HØGSKOLEN I HARSTAD

Skriftserie

An Introduction to Discrete Dynamical Systems

Arild Wikan

Høgskolen i Harstad
Skriftserien 2007/04
Harstad University College

HØGSKOLEN I HARSTAD

HARSTAD UNIVERSITY COLLEGE

Tittel/Title	Nummer/Number
An Introduction to Discrete Dynamical Systems	2007/04
	Sider/Pages
	157

Forfatter/Author

Arild Wikan

Institutt/Department

Institute of Business Administration and Social Sciences

Prosjekt/Project

Sammendrag/Abstract

In most textbooks on dynamical systems focus is on continuous systems which lead to the study of differential equations rather than on discrete systems which lead to the study of maps or difference equations. However, in several scientific branches like biology and ecology there is a growing understanding that several problems in fact should be modelled by use of discrete time models rather by use of their continuous counterparts, and as is well known, the dynamical outcomes from discrete models are often substantial different from the outcomes from continuous models.

Therefore, we have in this text included a rigorous treatment of both linear and nonlinear maps (difference equations). In some sections we consider solution methods of specific equations or systems of equations, but most sections deal with qualitative theory of nonlinear maps which include topics like bifurcation theory, symbolic dynamics, Lyapunov exponent methods... Examples are mainly taken from biology and ecology. A few sections of discrete dynamic optimization problems are also included.

Stikkord/ Key Words	Discrete systems, solution methods, qualitative theory, stability, bifurcations, symbolic dynamics, optimization, population dynamics, age and stage structured models

ISBN 82-453-0215-5

ISSN 0807-2698

An Introduction to Discrete Dynamical Systems

Arild Wikan

An Introduction to Discrete Dynamical Systems

by

Arild Wikan

Harstad University College

Norway

and

Department of Economics and Management

University of Tromsø

Norway

Acknowledgements

My special thanks goes to Einar Mjølhus and Arne Eide. E. Mjølhus was the one who introduced me to the field, and later, both he and A. Eide have been coauthors on several research papers. Responses from K. Sydsæter, A. Seierstad and A. Strøm are also gratefully acknowledged.

I also wish to thank Liv Larssen for her excellent typing of this manuscript.

Financial support from Harstad University College is also gratefully acknowledged.

Finally I would like to thank my family for bearing over with me while writing this manuscript.

May 2007

Arild Wikan

Introduction

In most textbooks on dynamical systems focus is on continuous systems which lead to the study of differential equations rather than on discrete systems which lead to the study of maps or difference equations. This fact has in many respects an obvious historical background. Indeed, if we go back to the time of Newton, physical scientists were interested in problems within celestial mechanics, especially problems concerning the computation of planet motions, and the study of such kind of problems eventually lead to the study of differential equations. Later on, in other fields such as fluid mechanics, relativity, quantum mechanics, but also in other scientific branches like ecology, biology and economy it became clear that important problems could be formulated in an elegant and often simple way in terms of differential equations. However, to solve these (nonlinear) equations proved to be very difficult. Therefore, throughout the years, a rich and vast literature on dynamical systems has been established and the majority of the textbooks focuses on the continuous case.

The story of discrete systems is not that old. One major breakthrough came by Poincaré in the 1890's when he introduced the Poincaré map as a powerful tool in his qualitative approach towards the study of differential equations. Nearly fifty years later, Lewis and Leslie independently developed matrix models (often referred to as Leslie matrix models) in order to study populations with nonoverlapping age classes. These (mainly linear) difference equation models were almost forgotten in the years to come, but had their renaissance in the 70's and 80's when nonlinearities were included in the models. Examples of frequently quoted papers from that era are Guckenheimer et al. (1977) and the striped bass fishery model by Levin and Goodyear (1980). Later, through the work by Costantino, Cushing, Dennis and Desharnais (see Cushing (1998)) it became clear that small difference equation models indeed were capable not to analyse only, but also to predict nonstationary and chaotic behaviour in laboratory insect populations.

On the whole, there is a growing understanding in the biological and ecological communities that species which exhibit birth pulse fertilities (species that reproduce in a short time interval during a year) should be modelled by use of difference equations rather than differential equations, cf. the discussions in Cushing (1998) and Caswell (2001). Therefore, there is now much more interest of discrete dynamical systems than earlier.

Another important aspect which we also want to stress is the fact that in case of "low-dimensional problems" (problems with only one or two state variables) the possible dynamics found in nonlinear discrete models is much richer than in their continuous counterparts. Indeed, let us briefly illustrate this aspect through the following example:

Let $N = N(t)$ be the size of a population at time t. In 1837 Verhulst suggested that the change of N could be described by the differential equation (later known as the Verhulst equation)

$$\dot{N} = rN\left(1 - \frac{N}{K}\right) \tag{I1}$$

where the parameter r ($r > 0$) is the intrinsic growth rate at low densities and K is the carrying capacity. Now, define $x = N/K$. Then (I1) may be rewritten as

$$\dot{x} = rx(1 - x) \tag{I2}$$

which (as (I1) too) is nothing but a separable equation. Hence, it is straightforward to show that its solution becomes

$$x(t) = \frac{1}{1 - \frac{x_0 - 1}{x_0}e^{-rt}} \tag{I3}$$

where we also have used the initial condition $x(0) = x_0$. From (I3) we conclude that $x(t) \to 1$ as $t \to \infty$ which means that $x^* = 1$ is a stable fixed point of (I2). Moreover, regarding (I1) we have proved that the population N will settle at its carrying capacity K.

Next, let us turn to the discrete analogue of (I2). From (I2) it follows that

$$\frac{x_{t+1} - x_t}{\Delta t} = rx_t(1 - x_t) \tag{I4}$$

which implies

$$x_{t+1} = x_t + r\Delta t x_t - r\Delta t x_t^2 = (1 + r\Delta t)x_t\left(1 - \frac{r\Delta t}{1 + r\Delta t}x_t\right) \tag{I5}$$

and through the definition $y = r\Delta t(1 + r\Delta t)^{-1}x$ we easily obtain

$$y_{t+1} = \mu y_t(1 - y_t) \tag{I6}$$

where $\mu = 1 + r\Delta t$.

The "sweet and innocent-looking" equation (I6) is often referred to as the quadratic or the logistic equation. Its possible dynamical outcomes were presented by Sir Robert May in an influential review article called "Simple mathematical models with very complicated dynamics" in Nature (1976). There, he showed, depending on the value of the parameter μ, that the asymptotic behaviour of (I6) could be a stable fixed point (just as in (I2)), but also periodic solutions of both even and odd periods as well as chaoic behaviour. Thus the dynamic outcome of (I6) is richer and much more complicated than the behaviour of the continuous counterpart (I2).

In many respects, one of the major motivations for writing this text comes from the findings presented above.

Consequently, in *Part I*, we will develop the necessary qualitative theory which will enable us to understand the complex nature of first order nonlinear difference equations (or maps). Definitions, theorems and proofs shall be given in a general context, but most examples are taken from biology and ecology. Equation (I6) will on many occasions serve as a running example throughout the text. In *Part II* the theory will be extended to n-dimensional maps (or systems of difference equations). Here too, the theory will be illustrated and exemplified by use of population models from biology and ecology, In particular, Leslie matrix models and their relatives, stage structured models, shall frequently serve as examples. As a result of a request we have also included an introduction to discrete dynamic optimization problems which is presented in Part III. Finally, we want to repeat and stress that this is a Mathematics text so in order to be well prepared the potential reader should also have a background from a calculus course and also a prerequisite of topics from linear algebra, especially some knowledge of real and complex eigenvalues and associated eigenvectors. Regarding section 2.5 where the Hopf bifurcation is presented, the reader would also benefit from a somewhat deeper comprehension of complex numbers, This is all that is necessary really in order to establish the machinery we need in order to study the fascinating behaviour of nonlinear maps.

Contents

Part I
One-dimensional maps

$$f : \mathbb{R} \to \mathbb{R} \quad x \to f(x)$$

1.1 Preliminaries and definitions

Let $I \subset \mathbb{R}$ and $J \subset \mathbb{R}$ be two intervals. If f is a map from I to J we will express that as $f : I \to J$, $x \to f(x)$. Sometimes we will also express the map as a difference equation $x_{t+1} = f(x_t)$. If the map f depends on a parameter u we write $f_u(x)$ and say that f is a one-parameter family of maps.

For a given x_0, successive iterations of map f (or the difference equation $x_{t+1} = f(x_t)$) give: $x_1 = f(x_0)$, $x_2 = f(x_1) = f(f(x_0)) = f^2(x_0)$, $x_3 = f(x_2) = f(f^2(x_0)) = f^3(x_0) \ldots$, so after n iterations $x_{n+1} = f^n(x_0)$. Thus, the orbit of a map is a sequence of points $\{x_0, f(x_0), \ldots, f^n(x_0)\}$ which we for simplicity will write as $\{f^n(x_0)\}$. This is in contrast to the continuous case (differential equation) where the orbit is a curve.

Regarding differential equations it is a well-known fact that most classes of equations may not be solved explicitly. The same is certainly true for maps. However, the map $x \to f(x) = ax + b$ where a and b are constants is solvable.

Theorem 1.1.1. The difference equation

$$x_{t+1} = ax_t + b \tag{1.1.1}$$

has the solution

$$x_t = a^t\left(x_0 - \frac{b}{1-a}\right) + \frac{b}{1-a}, \qquad a \neq 1 \tag{1.1.2a}$$

$$x_t = x_0 + bt\,, \qquad a = 1 \tag{1.1.2b}$$

where x_0 is the initial value. -

Proof. From (1.1.1) we have $x_1 = ax_0 + b \Rightarrow x_2 = ax_1 + b = a(ax_0 + b) + b = a^2x_0 + (a+1)b \Rightarrow x_3 = ax_2 + b = \ldots = a^3x_0 + (a^2 + a + 1)b$. Thus assume $x_k = a^kx_0 + (a^{k-1} + a^{k-2} + \ldots + a + 1)b$. Then by induction: $x_{k+1} = ax_k + b = a\left[a^kx_0 + (a^{k-1} + a^{k-2} + \ldots + a + 1)b\right] + b = a^{k+1}x_0 + (a^k + a^{k-1} + \ldots + a + 1)b$. If $a \neq 1$: $1 + a + \ldots + a^k = (1 - a^t)(1 - a)^{-1}$ so the solution becomes

$$x_t = a^tx_0 + \frac{1-a^t}{1-a}b = a^t\left(x_0 - \frac{b}{1-a}\right) + \frac{b}{1-a}$$

If $a = 1$: $1 + a + \ldots + a^{t-1} = t \cdot 1 = t$

$$x_t = x_0 + bt$$

□

Regarding the asymptotic behaviour (long-time behaviour) we have from Theorem 1.1.1: If $|a| < 1$ $\lim_{t\to\infty} x_t = b/(1-a)$. (If $x_0 = b/(1-a)$ this is true for any $a \neq 1$.) If $a > 1$ and $x_0 \neq b/(1-a)$ the result is exponential growth or decay, and finally, if $a < -1$ divergent oscillations is the outcome.

If $b = 0$, (1.1.1) becomes

$$x_{t+1} = ax_t \tag{1.1.3}$$

which we will refer to as the linear difference equation. The solution is

$$x_t = a^t x_0 \tag{1.1.4}$$

Hence, whenever $|a| < 1$, $x_t \to 0$ asymptotically (as a convergent oscillation if $-1 < a < 0$). $a > 1$ or $a < -1$ gives exponential growth or divergent oscillations respectively.

Exercise 1.1.1. Solve and describe the asymptotic behaviour of the equations:

a) $x_{t+1} = 2x_t + 4$, $x_0 = 1$,

b) $3x_{t+1} = x_t + 2$, $x_0 = 2$. □

Exercise 1.1.2. Denote $x^* = b/(1-a)$ where $a \neq 1$ and describe the asymptotic behaviour of equation (1.1.1) in the following cases:

a) $0 < a < 1$ and $x_0 < x^*$,

b) $-1 < a < 0$ and $x_0 < x^*$,

c) $a > 1$ and $x_0 > x^*$. □

Equations of the form $x_{t+1} + ax_t = f(t)$, for example $x_{t+1} - 2x_t = t^2 + 1$, may be regarded as special cases of the more general situation

$$x_{t+n} + a_1 x_{t+n-1} + a_2 x_{t+n-2} + \cdots + a_n x_t = f(t)\,, \quad n = 1, 2, \ldots$$

Such equations are treated in Section 2.1 (cf. Theorem 2.1.6, see also examples following equation (2.1.6) and Exercise 2.1.5).

—

When the map $x \to f(x)$ is nonlinear (for example $x \to 2x(1-x)$) there are no solution methods so information of the asymptotic behaviour must be obtained by use of qualitative theory.

Definition 1.1.1. A fixed point x^* for the map $x \to f(x)$ is a point which satisfies the equation $x^* = f(x^*)$. □

Fixed points are of great importance to us and the following theorem will be very useful.

Theorem 1.1.2.

a) Let $I = [a, b]$ be an interval and let $f : I \to I$ be continuous. Then f has at least one fixed point in I.

b) Suppose in addition that $|f'(x)| < 1$ for all $x \in I$. Then there exists a unique fixed point for f in I, and moreover

$$|f(x) - f(y)| < |x - y|$$

□

Proof.

a) Define $g(x) = f(x) - x$. Clearly, $g(x)$ too is continuous. Suppose $f(a) > a$ and $f(b) < b$. Then $g(a) > 0$ and $g(b) < 0$ so the intermediate value theorem from elementary calculus directly gives the existence of c such that $g(c) = 0$. Hence, $c = f(c)$.

b) From a) we know that there is at least one fixed point. Suppose that both x and y ($x \neq y$) are fixed points. Then according to the mean value theorem from elementary calculus there exists c between x and y such that $f(x) - f(y) = f'(c)(x - y)$. This yields (since $x = f(x)$, $y = f(y)$) that

$$f'(c) = \frac{f(x) - f(y)}{x - y} = 1$$

This contradicts $|f'(x)| < 1$. Thus $x = y$ so the fixed point is unique. Further from the mean value theorem:

$$|f(x) - f(y)| = |f'(c)|\,|x - y| < |x - y|$$

□

Definition 1.1.2. Consider the map $x \to f(x)$. The point p is called a periodic point of period n if $p = f^n(p)$. The least $n > 0$ for which $p = f^n(p)$ is referred to as the prime period of p.

Note that a fixed point may be regarded as a periodic point of period one. □

Exercise 1.1.3. Find the fixed points and the period two points of $f(x) = x^3$.

□

Definition 1.1.3. If $f'(c) = 0$, c is called a critical point of f. c is nondegenerate if $f''(c) \neq 0$, degenerate if $f''(c) = 0$. □

The derivative of the n-th iterate $f^n(x)$ is easy to compute by use of the chain rule. Observe that $f^n(x) = f(f^{n-1}(x))$, $f^{n-1}(x) = f(f^{n-2}(x)) \ldots$, $f^2(x) = f(f(x))$. Consequently:

$$f^{n\prime}(x) = f'(f^{n-1}(x))f'(f^{n-2}(x)) \ldots f'(x)' \tag{1.1.5}$$

(1.1.5) enables us to compute the derivative of points on a periodic orbit in an elegant way. Indeed, suppose the three cycle $\{p_0, p_1, p_2\}$ where $p_1 = f(p_0)$, $p_2 = f(p_1) = f^2(p_0)$ and $f^3(p_0) = p_0 \ldots$. Then

$$f^{3\prime}(p_0) = f'(p_2)f'(p_1)f'(p_0) \tag{1.1.6}$$

Obviously, if we have n periodic points $\{p_0, \ldots, p_{n-1}\}$ the corresponding formulae is

$$f^{n\prime}(p_0) = \prod_{i=0}^{n-1} f'(p_i) \tag{1.1.7}$$

(Later on we shall use the derivative in order to decide whether a periodic orbit is stable or not. (1.1.7) implies that all points on the orbit is stable (unstable) simultaneously.)

We will now proceed by introducing some maps (difference equations) that have been frequently applied in population dynamcis. Examples that show how to compute fixed points, periodic points, etc., will be taken from these maps. Some computations are performed in the next section, others are postponed to Section 1.3.

1.2 One-parameter family of maps

Here we shall briefly present some one-parameter family of maps which have often been applied in population dynamical studies. Since x is supposed to be the size of a population, $x \geq 0$.

The map

$$x \to f_\mu(x) = \mu x(1-x) \tag{1.2.1}$$

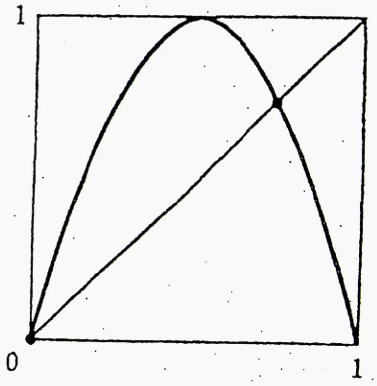
is often referred to as the quadratic or the logistic map. The parameter μ is called the intrinsic growth rate. Clearly $x \in [0,1]$, otherwise $x_t > 1 \Rightarrow x_{t+1} < 0$. If $\mu \in [0,4]$ any iterate of f_μ will remain in $[0,1]$. Further we may notice that $f_\mu(0) = f_\mu(1) = 0$ and $x = c = 1/2$ is the only critical point. Hence (1.2.1) is a unimodal map on the unit interval.

The map

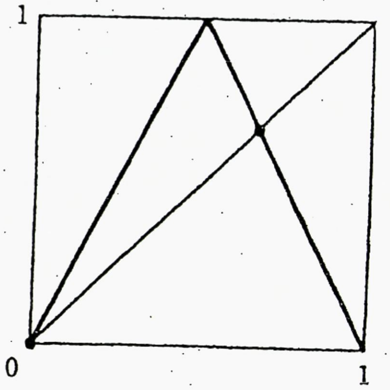
$$x \to f_r(x) = xe^{r(1-x)} \tag{1.2.2}$$

is called the Ricker map. Unlike the quadratic map, $x \in [0, \to\rangle$. The parameter r is positive.

Exercise 1.2.1. Show that the fixed points of (1.2.2) are 0 and 1 and that the critical point is $1/r$. □



(a)



(b)

Figure 1: The graphs of the functions: (a) $f(x) = 4x(1-x)$ (cf. (1.2.1)), and (b) the tent function (cf. (1.2.4) where $a = 2$).

The property that $x \in [0, \rightarrow\rangle$ makes (1.2.2) much more preferable to biologists than (1.2.1).

The map

$$x \to f_{a,b}(x) = \frac{ax}{(1+x)^b} \tag{1.2.3}$$

where $a > 1$, $b > 1$ is a two-parameter family of maps and is called the Hassel family.

Exercise 1.2.2. Show that $x = 0$ and $x = a^{1/b} - 1$ are the fixed points of (1.2.3) and that $c = 1/(b-1)$ is the critical point. □

The map

$$x \to f_a(x) = \begin{cases} ax & 0 \le x \le 1/2 \\ a(1-x) & 1/2 < x \le 1 \end{cases} \tag{1.2.4}$$

where $a > 0$ is called the tent map for obvious reasons. We will pay special attention to the case $a = 2$. Note that $f_a(x)$ attains its maximum at $x = 1/2$ but that $f'(1/2)$ does not exist.

All functions defined in (1.2.1)–(1.2.4) are called one-humped functions for obvious reasons. In Figure 1a we show the graph of the quadratic functions (1.2.1) ($\mu = 4$) and in Figure 1b the "tent" function (1.2.4) ($a = 2$). In both figures we have also drawn the line $y = x$ and we have marked e fixed points of the maps with dots.

As we have seen, maps (1.2.1)–(1.2.4) share much of the same properties. Our next goal is to explore this fact further.

