denote $x^* = (x_1^*, x_2^*)$ and $y(t) = x(t) - x^*$. Then $y(t)$ is a solution of (4.1) defined for $t \geq t_0$. From the proof of Theorem 5.1, we can easily see that $V'(y(t)) < 0$ for any $t \geq t_0$ for which the inequality $V(y(t)) > -\frac{\varrho}{2\beta d}$ is valid. This implies that $V(y(t))$ is decreasing for all $t \geq t_0$ with the mentioned property.

Choose $\varepsilon > -\frac{\varrho}{2\beta d}$ arbitrary. Since any set $V(y) \leq R$, where $R \geq -\frac{\varrho}{2\beta d}$, is positively invariant, it is sufficient to show that there exists a $t^* \geq t_0$ such that $V(y(t^*)) < \varepsilon$.

Suppose on the contrary that there is not such a t^*. Then, in view of the monotonicity of $V(y(t))$, the limit

$$\lim_{t \to \infty} V(y(t)) =: \xi$$

exists and

$$V(y(t)) \geq \xi \geq \varepsilon$$

for all $t \geq t_0$. Calculating the derivative $V'(y(t))$ similarly as in the proof of Theorem 5.1, we obtain the estimation

$$V'(y(t)) \leq -2WV(y(t)) + \varrho \leq 2W\xi + \varrho = 2\beta d\xi + \varrho =: \kappa < 0$$

for $t \geq t_0$. The integration over $[t_0, t]$ yields

$$V(y(t)) - V(y(t_0)) \leq \kappa(t - t_0)$$

for all $t \geq t_0$. Consequently $V(y(t)) \to -\infty$ for $t \to \infty$, which is contradiction with the non-negativity of $V(y(t))$.

Corollary 5.3. *Let $x^* = (x_1^*, x_2^*)$ be a critical point of the system (1.1). If*

$$\max_{y_1 \in IR} y_1^2 \left(a_3 y_1^2 + (a_2 + 3a_3 x_1^*)y_1 + a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2} - \frac{\beta}{\alpha}d \right) = 0,$$

10

then the unique solution corresponding to the critical point x^* is globally asymptotically stable.

P r o o f. If (5.1) holds, then $\varrho = 0$ using the notation from the proof of Theorem 5.1. Note $R \geq 0$ is fixed, but arbitrary. Then the inequality

$$V'(y(t)) \leq -2WR + \varrho = -2WR$$

yields that the orbital derivative of V is negative definite and this guarantees global asymptotical stability of the trivial solution.

Remark 5.4. Since the globally attractive set E from the Theorem 5.2 is depends on the critical point $P = (x_1^*, x_2^*)$, we denote $E = E(x_1^*, x_2^*)$ the globally attractive set related to the point P in the further section.

Definition 5.5. Let
(5.2) $\dot{x} = f(x)$

be an autonomous system, where $x = (x_1, \ldots, x_n) \in \Omega \subseteq IR^n$ and the vector function $f = (f_1, \ldots, f_n) : \Omega \to IR^n$ has continuous derivative. Let x^* be a critical point of (5.2) and $Df(x^*)$ be the Jacobi's matrix of the right-hand side of the system (5.2) at the point x^*. If the real parts of all eigenvalues of $Df(x^*)$ are non-zero, then the critical point is called hyperbolic.

6.   Global properties of the system (1.1)

Attributes of a structurally stable system (1.1) are studied in this section. For the precise definition of structural stability see [9]. Structural stability guarantees especially that all critical points are hyperbolic (see Peixoto's theorem in [9]). The main reason for studying the structurally stable system (1.1) is its potential application in economics (see the section 7), because the economic values cannot be measured precisely. Studying a structurally unstable system is purposeless in these applications, since small changes in the system make even qualitative changes of the phase portrait and different behaviour of described variables occures.

Lemma 6.1. *If the system (1.1) with assumptions (1.2) and (1.3) is structurally stable, then the following three statements hold.*

 (i) *The system has one or three critical points.*

 (ii) *If there is exactly one critical point (x_1^*, x_2^*), then (x_1^*, x_2^*) is a node or focus.*

 (iii) *If there are three critical points (x_1^1, x_2^1), (x_1^2, x_2^2) and (x_1^3, x_2^3), such that $x_1^1 < x_1^2 < x_1^3$, then (x_1^2, x_2^2) is a saddle point and (x_1^1, x_2^1), (x_1^3, x_2^3) are unstable nodes or foci.*

P r o o f. (i) Critical points of the system (1.1) are the intersections of null-curves $x_2 = \varphi(x_1)$, $x_2 = \psi(x_1)$. Since the first one is a cubic parabola and the second one is a line, the null-curves can intersect at three points at most. The condition (1.2) guarantees existence of at least one critical point.

Suppose there are exactly two critical points. Then the null curves are tangent to each other at some critical point (x_1^*, x_2^*), i.e.

$$\frac{d\varphi}{dx_1}(x_1^*) = -\frac{a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}}{b} = -\frac{c}{d} = \frac{d\psi}{dx_1}(x_1^*).$$

It follows that

$$\det A = 0$$

at the point (x_1^*, x_2^*) and the characteristic polynom satisfies

$$p_A(\lambda) = \lambda^2 - \lambda \operatorname{tr} A + \det A = \lambda(\lambda - \operatorname{tr} A).$$

That is a contradiction with structural stability, since zero is an eigenvalue of the matrix A at the point (x_1^*, x_2^*) and therefore the critical point (x_1^*, x_2^*) is not hyperbolic. Hence, null curves intersects at one or at three points.

(ii) Let (x_1^*, x_2^*) be a unique critical point of (1.1). If x_1 lies outside the interval (M, N), then

$$\frac{d\varphi}{dx_1}(x_1^*) < 0$$

and that implies $\operatorname{tr} A < 0$. Hence, (x_1^*, x_2^*) can be a stable node or focus.

If x_1 lies inside the interval (M, N), then it is evident that the following inequality is satisfied

(6.1) $$\frac{d\varphi}{dx_1}(x_1^*) = -\frac{a_1 + 2a_2 x_1^* + 3a_3 x_1^{*2}}{b} \leq -\frac{c}{d} = \frac{d\psi}{dx_1}(x_1^*).$$
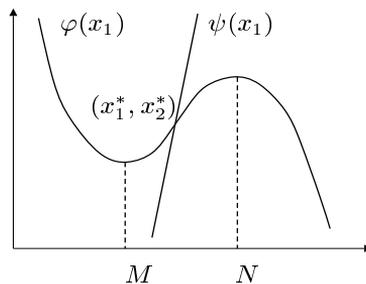
Fig.3

The inequality (6.1) holds if and only if $\det A \geq 0$ at the point (x_1^*, x_2^*). Similarly as in the proof of (i), structural stability of the system (1.1) excludes $\det A = 0$ and the condition (3.2) holds therefore. Suppose $\operatorname{tr} A = 0$, then the characteristic polynom

$$p_A(\lambda) = \lambda^2 - \lambda \operatorname{tr} A + \det A = \lambda^2 + \det A$$

12

has two purely imaginary eigenvalues. This is a contradiction with the structural stability of the system (1.1). Consequently, one of the conditions (3.1) or (3.3) holds and the critical point (x_1^*, x_2^*) can be either a node or a focus.

(iii) In the case that the null-curves intersect at three points, the "middle" critical point (x_1^2, x_2^2) is a saddle, since the inequality

$$\frac{d\varphi}{dx_1}(x_1^2) = -\frac{a_1 + 2a_2 x_1^2 + 3a_3 x_1^{2^2}}{b} > -\frac{c}{d} = \frac{d\psi}{dx_1}(x_1^2)$$

equivalent to (3.4) holds at this point. The oposite inequality holds for the other two points (x_1^1, x_2^1) and (x_1^3, x_2^3), i.e. the determinant $\det A$ is positive at these points. Reasonings analogous to the part (ii) of the proof lead to the consequence that they are nodes or foci necessarily.

Theorem 6.2. *If the system (1.1) with assumptions (1.2) and (1.3) is structurally stable, then the following three statements hold.*

(i) *If system (1.1) has exactly one critical point $P = (x_1^*, x_2^*)$, then either the constant solution coresponding with P is globally asymptotically stable, or $E(x_1^*, x_2^*)$ contains a stable limit cycle.*

(ii) *If system (1.1) has three unstable critical points $P_1 = (x_1^1, x_2^1)$, $P_2 = (x_1^2, x_2^2)$ and $P_3 = (x_1^3, x_2^3)$, where $x_1^1 < x_1^2 < x_1^3$, then $E(x_1^2, x_2^2)$ contains a stable limit cycle.*

(iii) *If system (1.1) has exactly one unstable critical point (x_1^*, x_2^*), then $E(x_1^*, x_2^*)$ contains a stable limit cycle.*

(iv) *If system (1.1) has exactly one critical point (x_1^*, x_2^*) and there exists $R > -\frac{\varrho}{2\beta d}$, ϱ given by (5.1), such that the inequality*

$$\alpha(a_1 + 2a_2 x_1 + 3a_3 x_1^2) + \beta d \leq 0$$

is fulfilled in the set $H = \left\{ (x_1, x_2) \in I\!R^2 \mid \frac{\beta c}{2}(x_1 - x_1^*)^2 - \frac{\alpha b}{2}(x_2 - x_2^*)^2 < R \right\}$, then the constant solution $x_1 = x_1^*$, $x_2 = x_2^*$ is globally asymptotically stable.

Proof.

(i) Let the system (1.1) has exactly one hyperbolic critical point $P = (x_1^*, x_2^*)$. According to Lemma 6.1, part (ii), P is a node or focus. Hence, $\det A > 0$ holds at P. Since P is hyperbolic, the trace $\operatorname{tr} A$ is non-zero at the point P.

Suppose there is no limit cycle. We will show that then the constant solution corresponding with P is globally asymptotically stable. Suppose on the contrary that there exists a trajectory C such that its ω-limit set $\omega(C+)$ is not equal to the point P. Since P is the unique critical point of the system (1.1), the ω-limit set $\omega(C+)$ does not contain any critical point. According to Poincaré-Bendixson theorem, $\omega(C+)$ is a closed trajectory. Clearly, it is contained in the globally attractive set $E(x_1^*, x_2^*)$ related to the point P. The closed trajectory $\omega(C+)$ surrounds P (the Poincaré's index of a closed curve is $+1$ and is equal to the

sum of indices of contained critical points). The ω-limit set of any trajectory entering the set $E(x_1^*, x_2^*)$ then does not contain any critical point and according to the Poincaré-Bendixson theorem there exists a limit cycle inside $E(x_1^*, x_2^*)$, which is a contradiction.

Otherwise, let there exists a limit cycle. It is contained in the globally attractive set $E(x_1^*, x_2^*)$ related to P and it surrounds P. The assumption of structural stability gives that there can be either an unstable or a stable limit cycle, semi-stable cycles cannot occur (see [3]). If the inequality $\operatorname{tr} A > 0$ holds at P, condition (3.3), P is a single unstable critical point (a node or a focus). Since P lies inside the globally attractive set $E(x_1^*, x_2^*)$ related to P, existence of a stable limit cycle around P follows from Poincaré-Bendixson theorem. If the inequality $\operatorname{tr} A < 0$ holds at P, condition (3.1), P is a single stable node or focus. Since the limit cycle contains the point P, the ω-limit set of any trajectory entering the set $E(x_1^*, x_2^*)$ does not contain any critical point. According to the Poincaré-Bendixson theorem, there exists a stable limit cycle inside $E(x_1^*, x_2^*)$. In this case, the structural stability implies co-existence a stable and an unstable limit cycle around the stable critical point P.

(ii) Since $E(x_1^2, x_2^2)$ is globally attractive, P_1, P_2 and P_3 are contained in E. According to Lemma 6.1, part (iii), P_2 is a saddle point, P_1 and P_3 are unstable nodes or foci. Therefore, there exists a trajectory C which remains in $E(x_1^2, x_2^2)$ and its ω-limit set $\omega(C+)$ does not contain any critical point. From Poincaré-Bendixson theorem follows that $E(x_1^2, x_2^2)$ contains a stable limit cycle.

(iii) Follows from (i).

(iv) Let f_1, f_2 be the right-hand sides of (1.1). Using the Bendixson criterion on non-existence of closed trajectories (see [16]), we have

$$\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} = \alpha(a_1 + 2a_2 x_1 + 3a_3 x_1^2) + \beta d \leq 0$$

in H, which is a simply connected domain. Hence, any limit cycle in H cannot exist. Since the globally attractive set $E(x_1^*, x_2^*)$ is a subset of H, there are no limit cycles in $E(x_1^*, x_2^*)$. The statement follows from (i).

7.   Applications in Economics

In this section, a possible application of the previous statements to the neo-keynesian macroeconomic model IS-LM is given. The point is to summarize the purpose of each section and show the possibilities of using the presented results.

Neo-keynesian macroeconomic model IS-LM can be formulated as a planar dynamic system given by equations

$$(7.1) \qquad \begin{aligned} \dot{y} &= \alpha[i(y, r) - s(y, r)], \\ \dot{r} &= \beta[l(y, r) - m], \end{aligned}$$

14

where $\alpha, \beta > 0$ are parameters and the following notation is used:

Y ... real net product, $Y \in I\!R_+$,

y ... $y = \ln Y$, $y \in I\!R$,

r ... real interest rate, $r \in I\!R$,

$I(Y,r)$ ... real investments,

$i(y,r)$ ... propensity to invest or real investment-real net product ratio i.e. $\frac{I}{Y}$,

$S(Y,r)$ ... real savings,

$s(y,r)$ ... propensity to save or real saving-real income ratio i.e. $\frac{S}{Y}$,

$L(Y,R)$ ... demand for money,

$l(y,r)$ ... $l(y,r) = L(e^y, r)$,

m ... real supply of money.

The continuity of all functions and their first derivatives in both variables y and r is assumed. In general, i and s are non-linear analytical functions. The derivatives of i, s, l functions are assumed to have the usual properties

$$(7.2) \qquad i_y > 0, \quad i_r < 0, \quad s_y > 0, \quad s_r > 0, \quad l_y > 0, \quad l_r < 0.$$

Economic observations yields that the i, s functions are of a sigmoid form (S-shaped) for any fixed r (for more economic details see [7], pages 442–443, or [15], pages 144–145). Furthermore, it is assumed that for some fixed r_0 curves i, s intersects at three points. That corresponds to existence of two points M, N, where $-\infty < M \leq N < \infty$ such that

$$\frac{\partial(i(y,r)-s(y,r))}{\partial y} < 0 \qquad \text{for} \quad y \in (-\infty, M) \cup (N, \infty),$$

$$\frac{\partial(i(y,r)-s(y,r))}{\partial y} > 0 \qquad \text{for} \quad y \in (M, N).$$
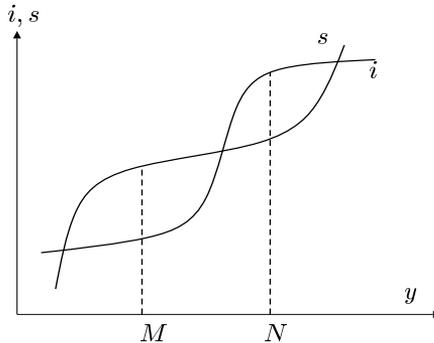
Fig.4

These assumptions on shape and position of functions $i(y,r)$, $s(y,r)$ are called the Caldor's assumption in economics.

This neo-keynesian model IS-LM has been studied by many economists. In their works, existence of an economic equilibrium is implicitly presumed although non of the above assumptions guarantees this existence. Such assumption can be represented by following two natural conditions:

$$\lim_{r \to \infty} [i(y,r) - s(y,r)] = -\infty, \qquad \lim_{r \to -\infty} [i(y,r) - s(y,r)] = \infty,$$

$$\text{for arbitrary fixed } y \in (-\infty, \infty),$$

(7.3)   and

$$\lim_{y \to \infty} l(y,r) = \infty, \qquad \lim_{y \to -\infty} l(y,r) = -\infty,$$

$$\text{for arbitrary fixed } r \in (-\infty, \infty).$$

Since null-curves of (7.1) correspond to the points of equilibrium on the goods market and money market, they are called the IS curve and LM curve in economy.

The IS-LM model (7.1) with assumptions (7.2), (7.3) and the Caldor's assumption is too general for precise mathematical analysis. In spite of it, some general results have been published, especially, we have to mention the proof of existence of a limit cycle by Torre (see [14]). Torre's results are direct applications of Andronov's theorems about bifurcation (see [2] and [3]).

Many economists and mathematics tried to apply the Poincaré-Bendixson theorem to the general system (7.1), but they never succeeded. In order to obtain the needed results, they had to choose one of two alternatives: either they changed some of the economic assumptions, or they had to specialize the situation by adding some other conditions. The first way was used by Chang and Smyth, who assumed $s_r < 0$ in (7.2) and that was why their work has not been accepted by economists. To the second group of works belong for example [11].

The applications of the results of this paper are of this second type too. The aproximation of the functions $i(y,r)$ and $s(y,r)$ by a cubic polynom and by a linear function, respectively, can give more precise information about the behaviour of economy than the general model. In concrete applications, the above general formulation of the dynamical IS-LM model together with related assumptions can be simplified to the form of the system (1.1), where

$$x_1 = y, \quad x_2 = r,$$
$$i(y,r) - s(y,r) = a_0 + a_1 x_1 + b x_2 + a_2 x_1^2 + a_3 x_1^3,$$
$$l(y,r) - m = c_0 + c x_1 + d x_2.$$

It is seen that assumption (1.2) of the system (1.1) corresponds with the assumption (7.2) of the system (7.1), the Caldor's assumption can be formulated into (1.3), while (7.3) are also fulfilled. Results from the previous sections lead to the following consequences.

The division from the section 3 specifies types of critical points. A critical point of (7.1) corresponds to an aggregate equilibrium in economy. The type of this equilibrium does not say anything more than how the economy will behave "near" the equilibrium. That is usually not sufficient, since we do not know whether we are sufficiently close to the equilibrium or not beforehand.

More useful consequences give results of sections 4, 5 and 6. We proved existence of a limit cycle in section 4 (using other apparatus, Torre have already done it in [14] for a general neo-keynesian model IS-LM). A limit cycle (especially a stable limit cycle) of (7.1) represents a business cycle in economy. In economic terminology, there

16

is no difference between oscilations with fixed or mildly varialble amplitude or period, tending to some constant amplitude or period. All these situations are called business cycles. Applying results of section 4, we get:

If the economy equilibrium $P = (y^*, r^*)$ satisfies $y^* \in (M, N)$, then the conditions (i) and (ii) from the section 4 are sufficient conditions for existence of a business cycle, which is not evoked by external influences, but which is entirely determined by internal structure of the system. The formula (4.2) then replies to the question, how the economy will behave if the initial values of y and r are sufficiently close to the limit cycle. Hence, the inner-determined business cycle occurs for values $\sigma < 0$, since in this case the limit cycle is stable and for all initial values of y and r in some neighbourhood of the cycle, the trajectory tends to it - a business cycle occurs in economy for all these situations.

Applying statements of section 5 and 6, we get essential consequences which need not the assumption $y^* \in (M, N)$. According to the Theorem 5.1, we can bound the level of product y and interest rate r. If the initial values of y and r lie inside the positively invariant set, they will stay there in the future. Corollary 5.3 gives a condition for global stability of the aggregate equilibrium. Another such condition is given by Theorem 6.2, part (iv). (Many economists presume that the economic equilibrium is globally stable always, i.e. they assume there exists some mechanism of adaptation in economy. This is true for a linear IS-LM model, with $a_2 = 0$, $a_3 = 0$. If the economy satisfies the Caldor's assumption, such mechanism need not exist.) Theorem 5.2 and Theorem 6.2 give a very useful and effective economic tool then.

If, for example, the central bank contemplates a monetary expansion (increase of the real supply of money m), the set E computed for this considered situation may help to predict the reaction of economy, although we do not presume existence of any mechanism of adaptation. Theorem 6.2 may be used for more profound description of this reaction. For example, (i) says that if there is exactly one critical point (IS curve and LM curve intersect in one point, it is a normal situation), then the economy gives signs of stability and there exists a kind of mechanism of adaptation, although the equilibrium may be unstable. The part (iii) then says that in this case a stable business cycle occurs. Parts (ii) or (iv) may be used in a similar way.

In section 6, we assume that the system (1.1), and hence (7.1), is structurally stable. This assumption is necessary in the economic model, since otherwise even small mistakes in econometric estimation of functions i, s and l could lead to incorrect results.

Except the mentioned Torre's results about existence of a limit cycle in IS-LM model, all economic applications are new to the best of my knowledge.

References

[1] AMANN H., Ordinary Differential Equations: An Introduction to Nonlinear Analysis, de Gruyter Studies in Mathematics **13**, ed. Walter de Gruyter, Berlin, New York, 1990

[2] ANDRONOV A.A., LEONTOVICZ E.A., GORDON I.I., MAIER A.G., Theory of Bifurcations of Dynamical Systems in the Plane, Nauka, Moskva, 1967 (in Russian)

[3] ANDRONOV A.A. AND CHAIKIN C.E., Theory of Oscillations, English Language Edition, Princeton University Press, Princeton, New Jersey, 1949

[4] CHERKAS L.A. AND ZHILEVICH L.J., Limit Cycles of a Quadratic Differential Equation, Differentsial'nye Uravnenija **10**, (1974), 947–949 (in Russian)

[5] CHICONE C., TIAN J.H., On General Properties of Quadratic Systems, Amer. Math. Monthly **89**, (1982), 167–179

[6] CODDINGTON E.A., LEVINSON N., Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955

[7] GANDOLFO G., Economic Dynamics, Third, Completely Revised and Enlarged Edition, Berlin, Heidelgerg, New York, Springer-Verlag, 1997

[8] HARTMAN P., Ordinary Differential Equations, ed. John Wiley & sons, New York, London, Sydney, 1964

[9] PEIXOTO M.M., Structural Stability on 2-dimensional Manifolds, Topology **1**, 1962, 101–120

[10] PERKO L., Differential Equations and Dynamical Systems, Second Edition, Berlin, Heidelgerg, New York, Springer-Verlag, 1996

[11] SCHINASI G.J., Fluctuations in a Dynamic, Intermediate-Run IS-LM Model: Applications of the Poincare-Bendixson Theorem, Journal of Economic Theory **28**, 1982, 369–375

[12] SHI S.-L., A Concrete Example of the Existence of Four Limit Cycles for Planar Quadratic Systems, Sci. Sinica **23**, (1980), 153–158

[13] SHU-XIANG Y., Limit Cycles of Quadratic Systems, Acta Math. Sinica **10**, (1977), 193–205 (in Chinese)

[14] TORRE V., Existence of Limit Cycles and Control in Complete Keynesian System by Theory of Bifurcations, Econometrica **45**, No. 6, (1977), 1457–66

[15] TU P.N.V., Dynamical Systems, An Introduction with Applications in Economics and Biology, Berlin, Heidelgerg, New York, Springer-Verlag, 1992

[16] VERHULST F., Nonlinear Differential Equations and Dynamical Systems, Berlin, Heidelgerg, New York, Springer-Verlag, 1990

Masaryk University, Dept. of Math., Janáčkovo nám. 2a, 662 95 Brno, Czech Republic

BIFURCATION DIAGRAM OF A CUBIC THREE-PARAMETER AUTONOMOUS SYSTEM

LENKA BARÁKOVÁ, EVGENII P. VOLOKITIN

ABSTRACT. In this paper, we study the cubic three-parameter autonomous planar system

$$\dot{x}_1 = k_1 + k_2 x_1 - x_1^3 - x_2,$$
$$\dot{x}_2 = k_3 x_1 - x_2,$$

where $k_2, k_3 > 0$. Our goal is to obtain a bifurcation diagram; i.e., to divide the parameter space into regions within which the system has topologically equivalent phase portraits and to describe how these portraits are transformed at the bifurcation boundaries. Results may be applied to the macroeconomical model IS-LM with Kaldor's assumptions. In this model existence of a stable limit cycles has already been studied (Andronov-Hopf bifurcation). We present the whole bifurcation diagram and among others, we prove existence of more difficult bifurcations and existence of unstable cycles.

1. INTRODUCTION

In the present paper we shall consider the real dynamical autonomous system

$$\dot{x}_1 = k_1 + k_2 x_1 - x_1^3 - x_2,$$
$$\dot{x}_2 = k_3 x_1 - x_2, \tag{1.1}$$

where $x_1, x_2 \in \mathbb{R}$ and $K = \{(k_1, k_2, k_3) \in \mathbb{R}^3 : k_2 > 0, k_3 > 0\}$ is a parameter space. Note that if $x_1(t)$, $x_2(t)$ are solutions of (1.1), $\tilde{x}_1(t) = -x_1(t)$, $\tilde{x}_2(t) = -x_2(t)$ are solutions of the system

$$\dot{x}_1 = -k_1 + k_2 x_1 - x_1^3 - x_2,$$
$$\dot{x}_2 = k_3 x_1 - x_2.$$

This implies that the bifurcation sets of (1.1) are symmetric with respect to the plane $k_1 = 0$, because the phase portraits of (1.1) with the parameters $(k_1, k_2, k_3) = (\tilde{k}_1, \tilde{k}_2, \tilde{k}_3)$ and $(k_1, k_2, k_3) = (-\tilde{k}_1, \tilde{k}_2, \tilde{k}_3)$ are symmetric about the origin. We

---

denote

$$A = \begin{pmatrix} k_2 - 3x_1^2 & -1 \\ k_3 & -1 \end{pmatrix},$$

$$\operatorname{tr} A = k_2 - 3x_1^2 - 1,$$

$$\det A = 3x_1^2 - k_2 + k_3,$$

$$p_A(\lambda) = \det(A - \lambda I) = \lambda^2 - \lambda \operatorname{tr} A + \det A,$$

where A is Jacobi's matrix of the system (1.1), its trace $\operatorname{tr} A$, determinant $\det A$ and characteristic polynomial $p_A(\lambda)$ are functions of variable x_1.

All equilibrium points (ξ_1, ξ_2) of the system (1.1) have to be solutions of the equations

$$k_1 + k_2 x_1 - x_1^3 - x_2 = 0,$$

$$k_3 x_1 - x_2 = 0,$$

which gives that ξ_1 has to satisfy the equality

$$k_1 + (k_2 - k_3)\xi_1 - \xi_1^3 = 0 \tag{1.2}$$

and $\xi_2 = k_3 \xi_1$. System (1.1) has from one to three equilibrium points.

Lemma 1.1. *Let (ξ_1, ξ_2) be an equilibrium point of (1.1). Then the set*

$$\{(x_1, x_2) \in \mathbb{R}^2 : k_3(x_1 - \xi_1)^2 + (x_2 - k_3\xi_1)^2 \le R\},$$

where

$$R = -k_3 \min_{x_1 \in \mathbb{R}} \{x_1^2(x_1^2 + 3\xi_1 x_1 - k_2 + 3\xi_1^2 - 1)\}$$

is globally attractive.

For the proof of the above lemma se [2, Theorems 5.1 and 5.2] .

Remark 1.2. A planar dynamical system

$$\dot{y} = \alpha[i(y, r) - s(y, r)],$$

$$\dot{r} = \beta[l(y, r) - m], \tag{1.3}$$

where $\alpha, \beta > 0$, may represent a macroeconomical model IS-LM (see [2]). The variable $y = \ln Y$ is the logarithm of the product (GNP), r is the interest rate. Functions i and s are propensities to invest and save, l and the constant m - demand and supply of money. Using basic economic properties of the functions i, s and l (including Kaldor's assumptions), we can concretize the system (1.3) to the most simple one - a cubic system

$$\dot{y} = \alpha(a_0 + a_1 y + br + a_2 y^2 + a_3 y^3),$$

$$\dot{r} = \beta(c_0 + cy + dr), \tag{1.4}$$

where $\alpha > 0$, $\beta > 0$, $b < 0$, $a_3 < 0$, $c > 0$, $d < 0$ and the quadratic equation $a_1 + 2a_2 x + 3a_3 x^2 = 0$ has two different real roots. System (1.4) can be replaced by the system (1.1) using some efficient transformation (see [8]).

The aim of this paper is to continue in the study of the dynamical system (1.4) from [2](the system (1.1) respectively) and to obtain deeper results concerning its stability, topological properties and types of bifurcations, especially existence and stability of limit cycles. From the economic point of view stable limit cycles

correspond to business cycles. Economists are used to presume that economic equilibrium is globally stable always, i.e. they assume there exists some mechanism of adaptation in economy. This is true for a linear IS-LM model, with $a_2 = 0$, $a_3 = 0$. If the economy satisfies the Kaldor's assumptions, such mechanism need not exist. This was pointed out already in the original Kaldor's paper [5], but dealing with this problem all authors provided just numerical results or made some other specific assumptions to the model and to the best of my knowledge never found any unstable cycle. Although the system (1.4) is "only" cubic, we will show that even more than one cycle can appear and surely it need not be stable. Moreover, the described cycles are not evoked by external influences, but they are entirely determined by internal structure of the system, which is a problem passed by so called "invisible hand" that should lead the economy to the globally stable equilibrium.

2. Local bifurcations

Lemma 2.1. *Let (ξ_1, ξ_2) be an equilibrium point of (1.1) and let*

$$k_2 = k_3 + 3\xi_1^2, \quad k_3 \neq 1.$$

Then the equilibrium point (ξ_1, ξ_2) is a saddle-node for $\xi_1 \neq 0$. The origin is topologically equivalent to a node in the case $\xi_1 = 0$.

Proof. After transformation of the equilibrium point (ξ_1, ξ_2) to the origin by the change of variables $u_1 = x_1 - \xi_1$, $u_2 = x_2 - \xi_2$ we get the system

$$\dot{u}_1 = k_3 u_1 - 3\xi_1 u_1^2 - u_1^3 - u_2,$$
$$\dot{u}_2 = k_3 u_1 - u_2.$$

For $k_3 \neq 1$, the following regular transformation

$$u_1 = y_1 + y_2, \quad u_2 = k_3 y_1 + y_2$$

(the matrix of the trasformation is given by the eigenvectors corresponding with one zero and one non-zero eigenvalues) and the time change $\tau = (k_3 - 1)t$ give the canonical form of system (1.1):

$$\dot{y}_1 = F(y_1, y_2),$$
$$\dot{y}_2 = y_2 - k_3 F(y_1, y_2),$$

where

$$F(y_1, y_2) = \frac{3\xi_1}{(k_3 - 1)^2}(y_1 + y_2)^2 + \frac{1}{(k_3 - 1)^2}(y_1 + y_2)^3.$$

Let $y_2 = \varphi(y_1)$ be a solution of the equation

$$y_2 - k_3 F(y_1, y_2) = 0$$

in the neighbourhood of the origin. We approximate this solution corresponding with the central manifold of the system by a Taylor expansion

$$\varphi(y_1) = \sum_{i=0}^{\infty} a_i y_1^i$$

in the neighbourhood of the origin and get

$$\sum_{i=0}^{\infty} a_i y_1^i = \frac{3k_3\xi_1}{(k_3 - 1)^2}\Big(y_1 + \sum_{i=0}^{\infty} a_i y_1^i\Big)^2 + \frac{k_3}{(k_3 - 1)^2}\Big(y_1 + \sum_{i=0}^{\infty} a_i y_1^i\Big)^3.$$

We equate coefficients of equal powers of x on the left and the righthand side and find

$$a_0 = 0, \quad a_1 = 0, \quad a_2 = \frac{3k_3\xi_1}{(k_3 - 1)^2} \neq 0.$$

The equilibrium point (ξ_1, ξ_2) of the system (1.1) is a saddle-node according to [1, Theorem 65 (par. 21)].

In the case that $\xi_1 = 0$, the system (1.1) has a unique equilibrium point $(0,0)$. We analogically aproximate the central manifold by the Taylor expansion with zero coefficients up to the second order (including) and get

$$a_3 = \frac{k_3}{(k_3 - 1)^2} > 0.$$

Consequently, the origin is topologically equivalent to a node according to [1, Theorem 65 (par. 21)].   □

Theorem 2.2. *The subset M_T of the parameter space K,*

$$M_T = \{(k_1, k_2, k_3) \in K : k_1 = -2\xi_1^3, k_2 = k_3 + 3\xi_1^2, k_3 \neq 1, \xi_1 \in \mathbb{R} - \{0\}\},$$

is a bifurcation set of codimension 1 - double equilibrium (also called "saddle-node bifurcation"). The double equilibrium point $(\xi_1, k_3\xi_1)$ is a saddle-node.

Proof. Let (ξ_1, ξ_2) be an equilibrium point of the system (1.1). The bifurcation "double equilibrium" occurres in the case that the parameters k_1, k_2, k_3 satisfy the following condition

$$3\xi_1^2 - k_2 + k_3 = 0. \tag{2.1}$$

In this case two equilibrium points coincide to one. So called non-degeneracy condition is $\xi_1 \neq 0$, because the equilibrium point is triple for $\xi_1 = 0$. Conditions (1.2) and (5) together with the non-degeneracy condition define the subset of K, where the system (1.1) has exactly two equilibrium points: the double equilibrium point $(\xi_1, k_3\xi_1)$ and the single equilibrium point $(-2\xi_1, -2k_3\xi_1)$. In the case that $k_3 = 1$, the double equilibrium point has two zero eigenvalues and bifurcation of codimension 2 takes place (this case is studied in Theorem 2.7).

The set M_T consists of two components M_{Tl} and M_{Tr}. They correspond with the case $\xi_1 < 0$ (the double equilibrium point lies left of the single one) and $\xi_1 > 0$ (the double equilibrium point lies right of the single one). These sets are symmetrical according to the axis $k_1 = 0$.

The closure of the set M_T divides the parameter space K into two sets M_1, M_3

$$M_1 = \{(k_1, k_2, k_3) \in K : k_1 = -2\xi_1^3, k_2 < k_3 + 3\xi_1^2, \xi_1 \in \mathbb{R}\},$$
$$M_3 = \{(k_1, k_2, k_3) \in K : k_1 = -2\xi_1^3, k_2 > k_3 + 3\xi_1^2, \xi_1 \in \mathbb{R}\}.$$

The set M_1 contsists of all the parameters from K, for which the system (1.1) has a unique equilibrium point (non-saddle), the set M_3 consists of those, for which the system (1.1) has 3 equilibrium points (non-saddle, saddle, non-saddle). While crossing the boundary M_T from the set M_3 to M_1, two equilibrium points coincide and disappear then. According to Lemma 2.1, the double equilibrium point is a saddle-node. A qualitative local change of the phase portraits occurres, a local bifurcation of codimension 1 - "saddle-node".   □
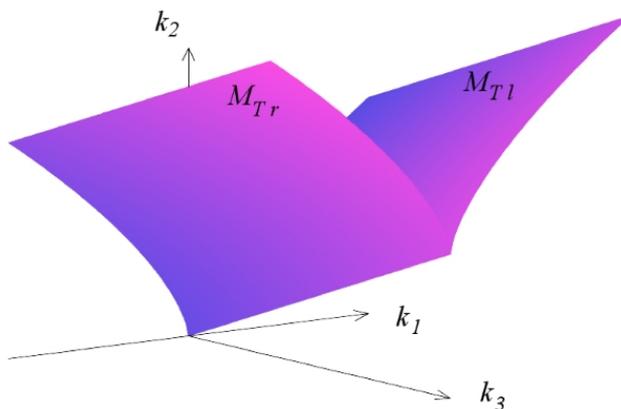
FIGURE 1. The set M_T.
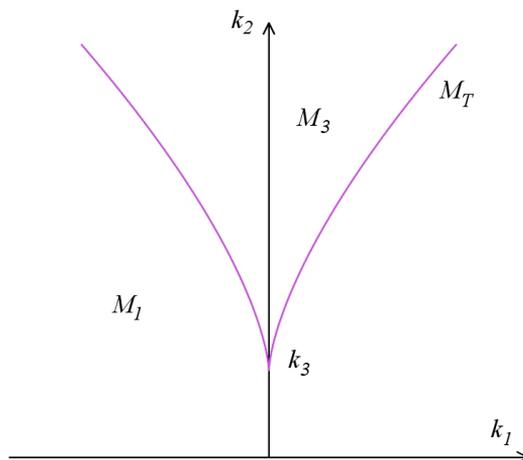
FIGURE 2. The section of M_T and the parameter space in k_3.

Theorem 2.3. *The subset M_H of the parameter space K,*

$$M_H = \{(k_1, k_2, k_3) \in K : k_1 = \xi_1(k_3 - 1 - 2\xi_1^2), k_2 = 1 + 3\xi_1^2, k_3 > 1, \xi_1 \in \mathbb{R}\},$$

is a bifurcation set corresponding with Andronov-Hopf bifurcation. The equilibrium point $(\xi_1, k_3\xi_1)$ is a multiple focus.

Proof. Let (ξ_1, ξ_2) be an equilibrium point of (1.1). The trace $\operatorname{tr} A = 0$ and the determinant $\det A > 0$ if and only if the Jacobi's matrix A has two purely imaginary

eigenvalues. We get the following conditions

$$k_1 + (k_2 - k_3)\xi_1 - \xi_1^3 = 0,$$
$$k_2 - 3\xi_1^2 - 1 = 0,$$
$$3\xi_1^2 - k_2 + k_3 > 0.$$

These three conditions define the set M_H (see fig. 3).
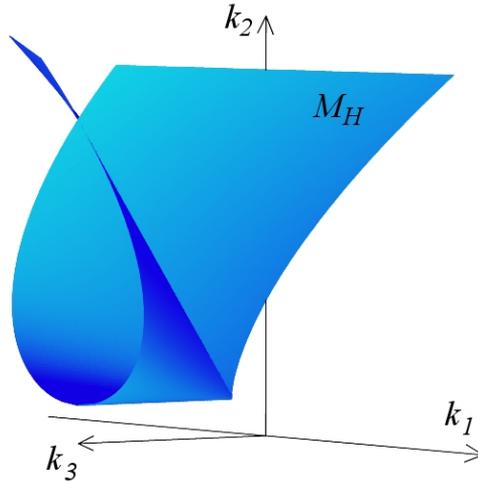
FIGURE 3. The set M_H.

The eigenvalues are purely imaginary on M_H,

$$\lambda_{1,2} = \pm i\omega, \quad \omega = \sqrt{\det A(\xi_1)} = \sqrt{k_3 - 1},$$

and the equilibrium point (ξ_1, ξ_2) is a multiple focus. While crossing the bound M_H, the equilibrium point may change its stability. We will compute the value of $\frac{d\operatorname{Re}\lambda_{1,2}}{dk_2}$ to describe the change of stability. Since

$$\frac{dp_A}{d\lambda} = 2\lambda - \operatorname{tr} A,$$

we have

$$\left.\frac{dp_A}{d\lambda}\right|_{M_H} = \pm i2\sqrt{k_3 - 1} \neq 0 \tag{2.2}$$

on the set M_H and we can apply the implicit function theorem and get

$$\left.\frac{d\lambda}{dk_2}\right|_{M_H} = -\frac{\frac{dp_A}{dk_2}}{\frac{dp_A}{d\lambda}}\bigg|_{M_H}. \tag{2.3}$$

The coordinates of the equilibrium point depend on the parameters. Let us denote $\xi_1 = \varphi(k_1, k_2, k_3)$. Then we get

$$\frac{dp_A}{dk_2} = -(\lambda + 1)\left(1 - 6\varphi\frac{\partial\varphi}{\partial k_2}\right). \tag{2.4}$$

Since the equality (1.2) gives

$$\varphi + (k_2 - k_3)\frac{\partial \varphi}{\partial k_2} - 3\varphi^2 \frac{\partial \varphi}{\partial k_2} = 0,$$

we can express the partial derivative $\frac{\partial \varphi}{\partial k_2}$ on the set M_H as

$$\frac{\partial \varphi}{\partial k_2} = \frac{\xi_1}{k_3 - 1}.$$

Using this expression, equalities (2.2) and (2.4) in (2.3), we get

$$\left.\frac{d\lambda}{dk_2}\right|_{M_H} = \frac{\left(1 - 6\frac{\xi_1^2}{k_3 - 1}\right)(\pm i \sqrt{k_3 - 1} + 1)}{\pm i \, 2\sqrt{k_3 - 1}}.$$

That yields

$$\left.\frac{d\,\mathrm{Re}\,\lambda}{dk_2}\right|_{M_H} = \frac{k_3 - 1 - 6\xi_1^2}{2(k_3 - 1)} = \frac{k_3 + 1 - 2k_2}{2(k_3 - 1)}. \tag{2.5}$$

Taking M_H as a parametric function of ξ_1, we have

$$\frac{dk_1}{d\xi_1} = k_3 - 1 - 6\xi_1^2.$$

The derivative $\frac{d\,\mathrm{Re}\,\lambda}{dk_2}$ is zero if and only if $\frac{dk_1}{d\xi_1} = 0$, that is in the case that the tangent to M_H is parallel to the axis k_2. In this situation, there is no crossing of M_H (just a contact) and there is also no change in stability of the focus. In the case that $k_3 + 1 > 2k_2$, a stable focus changes to an unstable focus, while crossing M_H in the direction of the axis k_2. In the opposite case, an unstable focus changes to a stable focus. (These results correspond to Theorem 2.5 on subcritical and supercritical bifurcation.)