Definition 1.2.1. Let $f : U \to U$ and $g : V \to V$ be two maps. If there exists a homeomorphism $h : U \to V$ such that $h \circ g = g \circ h$, then f and g are said to be topological equivalent. □

Remark. A function h is a homeomorphism if it is one-to-one, onto and continuous and that h^{-1} is also continuous. □

The important property of topological equivalent maps is that their dynamics is equivalent. Indeed, suppose that $x = f(x)$. Then from the definition, $h(f(x)) = h(x) = g(h(x))$, so if x is a fixed point of f, $h(x)$ is a fixed point for g. In a similar way, if p is a periodic point of f of period n (i.e. $f^n(p) = p$) we have from Definition 1.2.1 that $f = h^{-1} \circ g \circ h \Rightarrow f^2 = (h^{-1} \circ g \circ h) \circ (h^{-1} \circ g \circ h) = h^{-1} \circ g^2 \circ h$ so clearly $f^n = h^{-1} \circ g^n \circ h$. Consequently, $h(f^n(p)) = h(p) = g^n(h(p))$ so $h(p)$ is a periodic point of period n for g.

Proposition 1.2.1. The quadratic map $f : [0,1] \to [0,1]$ $x \to f(x) = 4x(1-x)$ is topological equivalent to the tent map

$$T : [0,1] \to [0,1] \quad x \to T(x) = \begin{cases} 2x & 0 \le x \le 1/2 \\ 2(1-x) & 1/2 < x \le 1 \end{cases}$$

□

Proof. We must find a function h such that $h \circ f = T \circ h$. Note that this implies that we also have $f \circ h^{-1} = h^{-1} \circ T$ where h^{-1} is the inverse of h.

Now, define $h^{-1}(x) = \sin^2(\pi x)/2$. Then

$$\begin{aligned} f \circ h^{-1} &= f\left(\sin^2 \frac{\pi x}{2}\right) = 4 \sin^2 \frac{\pi x}{2}\left(1 - \sin^2 \frac{\pi x}{2}\right) \\ &= 4 \sin^2 \frac{\pi x}{2} \cos^2 \frac{\pi x}{2} = \left(2 \sin \frac{\pi x}{2} \cos \frac{\pi x}{2}\right)^2 = \sin^2 \pi x \end{aligned}$$

$$h^{-1} \circ T = h^{-1}(2x) = \sin^2 \pi x \qquad 0 \le x \le \frac{1}{2}$$

$$h^{-1} \circ T = h^{-1}(2(1-x)) = \sin^2(\pi - \pi x) = \sin^2 \pi x \qquad \frac{1}{2} < x \le 1$$

Thus, $f \circ h^{-1} = h^{-1} \circ T$ which implies $h \circ f = T \circ h$ so f and T are topological equivalent. □

1.3 Fixed points and periodic points of the quadratic map

Most of the theory that we shall develop in the next sections will be illustrated by use of the quadratic map (1.2.1). In many respects (1.2.1) will serve as a running example. Therefore, in order to prepare the ground we are here going to list some main properties.

The fixed points are obtained from $x = \mu x(1-x)$. Thus the fixed points are $x^* = 0$ (the trivial fixed point) and $x^* = (\mu - 1)/\mu$ (the nontrivial fixed point). Note that the nontrivial fixed point is positive whenever $\mu > 1$. Assuming that (1.2.1) has periodic points of period two they must be found from $p = f_\mu^2(p)$ and since

$$f_\mu^2(p) = f(\mu p(1-p)) = \mu^2 p[1 - (\mu+1)p - 2\mu p^2 - \mu p^3]$$

the two nontrivial periodic points must satisfy the cubic equation

$$\mu^3 p^3 - 2\mu^3 p^2 + \mu^2(\mu+1)p + 1 - \mu^2 = 0 \tag{1.3.1}$$

Clearly, $p = (\mu-1)/\mu$ is a solution of (1.3.1) so after polynomial division we arrive at

$$\mu^2 p^2 - (\mu^2 + \mu)p + \mu + 1 = 0 \tag{1.3.2}$$

Thus, the periodic points are

$$p_{1,2} = \frac{\mu + 1 \pm \sqrt{(\mu+1)(\mu-3)}}{2\mu} \tag{1.3.3}$$

where $\mu > 3$ is a necessary condition for real solutions.

Period three points are obtained from $p = f_\mu^3(p)$ and must be found by means of numerical methods. (Newton's method works excellent.) (It is possible to show after a somewhat cumbersome calculation that the three periodic points do not exist unless $\mu > 1 + \sqrt{8}$.)

In general, it is a hopeless task to compute periodic points of period n for a given map when n becomes large. However, considering (1.2.1) it is in fact possible in the special case $\mu = 4$ as we now will demonstrate.

Consider the difference equation

$$x_{t+1} = 4x_t(1 - x_t) \tag{1.3.4}$$

Let $x_t = \sin^2 \varphi_t$. Then from (1.3.4):

$$\sin^2 \varphi_{t+1} = 4 \sin^2 \varphi_t \cos^2 \varphi_t = \sin^2 2\varphi_t$$

Further:

$$\begin{aligned} \sin^2 \varphi_{t+2} &= 4 \sin^2 \varphi_{t+1}(1 - \sin^2 \varphi_{t+1}) \\ &= 4 \sin^2 2\varphi_t \cos^2 2\varphi_t = \sin^2 2^2\varphi_t \end{aligned}$$

Thus, after n iterations

$$\sin^2 \varphi_{t+n} = \sin^2 2^n \varphi_t$$

and moreover:

$$\varphi_{t+n} \pm 2^n \varphi_t + l\pi$$

Now, if we have a period n orbit ($x_{t+n} = x_t$)

$$\sin^2 \varphi_{t+n} = \sin^2 \varphi_t$$

Hence:

$$\begin{aligned} \varphi_{t+n} = \pm\varphi_t + m\pi &\Leftrightarrow \pm 2^n \varphi_t + l\pi = \pm\varphi_t + m\pi \\ &\Leftrightarrow (2^n \pm 1)\varphi_t = (m - l)\pi \end{aligned}$$

so

$$\varphi_t = \frac{k\pi}{2^n \pm 1}$$

where $k = m - l$. Consequently, the periodic points are given by

$$p_i = \sin^2 \frac{k\pi}{2^n \pm 1} \tag{1.3.5}$$

Example 1.3.1. Compute all the period 1, period 2 and period 3 points of $f(x) = 4x(1-x)$. The period 1 points (which of course are the same as the fixed points) are

$$\sin^2 \frac{\pi}{2-1} = 0 \quad \text{and} \quad \sin^2 \frac{\pi}{2+1} = 0.75$$

The period 2 points are the period 1 points (which do not have prime period 2) plus the prime period 2 points.

$$\sin^2 \frac{\pi}{5} = 0.34549 \quad \text{and} \quad \sin^2 \frac{2\pi}{5} = 0.904508$$

(The latter points may of course also be obtained from (1.3.3).)

There are six points of prime period 3. The points

$$\sin^2 \frac{\pi}{7} = 0.188255\,, \quad \sin^2 \frac{2\pi}{7} = 0.611260 \quad \text{and} \quad \sin^2 \frac{4\pi}{7} = 0.950484$$

are the periodic points in one 3-cycle, while the points

$$\sin^2 \frac{\pi}{9} = 0.116977\,, \quad \sin^2 \frac{2\pi}{9} = 0.4131759 \quad \text{and} \quad \sin^2 \frac{4\pi}{9} = 0.969846$$

are the periodic points on another orbit. (The reason why it is one 2-cycle but two 3-cycles is strongly related to how they are created.) □

Example 1.3.2. Use (1.3.5) to find all the period 4 points of $f(x) = 4x(1-x)$. How many periodic points are there? □

Since $f(x) = 4x(1-x)$ is topological equivalent to the tent map we may use (1.3.5) together with Proposition 1.2.1 to find the periodic points of the tent map. Indeed, since $h^{-1}(x) = \sin^2(\pi x/2) \Rightarrow h(x) = (2/\pi)\arcsin\sqrt{x}$ (cf. the proof of Proposition 1.2.1) the periodic points p of $T(x)$ may be found from $T(h(p)) = T((2/\pi)\arcsin\sqrt{p})$. Thus the fixed points of the tent map are

$$T\left(\frac{2}{\pi}\arcsin\sqrt{0}\right) = \frac{4}{\pi}\arcsin 0 = 0$$

$$T\left(\frac{2}{\pi}\arcsin\sqrt{\frac{3}{4}}\right) = 2\left(1 - \frac{2}{\pi}\arcsin\sqrt{\frac{3}{4}}\right) = 0.6666$$

Exercise 1.3.1. Find the period 2 points of the tent map ($a = 2$). □

We shall close this section by computing numerically some orbits of the quadratic map for different values of the parameter μ:

$$\mu = 1.8 \text{ and } x_0 = 0.8 \text{ gives the orbit}$$

$$\{0.8\ \ 0.2880\ \ 0.3691\ \ 0.4192\ \ 0.4382\ \ 0.4431\ \ 0.4442\ \ 0.4444\ \ 0.4444\ \ldots\}$$

Thus the orbit converges towards the point 0.4444 which is nothing but the fixed point $(\mu - 1)/\mu$. In this case the fixed point is said to be locally asymptotic stable. (A precise definition will be given in the next section.)

$$\mu = 3.2 \text{ and } x_0 = 0.6 \text{ gives:}$$

$$\{0.6\ \ 0.7680\ \ 0.5702\ \ 0.7842\ \ 0.5415\ \ 0.7945\ \ 0.5225\ \ 0.7984\ \ 0.5151$$
$$0.7993\ \ 0.5134\ \ 0.7994\ \ 0.5131\ \ 0.7995\ \ 0.5130\ \ 0.7995\ \ 0.5130\ \ldots\}$$

Thus in this case the orbit does not converge towards the fixed point. Instead we find that the asymptotic behaviour is a stable periodic orbit of prime period 2. The points in the two-cycle are given by (1.3.3).

$$\mu = 4.0 \text{ and } x_0 = 0.30 \text{ gives}$$

$$\{0.30\ \ 0.84\ \ 0.5376\ \ 0.9943\ \ 0.02249\ \ 0.0879\ \ 0.3208\ \ 0.8716\ \ 0.4476\ \ 0.9890\ \ldots\}$$

Although care should be taken by drawing a conclusion after a few iterations only, the last example suggests that there are no stable periodic orbit when $\mu = 4$. (A formal proof of this fact will be given later.)

Exercise 1.3.2. Use a calculator or a computer to repeat the calculations above but use the initial values 0.6, 0.7 and 0.32 instead of 0.8, 0.6 and 0.3, respectively. Establish the fact that the long-time behaviour of the map when $\mu = 1.8$ or $\mu = 3.2$ is not sensitive to a slightly change of the initial conditions but that there is a strong sensitivity in the last case. □

1.4 Stability

Referring to the last example of the previous section we found that the equation $x_{t+1} = 1.8x_t(1-x_t)$ apparently possessed a stable fixed point and that the equation $x_{t+1} = 3.2x_t(1-x_t)$ did not. Both these equations are special cases of the quadratic family (1.2.1) so what the example suggests is that by increasing the parameter μ in (1.2.1) there exists a threshold value μ_0 where the fixed point of (1.2.1) loses its stability.

Now, consider the general first order nonlinear equation

$$x_{t+1} = f_\mu(x_t) \tag{1.4.1}$$

where μ is a parameter. The fixed point x^* satisfies $x^* = f_\mu(x^*)$.

In order to study the system close to x^* we write $x = x^* + h$ and expand f_μ in its Taylor series around x^* taking only the linear term. Thus:

$$x^* + h_{t+1} \approx f_\mu(x^*) + \frac{df}{dx}(x^*)h_t \tag{1.4.2}$$

which gives

$$h_{t+1} = \frac{df}{dx}(x^*)h_t \tag{1.4.3}$$

We call (1.4.3) the linearization of (1.4.1). The solution of (1.4.3) is given by (1.1.4). Hence, if $|(df/dx)(x^*)| < 1$, $\lim_{t\to\infty} h_t = 0$ whcich means that x_t will converge towards the fixed point x^*.

Now, we make the following definitions:

Definition 1.4.1. Let x^* be a fixed point of equation (1.4.1). If $|\lambda| = |(df/dx)(x^*)| \neq 1$ then x^* is called a hyperbolic fixed point. λ is called the eigenvalue. □

Definition 1.4.2. Let x^* be a hyperbolic fixed point. If $|\lambda| < 1$ then x^* is called a locally asymptotic stable hyperbolic fixed point. □

Example 1.4.1. Assume that $\mu > 1$ and find the parameter interval where the fixed point $x^* = (\mu - 1)/\mu$ of the quadratic map is stable.

Solution: $f_\mu(x) = \mu x(1-x)$ implies that $f'(x) = \mu(1-2x) \Rightarrow |\lambda| = |f'(x^*)| = |2-\mu|$. Hence from Definition 1.4.2, $1 < \mu < 3$ ensures that x^* is a locally asymptotic stable fixed point (which is consistent with our finding in the last example in the previous section). □

It is clear from Definition 1.4.2 that x^* is a *locally* stable fixed point. A formal argument that there exists an open interval U around x^* so that whenever $|f'(x^*)| < 1$ and $x \in U$ and that $\lim_{n\to\infty} f^n(x) = x^*$ goes like this:

By the continuity of f (f is C') there exists an $\varepsilon > 0$ such that $|f'(x)| < K < 1$ for $x \in [x^* - \varepsilon, x^* + \varepsilon]$. Successive use of the mean value theorem then implies

$$\begin{aligned}|f^n(x) - x^*| &= |f^n(x) - f^n(x^*)| = |f(f^{n-1}(x)) - f(f^{n-1}(x^*))| \\ &\leq K|f^{n-1}(x) - f^{n-1}(x^*)| \leq K^2|f^{n-2}(x) - f^{n-2}(x^*)| \\ &\leq \ldots \leq K^n|x - x^*| < |x - x^*| < \varepsilon\end{aligned}$$

so $f^n(x) \to x^*$ as $n \to \infty$.

Motivated by the preceding argument we define:

Definition 1.4.3. Let x^* be a hyperbolic fixed point. We define the local stable and unstable manifolds of x^*, $W^s_{\text{loc}}(x^*)$, $W^u_{\text{loc}}(x^*)$ as

$$W^s_{\text{loc}}(x^*) = \{x \in U / f^n(x) \to x^* \quad \text{as } n \to \infty \text{ and } f^n(x) \in U \text{ for all } n \geq 0\}$$

$$W^u_{\text{loc}}(x^*) = \{x \in U / f^n(x) \to x^* \quad \text{as } n \to -\infty \text{ and } f^n(x) \in U \text{ for all } n \leq 0\}$$

where U is a neighbourhood of the fixed point x^*. □

The definition of a hyperbolic stable fixed point is easily extended to periodic points.

Definition 1.4.4. Let p be a periodic point of (prime) period n so that $|f^{n'}(p)| < 1$. Then p is called an attracting periodic point. □

Example 1.4.2. Show that the periodic points 0.5130 and 0.7995 of $x_{t+1} = 3.2x_t(1-x_t)$ are stable and thereby proving that the difference equation has a stable 2-periodic attractor.

Solution: Since $f(x) = 3.2x(1-x) \Rightarrow f'(x) = 3.2(1-2x)$ we have from the chain rule (1.1.7) that $f^{2\prime}(0.5130) = f'(0.7995)f'(0.5130) = -0.0615$. Consequently, according to Definition 1.4.4, the periodic points are stable. □

Exercise 1.4.1. Use formulae (1.3.3) and compute the two-periodic points of the quadratic map in case of $\mu = 3.8$. Is the corresponding two-periodic orbit stable or unstable? □

Exercise 1.4.2. When $\mu = 3.839$ the quadratic map has two 3-cycles. One of the cycles consists of the points 0.14989, 0.48917 and 0.9593 while the other consists of the points 0.16904, 0.53925 and 0.95384. Show that one of the 3-cycles is stable and that the other one is unstable. □

Let us close this section by discussing the concept structural stability. Roughly speaking, a map f is said to be structurally stable if a map g which is obtained through a small perturbation of f has essentially the same dynamics as f, so intuitively this means that the distance between f and g and the distance between their derivatives should be small.

Definition 1.4.5. The C^1 distance between a map f and another map g is given by

$$\sup_{x\in\mathbb{R}}(|f(x) - g(x)|,\ |f'(x) - g'(x)|) \tag{1.4.4}$$

□

By use of Definition 1.4.5 we may now define structural stability in the following way:

Definition 1.4.6. The map f is said to be C^1 structurally stable on an interval I if there exists $\varepsilon > 0$ such that whenever (1.4.4) $< \varepsilon$ on I, f is topological equivalent to g. □

To prove that a given map is structurally stable may be difficult, especially in higher dimensional systems. However, our main interest is to focus on cases where a map is not structurally stable. In many respects maps with nonhyperbolic fixed points are standard examples of such maps as we now will demonstrate.

Example 1.4.3. When $\mu = 1$ the quadratic map is not structurally stable. Indeed, consider $x \to f(x) = x(1-x)$ and the perturbation $x \to g(x) = x(1-x)+\varepsilon$. Obviously, $x^* = 0$ is the fixed point of f and since $|\lambda| = |f'(0)| = 1$, x^* is a nonhyperbolic fixed point. Moreover, the C^1 distance between f and g is $|\varepsilon|$. Regarding g, the fixed points are easily found to be $\overline{x} = \pm\sqrt{\varepsilon}$. Hence, for $\varepsilon > 0$ there are two fixed points and $\varepsilon < 0$ gives no fixed points. Consequently, f is not structurally stable. □

Example 1.4.4. When $\mu = 3$ the quadratic map is not structurally stable. Let $x \to f(x) = 3x(1-x)$ and $x \to g(x) = 3x(1-x) + \varepsilon$ and again we notice that their C^1 distance is ε. Regarding f, the fixed points are $x_1^* = 0$ and $x_2^* = 2/3$. Further, $|\lambda_1| = |f'(0)| = 3$, $|\lambda_2| = |f'(2/3)| = 1$. Thus x_1^* is a repelling hyperbolic fixed point while x_2^* is nonhyperbolic. Considering g, the fixed points are $\overline{x}_1 = (1/3)(1-\sqrt{1+3\varepsilon})$ and $\overline{x}_2 = (1/3)(1+\sqrt{1+3\varepsilon})$. Note that $\varepsilon = 0 \Rightarrow \overline{x}_1 = x_1^*$, $\overline{x}_2 = x_2^*$.) Further, $|\sigma_1| = |g'(\overline{x}_1)| = |1+2\sqrt{1+3\varepsilon}|$ and $|\sigma_2| = |g'(\overline{x}_2)| = |1-2\sqrt{1+3\varepsilon}|$. Whatever the sign of ε, $\overline{x}_1$ is clearly a repelling fixed point (just as x_1^*) since $\sigma_1 > 1$. Regarding $\overline{x}_2$ it is stable in case of $\varepsilon < 0$ and unstable if $\varepsilon > 0$.

The equation $x = g^2(x)$ may be expressed as

$$-27x^4 + 54x^3 + (18\varepsilon - 36)x^2 + (8 - 18\varepsilon)x + 4\varepsilon - 3\varepsilon^2 = 0 \tag{1.4.5}$$

and since $\overline{x}_1$ and $\overline{x}_2$ are solutions of (1.4.5) we may use polynomial division to obtain

$$9x^2 - 12x - 3e + 4 = 0 \tag{1.4.6}$$

which has the solutions $x_{1,2} = (2/3)(1 \pm \sqrt{3\varepsilon})$. Thus there exists a two-periodic orbit in case of $\varepsilon > 0$.

Moreover, cf. (1.1.7) $g^{2'} = g'(x_1)g'(x_2) = 9(1-2x_1)(1-2x_2) = 1 - 48\varepsilon$ which implies that the two-periodic orbit is stable in case of $\varepsilon > 0$, ε small. Consequently, when $\varepsilon > 0$ there is a fundamental structurally difference between

f and g so f cannot be structurally stable. (Note that the problem is the nonhyperbolic fixed point, not the hyperbolic one.) □

As suggested by the previous examples a major reason why a map may fail to be structurally stable is the presence of the nonhyperbolic fixed point. Therefore it is in many respects natural to introduce the following definition:

Definition 1.4.7. Let x^* be a hyperbolic fixed point of a map $f : \mathbb{R} \to \mathbb{R}$. If there exists a neighbourhood U around x^* and an $\varepsilon > 0$ such that a map g is $C^1 - \varepsilon$ close to f on U and f is topological equivalent to g whenever (1.4.4) $< \varepsilon$ on this neighbourhood, then f is said to be C^1 locally structurally stable. □

There is a major general result on topological equivalent maps known under the name Hartman and Grobman's theorem. The "one-dimensional" formulation of this theorem (cf. Devaney, 1989) is:

Theorem 1.4.1. Let x^* be a hyperbolic fixed point of a map $f : \mathbb{R} \to \mathbb{R}$ and suppose that $\lambda = f'(x^*)$ such that $|\lambda| \neq 0, 1$. Then there is a neighbourhood U around x^* and a neighbourhood V of $0 \in \mathbb{R}$ and a homeomorphism $h : U \to \mathbb{R}$ which conjugates f on U to the linear map $l(x) = \lambda x$ on V. □

For a proof, cf. Hartman (1964).

Example 1.4.5. Consider $x \to f(x) = (5/2)x(1-x)$. The fixed point is $x^* = 3/5$ and is clearly hyperbolic since $\lambda = f'(x^*) = -1/2$. Therefore, according to Theorem 1.4.1, $f(x)$ on a neighbourhood about $3/5$ is topological equivalent to $l(x) = -(1/2)x$ on a neighbourhood about 0. □

1.5 Bifurcations

As we have seen, the map $x \to f_\mu(x) = \mu x(1-x)$ has a stable hyperbolic fixed point $x^* = (\mu-1)/\mu$ provided $1 < \mu < 3$. If $\mu = 3$, $\lambda = f'(x^*) = -1$, hence x^* is no longer hyperbolic. If $\mu = 3.2$ we have shown that there exists a stable 2-periodic orbit. Thus x^* experiences a fundamental change of structure when it fails to be hyperbolic which in our running example occurs when $\mu = 3$. Such a point will from now on be referred to as a bifurcation point. When $\lambda = -1$, as in our example, the bifurcation is called a flip or a period doubling bifurcation. If $\lambda = 1$ it is called a saddle-node bifurcation. Generally, we will refer to a flip bifurcation as supercritical if the eigenvalue λ crosses the value -1 outwards and that the 2-periodic orbit just beyond the bifurcation point is stable. Otherwise the bifurcation is classified as subcritical.

Theorem 1.5.1. Let $f_\mu : \mathbb{R} \to \mathbb{R}$, $x \to f_\mu(x)$ be a one-parameter family of maps and assume that there is a fixed point (x^*, μ_0) where the eigenvalue equals -1. Assume

$$a = \left(\frac{\partial f_\mu}{\partial \mu}\frac{\partial^2 f_\mu}{\partial x^2} + 2\frac{\partial^2 f_\mu}{\partial x \partial \mu}\right) = \frac{\partial f_\mu}{\partial \mu}\frac{\partial^2 f_\mu}{\partial x^2} - \left(\frac{\partial f_\mu}{\partial x} - 1\right)\frac{\partial^2 f}{\partial x \partial \mu} \neq 0 \text{ at } (x^*, \mu_0)$$

and

$$b = \left(\frac{1}{2}\left(\frac{\partial^2 f_\mu}{\partial x^2}\right)^2 + \frac{1}{3}\left(\frac{\partial^3 f_\mu}{\partial x^3}\right)\right) \neq 0 \text{ at } (x^*, \mu_0)$$

Then there is a smooth curve of fixed points of f_μ which is passing through (x^*, μ_0) and which changes stability at (x^*, μ_0). There is also a curve consisting of hyperbolic period-2 points passing through (x^*, μ_0). If $b > 0$ the hyperbolic period-2 points are stable, i.e. the bifurcation is supercritical. □

Proof. Through a coordinate transformation it suffices to consider f_μ so that for $\mu = \mu_0 = 0$ we have $f(x^*, 0) = x^*$ and $f'(x^*, 0) = -1$.

First we show that one without loss of generality may assume that $x^* = 0$. To this end, define $F(x, \mu) = f(x, \mu) - x$. Then $F'(x^*, \mu) = -2 \neq 0$ and by use of the implicit function theorem there exists a solution $\overline{x}(\mu)$ of $F(x, \mu) = 0$.

Next, define $g(y,\mu) = f(y+\overline{x}(\mu),\mu) - \overline{x}(\mu)$. Clearly, $g(0,\mu) = 0$ for all μ. Consequently, $y = 0$ is a fixed point so in the following it suffices to consider $x \to f(x)$ where $x^*(\mu) = 0$ and $f'(0,0) = -1$.

The Taylor expansion around $(x^*,\mu) = (0,0)$ is

$$
\begin{aligned}
g_\eta(\xi) &= -\xi + \frac{\partial f}{\partial \mu}\eta + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\xi^2 + 2\frac{\partial^2 f}{\partial x \partial \mu}\xi\eta\right) + \frac{1}{6}\frac{\partial^3 f}{\partial x^3}\xi^3 + \text{higher order} \\
&= -\xi + a\eta + b\xi^2 + c\eta\xi + d\xi^3 + \text{higher order}
\end{aligned}
$$

where the parameter η has the same weight as ξ^2. The composite $(g \circ g)(\xi)$ may be expressed as

$$
g_\eta^2(\xi) = \xi + \alpha\eta\xi + \beta\xi^3 + \text{higher order}
$$

Thus, in order to have a system to study we must assume $\alpha, \beta \neq 0$ which is equivalent to

$$
\begin{aligned}
\alpha = -(2c + 2ab) &= -\left(2\frac{\partial^2 f}{\partial x \partial \mu} + 2\frac{\partial f}{\partial \mu}\cdot\frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right) \neq 0 \\
\beta = -(2d + 2b^2) &= -\left(2\cdot\frac{1}{6}\frac{\partial^3 f}{\partial x^3} + 2\left(\frac{1}{2}\frac{\partial^2 f}{\partial x^2}\right)^2\right) \neq 0
\end{aligned}
$$

and we recognize the derivative formulaes as nothing but what is stated in the theorem.

Next, consider the truncated map

$$
\xi^2 \to h(\xi) = \xi + \alpha\eta\xi + \beta\xi^3
$$

Clearly, the fixed points are

$$
\overline{\xi}_1 = 0, \qquad \overline{\xi}_{2,3} = \pm\sqrt{-\frac{\alpha}{\beta}\eta}
$$

Further, $h'(\xi) = 1 + \alpha\eta + 3\beta\xi^2$ so $h'(\overline{\xi}_1) = 1 + \alpha\eta$ and $h'(\overline{\xi}_{2,3}) = 1 - 2\alpha\eta$. Thus we have the following configurations (see Figure 2), and we may conclude that the stable period-2 orbits corresponds to $\beta < 0$, i.e.

$$
\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}\right)^2 + \frac{1}{3}\frac{\partial^3 f}{\partial x^3} > 0
$$

□

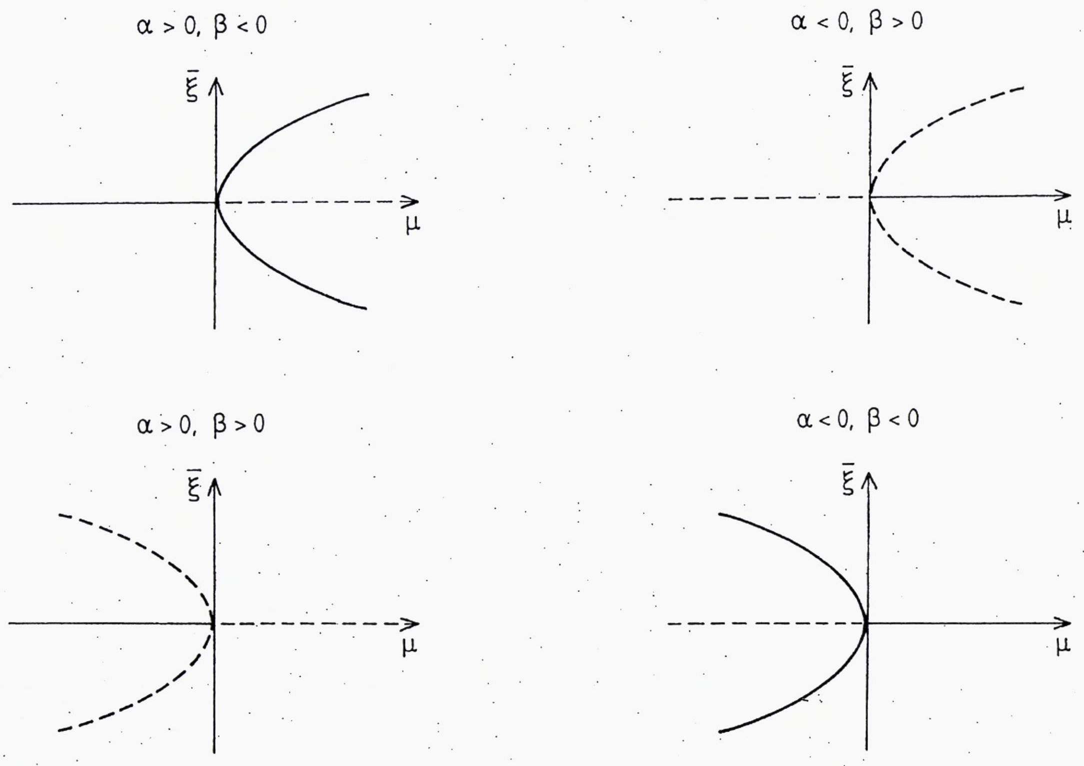


Figure 2: The possible configurations of $\xi^2 \to h(\xi) = \xi + \alpha\eta\xi + \beta\xi^3$.

Example 1.5.1. Show that the fixed point of the quadratic map undergoes a supercritical flip bifurcation at the threshold $\mu = 3$.

Solution: From the previous section we know that $x^* = 2/3$ and $f'(x^*) = -1$ when $\mu = 3$. We must show that the quantities a and b in Theorem 1.5.1 are different from zero and larger than zero respectively. By computing the various derivatives at $(x^*, \mu_0) = (2/3, 3)$ we obtain:

$$a = \frac{2}{9}(-6) + 2\left(-\frac{1}{3}\right) = -2 \neq 0 \quad \text{and} \quad b = \frac{1}{2}(-6)^2 + \frac{1}{3} \cdot 0 = 18 > 0$$

Thus the flip bifurcation is supercritical. When x^* fails to be stable, a stable period-2 orbit is established. □

Exercise 1.5.1. Show that the Ricker map $x \to x\exp[r(1-x)]$, cf. (1.2.2), undergoes a supercritical flip bifurcation at $(x^*, r) = (1, 2)$. □

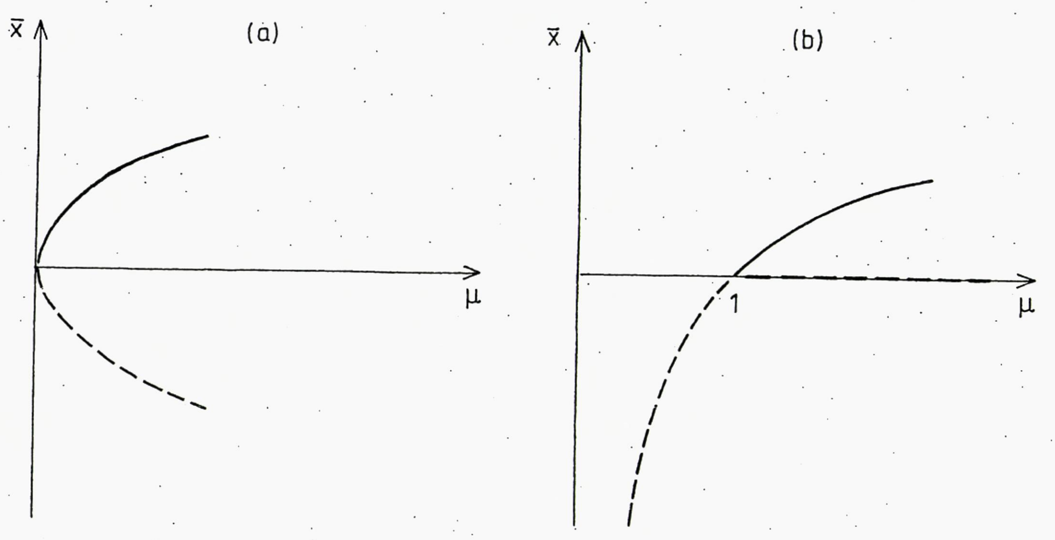


Figure 3: (a) The bifurcation diagram (saddle node) for the map $x \to x + \mu - x^2$. (b) The bifurcation diagram (transcritical) for the map $x \to \mu x(1-x)$.

As is clear from Definition 1.4.1 a fixed point will also lose its hyperbolicity if the eigenvalue λ equals 1. The general case then is that x^* will undergo a saddle-node bifurcation at the threshold where hyperbolicity fails. We shall now describe the saddle-node bifurcation.

Consider the map

$$x \to f_\mu(x) = x + \mu - x^2 \tag{1.5.1}$$

whose fixed points are $x^*_{1,2} = \pm\sqrt{\mu}$. Hence, when $\mu > 0$ there are two fixed points which equals when $\mu = 0$. If $\mu < 0$ there are no fixed points. In case of $\mu > 0$, μ small, we have $f'_\mu(x^*_1 = \sqrt{\mu}) = 1 - 2\sqrt{\mu} < 1$, hence $x^*_1 = \sqrt{\mu}$ is stable. On the other hand: $f'_\mu(x^*_2 = -\sqrt{\mu}) = 1 + 2\sqrt{\mu} > 1$, consequently x^*_2 is unstable. Thus a saddle-node bifurcation is characterized by that there is no fixed point when the parameter μ falls below a certain threshold μ_0. When μ is increase to μ_0, $\lambda = 1$, and two branches of fixed points are born, one stable and one unstable as displayed in the bifurcation diagram, see Figure 3a.

The other possibilities at $\lambda = 1$ are the pitchfork and the transcritical bifurcations. The various configurations for the pitchfork are given at the end of the proof of Theorem

1.5.1 (see Figure 2). A typical configuration in the transcritical case is shown in Figure 3b as a result of considering the quadratic map at $(x^*, \mu_0) = (0, 1)$.

Exercise 1.5.2. Do the necessary calculations which leads to Figure 3b. □

1.6 The flip bifurcation sequence

We shall now return to the flip bifurcation. First we consider the quadratic map. In the previous section we used Theorem 1.5.1 to prove that the quadratic map $x \to \mu x(1-x)$ undergoes a supercritical flip bifurcation at the threshold $\mu = \mu_0 = 3$. This means that in case of $\mu > \mu_0$, $|\mu - \mu_0|$ small, there exists a stable 2-periodic orbit and according to our findings in Section 1.3 the periodic points are given by (1.3.3), namely

$$p_{1,2} = \frac{\mu + 1 \pm \sqrt{(\mu+1)(\mu-3)}}{2\mu}$$

The period 2 orbit will remain stable as long as

$$|f'(p_1)f'(p_2)| < 1$$

cf. Section 1.4. Thus, in our example,

$$|\mu(1-2p_1)\mu(1-2p_2)| < 1$$

i.e.

$$|1 - (\mu+1)(\mu-3)| < 1 \tag{1.6.1}$$

from which we conclude that the 2-periodic orbit is stable as long as

$$3 < \mu < 1 + \sqrt{6} \tag{1.6.2}$$

Since $\lambda = f^{2'} = f'(p_1)f'(p_2) = -1$ when $\mu_1 = 1 + \sqrt{6}$ there is a new flip bifurcation taking place at μ_1 which in turn leads to a 4-periodic orbit. We also notice that while the fixed

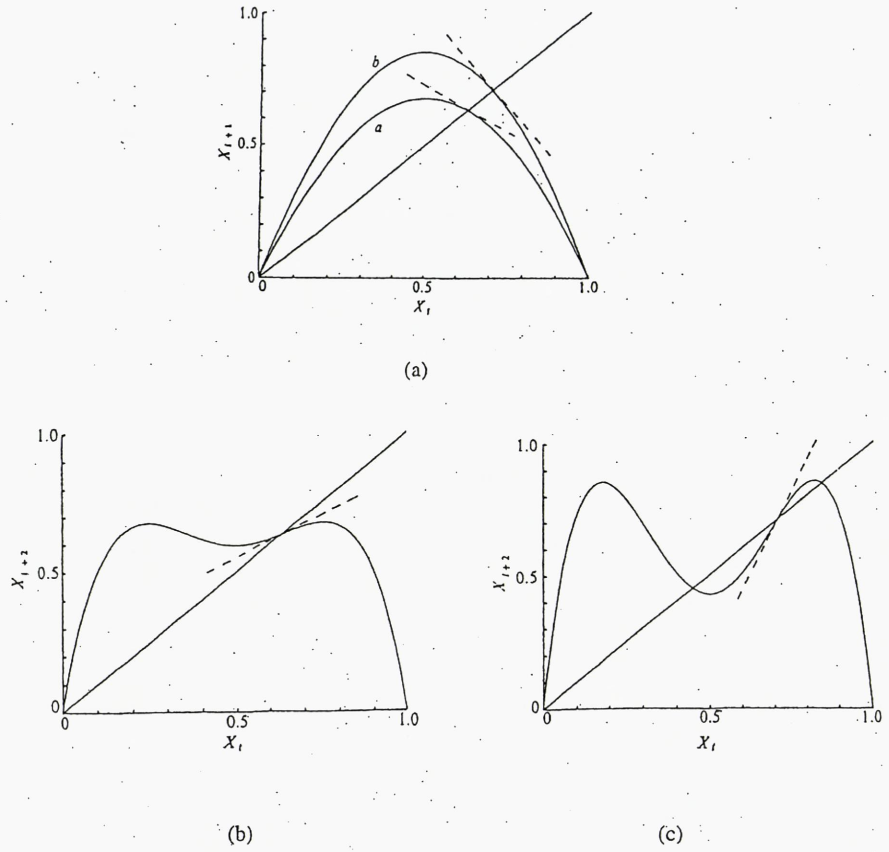


Figure 4: (a) The quadratic map in the cases $\mu = 2.7$ and $\mu = 3.4$. (b) and (c) The second iterate of the quadratic map in the cases $\mu = 2.7$ and $\mu = 3.4$, respectively.

point $x^* = (\mu - 1)/\mu$ is stable in the open interval $I = (2, 3)$, the length of the interval where the 2-periodic orbit is stable is roughly $(1/2)I$.

In Figure 4a we show the graphs of the quadratic map in the cases $\mu = 2.7$ (curve a) and $\mu = 3.4$ (curve b) respectively, together with the straight line $x_{t+1} = x_t$. $\mu = 2.7$ gives a stable fixed point x^* while $\mu = 3.4$ gives an unstable fixed point. These facts are emphasized in the figure by drawing the slopes (indicated by dashed lines). The steepness of the slope at the fixed point of curve a is less than $-45°$, $|\lambda| < 1$, while $\lambda < -1$ at the unstable fixed point located on curve b.

In general, if $f_\mu(x)$ is a single hump function (just as the quadratic map displayed in Figure 4a) the second iterate $f_\mu^2(x)$ will be a two-hump function. In Figures 4b and 4c we show the relation between x_{t+2} and x_t. Figure 4b corresponds to $\mu = 2.7$, Figure 4c corresponds to $\mu = 3.4$. Regarding 4b the steepness of the slope is still less than 45° so the fixed point is stable. However, in 4c the slope at the fixed point is steeper than 45°, the fixed point is unstable and we see two new solutions of period 2.

Let us now explore this mechanism analytically: Suppose that we have an n-periodic orbit consisting of the points $p_0, p_1 \dots p_{n-1}$ such that

$$p_i = f_\mu^n(p_i) \tag{1.6.3}$$

Then by the chain rule (cf. (1.1.7))

$$f_\mu^{n\prime}(p_0) = \prod_{i=0}^{n-1} f'_\mu(p_i) = \lambda^n(p_0) \tag{1.6.4}$$

Hence, if $|\lambda^n(p_0)| < 1$ the n-periodic orbit is stable, if $|\lambda^n(p_0)| > 1$ the orbit is unstable.