While crossing the bifurcation bound M_H, the focus changes its stability and a limit cycle arises in its neighbourhood. There occurres a local qualitative change of the phase portraits called Andronov-Hopf bifurcation.

The set M_H is divided by the set M_T into three parts M_{Hr}, M_{Hl} and M_{Hu} (see fig. 4).

These sets correspond with Andronov-Hopf bifurcation of the right, left (in the case of three equilibrium points) and unique equilibrium point. $\square$

Remark 2.4. Stability of the limit cycle depends on stability of the multiple focus and is determined by the sign of the first Lyapunov number of this multiple focus. The cycle is stable for $l_1 < 0$ and unstable for $l_1 > 0$. Parameters corresponding with zero values of the first Lyapunov number l_1 determine a subset of codimension 2 of M_H - degenerate Andronov-Hopf bifurcation.

Theorem 2.5. *The subset M_{DH} of the parameter space K*

$$M_{DH} = \{(k_1, k_2, k_3) \in K : k_1 = 4\xi_1^3, k_2 = 1 + 3\xi_1^2, k_3 = 1 + 6\xi_1^2, \xi_1 \in \mathbb{R} - \{0\}\}$$

is a bifurcation set of codimension 2 corresponding with degenerate Andronov-Hopf bifurcation.

Proof. Let (ξ_1, ξ_2) be an equilibrium point of the system (1.1). We transform the system (1.1) by a substitution $u = x - \xi_1$, $v = k_1 + k_2 x - x^3 - y$ to an equivalent
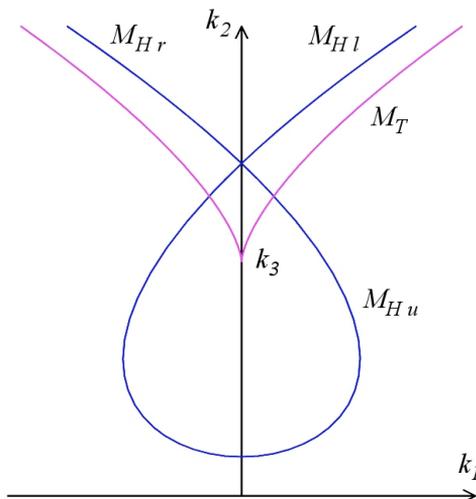
FIGURE 4. The section of M_T and M_H in $k_3 > 1$.

system of Lienard's type

$$\dot{u} = v,$$
$$\dot{v} = p(u) + q(u)v \equiv p_0 + p_1 u + p_2 u^2 + p_3 u^3 + (q_0 + q_1 u + q_2 u^2)v, \tag{2.6}$$

where

$$p_0 = k_1 + (k_2 - k_3)\xi_1 - \xi_1^3, \quad p_1 = k_2 - k_3 - 3\xi_1^2, \quad p_2 = -3\xi_1, \quad p_3 = -1,$$
$$q_0 = -1 + k_2 - 3\xi_1^2, \quad q_1 = -6\xi_1, \quad q_2 = -3. \tag{2.7}$$

Since (1.2) holds for the equilibrium point (ξ_1, ξ_2), we have $p_0 = 0$, system (10) has an equilibrium point at the origin. The origin is a multiple focus if and only if $p_1 < 0$ and $q_0 = 0$. According to [8] or [3], we can express the first and the second Lyapunov numbers as

$$l_1 = p_2 q_1 - p_1 q_2, \quad l_2 = -p_3 q_2.$$

Consequently from (2.7)

$$l_1 = 3(k_2 - k_3 + 3\xi_1^2), \quad l_2 = -3.$$

Since $\operatorname{tr} A = 0$ on M_H, we get

$$l_1 = 3(1 - k_3 + 6\xi_1^2) = 3(2k_2 - k_3 - 1), \quad l_2 = -3 \neq 0.$$

The condition $l_1 = 0$ determines the subset M_{DH} on M_H (see fig. 5) that corresponds with the degenerate Andronov-Hopf bifurcation of codimension 2 (since $l_2 \neq 0$). The curve M_{DH} divides the surface M_H into parts M_{DH-} corresponding with the supercritical bifurcation ($l_1 < 0$, a stable limit cycle occurres) and M_{DH+} corresponding with the subcritical bifurcation ($l_1 > 0$, an unstable limit cycle occurres). □
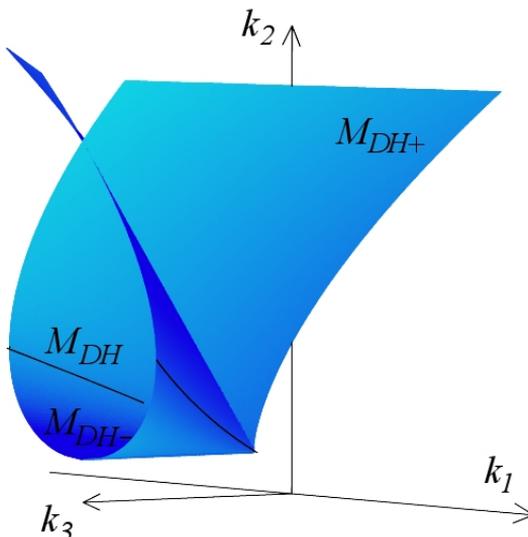
FIGURE 5. The set M_{DH}.

Remark 2.6. The set M_{DH-} is entirely contained in the set M_{Hu}, which imply that the stable limit cycle (caused by Andronov-Hopf bifurcation) may occur only in the case of the unique equilibrium point.

Theorem 2.7. *The subset M_{BT} of the parameter space K,*

$$M_{BT} = \{(k_1, k_2, k_3) \in K : k_1 = -2\xi_1^3, k_2 = 1 + 3\xi_1^2, k_3 = 1, \xi_1 \in \mathbb{R} - \{0\}\},$$

is a bifurcation set of codimension 2 corresponding with Bogdanov-Takens bifurcation.

Proof. Let (ξ_1, ξ_2) be an equilibrium point of the system (1.1). The bifurcation set of codimension 2 corresponding with Bogdanov-Takens bifurcation includes such parameters from K that both eigenvalues of Jacobi's matrix A are zero. The set M_{BT} is determined by two conditions $\det A = 0$ and $\operatorname{tr} A = 0$. The set M_{BT} lies in the intersection of the closure of M_H and the set M_T. In the case $\xi_1 = 0$, that is for $k_1 = 0, k_2 = k_3 = 1$, bifurcation of higher codimension occurres. Further analysis of this bifurcation is presented in Theorem 3.2. □

Theorem 2.8. *The subset M_C of the parameter space K,*

$$M_C = \{(k_1, k_2, k_3) \in K : k_1 = 0, k_2 = k_3, k_3 \neq 1\},$$

is a bifurcation set of codimension 2 - triple equilibrium point. The unique equilibrium point $(0,0)$ of (1.1) is topologically equivalent to a stable node for $k_3 < 1$, or an unstable node surrounded by a stable limit cycle for $k_3 > 1$.

Proof. The Jacobi's matrix on M_C is

$$A = \begin{pmatrix} k_2 & -1 \\ k_2 & -1 \end{pmatrix}$$

and its eigenvalues are $\lambda_1 = 0$ and $\lambda_2 = k_2 - 1$. The origin is the unique equilibrium point of (1.1) and it is stable for $k_3 < 1$, unstable for $k_3 > 1$. The unstable unique equilibrium is surrounded by a stable limit cycle according to Lemma 1.1 on existence of a globally attractive set and the Poincaré's theorem. The origin is topologically equivalent to a node according to Lemma 2.1.   □

3. Non-local bifurcations

In contradiction to local bifurcations, where the bifurcation sets could be expressed explicitly, bifurcation sets corresponding with non-local bifurcations can only be studied numerically or can be approximated with accuracy to a particular order in the neighbourhood of some important bifurcation points.

Non-local bifurcation of codimension 1 - multiple cycle. The curve M_{DH} is a boundary of a surface M_D corresponding with non-local bifurcation of codimension 1 - multiple cycle. While crossing the set M_D, two limit cycles (stable and unstable) merge into one semi-stable cycle that disappears then. Closures of sets M_D and M_H are tangent to each other in each point of the curve M_{DH}. The following schematic figure 6 shows the lay-out of the sets M_H, M_T and M_D only. They are figured by their intersections with the plane $k_3 = \text{const.} > 1$. The numerical computations shows, that these sets lie closely to each other and there are technical problems with their rendering on the same scale.
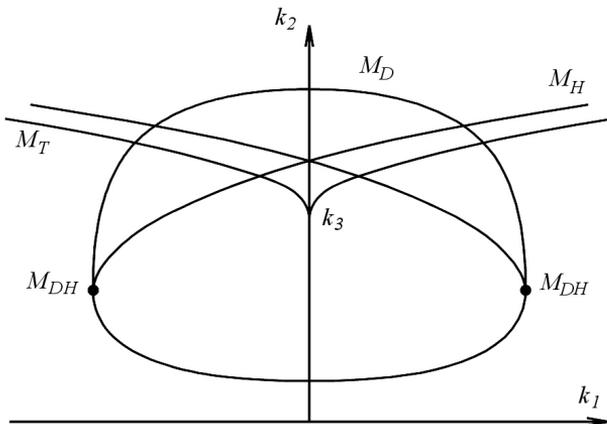
FIGURE 6. The section of M_{DH} in $k_3 > 1$.

Non-local bifurcation of codimension 1 - separatrix loop. The curve M_{BT} is a boundary of the surface M_L corresponding with non-local bifurcation of codimension 1 - separatrix loop. The surface M_L is tangent to M_T and M_H at each point of M_{BT}. The set M_L is contained in the half-space $k_3 > 1$ and consists of two components M_{Lr} and M_{Ll} corresponding with existence of the separatrix loop surrounding the right or the left equilibrium point respectively. While crossing the bound M_L, the unstable limit cycle (originated near M_H in consequence of the subcritical Andronov-Hopf bifurcation) merge into the separatrix loop and splits.

Let (ξ_1, ξ_2) be the right double equilibrium point of the system (1.1). Then the parameters of the system (1.1) lie in the set M_{BT} (Bogdanov-Takens bifurcation) and the coordinates of the double equilibrium point satisfy

$$\xi_1 = \sqrt{\frac{k_2 - 1}{3}}, \quad \xi_2 = k_3\sqrt{\frac{k_2 - 1}{3}}$$

according to Theorem 2.7. Using the following substitution

$$x = x_1 - \sqrt{\frac{k_2 - 1}{3}}, \quad y = x_2 - k_3\sqrt{\frac{k_2 - 1}{3}},$$

we transform the system (1.1) into a system

$$\dot{x} = k_1 + \sqrt{\frac{k_2 - 1}{3}}\left(k_2 - k_3 - \frac{k_2 - 1}{3}\right) + x - \sqrt{3(k_2 - 1)}x^2 - x^3 - y, \quad (3.1)$$
$$\dot{y} = k_3 x - y.$$

The origin is a double equilibrium point of the system (12) with two zero eigenvalues for parameters from M_{BT}.

System (3.1) can be transformed by the linear transformation $x_1 = y$, $x_2 = k_3 x - y$ into the system

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = h_{00} + h_{10}x_1 + \frac{1}{2}h_{20}x_1^2 + h_{11}x_1 x_2 + \frac{1}{2}h_{02}x_2^2 + R(x_1, x_2, k_1, k_2, k_3), \quad (3.2)$$

where

$$h_{00} = k_3\left(k_1 + \sqrt{\frac{k_2 - 1}{3}}\left(k_2 - k_3 - \frac{k_2 - 1}{3}\right)\right), \quad h_{10} = 1 - k_3,$$
$$h_{20} = -\frac{2}{k_3}\sqrt{3(k_2 - 1)}, \quad h_{11} = -\frac{2}{k_3}\sqrt{3(k_2 - 1)},$$
$$h_{02} = -\frac{2}{k_3}\sqrt{3(k_2 - 1)}, \quad R(x_1, x_2, k_1, k_2, k_3) = -\frac{(x_1 + x_2)^3}{k_3^2}.$$

This transformation keeps the equilibrium point at the origin as well as its zero eigenvalues. In the further analysis, we will study system (3.2) instead of the equivalent system (1.1).

Remark 3.1. For $(k_1, k_2, k_3) \in M_{BT}$, the following statements hold

$$h_{00} = 0, \quad h_{10} = 0, \quad h_{11} = h_{20} = h_{02} \neq 0.$$

Theorem 3.2. *The system* (3.2) *can be transformed by a smooth non-degenerate change of parameters to the Bogdanov-Takens normal canonical form*

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = \beta_1 + \beta_2 x_1 + x_1^2 + x_1 x_2 + O(\|x\|^3), \quad (3.3)$$

where

$$\beta_1 = \frac{h_{11}}{(-h_{10} + \frac{1}{4}h_{02}h_{00} + \frac{1}{2})^3}h_{00},$$
$$\beta_2 = \frac{1}{(-h_{10} + \frac{1}{4}h_{02}h_{00} + \frac{1}{2})^2}(h_{10} - h_{00}h_{02}). \quad (3.4)$$

In the neighbourhood of the Bogdanov-Takens curve M_{BT} corresponding with the right double equilibrium point, the set M_{Lr} can be expressed at the form

$$M_{Lr} = \left\{ (k_1, k_2, k_3) \in \mathbb{R}^3 : \beta_2 < 0,\ \beta_1 = -\frac{6}{25}\beta_2^2 + o(\beta_2^2) \right\}. \tag{3.5}$$

The set M_{Ll} is symmetrical to M_{Lr} according to the plane $k_1 = 0$.

Proof. The change of time $dt = (1 - \frac{h_{02}}{2}x_1)d\tau$ and the substitution

$$u_1 = x_1, \quad u_2 = x_2 - \frac{h_{02}}{2}x_1 x_2$$

eliminates the term with x_2^2. We get a system of the form

$$\dot{u}_1 = u_2,$$
$$\dot{u}_2 = \nu_1 + \nu_2 u_1 + C_1 u_1^2 + C_2 u_1 u_2 + O(\|u\|^3),$$

where

$$\nu_1 = h_{00}, \quad \nu_2 = h_{10} - h_{00}h_{02}, \quad C_1 = -h_{02}h_{10} + \frac{1}{4}h_{02}^2 h_{00} + \frac{1}{2}h_{20}, \quad C_2 = h_{11}.$$

Note that $C_1 = \frac{1}{2}h_{20} \neq 0$ on M_{BT} according to Remark 3.1. Introducing a new time (denoted again with t)

$$t = \left| \frac{C_2}{C_1} \right| \tau$$

and new variables (denoted again with x_1 and x_2)

$$x_1 = \frac{C_2^2}{C_1}u_1, \quad x_2 = \mathrm{sgn}\left(\frac{C_2}{C_1}\right)\frac{C_2^3}{C_1^2}u_2,$$

we get the Bogdanov-Takens normal canonical form (3.3), where

$$\beta_1 = \frac{h_{11}^4}{(-h_{02}h_{10} + \frac{1}{4}h_{02}^2 h_{00} + \frac{1}{2}h_{20})^3} h_{00},$$

$$\beta_2 = \frac{h_{11}^2}{(-h_{02}h_{10} + \frac{1}{4}h_{02}^2 h_{00} + \frac{1}{2}h_{20})^2}(h_{10} - h_{00}h_{02}).$$

With respect to the fact that $h_{20} = h_{11} = h_{02}$, we get the expressions (3.4).

The coefficient of the term with $x_1 x_2$ corresponds to

$$s = \mathrm{sgn}\left(\frac{C_2}{C_1}\right)\Big|_{M_{BT}} = \mathrm{sgn}\left(\frac{h_{11}}{-h_{02}h_{10} + \frac{1}{4}h_{02}^2 h_{00} + \frac{1}{2}h_{20}}\right)\Big|_{M_{BT}}.$$

According to Remark 3.1, we have $s = \mathrm{sgn}\, 2 = 1$. The Bogdanov-Takens bifurcation is non-degenerate, since

$$h_{11} = -2\sqrt{3(k_2 - 1)} = -6\xi_1 \neq 0$$

and $h_{20} \neq 0$ on M_{BT}. The change of parameters is invertible in the neighbourhood of the origin. It can be verified by a direct computation of the following determinants and finding

$$\begin{vmatrix} \frac{\partial\beta_1}{\partial k_1} & \frac{\partial\beta_1}{\partial k_2} \\ \frac{\partial\beta_2}{\partial k_1} & \frac{\partial\beta_2}{\partial k_2} \end{vmatrix} \neq 0, \quad \begin{vmatrix} \frac{\partial\beta_1}{\partial k_2} & \frac{\partial\beta_1}{\partial k_3} \\ \frac{\partial\beta_2}{\partial k_2} & \frac{\partial\beta_2}{\partial k_3} \end{vmatrix} \neq 0, \quad \begin{vmatrix} \frac{\partial\beta_1}{\partial k_3} & \frac{\partial\beta_1}{\partial k_1} \\ \frac{\partial\beta_2}{\partial k_3} & \frac{\partial\beta_2}{\partial k_1} \end{vmatrix} \neq 0.$$

This fact implies that the change of parameters cause no degeneration of the bifurcation manifold according to the parameter space. (In the bifurcation theory this regularity of the parameter transformation is called the transversality condition.)

The expression for the set M_L can be found in [6, Theorem 8.5, Appendix] or in [4]. The set M_{Ll} has to be symmetric to M_{Lr} about to the plane $k_1 = 0$.   □

Non-local bifurcation of codimension 2 - two separatrix loops. The curve M_{LL}, which is an intersection of the sets M_{Lr} and M_{Ll} and lies in the plane $k_1 = 0$ (because of the symmetry of the parameter portrait) corresponds with the non-local bifurcation of codimension 2 - two separatrix loops. Two separatrix loops surround both the right and the left equilibrium points (see fig. 7).
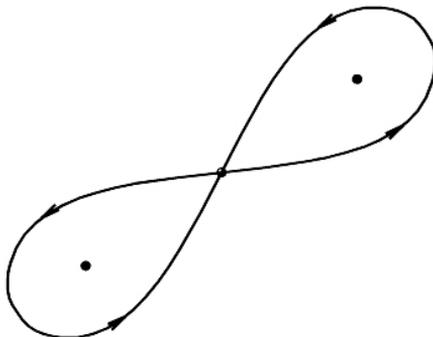
FIGURE 7. Structurally unstable two separatrix loops.

Non-local bifurcation of codimension 1 - "big separatrix loop". According to [7], the curve M_{LL} is a boundary of a bifurcation set M_{BL} corresponding with non-local bifurcation of codimension 1 - „big separatrix loop". While crossing the set M_{BL}, separatrix loop surrounding both equilibrium points appears and consequently gives to arise to an unstable limit cycle containing the saddle and both remaining equilibrium points in its interior (see fig. 8).
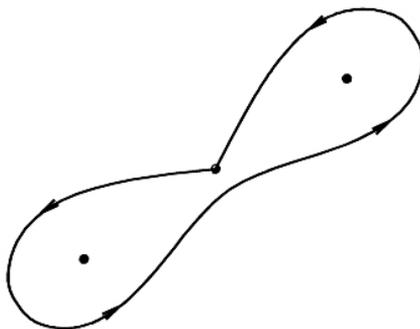
FIGURE 8. A structurally unstable big separatrix loop.

Figure 9 presents the lay-out of the sets M_T, M_H, M_L and M_{BL}, showing the section of the parameter space K by the plane $k_3 = const. > 1$, near 1.
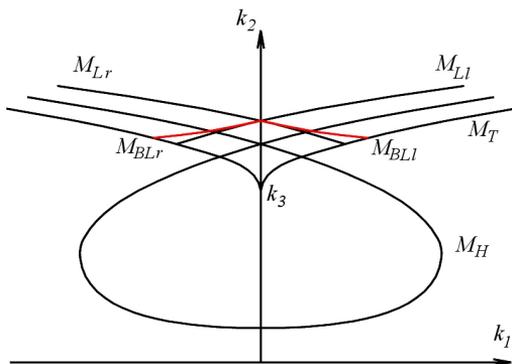
FIGURE 9. The section of M_L and M_{BL} in $k_3 > 1$.

4. GLOBAL BIFURCATION DIAGRAM

The bifurcation sets described above divide the parameter space K into parts, where the phase portraits of system (1.1) are topologically equivalent and structurally stable. The bifurcation sets contain those parameters, for which the phase portraits are structurally unstable.

Figure 10 shows a section of the global bifurcation diagram by the plane $k_3 =$const. for $k_3 \in (0, 1]$, and figure 11 this section for $k_3 > 1$. Figure 12 shows the structurally stable phase portraits corresponding to the marked regions for $k_1 < 0$. The half-space $k_1 > 0$ is symmetrical to the opposite one and the phase portraits are symmetrical according to the origin.
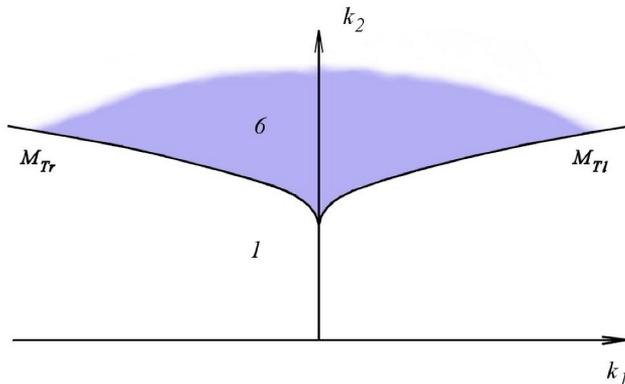
FIGURE 10. The section of the bifurcation diagram in $k_3 \in (0, 1]$.

FIGURE 11. The section of the bifurcation diagram in $k_3 > 1$.

FIGURE 12. Phase portraits corresponding to the marked regions.

References

[1] Andronov A.A., Leontovich E.A., Gordon I.I., Mayer A.G.: *Kachestvennaya teoriya dinamicheskikh sistem vtorovo poryadka*, Nauka, Moskva, 1966.

[2] Baráková L.: *Asymptotic Behaviour and Existence of a Limit Cycle in Cubic Autonomous Systems*, Demonstratio Mathematica Vol. 34 No. 3 (2001), 559-576.

[3] Bautin N.N.: *Povedeniye dinamicheskikh sistem vblizi granic oblasti ustoychivosti*, Ogiz Gostexizdat, Leningrad-Moskva, 1949.

[4] Bogdanov R.I.: *Versalnaya deformatsiya osoboy tochki vektornovo polya na ploskosti v sluchae nulevikh sobstvennikh chisel*, Tr. sem. im. Petrovskovo, No. 2 (1976), 37-65.

[5] Kaldor N.: *A Model of the Trade Cycle*, Econ. Jour. 50 (1940), 78-92.

[6] Kuznetsov Y.A., *Elements of Applied Bifucation Theory*, Second Edition, Applied Mathematical Sciences 112, Berlin, Heidelgerg, New York, Springer-Verlag, 1995, 1998.

[7] Turaev D.V.: ,,*Bifurkatsii dvumernikh dinamicheskikh sistem, blizkich k sisteme s dvumya petlyami separatris*", Uspechi mat. nauk Vol. 40, No. 6 (1985), 203-204.

[8] Volokitin E.P., Treskov S.A.: *Bifurkatsionnaya diagramma kubicheskoy sistemi Lienardovskovo tipa*, Sibirskij zhurnal industrialnoy matematiki Vol. 5, No. 3(11) (2002), 67-75.

Lenka Baráková

Mendel University, Dept. of Math., Zemedelska 1, 613 00 Brno, Czech Rep.
   E-mail address: barakova@mendelu.cz

Evgenii P. Volokitin

Sobolev Institute of Mathematics, Novosibirsk, 630090, Russia
   E-mail address: volok@math.nsc.ru

BIFURCATION ROUTES TO CHAOS IN AN EXTENDED VAN DER POL'S EQUATION APPLIED TO ECONOMIC MODELS

LENKA PŘIBYLOVÁ

ABSTRACT. In this paper a 3-dimensional system of autonomous differential equations is studied. It can be interpreted as an idealized macroeconomic model with foreign capital investment introduced in [9] or an idealized model of the firm profit introduced in [3]. The system has three endogenous variables with only one non-linear term and can be also interpreted as an extended van der Pol's equation. It's shown that this simple system covers several types of bifurcations: both supercritical and subcritical Hopf bifurcation and generalized Hopf bifurcation as well, the limit cycle exhibits period-doubling bifurcation as a route to chaos. Some results are analytical and those connected with chaotic motion are computed numerically with continuation programs Content, Xppaut and Maple. We present conditions for stability of the cycles, hysteresis, explore period doubling and using Poincaré mapping show a three period cycle that implies chaos.

1. INTRODUCTION

In this paper we consider the autonomous system of three differential equations

$$\dot{x} = ay + px(\kappa - y^2),$$
$$\dot{y} = v(x + z), \tag{1.1}$$
$$\dot{z} = mx - ry,$$

where x, y, z are real endogenous variables and a, p, v, m and r real exogenous parameters. This system may be interpreted as an extension of the van der Pol's equation

$$\dot{x} = ay + px(\kappa - y^2),$$
$$\dot{y} = vx. \tag{1.2}$$

1.1. **Two economic models.** Vosvrda [9] introduced an idealized macroeconomic model with foreign capital investment in the form

$$\dot{S} = aY + pS(\kappa - Y^2),$$
$$\dot{Y} = v(S + F), \tag{1.3}$$
$$\dot{F} = mS - rY,$$

where $S(t)$ are savings of households, $Y(t)$ is the Gross Domestic Product, $F(t)$ is the foreign capital inflow, t is the time and dot denotes the derivative with respect to t. Positive parameters represent corresponding ratios: a is the variation of the marginal propensity to savings, p is the ratio of the capitalized profit, v is the output/capital ratio, κ is the potential GDP (it can be set to 1, as a unit of GDP — Y, S, F then represent the percentage part of the potential GDP), m is the capital inflow/savings ratio and r is is the debt refund/output ratio. From the economic point of view, the condition

$$\frac{a}{vr} > 1. \tag{1.4}$$

guarantees the ability of the economy to refund the debt. Another — stronger condition — can be $a > r$ ($v \in (0,1)$, normally much lower then 1, see [4]), which implies the condition (1.4). The condition $a > r$ together with (1.4) was used in [9]. We will see that the equilibria are stable for $a > r$, although orbitally stable and even unstable cycles can occur in this economically "normal" case. In the section 8 we let some of the parameters cross the zero axis and accept their negativity to explain behaviour of trajectories in the positive neighbourhood of zero.

¿From the mathematical point of view, S. Bouali [3] introduced the same system as an idealized economic model of firm profit in the form

$$\dot{R} = aP + pR(\kappa - P^2),$$
$$\dot{P} = v(R + F), \tag{1.5}$$
$$\dot{F} = mR - rP,$$

where P is a firm profit, R are reinvestments and F represent debts, coefficients a, p, v, m and r are corresponding rates or proportions.

2. Equilibria and its stability

The system (1.1) is antisymmetric. If $(x(t), y(t), z(t))$ is a solution of (1.1), than $(-x(t), -y(t), -z(t))$ is also its solution. Solving the system

$$0 = ay + px(\kappa - y^2),$$
$$0 = v(x + z), \tag{2.1}$$
$$0 = mx - ry,$$

we find, that the system (1.1) has three equilibria: $E_0 = (x_0, y_0, z_0) = (0, 0, 0)$,

$$E_1 = (x_1, y_1, z_1) = \left(-\sqrt{\frac{amr + p\kappa r^2}{pm^2}}, \sqrt{\frac{amr + p\kappa r^2}{pm^2}}, \sqrt{\frac{am + p\kappa r}{pr}}\right),$$

$$E_2 = (x_2, y_2, z_2) = \left(\sqrt{\frac{amr + p\kappa r^2}{pm^2}}, -\sqrt{\frac{amr + p\kappa r^2}{pm^2}}, -\sqrt{\frac{am + p\kappa r}{pr}}\right).$$

We will study the equilibria E_0 and E_1, since results for E_2 are analogous to the antisymmetric equilibrium E_1. Since the Jacobian matrix has the form

$$J = \begin{pmatrix} p(\kappa - y^2) & a - 2pxy & 0 \\ v & 0 & v \\ m & -r & 0 \end{pmatrix}, \tag{2.2}$$

we can get the corresponding eigenvalues by solving cubic equations. These formulas are very complicated, but we can use them for numerical examples. On the other hand, we can get information about stability of both equilibria right from the characteristic polynomial

$$p(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3, \tag{2.3}$$

where $a_1 = p(y^2 - \kappa)$, $a_2 = v(r - a + 2pxy)$ and $a_3 = v(pry^2 + 2mpxy - p\kappa r - am)$:
1. Equilibrium $E_0 = (0, 0, 0)$: Coefficients of the characteristic polynomial are

$$a_1 = -p\kappa < 0, \; a_2 = v(r - a), \; a_3 = -v(rp\kappa + am) < 0. \tag{2.4}$$

From the Hurwitz criterion it yields that in the case $a \geq r$, the characteristic polynomial has one positive root. In the case that $a < r$, the polynomial has three or one positive root. The trivial equilibrium E_0 has at least one positive eigenvalue, so it can never be stable.
2. Equilibrium $E_{1,2}$: Coefficients of the characteristic polynomial are

$$a_1 = \frac{am}{r} > 0, \; a_2 = v(r + a + \frac{2rp\kappa}{m}) > 0, \; a_3 = 2v(rp\kappa + am) > 0. \tag{2.5}$$

Lemma 2.1. *Characteristic polynomial is Hurwitzian (that is all eigenvalues have negative real parts) if and only if $a > r$.*

Proof. The characteristic polynomial is Hurwitzian if and only if the determinant

$$D_3 = \begin{vmatrix} a_1 & 1 & 0 \\ a_3 & a_2 & a_1 \\ 0 & 0 & a_3 \end{vmatrix}$$

and all its main subdeterminants are positive. The subdeterminant $D_1 = a_1 = am/r > 0$, subdeterminant $D_2 = \begin{vmatrix} a_1 & 1 \\ a_3 & a_2 \end{vmatrix} = a_1a_2 - a_3$ and the determinant $D_3 = a_3(a_1a_2 - a_3)$ are positive if and only if

$$a_1a_2 - a_3 = \frac{am}{r}v(r + a + \frac{2rp\kappa}{m}) - 2v(rp\kappa + am)$$
$$= (a - r)\Big(\frac{amv}{r} + 2vp\kappa\Big) > 0,$$

that is if and only if $a > r$. □

Both the equilibria cannot change its stability on the real axis of the eigenvalues plane, since the characteristic polynomial cannot have positive real root. It is possible, if the eigenvalues crossed the imaginary axis. This type of bifurcation is called Hopf, and it will be discussed in the next section.

3. Hopf bifurcation and stability of limit cycles

If we look for the Hopf bifurcation, we have to find two purely imaginary eigenvalues $\lambda_{1,2} = i\omega$ of the characteristic polynomial. Denote the third real eigenvalue λ_3. Since substituting into (2.3) gives

$$(\lambda - i\omega)(\lambda + i\omega)(\lambda - \lambda_3) = \lambda^3 - \lambda^2\lambda_3 + \lambda\omega^2 - \omega^2\lambda_3,$$

the necessary condition for Hopf bifurcation is

$$a_1a_2 = a_3. \tag{3.1}$$

Solving a system of five equations (2.1), (3.1) with substitution from (2.3) and $a_2 = \omega^2$ according to three variables x, y, z and two parameters a, m, we get two solutions. For the trivial equilibrium E_0, we necessarily have

$$m = -p\kappa. \tag{3.2}$$

We can exclude this case since all parameters are positive for now (we will return to this bifurcation later). The next solution is equilibrium

$$E_{1,2} = \left(\pm\frac{r}{m}\sqrt{\frac{m+p\kappa}{p}}, \pm\sqrt{\frac{m+p\kappa}{p}}, \mp\frac{r}{m}\sqrt{\frac{m+p\kappa}{p}} \right) \tag{3.3}$$

with parameters satisfying $a = r$ and $m = \frac{2vrp\kappa}{\omega^2 - 2vr}$, which gives

$$\omega = \sqrt{\frac{2vr(p\kappa + m)}{m}}. \tag{3.4}$$

From the previous results we know that the third real eigenvalue has to be negative and we get its value $\lambda_3 = -m$.

We will show that the Hopf bifurcation appears for parameter a while crossing the critical value r. Using the implicit function theorem we can compute the derivative of the complex eigenvalue λ with respect to a for the equilibrium E_1 (using (2.5)):

$$\frac{d\lambda}{da} = -\frac{\frac{dp(\lambda)}{da}}{\frac{dp(\lambda)}{d\lambda}} = -\frac{\lambda^2\frac{m}{r} + \lambda v + 2vm}{3\lambda^2 + 2\lambda\frac{am}{r} + v(r+a) + \frac{2vrp\kappa}{m}}. \tag{3.5}$$

Substituting $a = r$ and $\lambda_{1,2} = \pm i\omega$ into (3.5), we evaluate

$$\mathrm{Re}\,\frac{d\lambda}{da}\big|_{a=r} = -\frac{(2mv - \frac{m\omega^2}{r})(-3\omega^2 + 2vr + \frac{2vrp\kappa}{m}) + 2v\omega^2 m}{(-3\omega^2 + 2vr + \frac{2vrp\kappa}{m})^2 + 4\omega^2 m^2} \tag{3.6}$$

The denominator of this fraction is positive, substituting (3.4) into the nominator we have

$$\mathrm{sgn}\,\mathrm{Re}\,\frac{d\lambda}{da}\big|_{a=r} = \mathrm{sgn}\,\frac{-4rv^2(m+p\kappa)(m+2p\kappa)}{m} < 0. \tag{3.7}$$

The transversality condition for Hopf bifurcation is fulfilled and according to [6] the Hopf bifurcation gives rise to a limit cycle near the equilibria at the critical value $a = r$.

According to the Lemma (2.1), the equilibrium $E_{1,2}$ change its unstability to stability as the parameter a grows and crosses the critical value r. Locally we can educe the same result from (3.7). Clearly, the following theorem holds:

Theorem 3.1. *Equilibria $E_{1,2}$ of the system* (1.1) *are stable for $a > r$ and lose their stability if a parameter crosses the Hopf bifurcation manifold $a = r$.*

Note that the condition (1.4) is satisfied on the Hopf bifurcation hyperplane, unlike $a > r$ used in [9].

In contrast to [9], we can prove that the Hopf bifurcation (in both previous cases) can give rise to both the stable and the unstable cycles. Following the projection method for center manifold computation (see [6]), we can get formula for the first Lyapunov coefficient l_1. Negative sign of this coefficient implies stability, positive sign unstability of the arising limit cycle near the equilibrium. Symbolically this formula is very complicated, so we computed the Hopf bifurcation curves numerically using the continuation program Content.

The bifurcation border GH of the generalized Hopf bifurcation ($l_1 = 0$) has typical form presented on figures 1, 2 and 3. The parameter κ can be set to 1 as it was explained earlier, parameters v, p, m correspond to ratios and the bifurcation diagrams are similar for all parameter values in the whole interval $\langle 0, 1 \rangle$ (here and below, the concrete values of the parameters were chosen arbitrarily, but to follow examples in [9], where misleading results were made).
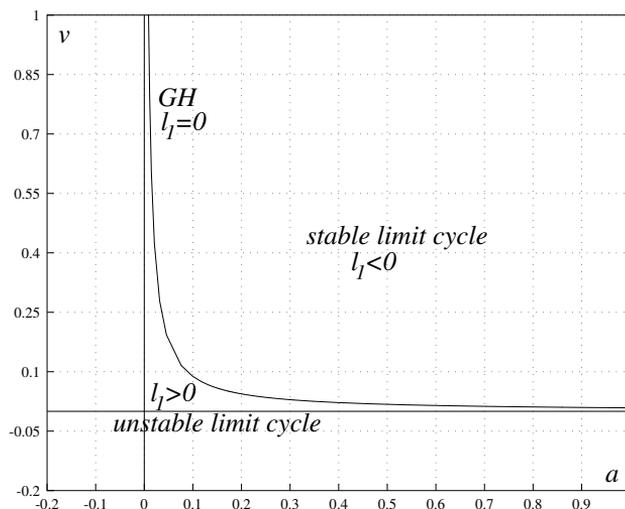
FIGURE 1. $a = r$, $p = 0.1$, $\kappa = 1$, $m = 0.19$.
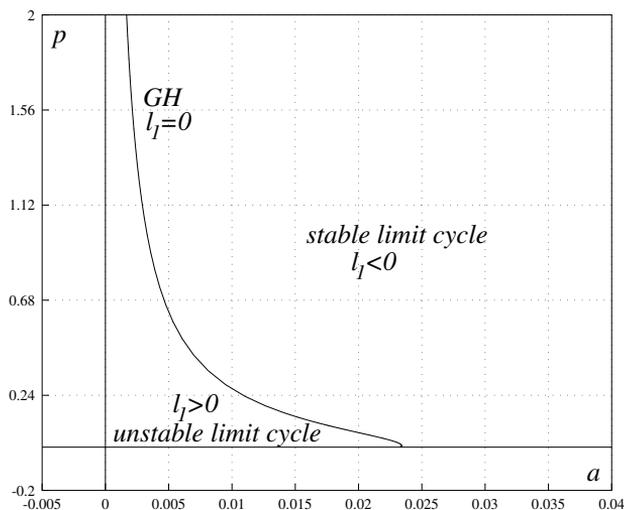
FIGURE 2. $a = r$, $v = 0.5$, $\kappa = 1$, $m = 0.19$.
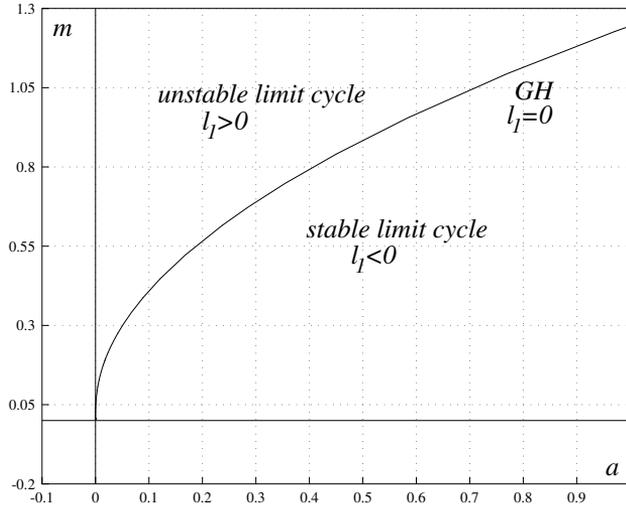
FIGURE 3. $a = r$, $v = 0.5$, $\kappa = 1$, $p = 0.1$.

For illustrating the diagrams, we present typical diagram 4 of the Hopf bifurcation in the plane a, r.
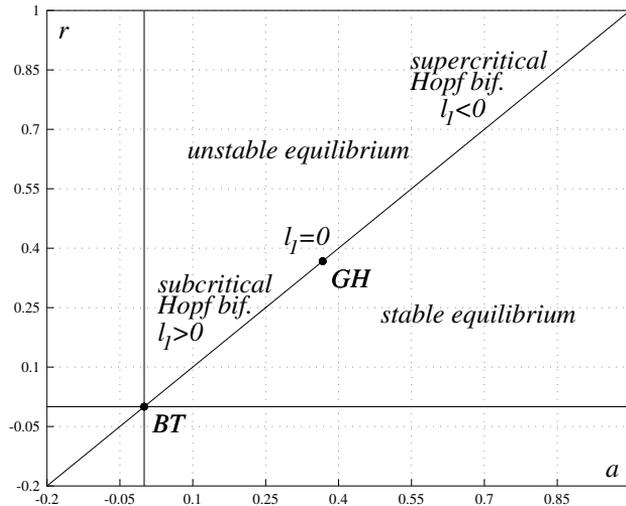
FIGURE 4. The typical Hopf bifurcation curve.

4. FOLDING OF CYCLES AND PERIOD DOUBLING

In this section typical bifurcation diagrams will be presented. The parameters in examples are chosen to correspond the previous bifurcation diagrams 1, 2, 3 and

4. By a convention, solid lines represent stable equilibria, dashed lines are unstable equilibria, solid circles correspond to orbitally stable cycles and empty circles to the unstable ones.