Next, consider the $2n$-periodic orbit

$$p_i = f_\mu^{2n}(p_i) = f_\mu^n(f_\mu^n(p_i))$$

By appealing once more to the chain rule we obtain

$$f_\mu^{2n\prime}(p_0) = \left(\prod_{i=0}^{n-1} f'_\mu(p_i)\right)^2 = \lambda^{2n}(p_0) \tag{1.6.5}$$

This allows us to conclude that if the n-point cycle is stable (i.e. $|\lambda^n| < 1$) then $\lambda^{2n} < 1$ too. On the other hand, when the n-cycle becomes unstable (i.e. $|\lambda^n| > 1$) then $\lambda^{2n} > 1$ too. So what this argument shows is that when a periodic point of prime period n becomes unstable it bifurcates into two *new* points which are initially stable points of period $2n$ and obviously there are $2n$ such points. This is the situation displayed in Figure 4c. So what the argument presented above really says is that as the parameter μ of the map $x \to f_\mu(x)$ is increased periodic orbits of period $2, 2^2, 2^3, \dots$ and so on are created through successive flip bifurcations. This is often referred to as the flip bifurcation sequence. Initially, all the 2^k cycles are stable but they become unstable as μ is further increased.

—

As already mentioned, if $f_\mu(x)$ is a single-hump function, then $f_\mu^2(x)$ is a two- hump function. In the same way, $f_\mu^3(x)$ is a four-hump function and in general f_μ^p will have 2^{p-1} humps. This means that the parameter range where the period 2^p cycles are stable shrinks through further increase of μ. Indeed, the μ values at successive bifurcation points act more or less as terms in a geometric series. In fact, Feigenbaum (1978) demonstrated the existence of a universal constant δ (known as the Feigenbaum number or the Feigenbaum geometric ratio) such that

$$\lim_{n\to\infty} \frac{\mu_{n+1} - \mu_n}{\mu_{n+2} - \mu_{n+1}} = \delta = 4.66920 \tag{1.6.6}$$

where μ_n, μ_{n+1} and μ_{n+2} are the parameter values at three consecutive flip bifurcations. From this we may conclude that there must exist an accumulation value μ_a where the series of flip bifurcations converge. (Geometrically, this may happen as a "valley" of some iterate of f_μ deepens and eventually touches the 45° line (cf. Figure 4c), then a saddle-node bifurcation ($\lambda = 1$) will occur.)

Regarding our running example $x \to \mu x(1-x)$ we have proved that the first flip bifurcation occurs at $\mu = 3$ and the second at $\mu = 1 + \sqrt{6}$. The point of accumulation for the flip bifurcations μ_a is found to be $\mu_a = 3.56994$.

Exercise 1.6.1. Identify numerically the flip bifurcation sequence for the Ricker map (1.2.2). □

In the next sections we will describe the dynamics beyond the point of accumulation μ_a for the flip bifurcations.

1.7 Period 3 implies chaos. Sarkovskii's theorem

Referring to our running example (1.2.1), $x \to \mu x(1-x)$ we found in the previous section that the point of accumulation for the flip bifurcation sequence $\mu_a \approx 3.56994$. We urge the reader to use a computer or a calculator to identify numerically some of the findings presented below. $\mu \in [\mu_a, 4]$.

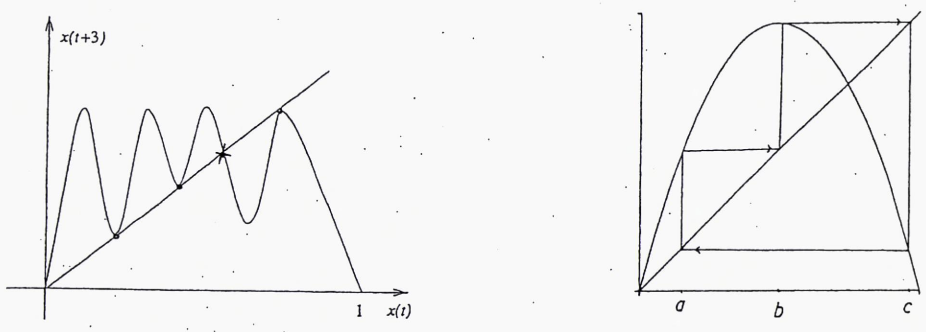


Figure 5: A 3-cycle generated by the quadratic map.

When $\mu > \mu_a$, $\mu - \mu_a$ small, there are periodic orbits of even period as well as aperiodic orbits. Regarding the periodic orbits, the periods may be very large, sometimes several thousands which make them indistinguishable from aperiodic orbits. Through further increase of μ odd period cycles are detected too. The first odd cycle is established at $\mu = 3.6786$. At first these cycles have long periods but eventually a cycle of period 3 appears. In case of (1.2.1) the period-3 cycle occurs for the first time at $\mu = 3.8284$. This is displayed in Figure 5. (The point marked with a cross is the initially fixed point $x^* = (\mu - 1)/\mu$ which became unstable at $\mu = 3$. It is also clear from the figure that the 3-cycle is established as the third iterate of (1.2.1) undergoes a saddle-node bifurcation.

In the bifurcation diagram, Figure 6, we display the dynamics of the quadratic map in the interval $2.9 \leq \mu \leq 4$. The stable fixed point ($\mu < 3$) as well as the flip bifurcation sequence is clearly identified. Also the period-3 "window" is clearly visible. Our goal in this and in the next sections is to give a thorough description of the dynamics beyond μ_a.

Theorem 1.7.1. Let $f_\mu : \mathbf{R} \to \mathbb{R}$, $x \to f_\mu(x)$ be continuous. Suppose that f_μ has a periodic point of period 3. Then f_μ has periodic points of all other periods. □

Remark: Theorem 1.7.1 was first proved in 1975 by Li and Yorke under the title "Period three implies chaos". Since there is no unique definition of the concept chaos many authors today prefer to use the concept "Li and Yorke

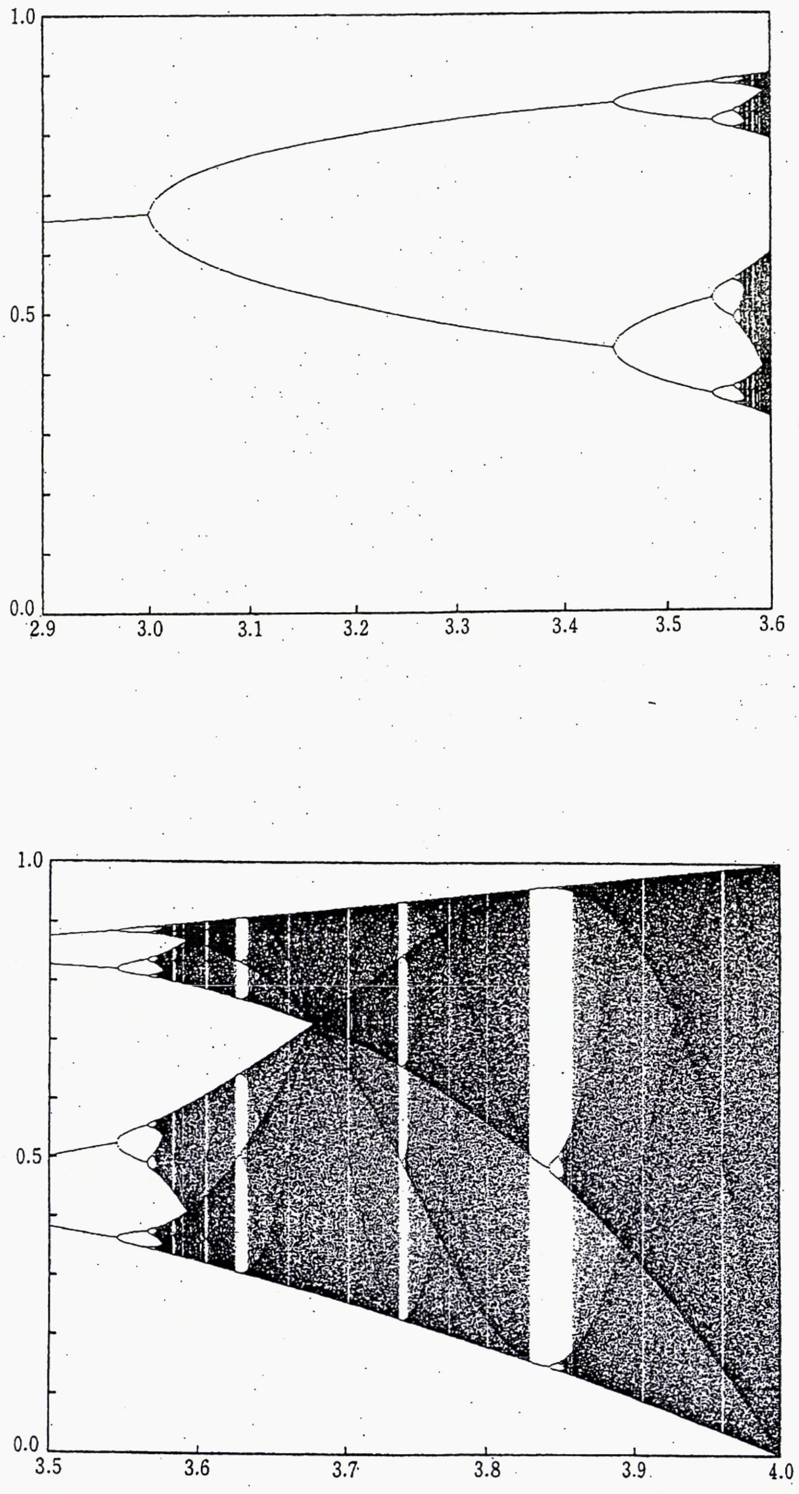


Figure 6: The bifurcation diagram of the quadratic map in the parameter range $2.9 \leq \mu \leq 4$.

chaos" when they refer to Theorem 1.7.1. The essence of Theorem 1.7.1 is that once a period-3 orbit is established it implies periodic orbits of all other periods. Note, however, that Theorem 1.7.1 does not address the question of stability. We shall deal with that in the next section. □

We will now prove Theorem 1.7.1. Our proof is based upon the proof in Devaney (1989), not so much upon the original proof by Li and Yorke (1975).

Proof. First, note that (1): If I and J are two compact intervals so that $I \subset J$ and $J \subset f_\mu(I)$ then f_μ has a fixed point in I. (2): Suppose that $A_0, A_1, \ldots, A_n$ are closed intervals and that $A_{i+1} \subset f_\mu(A_i)$ for $i = 0, \ldots, n-1$. Then there is at least one subinterval J_0 of A_0 which is mapped onto A_1. There is also a similar subinterval in A_1 which is mapped onto A_2 so consequently there is a $J_1 \subset J_0$ so that $f(J_1) \subset A_1$ and $f_\mu^2(J_1) \subset A_2$. Continuing is this fashion we find a nested sequence of intervals which map into the various A_i in order. Therefore there exists a point $x \in A_0$ such that $f_\mu^i(x) \in A_i$ for each i. We say that $f_\mu(A_i)$ covers A_{i+1}.

Now, let a, b and $c \in \mathbb{R}$ and suppose $f_\mu(a) = b$, $f_\mu(b) = c$ and $f_\mu(c) = a$. We further assume that $a < b < c$. Let $I_0 = [a, b]$ and $I_1 = [b, c]$, cf. Figure 5. Then from our assumptions $I_1 \subset f(I_0)$ and $I_0 \vee I_1 \subset f(I_1)$. The graph of f_μ, cf. Figure 5, shows that there must be a fixed point of f_μ between b and c. Similarly, f_μ^2 must have fixed points between a and b and at least one of them must have period 2. Therefore we let $n \geq 2$. Our goal is to produce a periodic point of prime period $n > 3$. Inductively, we define a nested sequence of intervals $A_0, A_1, \ldots, A_{n-2} \subset I_1$ as follows. Let $A_0 = I_1$. Since $I_1 \subset f(I_1)$ there is a subinterval $A_1 \subset A_0$ such that $f_\mu(A_1) = A_0 = I_1$. Then there is also a subinterval $A_2 \subset A_1$ such that $f_\mu(A_2) = A_1$ which implies $f_\mu^2(A_2) = f_\mu(f_\mu(A_2)) = f_\mu(A_1) = A_0 = I_1$. Continuing in this way there exists $A_{n-2} \subset A_{n-3}$ such that $f_\mu(A_{n-2}) = f_\mu(A_{n-3})$ so according to (2), if $x \in A_{n-2}$ then $f_\mu(x), f_\mu^2(x), \ldots, f_\mu^{n-1}(x) \in A_0$ and indeed $f_\mu^{n-2}(A_{n-2}) = A_0 = I_1$.

Now, since $I_0 \subset f_\mu(I_1)$ there exists a subinterval $A_{n-1} \subset A_{n-2}$ such that $f_\mu^{n-1}(A_{n-1}) = I_0$. Finally, since $I_1 \subset f_\mu(I_0)$ we have $I_1 \subset f_\mu^n(A_{n-1})$ so that $f_\mu^n(A_{n-1})$ covers A_{n-1}. Therefore, according to (1) f_μ^n has a fixed point p in A_{n-1}.

Finally, we claim that p has prime period n. Indeed, the first $n-2$ interations of p is in I_1, the $(n-1)$st lies in I_0 and the n-th is p again. If $f_\mu^{n-1}(p)$ lies in the interior of I_0 it follows that p has prime period n. If $f_\mu^{n-1}(p)$ lies on the boundary, then $n = 2$ or 3 and again we are done. □

Theorem 1.7.1 is a special case of Sarkovskii's theorem which came in 1964. However, it was written in Russian and published in an Ukrainian mathematical journal so it was not discovered and recognized in Western Europe and the U.S. prior to the work of Li and Yorke. We now state Sarkovskii's theorem:

Theorem 1.7.2. We order the positive integers as follows:

$$1 \lhd 2 \lhd 2^2 \lhd \ldots \lhd 2^m \lhd 2^k(2n+1) \lhd \ldots \lhd 2^k \cdot 3 \lhd \ldots 2 \cdot 3 \lhd 2n-1 \lhd \ldots \lhd 9 \lhd 7 \lhd 5 \lhd 3$$

Let $f_\mu : I \to I$ be a continuous map of the compact interval I into itself. If f_μ has a periodic point of prime period p, then it also has periodic points for any prime period $q \lhd p$. □

Proof. Cf. Devaney (1989) or Katok and Hasselblatt (1995). □

Clearly, Theorem 1.7.1 is a special case of Theorem 1.7.2. Also note that the first part in the Sarkovskii ordering $(1 \lhd 2 \lhd 2^2 \ldots \lhd 2^m)$ corresponds to the flip bifurcation sequence as demonstrated through our treatment of the quadratic map. As the parameter u in (1.2.1) is increased beyond the point of accumulation for the flip bifurcations. Sarkovskii's theorem says that we approach a situation where there are an infinite number of periodic orbits.

1.8 The Schwarzian derivative

In the previous section we established through Theorems 1.7.1 and 1.7.2 that a map may have an infinite number of periodic orbits. Our goal in this section is to prove that in fact only a few of them are attracting (or stable) periodic orbits.

Definition 1.8.1. Let $f : I \to I$ be a C^3 function. The Schwarzian derivative Sf of f is defined as

$$Sf(x) = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2 \tag{1.8.1}$$

□

Regarding $f_\mu(x) = \mu x(1-x)$ we easily find that $Sf_\mu(x) = -6/(1-2x)^2$. Note that $Sf_\mu < 0$ everywhere except at the critical point $c = 1/2$. (However, we may define $Sf_\mu(1/2) = -\infty$.)

The main result in this section is the following theorem which is due to Singer (1978):

Theorem 1.8.1. Lef f be a C^3 function with negative Schwarzian derivative. Suppose that f has one critical point c. Then f has at most three attracting periodic orbits. □

Proof. The proof consists of three steps.

(1) First we prove that if f has negative Schwarzian derivative then all f^n iterates also have negative Schwarzian derivatives.

To this end, assume $Sf < 0$ and $Sg < 0$. Our goal is to show that $S(f \circ g) < 0$. Successive use of the chain rule gives:

$$(f \circ g)'(x) = f'(g(x))g'(x)$$
$$(f \circ g)''(x) = f''(g(x))(g'(x))^2 + f'(g(x))g''(x)$$
$$(f \circ g)'''(x) = f'''(g(x))(g'(x))^3 + 3f''(g(x))g'(x)g''(x) + f'(g(x))g'''(x)$$

Then (omitting function arguments) Definition 1.8.1 gives

$$S(f \circ g) = \frac{f'''g'^3 + 3f''g'g'' + fg'''}{f'g'} - \frac{3}{2}\left(\frac{f''g'^2 + f'g''}{f'g'}\right)^2$$

which after some rearrangements may be written as

$$\left(\frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2\right) g'^2 + \frac{g'''}{g'} - \frac{3}{2}\left(\frac{g''}{g'}\right)^2 = Sf(g(x))(g'(x))^2 + Sg(x)$$

Thus $S(f \circ g)(x) < 0$ which again implies $Sf^n < 0$.

(2) Next we show that if $Sf < 0$ then $f'(x)$ cannot have a positive local minimum.

To this end, assume that d is a critical point of $f'(x)$. Then $f''(d) = 0$, and since $Sf < 0$ it follows from Definition 1.8.1 that $f'''/f' < 0$ so $f'''(d)$ and $f'(d)$ have opposite signs. Graphically, it is then obvious that $f'(x)$ cannot have a positive local minimum, and in the same way it is also clear that $f'(x)$ cannot have a negative local maximum. Consequently, between any two consecutive critical points d_1 and d_2 of f' there must be a critical point c of f such that $f'(c) = 0$.

(3) By considering $f^{n\prime}(x) = 0$ it follows directly from the chain rule that if $f(x)$ has a critical point then $f^n(x)$ will have a critical point too. Finally, let p be a point of period k on the attracting orbit and let $I = (a, b)$ be the largest open interval around p where all points approach p asymptotically. Then $f(I) \subset I$ and $f^k(I) \subset I$. Regarding the end points a and b we have: If $f(a) = f(b)$ then of course there exists a critical point. If $f(a) = a$ and $f(b) = b$ (i.e. that the end points are fixed points) it is easy to see graphically that there exist points u and v such that $a < u < p < v < b$ with properties $f'(u) = f'(v) = 1$. Then from (2) and the fact that $f'(p) < 1$ there must be a critical point in (u, v). In the last case $f(a) = b$ and $f(b) = a$ we arrive at the same conclusion by considering the second iterate f^2. □

Example 1.8.1. Assume $x \in [0,1]$ and let us apply Theorem 1.8.1 on the quadratic map $x \to f_\mu(x) = \mu x(1-x)$. For a fixed $\mu \in (1,3)$ the fixed point $x^* = (\mu - 1)/\mu$ is stable, and since $f_\mu(0) = f_\mu(1) = 0$ and the fact that 0 is repelling there is one periodic attractor, namely the period-1 attractor x^* which attracts the critical points $c = 1/2$.

When $\mu \in [3,4]$ both x^* and 0 are unstable fixed points. Thus according to Theorem 1.8.1 there is at most one attracting periodic orbit in this case. (Prior to μ_a there is exactly one periodic attractor.) When $\mu = 4$ the critical point is mapped on the origin through two iterations so there are no attracting periodic orbits in the case. □

Example 1.8.2. Let us close this section by giving an example which shows that Theorem 1.8.1 fails if the Schwarzian derivative is not negative. The following example is due to Singer (1978). Consider the map

$$x \to g(x) = -13.30x^4 + 28.75x^3 - 23.31x^2 + 7.86x \tag{1.8.2}$$

The map has one fixed point $x^* = 0.7263986$, and by considering $g^2(x) = x$ there is also one 2-periodic orbit which consists of the points $p_1 = 0.3217591$ and $p_2 = 0.9309168$.

Moreover: $\lambda_1 = g'(x^*) = -0.8854$ and $\sigma = g'(p_1)g'(p_2) = -0.06236$. Thus both the fixed point and the 2-periodic orbit are attracting.

The critical point of g is $c = 0.3239799$ and is attracted to the period-2 orbit so it does not belong to $W^s_{\text{loc}}(x^*)$, cf. Definition 1.4.3. The reason that x^* is not attracting c is that $Sg(x^*) = 8.56 > 0$ thus the assumption $Sg(x) < 0$ in Theorem 1.8.1 is violated. □

Exercise 1.8.1. Compute the Schwarzian derivative when $f(x) = x^n$. □

Exercise 1.8.2. Show that $Sf(x) < 0$ when f is given by (1.2.2) (the Ricker case). □

1.9 Symbolic dynamics I

Up to this point we have mainly been concerned with fixed points and periodic orbits. The main goal of this section is to introduce a useful tool called symbolic dynamics which will help us to describe and understand dynamics of other types than we have discussed previously. To be more concrete, we shall in this section analyse the quadratic map $x \to \mu x(1-x)$ where $\mu > 2+\sqrt{5}$ on the interval $I = [0,1]$, and as it will become clear, although almost all points in I eventually will escape I, there exists an invariant set Λ of points which will remain in I. We shall use symbolic dynamics to describe the behaviour of these points.