Example 1. Let us focus on the example with parameters $r = 0.25$, $p = 0.1$, $v = 0.5$, $\kappa = 1$ and $m = 0.19$. From (3.7) it yields that $\operatorname{sgn} \frac{d \operatorname{Re} \lambda}{da} \doteq -0.01$ and as parameter a grows and cross $r = 0.25$, $\operatorname{Re} \lambda$ decreases very slowly. Since the Lyapunov coefficient is $l_1 \doteq -0.0318$, it give rise to an orbitally stable cycle from the stable equilibrium. The branches of the periodic solutions are presented on the figure 5.
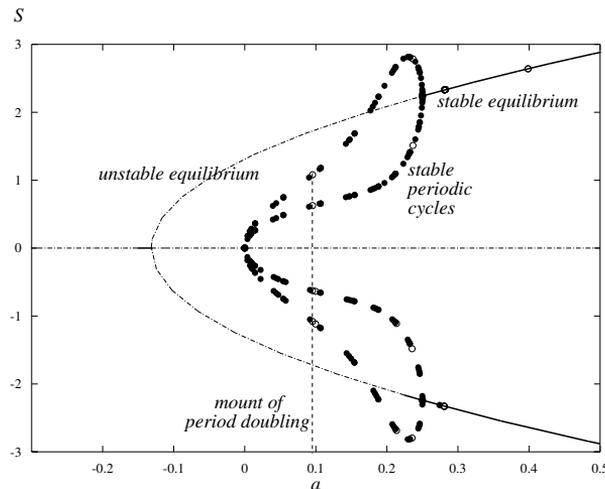
FIGURE 5. The typical Hopf bifurcation curve with arising stable cycle and period doubling of the cycle for $a \doteq 0.09537$.

Example 2. On the other hand, for parameters $r = 0.25$, $p = 0.01$, $v = 0.031847$, $\kappa = 1$ and $m = 0.19$ (this example was already shown in [9], example 3, but it was improperly interpreted as supercritical Hopf bifurcation with a stable mounting cycle) the Lyapunov coefficient is $l_1 \doteq 0.000278$ and so the subcritical Hopf bifurcation takes place here, an unstable cycle mount from the unstable equilibrium. On the figure 6 the branches of the periodic solutions are presented.

You can see the Hopf bifurcation critical point (HB) for $a = 0.25$ (period 48.53) and unstable cycle branch around a stable equilibrium E_1 nearby. This branch folds (LP) at $a \doteq 0.2695$ (period 51.92) and changes its stability. This is the reason, why some solutions tend to a stable cycle and others to a stable equilibrium. This coexistence of two attractors (a stable fixed point and a stable cycle) is called hysteresis. When we continue further, the cycle on the stable cycle branch bifurcates by period doubling (PD_1) at $a \doteq 0.1693$ (period 61.78). The unstable part of the branch give rise to a chaotic motion (a strange attractor). This phenomenon will be discussed in the next section. A stable cycle (period 69.83) mount again for $a \doteq 0.03342$ from the stable branch (PD_2) near zero.
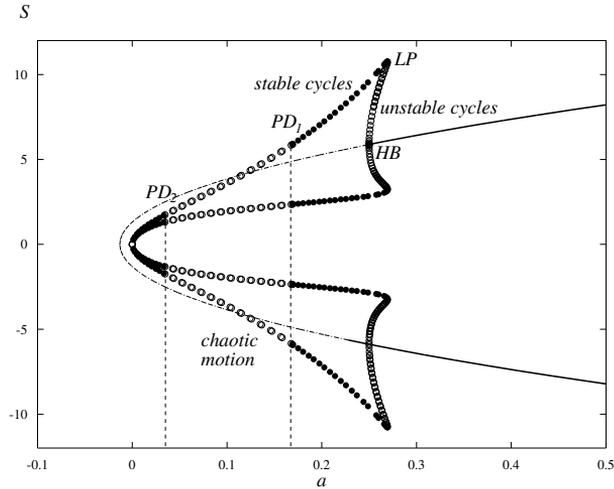
FIGURE 6. The typical Hopf bifurcation curve with arising unstable cycle.

The stable limit cycle surrounding the unstable one near the subcritical bifurcation point together with the trajectory tending to the stable equilibrium E_1 is shown on the figure 7. It demonstrates the above Example 2. for $a = 0.265$.

5. PERIOD DOUBLING ROUTE TO CHAOS

Period doubling of the limit cycle can be shown in the previous Example 2 (using continuation programs, for example Xppaut). First branch bifurcates for $a_1 \doteq 0.1692534$ (see figure 6), the next doubling takes place for $a_2 \doteq 0.1415250$ (see figure 8) and then for $a_3 \doteq 0.1346818$ (see figure 9), $a_4 \doteq 0.1331614$, $a_5 \doteq 0.1328327$, $a_6 \doteq 0.1327623$, $a_7 \doteq 0.1327473$, ...

The ratio of distances between period doubling bifurcation points is $\delta_1 = \frac{a_1 - a_2}{a_2 - a_3} \doteq \frac{0.0277284}{0.0068432} \doteq 4.052$, $\delta_2 = \frac{a_2 - a_3}{a_3 - a_4} \doteq \frac{0.0068432}{0.0015204} \doteq 4.501$, $\delta_3 \doteq 4.625$, $\delta_4 \doteq \delta_5 \doteq 4.669$ which is already pretty close to Feigenbaum constant.

This period doubling of the limit cycle lead to a chaotic motion in a bounded region — existence of a strange attractor. A chaotic region is characterized by a sensitive dependence on initial conditions, and arbitrarily close initial conditions lead to evolutions that diverge exponentially fast with time. This divergence is characterized by the maximal Lyapunov exponent (see precise definition in [1] for example). Briefly speaking, Lyapunov exponents measure average rate of divergence of nearby trajectories (for $\lambda > 0$) or convergence (for $\lambda < 0$) respectively. In our case of a three dimensional system, we have three Lyapunov exponents (for a trajectory or an attractor respectively), for a fixed point the signs are $(-, -, -)$, for a stable cycle $(0, 0, -)$ and for a strange attractor $(+, 0, -)$. The maximal Lyapunov exponent is negative for a fixed point, zero for a stable cycle and positive for chaotic strange attractor.

Changing parameter a from 0 to 0.3 (200 stps) in the Example 2, we computed the maximal Lyapunov exponent as a function of a. We used program Xppaut
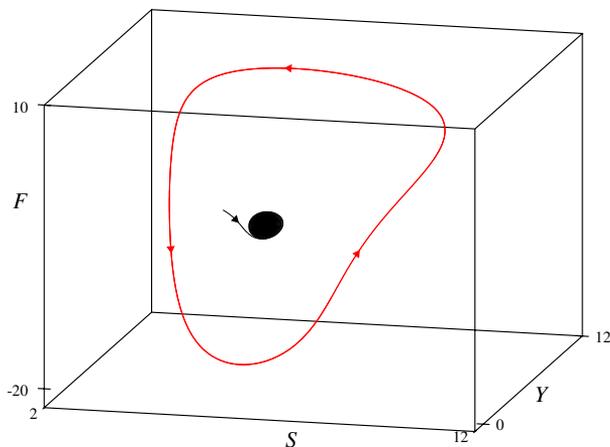
FIGURE 7. The stable cycle and a trajectory (converging to a stable equilibrium) repelling from the unstable cycle near the subcritical Hopf bifurcation due to folding. For a similar figure you can use parameters from the Example 2. and $a \in (0.25, 0.2695)$ and initial conditions near E_1 for a trajectory tending to the stable equilibrium E_1 and for the stable cycle use some farther initial condition, for example $x(0) = 10$, $y(0) = 10$, $z(0) = 10$ and $t > 6000$.
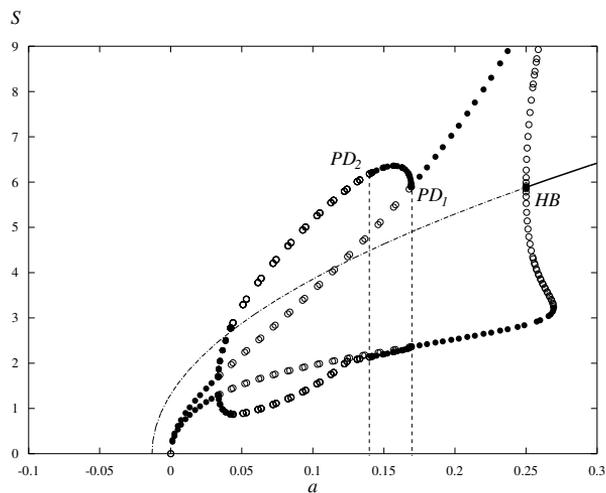
FIGURE 8. Second branch of the period doubling.

(Xppaut computes the maximal Lyapunov exponent along a computed trajectory by linearizing in each point of the trajectory, advancing one time step using a normalized vector, computing the expansion, and summing the log of the expansion.
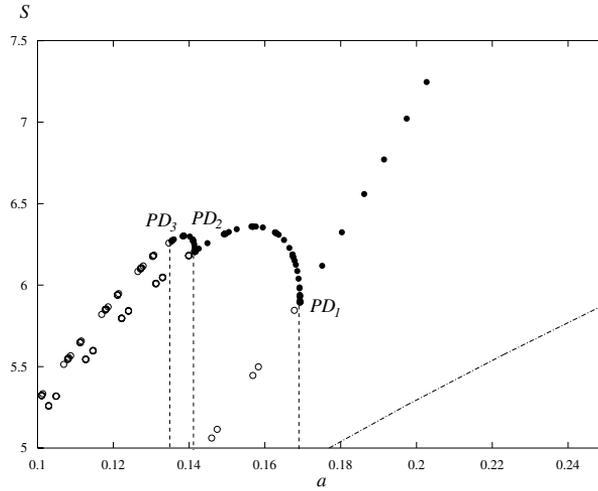
FIGURE 9. Third branch of the period doubling.

The average of this over the trajectory is a rough approximation of the maximal
Lyapunov exponent. For further details see methods presented in [2], [11] or [8].)
The maximal Lyapunov exponent is computed numerically (using Xppaut) in finite
number of points on trajectory converging to the equilibrium, periodic orbit or a
strange attractor, starting at some initial point within the basin of attraction with
first amount of transient iterations being discarded to converge to an attractor.
The results are presented on the next figures.
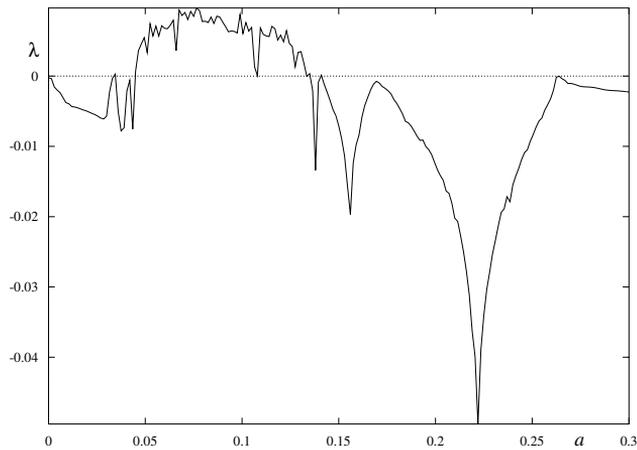
FIGURE 10. Estimation of the maximal Lyapunov exponent for
$a \in \langle 0, 0.3 \rangle$, $r = 0.25$, $p = 0.01$, $v = 0.031847$, $\kappa = 1$ and $m =$
0.19. Initial conditions for all computed trajectories: $x(0) = 0.45$,
$y(0) = 0.8$ and $z(0) = 0.35$ for $t \in \langle 3000, 6000 \rangle$

On the figure 10 you can see a range, where the maximal Lyapunov exponent is positive and that implies existence of a chaotic strange attractor for these values of parameters.

Changing the parameter v from 0 to 0.3 for $a = 0.1$ we can see a range, where the maximal Lyapunov exponent is positive again on the figure 11.
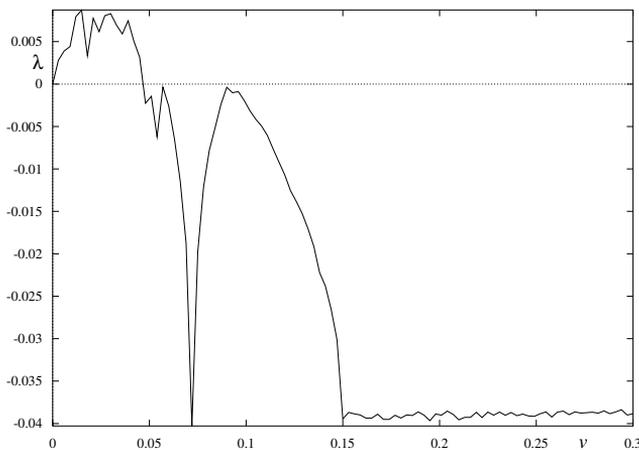
FIGURE 11. Estimation of the maximal Lyapunov exponent for $v \in \langle 0, 0.3 \rangle$, $a = 0.1$, $r = 0.25$, $p = 0.01$, $\kappa = 1$ and $m = 0.19$. Initial conditions for all computed trajectories: $x(0) = 0.45$, $y(0) = 0.8$ and $z(0) = 0.35$ for $t \in \langle 3000, 6000 \rangle$

Changing parameter p from 0 to 1 for fixed parameters $a = 0.1$, $r = 0.25$, $v = 0.031847$, $\kappa = 1$ and $m = 0.19$ and initial conditions $x = 0.45$, $y = 0.8$, $z = 0.35$ and $t \in \langle 3000, 6000 \rangle$, we got different results, since the maximal Lyapunov exponent is always either positive or zero (numerical programs of course give negative results, but it's due to the finite number of computed points and the numerical methods), that is either chaotic or periodic trajectories occur. For example, for $p = 0.2$ and $p = 0.24$ the maximal Lyapunov exponent is positive, but for $p \doteq 0.22$ zero (see figure 12) and the trajectory converges to a stable limit cycle (see figures 13 and 14). This phenomenon is caused by period-doubling route to chaos and a fractal structure of the bifurcation diagram with stable areas (stable areas of the Poincaré section more precisely). The maximal Lyapunov exponent is tending to zero from $p \doteq 0.67$ (see figure 15) and periodic orbits occur. Maximal Lyapunov exponent takes positive values for a big range of the $(0, 1)$ interval (see also figure 21) and that means this parameter cannot be used for controlling the system.

The last figure 16 presents dependence of the maximal Lyapunov exponent on the parameter m. Maximal Lyapunov exponent takes positive values as m goes to 1 (see also figure 22) and that means this parameter cannot be used for controlling the system.
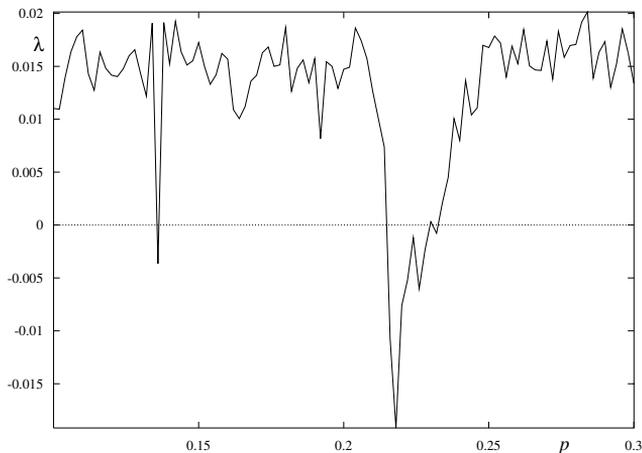
FIGURE 12. Estimation of the maximal Lyapunov exponent for $p \in \langle 0.1, 0.3 \rangle$, $a = 0.1$, $r = 0.25$, $v = 0.031847$, $\kappa = 1$ and $m = 0.19$. Initial conditions for all computed trajectories: $x(0) = 0.45$, $y(0) = 0.8$ and $z(0) = 0.35$ for $t \in \langle 3000, 6000 \rangle$
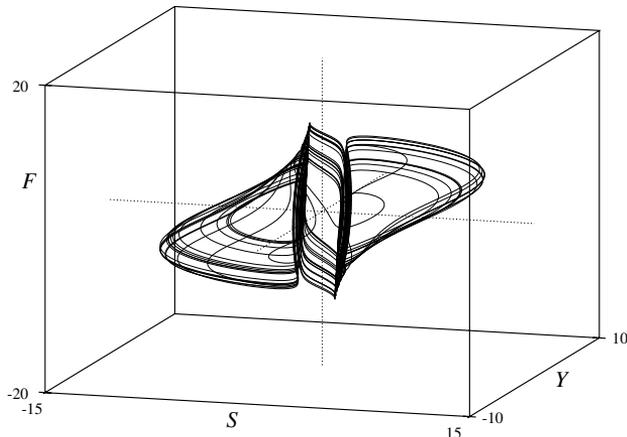
FIGURE 13. Chaotic orbit for positive maximal Lyapunov exponent for $p = 0.2$.

6. POINCARÉ SECTION AND A PERIOD THREE ORBIT

Let us take a look at the Poincaré section. A Poincaré map consists of a discrete set of values picked whenever one of the variables passes through some prescribed value. We chose to plot successive maxima of the variable y.

We computed local maxima $y(t_n) = y_n$ of the trajectory from the Example 2. with initial conditions $x(0) = 0.45$, $y(0) = 0.8$ and $z(0) = 0.35$ for $t \in \langle 3000, 30000 \rangle$
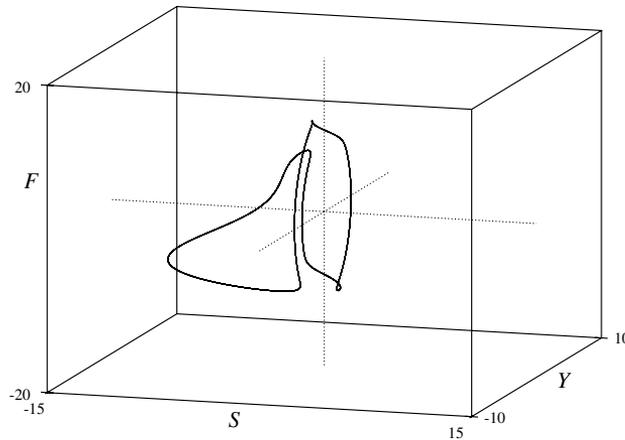
FIGURE 14. Periodic orbit for $p = 0.22$
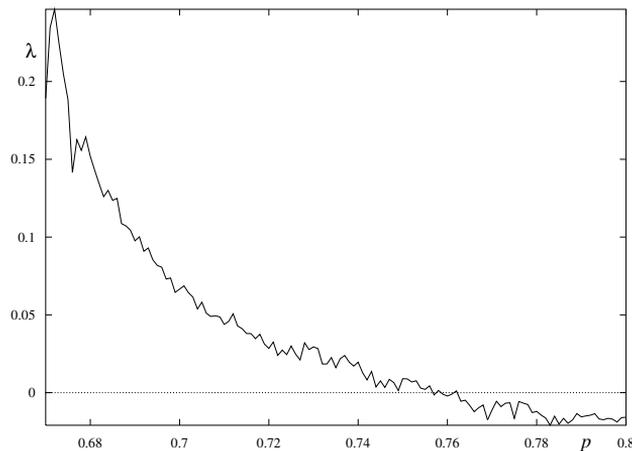
FIGURE 15. Estimation of the maximal Lyapunov exponent for
$p \in \langle 0.67, 0.8 \rangle$, $a = 0.1$, $r = 0.25$, $v = 0.031847$, $\kappa = 1$ and $m = 0.19$. Initial conditions for all computed trajectories: $x(0) = 0.45$,
$y(0) = 0.8$ and $z(0) = 0.35$ for $t \in \langle 3000, 6000 \rangle$

and got a sequence y_n. When we plotted this sequence to y_n vs. y_{n+1} plane (Ruelle plot in Xppaut), we got a black curve on the figure 17. The intersection with the first quadrant axes is a fixed point $y_n = y_{n+1}$ (FP) - a period one orbit. The sequence y_n plotted in y_n vs. y_{n+3} plane, gave another curve (blue). The intersections with the first quadrant axes different from FP are fixed point $y_n = y_{n+3}$ - corresponding to period three orbits. According to [7], the trajectory is chaotic. An example of a trajectory with a period three ($x(0) \doteq 1.9821$, $y(0) \doteq 5.8385$, $z(0) \doteq -1.9821$) found due to this Poincaré mapping is on figure 18.
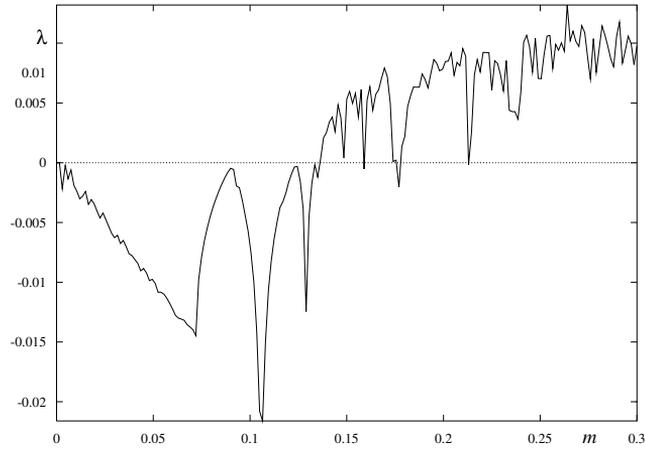
FIGURE 16. Estimation of the maximal Lyapunov exponent for $m \in \langle 0, 0.3 \rangle$, $a = 0.1$, $r = 0.25$, $v = 0.031847$, $\kappa = 1$ and $p = 0.01$. Initial conditions for all computed trajectories: $x = 0.45$, $y = 0.8$ and $z = 0.35$ for $t \in \langle 3000, 6000 \rangle$
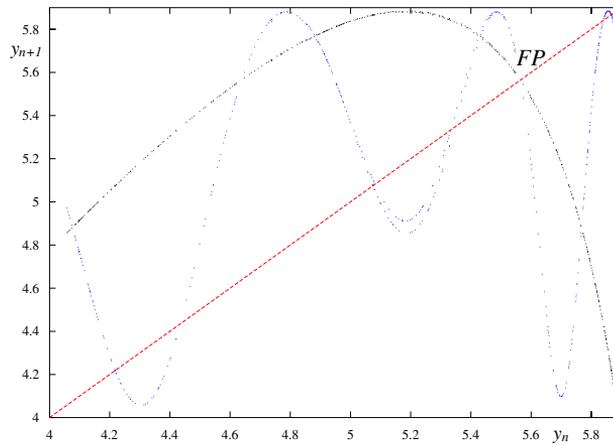
FIGURE 17. Fixed point of the Poincaré mapping corresponding to a period one orbit and intersections corresponding to period three orbits.

Compared to computing maximal Lyapunov exponent, this evidence of chaos can be used for single trajectories only and we have no image of the parameter dependence. On the other hand, maximal Lyapunov exponent estimation has an error, caused by finiteness of computed points and even may break down during reorthogonalization of the matrix Q (while using the standard QR method of computing submitted by Benettin et al. in [2]), since the computed matrix Q may
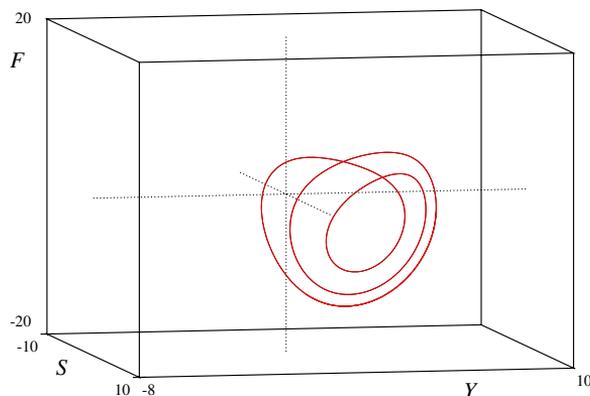
FIGURE 18. A period three orbit.

deviate from the origin one during the Gram-Schmidt orthogonalization type procedure. This last mentioned problem has been studied and overcame by Udwadia and Bremen using Cayley transformation (see [10]).

7. 0-1 TEST FOR CHAOS

Recently, a new test for determining chaos was introduced by Gottwald and Melbourne. In contrast to the usual method of computing the maximal Lyapunov exponent, their method is applied directly to the time series data and does not require phase space reconstruction. We computed time series corresponding to solutions x of the system 1.1 for parameters a, v, p and m by the fourth-order classical Runge-Kutta method with timestep 0.25 and initial conditions $x = 0.45$, $y = 0.8$ and $z = 0.35$. We used sampling time $\tau_s = 12$ (using $\tau_s = 5$ the data are oversampled still) and got time series $x(t_j)$, $t_j = j\tau_s$ for $j = 1, \ldots, 2000$. According to [5], we computed 100 times $p_c(n) = \sum_{j=1}^{n} x(t_j) \cos jc$ and $q_c(n) = \sum_{j=1}^{n} x(t_j) \sin jc$, $j = 1, \ldots, 2000$ for arbitrary chosen $c \in (0, \pi)$ and computed median K of the asymptotic growth rates K_c of the mean-square-displacement

$$M_c(n) = \lim_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} [p_c(j+n) - p_c(j)]^2 + [q_c(j+n) - q_c(j)]^2.$$

Since $N = 2000$ and for the estimation of the limit we need $n << N$, we used the last $n_{cut} = 200$. For computing K_c we used the correlation method. The median $K \approx 0$ indicates regular dynamics, $K \approx 1$ indicates chaos. Results are displayed on figures 19, 20, 21 and 22 and correspond to the estimations of the maximal Lyapunov exponent on figures 10, 11, 12 and 16.
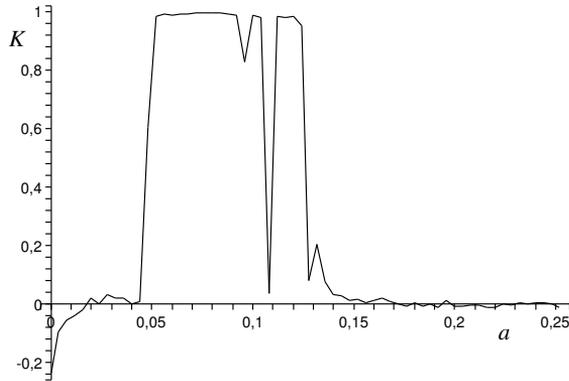
FIGURE 19. 0-1 test for chaos: chaos range for the parameter a, $r = 0.25$, $p = 0.01$, $v = 0.031847$, $\kappa = 1$ and $m = 0.19$.
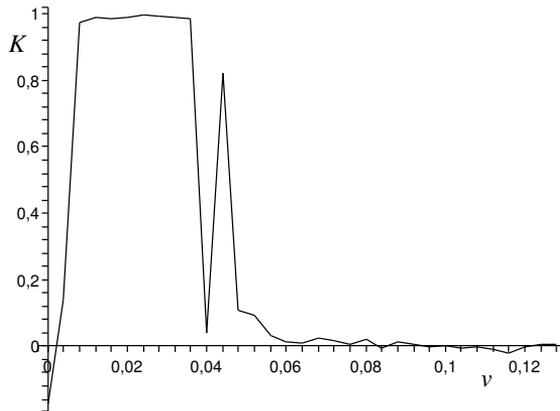
FIGURE 20. 0-1 test for chaos: chaos range for the parameter v, $a = 0.1$, $r = 0.25$, $p = 0.01$, $\kappa = 1$ and $m = 0.19$.

8. HOPF BIFURCATION NEAR ZERO

Now we will turn our attention to the zero neighbourhood. It was already mentioned above, that for the trivial equilibrium E_0 (always unstable), the necessary condition (3.2) for Hopf bifurcation is $m = -p\kappa$, while $\lambda = \pm i\omega$, where $\omega^2 = v(r - a) > 0$, that is for $r > a$ (since $v > 0$). The economic system can never reach the critical value, since we assumed the parameters to be positive. On the other hand, local behaviour of the trajectories near this critical value may coincide with the zero neighbourhood and affect the properties of the positive quadrant.
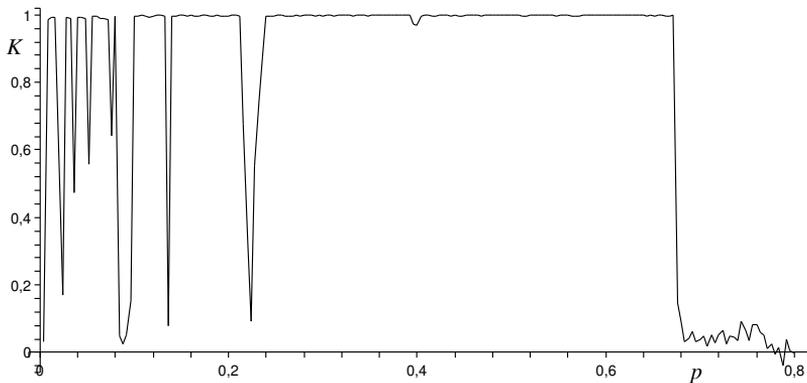
FIGURE 21. 0-1 test for chaos: chaos range for the parameter p, $a = 0.1$, $r = 0.25$, $v = 0.031847$, $\kappa = 1$ and $m = 0.19$.
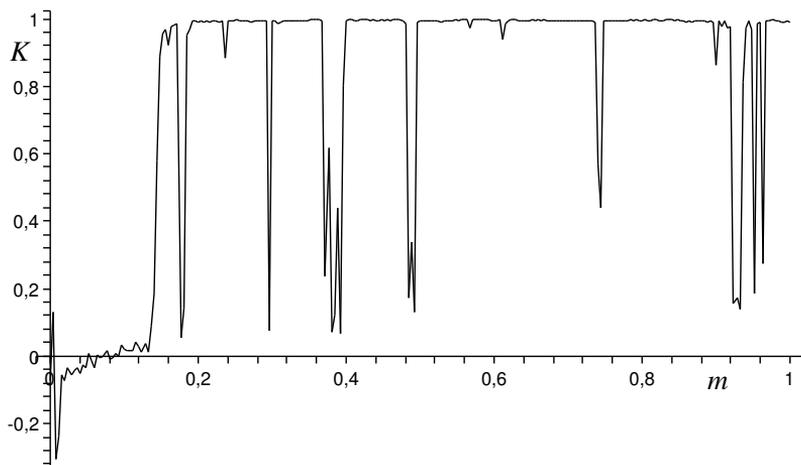
FIGURE 22. 0-1 test for chaos: chaos range for the parameter m, $a = 0.1$, $r = 0.25$, $v = 0.031847$, $\kappa = 1$ and $p = 0.01$.

Due to this we let some necessary parameters reach and even cross zero to negative values in the next part.

Using the implicit function theorem we can compute the derivative of the complex eigenvalue λ according to m for the equilibrium E_0 (using (2.4)) $\omega^2 = v(r - a)$:

$$\frac{d\lambda}{dm} = -\frac{\frac{dp(\lambda)}{dm}}{\frac{dp(\lambda)}{d\lambda}} = -\frac{-av}{3\lambda^2 + 2m\lambda + \omega^2}. \tag{8.1}$$

Substituting $\lambda_{1,2} = \pm i\omega$, we evaluate

$$\text{sgn}\,\text{Re}\,\frac{d\lambda}{dm} = \text{sgn}\,\frac{-av}{2(\omega^2 + m^2)} < 0. \tag{8.2}$$

The transversality condition for Hopf bifurcation is fulfilled and according to [6] the Hopf bifurcation gives rise to a limit cycle near the trivial equilibrium at the critical value $m = -p\kappa$ for $a < r$.

Theorem 8.1. *The limit cycle mounting from the trivial equilibrium near $m = -p\kappa$ for $a < r$ is unstable for $a > 0$, $v > 0$ and $p > 0$ (subcritical Hopf bifurcation). Generalized Hopf bifurcation of the trivial equilibrium occurs on the parametric manifold*

$$M_{GH} = \{(a, p, \kappa, v, m, r) \in \mathbb{R}^6 : m = -p\kappa, a = 0, v > 0, r > 0, p > 0\} \tag{8.3}$$

Proof. We already proved the Hopf bifurcation occures for $m = -p\kappa$ for $a < r$. To prove that its a subcritical type, we have to compute the first lyapunov coefficient l_1 using the projection method (see [6]). The Jacobi's matrix at zero for the Hopf critical parameters has the form

$$J_{\text{crit}} = \begin{pmatrix} -m & a & 0 \\ v & 0 & v \\ m & -r & 0 \end{pmatrix}$$

and its eigenvalues are $(i\omega, -i\omega, -m)$, where $\omega = \sqrt{v(r - a)}$. From this we get assumptions $v > 0$ and $r > a$ (the opposite signs are irrelevant for the economic model). For an invariant expression for the first Lyapunov coefficient we need to find eigenvectors Q and P such that $JQ = i\omega Q$, $J^T P = -i\omega P$ and $\langle P, Q \rangle = 1$. One of the eigenvectors corresponding to $\lambda = i\omega$ is

$$Q = \left(a, m + i\omega, -r + i\frac{\omega m}{v}\right).$$

The corresponding normalized eigenvector is $P = \frac{1}{c}\left(1, -\frac{i\omega}{v}, 1\right)$, where

$$c = \langle(1, -\frac{i\omega}{v}, 1), Q\rangle = 2(a - r) + 2i\frac{\omega m}{v} \neq 0$$

(as in [6] we use an unusual scalar product definition $\langle x, y \rangle = \sum \overline{x_i} y_i$ to keep the Kuznetsov's notation). According to [6]

$$\text{sgn}\,l_1 = \text{sgn}\,\frac{1}{2\omega}\,\text{Re}\langle P, C(Q, Q, \bar{Q})\rangle,$$

where for $i = 1, \ldots 3$ F_i stays for nonlinear part of the right hand side of (1.1) and we define

$$C_i(x, y, z) = \sum_{j,k,l=1}^{3} \frac{\partial^3 F_i(\xi)}{\partial \xi_j \partial \xi_k \partial \xi_l}|_{\xi=0} x_j y_k z_l,$$

that is, $C_1(x, y, z) = -4p(x_1 y_2 z_2 + x_2 y_1 z_2 + x_2 y_2 z_1)$, $C_2(x, y, z) = 0$ and $C_3(x, y, z) = 0$. (The general formula for l_1 contains also quadratic members of the nonlinear part of the right hand side of (1.1), but it vanishes in our case). From that we get

$$\text{sgn}\,l_1 = \text{sgn}\,\frac{apv}{\omega}.$$

The generalized Hopf bifurcation occures for $l_1 = 0$, that is only for $a = 0$. We have to exclude cases $p = 0$, when the system (1.1) is strictly linear, and of course

$v = 0$, when $\lambda_{1,2} = \pm i\omega = 0$. Consequently, $l_1 > 0$ for $r > a > 0$, $v > 0$ and $p > 0$ and subcritical Hopf bifurcation occurs near zero for $m = -p\kappa$. $\square$

As a consequence of (8.2) and the Theorem 8.1 we see, that unstable cycle arises on the right hand side of m, that is for $m > -p\kappa$, since the real part of the complex conjugate eigenvalues decreases from positive to negative values as m increases and an unstable equilibrium $(0,0,0)$ become stable on the center manifold with an unstable cycle in its local neighbourhood. This is the reason, why we included study of the economically impossible Hopf bifurcation of the zero equilibrium into account - an unstable limit cycle may continue (and already does) in the positive parametric space. You can see the bifurcation diagram on the figure 23
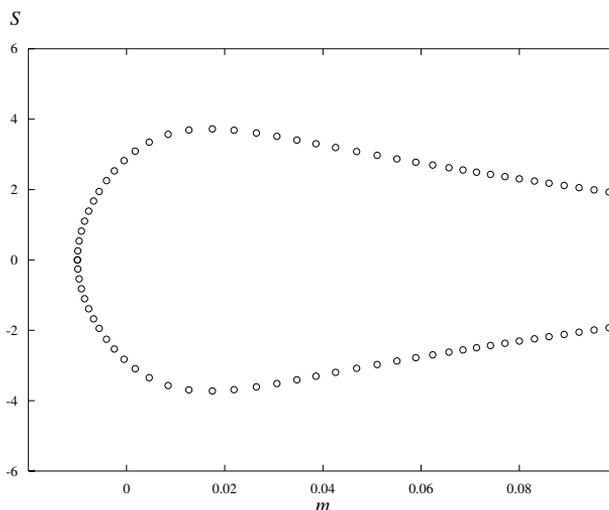
FIGURE 23. The typical Hopf bifurcation diagram with arising unstable cycle near zero for parameters $a = 0.1$, $r = 0.25$, $p = 0.01$, $v = 0.031847$, $\kappa = 1$ and $m \in (-0.02, 0.1)$.

9. CONCLUSIONS

In this paper, we studied several types of bifurcations in the system (1.1) that are connected with periodic and non-periodic bounded trajectories that may represent economic cycle and non-periodic chaotic oscillations in macroeconomic quantities. From the economic point of view, condition (1.4) does not guarantee existence of a stable equilibrium. Nor even stronger condition $a > r$, that according to Theorem 3.1 guarantees existence of a stable equilibria, does not guarantee asymptotic tending towards one of the stable fixed points, since for a set of parameters the first Lyapunov coefficient is positive and folding of the unstable cycle into a stable one give rise to a large basin of attraction of a stable periodic solution for $a > r$ close to r. This result is in the contrary to the conclusions made in [9]. If the debt/output ratio is less than the variation of the marginal propensity to savings then the economic equilibrium is locally stable, but the economy near a stable state

may be on a non-local destabilizing cycle asymptotically tending to a non-local stable cycle. As we can see from results described by the figure 1, for normal values of capital-output ratio $v << 1$, the first Lyapunov coefficient is positive for quite big range of parameter a, even the less is the value of capital-output ratio v, the bigger "unstable" range we get for a. For normal economic parameters $a >> r$ the equilibria $E_{1,2}$ are always stable and attracting, no cycles or chaos appear in the system. But for low v, p or large m, there exist trajectories corresponding to unstable trade cycles that may even change into non-periodic bounded chaotic unpredictable regime. Due to this we cannot agree with the conclusion in [9] that if the capital inflow/savings ratio is less than double the ratio of capitalized profit then the system is in a stable state. You can see on the figures 2 and 3 ($m = 0.19$, $p = 0.1$) that for low variation of the marginal propensity to savings this is not true, unstable cycles occurs. As it can be seen from figures 16 and 22, parameter m, that is capital inflow/savings ratio, cannot be used for controlling the system at all. Similarly p, the ratio of the capitalized profit, cannot be used for controlling (see figures 12, 15 and 21). The first conclusion in [9] is true: an increasing of the capitalization of the profits demonstrates well-known results in economics that the capitalization of profits causes the stabilization of the system.

It should be taken into account, that the model of foreign financing is highly simplified and therefore the conclusions may represent possibilities of the real economic behaviour only. On the other hand the model is "nearly linear" - and linear models are often used in economics to show "the invisible hand of the market" that leads the economy to the stable and predictable, equalized quantities. The one cubic term included in the first equation (also included in the model without foreign financing with no bifurcations at all, for analysis of this model see [9]) give rise to a great deal of nonlinear dynamics - to periodic and chaotic motion with no invisible hand to lead.

The mathematical results may be applied to the second model of firm introduced by [3] that is based on the Hunt's hypothesis that the call for loans pushes the profit ratio of stockholderscapital. Bouali shows various periodicity and chaos in the system without deeper mathematical background. The results derived here explain the dynamics in this model of firm. In order to illustrate farther economic applicability of these results, we cite the economic conclusions made by Bouali. "... rules and principles of finance governance built in static framework may lose their validity. The findings of a well corporate debt policy connected to a well dividend policy may lead to an unpredictable and hazardous motion of the profit. In our 3D system, the rise of the loss level is an endogenous outcome of the borrowing policy and is not determined by a shock of economic recession. Against the common sense, the profit motion is worsened by the braking of dividend distribution!"

The existence of hysteresis and various types of bifurcations and chaos open a question, whether the stabilization policy is efficient. The policy advice is based mostly on linear models while the economy is actually characterized by significant nonlinearities. Linear modelling of a system with existence of significant nonlinearity in the data may provide misleading results. In the case of hysteresis, stabilization policy may lead to destabilizing trade cycles instead of tending to an equilibrium or to chaos. In the case of various cycle bifurcations the data itself cannot be correctly analyzed by methods based on linear modelling and in the case of chaos,

the precision of a forecast with a very small error in the initial conditions worsens exponentially over time in addition to this previously mentioned failings.

References

[1] Alligood, K.T., Sauer, T.D., Yorke, J.A. (1996) *Chaos - an Introduction to Dynamical Systems*, Berlin, Heidelegrg, New York, Springer-Verlag

[2] Benettin, B., Galgani, L., Giorgilli, A. and Strelcyn, J. M. (1980) Lyapunov Characteristic Exponents for Smooth Dynamical Systems and Hamiltonian Systems; a Method for Computing All of Them, Part I. and Part II., *Meccanicca* 15, 9-30

[3] Bouali, S. (2002) The Hunt Hypothesis and the Dividend Policy of the Firm. The Chaotic Motion of the Profits., *8th International Conference of the Society for Computational Economics Computing in Economics and Finance*, Aix-en-Provence, France, June 27-29

[4] D'adda, C., Scorcu, A.E. (2003) On the time stability of the output-capital ratio, *Economic Modelling* 20, issue 6, 1175-1189

[5] Gottwald, A G., Melbourne, I. (2009) On the Implementation of the 01 Test for Chaos, *SIAM J. Appl. Dyn. Syst.* 8, issue 1, pp. 129-145

[6] Kuznetsov, Y.A. (1998) *Elements of Applied Bifucation Theory*, Second Edition, Applied Mathematical Sciences 112, Berlin, Heidelegrg, New York, Springer-Verlag

[7] Li, T. Y., Yorke, J.A. (1975) Period Three Implies Chaos, *American Mathematical Monthly* 82, 985-992

[8] Eckmann, J. P., Ruelle, D. (1985) Ergodic Theory of Chaos and Strange Attractors, *Rev. Mod. Phys.* 57, 617-656.

[9] Vošvrda, M. (2001) Bifurcation Routes and Economic Stability, *Bulletin of the Czech Econometric Society* 14, 43-60

[10] Udwadia, F. E., von Bremen H. F. (2001) Computation of Lyapunov Characteristic Exponents for Continuous Dynamical Systems, *Math. Phys.*, Vol. 53, 23-146

[11] Wolf A., Swift, J. B., Swinney, H. L. and Vastano, J. A. (1985) Determining Lyapunov Exponents from a Time Series, *Physica D* 16, 285-317.

Lenka Přibylová

Dept. of Applied Mathematics, Masaryk University, Janáčkovo nám. 2a, 602 00 Brno, Czech Republic

E-mail address: pribylova@math.muni.cz

The Solow-Swan model generalization with non-constant labor growth rate

LENKA PŘIBYLOVÁ

Department of Mathematics and Statistics, Faculty of Science, Masaryk University Brno, Botanická 2, 611 37 Brno, Czech Republic.
e-mail: pribylova@math.muni.cz

Abstract. Generalization of the neoclassical one-sector Solow-Swan model and its modification of Kaldor-Pasinetti type are introduced. The generalization is due to non-constant labor growth rate. Both continuous-time and discrete-time models are derived. Concrete examples with complex dynamics are presented for models with differential savings of the Kaldor-Pasinetti type and Beverton-Holt and logistic version of the labor dynamics. We found that instabilities and complex dynamics may be driven by the capital–labor ratio dynamics, labor dynamics or both together.

Keywords: Solow-Swan model, population growth, stability, chaos

JEL classification: O41, C62
AMS classification: 37D45, 91B62, 92D25

1. Introduction

The standard dynamic economic one-sector neoclassical Solow-Swan model ([12], [13] or [2]) show that the system monotonically converges to the equilibrium and neither cycles nor complex dynamics can be observed.

In standard economic theory, the labor growth rate is assumed to be constant, although this is clearly unrealistic, since the consequence of the constant labor growth rate is unbounded exponential population growth. In population dynamics it's usual to consider some carrying capacity of the environment, the maximal size of the population that the environment can sustain in the long term (see [11] or [14]). There are few papers that deal with non-constant labor growth rate (for example [1] introduced the Solow-Swan model with Richards' law labor dynamic, [6] used a continuous logistic growth rate dynamics and [5] tried to deal with Beverton-Holt discrete labor dynamics, but made a mistaken economic derivation of the equations, since the Beverton-Holt model describes population (or density), not the growth rate dynamics).

In this paper the Solow-Swan model and its modification of Kaldor-Pasinetti type (discrete and continuous) are generalized by implementation of the labor dynamics.

2.   The continuous-time model

Let us consider a standard neoclassical one-sector continuous growth model

$$\dot{K} = I - \delta K, \tag{1}$$

where capital K depreciates at the constant rate $\delta > 0$ and increases with investment I that is considered to be the saved part of the produced output Y given by the production function

$$Y = F(K, L). \tag{2}$$

In the basic Solow-Swan model the saving rate $s \in (0, 1)$ is supposed to be constant and the equation (1) has the form

$$\dot{K} = sF(K, L) - \delta K. \tag{3}$$

The production function F satisfy the following conditions:

$$F(K, L) > 0 \quad \text{for all} \quad K > 0 \text{ and } L > 0, \tag{4}$$

$$\frac{\partial F}{\partial K} > 0, \ \frac{\partial F}{\partial L} > 0, \ \frac{\partial^2 F}{\partial K^2} < 0 \text{ and } \frac{\partial^2 F}{\partial L^2} < 0 \quad \text{for all} \quad K > 0 \text{ and } L > 0 \tag{5}$$

and F exhibits constant return to scale

$$F(\lambda K, \lambda L) = \lambda F(K, L) \quad \text{for all} \quad \lambda > 0. \tag{6}$$

The constant return to scale condition is also called linear homogenity condition, since from the Euler's theorem, the equation $F(K, L) = K\frac{\partial F}{\partial K} + L\frac{\partial F}{\partial L}$ holds for the marginal product of capital and labor. The constant return to scale condition allows the intensive form

$$y = \frac{Y}{L} = F(K/L, 1) = f(k),$$

where $k = \frac{K}{L}$ is the capital–labor ratio and y is per capita output.

Moreover the production function F usually satisfies Inada conditions

$$\lim_{K \to 0^+} F_K = \lim_{L \to 0^+} F_L = \infty, \ \lim_{K \to \infty} F_K = \lim_{L \to \infty} F_L = 0, \tag{7}$$

where F_K and F_L are marginal product of capital and labor respectively (as functions of K and L). These conditions can be written in the intensive form as

$$\lim_{k \to 0^+} f'(k) = \infty, \ \lim_{k \to \infty} f'(k) = 0. \tag{8}$$

Weak Inada conditions in the intensive form may be written as

$$\lim_{k \to 0^+} \frac{f(k)}{k} = \infty, \ \lim_{k \to \infty} \frac{f(k)}{k} = 0. \tag{9}$$

It is clear that differentiable per capita production function f that satisfy Inada conditions satisfy also weak conditions. It's not true in the other direction. For example the per capita production function $f(k) = 2 - e^{-k}$ satisfies weak Inada

conditions and $\lim_{k \to 0^+} f'(k) = 1$. Notice that in this case also capital is not essential for production, that is $F(0, L) = Lf(0) > 0$ is allowed.

Dividing the equation (3) by L now gives (1) in the intensive form

$$\dot{k} = G(k) - (\delta + n(t))k, \qquad (10)$$

where $G(k) = sf(k)$ and $n(t) = \frac{\dot{L}}{L}$ is the population (labor) growth rate. In standard economic growth theory, $n(t)$ is assumed to be constant which implies unrealistic exponential growth and the equation (10) doesn't depend on labor explicitly.

Following Böhm and Kaas [4] a generalized model of Kaldor-Pasinetti type can be considered. We expect two types of agents, workers and shareholders, with generally different but constant saving rates s_w and s_r. Since $\frac{\partial F}{\partial K} = f'(k) > 0$ and $\frac{\partial F}{\partial L} \equiv w(k) = f(k) - kf'(k) > 0$, we can write the intensive form of the equation (1) as

$$\dot{k} = G(k) - (\delta + n(t))k, \qquad (11)$$

where $G(k) = s_w w(k) + s_r k f'(k)$. For $s_w = s_r = s$, the equation is simplified into (10).

For both models (10) and (11) we will consider a non-constant labor growth rate $n(t)$, more precisely the labor dynamics will be described by the equation of population dynamics

$$\dot{L} = n(t)L, \qquad (12)$$

with generally non-constant $n(t)$ and in most population models $n(t) > 0$. We will also expect existence of so called carrying capacity L_∞ of the environment that guarantees the upper bound of the labor (population). We will assume $0 < L_0 \leq L_\infty$. These assumptions are more realistic in the long run than the standard usage of constant $n(t) = n$, which implies unlimited growth of the population (see for example [14], [11] or [3]). Typical growth of the population (labor) is sigmoidal, roughly we can say that the population grows slowly near zero, for the middle values more quickly and then it deflates as it tends to the carrying capacity. A common case is described by logistic differential equation

$$\dot{L} = n_0 L \left(1 - \frac{L}{L_\infty}\right), \qquad (13)$$

with explicit solution $L(t) = \frac{L_\infty L_0 e^{n_0 t}}{L_\infty + L_0(e^{n_0 t} - 1)}$, where L_0 is initial labor (population) and the growth rate n_0 is positive. It's clear that the steady state $L(t) \equiv L_\infty$ is globally asymptotically stable, while $n(t) \to 0$ as the solution tends to L_∞.

Theorem 1. *The steady state k^* of the equation (11) with logistic labor dynamic satisfy condition*

$$\frac{G(k^*)}{k^*} = \delta.$$

If $s_r \geq s_w$, the weak Inada conditions (9) guarantee existence of exactly one non-trivial equilibrium and the steady state is locally asymptotically stable.

Proof. The long-run dynamics of (11) with logistic labor dynamic is described by the dynamics on the attractor $L = L_\infty$, that is by the equation $\dot{k} = G(k) - \delta k$. If $s_r \geq s_w$ then $\frac{G(k^*)}{k^*}$ is strictly decreasing since

$$\left(\frac{G(k)}{k}\right)' = s_w \frac{f'(k)k - f(k)}{k^2} + (s_r - s_w)f''(k) < 0$$

and

$$\lim_{k \to 0^+} \frac{G(k)}{k} = \infty, \ \lim_{k \to \infty} \frac{G(k)}{k} = 0,$$

that guarantees existence and uniqueness of the non-zero equilibrium. If $s_r \geq s_w$ and the weak Inada conditions (9) are satisfied, we have

$$\begin{aligned} G'(k^*) - \delta &= s_w w'(k^*) + s_r f'(k^*) + s_r k^* f''(k^*) = \\ &= (s_r - s_w)k^* f''(k^*) - s_w \frac{f(k^*) - k^* f'(k^*)}{k^*} < 0. \end{aligned}$$

Consequently the steady state k^* of the equation $\dot{k} = G(k) - \delta k$ is locally asymptotically stable.

Notice that if the saving propensities satisfy $s_r < s_w$, more equilibria can occur. This give rise to room for local bifurcations of saddle-node type. The following example shows a concrete case of near-Leontief production function

$$f(k) = a\left(k + \alpha \ln\left(\frac{1 + e^{\frac{-b}{\alpha a}}}{1 + e^{\frac{ak-b}{\alpha a}}}\right)\right) + c, \qquad (14)$$

as an approximation of Leontief production function $f(k) = \min(ak, b) + c$, $a, b, c > 0$ for $\alpha > 0$ near zero. It can be easily verified that this family of functions satisfy the weak Inada conditions. For parameters $\alpha = 0.01$, $a = 0.2$, $b = 1$, $c = 0.01$, $L_\infty = 10$, $n_0 = 0.4$, , $\delta = 0.1$ and $s_w = 0.8 > s_r = 0.4$, you can see that nullclines cross three times, the outer equilibria are stable, the middle one is an unstable saddle. In this case there is switch (on the saddle separatrix) between behaviour of the trajectories as you can see on the figure 1.

The saddle-node bifurcation takes place here for both the parameters s_w and s_r (see the figure 2).

3.   The discrete-time model

A standard neoclassical one-sector discrete growth model has similar form

$$K_{t+1} - K_t = I_t - \delta K_t, \qquad (15)$$

with the production function $Y = F(K, L)$ and the constant saving rate $s \in (0, 1)$ the form

$$K_{t+1} - K_t = sF(K_t, L_t) - \delta K_t. \qquad (16)$$

Dividing the equation (3) by L_t gives the intensive form of (15)

$$k_{t+1}n_t = G(k_t) + (1 - \delta)k_t, \qquad (17)$$

4
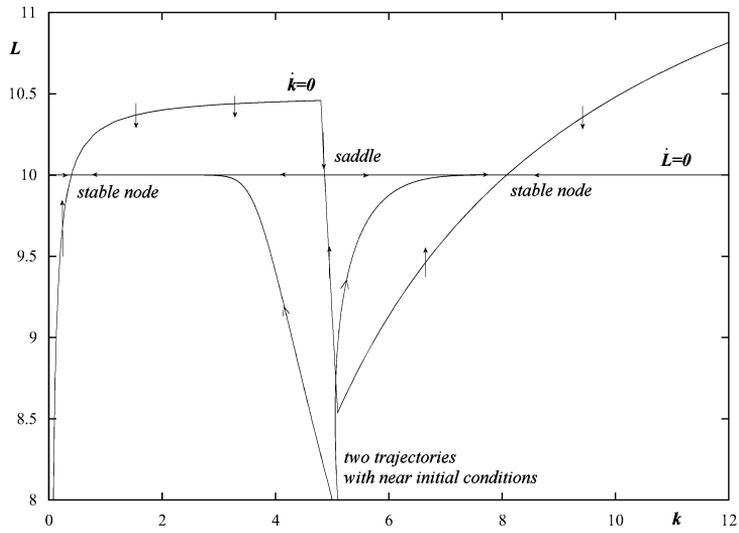
Figure 1: Nullclines and near starting trajectories
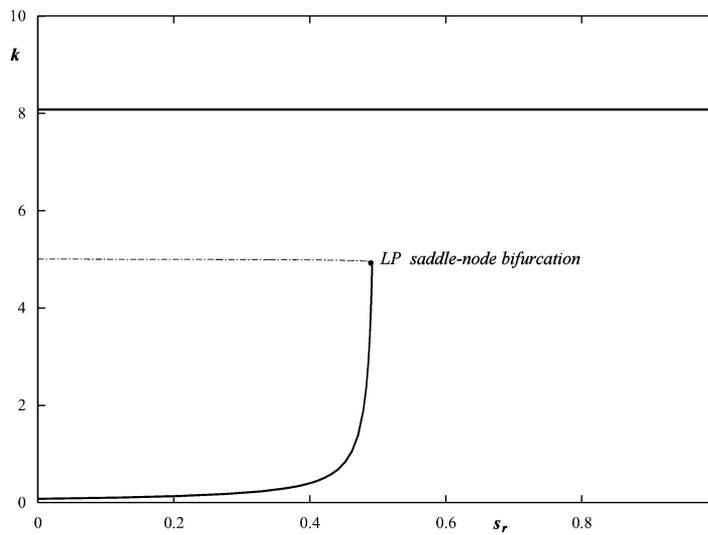
Figure 2: Bifurcation diagram for s_r, solid line is for stable, dashed line for unstable equilibrium.

where $G(k_t) = s_w w(k_t) + s_r k_t f'(k_t)$ and $n_t = \frac{L_{t+1}}{L_t}$ is the population (labor) growth rate. Besides the general case we will consider two basic discrete labor growth models. The first one is described by Beverton-Holt equation

$$L_{t+1} = \frac{n_1 L_\infty L_t}{L_\infty + (n_1 - 1)L_t},$$
(18)

for $n_1 > 1$, with explicit solution (see [3])

$$L_t = \frac{L_\infty L_0}{L_0 + (L_\infty - L_0)n_1^{-t}}.$$

Because of this structure, the model can be considered as the discrete-time analogue of the previously mentioned continuous-time logistic equation for population growth introduced by Verhulst [14]. It is clear that the fixed point $L_t = L_\infty$ is globally asymptotically stable. The second labor growth model is described by the logistic equation

$$L_{t+1} - L_t = n_2 L_t \left(1 - \frac{L_t}{L_\infty}\right)$$
(19)

or equivalently in the form analogous to (12)

$$L_{t+1} = \left(1 + n_2 \left(1 - \frac{L_t}{L_\infty}\right)\right) L_t,$$
(20)

where $n_2 > 0$. Brianzoni, Mammana and Michetti in [5] tried to derive this model with Beverton-Holt labor dynamics, but they used it in the form

$$n_{t+1} = \frac{rh}{h + (r - 1)n_t} n_t,$$

where $n_t = \frac{L_{t+1} - L_t}{L_t}$ was the growth rate (according to the derivation of their intensive form), h and r positive constants. This dynamics cannot be used for growth rates, but for labor L_t as we did in (18). It is usually used in biology for describing population or its density, the model has two steady states $n_t = 0$ and $n_t = h$ (so called carrying capacity, in our case L_∞), the second obviously cannot be a growth rate. If n_t meant the labor (and not rates) in [5], then their intensive form of the capital equation is not correct. In short, their system is mixed up.

    Correctly the equations of the generalized model of Kaldor-Pasinetti type with non-constant labor force growth rate are generally

$$T := \begin{cases} L_{t+1} = T_1(L_t) = n_t L_t \\ k_{t+1} = T_2(L_t, k_t) = \frac{1}{n_t}(G(k_t) + (1 - \delta)k_t) = \frac{1}{\varphi(L_t)}(G(k_t) + (1 - \delta)k_t), \end{cases}$$
(21)

where $n_t = \varphi(L_t) \neq 0$ (more generally we could take $n_t = \phi(t)$, but this is not that case). The Beverton-Holt version of the model is

$$T_{BH} := \begin{cases} L_{t+1} = \frac{n_1 L_\infty}{L_\infty + (n_1 - 1)L_t} L_t \\ k_{t+1} = \frac{L_\infty + (n_1 - 1)L_t}{n_1 L_\infty}(G(k_t) + (1 - \delta)k_t), \end{cases}$$
(22)

6

where $n_t = \varphi(L_t) = \frac{L_\infty + (n_1 - 1)L_t}{n_1 L_\infty}$, and the logistic version is

$$T_{log} := \begin{cases} L_{t+1} = \left(1 + n_2\left(1 - \frac{L_t}{L_\infty}\right)\right)L_t \\ k_{t+1} = \left(1 + n_2\left(1 - \frac{L_t}{L_\infty}\right)\right)^{-1}(G(k_t) + (1 - \delta)k_t), \end{cases} \tag{23}$$

where $n_t = \varphi(L_t) = \left(1 + n_2\left(1 - \frac{L_t}{L_\infty}\right)\right)$.

4. Near-Leontief production functions

Böhm and Kaas in [4] proved that chaotic behaviour exhibits for the equation

$$k_{t+1} = \frac{1}{1 + n}(G(k_t) + (1 - \delta)k_t) \tag{24}$$

with constant labor growth rate n and near-Leontief production function

$$f(k) = a\left(k + \alpha \ln\left(\frac{1 + e^{\frac{-b}{\alpha a}}}{1 + e^{\frac{ak-b}{\alpha a}}}\right)\right) + c, \tag{25}$$

as an approximation of Leontief production function $f(k) = \min(ak, b) + c$, $a, b, c > 0$ for $\alpha > 0$ near zero. It can be easily verified that this family of functions satisfy the weak Inada conditions. Böhm and Kaas presented a typical bifurcation diagram with period doubling route to chaos. This can be clearly observed for Beverton-Holt version (22) with $n_1 = 1$ and logistic version (23) with $n_2 = 0$ (see figure 3).

The question to discuss is what happens if $n_t \neq 0$ or constant.

Theorem 2. *The Beverton-Holt version (22) of the generalized Solow-Swan model with non-constant labor growth rate exhibits the same long-run behaviour as the Böhm and Kaas model (24) with $n = 0$.*

Proof. Since $L_{t+1} = T_1(L_t)$ has globally asymptotically stable fixed point $L_t = L_\infty$, all solutions tend to the attractor L_∞ with $n_t = \varphi(L_t) \to 1$. The long-run dynamics of (22) is described by the dynamics on the attractor, where $n_t = 1$, and is given by the second equation $k_{t+1} = T_2(L_\infty, k_t) = G(k_t) + (1 - \delta)k_t$ that is equation (24) with $n = 0$.

As a consequence of this theorem and Proposition 5 from [4] we have

Corollary 3. *For all $0 < \delta < 1$, $s_r > 0$ and $s_w = s_r - \varepsilon$ with $\varepsilon > 0$, there exists an open set of near-Leontief production functions (25) such that (22) exhibits chaos.*

We get an analogous result for the logistic version (23) with $n_2 \in (0, 2)$ ($\mu \in (1, 3)$), since $L_t = L_\infty$ is globally asymptotically attracting fixed point.

Corollary 4. *For all $n_2 \in (0, 2)$, $0 < \delta < 1$, $s_r > 0$ and $s_w = s_r - \varepsilon$ with $\varepsilon > 0$, there exists an open set of near-Leontief production functions (25) such that (23) exhibits chaos.*

7
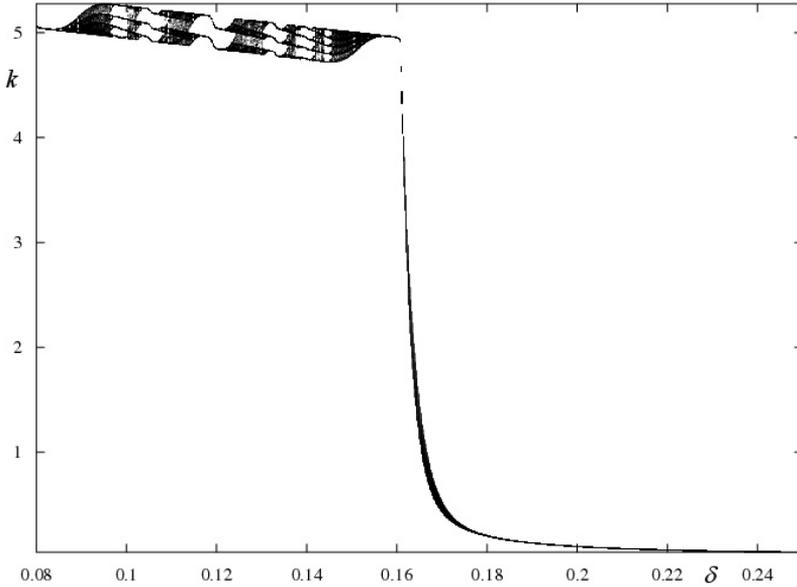
Figure 3: Bifurcation diagram of Beverton-Holt version (22) with $n_1 = 1$ or logistic version (23) with $n_2 = 0$ with parameters $a = 0.2$, $b = 1$, $c = 0.01$, $s_r = 0.8$, $s_w = 0.4$, $L_\infty = 10$ and i.c. $k_0 = 5$, $L_0 = 1$.

The bifurcation diagram for the case of $n_2 = 1.3$ is on the figure 4.You can see that it's the same dynamics as on the figure 3 (with constant labor growth rate).

The most interesting behaviour of the system (23) is chaos driven by both equations - the labor dynamics and the capital dynamics too. You can see such an example on the next figures. For $n_2 = 2$ a stable 2-cycle of labor dynamics appears and splits into 4-cycle at $n_2 = \sqrt{6}$ (see figures 5 and 6) and period doubling routes to chaos at $n_2 \doteq 2.5699457$ (see figures 7, 8 and 10). A big window with 3-cycle is around $n_2 = 2.83$ (the figure 9). The bifurcation diagrams show splitting the stable branch together with the fractal attractor (you can see the self-similarity and dividing).

Bifurcation diagrams are made in XPPAUT. We varied the parameter δ with fixed other parameters, usually we counted the attractor for 2000 values of parameter δ in each figure. For each δ we omitted the first 500 iterates to get close enough to the attractor and plotted points to the 750th iterate.

5.   Conclusions

In this paper we assumed that the labor is proportional to the population and consequently it can be modelled by Beverton-Holt or logistic growth model. Commonly exponential growth model would almost always be used. Arguments that speak in favour of the first two models are at least two. First is theoretical: exponential growth implies unlimited growth of population and labor which is surely
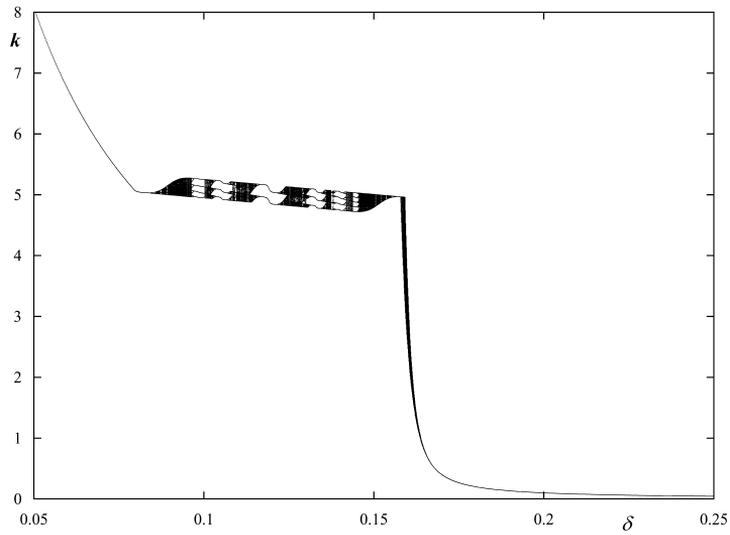
Figure 4: Bifurcation diagram of logistic version (23) for $n_2 = 1.3$ with parameters $a = 0.2$, $b = 1$, $c = 0.01$, $s_r = 0.8$, $s_w = 0.4$, $L_\infty = 10$ and i.c. $k_0 = 5$, $L_0 = 1$.
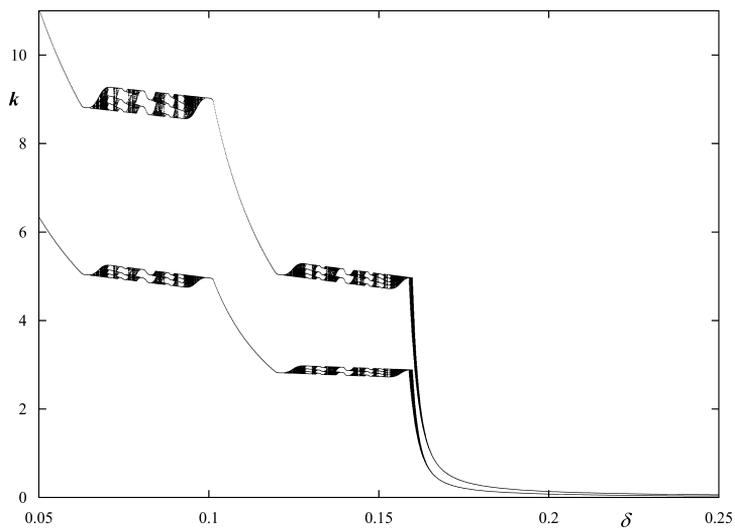
Figure 5: Bifurcation diagram of logistic version (23) for parameter $n_2 = 2.3$ and other as previous, 2-cycle
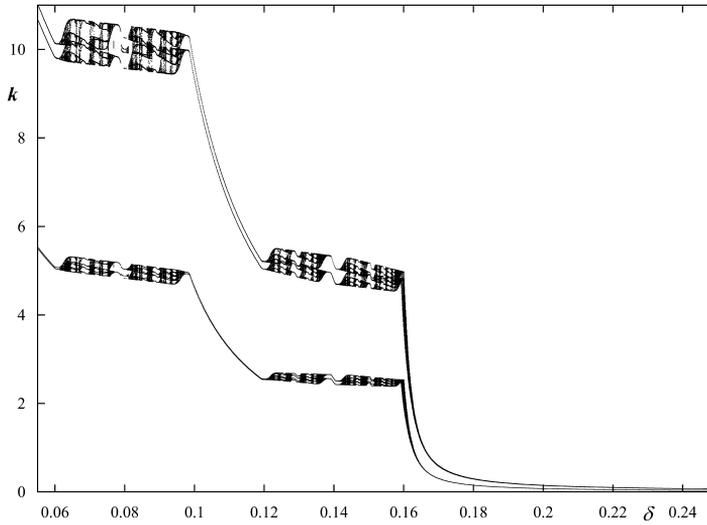
Figure 6: Bifurcation diagram of logistic version (23) for parameter $n_2 = 2.45$ and other as previous, starting 4-cycle
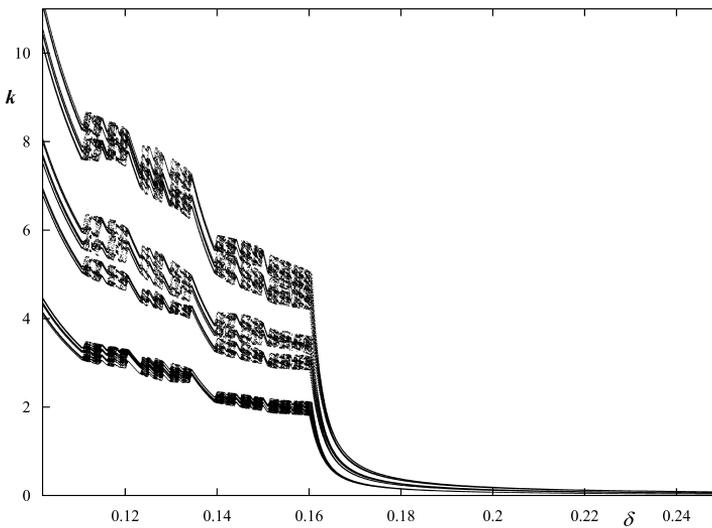
Figure 7: Bifurcation diagram of logistic version(23) for parameters $n_2 = 2.57$ and other as previous, starting chaos

10
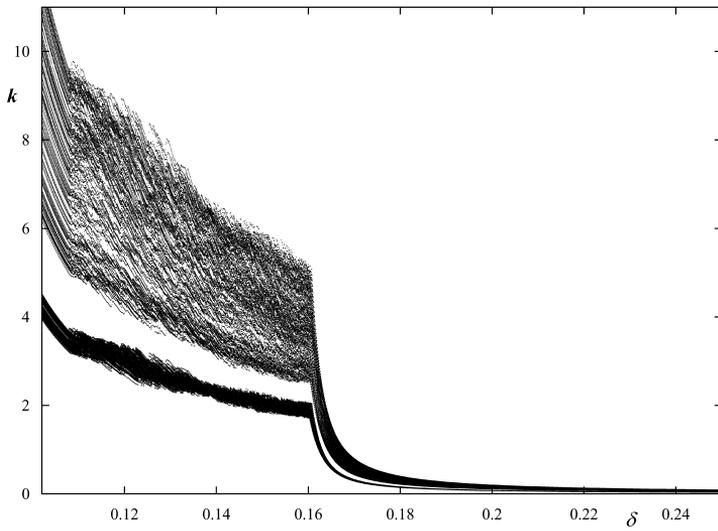
Figure 8: Bifurcation diagram of the logistic version (23) for parameters $n_2 = 2.6$ and other as previous, chaos
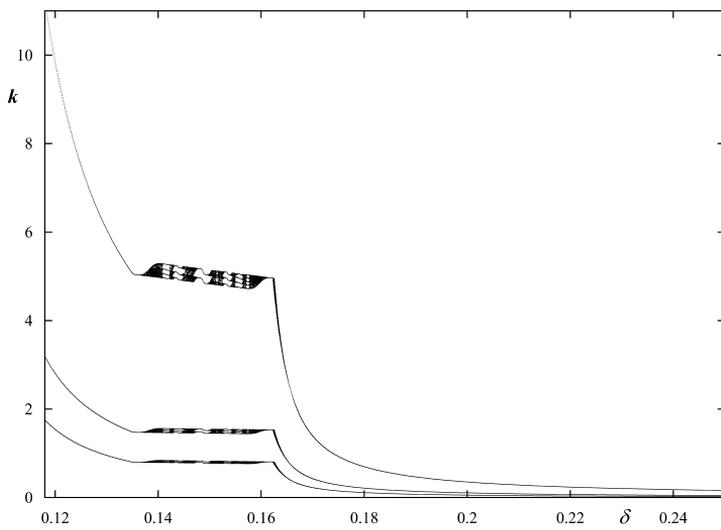
Figure 9: Bifurcation diagram of the logistic version (23) for parameters $n_2 = 2.83$ and other as previous, 3-cycle
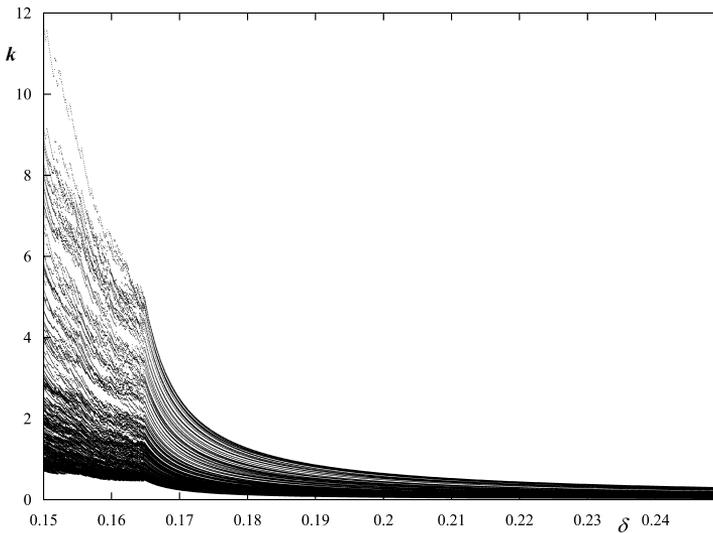
Figure 10: Bifurcation diagram of the logistic version (23) for parameters $n_2 = 2.9$ and other as previous, chaos

impossible. Second reason is based on real observations, for example population of Czech republic from 1950 to 2008 has decreasing growth rates and is better to fit by Beverton-Holt or logistic trend (see World Population Prospects: The 2008 RevisionPopulation Database at http://esa.un.org/unpp/). Modelling the labor dynamics by these equations is of course a very simplifying assumption, since the labor market responds to many other variables, but it is still more realistic than the assumption of the constant growth rate.

The results of this paper show that the labor dynamics influence the capital–labor ratio dynamics in more complex way than it is in the basic one-sector models of economic growth with constant labor growth rate, that is with exponential labor dynamics. For the case of the labor tending to the globally asymptotically stable fixed point it proves that the long run behaviour of the capital–labor ratio is the same as formerly studied models with constant, but zero, labor growth rate. It has to be mentioned that in the case of constant labor growth rate there is not a (positive) fixed point of the labor equation (the labor grows exponentially) although the capital-labor ratio has an equilibrium. This case is structurally unstable, since varying of the labor directly influence it. That is one of the reasons why the dynamics of labor should be included in the model.

For the constant labor growth rate complex dynamics caused by the shape of the production function has already been studied (see [7] or [4]). This paper shows that labor dynamics should be taken into account, since it influences the capital dynamics in an essential way (comparing mentioned [7] or [4] examples with kinked production function it is a different way, complex dynamics may take place even in the case that

is stable if constant labor growth rate is assumed). Usually saving rate is crucial for the stabilization of the capital dynamics. As we have seen in the Corollaries 3 and 4 even complex behaviour may arise in case that shareholders saving rate exceeds the workers saving rate, moreover it does not depend on how high or low the rates are, so for the stabilization policy we should differentiate them. There may be also cases of instabilities (cycles or chaos) such that saving rates cannot influence the dynamics at all, since it is based on the behaviour of the labor. The stabilization policy should influence the labor market in such case instead of interest rate change.

References

[1] Accinelli, E., and Brida, J.G.: Re-formulation of the Solow economic growth model whit the Richards population growth law, *Economics Working Paper Archive* **0508006** (2005), http://econwpa.wustl.edu/listings/ge/0510.html.

[2] Barro, R.J., and Sala-i-Martin, X.: *Economic growth*. McGraw-Hill, Inc., 1995.

[3] Beverton, R. J. H., Holt, S. J.: On the dynamics of exploited fish populations, *Fishery Investigations* **19** (1957), 1–533.

[4] Böhm, V., and Kaas, L.: Differential savings, factor shares, and endogenous growth cycles, *Journal of Economic Dynamics & Control*, **24** (2000), 965–980.

[5] Brianzoni, S., Mammana, C., and Michetti, E.: Complex Dynamics in the Neoclassical Growth Model with Differential Savings and Non-Constant Labor Force Growth, *Studies in Nonlinear Dynamics & Econometrics* **11**, issue 3 (2007), 1407–1407.

[6] Brianzoni, S., Mammana, C., and Michetti, E.: Nonlinear dynamics in a business-cycle model with logistic population growth, *Chaos, Solitons & Fractals* **40** (2009), 717-730.

[7] Day, R.: *Complex Economic Dynamics II*. MIT Press, 2000.

[8] Kaldor, N.: Alternative theories of distribution, *Review of Economic Studies* **23** (1956), 83–100.

[9] Kaldor, N.: A model of economic growth, *Economic Journal* **67** (1957), 591–624.

[10] Pasinetti, L. L.: Rate of profit and income distribution in relation to the rate of economic growth, *Review of Economic Studies* **29** (1962), 267–279.

[11] Smith, J.M.: *Models in ecology*. Cambridge University Press, Cambridge, 1974.

[12] Solow, R. M.: A contribution to the theory of economic growth, *Quarterly Journal of Economics* **70** (1956), 65–94.

[13] Swan, T. W.: Economic growth and capital accumulation, *Economic Record* **32** (1956), 334–361.

[14] Verhulst, P.F.: Notice sur la loi que la population suit dans son accroisement, *Correspondences Mathematiques et Physiques*, **10** (1838), 113–121.

Appendix A

Collection of author's work, published articles, posters and preprints related to covid-19 epidemic modeling

Articles in international journals:

- Martin Šmíd, et al. Protection by vaccines and previous infection against the Omicron variant of SARS-CoV-2, The Journal of Infectious Diseases, jiac161, **2022**.

- Luděk Berec, et al. Protection provided by vaccination, booster doses and previous infection against covid-19 infection, hospitalisation or death over time in Czechia, PloS ONE 17.7 : e0270801 **2022**.

Real-time model of covid-19 epidemics in the Czech Republic (software):

- Tomáš Pavlík, et al. MAMES – Monitoring, analysis, and management of epidemic situations
  https://webstudio.shinyapps.io/MAMES/, **2021**.

- Lenka Přibylová and Veronika Eclerová. Model ZSEIAR(S) forecasts submitted to ECDC Forecast Hub (MUNI_DMS-SEIAR), **2021-2022**

Conferences and posters:

- Lenka Přibylová and Veronika Eclerová. Model ZSEIAR, In: V4-UK Frontiers of Science 2021 (Royal Society UK), **2021   invited**

- Veronika Eclerová and Lenka Přibylová. Ascertainment rate estimate from hospital data used in modeling COVID-19 epidemics, In: Epidemics8, **2021**

Preprints:

- Lenka Přibylová and Veronika Eclerová. SEIAR model with asymptomatic cohort and consequences to efficiency of quarantine government measures in COVID-19 epidemic. arXiv preprint arXiv:2004.02601. **2020**   Google Scholar citations: 14

- Katharine Sherratt, et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. medRxiv **2022**.

Protection by Vaccines and Previous Infection Against the Omicron Variant of Severe Acute Respiratory Syndrome Coronavirus 2

Martin Šmíd,[1,2,] Luděk Berec,[2,3,4] Lenka Přibylová,[5] Ondřej Májek,[6,7] Tomáš Pavlík,[6,7] Jiří Jarkovský,[6,7] Jakub Weiner,[1,2] Tamara Barusová,[8,9] and Jan Trnka[10]

[1]Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic; [2]Centre for Modelling of Biological and Social Processes, Prague, Czech Republic; [3]Centre for Mathematical Biology, Institute of Mathematics, Faculty of Science, University of South Bohemia, Budjovice, Czech Republic; [4]Biology Centre, Institute of Entomology, Department of Ecology, Czech Academy of Sciences, Budjovice, Czech Republic; [5]Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic; [6]Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic; [7]Department of Data Analysis, Institute of Health Information and Statistics of the Czech Republic, Prague, Czech Republic; [8]First Faculty of Medicine, Charles University, Prague, Czech Republic; [9]Department of Statistical Modelling, Czech Academy of Sciences, Institute of Computer Science, Prague, Czech Republic; and [10]Department of Biochemistry, Cell and Molecular Biology, Third Faculty of Medicine, Charles University, Prague, Czech Republic

Background. The Omicron variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) evades immunity conferred by vaccines and previous infections.

Methods. We used a Cox proportional hazards model and a logistic regression on individual-level population-wide data from the Czech Republic to estimate risks of infection and hospitalization, including severe states.

Results. A recent ($\leq$2 months) full vaccination reached vaccine effectiveness (VE) of 43% (95% confidence interval [CI], 42%–44%) against infection by Omicron compared to 73% (95% CI, 72%–74%) against Delta. A recent booster increased VE to 56% (95% CI, 55%–56%) against Omicron infection compared to 90% (95% CI, 90%–91%) for Delta. The VE against Omicron hospitalization of a recent full vaccination was 45% (95% 95% CI, 29%–57%), with a recent booster 87% (95% CI, 84%–88%). The VE against the need for oxygen therapy due to Omicron was 57% (95% CI, 32%–72%) for recent vaccination, 90% (95% CI, 87%–92%) for a recent booster. Postinfection protection against Omicron hospitalization declined from 68% (95% CI, 68%–69%) at $\leq$6 months to 13% (95% CI, 11%–14%) at >6 months after a previous infection. The odds ratios for Omicron relative to Delta were 0.36 (95% CI, .34–.38) for hospitalization, 0.24 (95% CI, .22–.26) for oxygen, and 0.24 (95% CI, .21–.28) for intensive care unit admission.