First we need some definitions. Consider $x \to f(x)$. Suppose that $f(x)$ can take its values on two disconnected intervals I_1 and I_2 only. Define an infinite forward-going sequence of 0's and 1's $\{a_k\}_{k=0}^{\infty}$ so that

$$a_k = 0 \quad \text{if} \quad f^k(x_0) \in I_1 \tag{1.9.1a}$$

$$a_k = 1 \quad \text{if} \quad f^k(x_0) \in I_2 \tag{1.9.1b}$$

Thus what we really do here is to represent an orbit of a map by an infinite sequence of 0's and 1's.

Definition 1.9.1.

$$\Sigma_2 = \{\overline{a} = (a_0 a_1 a_2 \ldots)/a_k = 0 \ \text{ or } \ 1\} \tag{1.9.2}$$

□

We shall refer to Σ_2 as the sequence space.

Definition 1.9.3. The itinerary of x is a sequence $\phi(x) = a_0 a_1 \ldots$ where a_k is given by (1.9.1). □

We now define one of the cornerstones of the theory of symbolic dynamics.

Definition 1.9.4. The shift map $\sigma : \Sigma_2 \to \Sigma_2$ is given by

$$\sigma(a_0a_1a_2a_3\ldots) = a_1a_2a_3\ldots \tag{1.9.3}$$

$\square$

Hence the shift map deletes the first entry in a sequence and moves all the other entries one place to the left.

Example 1.9.1. $\overline{a} = (1111\ldots)$ represents a fixed point under σ since $\sigma(\overline{a}) = \sigma^n(\overline{a}) = (111\ldots)$. Suppose $\overline{a} = (001, 001, 001, \ldots)$. Then $\sigma(\overline{a}) = (010, 010, 010, \ldots)$, $\sigma^2(\overline{a}) = (100, 100, 100, \ldots)$ and $\sigma^3(\overline{a}) = (001, 001, 001, \ldots) = \overline{a}$. Thus $\overline{a} = (001, 001, 001, \ldots)$ represents a periodic point of period 3 under the shift map.

$\square$

The previous example may obviously be generalized. Indeed, if $\overline{a} = (a_0a_1\ldots a_{n-1}, a_0a_1\ldots a_{n-1}, \ldots)$ there are 2^n periodic points of period n under the shift map since each entry in the sequence may have two entries 0 or 1.

Definition 1.9.5. Let U be a subset of a set S. U is dense in S if the closure $\overline{U} = S$. $\square$

Definition 1.9.6. If a set S is closed, contains no intervals and no isolated points it is called a Cantor set. $\square$

Proposition 1.9.1. The number of periodic points $P_{\text{er}}(\sigma) = 2^n$ is dense in Σ_2. $\square$

Proof. Let $\overline{a} = (a_0a_1a_2\ldots)$ be in Σ_2 and suppose that $\overline{b} = (a_0\ldots a_{n-1}, a_0\ldots a_{n-1}\ldots)$ represent the 2^n periodic points. Our goal is to prove that $\overline{b}$ converges to $\overline{a}$. By use of the usual distance function in a sequence space, $d[\overline{a}, \overline{b}] = \Sigma(|a_i - b_i|/2^i)$ we easily find that $d[\overline{a}, \overline{b}] < 1/2^n$. Hence $\overline{b} \to \overline{a}$. $\square$

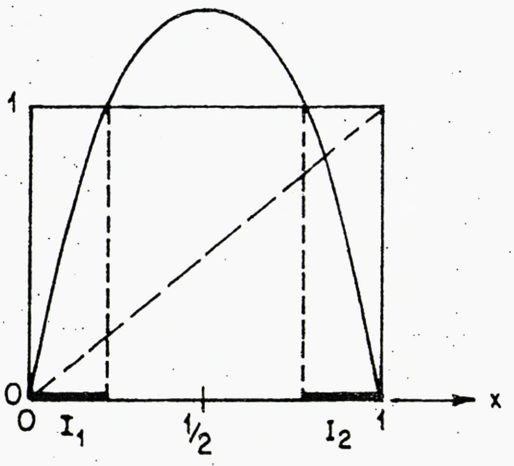


Figure 7: The quadratic map in the case $\mu > 2+\sqrt{5}$. Note the subintervals I_1 and I_2 where $f_\mu(x) = \mu x(1-x) \leq 1$.

We now have the necessary machinery we need in order to analyse the quadratic map in case of $\mu > 2+\sqrt{5}$.

Let $x \to f(x) = \mu x(1-x)$ where $\mu > 2+\sqrt{5}$. From the equation $\mu x(1-x) = 1$ we find $x = 1/2 + 1/2\sqrt{1-4\mu}$. Hence in the intervals $I_1 = [0, 1/2 - 1/2\sqrt{1-4\mu}]$ and $I_2 = [1/2+1/2\sqrt{1-4\mu}]$, $f \leq 1$, cf. Figure 7. Moreover, $|f'(x)| = |\mu - 2\mu x|$ and whenever $\mu > 2+\sqrt{5}$ we find that $|f'(x)| > \lambda > 1$.

Denote $I = [0,1]$. Then $I \cap f^{-1}(I) = I_1 \cup I_2$ so if $x \in I - (I \cap f^{-1}(I))$ we have $f > 1$ (cf. Figure 7) which implies $f^2 < 0$ and consequently $f^n \to -\infty$. All the other points will remain in I after one iteration. The second observation is that $f(I_1) = f(I_2) = I$ so there must be a pair of open intervals, one in I_1 and one in I_2, which is mapped into $I - (I_1 \cup I_2)$ such that all points in these two intervals will leave I after two iterations. Continuing in this way by removing pairs of open intervals (i.e. first the interval $I - (I_1 \cup I_2)$, then two intervals, one in I_1 (J_1) and one in I_2 (J_2), then 2^2 open intervals, two from $I_1 - J_1$, two from $I_2 - J_2 \ldots$ and finally 2^n intervals) from closed intervals we are left with a closed set Λ which is I minus the union of all the $2^{n+1}-1$ open sets. Hence Λ consists of the points that remain in I after n iterations, $\Lambda \subset I \cap f^{-1}(I)$ and Λ consists of 2^{n+1} closed intervals.

Now, associate to each $x \in \Lambda$ a symbol sequence $\{a_i\}_{i=1}^{\infty}$ of 0's and 1's such that $a_k = 0$ if $f^k(x) \in I_1$ and $a_k = 1$ if $f^k(x) \in I_2$.

Next, define

$$I_{a_0 \ldots a_n} = \{x \in I / x \in I_{a_0}, f(x) \in I_{a_1} \ldots f^n(x) \in I_{a_n}\} \tag{1.9.4}$$

as one of the 2^{n+1} closed subintervals in Λ. Our first goal is to show that $I_{a_0 \ldots a_n}$ is non-empty when $n \to \infty$. Indeed,

$$\begin{aligned} I_{a_0 \ldots a_n} &= I_{a_0} \cap f^{-1}(I_{a_1}) \cap \ldots \cap f^{-n}(I_{a_n}) \\ &= I_{a_0} \cap f^{-1}(I_{a_1 \ldots a_n}) \end{aligned} \tag{1.9.5}$$

I_{a_1} is nonempty. Then by induction $I_{a_1 \ldots a_n}$ is non-empty, and moreover, since $f^{-1}(I_{a_1 \ldots a_n})$ consists of two closed subintervals it follows that $I_{a_0} \cap f^{-1}(I_{a_1 \ldots a_n})$ consists of one closed interval. A final observation is that

$$\begin{aligned} I_{a_0 \ldots a_n} &= I_{a_0} \cap \ldots \cap f^{-(n-1)}(I_{a_{n-1}}) \cap f^{-n}(I_{a_n}) \\ &= I_{a_0 \ldots a_{n-1}} \cap f^{-n}(I_{a_n}) \subset I_{a_0 \ldots a_{n-1}} \end{aligned}$$

Consequently, $I_{a_0 \ldots a_n}$ is non-empty. Clearly the length of all sets $I_{a_0 \ldots a_n}$ approaches zero as $n \to \infty$ which allows us to conclude that the itinerary $\phi(x) = a_0 a_1 \ldots$ is unique.

We now proceed by showing that Λ is a Cantor set. Assume that Λ contains an interval $[a, b]$ where $a \neq b$. For $x \in [a, b]$ we have $|f'(x)| > \lambda > 1$ and by the chain rule $|f^{n\prime} x)| > \lambda^n$. Let n be so large that $\lambda^n |b - a| > 1$. Then from the mean value theorem $|f^n(b) - f^n(a)| \geq \lambda^n |b - a| > 1$ which means that $f^n(b)$ or $f^n(a)$ (or both) are located outside I. This is of course a contradiction so Λ contains no intervals.

To see that Λ contains no isolated points it suffices to note that any end point of the $2^{n+1} - 1$ open intervals eventually goes to 0 and since $0 \in \Lambda$ these end points are in Λ too. Now, if $y \in \Lambda$ is isolated all points in a neighbourhood of y eventually will leave I which means that they must be elements of one of the $2^{n+1} - 1$ open sets which are removed from I. Therefore, the only possibility such that $y \in \Lambda$ is that there is a sequence of end points converging towards y so y cannot be isolated.

From the discussion above we conclude that the quadratic map where $\mu > 2 + \sqrt{5}$ possesses an invariant set Λ, a Cantor set, of points that never leave I under iteration.

Λ is a repelling set. Our final goal is to show that the shift map σ defined on Σ_2 is topological equivalent to f defined on Λ.

Let $f : \Lambda \to \Lambda$, $f(x) = \mu x(1-x)$, $\sigma : \Sigma_2 \to \Sigma_2$, $\sigma(a_0a_1a_2\ldots) = a_1a_2\ldots$ and $\phi : \Lambda \to \Sigma_2$, $\phi(x) = a_0a_1a_2\ldots$. We want to prove that $\phi \circ f = \sigma \circ \phi$.

Observe that

$$\phi(x) = a_0a_1a_2\ldots = \bigcap_{n\geq 0} I_{a_0a_1a_2\ldots a_n\ldots}$$

Further

$$I_{a_0a_1\ldots a_n} = I_{a_0} \cap f^{-1}(I_{a_1}) \cap \ldots \cap f^{-n}(I_{a_n})$$

so

$$f(I_{a_0a_1\ldots a_n}) = f(I_{a_0}) \cap (I_{a_1}) \cap \ldots \cap f^{-n+1}(I_{a_n}) = I_{a_1} \cap \ldots \cap f^{-n+1}(I_{a_n}) = I_{a_1\ldots a_n}$$

This implies that

$$\phi(f(x)) = \phi\left(f\left(\bigcap_{n\geq 0} I_{a_0\ldots a_n}\right)\right) = \phi\left(\bigcap_{n\geq 1} I_{a_1\ldots a_n}\right) = \sigma(\phi(x))$$

Thus, f and σ are topological equivalent maps.

1.10 Symbolic dynamics II

In Section 1.8 we proved that if a map $f : I \to I$ with negative Schwarzian derivative possessed an attracting periodic orbit then there was a trajectory from the critical point c to the periodic orbit. Our goal here is to extend the theory of symbolic dynamics by assigning a symbol sequence to c or more precisely to $f(c)$. We will assume that f is unimodal. The theory will mainly be applied on periodic orbits.

Note, however, that the purpose of this section is somewhat different than the others so readers who are not too interested in symbolic dynamics may skip this section and proceed directly to the next where chaos is treated.

Definition 1.10.1. Let $x \in I$. Define the itinerary of x as $\phi(x) = a_0 a_1 a_2 \ldots$ where

$$a_j = \begin{cases} 0 & \text{if} \quad f^j(x) < c \\ 1 & \text{if} \quad f^j(x) > c \\ C & \text{if} \quad f^j(x) = c \end{cases} \tag{1.10.1}$$

□

What is new here really is that we associate a symbol C to the critical point c. Also note that we may define two intervals $I_0 = [0, c\rangle$ and $I_1 = \langle c, 1]$ such that f is increasing on I_0 and decreasing on I_1.

Definition 1.10.2. The kneading sequence is defined as the itinerary of $f(c)$, i.e.

$$K(f) = \phi(f(c)) \tag{1.10.2}$$

□

Example 1.10.1.

1) Suppose that $x \to f(x) = 2x(1-x)$. Then $c = 1/2$ and $f(c) = 1/2$, $f^2(c) = 1/2 \ldots f^j(c) = 1/2$ so the kneading sequence becomes $K(f) = (C\,C\,C\,C\,\ldots)$ which also may be written as $(C\,C\,\overline{C}\ldots)$ where the bar refers to repetition.

2) Suppose that $x \to f(x) = 4x(1-x)$. $c = 1/2$, $f(c) = 1$, $f^2(c) \ldots = f^j(c) = 0$ so $K(f) = (1\,0\,0\,\overline{0}\ldots)$. □

An unimodal map may of course have several itineraries.

Example 1.10.2. By use of a calculator we easily find that the possible itineraries of $x \to 2x(1-x)$ are

$$(0\,0\ldots0\,C\,C\,\overline{C}\ldots)\ (C\,C\,\overline{C}\ldots)\ (1\,0\ldots0\,C\,C\,\overline{C}\ldots)\ (0\,0\,\overline{0}\ldots)\ (1\,0\,0\,\overline{0}\ldots)$$

(The last two itineraries correspond to the orbits of $x_0 = 0$ and $x_1 = 1$ respectively. Note that the critical point is the same as the stable fixed point x^* in this example.

In case of $x \to 3x(1-x)$ we obtain the sequences

$$(0\,0\ldots 0\,1\,1\,\overline{1}\ldots)\ (C\,1\,1\,\overline{1}\ldots)(1\,1\,\overline{1}\ldots)$$

$$(1\,0\ldots 0\,1\,1\,\overline{1}\ldots)\ (0\,0\,\overline{0}\ldots)\ (1\,0\,0\,\overline{0}\ldots)$$

$$(0\,C\,1\,1\,\overline{1}\ldots)\ (1\,C\,1\,1\,\overline{1}\ldots)$$

where the last two itineraries correspond to the orbits of $x_0 = (1/6)(3-\sqrt{3})$ and $x_0 = (1/6)(3+\sqrt{3})$ respectively. □

The reader should also have in mind that periodic orbits with different periods may share the same itinerary.

Indeed, consider $x \to 3.1x(1-x)$. Then $x^* = 0.6774 > c = 1/2$ so the itinerary of the fixed point becomes $\phi(x^*) = (1\,1\,\overline{1}\ldots)$. However, there is also a two-periodic orbit whose periodic points are (cf. formulae (1.3.3)) $p_1 = 0.7645$, $p_2 = 0.5581$. Again we observe that $p_i > c$ so the itinerary of any of the two-periodic points is also $(1\,1\,\overline{1}\ldots)$. (When μ becomes larger than 3.1 one of the periodic points eventually will become smaller than c which results in the itinerary $(1\,0\,1\,0\,1\,0\ldots)$ or $(0\,1\,0\,1\,0\,1\ldots)$.)

Our next goal is to establish an ordering principle of the possible itineraries of a given map. Let $\overline{a} = (a_0a_1a_2\ldots)$ and $\overline{b} = (b_0b_1b_2\ldots)$. If $a_i = b_i$ for $0 \le i < n$ and $a_n \ne b_n$ we say that the sequences have discrepancy n. Let $S_n(\overline{a})$ be the number of 1's among $a_0a_1\ldots a_n$ and assume $0 < C < 1$.

Definition 1.10.3. Suppose that $\overline{a}$ and $\overline{b}$ have discrepancy n. We say that $\overline{a} \prec \overline{b}$ if

$$S_{n-1}(\overline{a}) \text{ is even and } a_n < b_n \tag{1.10.3a}$$

$$S_{n-1}(\overline{a}) \text{ is odd and } a_n > b_n \tag{1.10.3b}$$

□

Example 1.10.3. Due to a) we have the following order:

$$(1\,1\,0\ldots) \prec (1\,1\,C\ldots) \prec (1\,1\,1\ldots)$$

Due to b) we have

$$(1\,1\,0\ldots) \prec (1\,0\,1\ldots) \prec (1\,0\,0\ldots)$$

□

Also note that any two sequences with discrepancy 0 are ordered such that the sequence which has 0 as the first entry is of lower order than the one with C or 1 as the first entry. Thus:

$$(0\,1\ldots) \prec (C\,1\ldots) \prec (1\,1\ldots)$$

Exercise 1.10.1. Let $\overline{a} = (0\,1\,1\,0\,1\,1\ldots)$ be a repeating sequence. Compute $\sigma(\overline{a})$ and $\sigma^2(\overline{a})$ and verify the ordering $\overline{a} \prec \sigma(\overline{a}) \prec \sigma^2(\overline{a})$. □

The following theorem (due to Milner and Thurston) relates the ordering of two symbol sequences to the values of two points in an interval.

Theorem 1.10.1. Let $x, y \in I$

a) If $\phi(x) \prec \phi(y)$ then $x < y$

b) If $x < y$ then $\phi(x) \preceq \phi(y)$

□

Proof. Suppose that $\phi(x) = (a_0a_1a_2\ldots)$ and $\phi(y) = (b_0b_1b_2\ldots)$ and let n be the discrepancy of $\phi(x)$ and $\phi(y)$. First, suppose $n = 0$. Then $x < y$ since $0 < C < 1$. Next, suppose that a) is true with discrepancy $n - 1$. Our goal is to show that a) also is true with discrepancy n. By use of the shift we have $\phi(f(x)) = (a_1a_2a_3\ldots)$ and $\phi(f(y)) = (b_1b_2b_3\ldots)$. Suppose $a_0 = 0$. Then $\phi(f(x)) \prec \phi(f(y))$ since the number of 1's before the discrepancy is as before. Therefore $f(x) < f(y)$ but since f is increasing on $[0, c\rangle$ it follows that $x < y$. Next, assume $a_0 = 1$. Then $\phi(f(x)) \succ \phi(f(y))$ since the number of 1's among the a_i's ($i \geq 1$) has been reduced by one. Therefore $f(x) > f(y)$ which implies that $x < y$ since f decreases on $\langle c, 1]$. If $a_0 = C$ we have $x = y = c$.

Regarding b) suppose $x < y$ and assume that $\phi(x)$ and $\phi(y)$ has discrepancy n. First, note that if $x < c < y$ we have directly $\phi(x) < \phi(y)$. Otherwise (i.e. $x < y < c$ or $c < x < y$) note that f^i is monotone in $[x, y]$ for $i \leq n$. Since the number of 1's (cf. the chain rule) directly says if f^n is increasing or decreasing it is easily verified that $\phi(x) \leq \phi(y)$. □

Theorem 1.10.2. Let $x = \varphi(\overline{a}) = a_0a_1a_2\ldots$ and suppose that $x \to f(x)$ unimodal. Then $\phi(\sigma^n\varphi(\overline{a})) \preceq K(f(c))$ for $n \geq 1$. □

Proof. Since the maximum of f is $f(c)$ we have $f(x) < f(c)$ and $f^n(x) \leq f(c)$. Moreover, $\sigma x = \sigma(\varphi(\overline{a})) = a_1a_2\ldots = \varphi(f(x))$ so inductively $\sigma^n x = \varphi(f^n(x))$. Therefore, according to Theorem 1.10.1

$$\phi(\sigma^n\varphi(\overline{a})) \preceq \phi(f(c)) = K(f(c))$$

□

The essence of Theorem 1.10.2 is that any sequence $\overline{a}$ such that $\phi(x) = \overline{a}$ has lower order than the kneading sequence.

Now, consider periodic orbits. In order to simplify notation, repeating sequences (corresponding to periodic points) of the form $\overline{a} = (a_0a_1\ldots a_n\, a_0a_1\ldots a_n\, a_0a_1\ldots a_n\ldots) = (a_0a_1\ldots a_n\, \overline{a_1a_1\ldots a_n})$ will from now on be written as $\overline{a} = (a_0a_1\ldots a_n)$.