Conclusions. Recent vaccination still brings substantial protection against severe outcome for Omicron.

Keywords. COVID-19; postinfection immunity; vaccine effectiveness; SARS-CoV-2; Omicron variant; hospitalization.

The B.1.1.529 (Omicron) variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) was first detected in South Africa in November 2021, immediately designated a variant of concern by the World Health Organization [1], and thereafter seen to spread quickly throughout most of the world. This rapid spread was at least in part brought about by a degree of immune evasion due to a large number of mutations in the viral S-protein, which led to changes in epitopes recognized by antibodies elicited by vaccination or previous infection [2]. Together with nonpharmacological interventions, such as face masks, distancing, ventilation of interior spaces testing, and isolating, vaccination is among the most effective means of individual and collective protection from the impacts of the pandemic. The immune evasion by the Omicron variant thus caused concern and led to much interest in both laboratory and real-life epidemiological data that could accurately measure this phenomenon.

Since 27 December 2020 the inhabitants of the Czech Republic have been receiving coronavirus disease 2019 (COVID-19) vaccines, the largest number vaccinated with the messenger RNA (mRNA) vaccine BNT162b2 (Pfizer/BioNTech), followed by mRNA-1273 (Moderna) and the adenovirus-based vector vaccines ChAdOx1 nCoV-19 (AstraZeneca) and Ad26.CoV2.S (Johnson & Johnson) [3]. By the end of our study period on 13 February 2022, 68% of the population had a complete vaccination and 39% had received a booster dose [3].

The first case of the Omicron variant in the Czech Republic was detected at the end of November 2021; its proportion of recorded cases rapidly rose and by 10 January 2022 it became the dominant variant (Figure 1). An increasing number of infections among fully vaccinated and reinfections indeed suggests that immune evasion poses a significant risk to further COVID-19 development [3].
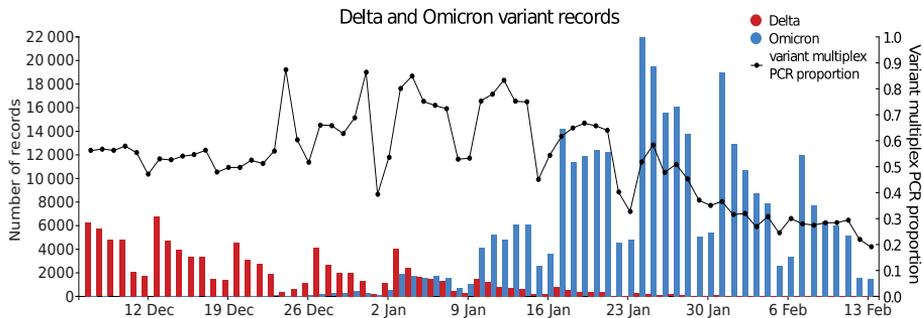
Figure 1.   Number of recorded cases with assigned Delta and Omicron variant and the proportion of polymerase chain reaction (PCR)–positive tests tested for viral variants using multiplex PCR.

In this study, we estimate how the protection due to vaccination or previous SARS-CoV-2 infection against COVID-19 infection, hospital admission, oxygen therapy and intensive care unit (ICU) admission varies in relation to the virus variant and time elapsed for the entire population of the Czech Republic.

METHODS

Study Population and Data Sources

The analyses are based on data from the Czech National Information System of Infectious Diseases (ISID), which includes records of all individuals who tested positive for SARS-CoV-2 in the Czech Republic since the beginning of the COVID-19 pandemic, including children [4]. This database is overseen by the Czech Ministry of Health and operated by the Institute of Health Information and Statistics of the Czech Republic. Data are routinely collected in compliance with Czech legal regulations (Act on the Protection of Public Health). The Director of the Institute of Health Information and Statistics of the Czech Republic has granted that there is no need for ethical approval of the retrospective analyses presented in this article.

The ISID database collects demographic data (age, sex and region of residence), dates of vaccination, including the vaccine types for each dose, and dates of infection and potential reinfection, hospitalization including treatment type, and the date of potential death with COVID-19. The data recorded in the study period include information on whether the infection is caused by the Omicron, Delta, or some other variant, or that a variant discrimination was not performed (Figure 1). The information on the variant is based on results of multiplex polymerase chain reaction (PCR) or viral genome sequencing, which are available only for a subset of all PCR-positive cases. The variants were identified using the definition of viral S-protein mutations according to the European Centre for Disease Prevention and Control [5]; the algorithm was tailored to multiplex PCR kits used in the Czech Republic in collaboration with the

National Institute of Public Health and the National Reference Laboratory [6]. Additional information on deaths from any cause comes from the Death Certificate System; these data are used for censoring purposes only.

Study Endpoints

We studied 4 types of events: (1) SARS-CoV-2 infection, defined as a PCR-confirmed positive test of any type of sample regardless of the presence of symptoms; (2) hospital admission of a person who tested positive on a PCR test within 2 weeks after the confirmed infection or earlier; (3) use of any type of oxygen therapy (nasal oxygen, noninvasive ventilation, invasive mechanical ventilation, high-flow nasal oxygen, and extracorporeal membrane oxygenation); and (4) admission to ICU during the hospitalization. All events were related to the date of infection report.

We examined events during the 2-month period from 7 December 2021 to 13 February 2022, during which Delta and Omicron switched dominance in the Czech Republic (Figure 1).

Statistical Analysis

A Cox regression with time-varying covariates was applied to estimate hazard ratios (HRs) for the outcomes of interest separately for each viral variant. In these analyses, the infections by the variant other than the examined one and the infections lacking variant assignment were censored at the time of infection. Analogously to Tartof et al [7], we used calendar time instead of time from event occurrence as the time scale. Thus, the time course of individual cases was modeled by means of "switching" dummy variables, corresponding to the development of the immune status after vaccination or past infection in 61-day periods for vaccination and 121-day periods for the time from the last infection. The control variables included age group and sex.

The protection provided by vaccine (vaccine effectiveness [VE]) or previous infection is calculated by comparing hazards of the vaccinated and/or immunized individuals to those of the
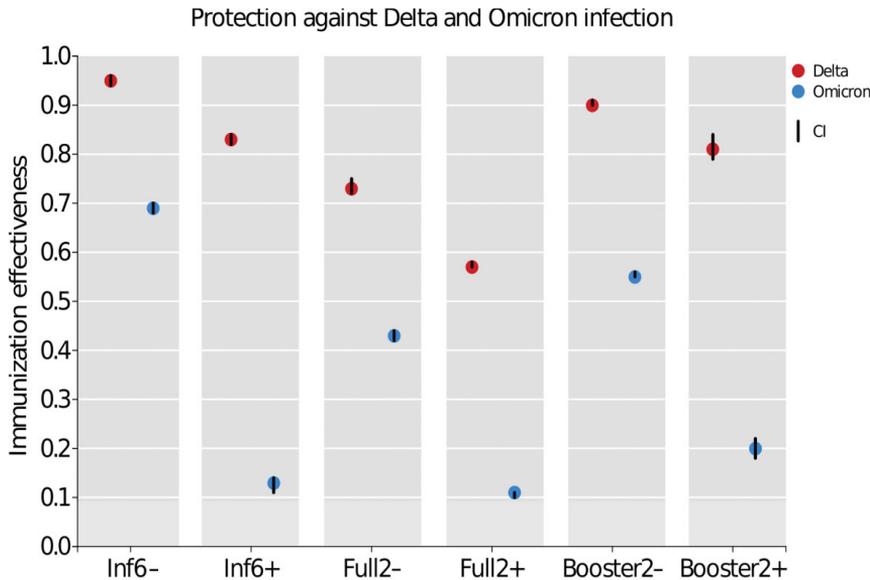
Protection against Delta and Omicron infection

Figure 2. Protection provided by vaccination or previous infection against infection by the Omicron and Delta variants of the severe acute respiratory syndrome coronavirus 2. Point estimates of protection with 95% confidence interval are shown. Abbreviations: Booster2−, booster dose ≤2 months ago; Booster2+, booster dose >2 months ago; CI, confidence interval; Full2−, complete vaccination ≤2 months ago; Full2+, complete vaccination >2 months ago; Inf6−, previous infection ≤6 months ago; Inf6+, previous infection >6 months ago.

"control group"—those who have not been vaccinated and infected so far and subtracted from 1 using the equation:

$$\text{Protection(VE)} = 1 - \frac{\text{Hazard}_{\text{protected}}}{\text{Hazard}_{\text{unprotected}}}.$$

Furthermore, we examine the postinfection immunity by estimating HRs of infection of previously unvaccinated individuals in relation to time elapsed from the infection.

By using calendar time we were able to incorporate automatically the changing conditions of the epidemic, including nonpharmacological measures, seasonal effects, the ratio of discriminated samples, and the proportion of the virus variant, as all of these phenomena can be included in the underlying baseline hazard function.

To examine the probabilities of hospitalization, oxygen therapy, and ICU admission for an infected individual, we use the logistic regression with the event of interest as the outcome and with immunity status at the time of infection, age group, and sex as the covariates. We compare the probabilities of the outcome for both variants by means of the dummy corresponding to the virus variant.

All calculations were performed using the R software. The algorithm used to transform data from the database into the package command inputs was coded in C++. See Supplementary Material 1 for details.

RESULTS

Protection Against Infection

First we looked at the protection conferred by vaccination or a previous infection against a new infection, since the protection against infection represents the potential to protect other risk groups in the population. The protection after vaccination against the Omicron variant reached 43% (95% confidence interval [CI], 42%–44%) shortly after completing the full vaccination scheme, falling to 9% (95% CI, 8%–10%) after >2 months. This protection increased to 56% (95% CI, 55%–56%) shortly after receiving a booster dose, followed by a decline to 21% (95% CI, 19%–23%) after >2 months. These numbers strongly contrast with the protection against the Delta variant, which was consistently higher at 73% (95% CI, 72%–74%), 57% (95% CI, 56%–58%), 90% (95% CI, 90%–91%), and 82% (95% CI, 79%–84%), respectively. Similar degrees of protection against infection are conferred also by postinfection immunity: 68% (95% CI, 68%–69%) shortly after a previous infection (2–6 months; a positive test during the first 2 months after an infection is not considered a reinfection by definition) and 13% (95% CI, 11%–14%) after 6 months for Omicron, compared with 95% (95% CI, 94%–96%) shortly after infection and 83% (95% CI, 82%–84%) after 6 months for Delta (Figure 2). Based on the past prevalence of viral variants, it can be expected

Table 1.   Protection Due to Various Combinations of Past Infection Preceding Vaccination Against Infection for the Omicron and the Delta Variants of the Severe Acute Respiratory Syndrome Coronavirus 2

		Vaccination			
VOC	Infection	Booster2−	Full2−	Booster2+	Full2+
Omicron	Inf6−	92% (89%–94%)	82% (75%–87%)	82% (72%–89%)	86% (85%–88%)
	Inf6+	74% (73%–75%)	77% (76%–78%)	48% (45%–52%)	45% (44%–46%)
Delta	Inf6−	95% (66%–99%)	100% (no case)	100% (no case)	97% (94%–98%)
	Inf6+	98% (98%–99%)	98% (97%–98%)	94% (89%–97%)	96% (95%–96%)

Data show protection by vaccination following past infection (95% confidence interval).

Abbreviations: Booster2−, booster dose ≤2 months ago; Booster2+, booster dose >2 months ago; Full2−, complete vaccination ≤2 months ago; Full2+, complete vaccination >2 months ago; Inf6−, previous infection ≤6 months ago; Inf6+, previous infection >6 months ago; VOC, variant of concern.

that the infections older than 6 months were mostly due to the original Wuhan, D614G, and Alpha variants, whereas the more recent ones were predominantly due to Delta. As we show in Supplementary Material 2, Sections 11 and 12, explicit accounting for the vaccine type (BNT162b2 by Pfizer/BioNTech and mRNA-1273 by Moderna) gave values of effectiveness comparable with the analyses of pooled data reported here in the main text.

We had enough data to examine all the combinations in which a previous infection preceded vaccination. As expected, protection declined with time elapsed from the previous infection or vaccination (Table 1). Regarding protection against the Delta variant, any combination provided ≥95% protection against infection (Table 1). This protection also remained quite high against Omicron when the previous infection was recent, falling to lower values for an older previous infection, but even then the protection was significantly higher than that provided by a vaccination or previous infection alone (Table 1). We also analyzed cases when a vaccination preceded an infection followed by a reinfection. In the case of reinfections caused by Delta, against which the achieved protection was generally high at 96% (95% CI, 90%–98%), the exact order of events did not appear to matter. For reinfections caused by Omicron, against which protection is generally lower, the cases where a previous infection followed a vaccination appeared to provide a higher level of protection than the inverse sequence:

Protection provided by the complete vaccination >2 months ago/previous infection ≤6 months ago combination was 89% (95% CI, 88%–91%) as compared to 86% (95% CI, 85%–88%) for the previous infection ≤6 months ago/complete vaccination >2 months ago combination.

A finer-grained analysis of temporal dynamics of immunity waning after a previous infection was then conducted specifically for individuals who were previously infected but remained nonvaccinated. Against Omicron, the protection was estimated as 69% (95% CI, 68%–69%) for 2–6 months after previous infection, 48% (95% CI, 46%–50%) for 7–10 months, 34% (95% CI, 33%–35%) for 11–14 months, and 17% (95% CI, 15%–18%) for ≥14 months after previous infection. For Delta, in contrast, these numbers were 93% (95% CI, 91%–94%), 91% (95% CI, 90%–92%), 86% (95% CI, 85%–86%), and 79% (95% CI, 77%–81%), respectively.

Protection Against Hospitalization

A qualitatively similar pattern yet quantitatively consistently higher protection is seen against hospitalization, a need for oxygen therapy, and a need for intensive care (Table 2). For example, a recent booster dose provides 86% protection against hospitalization, 90% against a need for oxygen therapy, and 83% against a need for intensive care when infected by the Omicron variant. Moreover, all combinations of previous

Table 2.   Vaccine Effectiveness and Protection Provided by Postinfection Immunity Against a Need for Hospitalization, Oxygen Therapy, or Intensive Care, for the Omicron and Delta Variants of the Severe Acute Respiratory Syndrome Coronavirus 2

Vaccination or Infection	Hospitalization		Oxygen Therapy		Intensive Care	
	Omicron	Delta	Omicron	Delta	Omicron	Delta
Full2−	45% (29%–57%)	73% (69%–76%)	57% (32%–72%)	82% (76%–87%)	58% (3%–82%)	84% (72%–91%)
Full2+	29% (21%–37%)	77% (76%–79%)	32% (20%–43%)	82% (80%–83%)	37% (12%–55%)	86% (83%–88%)
Booster2−	86% (84%–88%)	97% (97%–98%)	90% (87%–92%)	98% (98%–98%)	83% (75%–89%)	98% (97%–99%)
Booster2+	79% (75%–82%)	96% (94%–97%)	85% (80%–88%)	97% (95%–98%)	60% (37%–74%)	97% (92%–99%)
Inf6−	73% (55%–84%)	100% (no case)	81% (40%–94%)	100% (no case)	83% (0–98%)	100% (no case)
Inf6+	66% (54%–75%)	94% (91%–96%)	88% (72%–94%)	98% (95%–99%)	66% (15%–86%)	97% (90%–99%)

Data show vaccine effectiveness or protection by postinfection immunity (95% confidence interval) against need for hospitalization, oxygen therapy, or intensive care.

Abbreviations: Booster2−, booster dose ≤2 months ago; Booster2+, booster dose >2 months ago; Full2−, complete vaccination ≤2 months ago; Full2+, complete vaccination >2 months ago; Inf6−, previous infection ≤6 months ago; Inf6+, previous infection >6 months ago.

Table 3. Protection Due to Various Combinations of Past Infection Preceding Vaccination Against Hospitalization for the Omicron and the Delta Variants of the Severe Acute Respiratory Syndrome Coronavirus 2

VOC	Infection	Vaccination			
		Booster2−	Full2−	Booster2+	Full2+
Omicron	Inf6−	100% (no case)	100% (no case)	71% (0–96%)	93% (49%–99%)
	Inf6+	95% (78%–99%)	94% (77%–95%)	90% (64%–98%)	73% (78%–99%)
Delta	Inf6−	100% (no case)	100% (no case)	100% (no case)	100% (no case)
	Inf6+	99% (99%–100%)	97% (91%–99%)	98% (85%–100%)	98% (98%–100%)

Data show protection by vaccination following past infection (95% confidence interval).

Abbreviations: Booster2−, booster dose ≤2 months ago; Booster2+, booster dose >2 months ago; Full2−, complete vaccination ≤2 months ago; Full2+, complete vaccination >2 months ago; Inf6−, previous infection ≤6 months ago; Inf6+, previous infection >6 months ago; VOC, variant of concern.

infection and recent vaccination present in our data appear to provide nearly complete protection against Omicron as regards hospitalization (Table 3) as well as oxygen therapy or intensive care (often no cases have been observed for such situations; see Supplementary Material 2, Sections 7–10).

Risk of a Severe Outcome for Omicron vs Delta

Finally, our logistic regression analyses show that once infected, the odds ratio is 0.36 (95% CI, .34–.38) for hospitalization with Omicron relative to Delta; 0.24 (95% CI, .22–.26) for a need of oxygen therapy with Omicron relative to Delta; and 0.24 (95% CI, .21–.28) for a need of intensive care with Omicron relative to Delta. Moreover, once hospitalized, the odds ratio is 0.44 (95% CI, .39–.49) for a need of oxygen therapy with Omicron relative to Delta, and 0.64 (95% CI, .52–.72) for a need of intensive care with Omicron relative to Delta (see Supplementary Material 2, Sections 15–19, for further details).

DISCUSSION

Our data support the existing evidence that the Omicron variant of SARS-CoV-2, to a significant extent, evades both postvaccination and postinfection immunity [2, 8–11]. The VE levels of all the vaccines used in the Czech Republic are lower for Omicron compared to Delta. As we previously observed with Alpha and Delta [12], the protection against infection by the Omicron variant also wanes over time. However, a booster vaccine dose provides robust and lasting, or slowly waning, protection against hospitalization, the need for oxygen therapy, and intensive care. The combined postinfection and postvaccination immunity is the most protective regardless of the exact sequence of events, suggesting that the best protective strategy before a coming wave is to vaccinate all individuals, whether previously vaccinated or with a previous COVID-19 infection.

We are aware of the complicated interpretation of the hospitalization data for the Omicron wave: The very high basic reproduction number (R_0) of this variant [13] translated into the very high prevalence of infection in the population at the peak of the epidemic wave; and a much higher proportion of hospitalized patients with COVID-19 as a concomitant finding rather than the reason for admission. We therefore analyzed separately the need for oxygen therapy and ICU admission as a more relevant measure of severe outcomes due to the Omicron infection.

Compared to the Delta variant, the protection provided by the postinfection or postvaccination immunity is lower against the Omicron variant, but at the same time the Omicron variant appears less severe than the Delta variant and the odds ratio for oxygen therapy or ICU admission both approximately equal about one-quarter compared to the Delta variant.

A common limitation of studies like ours is the fact that only a certain proportion of infections is reported (ascertainment rate). We believe this phenomenon does not significantly affect our estimates of VE, assuming that the ascertainment rate is the same for the vaccinated and the unvaccinated alike and we have no evidence to the contrary. A potentially low ascertainment rate could also distort our estimates of the protection by the postinfection immunity; in particular, if there had been many undetected individuals with postinfection immunity in the control group, the infection risk of the virgin population would have been underestimated and, consequently, the protection by infection underestimated as well. Our results should be interpreted in terms of reported infections only.

In all of our analyses we used age and sex as control variables; however, with some caution they can also be understood as risk factors. In this respect, our results generally confirm the common knowledge that the risk of various severe outcomes grows exponentially with the person's age – this is clearly illustrated by the linear increase of log-HRs for both variants (see Supplementary Material 2, Sections 5–19). The age-related risk of (re-)infection, on the other hand, appears to be the highest for children and people in the working age. This pattern is more pronounced for the Omicron variant. However, it is not clear to what extent the pattern is caused by behavioral causes and/or the current epidemic situation rather than biological causes.

Supplementary Data

Supplementary materials are available at *The Journal of Infectious Diseases* online (http://jid.oxfordjournals.org/). Supplementary materials consist of data provided by the author

that are published to benefit the reader. The posted materials are not copyedited. The contents of all supplementary data are the sole responsibility of the authors. Questions or messages regarding errors should be addressed to the author.

Notes

Data sharing. Data reported in this study and used for the analyses are not public. De-identified individual-level data are available to the scientific community. Requests, together with a short description of their analysis proposals, should be submitted to the Institute of Health Information and Statistics of the Czech Republic (www.uzis.cz/index-en.php), where they will be assessed based on relevance and scientific merit.

Financial support. No external funding was used to conduct this study.

Potential conflicts of interest. The authors: No reported conflicts of interest.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. World Health Organization. Classification of Omicron (B.1.1.529): SARS-CoV-2 variant of concern. **2021**. https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern. Accessed 20 February 2022.
2. McCallum M, Czudnochowski N, Rosen LE, et al. Structural basis of SARS-CoV-2 Omicron immune evasion and receptor engagement. Science **2022**; 375:864–8.
3. Ministry of Health of the Czech Republic. COVID-19: an overview of the actual situation in the Czech Republic [in Czech]. **2020**. onemocneni-aktualne.mzcr.cz/covid-19. Accessed 20 February 2022.
4. Komenda M, Bulhart V, Karolyi M, et al. Complex reporting of the COVID-19 epidemic in the Czech Republic: use of an interactive web-based app in practice. J Med Internet Res **2020**; 22:e19367.
5. European Centre for Disease Prevention and Control. SARS-CoV-2 variants of concern. **2022**. https://www.ecdc.europa.eu/en/covid-19/variants-concern. Accessed 20 February 2022.
6. The Czech National Institute of Public Health. Charakterizace viru SARS-CoV-2 v České Republice dle diskriminačních PCR a celogenomové sekvenace [in Czech]. **2022**. http://www.szu.cz/uploads/Epidemiologie/Coronavirus/WGS_covid/2022_tydenni_hlaseni/SARS_CoV_2_podrobna_zprava_SZU_2022_01_31.pdf.
7. Tartof S, Slezak J, Fischer H, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet **2021**; 398:1407–16.
8. Dejnirattisai W, Huo J, Zhou D, et al. SARS-CoV-2 Omicron-B.1.1.529 leads to widespread escape from neutralizing antibody responses. Cell **2022**; 185:467–84.e15.
9. Hoffmann M, Krüger N, Schulz S, et al. The Omicron variant is highly resistant against antibody-mediated neutralization: implications for control of the COVID-19 pandemic. Cell **2022**; 185:447–56.e11.
10. Cui Z, Liu P, Wang N, et al. Structural and functional characterizations of infectivity and immune evasion of SARS-CoV-2 Omicron. Cell **2022**;185:860–71.e13.
11. Cao Y, Wang J, Jian F, et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. Nature **2021**; 602:657–63.
12. Berec L, Šmíd M, Přibylová L, et al. Real-life protection provided by vaccination, booster doses and previous infection against covid-19 infection, hospitalisation or death over time in the Czech Republic: a whole country retrospective view. medRxiv [Preprint]. Posted online 12 December **2021**. doi:10.1101/2021.12.10.21267590.
13. Nishiura H, Ito K, Anzai A, Kobayashi T, Piantham C, Rodríguez-Morales AJ. Relative reproduction number of SARS-CoV-2 Omicron (B.1.1.529) compared with Delta variant in South Africa. J Clin Med **2022**; 11:30.

PLOS ONE

RESEARCH ARTICLE

Protection provided by vaccination, booster doses and previous infection against covid-19 infection, hospitalisation or death over time in Czechia

Luděk Berec[1,2,3]*, **Martin Šmíd**[3,4], **Lenka Přibylová**[5], **Ondřej Májek**[6,7], **Tomáš Pavlík**[6,7], **Jiří Jarkovský**[6,7], **Milan Zajíček**[4], **Jakub Weiner**[3,8], **Tamara Barusová**[9,10], **Jan Trnka**[11]

1 Centre for Mathematical Biology, Institute of Mathematics, Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic, **2** Department of Ecology, Czech Academy of Sciences, Biology Centre, Institute of Entomology, České Budějovice, Czech Republic, **3** Centre for Modelling of Biological and Social Processes, Praha, Czech Republic, **4** Czech Academy of Sciences, Institute of Information Theory and Automation, Praha, Czech Republic, **5** Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic, **6** Institute of Biostatistics and Analyses, Faculty of Medicine, Masaryk University, Brno, Czech Republic, **7** Institute of Health Information and Statistics of the Czech Republic, Praha, Czech Republic, **8** Siesta Labs, Praha, Czech Republic, **9** First Faculty of Medicine, Charles University, Praha, Czech Republic, **10** Department of Statistical Modelling, Czech Academy of Sciences, Institute of Computer Science, Praha, Czech Republic, **11** Department of Biochemistry, Cell and Molecular Praha, Czech Republic

* lberec@prf.jcu.cz

Abstract

Studies demonstrating the waning of post-vaccination and post-infection immunity against covid-19 generally analyzed a limited range of vaccines or subsets of populations. Using Czech national health data from the beginning of the covid-19 pandemic till November 20, 2021 we estimated the risks of reinfection, breakthrough infection, hospitalization and death by a Cox regression adjusted for sex, age, vaccine type and vaccination status. Vaccine effectiveness against infection declined from 87% at 0-2 months after the second dose to 53% at 7-8 months for BNT162b2 vaccine, from 90% at 0-2 months to 65% at 7-8 months for mRNA-1273, and from 83% at 0-2 months to 55% at 5-6 months for the ChAdOx1-S. Effectiveness against hospitalization and deaths declined by about 15% and 10%, respectively, during the first 6-8 months. Boosters (third dose) returned the protection to the levels observed shortly after dose 2. In unvaccinated, previously infected individuals the protection against infection declined from 97% after 2 months to 72% at 18 months. Our results confirm the waning of vaccination-induced immunity against infection and a smaller decline in the protection against hospitalization and death. Boosting restores the original vaccine effectiveness. Post-infection immunity also decreases over time.

Introduction

The availability of vaccines brought about a breakthrough in the fight against the coronavirus disease 2019 (covid-19) worldwide. In the light of the economic and social costs already caused by covid-19, and the widespread aversion to any serious limitations of people's daily lives due to lockdowns, vaccination is undoubtedly a key tool for the containment of the pandemic and for the limiting of its devastating impact on lives and health of people around the globe. Proving in the clinical studies and the weeks and months of their real-world application their high effectiveness against SARS-CoV-2 infection, symptomatic covid-19 illness, need of a hospital admission, the probability of severe symptoms, and death [1–4], their continued impact now starts to be challenged by an increasing proportion of breakthrough infections and illnesses in fully vaccinated individuals [5, 6].

In the Czech Republic, vaccination started on December 27, 2020 initially with the mRNA-based vaccine BNT162b2 (Pfizer/BioNTech), followed by mRNA-1273 (Moderna), and the adenovirus-based vector vaccines ChAdOx1-S (AstraZeneca) and Ad26.COV2-S (Johnson&Johnson). The administration of booster doses then started on September 20, 2021 and was initially open to all individuals who completed their vaccination 8 months or longer ago with only BNT162b2 and mRNA-1273 as allowed boosting vaccines. The waiting period was shortened to 6 months on October 29, 2021. Until November 1, 2021 a full 100 µg dose of Moderna was being administered as a booster dose and since that day only a 50 µg dose was used. We emphasize that by complete vaccination we mean two doses of vaccine (and just one for Ad26. COV2-S) and by booster we mean the third vaccine dose.

Central Europe experienced another wave of SARS-CoV-2 infections in the autumn of 2021 despite the substantial proportion of vaccinated and/or recovered population. This wave was accompanied by a non-negligible proportion of new infections in vaccinated individuals including the need for hospital admission and, in a relatively few cases, for intensive care. The appearance of breakthrough infections, though not unexpected, has complicated the public health messaging related to the importance of vaccination and calls for a better understanding of the temporal dynamics of post-vaccination immunity in real-world settings. Post-infection immunity is another important factor determining individual risk. SARS-CoV-2 reinfections have been reported as relatively rare events, yet the post-infection immunity appears to wane, too [7–10]. None of these studies addressed longer-term dynamics of post-infection immunity and their relationship to post-vaccination immunity.

Materials and methods

Study population and data sources

The analyses are based on data from the Czech National Information System of Infectious Diseases (ISID), which includes records of all individuals tested positive for SARS-CoV-2 in the Czech Republic since the beginning of covid-19 pandemic [11]. This database is overseen by the Czech Ministry of Health and operated by the Institute of Health Information and Statistics of the Czech Republic. The ISID data is routinely collected in compliance with Czech legal regulations (Act on the Protection of Public Health). The Director of the Institute of Health Information and Statistics of the Czech Republic has granted that there is no need for ethical approval of the retrospective analyses presented in this paper. Among other things, the ISID database covers demographic data, dates of vaccination, including the vaccine types for each dose, and dates of infection and potential reinfection, including information on dates of hospital admission with covid-19, and death with covid-19. Additional information on deaths from

any cause come from the Death Certificate System; these data are used for censoring purposes only.

In total, our dataset contains 7,428,968 valid records of vaccinated and/or SARS-CoV-2 positive persons (additional 8,834 cases lack information on sex or age and 216 other cases contain data errors, see S2 Table in the S2 File). We further excluded 16,399 persons who were recorded to die by the start of vaccination (December 26, 2020). As the source dataset consists only of those who were tested positive and/or were vaccinated, we completed the sample to the whole population such that the added subjects were neither tested positive nor vaccinated. In particular, we completed each sex-age category to the numbers reported by the Czech Statistical Office by December 31, 2020—10,701,777 inhabitants; consequently, our sample truly reflected the sex and age structure of the whole population, containing all the positive and/or vaccinated individuals. We neglected births and deaths of the added persons.

Vaccine types and vaccination and infection dynamics

In the Czech Republic, all EMA-approved Covid-19 vaccines have been distributed and used. They were provided to all individuals at no cost following the Czech public health insurance system. Starting on December 27, 2020, workers in the critical infrastructure were vaccinated first, followed since January 15, 2021 by persons of age 80 and older (S1 Table in S2 File). As of November 20, 2021, the national Institute of Health Information and Statistics reported 6,287,356 individuals completing the vaccination (58.75% of the population and 67.36% of persons of age 12 years and older); see Fig 1.

The covid-19 epidemic in the Czech Republic started with the first three cases reported on March 1, 2020 and was initially fueled by Czech citizens returning from the alpine ski resorts of Italy and Austria. Since then the country saw five waves of covid-19 spread. As of November 20, 2021, 1,996,080 individuals were infected with SARS-CoV-2 virus, of which 12,894 (0.65%) were reinfected; see S1 Fig in S2 File. See also S2 and S3 Tables and the accompanying S2 and S3 Figs in S2 File for an overview of the numbers of infection-related outcomes and vaccines applied within various age cohorts.

Statistical analysis

We separately studied three types of events: (i) SARS-CoV-2 infection defined as a PCR confirmed positive test of a person from any sample regardless of the presence of symptoms, (ii) hospital admission of a person tested positive via a PCR test (within two weeks before hospital admission and whenever during hospitalization), and (iii) death due to covid-19.

A Cox regression with time-varying covariates was applied to estimate hazard ratios (HRs) for the outcomes of interest. Analogously to [12], we used calendar time instead of time from event occurrence as the time scale. Thus, the time course of individual cases was modelled using "switching" dummy variables, corresponding to the stages of the process the subject goes through. The vaccine effectiveness is calculated by comparing hazards of the vaccinated individuals to those of the control group—those who have not been vaccinated and infected so far—individually for each vaccine type [12]. By using calendar time, we could control for changing epidemic conditions, including non-pharmaceutical measures, seasonal effects and viral variants; these phenomena can then be encompassed in the baseline hazard function. Subjects were withdrawn from the study at the time of their (covid or non-covid) death.

Time zero corresponded to the day before the start of vaccination (26 December 2020) for the analyses of vaccine-induced immunity, and the onset of epidemic in the Czech Republic for the analyses of infection-induced immunity. Moreover, we estimated how HRs of infection after vaccination depended on time after the vaccine application (adjusted for sex, age and
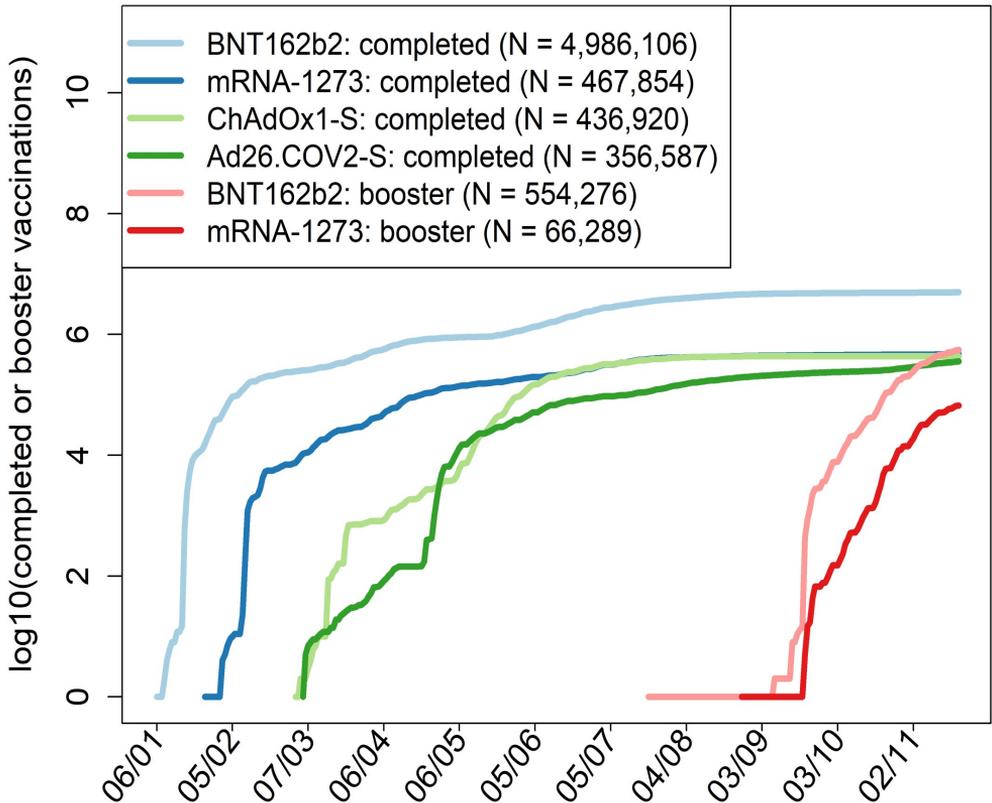
Fig 1. Dynamics of vaccination in the Czech Republic. Specific days in which vaccination was open to an age group or professional or other category are specified in S1 Table in S2 File.

https://doi.org/10.1371/journal.pone.0270801.g001

time since the last infection), how HRs of hospital admission or death depended on time after the vaccine application (adjusted for sex and age), and how HRs of reinfection in unvaccinated individiuals depended on time since the previous infection (adjusted for sex and age). In all these cases, we estimated the vaccine effectiveness (VE, regarding a previous infection as a "vaccine") as $VE = 1 − HR$ [12–14].

We aggregate the time delays in two-month (61 days) periods. We consider one such period after the first dose, four periods after the second one and a single period after the booster dose. When a new dose is applied to a person, (s)he is no longer regarded to be in any period corresponding to the previous dose, but enters the first period corresponding to the new dose. In line with the Czech vaccination recognition policy, the first period corresponding to any of the first two doses starts two weeks after the dose application, while for boosters this interval is just 7 days. For the reinfections, we consider nine two-month periods.

We examine boosting effects by the mRNA vaccines BNT162b2 and mRNA-1273. Since 93% of BNT162b2 boosters is preceded by the BNT162b2 second dose and 73% of mRNA-1273 boosters is preceded by the mRNA-1273 second dose, but the vaccine type used for the first two doses does not play a role in which type is applied as the booster, we estimate the

boosting effects of the mRNA vaccines both with and without individual vaccination history. Subjects with the alleged application of ChAdOx1-S (in total 176) or Ad26.COV2-S (in total 332) boosters are withdrawn from the study of booster dose effectiveness, as these records are most likely data entry errors.

To analyse a possible impact of the delta variant to breakthrough infections, we performed an alternative analysis including dummy variables indicating time period starting on July 1st, 2021 when the delta variant started to dominate in Czechia (virus.img.cas.cz/lineages). We performed three such comparisons, each concerning only age cohorts which started to be extensively vaccinated at similar time (S1 Table in S2 File); only this way could guarantee to some extent that the estimation of the delta effect would not collide with that of immunity waning.

All calculations were performed using the R software (package survival). The algorithm used to transform data from the database into the package command inputs was coded in C++. See Supporting information for details.

Results

Since December 26, 2020 to November 20, 2021, 6,287,356 individuals received complete vaccination (58.75% of the population and 67.36% of persons of age 12 years and older). In this period a total of 1,335,055 individuals were infected, of which 96,237 (7.21%) were hospitalized and 20,809 (1.56%) died because of covid-19 (S2 Table and S2 Fig in S2 File). Among vaccinated individuals by far the largest group of 5,011,115 persons (79.7%) received BNT162b2, followed by 469,605 persons (7.47%) vaccinated with mRNA-1273, 436,575 persons (6.94%) with ChAdOx1-S and 370,061 persons (5.89%) with the one-dose Ad26.COV2-S vaccine (S3 Table and S3 Fig in S2 File). The 693,071 booster doses administered in this period comprised 617,002 doses of BNT162b2 and 76,069 doses of mRNA-1273 (S3 Table and S3 Fig in S2 File). We emphasize again that by complete or full vaccination we mean two doses of vaccine (and just one for Ad26.COV2-S) and by booster we mean the third vaccine dose.

Using a Cox regression model we estimated changes in vaccine effectiveness over time at two-month intervals (Fig 2, S4 Table in S2 File). The vaccine effectiveness against any PCR-confirmed SARS-CoV-2 infection declined for BNT162b2 from 87% (95% CI 86–87) 0–2 months after the second dose to 53% (95% CI 52–54) at 7–8 months, for mRNA-1273 from 90% (95% CI 89–91) at 0–2 months to 65% (95% CI 63–67) at 7–8 months, and for ChAdOx1-S from 83% (95% CI 80–85) at 0–2 months to 55% (95% CI 54–56) at 5–6 months. Interestingly, the estimated effectiveness for the Ad26.COV2-S vaccine (68% (95% CI 66–70) at 0–2 months and 67% (95% CI 65–69) at 5–6 months) did not seem to exhibit any significant decline over the study period but notably starts at a significantly lower effectiveness. The effectiveness estimates for ChAdOx1-S and Ad26.COV2-S at 7–8 months after the completion of vaccination exhibit very large uncertainty due to a low number of events as most people completed their vaccination with these vaccines much later, and are therefore only shown in S4 Table in S2 File.

A similar trend can be seen in the estimation of vaccine effectiveness against hospital admissions and deaths. For hospital admission, the vaccine effectiveness declined for BNT162b2 from 90% (95% CI 89–91) at 0–2 months after dose 2 to 75% (95% CI 73–76) at 7–8 months, for mRNA-1273 from 94% (95% CI 92–96) to 81% (95% CI 78–84), and for ChAdOx1-S from 87% (95% CI 81–91) at 0–2 months to 70% (95% CI 68–72) at 5–6 months (Fig 2 red curves, S4 Table in S2 File). In the case of protection from death the model estimated for BNT162b2 a decrease from 92% (95% CI 90–93) at 0–2 months to 83% (95% CI 81–86) at 7–8 months, from 96% (95% CI 91–98) to 88% (95% CI 82–92) for mRNA-1273 within the first 8
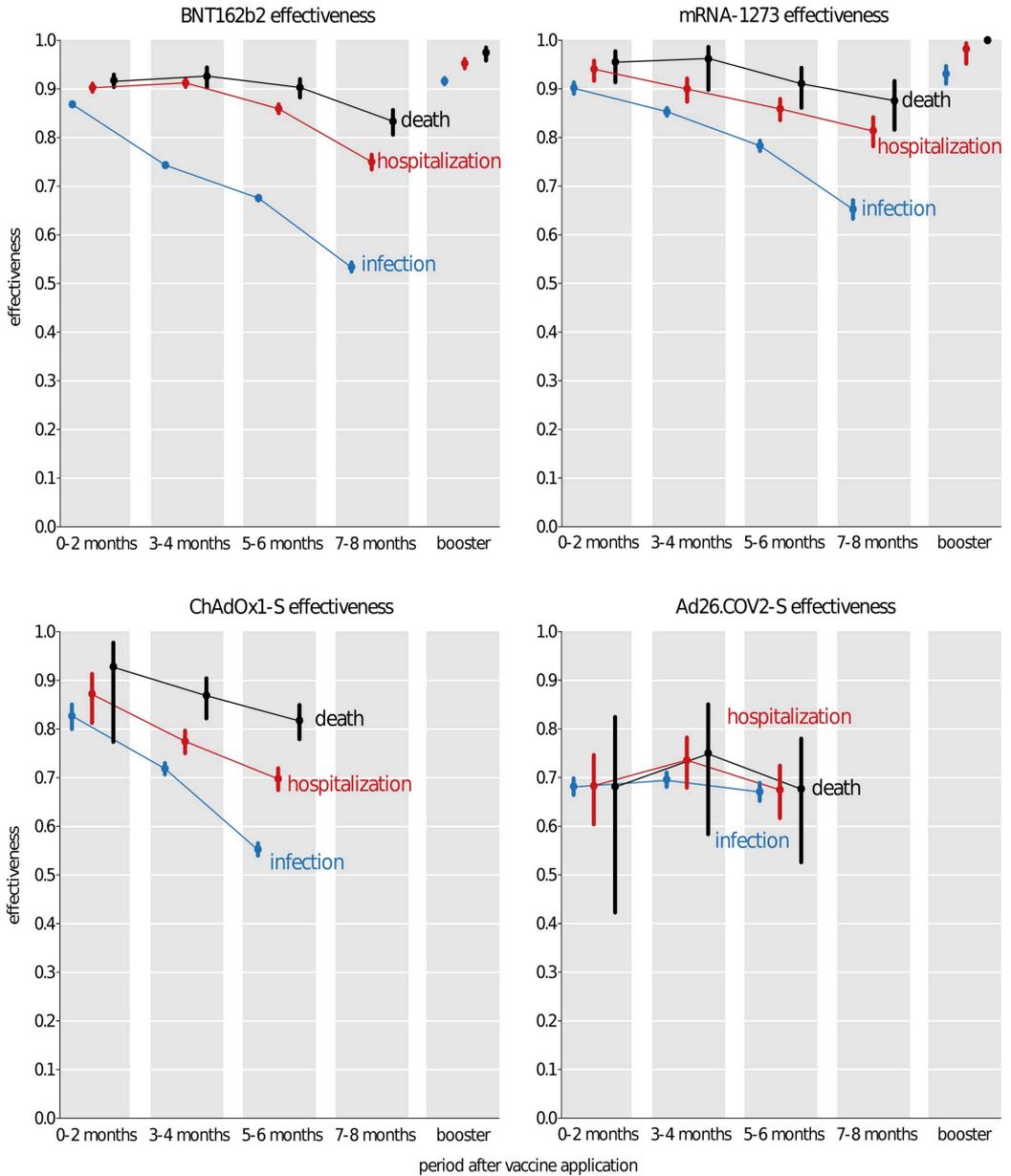
Fig 2. Vaccine effectiveness against infection. Vaccine-acquired immunity against infection with respect to the delay from the full vaccine application, including the effect of a booster vaccine dose.

https://doi.org/10.1371/journal.pone.0270801.g002

months and from 93% (95% CI 77–98) to 82% (95% CI 78–85) for ChAdOx1-S within the first 6 months after application (Fig 2 black curves, S4 Table in S2 File). Ad26.COV2-S once again exhibits virtually no decline either in the protection against hospitalization starting from 68% (95% CI 60–75) at 2 months to 67% (95% CI 62–72) at 5–6 months, or deaths starting from 68% (95% CI 42–82) and reaching 68% (95% CI 53–78) at 5–6 months (Fig 2, S4 Table in S2 File).

To evaluate the differences between the individual vaccines, we statistically tested whether the corresponding covariates differ significantly; that is, whether their differences are significantly different from zero. To this end, we estimated the distribution of the differences by means of the estimator's covariance matrix and checked for statistical significance via a Z-test. Fig 3 summarizes the results. We see, for example, that the BNT162b2 booster is quite superior over all other covariates except the mRNA-1273 booster (quite superior to all covariates).

In June-July 2021 the alpha variant of the SARS-CoV-2 virus was largely superseded by the delta variant in Czechia (virus.img.cas.cz/lineages). Therefore, we attempted to disentangle the effects of immunity waning and immunity evasion due to the delta variant on the observed changes in vaccine effectiveness. Evaluating the extra risk of breakthrough infection due to the delta variant, for consistency estimated just for the age cohorts that started to be vaccinated at about the same time, we found a consistent and significant increase in the risk for BNT162b2, mostly significant increase for mRNA-1273 and ChAdOx1-S, and inconclusive results for Ad26.COV2-S (Table 1). Note that these differences do not represent the infection risk increase due to delta; they represent the additional risk increase of a vaccinated individual over the generally higher infectiousness of the delta variant compared to alpha.

Regardless of the original vaccine used for the initial vaccination schedule a BNT162b2 booster dose enhances protection against infection to 92% (95% CI 91–92), against hospital admission to 95% (95% CI 94–96), and against death to 97% (95% CI 96–98) (Fig 2). A mRNA-1273 booster dose reaches 93% (95% CI 91–95) protection against infection, 98% (95% CI 95–99) against hospital admission, and close to 100% against death (Fig 2). Combining primary and booster mRNA-based vaccines, boosted effectiveness reached > 91%. The combination of ChAdOx1-S primary and an mRNA booster showed a somewhat lower effectiveness but these estimates are less certain due to a low number of observations (Table 2).

To study reinfections, we used data on PCR-confirmed infections since the beginning of covid-19 pandemic in the Czech Republic; 1,999,315 individuals were infected with SARS-CoV-2 virus until 20th November 2021, of which 12,894 (0.64%) were reinfected. Previous SARS-CoV-2 infection in our population conferred a high and fairly stable level of protection against infection lasting for more than 18 months. In unvaccinated but previously covid-19-positive individuals protection against PCR-confirmed covid-19 infection declined from close to 97% (95% CI 97–97) at 2–4 months through 91% (95% CI 90–91) at 5–6 months down to 83% (95% CI 82–84) at 11–12 months and 72% (95% CI 65–78) at 17–18 months (Fig 4, S5 Table in S2 File).

Discussion

Our results show a gradual decrease in protective effectiveness of three out of four vaccines used in Czechia to vaccinate against covid-19. The observed decrease was the fastest for protection against infections followed by hospital admissions, while the protection from covid-19-related death was the least affected by the time elapsed from the completion of primary vaccination schedule.

There are several plausible explanations for this decrease and for a corresponding rise in breakthrough infections. One is waning of the immunity conferred by the vaccines,
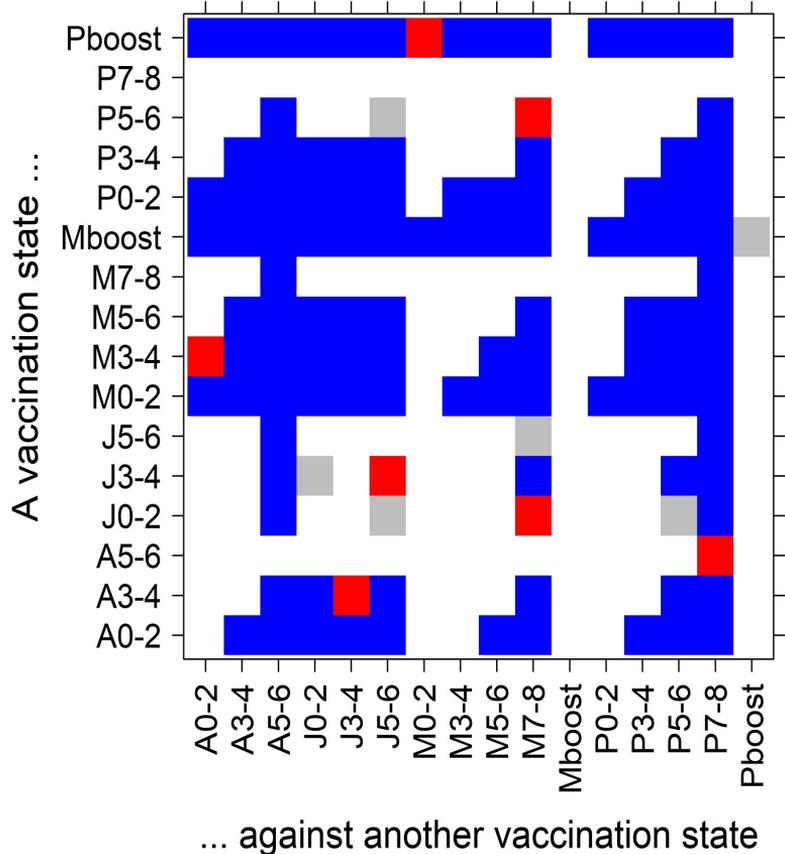
Fig 3. Estimating potential statistical differences between the vaccines. A Z-test has been performed to test for those differences. For each pair of covariates (each covariate is characterized by the vaccine and the time interval since completing the corresponding vaccination scheme), a color is assigned to indicate a degree of statistical significance: blue for 1% ($|Z| > 2.576$), red for 5% ($2.576 \geq |Z| > 1.960$), and gray for $|Z| \leq 1.960$. Moreover, only pairs with positive values of the test statistic Z are plotted, indicating a positive difference between a respective y-axis covariate and x-axis covariate (values symmetric around the diagonal are negative with the same absolute value). The axis labels are composed of a capital letter (P = BNT162b2 vaccine, M = mRNA-1273 vaccine, A = ChAdOx1-S vaccine, and J = Ad26.COV2-S vaccine) and a number range (months since full vaccination) or 'boost' (3rd vaccine dose).

https://doi.org/10.1371/journal.pone.0270801.g003

Table 1. Estimated increase of breakthrough infection hazard ratios (HRs) in times of the SARS-CoV-2 delta variant dominance for age groups having started vaccination in the same month.

Vaccine	March (age 70–80y)		April (age 55–69y)		May (age 35–54y)	
	HR	95% CI	HR	95% CI	HR	95% CI
BNT162b2	1.28	1.09–1.52	1.04	0.95–1.14	1.33	1.27–1.40
mRNA-1273	0.82	0.41–1.67	1.56	1.08–2.25	1.59	1.29–1.98
ChAdOx1-S	1.64	1.05–2.57	1.12	0.74–1.70	1.24	0.82–1.86
Ad26.COV2-S	2.70	0.37–19.63	0.40	0.20–0.78	0.91	0.34–2.43

https://doi.org/10.1371/journal.pone.0270801.t001

Table 2. Vaccine effectiveness against infection after administering the booster vaccine dose for various possible combinations of primary (columns) and booster (rows) vaccines (with the exception of Janssen due to insufficient data). Hazard ratios (HRs) are given.

Vaccine	BNT162b2		mRNA-1273		ChAdOx1-S	
	VE	95% CI	VE	95% CI	VE	95% CI
BNT162b2	0.92	0.91–0.92	0.94	0.91–0.96	0.82	0.68–0.9
mRNA-1273	0.92	0.88–0.95	0.94	0.91–0.95	0.91	0.63–0.98

https://doi.org/10.1371/journal.pone.0270801.t002

documented for a range of commonly used vaccines and demonstrated for covid-19 vaccines in an increasing number of recent studies [12–14]. The other possibility is the effect of the delta variant, shown to evade to some extent the vaccine-induced immunity [12, 15, 16]. Our sub-analysis of the data suggests but a modest overall effect of the delta variant on the vaccine effectiveness in the studied period. Lacking individual-level information about specific variants causing breakthrough infections we used an indirect method of time dummies corresponding to the period of delta variant dominance. Since this approach may be affected by co-linearity ("influence" of the absolute time will be mismatched with the waning), we present these estimates just as secondary results; the primary effectiveness estimates are averaged over the viral variants. However, it is likely that long-term estimates of vaccine effectiveness correspond to the period of dominating delta variant. This is not an issue for boosters, which have been given only after delta reached overwhelming dominance.

A largely unstudied factor that could affect the observed vaccine effectiveness are changes in behaviour of vaccinated vs. unvaccinated persons and a possible effect of infection control measures due to differential access of the unvaccinated individuals to many social activities or due to differential testing strategies, since vaccinated people and people within 6 months after their PCR-positivity have not been required to undergo testing as often as the others. Indeed, analyses of vaccine effectiveness and its temporal dynamics generally assume that both vaccinated and unvaccinated persons behave similarly and we assume this in our study as well. This possible limitation of this study is applicable mainly to the endpoint of confirmed infection and to the lesser extent to the hospitalisation or death endpoints.

The one-dose Janssen vaccine appears in our analysis to defy the general trend of protection decay. While it starts at a significantly lower effectiveness, it holds it over the all 6 months we consider. To our knowledge, this somewhat counter-intuitive result has not yet been reported and as such is not easy to interpret. However, since this vaccine was introduced to the Czech Republic much later than the other three vaccines and only one dose is required for complete vaccination, it is plausible that this vaccine was mostly chosen by people with different social and behavioural characteristics compared to the two-dose vaccines. Since we cannot support this suggestion with data, we leave this as a suggestion for further studies.

We show that administration of booster doses of two approved mRNA vaccines brings the observed effectiveness to above 90% for infections, hospital admissions and deaths alike. Booster doses are highly efficient for preventing serious or fatal infections. Although our results are in a general agreement with the study on protective effect of vaccine booster in Israel [17], we cover a more extensive period of booster applications, use of the mRNA-1273 vaccine as a booster, and do not limit ourselves to any specific age group.

Protection afforded by previous covid-19 infection declines over time, too, but at a slower rate compared to the post-vaccination immunity. Whereas several studies consistently report that protection against reinfection declines [7–10], we are the first to describe the long-term temporal dynamics of infection-induced immunity against SARS-CoV-2 reinfections. We note that this finding relates only to directly confirmed primary infections (possibly associated
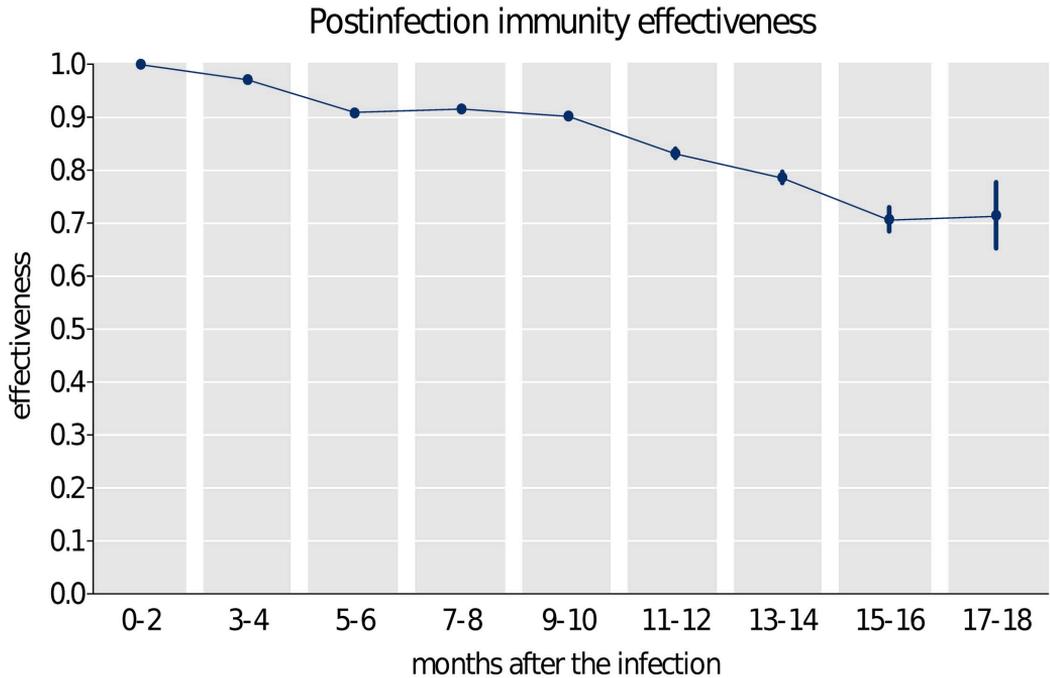
Postinfection immunity effectiveness

Fig 4. Infection-acquired immunity against reinfection with respect to the delay from the prior infection. The delay 0–2 months is not considered as a new infection which implies 100% effectiveness by definition.

https://doi.org/10.1371/journal.pone.0270801.g004

with test-seeking behaviour and severity of the disease) and may not be translatable to evidence of previous infection from antibody testing.

In this study we used a Cox model with calendar time, which has an obvious advantage—our results are independent of factors that influence the risk of all the subjects equally, such as a change in the basic reproduction number, viral prevalence in the population, non-pharmaceutical interventions, weather or seasonal influences, or a dominant virus variant—all these factors can be included in the baseline hazard function of the Cox model. All this makes our findings comparable with and transferable to other contexts. Indeed, similar studies have come to similar conclusions [12]. Our results are also robust with respect to under-reporting provided the reporting rate is the same for all subjects, because then the Cox regression equation is only multiplied by a constant, so the estimation of the HRs remains correct.

The model, however, has some limitations. Importantly, the dependence of individual hazard function on covariates may be non-log-linear. This happens, for instance, when the detection rate depends on a characteristic of a subject—e.g. unvaccinated are being tested more often than the vaccinated. If this were true the vaccine effectiveness would be overestimated, yet the estimates of the HR increase (i.e. VE decline) over time would still be valid (provided that the testing propensity does not change in time). Equally such a case could arise if the vaccinated behaved more riskily than the unvaccinated—the effectiveness then would be underestimated, yet the estimates of the relative increase of HR (and the consequent decrease of VE) would again be valid provided that this behaviour does not change substantially over the study

period. It is also worth noting that unlike infections the hospitalization and mortality data are less likely to suffer from the aforementioned bias as the facts of hospitalizations and deaths depend much less on test seeking behavior.

In addition, as everywhere, only a certain fraction of infections are reported and hence the results could possibly be distorted by this fact; that fraction of unreported infections is called the ascertainment rate. This potential distortion is less severe for the estimates of VE if the ascertainment rate is comparable in the group of the vaccinated individuals and in the control group of the unvaccinated virgin population, since biases then cancel out. However, the problem could be more serious for the estimates of post-infection immunity protection; here, even if the ascertainment rates were comparable in the group of once infected and in the control group of covid-19 virgins, there would be an undetected part of formerly infected and hence protected in the control group. Therefore, the risk of infection in the truly virgin population would be underestimated, while the estimate of the infection risk in the treatment group would not be biased, as all have been infected here. As a consequence, the protection by the past infection is likely underestimated. Determining size of this possible bias is complicated e.g. by the fact that the unreported cases are more likely to be mild ones, providing less protection.

Concerns may arise as to what extent are the results affected by the changing environment due to such aspects as weather, counter-epidemic measures, people's behavior, etc. Thanks to the fact that we use calendar instead of relative time in our analyses, these changes are handled by the baseline hazard function, provided that the environment affects the "treatment" groups in the same way as the "control" group. This can be the case for both the weather (changing the amount of time spent indoors) and the overall counter-epidemic measures (reducing the number of risk contacts). On the other hand, as already emphasized above, differences in behavioral effects cannot be fully excluded (vaccinated or even unvaccinated people may behave in riskier ways); thus, the VE should be understood including the potential behavioral responses. Also, it should be stressed that the fact that the environmental effects "cancel out" in the Cox model with calendar time (the results do not depend on the baseline hazard in any way) effectively precludes studying these effects by means of this model.

Conclusion

We used a comprehensive national population-based database containing individual level data about all detected SARS-CoV-2 infection cases to estimate many important characteristics of the post-vaccination and post-infection immunity in the population of the Czech Republic, covering all four vaccines currently approved in the EU and the protection from infection, hospital admission and death. The results strongly advocate for a timely and widespread administration of the third (i.e., booster) dose. Covid-19 will undoubtedly continue to disrupt everyday lives and cause suffering and loss of life around the globe and vaccine effectiveness data such as the ones presented in this study can bring an important insight for policy makers in order to limit the worst impacts of the current pandemic.

Supporting information

S1 File. More information on the use of the Cox regression method. Details on the statistical model used to process infection and vaccination data and to produce the main results of this study.
(PDF)

S2 File. Supporting Figures and Tables. S1-S3 Figs and S1-S5 Tables that are referred to within the main text. S6-S10 Tables indicating numbers of respective cases behind results plotted in Figs 2 and 4 in the main text.
(PDF)

S3 File. Detailed age- and gender-structured data on temporal dynamics of vaccination and infection in the Czech Republic during the study period. Data presented on a weekly basis, with the date in the first column indicating the middle of the respective week (week is not a calendar week, but from Saturday till Friday, since vaccination started on Saturday, December 26, 2020).
(ODS)

Author Contributions

Conceptualization: Luděk Berec, Martin Šmíd, Lenka Přibylová, Tomáš Pavlík, Tamara Barusová, Jan Trnka.

Data curation: Martin Šmíd, Ondřej Májek, Jiří Jarkovský, Milan Zajíček, Jakub Weiner.

Formal analysis: Martin Šmíd.

Methodology: Ondřej Májek, Tomáš Pavlík, Tamara Barusová.

Resources: Ondřej Májek, Jiří Jarkovský.

Software: Milan Zajíček, Jakub Weiner.

Supervision: Luděk Berec.

Visualization: Lenka Přibylová.

Writing – original draft: Luděk Berec, Martin Šmíd, Jan Trnka.

Writing – review & editing: Lenka Přibylová, Ondřej Májek, Tomáš Pavlík, Jiří Jarkovský, Milan Zajíček, Jakub Weiner, Tamara Barusová.

References

1.   Polack F, Thomas S, Kitchin N, Absalon J, Gurtman A, et al. Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. N Engl J Med. 2020; 383:2603–2615. https://doi.org/10.1056/NEJMoa2034577 PMID: 33301246

2.   Baden L, Sahly HE, Essink B, Kotloff K, Frey S, et al. Efficacy and Safety of the mRNA-1273 SARS-CoV-2 Vaccine. N Engl J Med. 2021; 384:403–416. https://doi.org/10.1056/NEJMoa2035389 PMID: 33378609

3.   Voysey M, Clemens SC, Madhi S, Weckx L, Folegatti P, Aley P, et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. Lancet. 2021; 397:99.111. https://doi.org/10.1016/S0140-6736(20)32661-1 PMID: 33306989

4.   Sadoff J, Gray G, Vandebosch A, Cárdenas V, Shukarev G, et al. Safety and Efficacy of Single-Dose Ad26.COV2.S Vaccine against Covid-19. N Engl J Med. 2021; 384:2187–2201. https://doi.org/10.1056/NEJMoa2101544 PMID: 33882225

5.   Brown C, Vostok J, Johnson H, Burns M, Gharpure R, et al. Outbreak of SARS-CoV-2 Infections, Including COVID-19 Vaccine Breakthrough Infections, Associated with Large Public Gatherings—Barnstable County, Massachusetts, July 2021. MMWR Morb Mortal Wkly Rep. 2021; 70:1059–1062. https://doi.org/10.15585/mmwr.mm7031e2 PMID: 34351882

6.   Levine-Tiefenbrun M, Yelin I, Alapi H, Katz R, Herzel E, Kuint J, et al. Viral loads of Delta-variant SARS-CoV-2 breakthrough infections after vaccination and booster with BNT162b2. Nat Med. PMID: 34728830

7. Abu-Raddad L, Bertollini R. Severity of SARS-CoV-2 Reinfections as Compared with Primary Infections. N Engl J Med. https://doi.org/10.1056/NEJMc2108120 PMID: 34818474

8. Cavanaugh A, Spicer K, Thoroughman D, Glick C, Winter K. Reduced Risk of Reinfection with SARS-CoV-2 After COVID-19 Vaccination. Kentucky, May–June 2021. MMWR Morb Mortal Wkly Rep. 2021; 70:1081–1083. https://doi.org/10.15585/mmwr.mm7032e1 PMID: 34383732

9. Hansen C, Michlmayr D, Gubbels S, Molbak K, Ethelberg S. Assessment of protection against reinfection with SARS-CoV-2 among 4 million PCR-tested individuals in Denmark in 2020: a population-level observational study. Lancet. 2021; 397:1204–1212. https://doi.org/10.1016/S0140-6736(21)00575-4 PMID: 33743221

10. Townsend J, Hassler H, Wang Z, Miura S, Singh J, Kumar S, et al. The durability of immunity against reinfection by SARS-CoV-2: a comparative evolutionary study. Lancet Microbe. https://doi.org/10.1016/S2666-5247(21)00219-6 PMID: 34632431

11. Komenda M, Bulhart V, Karolyi M, Jarkovský J, Mužík J, et al. Complex Reporting of the COVID-19 Epidemic in the Czech Republic: Use of an Interactive Web-Based App in Practice. J Med Internet Res. 2020; 22:e19367. https://doi.org/10.2196/19367 PMID: 32412422

12. Tartof S, Slezak J, Fischer H, Hong V, Ackerson B, Ranasinghe O, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet. 2021; 398:1407–1416. https://doi.org/10.1016/S0140-6736(21)02183-8 PMID: 34619098

13. Chemaitelly H, Tang P, Hasan M, AlMukdad S, Yassine H, et al. Waning of BNT162b2 Vaccine Protection against SARS-CoV-2 Infection in Qatar. N Engl J Med. 2021; https://doi.org/10.1056/NEJMoa2114114

14. Cohn B, Circillo P, Murphy C, Krigbaum N, Wallace A. SARS-CoV-2 vaccine protection and deaths among US veterans during 2021. Science. PMID: 34735261

15. Lopez Bernal J, Andrews N, Gower C, Gallagher E, Simmons R, Thelwall S, et al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. New Engl J Med. 2021; 385(7):585–594. https://doi.org/10.1056/NEJMoa2108891 PMID: 34289274

16. Tang P, Hasan M, Chemaitelly H, Yassine H, Benslimane F, Al Khatib H, et al. BNT162b2 and mRNA-1273 COVID-19 vaccine effectiveness against the SARS-CoV-2 Delta variant in Qatar. Nat Med. 2021; https://doi.org/10.1038/s41591-021-01583-4

17. Bar-On Y, Goldberg Y, Mandel M, Bodenheimer O, Freedman L, Kalkstein N, et al. Protection of BNT162b2 Vaccine Booster against Covid-19 in Israel. N Engl J Med. 2021; 385:1393–1400. https://doi.org/10.1056/NEJMoa2114255 PMID: 34525275

Project MAMES (in colaboration with IBA and IHIS (ÚZIS])

European Forecast Hub modeling teams under ECDC for covid-19 forecasting
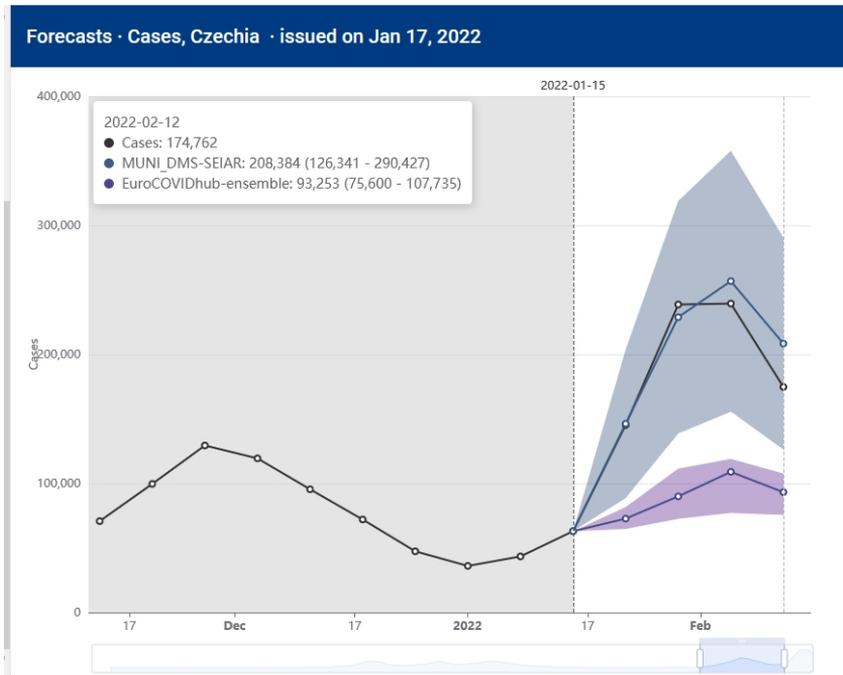

Figure A.1: 4-week prediction of the omicron wave of our model (blue) and ECDC ensemble model (violet). Black line corresponds to reported cases.

Description of model performance evaluation: The weighted interval score (WIS) is a proper scoring rule (i.e., it cannot be "cheated") that is suited to scoring forecasts in an interval format. It generalizes the absolute error (i.e. lower values are better) and has three components: dispersion, underprediction and overprediction. Dispersion is a weighted average of the widths of the submitted prediction intervals. Over- and underprediction (overpred/underpred) penalties are added whenever an observation falls outside of a reported central prediction interval, with the strength of the penalty depending on the nominal level of the interval and how far outside of the interval the observation fell. Note that the average WIS can refer to different sets of targets for different models and therefore cannot always be compared across models. Such comparisons should be done based on the relative skill. The Relative WIS (column rel_wis) is a

relative measure of forecast performance which takes into account that different teams may not cover the exact same set of forecast targets (i.e., weeks and locations). Loosely speaking, a relative WIS of X means that averaged over the targets a given team addressed, its WIS was X times higher/lower than the performance of the baseline model described in [6]. Smaller values are thus better and a value below one means that the model has above-average performance. The relative WIS is computed using a 'pairwise comparison tournament' where for each pair of models a mean score ratio is computed based on the set of shared targets. The relative WIS is the geometric mean of these ratios. Details on the computation can be found in [6]. This metric is calculated for all models that provide a full set of quantiles. The Relative Absolute Error (column rel_ae) is the relative absolute error of the predictive point forecasts, that is the predicted value that individual forecasts deem the most likely. Coverage (50% Cov. / 95% Cov.) is the proportion of observations that fell within a given prediction interval. Bias (bias) is a measure between -1 and 1 that expresses the tendency to underpredict (-1) or overpredict (1). Our model MUNI_DMS-SEIAR for the Czech Republic is being submitted to ECDC Forecast Hub weekly for around a year with the best performance evaluations for cases and hospitalizations (even above ensemble performance):

model	week	num.	rel_wis	rel_ae	50%	95%	bias
MUNI_DMS-SEIAR	1	10	0.56	0.41	1	1	-0.04
EuroCOVIDhub-ensemble	1	10	0.73	0.78	0.1	0.5	0.09
MUNI_DMS-SEIAR	1	47	0.55	0.5	0.66	0.94	-0.09
EuroCOVIDhub-ensemble	1	49	0.69	0.74	0.63	0.96	0
MUNI_DMS-SEIAR	2	10	0.52	0.64	0.8	1	0.05
EuroCOVIDhub-ensemble	2	10	0.95	1.15	0.1	0.6	0.14
MUNI_DMS-SEIAR	2	46	0.47	0.54	0.48	0.76	-0.03
EuroCOVIDhub-ensemble	2	48	0.78	0.88	0.44	0.85	0.03
MUNI_DMS-SEIAR	3	9	0.79	1.14	0.56	0.89	0.01
EuroCOVIDhub-ensemble	3	10	1.07	1.23	0.1	0.3	0.06
MUNI_DMS-SEIAR	3	45	0.65	0.82	0.27	0.69	0.02
EuroCOVIDhub-ensemble	3	47	0.83	0.91	0.38	0.74	0
MUNI_DMS-SEIAR	4	9	0.82	1.08	0.22	0.67	-0.07
EuroCOVIDhub-ensemble	4	10	1.13	1.19	0.1	0.4	-0.04
MUNI_DMS-SEIAR	4	44	0.69	0.81	0.16	0.55	0
EuroCOVIDhub-ensemble	4	46	0.88	0.91	0.35	0.8	-0.01

Table A.1: European COVID-19 Forecast Hub Evaluation Report for Czechia, accessed February 17, 2022. Column "week" addresses the forecast type, "num." number of last forecasts included in the evaluation, the next columns belong to the above described evaluation metrics.

Invited poster to V4-UK Frontiers of Science 2021 (Royal Society UK)

Model ZSEIAR

Lenka Přibylová (email: pribylova@math.muni.cz), Veronika Hajnová (email: hajnova@math.muni.cz)

Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic

METHODICS

Differential equations

$$\dot{Z} = -\varepsilon Z/N,$$
$$\dot{S} = -\frac{\beta}{N-Z}S(I+A) + \varepsilon Z/N,$$
$$\dot{E} = \frac{\beta}{N-Z}S(I+A) - \gamma E,$$
$$\dot{I} = \gamma p E - \mu_1 I,$$
$$\dot{A} = \gamma(1-p)E - \mu_2 A,$$
$$\dot{Q} = \mu_1 I - \nu Q,$$
$$\dot{R} = \nu Q,$$

- Z not affected population size
- S susceptible
- E exposed
- I detected infectious
- A undetected infectious
- Q izolated infectious
- R removed detected
- N population size
- p ascertainment rate
- $\varepsilon, \beta, \gamma, \mu_1, \mu_2, \nu$ parameters
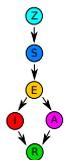
Optimization estimates the size of the affected clusters (dependencies such as seasonality, the degree of influence of government measures or changes in people's behavior caused by fear or disinformation are included in the estimate of affected clusters), we use estimates of latent or infectious periods from the literature, estimates of computable periods as isolation time or time to hospitalization or death from ÚZIS dataset and a 14-day moving average ascertainment rate estimate. Scenarios of ZSEIAR model can be generated using MAMES application [?], for detailed documentation see [?].

Mobility dependence

Transmissibility rate β is assumed to be strictly dependent on the number of contacts or mobility.
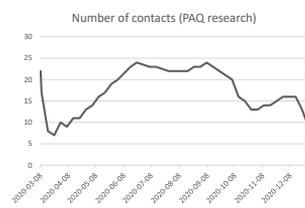
Figure 1: Number of contacts according to social studies [1] during COVID-19 epidemic in the Czech Republic, 2020
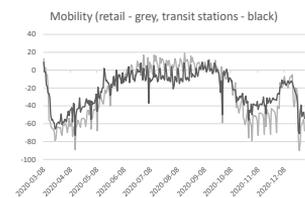
Figure 2: Mobility [2] during COVID-19 epidemic in the Czech Republic, 2020

Ascertainment rate estimate

The rate is based on the Bayes rule for conditional probabilities and on the assumption that the average infected person's probability of hospitalization is given or estimated.

$$p = P(Det) = \frac{P(Det|H)P(H)}{P(H|Det)}$$

$P(Det|H)$ – the probability that a person hospitalized with COVID-19 was previously detected; the proportion of patients detected prior to admission to hospital and all patients hospitalized with COVID-19 (including those not detected prior to admission to hospital) relative to the date of the positive test report.
$P(H|Det)$ – the probability that if an individual was detected, he or she will be hospitalized; the proportion of all reported hospitalized patients detected before admission to the hospital, and all detected persons except those detected only in the hospital.
$P(H)$ – the probability that SARS-CoV-2 positive individual is/was/will be hospitalized (regardless of whether was detected or not); estimated to the age structure of the epidemics affected.

Hospitalization probability age structure dependence

Age structure in the Czech Republic attained 20% share of the over 65 year-old population. Patients hospitalized with COVID-19 over the age of 65 had a long-term ratio of around 3/4 during the autumn 2020, so a rough estimate of probability $P(H_{65+})$ of hospitalization with COVID-19 for a person over 65 is around twelve times higher than probability $P(H_{65-})$ of hospitalization with COVID-19 for a person under 65.

$$P(H) = p_{65-}^+ P(H_{65-}) + p_{65+}^+ P(H_{65+}) =$$
$$= \left(1 + 11\frac{p_{65+}^+}{p_{65-}^+}\right) P(H_{65-}),$$

where p_{65+}^+ and p_{65-}^+ are 7-day moving averages of the senior and non-senior population ratio in the reported positive cases.

Early warning

A significant decrease in the ascertainment rate estimate is a signal of tracing and testing overload collapse.
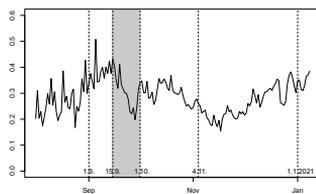

Figure 3: The ascertainment rate estimate shows the collapse of tracing by regional hygiene stations, which occurred in the second half of September.

Acknowledgements

REAL DATA FIT

Spring and summer 2020

Simulation of a spring outbreak in the Czech Republic with the constant growth rate of affected clusters copies the real data due to the ascertainment rate estimate that shapes the local outbreak in the OKD mines that was the main source of the infected in the Czech Republic in July. In August, the optimization of affected clusters abruptly changes and is no longer constant as the real data start to increase exponentially.
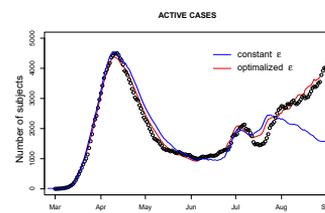

Figure 4: Active cases – real data (circles), ZSEIAR model with constant ε from April (blue), and optimized ε (red)
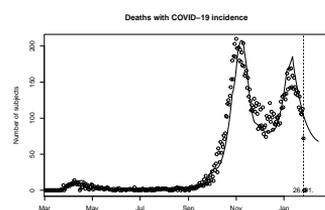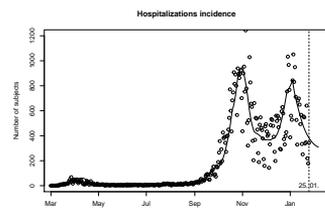
Fall 2020

Figure 5: Optimalized fit (line) to real data (circles) during the year 2020 – incidence of admissions to hospitals and deaths.

References

[1] PAQ research, IDEA AntiCovid, https://zivotbehempandemie.cz/kontakty, 2020 (accessed January 21, 2021)

[2] Google Mobility reports, https://github.com/ActiveConclusion/COVID19_mobility/blob/master/google_reports/mobility_report_europe.xlsx, 2020 (accessed January 21, 2021)

[3] MAMES application, https://webstudio.shinyapps.io/MAMES/

[4] ZSEIAR model (documentation in Czech), https://is.muni.cz/www/98951/47857356/mames/MAMES_ZSEIAR_metodika.pdf, 2021

Ascertainment rate estimate from hospital data used in modelling COVID-19 epidemics

Veronika Eclerová (email: eclerova@math.muni.cz), Lenka Přibylová (email: pribylova@math.muni.cz)

Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic

METHODS

Differential equations

$$\dot{Z} = -\varepsilon Z/N,$$
$$\dot{S} = -\frac{\beta}{N-Z}S(I+A) + \varepsilon Z/N,$$
$$\dot{E} = \frac{\beta}{N-Z}S(I+A) - \gamma E,$$
$$\dot{I} = \gamma p E - \mu_1 I,$$
$$\dot{A} = \gamma(1-p)E - \mu_2 A,$$
$$\dot{Q} = \mu_1 I - \nu Q,$$
$$\dot{R} = \nu Q,$$

$N = 10^7$ population size
$Z = Z(t)$ unaffected at time t
$S = S(t)$ susceptible at time t
$E = E(t)$ exposed at time t
$I = I(t)$ reported infectious at time t
$A = A(t)$ unreported infectious at time t
$Q = Q(t)$ isolated reported infectious at time t
$R = R(t)$ reported recovered or deceased at time t
$\varepsilon = \varepsilon(t)$ rate of cluster growth at time t
$p = p(t)$ ascertainment rate at time t
$\beta = \beta(t)$ transmission rate at time t
$\gamma = 1/4$ reciprocal of the mean incubation period
$\mu_1 = \mu_1(t)$ reciprocal of the mean infectious period of I
$\mu_2 = 1/3$ reciprocal of the mean infectious period of A
$\nu = \nu(t)$ reciprocal of the mean quarantine duration

Optimization of $\varepsilon(t)$: dependencies such as seasonality, the degree of influence of government measures, or changes in people's behavior caused by fear or disinformation are included in the estimate of affected clusters

Estimation from data

$\beta(t)$ - Google mobility report
$p(t)$, $\mu_1(t)$, $\nu(t)$ - ÚZIS dataset

Ascertainment rate estimate

The rate is based on the Bayes rule for conditional probabilities and on the assumption that the average infected person's probability of hospitalization is given or estimated.

$$p = P(Det) = \frac{P(Det|H)P(H)}{P(H|Det)}$$

$P(Det|H)$ – the probability that a person admitted to a hospital with COVID-19 was previously detected at time t; to estimate this probability, we use the 7-day moving proportion of patients reported prior to admission to the hospital from all patients hospitalized with COVID-19 (including those not detected prior to admission to hospital) with respect to the date of the positive test report
$P(H|Det)$ – the probability that if an individual was detected at time t, he or she will be hospitalized; to estimate this probability we use the 7-day moving proportion of all reported hospitalized patients detected before admission to the hospital, from all that time already detected subjects with respect to the date of the positive test report, that is except those detected in the hospital afterwards who are a part of the undetected compartment at time t
$P(H)$ – the probability that SARS-CoV-2 positive individual is/was/will be hospitalized (regardless of whether was detected or not) at time t; should be derived for each community/country separately since it depends highly on the population age structure, one possible estimation method is described in the next subsection

Hospitalization probability age structure dependence

Age structure in the Czech Republic attained 20% share of the over 65 year-old population. Patients hospitalized with COVID-19 over the age of 65 had a long-term ratio of around 3/4 during the autumn 2020, so a rough estimate of probability $P(H_{65+})$ of hospitalization with COVID-19 for a person over 65 is around twelve times higher than probability $P(H_{65-})$ of hospitalization with COVID-19 for a person under 65.

$$P(H) = p_{65-}^+ P(H_{65-}) + p_{65+}^+ P(H_{65+})$$
$$= P(H_{65-})\left(1 + 11p_{65+}^+\right)$$
$$= \frac{1}{160}(1 + 11p_{65+}^+),$$

where p_{65+}^+ and p_{65-}^+ are 7-day moving averages of the senior and non-senior population ratio in the reported positive cases.