We also define a sequence $\hat{a} = (a_0\ldots a_{n-1}\hat{a}_n)$ where $\hat{a}_n = 1$ if $a_n = 0$ or $\overline{\hat{a}}_n = 0$ if $a_n = 1$. If $\overline{b} = (b_0b_1\ldots b_m)$, $\overline{a}\cdot\overline{b} = (a_0a_1\ldots a_n\, b_0b_1\ldots b_m)$.

Suppose that there exists a parameter value μ such that there are two periodic orbits γ_1 and γ_2 of the same prime period. We say that the orbit γ_1 is larger than the orbit γ_2 if γ_1 contains a point p_m which is larger than all the points of γ_2. Note that, according to Theorem 1.10.1, the itinerary of p_m satisfies $\phi(p_i) \preceq \phi(p_m)$ where p_i are any of the other periodic points contained in γ_1.

Our main interest is the ordering of itineraries of periodic points p which satisfy:

(A) The periodic point p shall be the largest point contained in the orbit.

(B) Every other periodic orbit of the same prime period must have a periodic point which is larger than p.

Before we continue the discussion of (A) and (B) let us state a useful lemma.

Lemma 1.10.1. Given two symbol sequences $\overline{a} = (a_0a_1a_2\ldots)$ and $\overline{b} = (b_0b_1b_2\ldots)$.

Suppose that $a_0 = b_0 = 1$ and $a_1 = b_1 = 0$. $a_j = b_j = 1$ for $2 \leq j \leq l$, $a_l = 0$, $b_l = 1$.

If l is even then $\overline{b} \prec \overline{a}$. If l is odd then $\overline{a} \prec \overline{b}$. □

Proof. Assume l even. Then the number of 1's before the discrepancy is odd and since $b_l > a_l$ Definition 1.10.3 gives that $\overline{b} \prec \overline{a}$.

If l is odd the number of 1's before the discrepancy is even and since $a_l = 0 < b_l = 1$, $\overline{a} \prec \overline{b}$ according to the definition. □

A consequence of this theorem is that sequences than begin with 1 0 are of larger order than sequences which begin with 1 1. In the same way, a sequence which first entries are 1 0 0 is larger than one which begins with 1 0 1.

Now, consider the quadratic map $x \to \mu x(1-x)$. Whenever $\mu > 2$ the fixed point $x^* = (\mu - 1)/\mu > c = 1/2$ so the (repeating) itinerary becomes $\phi(x^*) = (1)$. When x^* bifurcates at the threshold $\mu = 3$, the largest point p_1 contained in the 2-cycle is always larger than c, hence the itinerary of p_1 starts with 1 in the first entry. Therefore, when $\mu > 3$, there may be two possible itineraries (1 0) and (1 1) and clearly (1 1) $\prec$ (1 0). We are interested in (1 0). Considering the 4-cycle which is created through another flip bifurcation the itinerary of the largest point contained in the cycle which we seek is (1 0 1 1) which is of larger order than the other alternatives.

Regarding odd periodic orbits, remember that they are established through saddle-node bifurcations, thus two periodic orbits, one stable and one unstable, are established at the bifurcation. Considering the stable 3-cycle at $\mu = 3.839$ (see Exercise 1.4.2 or the bifurcation diagram, Figure 6) two of the points in the cycle 0.14989 and 0.48917 are smaller than c while the third one 0.95943 is larger. Hence the itinerary of largest order of 0.95493 is $(1\,0\,0)$. Referring to Exercise 1.4.2 the largest point contained in the unstable 3-cycle is 0.95384 and the other points are 0.16904 and 0.53392. Hence the itinerary of 0.95384 of largest order is $(1\,0\,1)$ and according to (A) and (B) this is the itinerary we are looking for, not the itinerary $(1\,0\,0)$.

Therefore, the itineraries we seek are the ones that satisfy (A) and (B) and correspond to periodic points which are established through flip or saddle-node bifurcations as the parameter in the actual family is increased. (A final observation is that sequences which contain the symbol C are out of interest since they violate (B).)

Now, cf. our previous discussion, define the repeating sequences:

$$S_0 = (1) \qquad S_1 = (1\,0) \qquad S_2 = (1\,0\,1\,1) \qquad S_3 = (1\,0\,1\,1\,1\,0\,1\,0)$$

and

$$S_{j+1} = S_j \cdot \hat{S}_j \tag{1.10.4}$$

Clearly, the sequence S_j has prime period 2^j so it represents a periodic point with the same prime period.

Another important property is that S_j has an odd number of 1's. To see this, note that $S_0 = (1)$ has an odd number of 1's. Next, assume that $S_k = (S_0 \dots S_{k-1}\,1)$ has an odd number of 1's. Then $\hat{S}_k = (S_0 \dots S_{k-1}0)$ has an even number of 1's so the concatenation $\hat{S}_{k+1} = S_k \cdot \hat{S}_k$ clearly has an odd number of 1's. (If S_k has a 0 at entry S_k we arrive at the same conclusion.) We have also that

$$\hat{S}_{j+1} = S_j \cdot S_j = S_j \tag{1.10.5}$$

Indeed, suppose $S_k = (S_0 \dots S_k)$. Then $S_{k+1} = S_k \cdot \hat{S}_k = (S_0 \dots S_k\; S_0 \dots \hat{S}_k)$ so $\hat{S}_{k+1} = (S_0 \dots S_k\; S_0 \dots S_k) = S_k \cdot S_k = S_k$.

Lemma 1.10.2. The sequences defined through (1.10.4) have the ordering

$$S_0 \prec S_1 \prec S_2 \prec S_3 \prec \dots$$

□

Proof. Assume that $S_j = (S_0 \dots S_{j-1}S_j)$. If $S_j = 1$ there must be an even number of 1's among $(S_0 \dots S_{j-1})$ so according to Definition 1.10.3a $\hat{S}_j \prec S_j$. If $S_j = 0$ there is an odd number of 1's among $(S_0 \dots S_{j-1})$ so according to Definition 1.10.3b $\hat{S}_j \prec S_j$ also here. Therefore, by use of (1.10.5), we have $S_j \succ \hat{S}_j = S_{j-1} \cdot S_{j-1} = S_{j-1}$. □

Let us now turn to periodic orbits of odd period. The following lemma is due to Guckenheimer.

Lemma 1.10.3. The largest point p_m in the smallest periodic orbit of odd period n has itinerary $\phi(p_m) = \overline{a}$ such that $a_i = 0$ if $i = 1(\text{mod}\, n)$ and $a_i = 1$ otherwise. □

Example 1.10.4. If $n = 3$, $\phi(p_m) = (1\,0\,1\;1\,0\,1\;1\,0\,1 \dots) = (1\,0\,1)$ which is in accordance with our previous discussion of 3-cycles. □

Proof. Suppose that we have a sequence $\overline{a}$ and that there exists a number k such that $a_k = 1$ and $a_{k+1} = a_{k+2} = 0$. Then by applying the shift map k times we arrive at $\sigma^k(\overline{a}) = (1\,0\,0 \dots)$ which according to Lemma 1.10.1 has larger order than any sequence with isolated 0's. Hence the sequence $\sigma^k(\overline{a})$ violates (A) and (B).

Therefore, the argument above shows that the sequence we are looking for in this lemma must satisfy that if $a_k = 0$ then both a_{k-1} and a_{k+1} must equal 1. Consequently there are blocks in $\overline{a}$ of even length where the first and last entry of the blocks consist of 0 and the intermediate elements of 1's. As a consequence of Lemma 1.10.1 the longer these blocks are the smaller is the order of the sequence. Note that the blocks in this lemma have maximum length $n + 1$ for a periodic sequence of period n. □

Example 1.10.5. $(1\,\underbrace{0\,1\,1\,0}\,1\,1\,0\,1)$ is a 3-cycle where the length of the block is 4.

$(1\ \underbrace{0\,1\,1\,1\,1\,0}\ 1\,1\,1)$ is a 5-cycle where the length of the block is 6. Clearly, the order of the 5-cycle is smaller than the order of the 3-cycle. □

Lemma 1.10.4. Let $n > 1$ be an odd number. Then there is a periodic orbit of period $n + 2$ which is smaller than all periodic orbits of period n. □

Proof. The lemma is an immediate consequence of how the itinerary in Lemma 1.10.3 is defined combined with the results of Lemma 1.10.1. □

We now turn to orbits of even period where the period is $2^n \cdot m$ where $m > 1$ is an odd number. The fundamental observation regarding the associated symbol sequences is that they may be written as $S_{j+1}S_j \ldots S_j$ or $S_j\hat{S}_jS_j \ldots S_j$ where the number of S_j blocks following S_{j+1} (or $\hat{S}_j$) is $m - 2$. (See Guckenheimer (1977) for further details.)

Example 1.10.6. If $n = 2$ (cf. 1.10.4) and $m = 3$ we have the sequence $(1\,0\,1\,1\,1\,0\,1\,0\,1\,0\,1\,1)$ and if $n = 1$ and $m = 5$ we arrive at $(1\,0\,1\,1\,1\,0\,1\,0\,1\,0)$. □

Lemma 1.10.5. Let P be a periodic orbit of odd period k. Then there exists a periodic orbit of even period $l = 2^n \cdot m$ where $m > 1$ is odd which is smaller than any odd period orbit. □

Proof. From Lemma 1.10.4 we have that the longer the odd period is the smaller is the ordering of the associated symbol sequence. From Lemma 1.10.3 it follows that such a symbol sequence may be written as $(1\,0\,1\,1\,1\ \ldots 1\,1\,1\,0\,1\,1\,1 \ldots)$. Therefore by comparing an even period sequence with the odd one above it is clear that the even period sequence has 0 as entry at the discrepancy. If the even period is 2 it is two 1's before the discrepancy. If the even period is larger there are three consecutive 1's just prior to the 0 and since the first entry of the sequence is 1 there is an even number of 1's before the discrepancy also here and the result of the lemma follows. □

We need one more lemma which deals with periodic orbits of even period.

Lemma 1.10.6. Let $u = 2^n \cdot l$, $v = 2^n \cdot k$ and $w = 2^m \cdot r$ where l, k and r are odd numbers.

a) Provided $1 < k < l$ there are repeating symbol sequences of period u which has smaller order than any repeating symbol sequence of period v.

b) Provided $m > n$ there are repeating symbol sequences of period w which has smaller order than any repeating symbol sequence of period v. □

Sketch of proof. Regarding a) consider S_j such that j is odd. Then by carefully examining the various sequences we find that the discrepancy occurs at entry $2^j(k+2)$ in the repeating sequence of the $2^n \cdot k$ periodic point and it happens as the last entry of the $\hat{S}_j$ block (which of course is 1 since j is odd) differs from the same entry in the $2^n \cdot l$ sequence. Now, since $S_j\hat{S}_j$ has an odd number of 1's the number of 1's before the discrepancy is even, so according to Definition 1.10.3a we have that sequences of period $2^n \cdot l$ are smaller than any sequence of period $2^n \cdot k$. (The case that j is even is left to the reader.)

Regarding b), scrutinizing sequences $\overline{a}$ of period $2^m \cdot k$ it is clear that all of them have 1 0 1 1 as the first entries and that $a_i = 1$ if i is even and $a_i = 0$ if $i = 1 (\mathrm{mod}\, 4)$. Moreover, assuming $k > r$ whenever $m > n$ we find that at discrepancy the sequence of period w has 1 as its element and in fact it is the last 1 in 1 0 1 1. Now, since $S_j\hat{S}_jS_j \dots S_j$ has an even number of 1's the observation above implies that the sequence of period $2^n \cdot k$ must have an even number of 1's before the discrepancy so the result follows. □

Now at last, combining the results from Lemmas 1.10.1–1.10.6 we have established the following ordering for the itineraries of periodic points that satisfy (A) and (B):

$$2 \prec 2^2 \prec 2^3 \prec \dots \prec 2^n \prec 2^n(2l+1) \prec 2^n(2l-1) \prec \dots \prec 2^n \cdot 5 \prec 2^n \cdot 3 \prec$$
$$2^{n-1}(2l+1) \prec \dots \prec 2^{n-1} \cdot 3 \prec \dots \prec (2l+1) \prec (2l-1) \dots \prec 5 \prec 3$$

which is nothing but the ordering we find in Sarkovskii's theorem.

We do not claim that we actually have proved the theorem in all its details, our main purpose here have been to show that symbolic dynamics is a powerful tool when dealing with periodic orbits. For further reading, also of other aspects of symbolic dynamics we refer to Guckenheimer and Holmes (1990), Devaney (1989) and Collet and Eckmann (1980).

1.11 Chaos

As we have seen, the dynamics of $x \to \mu x(1-x)$ differs substantially depending on the value of the parameter μ. For $2 < \mu < 3$ there is a stable nontrivial fixed point, and in case of larger values of μ we have detected periodic orbits both of even and odd period. If $\mu > 2 + \sqrt{5}$ the dynamics is aperiodic and irregular and occurs on a Cantor set Λ and points $x \in (I \setminus \Lambda)$ approaches $-\infty$. (I is the unit interval.)

In this section we shall deal with the concept chaos. Chaos may and has been defined in several ways. We have already used the concept when we stated "Period three implies chaos".

Referring to the examples and exercises at the end of Section 1.3 we found that whenever the long-time behaviour of a system was a stable fixed point or a stable periodic orbit there was no sensitive dependence on the initial condition x_0. However, when $x \to f(x) = 4x(1-x)$ we have proved that there is no stable periodic orbit and moreover, we found a strong sensitivity on the initial condition. Assuming $x \in [0,1]$ and that $x_0 = 0.30$ is one initial condition and $x_{00} = 0.32$ is another we have $|x_0 - x_{00}| = 0.02$ but most terms $|f^k(x_0) - f^k(x_{00})| > 0.02$ and for some k ($k = 9$) $|f^k(x_0) - f^k(x_{00})| \approx 1^-$ which indeed shows a strong sensitivity.

Motivated by the example above, if an orbit of a map $f : I \to I$ shall be denoted as chaotic it is natural to include that f has sensitive dependence on the initial condition in the definition. It is also natural to claim that there is no convergence to any periodic orbit which is equivalent to, say, that periodic orbits must be dense in I. Our goal is to establish a precise definition of the concept chaos but before we do that let us first illustrate what we have discussed above by two examples.

Example 1.11.1. This is a "standard" example which may be found in many textbooks. Consider the map $h : S' \to S'$, $\theta \to h(\theta) = 2\theta$. ($h$ is a map from the circle to the circle.) Clearly, h is sensitive to initial conditions since the arc length between nearby points is doubled under h. Regarding the dense property, observe that $h^n(\theta) = 2^n\theta$ so any periodic points must be obtained from the relation $2^n\theta = \theta + 2k\pi$ or $\theta = 2k\pi/(2^n - 1)$ where the integer k satisfies $0 \leq k \leq 2^n$. Hence in any neighbourhood of a point in S there is a periodic point so the periodic points are dense so h does not converge to any stable periodic orbit. Consequently, h is chaotic on S'. □

Example 1.11.2. Consider $x \to f(x) = \mu x(1 - x)$ where $\mu > 2 + \sqrt{5}$. We claim that f is chaotic on the Cantor set Λ. In order to show sensitive dependence on the initial condition let δ be less than the distance between the intervals I_0 and I_1 (cf. Figure 7). Next, assume $x, y \in \Lambda$ where $x \neq y$. Then the itineraries $\phi(x) \neq \phi(y)$ so after, say, k iterations $f^k(x)$ is in I_0 (I_1) and $f^k(y)$ is in I_1 (I_0). Thus $|f^k(x) - f^k(y)| > \delta$ which establishes the sensitive dependence.

Since $f : \Lambda \to \Lambda$ is topological equivalent to the shift map $\sigma : \Sigma_2 \to \Sigma_2$ it suffices to show that the periodic points of σ are dense in Σ_2. Let $\overline{a} = (a_1 \ldots a_n)$ be a repeating sequence of a periodic point and let $\overline{b} = (a_1 a_2 a_3 \ldots)$ be the sequence of an arbitrary point and note that $\sigma^n(\overline{a}) = \overline{a}$. By use of the distance d between two symbol sequences one easily obtains $d[\overline{a}, \overline{b}] < 1/2^n$ so in any neighbourhood of an arbitrary sequence (point) there is a periodic sequence (periodic point). Hence periodic points of f are dense (and unstable). □

In our work towards a definition of chaos we will now focus on the sensitive dependence on the initial condition.

If a map $f : \mathbb{R} \to \mathbb{R}$ has a fixed point we know from Section 1.4 that if the eigenvalue λ of the linearized system satisfies $-1 < \lambda < 1$ the fixed point is stable and not sensitive to changes of the initial condition. If $|\lambda| > 1$ one may measure the degree of sensitivity by the size of $|\lambda|$. We may use the same argument if we deal with periodic orbits of period k except that we on this occasion consider the eigenvalue of every periodic point contained on the orbit. If a system is chaotic it is natural to consider the case $k \to \infty$ since we may think of a chaotic orbit as one having an infinite period. Therefore, define

$$\eta = \lim_{k\to\infty} \left| \frac{d}{dx} f^k(x)_{x=x_0} \right|^{1/k} \tag{1.11.1}$$

where we have used the k'th root in order to avoid problems in order to obtain a well defined limit. If x_0 is a fixed point $\lambda = |(df/dx)(x = x_0)|$. For a general orbit starting at x_0 we may think of η as an average measure of sensitivity (or insensitivity) over the whole orbit. Let $L = \ln \eta$, that is

$$L = \lim_{k\to\infty} \ln \left| \frac{d}{dx} f^k(x_0) \right|^{1/k} = \lim_{k\to\infty} \frac{1}{k} \sum_{n=0}^{k-1} \ln |f'(x = x_n)| \tag{1.11.2}$$

The number L is called the Lyapunov exponent and if $L > 0$ (which is equivalent to $|\lambda| > 1$) we have sensitive dependence on the initial condition. By use of L we may now define chaos.

Definition 1.11.1. The orbit of a map $x \to f(x)$ is called chaotic if

1) It possesses a positive Lyapunov exponent, and

2) it does not converge to a periodic orbit (that is, there does not exist a periodic orbit $y_t = y_{t+T}$ such that $\lim_{t\to\infty} |x_t - y_t| = 0$.) □

Note that 2) is equivalent to, say, that periodic orbits are dense.

In most cases the Lyapunov exponent must be computed numerically and in cases where L is slightly larger than zero such computations have to be performed by some care due to accumulation effects of round-off errors. Note, however, that there exists a theorem saying that L is stable under small perturbations of an orbit.

Example 1.11.3. Compute L for the map $h : S' \to S'$, $h(\theta) = 2\theta$. In this case $h' = 2$ for all points on the orbit so

$$L = \lim_{k\to\infty} \frac{1}{k} \sum_{n=0}^{k-1} \ln |h'(x = x_n)| = \lim_{k\to\infty} \frac{1}{k} \cdot k \ln 2 = \ln 2 > 0$$

and since the periodic orbits are dense h is chaotic. □

Example 1.11.4. Compute L for the two periodic orbit of $x \to f(x) = \mu x(1-x)$ where $3 < \mu < 1+\sqrt{6}$. Referring to formulae (1.3.3) the periodic points are

$$p_{1,2} = \frac{\mu + 1 \pm \sqrt{(\mu+1)(\mu-3)}}{2\mu}$$

Thus,

$$\begin{aligned} L &= \lim_{k\to\infty} \frac{1}{k} \{\ln |f'(x=p_1)| + \ln |f'(x=p_2)| + \ln |f'(x=p_1)| + \ldots + \ln |f'(x=p_2)|\} \\ &= \lim_{k\to\infty} \frac{1}{k} \left\{ \frac{k}{2} \ln |f'(x=p_1)| + \frac{k}{2} \ln |f'(x=p_2)| \right\} \\ &= \frac{1}{2} \ln |f'(x=p_1) f'(x=p_2)| \end{aligned}$$

Since

$$f'(x=p_1)f'(x=p_2) = \mu(1-2p_1)\mu(1-2p_2) = 1 - (\mu+1)(\mu-3)$$

it follows that $L = (1/2) \ln |1-(\mu+1)(\mu-3)|$ and as expected $L < 0$ whenever $3 < \mu < 1+\sqrt{6}$. (Note that if $\mu > 1+\sqrt{6}$ then $L > 0$ but the map is of course not chaotic since there in this case (provided $|\mu - (1+\sqrt{6})|$ small) exists a stable 4-periodic orbit with negative L.) □

Example 1.11.5. Show that the Lyapunov exponents of almost all orbits of the map $f : [0,1] \to [0,1]$, $x \to f(x) = 4x(1-x)$ is $\ln 2$.

Solution: From Proposition 1.2.1 we know that $f(x)$ is topological equivalent to the tent map $T(x)$. The "nice" property of $T(x)$ which we shall use is that $T'(x) = 2$ for all $x \neq c = 1/2$. Moreover, $h \circ f = T \circ h$ implies that $h'(f(x))f'(x) = T'(h(x))h'(x)$ so

$$f'(x) = \frac{T'(h(x))h'(x)}{h'(f(x))}$$

We are now ready to compute the Lyapunov exponent:

$$\begin{aligned} L &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x = x_i)| \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| \frac{T'(h(x_i))h'(x_i)}{h'(f(x_i))} \right| \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |T'(h(x_i))| + \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \{\ln |h'(x_i)| - \ln |h'(f(x_i))|\} \end{aligned}$$

Since $x_{i+1} = f(x_i)$ the latter sum may be written as

$$\lim_{n\to\infty} \frac{1}{n} \{\ln |h'(x_0)| - \ln |h'(x_n)|\}$$

which is equal to zero for almost all orbits. Thus, for almost all orbits:

$$L = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |T'(h(x_i))| = \lim_{n\to\infty} \frac{1}{n} \cdot n \ln 2 = \ln 2$$

□

For comparison reasons we have also computed L numerically with initial value $x_0 = 0.30$ in the example above. Denoting the Lyapunov exponent of n iterations for L_n we find $L_{100} = 0.67547$, $L_{1000} = 0.69227$ and $L_{5000} = 0.69308$ so in this example we do not need too many terms in order to show that $L > 0$.

A final comment is that since we have proved earlier (cf. Example 1.8.1) that the quadratic map does not possess any stable orbits in case of $\mu = 4$, Definition 1.11.1 directly gives that almost all orbits of the map are chaotic.

1.12 Superstable orbits and a summary of the dynamics of the quadratic map

The quadratic map has two fixed points. One is the trivial one $x^* = 0$ which is stable if $\mu < 1$ and unstable if $\mu > 1$. If $\mu > 1$ the nontrivial fixed point is $x^* = (\mu - 1)/\mu$ and as we have shown this fixed point is stable whenever $1 < \mu < 3$. Whenever $\mu > 2$ the fixed point is larger than the critical point c. At $\mu = 3$ the map undergoes a supercritical flip bifurcation and in the interval $3 < \mu < 1 + \sqrt{6}$ the quadratic map possesses a stable period-2 orbit which has a negative Lyapunov exponent. The periodic points are given by formulae (1.3.3).

At the threshold $\mu = 1+\sqrt{6}$ there is a new (supercritical) flip bifurcation which creates a stable orbit of period 2^2 and through further increase of μ stable orbits of period 2^k are established. However, the parameter intervals where the period 2^k cycles are stable shrinks as μ is enlarged so the μ values at the bifurcation points act more or less as terms in a geometric series. By use of the Feigenbaum geometric ratio one can argue that there exists an accumulation value μ_a for the series of flip bifurcations. Regarding the quadratic map, $\mu_a = 3.56994$. In the parameter interval $\mu_a < \mu \leq 4$ we have seen that the dynamics is much more complicated.

Still considering periodic orbits, Sarkovskii's theorem tells us that periodic orbits occur in a definite order so beyond μ_a there are periodic orbits of periods given by Theorem 1.7.2 (see also Section 1.10). Even in cases where such orbits are stable they may be difficult to distinguish from non-periodic orbits due to the long period. In many respects the ultimate event occurs at the threshold $\mu = 1 + \sqrt{8}$ where a 3-periodic orbit is created because period 3 implies orbits of all other periods which is the content both in Li and Yorke and in Sarkovskii's theorem.

Chaotic orbits may be captured by use of Lyapunov exponents. In Figure 8 we show the value of the Lyapunov exponent L for $\mu \in [\mu_a, 4]$. $L < 0$ corresponds to stable periodic orbits, $L > 0$ corresponds to chaotic orbits. (Figure 8 should be compared to the bifurcation diagram, Figure 6.) The regions where we have periodic orbits are often

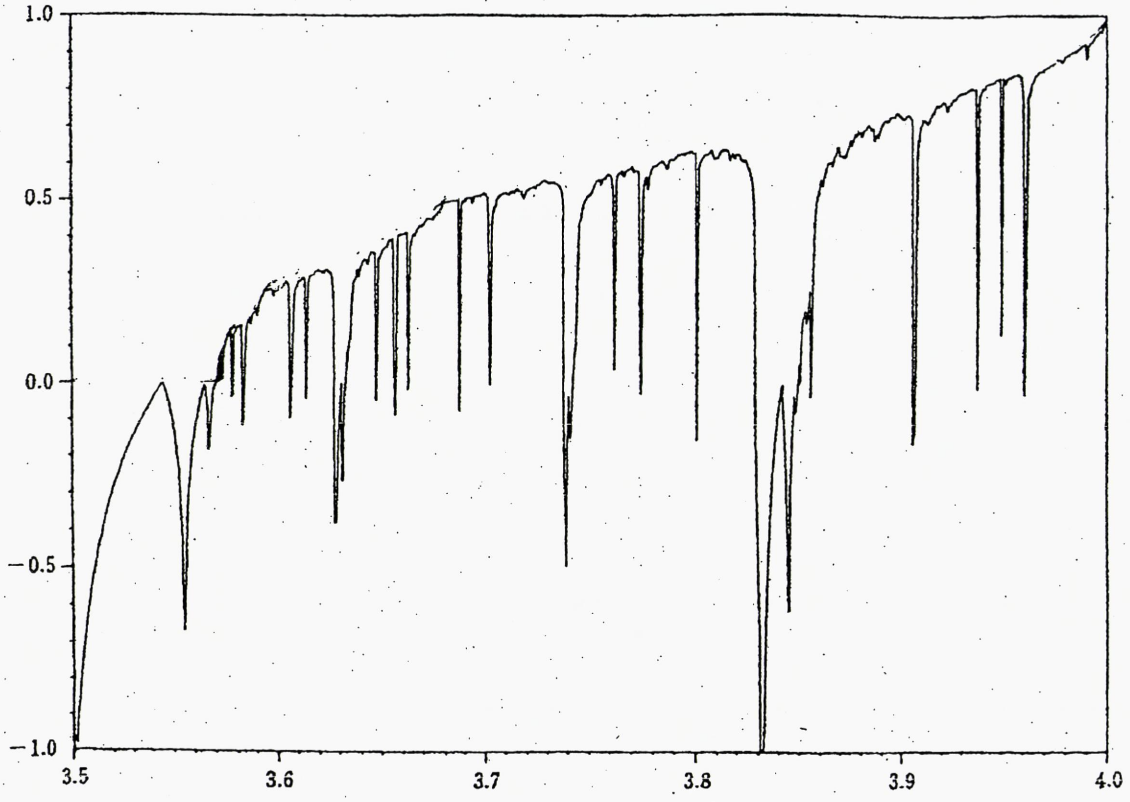


Figure 8: The value of the Lyapunov exponent for $\mu \in [\mu_a, 4]$. $L < 0$ corresponds to stable periodic orbits. $L > 0$ corresponds to chaotic orbits.

referred to as windows. The largest window found in Figure 6 (or 8) is the period 3 window. The periodic orbits in the interval $3 < \mu < \mu_a$ are created through a series of flip bifurcations. However, the period-3 orbit is created through a saddle-node bifurcation. In fact, every window of periodic orbits beyond μ_a is created in this way so just beyond the bifurcation value there is one stable and one unstable orbit of the same period. (If μ is slightly larger than $1+\sqrt{8}$ there is one stable and one unstable orbit of period 3.) Within a window there may be flip bifurcations before chaos is established again, cf. Figure 6. Since the quadratic map has negative Schwarzian derivative there is at most one stable periodic orbit for each value of μ.

There is a way to locate the periodic windows. The vital observation is that at the critical point c, $f'(c) = 0$, so accordingly $\ln|f'(c)| = -\infty$ which implies $L < 0$ and consequently a stable periodic orbit. Also, confer Singer's theorem (Theorem 1.8.1).

Definition 1.12.1. Given a map $f : I \to I$ with one critical point c. Any periodic orbit π passing through c is called a superstable orbit. $\square$

Hence, by searching for superstable orbits one may obtain a representative value of the location of a periodic window. Indeed, any superstable orbit of period n must satisfy the equation

$$f_\mu^n(c) = c \tag{1.12.1}$$

Example 1.12.1. Consider the quadratic map and let us find the value of μ such that $f_\mu^3(1/2) = 1/2$.

We have

$$\begin{aligned} c = \frac{1}{2} &\Rightarrow f_\mu(c) = \frac{1}{4}\mu \Rightarrow f_\mu^2(c) = \frac{1}{4}\mu^2 - \frac{1}{16}\mu^3 \\ &\Rightarrow f_\mu^3(c) = \left(\frac{1}{4}\mu^3 - \frac{1}{16}\mu^4\right)\left\{1 - \left(\frac{1}{4}\mu^2 - \frac{1}{16}\mu^3\right)\right\} \end{aligned}$$

Hence, the equation $f_\mu^3(1/2) = 1/2$ becomes

$$\mu^7 - 8\mu^6 + 16\mu^5 + 16\mu^4 - 64\mu^3 + 128 = 0 \tag{1.12.2}$$

By inspection, $\mu = 2$ is a solution of (1.12.2) so after dividing by $\mu - 2$ we arrive at

$$\mu^6 - 6\mu^5 + 4\mu^4 + 24\mu^3 - 16\mu^2 - 32\mu - 64 = 0 \tag{1.12.3}$$

This equation may be solved numerically by use of Newton's method and if we do that we find that the only solution in the interval $\mu_a \leq \mu \leq 4$ is $\mu = 3.83187$. Therefore, there is only one period-3 window and the location clearly agrees both with the bifurcation diagram, Figure 6 and Figure 8. In the same way, by solving $f_\mu^4(1/2) = 1/2$ one finds that the only solution which satisfies $\mu_a < \mu < 4$ is $\mu = 3.963$ which shows that there is also only one period-4 window. However, if one solves $f_\mu^5(1/2) = 1/2$ one obtains three values which means that there exists three period-5 windows. The first one occurs around $\mu_1 = 3.739$ and is visible in the bifurcation diagram, Figure 6. The others have almost no widths, the values that correspond to the superstable orbits are $\mu_2 = 3.9057$ and $\mu_3 = 3.9903$. □

Referring to the numerical examples given at the end of Section 1.3 where $\mu < \mu_a$ we observed a rapid convergence towards the 2-period orbit independent on the choice of initial value. Within a periodic window in the interval $[\mu_a, 4]$ the dynamics may be much more complicated. Indeed, still considering the period-3 window, we have according to the Li and Yorke theorem that there are also periodic orbits of any period, although invisible to a computer. (The latter is a consequence of Singer's theorem.) If we consider an initial point which is not on the 3-periodic orbit we may see that it behaves irregularly through lots of iterations before it starts to converge, and moreover, if we change the initial point somewhat it may happen that it is necessary to perform an even larger amount of iterations before we are able to detect any convergence towards the 3-cycle. Hence, the dynamics within a periodic window in the interval $[\mu_a, 4]$ is in general much more complex than in the case of periodic orbits in the interval $[3, \mu_a]$ due to the presence of an (infinite) number of unstable periodic points.

By carefully scrutinizing the periodic windows one may find numerically that the sum of the widths of all the windows is roughly 10% of the length of the interval $[\mu_a, 4]$. In the remaining part of the interval the dynamics is chaotic. If we want to give a thorough description of chaotic orbits we may use symbolic dynamics in much of a similar way as we did in Sections 1.9 and 1.10. Here we shall give a more heuristic approach only. If μ is not close to a periodic window, orbits are irregular and there is almost no sign of periodicity. However, if μ is close to a window, for example, if μ is smaller but close to $1 + \sqrt{8}$ (the threshold value for the period-3 window) one finds that an orbit seems to consist of two parts, one part with appears to be almost 3-periodic and another irregular part where the point x may take almost any value in $(0, 1)$. The almost 3-periodic part of the orbit is established when the orbit becomes close to the diagonal line $x_{t+1} = x_t$. Then, since μ is close to $1 + \sqrt{8}$ the orbit may stay close to the diagonal for several iterations before it moves away. Therefore, a typical orbit close to a periodic window consists of an irregular part which after a finite number of iterations becomes almost periodic and again turns irregular in a repeating fashion.

Part II
n-dimensional maps
$$f : \mathbb{R}^n \to \mathbb{R}^n \quad \mathbf{x} \to f(\mathbf{x})$$

2.1 Higher order difference equations

Consider the second order difference equation

$$x_{t+2} + a_t x_{t+1} + b_t x_t = f(t) \tag{2.1.1}$$

If $f(t) \neq 0$, (2.1.1) is called a nonhomogeneous difference equation. If $f(t) = 0$, that is

$$x_{t+2} + a_t x_{t+1} + b_t x_t = 0 \tag{2.1.2}$$

we have the associated homogeneous equation.

Theorem 2.1.1. The homogeneous equation (2.1.2) has the general solution

$$x_t = C_1 u_t + C_2 v_t$$

where u_t and v_t are two linear independent solutions and C_1, C_2 arbitrary constants.

Proof. Let $x_t = C_1 u_t + C_2 v_t$. Then $x_{t+1} = C_1 u_{t+1} + C_2 v_{t+1}$ and $x_{t+2} = C_1 u_{t+2} + C_2 v_{t+2}$ and if we substitute into (2.1.2) we obtain

$$C_1(u_{t+2} + a_t u_{t+1} + b_t u_t) + C_2(v_{t+2} + a_t v_{t+1} + b_t v_t) = 0$$

which clearly is correct since u_t and v_t are linear independent solutions. □

Regarding (2.1.1) we obviously have:

Theorem 2.1.2. The nonhomogeneous equation (2.1.1) has the general solution

$$x_t = C_1 u_t + C_2 v_t + u_t^*$$

where $C_1 u_t + C_2 v_t$ is the general solution of the associated homogeneous equation (2.1.2) and u_t^* is any particular solution of (2.1.1).

Just as in case of differential equations there is no general method of how to find two linear independent solutions of a second order difference equation. However, if the coefficients a_t and b_t are constants then it is possible.

Indeed, consider

$$x_{t+2} + ax_{t+1} + bx_t = 0 \tag{2.1.3}$$

where a and b are constants. Suppose that there exists a solution of the form $x_t = m^t$ where $m \neq 0$. Then $x_{t+1} = m^{t+1} = mm^t$ and $x_{t+2} = m^2 m^t$ so (2.1.3) may be expressed as

$$(m^2 + am + b)m^t = 0$$

which again implies that

$$m^2 + am + b = 0 \tag{2.1.4}$$

(2.1.4) is called the characteristic equation and its solution is easily found to be

$$m_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b} \tag{2.1.5}$$

Now we have the following result regarding the solution of (2.1.3) which we state as a theorem:

Theorem 2.1.3.

1) If $(a^2/4) - b > 0$, the characteristic equation have two real solutions m_1 and m_2. Moreover, m_1^t and m_2^t are linear independent so according to Theorem 2.1.1 the general solution of (2.1.3) is

$$x_t = C_1 m_1^t + C_2 m_2^t \quad \text{where} \quad m_{1,2} = -\frac{a}{2} \pm \sqrt{\frac{a^2}{4} - b}$$

2) The case $(a^2/4) - b = 0$ implies that $m = -a/2$. Then m^t and tm^t are two linear independent solutions of (2.1.3) so the general solution becomes:

$$x_t = C_1 m^t + C_2 t m^t = (C_1 + C_2 t)m^t \quad \text{where} \quad m = -a/2$$

(In order to see that tm^t really is a solution of (2.1.3) note that if $a^2/4 = b$, then (2.1.3) may be expressed as (*) $x_{t+2} + ax_{t+1} + (a^2/4)x_t = 0$. Now, assuming that $x_t = t(-a/2)^t$ we have $x_{t+1} = -(a/2)(t+1)(-a/2)^t$, $x_{t+2} = (a^2/4)(t+2)(-a/2)^t$ and by inserting into (*) we obtain $(a^2/4)[t+2-2(t+1)+t](-a/2)^t = 0$ which proves what we want.)

3) Finally, if $(a^2/4) - b < 0$ we have

$$m = -\frac{a}{2} \pm \sqrt{-(b-(a^2/4)} = -\frac{a}{2} \pm \sqrt{b-(a^2/4)}\, i = \alpha + \beta i$$

From the theory of complex numbers we know that

$$\alpha + \beta i = r(\cos\theta + i\sin\theta)$$

where

$$r = \sqrt{\alpha^2 + \beta^2} = \sqrt{(-a/2)^2 + \sqrt{b-(a^2/4)}^2} = \sqrt{b}$$

and

$$\cos\theta = \frac{-a/2}{\sqrt{b}} \qquad \sin\theta = \frac{\sqrt{b-(a^2/4)}}{\sqrt{b}}$$

which implies that

$$m^t = [r(\cos\theta + i\sin\theta)]^t = r^t(\cos\theta + i\sin\theta)^t = r^t(\cos\theta\, t + i\sin\theta\, t)$$

where we have used Moivre's formulae in the last step. Since the real and imaginary parts of m^t are linear independent functions we express the general solution of (2.1.3) as

$$x_t = C_1 r^t \cos\theta\, t + C_2 r^t \sin\theta\, t$$

□

Example 2.1.1. Find the general solution of the following equations:

a) $x_{t+2} - 7x_{t+1} + 12x_t = 0$,

b) $x_{t+2} - 6x_{t+1} + 9x_t = 0$,

c) $x_{t+2} - x_{t+1} + x_t = 0$.

Solutions:

a) Assuming $x_t = m^t$ the characteristic equation becomes $m^2 - 7m + 12 = 0 \Leftrightarrow m_1 = 4$, $m_2 = 3$ so according to Theorem 2.1.3 the general solution is $x_t = C_1 \cdot 4^t + C_2 \cdot 3^t$.

b) The characteristic equation is $m^2 - 6m + 9 = 0 \Leftrightarrow m_1 = m_2 = 3$. Thus $x_t = C_1 \cdot 3^t + C_2 t \cdot 3^t = (C_1 + C_2 t)3^t$.

c) The characteristic equation becomes $m^2 - m + 1 = 0 \Leftrightarrow m = (1 \pm \sqrt{-3})/2 = \frac{1}{2} \pm \frac{1}{2}\sqrt{3}\,i$.

Further

$$r = \sqrt{\left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\sqrt{3}\right)^2} = 1$$

$$\cos\theta = \frac{\frac{1}{2}}{1} = \frac{1}{2} \qquad \sin\theta = \frac{\frac{1}{2}\sqrt{3}}{1} = \frac{1}{2}\sqrt{3} \Rightarrow \theta = \frac{\pi}{3}$$

Thus

$$x_t = C_1 1^t \cos\frac{\pi}{3}t + C_2 1^t \sin\frac{\pi}{3}t = C_1 \cos\frac{\pi}{3}t + C_2 \sin\frac{\pi}{3}t$$

□

Exercise 2.1.1. Find the general solution of the homogeneous equations:

a) $x_{t+2} - 12x_{t+1} + 36x_t = 0$,

b) $x_{t+2} + x_t = 0$,

c) $x_{t+2} + 6x_{t+1} - 16x_t = 0$. □

Exercise 2.1.2. Prove Moivre's formulae: $(\cos\theta + i\sin\theta)^t = \cos\theta\, t + i\sin\theta\, t$. (Hint: Use induction and trigonometric identities.) □

Definition 2.1.1. The equation $x_{t+2} + ax_{t+1} + bx_t = 0$ is said to be globally asymptotic stable if the solution x_t satisfies $\lim_{t\to\infty} x_t = 0$. □

Referring to Example 2.1.1 it is clear that none of the equations considered there are globally asymptotic stable. The solutions of the equations (a) and (b) tend to infinity as $t \to \infty$ and the solution of (c) does not tend to zero either.