Retrospective data analysis shows that people under 20 were hospitalized with almost zero probability and three age compartments (under 20, 20-65, and over 65) seems to be enough for the basic improvement of $P(H)$ estimate in the form

$$P(H) = \frac{1}{160}(1 - p_{20-}^+ + 11p_{65+}^+),$$

where p_{20-}^+, p_{20-65}^+, and p_{65+}^+ are 7-day moving averages of the young, middle-aged, and senior population ratio in the reported positive cases.

Spring and summer 2020

Simulation of a spring outbreak in the Czech Republic with the constant growth rate of affected clusters copies the real data due to the ascertainment rate estimate that shapes the local outbreak in the OKD mines that was the main source of the infected in the Czech Republic in July. In August, the optimization of affected clusters abruptly changes and is no longer constant as the real data start to increase exponentially.
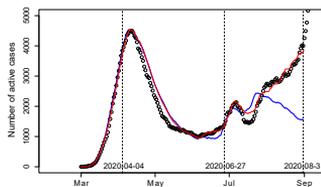

Figure 1: Active cases - real data (circles), model with constant ε from April (blue), and optimized ε (red)
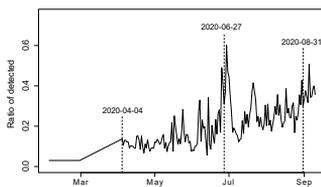

Figure 2: Ascertainment rate estimate

Fall 2020

Insufficience of tracing in the second half of September. The peak of the second wave is marked at 2020-11-04. Dates in chart denotes: 2020-09-01 – beginning of school year, 2020-09-15–2021-10-01 – tracing overload, 2021-11-04 – peak of the second wave, 2021-01-01 – new year.
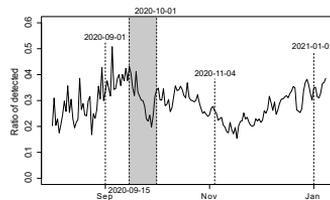

Figure 3: Ascertainment rate estimate

Spring 2020 - Summer 2021

Exposed are estimated using new cases divided by the ascertainment rate estimate compared to the introduction of NPIs: 2020-09-01 (beginning of the school year), 2020-10-22 (partial lockdown), 2020-11-14 (announcement of the first measure release after 2020-11-18), 2020-12-03 (shops reopening), 2020-12-27 (a partial lockdown), 2021-01-11 (partial school reopening), 2021-02-23 (mandatory respirators).
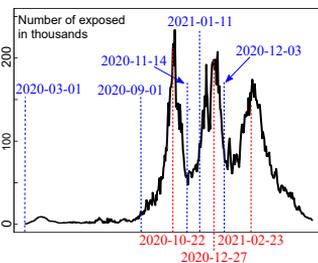

Figure 4: Exposed

Acknowledgements

References

[1] MAMES application, https://webstudio.shinyapps.io/MAMES/
[2] ZSEIAR model (documentation in Czech), https://is.muni.cz/www/98951/47857356/mames/MAMES_ZSEIAR_metodika.pdf, 2021
[3] European Covid-19 Forecast Hub, https://covid19forecasthub.eu/community.html, model MUNI_DMS-SEIAR, 2021

Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations

Sherratt, K. [1], Gruson, H. [1], Grah, R. [2], Johnson, H. [2], Niehus, R. [2], Prasse, B. [2], Sandman, F. [2], Deuschel, J. [3], Wolffram, D. [3], Abbott, S. [1], Ullrich, A. [4], Gibson, G. [5], Ray, EL. [5], Reich, NG. [5], Sheldon, D. [5], Wang, Y. [5], Wattanachit, N. [5], Wang, L. [6], Trnka, J. [7], Obozinski, G. [8], Sun, T. [8], Thanou, D. [8], Pottier, L. [9], Krymova, E. [10], Barbarossa, MV. [11], Leithäuser, N. [12], Mohring, J. [12], Schneider, J. [12], Wlazlo, J. [12], Fuhrmann, J. [13], Lange, B. [14], Rodiah, I. [14], Baccam, P. [15], Gurung, H. [15], Stage, S. [15], Suchoski, B. [15], Budzinski, J. [16], Walraven, R. [16], Villanueva, I. [17], Tucek, V. [18], Šmíd, M. [19], Zajícek, M. [19], Pérez Álvarez, C. [20], Reina, B. [20], Bosse, NI. [1], Meakin, S. [1], Alaimo Di Loro, P. [21], Maruotti, A. [21], Eclerová, V. [22], Kraus, A. [22], Kraus, D. [22], Pribylova, L. [22], Dimitris, B. [23], Li, ML. [23], Saksham, S. [23], Dehning, J. [24], Mohr, S. [24], Priesemann, V. [24], Redlarski, G. [25], Bejar, B. [26], Ardenghi, G. [27], Parolini, N. [27], Ziarelli, G. [27], Bock, W. [28], Heyder, S. [29], Hotz, T. [29], E. Singh, D. [30], Guzman-Merino, M. [30], Aznarte, JL. [31], Moriña, D. [32], Alonso, S. [33], Álvarez, E. [33], López, D. [33], Prats, C. [33], Burgard, JP. [34], Rodloff, A. [35], Zimmermann, T. [35], Kuhlmann, A. [36], Zibert, J. [37], Pennoni, F. [38], Divino, F. [39], Català, M. [40], Lovison, G. [41], Giudici, P. [42], Tarantino, B. [42], Bartolucci, F. [43], Jona Lasinio, G. [44], Mingione, M. [44], Farcomeni, A. [45], Srivastava, A. [46], Montero-Manso, P. [47], Adiga, A. [48], Hurt, B. [48], Lewis, B. [48], Marathe, M. [48], Porebski, P. [48], Venkatramanan, S. [48], Bartczuk, R. [49], Dreger, F. [49], Gambin, A. [49], Gogolewski, K. [49], Gruziel-Slomka, M. [49], Krupa, B. [49], Moszynski, A. [49], Niedzielewski, K. [49], Nowosielski, J. [49], Radwan, M. [49], Rakowski, F. [49], Semeniuk, M. [49], Szczurek, E. [49], Zielinski, J. [49], Kisielewski, J. [50], Pabjan, B. [51], Holger, K. [52], Kheifetz, Y. [52], Scholz, M. [52], Bodych, M. [53], Filinski, M. [53], Idzikowski, R. [53], Krueger, T. [53], Ozanski, T. [53], Bracher, J. [3], Funk, S. [1]

[1]London School of Hygiene & Tropical Medicine, [2]European Centre for Disease Prevention and Control (ECDC), [3]Karlsruhe Institute of Technology, [4]Robert Koch Institute, [5]University of Massachusetts Amherst, [6]Boston Children's Hospital and Harvard Medical School, [7]Charles University, [8]École Polytechnique Fédérale de Lausanne, [9]Éducation nationale, [10]Eidgenössische Technische Hochschule Zürich, [11]Frankfurt Institute for Advanced Studies, [12]Fraunhofer Institute for Industrial Mathematics, [13]Heidelberg University, [14]Helmholtz Centre for Infection Research, [15]IEM, Inc., [16]Independent, [17]Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Universitat Pompeu Fabra, [18]Institute of Computer Science of the CAS, [19]Institute of Information Theory and Automation of the CAS, [20]Inverence, [21]LUMSA University, [22]Masaryk University, [23]Massachusetts Institute of Technology, [24]Max-Planck-Institut für Dynamik und Selbstorganisation, [25]Medical University of Gdansk, [26]Paul Scherrer Institute, [27]Politecnico di Milano, [28]Technical University of Kaiserlautern, [29]Technische Universität Ilmenau, [30]Universidad Carlos III de Madrid, [31]Universidad Nacional de Educación a Distancia (UNED), [32]Universitat de Barcelona, [33]Universitat Politècnica de Catalunya, [34]Universität Trier, [35]University of Cologne, [36]University of Halle, [37]University of Ljubljana, [38]University of Milano-Bicocca, [39]University of Molise, [40]University of Oxford, [41]University of Palermo, [42]University of Pavia, [43]University of Perugia, [44]University of Rome "La Sapienza", [45]University of Rome "Tor Vergata", [46]University of Southern California, [47]University of Sydney, [48]University of Virginia, [49]University of Warsaw, [50]University of Warsaw, University of Bialystok, [51]University of Wroclaw, [52]Universtät Leipzig, [53]Wroclaw University of Science and Technology

Abstract

Background: Short-term forecasts of infectious disease burden can contribute to situational awareness and aid capacity planning. Based on best practice in other fields and recent insights in infectious disease epidemiology,

one can maximise the predictive performance of such forecasts if multiple models are combined into an ensemble. Here we report on the performance of ensembles in predicting COVID-19 cases and deaths across Europe between 08 March 2021 and 07 March 2022.

Methods: We used open-source tools to develop a public European COVID-19 Forecast Hub. We invited groups globally to contribute weekly forecasts for COVID-19 cases and deaths reported from a standardised source over the next one to four weeks. Teams submitted forecasts from March 2021 using standardised quantiles of the predictive distribution. Each week we created an ensemble forecast, where each predictive quantile was calculated as the equally-weighted average (initially the mean and then from 26th July the median) of all individual models' predictive quantiles. We measured the performance of each model using the relative Weighted Interval Score (WIS), comparing models' forecast accuracy relative to all other models. We retrospectively explored alternative methods for ensemble forecasts, including weighted averages based on models' past predictive performance.

Results: Over 52 weeks we collected and combined up to 28 forecast models for 32 countries. We found a weekly ensemble had a consistently strong performance across countries over time. Across all horizons and locations, the ensemble performed better on relative WIS than 84% of participating models' forecasts of incident cases (with a total N=862), and 92% of participating models' forecasts of deaths (N=746). Across a one to four week time horizon, ensemble performance declined with longer forecast periods when forecasting cases, but remained stable over four weeks for incident death forecasts. In every forecast across 32 countries, the ensemble outperformed most contributing models when forecasting either cases or deaths, frequently outperforming all of its individual component models. Among several choices of ensemble methods we found that the most influential and best choice was to use a median average of models instead of using the mean, regardless of methods of weighting component forecast models.

Conclusions: Our results support the use of combining forecasts from individual models into an ensemble in order to improve predictive performance across epidemiological targets and populations during infectious disease epidemics. Our findings further suggest that median ensemble methods yield better predictive performance more than ones based on means. Our findings also highlight that forecast consumers should place more weight on incident death forecasts than incident case forecasts at forecast horizons greater than two weeks.

Code and data availability: All data and code are publicly available on Github: covid19-forecast-hub-europe/euro-hub-ensemble.

Background

Epidemiological forecasts make quantitative statements about a disease outcome in the near future. Forecasting targets can include measures of prevalent or incident disease and its severity, for some population over a specified time horizon. Researchers, policy makers, and the general public have used such forecasts to understand and respond to the global outbreaks of COVID-19 [1]–[3].

Forecasters use a variety of methods and models for creating and publishing forecasts, varying in both defining the forecast outcome and in reporting the probability distribution of outcomes [4], [5]. Such variation makes it difficult to compare predictive performance between forecast models, and from there to derive objective arguments for using one forecast over another. This confounds the selection of a single representative forecast and reduces the reliability of the evidence base for decisions based on forecasts.

A "forecast hub" is a centralised effort to improve the transparency and usefulness of forecasts, by standardising and collating the work of many independent teams producing forecasts [6]. A hub sets a commonly agreed-upon structure for forecast targets, such as type of disease event, spatio-temporal units, or the set of quantiles of the probability distribution to include from probabilistic forecasts. For instance, a hub may collect predictions of the total number of cases reported in a given country for each day in the next two weeks. Forecasters can adopt this format and contribute forecasts for centralised storage in the public domain. This shared infrastructure allows forecasts produced from diverse teams and methods to be visualised and quantitatively compared on a like-for-like basis, which can strengthen public and policy use of disease forecasts.

The underlying approach to creating a forecast hub was pioneered in climate modelling and adapted for collaborative epidemiological forecasts of dengue [7] and influenza in the USA [6], [8]. This infrastructure was adapted for forecasts of short-term COVID-19 cases and deaths in the US [9], [10], prompting similar efforts in some European countries [11]–[13].

Standardising forecasts allows for combining multiple forecasts into a single ensemble with the potential for an improved predictive performance. Evidence from previous efforts in multi-model infectious disease forecasting suggests that forecasts from an ensemble of models can be consistently high performing compared to any one of the component models [7], [8], [14]. Elsewhere, weather forecasting has a long-standing use of building ensembles of models using diverse methods with standardised data and formatting in order to improve performance [15], [16].

The European COVID-19 Forecast Hub [17] is a project to collate short term forecasts of COVID-19 across 32 countries in the European region. The Hub is funded and supported by the European Centre for Disease Prevention and Control (ECDC), with the primary aim to provide reliable information about the near-term epidemiology of the COVID-19 pandemic to the research and policy communities and the general public [3]. Second, the Hub aims to create infrastructure for storing and analysing epidemiological forecasts made in real time by diverse research teams and methods across Europe. Third, the Hub aims to maintain a community of infectious disease modellers underpinned by open science principles.

We started formally collating and combining contributions to the European Forecast Hub in March 2021. Here, we investigate the predictive performance of an ensemble of all forecasts contributed to the Hub in real time each week, as well as the performance of variations of ensemble methods created retrospectively.

Methods

We developed infrastructure to host and analyse prospective forecasts of COVID-19 cases and deaths. The infrastructure is compatible with equivalent research software from the US [18], [19] and German and Polish COVID-19 [20] Forecast Hubs, and easy to replicate for new forecasting collaborations.

Forecast targets and models

We sought forecasts for the incidence of COVID-19 as the total reported number of cases and deaths per week. We considered forecasts for 32 countries in Europe, including all countries of the European Union, European Free Trade Area, and the United Kingdom. We compared forecasts against observed data reported for each country by Johns Hopkins University (JHU, [21]). JHU data sources included a mix of national and aggregated subnational data. We aggregated incidence over the Morbidity and Mortality Weekly Report (MMWR) epidemiological week definition of Sunday through Saturday.

Teams could express their uncertainty around any single forecast target by submitting predictions for up to 23 quantiles (from 0.01 to 0.99) of the predictive probability distribution. Teams could also submit a single point forecast. At the first submission we asked teams to add a pre-specified set of metadata briefly describing the forecasting team and methods (see supplementary information (SI)). No restrictions were placed on who could submit forecasts. To increase participation we actively contacted known forecasting teams across Europe and the US and advertised among the ECDC network. Teams submitted a broad spectrum of model types, ranging from mechanistic to empirical models, agent-based and statistical models, and ensembles of multiple quantitative or qualitative models (described at [22]). We maintain a full project specification with a detailed submissions protocol [23].

We collected forecasts submitted weekly in real time over the 52 week period from 08 March 2021 to 07 March 2022. Teams submitted at latest two days after the complete dataset for the latest forecasting week became available each Sunday. We implemented an automated validation programme to check that each new forecast conformed to standardised formatting. Forecast validation ensured a monotonic increase of predictions with each increasing quantile, integer-valued non-negative counts of predicted cases, as well as consistent date and location definitions.

Each week we used all available valid forecasts to create a weekly real-time ensemble model (referred to as "the ensemble" from here on), for each of the 256 possible forecast targets: incident cases and deaths in 32 locations over the following one through four weeks. The ensemble method was an unweighted average of all models' forecast values, at each predictive quantile for a given location, target, and horizon. From 08 March 2021, we used the arithmetic mean. However we noticed that including highly anomalous forecasts in a mean ensemble produced extremely wide uncertainty. To mitigate this, from 26th July 2021 onwards the ensemble instead used a median of all predictive quantiles.

We created an open and publicly accessible interface to the forecasts and ensemble, including an online visualisation tool allowing viewers to see past data and interact with one or multiple forecasts for each country and target for up to four weeks' horizon [24]. All forecast and meta data are freely available and held on Github [17] and Zoltar, a platform for hosting epidemiological forecasts [25], [26]. In the codebase for this study [27] we provide a simple method and instructions for downloading and preparing these data for analysis using R. We encourage other researchers to freely use and adapt this to support their own analyses.

Forecast evaluation

In this study we focus only on the comparative performance of forecasting models relative to each other. For each model, we evaluated performance in terms of both accuracy (coverage) and overall predictive performance (weighted interval score). We evaluated all previous forecasts against actual observed values for each model, stratified by the forecast horizon, location, and target. We calculated scores using the *scoringutils* R package [28]. We removed any forecast surrounding (both the week of, and the first week after) a strongly anomalous data point. We defined anomalous as where any subsequent data release revised that data point by over 5%.

We established the accuracy of each model's prediction boundaries as the coverage of the predictive intervals. We calculated coverage at a given interval level k, where $k \in [0, 1]$, as the proportion p of observations that fell within the corresponding central predictive intervals across locations and forecast dates. A perfectly calibrated model would have $p = k$ at all 11 levels (corresponding to 22 quantiles excluding the median). An underconfident model at level k would have $p > k$, i.e. more observations fall within a given interval than expected. In contrast, an overconfident model at level k would have $p < k$, i.e. fewer observations fall within a given interval than expected. We here focus on coverage at the $k = 0.5$ and $k = 0.95$ levels.

We also assessed the overall predictive performance of weekly forecasts using the weighted interval score (WIS) across all available quantiles. The WIS represents a parsimonious approach to scoring forecasts based on uncertainty represented as forecast values across a set of quantiles [29], and is a strictly proper scoring rule, that is, it is optimal for predictions that come from the data-generating model. As a consequence, the WIS encourages forecasters to report predictions representing their true belief about the future [30]. Each forecast for a given location and date is scored based on an observed count of weekly incidence, the median of the predictive distribution and the predictive upper and lower quantiles corresponding to the central predictive interval level.

Not all models provided forecasts for all locations and dates, and we needed to compare predictive performance in the face of various levels of missingness across each forecast target. Therefore we calculated a relative WIS. This is a measure of forecast performance which takes into account that different teams may not cover the same set of forecast targets (i.e., weeks and locations). The relative WIS is computed using a *pairwise comparison tournament* where for each pair of models a mean score ratio is computed based on the set of shared targets. The relative WIS of a model with respect to another model is then the ratio of their respective geometric mean of the mean score ratios, such that smaller values indicate better performance.

We scaled the relative WIS of each model with the relative WIS of a baseline model, for each forecast target, location, date, and horizon. The baseline model assumes case or death counts stay the same as the latest data point over all future horizons, with expanding uncertainty, described previously in [31]. Here we report the relative WIS of each model with respect to the baseline model.

Retrospective ensemble methods   We retrospectively explored alternative methods for combining forecasts for each target at each week. A natural way to combine probability distributions available in the quantile format used here is [32]

$$F^{-1}(\alpha) = \sum_{i=1}^{n} w_i F_i^{-1}(\alpha)$$

,

Where $F_1 \ldots F_n$ are the cumulative distribution functions of the individual probability distributions (in our case, the predictive distributions of each forecast model i contributed to the hub), w_i are a set of weights in $[0, 1]$; and α are the quantile levels such that

$$F^{-1}(\alpha) = \inf\{t : F_i(t) \geq \alpha\}$$

.

Different ensemble choices then mainly translate to the choice of weights w_i. An arithmetic mean ensemble uses weights at $w_i = 1/n$, where all weights are equal and sum up to 1.

Alternatively, we can choose a set of weights to apply to forecasts before they are combined. Numerous options exist for choosing these weights with the aim to maximise predictive performance, including choosing weights to reflect each forecast's past performance (thereby moving from an untrained to a trained ensemble). A straightforward choice is so-called inverse score weighting. In this case, the weights are calculated as

$$w_i = \frac{1}{S_i}$$

,

where S_i reflects the forecasting skill of forecaster i, normalised so that weights sum to 1. This method of weighting was found in the US to outperform unweighted scores during some time periods [33] but this was not confirmed in a similar study in Germany and Poland [11].

When constructing ensembles from quantile means, a single outlier can have an oversized effect on the ensemble forecast. Previous research has found that a median ensemble, replacing the arithmetic mean of each quantile with a median of the same values, yields competitive performance while maintaining robustness to outlying forecasts [34]. Building on this, we also created weighted median ensembles using the weights described above and a Harrel-Davis quantile estimator with a beta function to approximate the weighted percentiles [35]. We then compared the performance of unweighted and inverse relative WIS weighted mean and median ensembles.

Results

An example of weekly forecasts from the ensemble model is shown in Figure 1.

Over the whole study period, 26 independently participating forecasting teams contributed results from 28 unique forecasting models (see supplementary information (SI), Table 1). The number of models contributing to each ensemble forecast varied over time and by forecasting target (SI Figure 1). Not all modellers created forecasts for all locations, horizons, or variables. At most, 15 models contributed forecasts for cases in Germany at the 1 week horizon, with an accumulated 592 forecasts for that single target over the study period. In contrast, deaths in Finland at the 2 week horizon saw the smallest number of forecasts, with only 6 independent models contributing a total 24 forecasts. Similarly, not all teams forecast across all quantiles of the predictive distribution for each target, with only 23 models providing the full set of 23 quantiles. No ensemble forecast was composed of less than 3 independent models.

Using all models and the ensemble, we created 2106 forecasting scores where each score summarises a unique combination of forecasting model, variable, country, and week ahead horizon (SI Figure 2). We visually
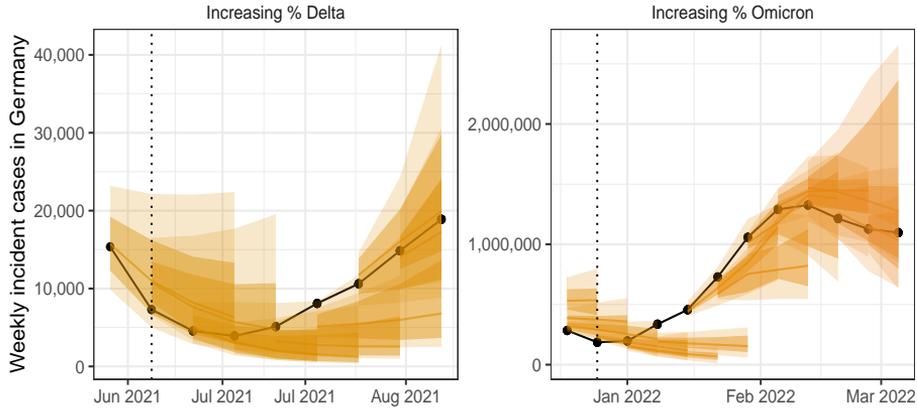
Figure 1: *Ensemble forecasts of weekly incident cases in Germany over periods of increasing SARS-CoV-2 variants Delta (B.1.617.2, left) and Omicron (B.1.1.529, right). Black indicates observed data. Coloured ribbons represent each weekly forecast of 1-4 weeks ahead (showing median, 50%, and 90% probability). For each variant, forecasts are shown over an x-axis bounded by the earliest dates at which 5% and 99% of sequenced cases were identified as the respective variant of concern, while vertical dotted lines indicate the approximate date that the variant reached dominance (>50% sequenced cases).*

compared the absolute performance of forecasts in predicting numbers of incident cases and deaths. We observed that forecasts predicted well in times of stable epidemic behaviour, while struggling to accurately predict at longer horizons around inflection points, for example during rapid changes in population-level behaviour or surveillance. Forecast models varied widely in their ability to predict and account for the introduction of new variants, giving the ensemble forecast over these periods a high level of uncertainty (Figure 1).

In relative terms, the ensemble of all models performed well compared to both its component models and the baseline. By relative WIS scaled against a baseline of 1 (where a score <1 indicates outperforming the baseline), the median score for participating models across all submitted forecasts was 1.04, while the median score of forecasts from the ensemble model was 0.71. Across all horizons and locations, the ensemble performed better on scaled relative WIS than 84% of participating model scores when forecasting cases (with a total N=862), and 92% of participating model scores for forecasts of incident deaths (N=746).

The performance of individual and ensemble forecasts varied by length of the forecast horizon (Figure 2). At each horizon, the typical performance of the ensemble outperformed both the baseline model and the aggregated scores of all its component models, although we saw wide variation between individual models in performance across horizons. Both individual models and the ensemble saw a trend of worsening performance at longer horizons when forecasting cases with the median scaled relative WIS of the ensemble across locations worsened from 0.62 for one-week ahead forecasts to 0.9 when forecasting four weeks ahead. Performance for forecasts of deaths was more stable over one through four weeks, with median ensemble performance moving from 0.69 to 0.76 across the four week horizons.

We observed similar trends in performance across horizon when considering how well the ensemble was calibrated with respect to the observed data. At one week ahead the case ensemble was well calibrated (ca. 50% and 95% nominal coverage at the 50% and 95% levels respectively). This did not hold at longer forecast horizons as the case forecasts became increasingly over-confident. Meanwhile, the ensemble of death forecasts was well calibrated at the 95% level across all horizons, and the calibration of death forecasts at the 50% level improved with lengthening horizons compared to being underconfident at shorter horizons.
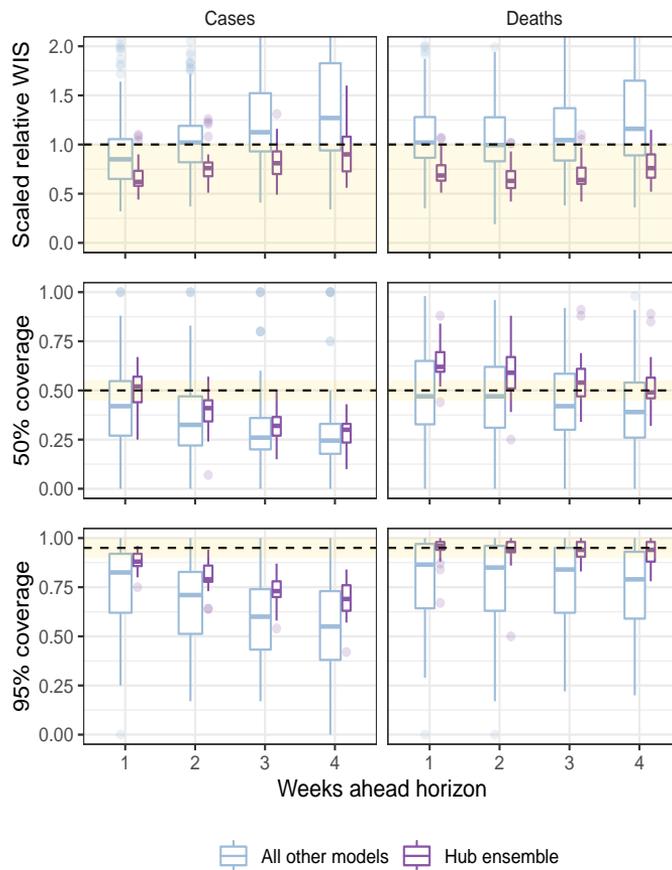
Figure 2: *Performance of short-term forecasts aggregated across all individually submitted models and the Hub ensemble, by horizon, forecasting cases (left) and deaths (right). Performance measured by relative weighted interval score scaled against a baseline (dotted line, 1), and coverage of uncertainty at the 50% and 95% levels. Boxplot, with width proportional to number of observations, show interquartile ranges with outlying scores as faded points. The target range for each set of scores is shaded in yellow.*
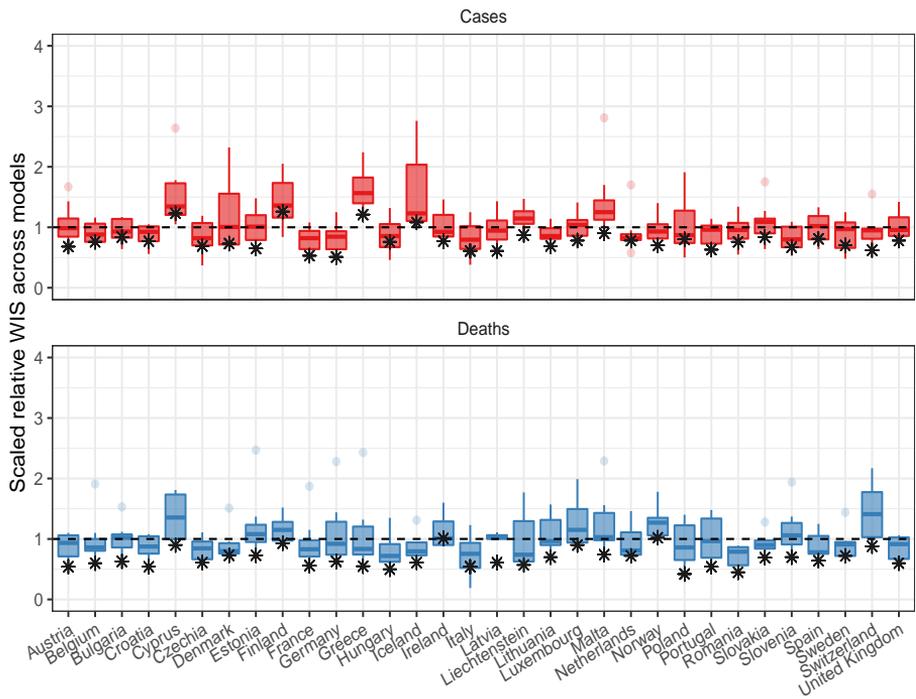
Figure 3: *Performance of short-term forecasts across models and median ensemble (asterisk), by country, forecasting cases (top) and deaths (bottom) for two-week ahead forecasts, according to the relative weighted interval score. Boxplots show interquartile ranges, with outliers as faded points, and the ensemble model performance is marked by an asterisk. y-axis is cut-off to an upper bound of 4 for readability.*

Table 1: Predictive performance of main ensembles, as measured by the scaled relative WIS.

Horizon	Weighted mean	Weighted median	Unweighted mean	Unweighted median
Cases				
1 week	0.59	0.62	0.59	0.61
2 weeks	0.67	0.67	0.67	0.67
3 weeks	0.79	0.70	0.81	0.71
4 weeks	1.06	0.75	1.09	0.79
Deaths				
1 week	0.63	0.59	1.00	0.59
2 weeks	0.57	0.54	0.81	0.53
3 weeks	0.64	0.56	0.83	0.54
4 weeks	0.83	0.64	0.82	0.62

The ensemble also performed consistently well in comparison to individual models when forecasting across countries (Figure 3). In total, across 32 countries forecasting for one through four weeks, when forecasting cases the ensemble outperformed 75% of component models in 21 countries, and outperformed all available models in 3 countries. When forecasting deaths, the ensemble outperformed 75% and 100% of models in 30 and 9 countries respectively. Considering only the the two-week horizon shown in Figure 3, the ensemble of case forecasts outperformed 75% models in 24 countries and all models in only 12 countries. At the two-week horizon for forecasts of deaths, the ensemble outperformed 75% and 100% of its component models in 30 and 26 countries respectively.

We considered alternative methods for creating ensembles from the participating forecasts, using either a mean or median to combine either weighted or unweighted forecasts (Table 1). Across locations we observed that the median outperformed the mean across all one through four week horizons and both cases and death targets, for all but cases at the 1 week horizon. This held regardless of whether the component forecasts were weighted or unweighted by their individual past performance. Between methods of combination, weighting made little difference to the performance of the median ensemble, but slightly improved performance of the mean ensemble.

Discussion

We collated 12 months of forecasts of COVID-19 cases and deaths across 32 countries in Europe, collecting from multiple independent teams and using a principled approach to standardising both forecast targets and the predictive distribution of forecasts. We combined these into an ensemble forecast and compared the relative performance of forecasts between models, finding that the ensemble forecasts outperformed most individual models across all countries and horizons over time.

Across all models we observed that forecasting changes in trend in real time was particularly challenging. Our study period included multiple fundamental changes in viral-, individual-, and population-level factors driving the transmission of COVID-19 across Europe. In early 2021, the introduction of vaccination started to change population-level associations between infections, cases, and deaths [36], while the Delta variant emerged and became dominant [37]. Similarly from late 2021 we saw the interaction of individually waning immunity during the emergence and global spread of the Omicron variant [38]. Neither the extent nor timing of these factors were uniform across European countries covered by the Forecast Hub [39]. This meant that the performance of any single forecasting model depended partly on the ability, speed, and precision with which it could adapt to new conditions for each forecast target.

We observed a contrast between a more stable performance of forecasting deaths further into the future compared to forecasts of cases. Previous work has found rapidly declining performance for case forecasts with increasing horizon [31], [40], while death forecasts can perform well with up to six weeks lead time

[41]. We can similarly link this to the specific epidemic dynamics in this study. COVID-19 has a typical serial interval of less than a week [42]. This implies that case forecasts of more than two weeks only remain valid if rates of transmission and detection remain stable over the entire forecast horizon. This is unlikely to have held given the rapid changes in epidemic dynamics across many countries in Europe. Meanwhile, we can interpret the higher reliability of death forecasts as due to the longer time lag between infection and death [43], and higher consistency of reporting deaths in surveillance data [44]. This allows forecasters to incorporate the effect of changes in transmission. Additionally, the performance of trend-based forecasts may have benefited from the slower changes to trends in incident deaths caused by gradually increasing vaccination rates.

We found the ensemble in this study continued to outperform both other models and the baseline at up to four weeks ahead. Our results support previous findings that ensemble forecasts are the best or nearly the best performing models with respect to absolute predictive performance and appropriate coverage of uncertainty [12], [14], [31]. While the ensemble was consistently high performing, it was not strictly dominant across all forecast targets, reflecting findings from previous comparable studies of COVID-19 forecasts [11], [45]. Our finding suggests the usefulness of an ensemble as a robust summary when forecasting across many spatio-temporal targets, without replacing the importance of communicating the full range of model predictions.

When exploring variations in ensemble methods, we found that the choice of median over means yielded the most consistent improvement in predictive performance, regardless of the method of weighting. Other work has supported the importance of the median in providing a stable forecast that better accounts for outlier forecasts than the mean [45], although though this finding may be dependent on the quality of the individual forecast submissions. In contrast, weighing models by past performance did not result in any consistent improvement in performance. This is in line with existing mixed evidence for any optimal ensemble method for combining short term probabilistic infectious disease forecasts. Many methods of combination have performed competitively in analyses of forecasts for COVID-19 in the US, including the simple mean and weighted approaches outperforming unweighted or median methods [33]. This contrasts with later analyses finding weighted methods to give similar performance to a median average [10], [45]. We can partly explain this inconsistency if performance of each method depends on the outcome being predicted (cases, deaths), its count (incident, cumulative) and absolute level, the changing disease dynamics, and the varying quality and quantity of forecasting teams over time.

We note several limitations in our approach to assessing the relative performance of an ensemble among forecast models. Our results are the outcome of evaluating forecasts against a specific performance metric and baseline, where multiple options for evaluation exist and the choice reflects the aim of the evaluation process. Further, our choice of baseline model affects the given performance scores in absolute terms, and more generally the choice of appropriate baseline for epidemic forecast models is not obvious when assessing infectious disease forecasts. The model used here is supported by previous work [31], yet previous evaluation in a similar context has suggested that choice of baseline affects relative performance in general [46], and future research should be done on the best choices of baseline models in the context of infectious disease epidemics.

Our assessment of forecast performance may further have been inaccurate due to limitations in the observed data against which we evaluated forecasts. We sourced data from a globally aggregated database to maintain compatibility across 32 countries [21]. However, this made it difficult to identify the origin of lags and inconsistencies between national data streams, and to what extent these could bias forecasts for different targets. In particular we saw some real time data revised retrospectively, introducing bias in either direction where the data used to create forecasts was not the same as that used to evaluate it. We attempted to mitigate this by using an automated process for determining data revisions, and excluding forecasts made at a time of missing, unreliable, or heavily revised data. More generally it is unclear if the expectation of observation revisions should be a feature built into forecasts. Further research is needed to understand the perspective of end-users of forecasts in order to assess this.

In an emergency setting, open access to visualised forecasts and underlying data is useful for researchers, policymakers, and the public. For forecast producers, an easily accessible comparison between results from different methods can highlight individual strengths and weaknesses and help prioritise new areas of work.

For forecast users, probabilistic information about the future can influence decisions in the present that can then change epidemic dynamics [1].

Existing participatory modelling efforts for COVID-19 have been useful for policy communication [2], while multi-country efforts have included only single models adapted to country-specific parameters [47]–[49]. By expanding participation to many modelling teams, our work was able to create robust ensemble forecasts across Europe while allowing comparison across forecasts built with different interpretations of current data, on a like for like scale in real time. At the same time, collating time-stamped predictions ensures that we can test true out-of-sample performance of models and avoid retrospective claims of performance. Testing the limits of forecasting ability with these comparisons forms an important part of communicating any model-based prediction to decision makers.

This study raises many further questions which could inform epidemic forecast modellers and users. The dataset created by the European Forecast Hub is an openly accessible, standardised, and extensively documented catalogue of real time forecasting work from a range of teams and models across Europe [24], and we recommend its use for further research on forecast performance. In the code developed for this study we provide a worked example of downloading and using both the forecasts and their evaluation scores [27].

Future work could explore the impact on forecast models of changing epidemiology at a broad spatial scale by combining analyses of trends and turning points in cases and deaths with forecast performance, or extending to include data on vaccination, variant, or policy changes over time. There is also much scope for future research into methods for combining forecasts to improve performance of an ensemble. This includes altering the inclusion criteria of forecast models based on different thresholds of past performance, excluding or including only forecasts that predict the lowest and highest values (trimming) [33], or using alternative weighting methods such as quantile regression averaging [12]. Exploring these questions would add to our understanding of real time performance, supporting and improving future forecasting efforts.

We see additional scope to adapt the Hub format to the changing COVID-19 situation across Europe. We have extended the Forecast Hub infrastructure to include short term forecasts for hospitalisations with COVID-19, which is a challenging task due to limited data across the locations covered by the hub. As the policy focus shifts from immediate response to anticipating changes brought by vaccinations or the geographic spread of new variants [39], we are also separately investigating models for longer term scenarios in addition to the short term forecasts in a similar framework to existing scenario modelling work in the US [50].

In conclusion, we have shown that during a rapidly evolving epidemic spreading through multiple populations, an ensemble forecast performed highly consistently across a large matrix of forecast targets, typically outperforming the majority of its separate component models and a naive baseline model. In addition, we have linked issues with the predictability of short-term case forecasts to underlying COVID-19 epidemiology, and shown that ensemble methods based on past model performance were unable to reliably improve forecast performance. Our work constitutes a step towards both unifying COVID-19 forecasts and improving our understanding of them.

References

[1]   P. van Basshuysen, L. White, D. Khosrowi, and M. Frisch, "Three Ways in Which Pandemic Models May Perform a Pandemic," *Erasmus Journal for Philosophy and Economics*, vol. 14, no. 1, 1, pp. 110-127-110-127, Jul. 2021, doi: 10.23941/ejpe.v14i1.582.

[2]   CDC, "Coronavirus Disease 2019 (COVID-19)," Feb. 11, 2020. https://www.cdc.gov/coronavirus/ 2019-ncov/science/forecasting/forecasting.html (accessed Jan. 09, 2022).

[3]   European Centre for Disease Prevention and Control, "Forecasting COVID-19 cases and deaths in Europe - new hub will support European pandemic planning," Apr. 22, 2021. https://www.ecdc. europa.eu/en/news-events/forecasting-covid-19-cases-and-deaths-europe-new-hub

[4]   J. Zelner, J. Riou, R. Etzioni, and A. Gelman, "Accounting for uncertainty during a pandemic," *PATTER*, vol. 2, no. 8, Aug. 2021, doi: 10.1016/j.patter.2021.100310.

[5]   L. P. James, J. A. Salomon, C. O. Buckee, and N. A. Menzies, "The Use and Misuse of Mathematical
       Modeling for Infectious Disease Policymaking: Lessons for the COVID-19 Pandemic," *Med Decis
       Making*, vol. 41, no. 4, pp. 379–385, May 2021, doi: 10.1177/0272989X21990391.

[6]   N. G. Reich *et al.*, "A collaborative multiyear, multimodel assessment of seasonal influenza fore-
       casting in the United States," *PNAS*, vol. 116, no. 8, pp. 3146–3154, Feb. 2019, doi:
       10.1073/pnas.1812594116.

[7]   M. A. Johansson *et al.*, "An open challenge to advance probabilistic forecasting for dengue epidemics,"
       PNAS, vol. 116, no. 48, pp. 24268–24274, Nov. 2019, doi: 10.1073/pnas.1909865116.

[8]   N. G. Reich *et al.*, "Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the
       U.S," *PLoS Comput Biol*, vol. 15, no. 11, p. e1007486, Nov. 2019, doi: 10.1371/journal.pcbi.1007486.

[9]   E. Y. Cramer *et al.*, "The United States COVID-19 Forecast Hub dataset," p. 2021.11.04.21265886,
       Nov. 2021, doi: 10.1101/2021.11.04.21265886.

[10]   E. L. Ray *et al.*, "Ensemble Forecasts of Coronavirus Disease 2019 (COVID-19) in the U.S." p.
       2020.08.19.20177493, Aug. 2020, doi: 10.1101/2020.08.19.20177493.

[11]   J. Bracher *et al.*, "A pre-registered short-term forecasting study of COVID-19 in Germany and Poland
       during the second wave," *Nat Commun*, vol. 12, no. 1, 1, p. 5173, Aug. 2021, doi: 10.1038/s41467-
       021-25207-0.

[12]   S. Funk *et al.*, "Short-term forecasts to inform the response to the Covid-19 epidemic in the UK,"
       medRxiv, p. 2020.11.11.20220962, Nov. 2020, doi: 10.1101/2020.11.11.20220962.

[13]   M. Bicher *et al.*, "Supporting COVID-19 Policy-Making with a Predictive Epidemiologi-
       cal Multi-Model Warning System," *medRxiv*, p. 2020.10.18.20214767, Apr. 2021, doi:
       10.1101/2020.10.18.20214767.

[14]   C. Viboud *et al.*, "The RAPIDD ebola forecasting challenge: Synthesis and lessons learnt," *Epidemics*,
       vol. 22, pp. 13–21, Mar. 2018, doi: 10.1016/j.epidem.2017.08.002.

[15]   R. Buizza, "Introduction to the special issue on '25 years of ensemble forecasting'," *Quarterly Journal
       of the Royal Meteorological Society*, vol. 145, no. S1, pp. 1–11, 2019, doi: 10.1002/qj.3370.

[16]   K. R. Moran *et al.*, "Epidemic Forecasting is Messier Than Weather Forecasting: The Role of Human
       Behavior and Internet Data Streams in Epidemic Forecast," *J Infect Dis*, vol. 214, pp. S404–S408,
       Dec. 2016, doi: 10.1093/infdis/jiw375.

[17]   European Covid-19 Forecast Hub, *European COVID-19 Forecast Hub.* covid19-forecast-hub-europe,
       2021.Available: https://github.com/covid19-forecast-hub-europe/covid19-forecast-hub-europe

[18]   E. Cramer *et al.*, "Reichlab/Covid19-forecast-hub: Release for Zenodo, 20210816," Aug. 2021, doi:
       10.5281/zenodo.5208210.

[19]   S. Y. Wang *et al.*, "Reichlab/covidHubUtils: Repository release for Zenodo," Aug. 2021, doi:
       10.5281/zenodo.5207940.

[20]   J. Bracher *et al.*, *The German and Polish COVID-19 Forecast Hub.* 2020.Available: https://github.
       com/KITmetricslab/covid19-forecast-hub-de

[21]   E. Dong, H. Du, and L. Gardner, "An interactive web-based dashboard to track COVID-19 in real
       time," *The Lancet Infectious Diseases*, vol. 20, no. 5, pp. 533–534, May 2020, doi: 10.1016/S1473-
       3099(20)30120-1.

[22]   European Covid-19 Forecast Hub, "Community." https://covid19forecasthub.eu/community.html

[23]   European Covid-19 Forecast Hub, "Covid19-forecast-hub-europe: Wiki." https://github.com/
       covid19-forecast-hub-europe/covid19-forecast-hub-europe

[24]   European Covid-19 Forecast Hub, "European Covid-19 Forecast Hub." https://covid19forecasthub.
       eu/index.html

[25]   EpiForecasts, "Project: ECDC European COVID-19 Forecast Hub - Zoltar," 2021. https://www.
       zoltardata.com/project/238

[26] N. G. Reich, M. Cornell, E. L. Ray, K. House, and K. Le, "The Zoltar forecast archive, a tool to standardize and store interdisciplinary prediction research," *Sci Data*, vol. 8, no. 1, 1, p. 59, Feb. 2021, doi: 10.1038/s41597-021-00839-5.

[27] *Predictive performance of multi-model ensemble forecasts of Covid-19 across European nations.* covid19-forecast-hub-europe, 2022.Available: https://github.com/covid19-forecast-hub-europe/euro-hub-ensemble

[28] Nikos I Bosse, Hugo Gruson, Sebastian Funk, EpiForecasts, and Sam Abbott, *Scoringutils: Utilities for Scoring and Assessing Predictions.* 2020.Available: https://github.com/epiforecasts/scoringutils

[29] J. Bracher, E. L. Ray, T. Gneiting, and N. G. Reich, "Evaluating epidemic forecasts in an interval format," *PLOS Computational Biology*, vol. 17, no. 2, p. e1008618, Feb. 2021, doi: 10.1371/journal.pcbi.1008618.

[30] T. Gneiting and A. E. Raftery, "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 359–378, Mar. 2007, doi: 10.1198/016214506000001437.

[31] E. Y. Cramer *et al.*, "Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US," *medRxiv*, p. 2021.02.03.21250974, Jan. 2021, doi: 10.1101/2021.02.03.21250974.

[32] C. Genest, "Vincentization Revisited," *The Annals of Statistics*, vol. 20, no. 2, pp. 1137–1142, 1992,Available: https://www.jstor.org/stable/2242003

[33] J. W. Taylor and K. S. Taylor, "Combining Probabilistic Forecasts of COVID-19 Mortality in the United States," *Eur J Oper Res*, Jun. 2021, doi: 10.1016/j.ejor.2021.06.044.

[34] E. L. Ray *et al.*, "Comparing trained and untrained probabilistic ensemble forecasts of COVID-19 cases and deaths in the United States," Jan. 28, 2022. Accessed: Mar. 30, 2022. [Online]. Available: http://arxiv.org/abs/2201.12387

[35] F. E. HARRELL and C. E. DAVIS, "A new distribution-free quantile estimator," *Biometrika*, vol. 69, no. 3, pp. 635–640, Dec. 1982, doi: 10.1093/biomet/69.3.635.

[36] European Centre for Disease Prevention and Control, "Interim guidance on the benefits of full vaccination against COVID-19 for transmission and implications for non-pharmaceutical interventions - 21 April 2021," ECDC, Stockholm, 2021.Available: https://www.ecdc.europa.eu/en/publications-data/interim-guidance-benefits-full-vaccination-against-covid-19-transmission

[37] European Centre for Disease Prevention and Control, "Threat Assessment Brief: Implications for the EU/EEA on the spread of the SARS-CoV-2 Delta (B.1.617.2) variant of concern," ECDC, Stockholm, Jun. 2021.Available: https://www.ecdc.europa.eu/en/publications-data/threat-assessment-emergence-and-impact-sars-cov-2-delta-variant

[38] European Centre for Disease Prevention and Control, "Assessment of the further spread and potential impact of the SARS-CoV-2 Omicron variant of concern in the EU/EEA, 19th update," Jan. 27, 2022. https://www.ecdc.europa.eu/en/publications-data/covid-19-omicron-risk-assessment-further-emergence-and-potential-impact

[39] European Centre for Disease Prevention and Control, "Overview of the implementation of COVID-19 vaccination strategies and deployment plans in the EU/EEA," ECDC, Stockholm, Nov. 2021.Available: https://www.ecdc.europa.eu/en/publications-data/overview-implementation-covid-19-vaccination-strategies-and-deployment-plans

[40] M. Castro, S. Ares, J. A. Cuesta, and S. Manrubia, "The turning point and end of an expanding epidemic cannot be precisely forecast," *Proceedings of the National Academy of Sciences*, vol. 117, no. 42, pp. 26190–26196, Oct. 2020, doi: 10.1073/pnas.2007868117.

[41] J. Friedman *et al.*, "Predictive performance of international COVID-19 mortality forecasting models," *Nat Commun*, vol. 12, no. 1, 1, p. 2609, May 2021, doi: 10.1038/s41467-021-22457-w.

[42] M. Alene, L. Yismaw, M. A. Assemie, D. B. Ketema, W. Gietaneh, and T. Y. Birhan, "Serial interval and incubation period of COVID-19: A systematic review and meta-analysis," *BMC Infectious Diseases*, vol. 21, no. 1, p. 257, Mar. 2021, doi: 10.1186/s12879-021-05950-x.

[43]   R. Jin, "The lag between daily reported Covid-19 cases and deaths and its relationship to age," *J Public Health Res*, vol. 10, no. 3, p. 2049, Mar. 2021, doi: 10.4081/jphr.2021.2049.

[44]   M. Català *et al.*, "Robust estimation of diagnostic rate and real incidence of COVID-19 for European policymakers," *PLOS ONE*, vol. 16, no. 1, p. e0243701, Jan. 2021, doi: 10.1371/journal.pone.0243701.

[45]   L. Brooks, "Comparing ensemble approaches for short-term probabilistic COVID-19 forecasts in the U.S." 2020. https://forecasters.org/blog/2020/10/28/comparing-ensemble-approaches-for-short-term-probabilistic-covid-19-forecasts-in-the-u-s/ (accessed Jul. 15, 2021).

[46]   J. Bracher *et al.*, "National and subnational short-term forecasting of COVID-19 in Germany and Poland, early 2021," p. 2021.11.05.21265810, Nov. 2021, doi: 10.1101/2021.11.05.21265810.

[47]   R. Aguas *et al.*, "Modelling the COVID-19 pandemic in context: An international participatory approach," *BMJ Global Health*, vol. 5, no. 12, p. e003126, Dec. 2020, doi: 10.1136/bmjgh-2020-003126.

[48]   K. Adib *et al.*, "A participatory modelling approach for investigating the spread of COVID-19 in countries of the Eastern Mediterranean Region to support public health decision-making," *BMJ Global Health*, vol. 6, no. 3, p. e005207, Mar. 2021, doi: 10.1136/bmjgh-2021-005207.

[49]   A. Agosto, A. Campmas, P. Giudici, and A. Renda, "Monitoring COVID-19 contagion growth," *Statistics in Medicine*, vol. 40, no. 18, pp. 4150–4160, 2021, doi: 10.1002/sim.9020.

[50]   R. K. Borchering, "Modeling of Future COVID-19 Cases, Hospitalizations, and Deaths, by Vaccination Rates and Nonpharmaceutical Intervention Scenarios — United States, April–September 2021," *MMWR Morb Mortal Wkly Rep*, vol. 70, 2021, doi: 10.15585/mmwr.mm7019e3.

SEIAR MODEL WITH ASYMPTOMATIC COHORT AND CONSEQUENCES TO EFFICIENCY OF QUARANTINE GOVERNMENT MEASURES IN COVID-19 OUTBREAK

A PREPRINT

Lenka Přibylová[*]
Department of Mathematics and Statistics
Masaryk University
Kotlářská 2, Brno, 611 37, Czech Republic
pribylova@math.muni.cz

Veronika Hajnová
Department of Mathematics and Statistics
Masaryk University
Kotlářská 2, Brno, 611 37, Czech Republic
hajnova@math.muni.cz

April 7, 2020

ABSTRACT

We present a compartmental SEIAR model of epidemic spread as a generalization of the SEIR model. We believe that the asymptomatic infectious cohort is an omitted part of the understanding of the epidemic dynamics of disease COVID-19. We introduce and derive the basic reproduction number as the weighted arithmetic mean of the basic reproduction numbers of the symptomatic and asymptomatic cohorts. Since the asymptomatic subjects people are not detected, they can spread the disease much longer, and this increases the COVID-19 R_0 up to around 9. We show that European epidemic outbreaks in various European countries correspond to the simulations with commonly used parameters based on clinical characteristics of the disease COVID-19, but R_0 is around three times bigger if the asymptomatic cohort is taken into account. Many voices in the academic world are drawing attention to the asymptomatic group of infectious subjects at present. We are convinced that the asymptomatic cohort plays a crucial role in the spread of the COVID-19 disease, and it has to be understood during government measures.

*K*eywords SEIR model · SEIAR model · quarantine · basic reproduction number · asymptomatic infectious cohort

1 Introduction

We started modeling the Wuhan COVID-19 outbreak with the standard SEIR model with WHO premised basic reproduction number R_0 between 2 and 3, the incubation period around 5 days, and the serial time around a week using GLEAMviz network simulator that includes populations, traffic, and measures. Soon we realized, the simulation predicts outbreaks in European countries too late. The outbreaks fitted better for R_0 much higher than it is generally considered. Afterwards, serious outbreaks in Italy, Spain, UK, and the US happened very quickly. The paper [1] appeared in Science, estimating 86% of undocumented infectious (published 16.3.2020). Padova University recently informed about an experiment in town Vò, where areal testing showed that the majority of the positive 3% of citizens were asymptomatic at the beginning of the outbreak. We made a hypothesis that the asymptomatic cohort can play a crucial role. This could also be a clue to understanding why early enough closing school policy contributes to slow down the outbreak. We propose a model that counts in with this asymptomatic infectious cohort, and we derive its basic reproduction number R_0. We proved that the basic reproduction number is a weighted average of reproduction numbers of both the symptomatic and asymptomatic cohorts, which seems to be very intuitive. On the other hand, this implies that in case of a disease with the majority of asymptomatic cases, this basic reproduction number R_0 is much higher than the one gained from standard estimates based only on the symptomatic cohort. This can explain the "invisible" and fast increase in COVID-19 outbreaks since it also acts through the asymptomatic cohort. There are already studies that

---

[*]corresponding author

estimate R_0 higher than WHO proposed (for example, Diamond Princess estimate 14.8 at [2]). Probably we can see just the tip of the iceberg. The consequences are very important since it gives a much shorter time to governments. Late measures have almost no effect.

2   SEIAR model and the basic reproduction number

We propose a simple model that generalizes the SEIR model commonly used for virus disease modeling. The total population is partitioned into the following compartments: S, Susceptible; E, Exposed; I, Infected (symptomatic infected); A, Asymptomatic (asymptomatic infected); R, Removed (healed or dead, no more infectious).

$$
\begin{align}
\dot{S} &= -\beta S(I + A), \tag{1}\\
\dot{E} &= \beta S(I + A) - \gamma E, \tag{2}\\
\dot{I} &= \gamma p E - \mu_1 I, \tag{3}\\
\dot{A} &= \gamma(1 - p)E - \mu_2 A, \tag{4}\\
\dot{R} &= \mu_1 I + \mu_2 A, \tag{5}
\end{align}
$$

where parameter β denotes the transmission rate (i.e., the probability of disease transmission in a single contact times the average number of contacts per person) due to contacts between a Susceptible subject and an Infected or an Asymptomatic subject[2]. Parameter β can be modified by quarantine government measures that increase social distancing as closing schools, remote working, using masks, or similar. Parameter γ usually denotes the probability rate at which the Exposed subject develops clinically relevant symptoms. The period $1/\gamma$ (days^{-1}) is called an incubation period. In the case of COVID-19, it is known that an Exposed subject is infectious one or two days before developing symptoms, and so we will define the Exposed compartment as a latent non-infectious compartment. Due to this, we assume that the period $1/\gamma$ is shorter than usually estimated (see [3], [4]). We assume that every Exposed subject becomes infectious, but subjects that enter I cohort later develop symptoms with probability p and those who not enter A cohort with probability $1 - p$. This assumption is based on recent studies [1] and experiments in the Italian town Vò. Parameters μ_1 or μ_2, respectively, denote the remove rate, so $\frac{1}{\mu_1}$ is the average period to isolation for Infected symptomatic subjects, and $\frac{1}{\mu_2}$ is the average recovery period for Asymptomatic subjects. We assume that $\frac{1}{\mu_2} > \frac{1}{\mu_1}$, since the COVID-19 patients are treated or isolated very quickly after developing symptoms. We omit the probability rate of becoming susceptible again after recovery, although this is not evident (mainly for Asymptomatic subjects). We assume $S + I + E + A + R = 1$ (a non-dimensionalized model with a constant incidence rate).

2.1   Unstability of the non-epidemic equilibrium – outbreak

The non-epidemic equilibrium of the system (1), (2), (3),(4) (we can separate independent equation (5)) is obviously $(1, 0, 0, 0)$. The Jacobian linearization matrix of the system (1), (2), (3),(4) is

$$
J = \begin{pmatrix}
-\beta(I + A) & 0 & -\beta S & -\beta S \\
\beta(I + A) & -\gamma & \beta S & \beta S \\
0 & \gamma p & -\mu_1 & 0 \\
0 & \gamma(1 - p) & 0 & -\mu_2
\end{pmatrix}
$$

and in the non-epidemic equilibrium it is

$$
J(1, 0, 0, 0) = \begin{pmatrix}
0 & 0 & -\beta & -\beta \\
0 & -\gamma & \beta & \beta \\
0 & \gamma p & -\mu_1 & 0 \\
0 & \gamma(1 - p) & 0 & -\mu_2
\end{pmatrix}. \tag{6}
$$

The Jacobian matrix (6) has one zero eigenvalue. The non-epidemic equilibrium loses stability as any eigenvalue of a submatrix

$$
A = \begin{pmatrix}
-\gamma & \beta & \beta \\
\gamma p & -\mu_1 & 0 \\
\gamma(1 - p) & 0 & -\mu_2
\end{pmatrix}
$$

---

[2]The model can be improved by incorporating different transmission rates for cohorts I and A as β_I and β_A, respectively. It is difficult to compare them since, on one side, people tend to avoid contact with subjects showing symptoms, but on the other side, an asymptomatic subject is less infectious more probably. Due to this dichotomy, we use one the mean β.

crosses imaginary axes and has positive real part. The characteristic polynomial of the matrix A is

$$p(\lambda) = \lambda^3 + (\gamma + \mu_1 + \mu_2)\lambda^2 + (-\beta\gamma + \gamma\mu_1 + \gamma\mu_2 + \mu_1\mu_2)\lambda + \beta\gamma p\mu_1 - \beta\gamma p\mu_2 - \beta\gamma\mu_1 + \gamma\mu_1\mu_2$$

and Routh-Hurwitz criterion

$$\gamma + \mu_1 + \mu_2 \; > \; 0, \tag{7}$$

$$-\det A = \beta\gamma p\mu_1 - \beta\gamma p\mu_2 - \beta\gamma\mu_1 + \gamma\mu_1\mu_2 \; > \; 0, \tag{8}$$

$$(\gamma + \mu_1 + \mu_2)(-\beta\gamma + \gamma\mu_1 + \gamma\mu_2 + \mu_1\mu_2) - (\beta\gamma p\mu_1 - \beta\gamma p\mu_2 - \beta\gamma\mu_1 + \gamma\mu_1\mu_2) \; > \; 0, \tag{9}$$

implies negative real parts of all the eigenvalues of the matrix A. The condition (7) is satisfied always. Violation of the condition (8) implies that at least one eigenvalue has positive real part. The condition (8) can be equivalently rewritten as

$$R_0 := \frac{\beta p}{\mu_1} + \frac{\beta(1-p)}{\mu_2} < 1 \tag{10}$$

and so the left-hand side as a weighted average of reproduction numbers of both the symptomatic and asymptomatic cohorts can be defined as $\tilde{R}_0$. If $\tilde{R}_0 > 1$ the epidemic outbreaks. Since the infectious period $1/\mu_2$ of an Asymptomatic subject is much longer than the infectious period $1/\mu_1$ of a symptomatic Infectious subject, and $1 - p$ is majority percentage for the COVID-19 disease, the derived R_0 is much higher than the one estimated using SEIR model ([5], [6], [7]). Violation of the condition (9) give birth to an unstable focus (two complex eigenvalues cross the imaginary axes).

2.2 Case Study: Numerical results for the COVID-19 Outbreak

We do not aim here to fit and predict the curve of the infected. We are trying to explain that the principal difference between SEIR and SEIAR models leads to principally different outcomes in for example times and rates of outbreaks, a percentage of the population affected. It also has implications for state measures that are being introduced to keep the disease contained and controlled under health system capacity.

For commonly used parameters of COVID-19 $\beta = 1$, $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$ and $p = 0.14$ (see [8], [5], [6], [7], [4]) we get $R_0 \doteq 9 > 1$ from (10). Condition (9) is satisfied.
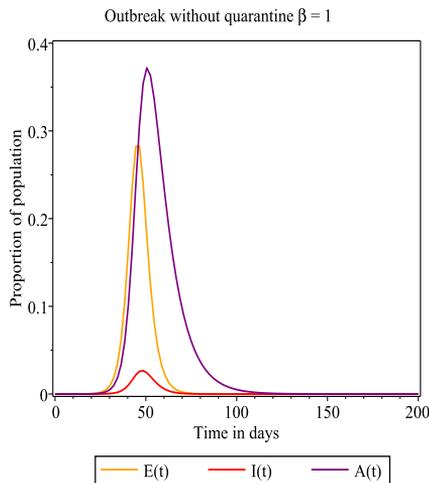

Figure 1: Dynamics of the model (1), (2), (3),(4). The peak time $t_p = 48$, a proportion of infectious in the population in the peak $I(48) = 0.0268$.

Figure 1 shows the disease dynamics in case that no strong restrictions are held (without social distancing and other measures except hospitalization or isolated home treatment). The health system has to contain around 2.7% of Infected

symptomatic population at once[3] at the moment of the peak (day 48 from the first Infected subject in a million population). Commonly, 20% of symptomatic Infected subjects are hospitalized, and 10% need intensive care ([9]). That is 0.27% of the whole population. 60% decrease of β (e.g., mask protection in the whole population) causes the peak to be halved in approximately doubled time. SEIR model without the Asymptomatic cohort gives five times higher and about 3 weeks later peak values (for $\beta = 1$), which seems to be very different from the actual dynamics of the disease COVID-19.

2.3 GLEAMviz simulation of COVID-19 pandemic

The GLEAMviz simulation software is a platform that gives a possibility to simulate the global pandemic on a network background. The Global Epidemic and Mobility Model (GLEAM) is a stochastic computational model that integrates high-resolution demographic and mobility data and uses a compartmental approach to define the epidemic characteristics of the infectious disease [10]. We simulated the pandemic starting in Wu-Han on the 1st of December using SEIAR model with parameters $\beta = 1$, $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$ and $p = 0.14$ and an exception that $\beta = 0.3$ from 20/01/2020 in the whole China (Figure 2). The outbreaks in Europe fit in time, unlike the much later onsets of the epidemics in case of using the SEIR model. Simulation-based on SEIAR showed good times of local epidemic outbreaks in Wu-Han, Italy, Germany, Czech Republic, and others (with differences in days, Figure 3) and much closer proportion of infected in a population (difference in the order). Due to the gradual cancellation of flights, air traffic slowed down, so we cannot simulate the real situation. Possibility to decrease air traffic during simulations could be a good software improvement.
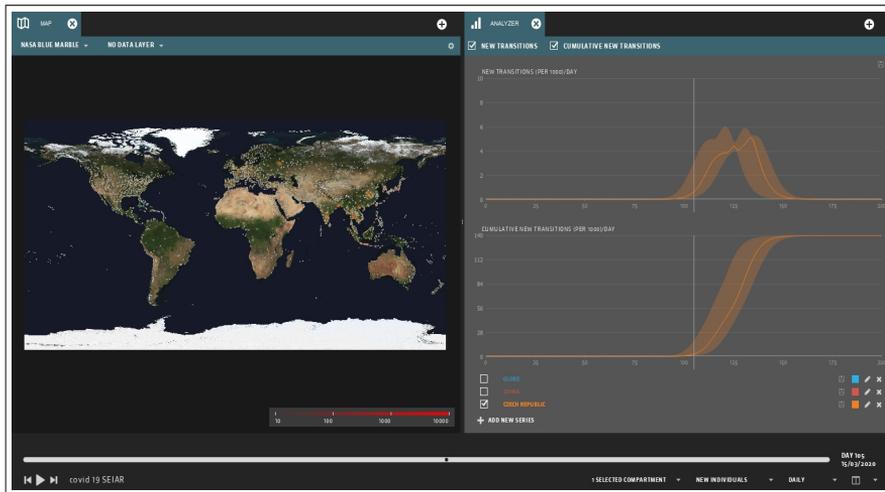

Figure 2: Global simulation of COVID-19 outbreak without European protective measures in GLEAMviz based on the model (1), (2), (3),(4) with parameters $\beta = 1$, $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$ and $p = 0.14$. Czech Republic – new transitions per 1000/day and cumulative transitions per 1000/day.

3   Efficiency of quarantine government measures with respect to the time of their introduction

One important information that governments need is to estimate the efficiency of introduced measures. For the SEIAR model, we will show the dependence of the peak height on the time of the introduction of quarantine and social measures that affect transmission rate β. It turns out that there is a time threshold for the introduction of quarantine and social

---

[3]This hypothetical case with no government interventions gives a peak around 280 000 people symptomatic altogether in Czech Republic.
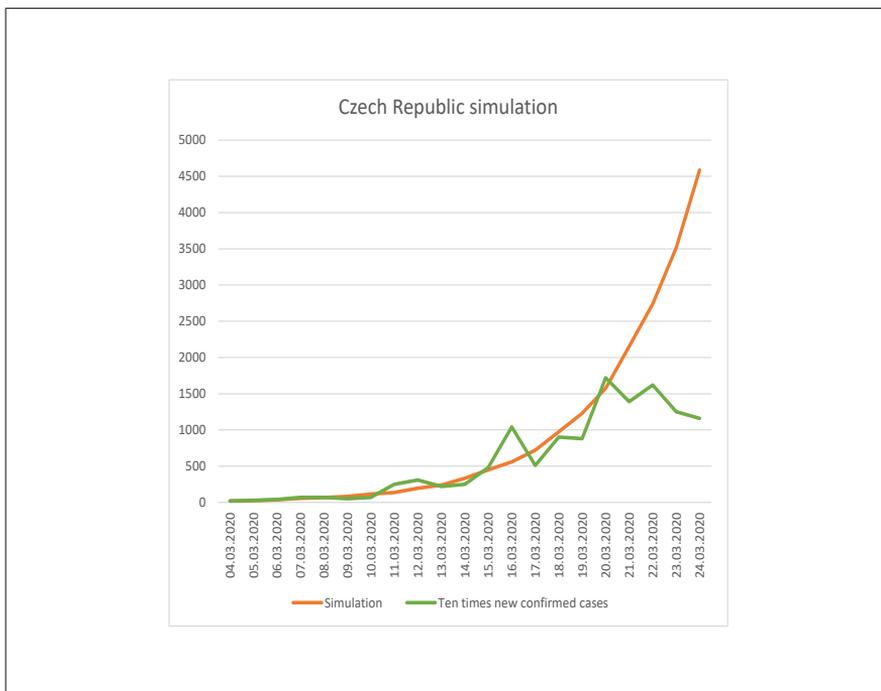
Figure 3: COVID-19 outbreak without European protective measures in GLEAMviz based on the model (1), (2), (3),(4) with parameters $\beta = 1$, $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$ and $p = 0.14$ – a comparison for Czech Republic new daily confirmed cases data set ([11]). Transmission rate β is decreased to 0.3 from 20/1/2020 for the whole China. Air traffic and local traffic is not decreased, there is no decrease of R_0 due to government measures in Europe, so the simulation fails after 20/03/2020 (around a week after first measures). Real data are 10 times lower (that could be due to the lack of testing at the beginning of the outbreak and undetected mild symptomatic cases) and delayed 16 days (due to the incubation period, developing symptoms period and testing period, plus possible time for undetected first cases). Simulation for SEIR model with the same parameters ($p = 1$, $R_0 = 3$) moves the outbreak more than a month later, and the infected cohort is more than five times bigger. This is evidence of the unsuitability of models without the asymptomatic infectious cohort since the parameters are based on clinical characteristics of the disease (commonly used, [8], [5], [6], [7], [4]).

measures for both SEIR and SEIAR models[4]. Early measures lead to a significant reduction of the epidemic peak, whereas after a certain threshold, the epidemic peak cannot be significantly affected.

We will show this threshold for parameters $\beta = 1$, $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$ and $p = 0.14$ used above . Figure 4 shows how the dynamics of the system change abruptly as β reduction from 1 to 0.3 is introduced at times $t = 30$ and $t = 40$. We have to highlight the fact that the time $t = 30$ is counted from the first Infected subject in a million population, but due to delay (the incubation period, few more days of asymptomatic disease, test duration to confirm positivity) first COVID-19 patients in a new country were confirmed after around 10 days.

3.1 Existence of a threshold

Depending on the time of the introduction of government measures that reduce β, we can draw the peak of the epidemic. The heatmap 5 may serve to gain a good insight into the effectiveness of government measures in terms of time and

---

[4]Model SEIR is a limit model SEIAR for $p = 1$.

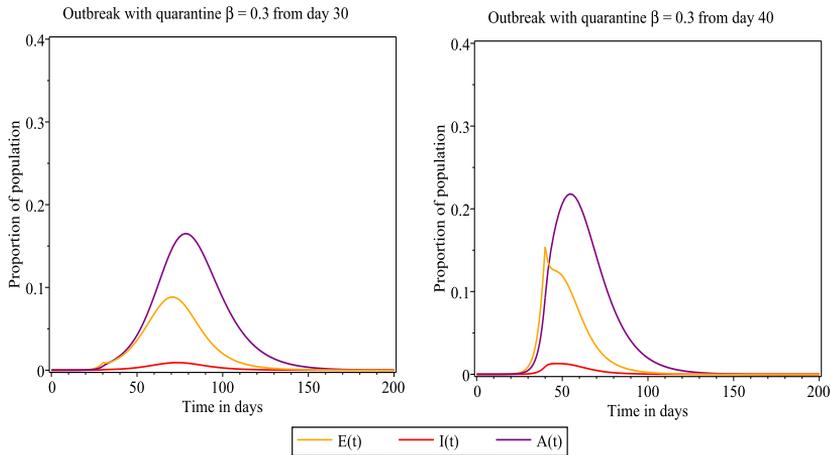Figure 4: β reduction from 1 to 0.3 is introduced at times $t = 30$ and $t = 40$. For $t = 30$ the peak time $t_p = 73$, a proportion of infectious in the population in the peak $I\,(48) = 0.0091$. For $t = 40$ the peak time $t_p = 47$, a proportion of infectious in the population in the peak $I\,(48) = 0.0130$.

strength. It displays a steep growth that is almost beta independent. This property is not exceptional for taken parameter values, but it is a generic behavior.

This could imply that countries that took action quickly (not later than 20-25 days from the first confirms) are probably on a good way to contain the outbreak. On the other hand, Italy, Spain, Switzerland, UK (with a lack of testing) or the US at highly populated areas face a sharp increase in the infected since the measures were introduced too late (this is of course not true for local areas with later outbreaks, where measures took place early enough).

3.2   Efficiency threshold estimation

It is challenging to estimate the efficiency of measures threshold during an ongoing outbreak. It is not possible to compare the time of the introduction of measures with the time of epidemic peak because we do not know it. A possible approach could be to compare the time of the fastest growth (denoted as Phase 2) with the time of introduction of measures.

Figure 6 shows the possibility to efficiently contain the outbreak by quarantine, social distancing, mask-wearing, and other government measures that decrease transmission rate β for various original R_0 (β respectively). The dashed line indicates the outbreak of the epidemic (the maximal steep), the dotted line is the peak time. The threshold is orange, but the measures have to be introduced a few days before it. Again, parameters $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$ with $p = 0.14$ are used. Figure 7 shows the quarantine efficiency as decrease of infected in peak (proportion of population) for various original R_0. Each column is labeled with a percentage that shows a relative difference with respect to the peak of the epidemic without a quarantine.

4   Conclusion

We presented a compartmental model of epidemic spread, which is a generalization of the SEIR model by adding an asymptomatic infectious cohort. The derived basic reproduction number is the weighted arithmetic mean of the basic reproduction numbers of the symptomatic and asymptomatic cohorts. In a limit case (for an empty asymptomatic cohort), the SEIAR model is the SEIR model with standard R_0. This model was developed based on studies and clinical characteristics of COVID-19, but can be used, for example, to model the dynamics of the H1N1 influenza epidemic and other human-to-human transmission diseases.
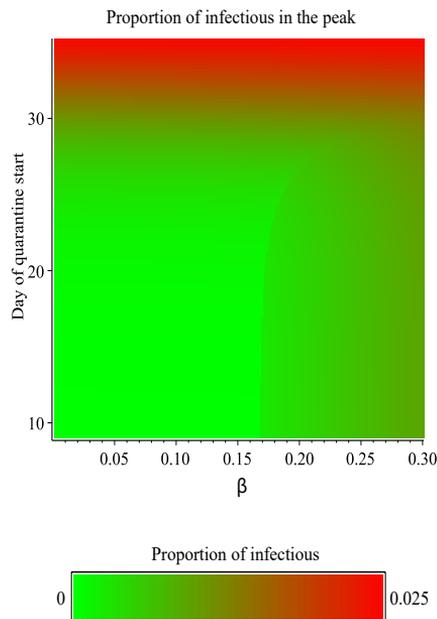
Figure 5: Proportion of infections in the peak depending on the strength of used measures (smaller β means larger measures), and a number of days, when quarantine starts, from the first detected case (approximately 10 days after infection started). Days are shown from day 9, which is an estimation of a start of the phase 2 (outbreak). In case of two peaks of the epidemic, it displays the first peak.

We are convinced that model SEIAR is the most simple model that can be used to simulate the epidemic dynamics of COVID-19. The standard SEIR model without an asymptomatic cohort does not match existing data, while the SEIAR model does quite well. Moreover, the onset of epidemics in different European countries corresponds to the simulations in GLEAMviz. The extremely rapid outbreaks in Italy and other countries, the functionality of only extreme measures, and the failure of measures that tried to protect and isolate only vulnerable groups are indirect supports for the suitability of the SEIAR model. Voices in the academic world are drawing attention to the asymptomatic group of infectious subjects. We are convinced that the asymptomatic cohort plays a crucial role in the spread of the COVID-19 disease.

The graph 5 indicates that all measures need to be taken early and vigorously to be effective. Late or insufficient measures have little effect on the outbreak. All measures have economic, social, logistical, and psychological consequences. Their balance is up to the authorities.

Acknowledgements

References

[1] Ruiyun Li, Sen Pei, Bin Chen, Yimeng Song, Tao Zhang, Wan Yang, and Jeffrey Shaman. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (sars-cov2). *Science*, 2020.
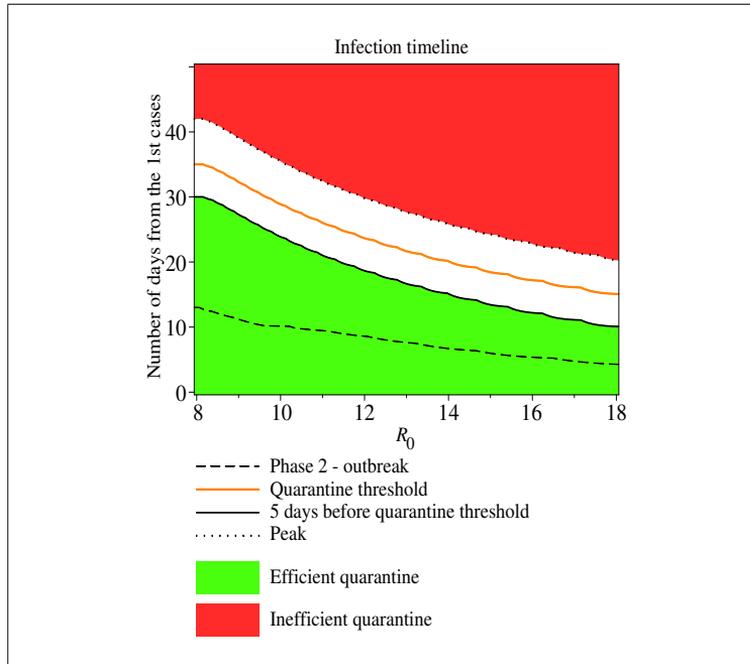
Figure 6: Graph of quarantine efficiency for drop of β to 0.3 from a level given by R_0 ($\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$, $p = 0.14$ β given by (10)). A period for efficient government measures is green – from the 1st positive case to few days before the orange threshold. The dashed line indicates the outbreak of epidemic (maximal steep – Phase 2). $R_0 = 9$ corresponds to clinical characteristics of COVID-19.

[2] Joacim Rocklöv, Henrik Sjödin, and Annelies Wilder-Smith. Covid-19 outbreak on the diamond princess cruise ship: estimating the epidemic potential and effectiveness of public health countermeasures. *Journal of Travel Medicine*, 2020.

[3] Cereda D, Tirani M, Rovida F, Demicheli V, Ajelli M, Poletti P, Trentini F, Guzzetta G, Marziano V, Barone A, Magoni M, Deandrea S, Diurno G, Lombardo M, Faccini M, Pan A, Bruno R, Pariani E, Grasselli G, Piatti A, Gramegna M, Baldanti F, Melegaro A, and Merler S. The early phase of the covid-19 outbreak in lombardy, italy, 2020.

[4] Wei-jie Guan, Zheng-yi Ni, Yu Hu, Wen-hua Liang, Chun-quan Ou, Jian-xing He, Lei Liu, Hong Shan, Chun-liang Lei, David SC Hui, et al. Clinical characteristics of coronavirus disease 2019 in china. *New England Journal of Medicine*, 2020.

[5] A Kucharski, T Russell, C Diamond, and Y Liu. Analysis and projections of transmission dynamics of ncov in wuhan. *CMMID repository*, 2, 2020.

[6] Qun Li, Xuhua Guan, Peng Wu, Xiaoye Wang, Lei Zhou, Yeqing Tong, Ruiqi Ren, Kathy SM Leung, Eric HY Lau, Jessica Y Wong, et al. Early transmission dynamics in wuhan, china, of novel coronavirus–infected pneumonia. *New England Journal of Medicine*, 2020.

[7] Joseph T Wu, Kathy Leung, and Gabriel M Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395 (10225):689–697, 2020.

[8] Stephen A Lauer, Kyra H Grantz, Qifang Bi, Forrest K Jones, Qulu Zheng, Hannah R Meredith, Andrew S Azman, Nicholas G Reich, and Justin Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 2020.
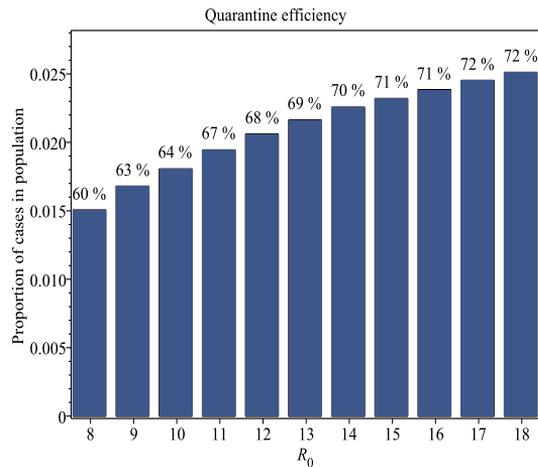
Figure 7: A difference between number of cases in the peak of the epidemic without a quarantine with $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$, $p = 0.14$ and R_0 given by (10), and the epidemic with a quarantine with $\gamma = 1/4$, $\mu_1 = 1/3$, $\mu_2 = 1/10$, $p = 0.14$, $\beta = 0.3$. The percentages show a relative difference with respect to the peak of the epidemic without a quarantine.

[9] Joint Mission WHO. Report of the who-china joint mission on coronavirus disease 2019 (covid-19), 2020. URL https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf. (accessed March 28, 2020).

[10] Wouter Van den Broeck, Corrado Gioannini, Bruno Gonçalves, Marco Quaggiotto, Vittoria Colizza, and Alessandro Vespignani. The gleamviz computational tool, a publicly available software to explore realistic epidemic spreading scenarios at the global scale. *BMC infectious diseases*, 11(1):37, 2011.

[11] Johns Hopkins Center for Systems Science and Engineering. Coronavirus covid-19 time series, 2020. URL https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series. (accessed March 28, 2020).