However, consider the equation $x_{t+2} - (1/6)x_{t+1} - (1/6)x_t = 0$. The characteristic equation is $m^2 - (1/6)m - (1/6)m = 0 \Leftrightarrow m_1 = 1/2$, $m_2 = -(1/3)$ so the general solution becomes $x_t = C_1(1/2)^t + C_2(-1/3)^t$.

Here, we obviously have $\lim_{t\to\infty} x_t = 0$ so according to Definition 1.2.1 the equation $x_{t+2} + (1/6)x_{t+1} - (1/6)x_t = 0$ is globally asymptotic stable.

Theorem 2.1.4. The equation $x_{t+2} + ax_{t+1} + bx_t = 0$ with associated characteristic equation $m^2 + am + b = 0$ is globally asymptotic stable if and only if all the roots of the characteristic equation have moduli strictly less than 1. □

Proof. Referring to Theorem 2.1.3, the cases (1) and (3) are clear (remember $|m| = r$ in (3)).

Considering (2): If $|m| < 1$

$$\lim_{t\to\infty} tm^t = \lim_{t\to\infty} \frac{t}{s^t}$$

where $s = 1/m$ and $s > 1$. Then by L'hopital's rule

$$\lim_{t\to\infty} \frac{t}{s^t} = \lim_{t\to\infty} \frac{1}{s^t \ln s} \to 0$$

and the results of Theorem 2.1.4 follows.

As we shall see later on, Theorem 2.1.4 will be useful for us when we discuss stability of nonlinear systems.

—

We close this section by considering the nonhomogeneous equation

$$x_{t+2} + ax_{t+1} + bx_t = f(t) \tag{2.1.6}$$

According to Theorem 2.1.2 the general solution of (2.1.6) is the sum of the general solution of the homogeneous equation (2.1.3) and a particular solution u_t^* of (2.1.6).

If $f(t)$ is a polynomial, say $f(t) = 2t^2 + 4t$ it is natural to assume a particular solution of the form $u_t^* = At^2 + Bt + C$.

If $f(t)$ is a trigonometric function, for example $f(t) = \cos u\, t$ we assume that $u_t^* = A\cos u\, t + B\sin u\, t$.

If $f_t = c^t$, assume $u_t^* = Ac^t$ (but see the comment following (2.1.7)).

Example 2.1.2. Solve the following equations:

a) $x_{t+2} + x_{t+1} + 2x_t = t^2$,

b) $x_{t+2} - 2x_{t+1} + x_t = 2\sin(\pi/2)t$,

Solutions:

a) The characteristic equation of the homogeneous equation becomes $m^2 - m - 2 = 0 \Leftrightarrow m_1 = 2$ and $m_2 = -1$ so the general solution of the homogeneous equation is $x_t = C_1{\cdot}2^t{+}C_2(-1)^t$. Assume $u_t^* = At^2{+}Bt{+}C$. Then $u_{t+1}^* = A(t+1)^2 + B(t+1) + C$, $u_{t+2}^* = A(t+2)^2 + B(t+2) + C$ which inserted into the original equation gives

$$A(t{+}2)^2{+}B(t{+}2){+}C{-}[A(t+1)^2 + B(t+1) + C]{-}2\,[At^2 + Bt + C] = t^2$$

$$\Leftrightarrow$$

$$-2At^2 + (2A - 2B)t + (3A + B - 2C) = t^2 + 0t + 0$$

and by equating terms of equal powers of t we have (1) $-2A = 1$, (2) $2A - 2B = 0$, and (3) $3A + B - 2C = 0$ from which we easily obtain $A = -1/2$, $B = -1/2$ and $C = -1$. Thus $u_t^* = -(1/2)t^2 - (1/2)t - 1$ and the general solution is $x_t = C_1 2^t + C_2(-1)^t - (1/2)t^2 - (1/2)t - 1$.

b) The solution of the characteristic equation becomes $m_1 = m_2 = 1 \Rightarrow$ homogeneous solution $(C_1 + C_2 t)1^t = C_1 + C_2 t$. Assume $u_t^* = A\cos(\pi/2)t + B\sin(\pi/2)t$. Then, $u_{t+1}^* = A\cos[(\pi/2)(t+1)] + B\sin[(\pi/2)(t+1)] =$

$A[\cos(\pi/2)t\cos(\pi/2) - \sin(\pi/2)t\sin(\pi/2)] + B[\sin(\pi/2)t\cos(\pi/2)+ \sin(\pi/2)\cos(\pi/2)] = -A\sin(\pi/2)t + B\cos(\pi/2)t$. In the same way, $u^*_{t+2} = -A\cos(\pi/2)t - B\sin(\pi/2)t$ so after inserting u^*_{t+2}, u^*_{t+1} and u^*_t into the original equation we arrive at

$$-2B\cos\frac{\pi}{2}t + 2A\sin\frac{\pi}{2}t = 0\cos\frac{\pi}{2}t + 2\sin\frac{\pi}{2}t$$

Thus $-2B = 0$ and $2A = 2 \Leftrightarrow A = 1$ and $B = 0$ so $u^*_t = \cos(\pi/2)t$. Hence, the general solution is $x_t = C_1 + C_2 t + \cos(\pi/2)t$. □

Finally, if $x_{t+2} + ax_{t+1} + bx_t = c^t$ we assume a particular solution of the form $u^*_t = Ac^t$. Then $u^*_{t+1} = Acc^t$ and $u^*_{t+2} = Ac^2c^t$ which inserted into the original equation yields

$$A(c^2 + ac + b)c^t = c^t$$

Thus, whenever $c^2 + ac + b \neq 0$ the particular solution becomes

$$u^* = \frac{1}{c^2 + ac + b}c^t \tag{2.1.7}$$

Note, however, that if c is a simple root of the characteristic equation, i.e. $c^2 + ac + b = 0$, then we try a solution of the form $u^*_t = Btc^t$ and if c is a double root, assume $u^*_t = Dt^2c^t$.

Example 2.1.3. Solve the equations:

a) $x_{t+2} - 4x_t = 3^t$,

b) $x_{t+2} - 4x_t = 2^t$,

Solutions:

a) The characteristic equation is $m^2 - 4 = 0 \Leftrightarrow m_1 = 2$, $m_2 = -2$ thus the homogeneous solution is $C_1 2^t + C_2(-2)^t$. Since 3 is not a root of $m^2 - 4 = 0$ we have directly from (2.1.7) that $u^*_t = (1/5)3^t$ so the general solution becomes $x_t = C_1 2^t + C_2(-2)^t + (1/5)3^t$.

b) The homogeneous solution is of course $C_1 2^t + C_2(-2)^t$ but since 2 is a simple root of $m^2 - 4 = 0$ we try a particular solution of the form $u_t^* = Bt2^t$. Then $u_{t+2}^* = 4B(t+2)2^t$ and by inserting into the original equation we arrive at

$$4B(t+2)2^t - 4Bt \cdot 2^t = 2^t$$

which gives $B = 1/8$. Thus $x_t = C_1 2^t + C_2(-2)^t + (1/8)t \cdot 2^t$. □

Exercise 2.1.3. Solve the problems:

a) $x_{t+2} + 2x_{t+1} - 3x_t = 2t + 5$,

b) $x_{t+2} - 10x_{t+1} + 25x_t = 5^t$, c) $x_{t+2} - x_{t+1} + x_t = 2^t$,

d) $x_{t+2} + 9x_t = 2^t$, e) $x_{t+2} - 5x_{t+1} - 6x_t = t \cdot 2^t$.

(Hint: Assume a particular solution of the form $(At + B) \cdot 2^t$.) □

Exercise 2.1.4. Consider the equation $x_{t+2} = x_{t+1} + x_t$ with initial conditions $x_0 = 0$, $x_1 = 1$.

a) Solve the equation.

b) Use a) and induction to prove that $x_t \cdot x_{t+2} - x_{t+1}^2 = (-1)^{t+1}$, $t = 0, 1, 2, \ldots$. □

Let us now turn to equations of order n, i.e. equations of the form

$$x_{t+n} + a_1(t)x_{t+n-1} + a_2(t)x_{t+n-2} + \cdots + a_{n-1}(t)x_{t+1} + a_n(t)x_t = f(t) \tag{2.1.8}$$

In the homogeneous case we have the following result:

Theorem 2.1.5. Assuming $a_n(t) \neq 0$, the general solution of

$$x_{t+n} + a_1(t)x_{t+n-1} + \cdots + a_n(t)x_t = 0 \tag{2.1.9}$$

is $x_t = C_1 u_{1,t} + \cdots + C_n u_{n,t}$ where $u_{1,t} \ldots u_{n,t}$ are linear independent solutions of the equation and $C_1 \ldots C_n$ arbitrary constants. □

Proof. Easy extension of the proof of Theorem 2.1.1. We leave the details to the reader. □

Regarding the nonhomogeneous equation (2.1.8) we have

Theorem 2.1.6. The solution of the nonhomogeneous equation (2.1.8) is

$$x_t = C_1 u_{1,t} + \cdots + C_n u_{n,t} + u_t^*$$

where u_t^* is a particular solution of (2.1.8) and $C_1 u_{1,t} + \cdots + C_n u_{n,t}$ is the general solution of (2.1.9). □

If $a_1(t) = a_1, ..., a_n(t) = a_n$ constants we arrive at

$$x_{t+n} + a_1 x_{t+n-1} + \cdots + a_n x_t = f(t) \tag{2.1.10}$$

and as in the second order case we may assume a solution $x_t = m^t$ of the homogeneous equation. This yields the n-th order characteristic equation

$$m^n + a_1 m^{n-1} + \cdots + a_{n-1} m + a_n = 0 \tag{2.1.11}$$

Appealing to the fundamental theorem of algebra we know that (2.1.11) has n roots. If a root is real with multiplicity 1 or complex we form linear independent solutions in exactly the same way as explained in Theorem 2.1.3. In case of real roots with multiplicity p, linear independent solutions are $m^t, tm^t, ..., t^{p-1}m^t$.

Example 2.1.4. Solve the equations:

a) $x_{t+3} - 2x_{t+2} + x_{t+1} - 2x_t = 2t - 4$,

b) $x_{t+3} - 6x_{t+2} + 12x_{t+1} - 8x_t = 0$.

Solutions:

a) The characteristic equation is $m^3 - 2m^2 + m - 2 = 0$. Clearly, $m_1 = 2$ is a solution and $m^3 - 2m^2 + m - 2 = (m-2)(m^2+1) = 0$. Hence the other roots are complex, $m_{2,3} = \pm i$. Following Theorem 2.1.3 $r = \sqrt{0^2 + 1^2} = 1$, $\cos\theta = 0/1 = 0$, $\sin\theta = 1/1 = 1 \Rightarrow \theta = \pi/2$ which implies the homogeneous solution $C_1 \cdot 2^t + C_2 \cos(\pi/2)t + C_3 \sin(\pi/2)t$. Assuming a particular solution $u_t^* = At + B$ we find after inserting into the original equation, $-2At - 2B = 2t - 4$ so $A = -1$ and $B = 2$. Consequently, according to Theorem 2.1.6, the general solution is $x_t = C_1 \cdot 2^t + C_2 \cos(\pi/2)t + C_3 \sin(\pi/2)t - t + 2$.

b) The characteristic equation becomes $m^3 - 6m^2 + 12m - 8 = 0 \Leftrightarrow (m-2)^3 = 0$. Hence, there is only one root, $m = 2$, with multiplicity 3. Consequently, $x_t = C_1 \cdot 2^t + C_2 t \cdot 2^t + C_3 t^2 \cdot 2^t$. □

Exercise 2.1.5. Find the general solution of the equations:

a) $x_{t+3} - 2x_{t+2} - 5x_{t+1} + 6x_t = 0$

b) $x_{t+4} - x_t = 2^t$

c) $x_{t+1} - 2x_t = 1 + t^2$

d) $x_{t+1} - 2x_t = 2^t + 3^t$

□

Definition 2.1.2. The equation $x_{t+n} + a_1 x_{t+n-1} + \cdots + a_n x_t = 0$ is said to be globally asymptotic stable if the solution x_t satisfies $\lim_{t\to\infty} x_t = 0$. □

Theorem 2.1.7. The equation $x_{t+n} + a_1 x_{t+n-1} + \cdots + a_n x_t = 0$ is globally asymptotic stable if all solutions of the characteristic equation (2.1.11) have moduli less than 1. □

It may be a difficult task to decide whether all roots of a given polynomial equation have moduli less than unity or not. However, there are methods and one of the most frequently used is the Jury criteria which we now describe.

Let

$$P(x) = x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_n \tag{2.1.12}$$

be a polynomial with real coefficients $a_1 \ldots a_n$. Define

$$b_n = 1 - a_n^2\,, \quad b_{n-1} = a_1 - a_n a_{n-1}, \cdots\; b_{n-j} = a_j - a_n a_{n-j},\; b_1 = a_{n-1} - a_n a_1$$

$$c_n = b_n^2 - b_1^2\,, \quad c_{n-1} = b_n b_{n-1} - b_1 b_2, \cdots c_{n-j} = b_n b_{n-j} - b_1 b_{j+1},\; c_2 = b_n b_2 - b_1 b_{n-1}$$

$$d_n = c_n^2 - c_2^2\,, \cdots\; d_{n-j} = c_n c_{n-j} - c_2 c_{j+2} \ldots d_3 = c_n c_3 - c_2 c_{n-1}$$

and so on.

$$w_n = v_n^2 - v_{n-3}^2\,, \quad w_{n-1} = v_n v_{n-1} - v_{n-3} v_{n-2}\,, \quad w_{n-2} = v_n v_{n-2} - v_{n-3} v_{n-1}$$

Theorem 2.1.8 (The Jury criteria). All roots of the polynomial equation $P(x) = 0$ where $P(x)$ is defined through (2.1.12) have moduli less than 1 provided:

$$P(1) > 0 \qquad (-1)^n P(-1) > 0$$

$|a_n| < 1$, $|b_n| > |b_1|$, $|c_n| > |c_2|$, $|d_n| > |d_3|$, $\cdots$ $|w_n| > |w_{n-2}|$. □

Remark. Instead of saying that all roots have moduli less than 1, an alternative formulation is to say that all roots are located inside the unit circle in the complex plane. □

Regarding the second order equation

$$x^2 + a_1 x + a_2 = 0 \tag{2.1.13}$$

the Jury criteria become

$$\begin{aligned} 1 + a_1 + a_2 &> 0 \\ 1 - a_1 + a_2 &> 0 \\ 1 - |a_2| &> 0 \end{aligned} \tag{2.1.14}$$

If we have a polynomial equation of order 3

$$x^3 + a_1 x^2 + a_2 x + a_3 = 0 \tag{2.1.15}$$

the Jury criteria may be cast in the form

$$\begin{aligned} 1 + a_1 + a_2 + a_3 &> 0 \\ 1 - a_1 + a_2 - a_3 &> 0 \\ 1 - |a_3| &> 0 \\ |1 - a_3^2| - |a_2 - a_3 a_1| &> 0 \end{aligned} \tag{2.1.16}$$

Evidently, the higher the order of the equation is, the more complicated are the Jury criteria. Therefore, unless the coefficients are very simple or on a special form the method does to work is the order of the polynomial becomes large.

Later, when we shall focus on stability problems of nonlinear maps (which involves the study of polynomial equations), we will also face the fact that the coefficients $a_1 \ldots a_n$ do not consist of numbers only but a mixture of numbers and parameters. In such cases, even (2.1.16) may be difficult to apply.

However, let us give one simple example of how the Jury criteria works.

Example 2.1.5. Show that $x_{t+3} - (2/3)x_{t+2} + (1/4)x_{t+1} - (1/6)x_t = 0$ is globally asymptotic stable.

Solution: According to Theorem 2.1.7 we must show that the roots of the associated characteristic equation $m^3 - (2/3)m^2 + (1/4)m - (1/6) = 0$ are located inside the unit circle. Defining $a_1 = -(2/3)$, $a_2 = 1/4$, $a_3 = -(1/6)$ the four left-hand sides of (2.1.16) become 1/12, 25/12, 5/6 and 5/6, respectively. Consequently, all the roots are located inside the unit circle so the difference equation is globally asymptotic stable. □

Another theorem (from complex function theory) that may be useful and which applies not to polynomial equations only is Rouche's theorem. (In the theorem below, $z = \alpha + \beta i$ is a complex number.)

Theorem 2.1.9 (Rouche's theorem). If $f(z)$ and $g(z)$ are analytic inside and on a simple closed curve C and if $|g(z)| < |f(z)|$ on C then $f(z)+g(z)$ and $f(z)$ and the same number of zeros inside C. □

Remark. If we take the simple closed curve C to be the unit circle $|z| = 1$, then we may use Theorem 2.1.9 in order to decide if all the roots of a given equation have moduli less than one or not. □

Example 2.1.6. Suppose that $a > e$ and show that the equation $az^n - e^z = 0$ has n roots located inside the unit circle $|z| = 1$.

Solution: Define $f(z) = az^n$, $g(z) = -e^z$ and consider $f(z) + g(z) = 0$. Clearly, the equation $f(z) = 0$ has n roots located inside the unit circle. On the boundary of the unit circle we have $|g(z)| = |-e^z| \leq e < a = |f(z)|$. Thus, according to Theorem 2.1.9, $f(z)$ and $f(z)+g(z)$ have the same number of zeros inside the unit circle, i.e. n zeros. □

2.2 Systems of linear difference equations. Linear maps from $\mathbb{R}^n$ to $\mathbb{R}^n$

In this section our purpose is to analyse linear systems. There are several alternatives when one tries to find the general solution of such systems. One possible method is to transform a system into one higher order equation and use the theory that we developed in the previous section. Other methods are based upon topics from linear algebra, and of particular relevance is the theory of eigenvalues and eigenvectors. Later when we turn to nonlinear systems and stability problems it will be useful for us to have a broad knowledge of linear systems so therefore we shall deal with several possible solution methods in this section.

Consider the system

$$\begin{aligned}
x_{1,t+1} &= a_{11}x_{1,t} + a_{12}x_{2,t} + \cdots + a_{1n}x_{n,t} + b_1(t) \\
x_{2,t+1} &= a_{21}x_{1,t} + a_{22}x_{2,t} + \cdots + a_{2n}x_{n,t} + b_2(t) \\
&\vdots \\
x_{n,t+1} &= a_{n1}x_{1,t} + a_{n2}x_{2,t} + \cdots + a_{nn}x_{n,t} + b_n(t)
\end{aligned} \tag{2.2.1}$$

Here, all coefficients $a_{11} \ldots a_{nn}$ are constants and if $b_i(t) = 0$ for all $1 \leq i \leq n$ we call (2.2.1) a linear autonomous system.

It is often convenient to express (2.2.1) in terms of vectors and matrices. Indeed, let $\mathbf{x} = (x_1, ..., x_n)^T$, $\mathbf{b} = (b_1, ..., b_n)^T$ and

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \tag{2.2.2}$$

Then, (2.2.1) may be written as

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \mathbf{b}_t \tag{2.2.3}$$

or in map notation

$$\mathbf{x} \to A\mathbf{x} + \mathbf{b} \tag{2.2.4}$$

First, let us show how one may solve a system by use of the theory from the previous section.

Example 2.2.1. Solve the system

$$(1)\ x_{t+1} = 2y_t + t$$

$$(2)\ y_{t+1} = x_t + y_t$$

Replacing t by $t+1$ in (1) gives

$$x_{t+2} = 2y_{t+1} + t + 1 \underset{(2)}{=} 2(x_t + y_t) + t + 1 = 2x_t + 2y_t + t + 1$$

Further, from (1): $2y_t = x_{t+1} - t$. Hence

$$x_{t+2} - x_{t+1} - 2x_t = 1$$

Thus, we have transformed a system to two first order equations into one second order equation, and by use of the theory from the previous section the general solution of the latter equation is easily found to be

$$x_t = C_1 \cdot 2^t + C_2(-1)^t - 1/2$$

y_t may be obtained from (1):

$$\begin{aligned} y_t = \frac{1}{2}(x_{t+1} - t) &= \frac{1}{2}\left(C_1 2^{t+1} + C_2(-1)^{t+1} - \frac{1}{2} - t\right) \\ &= C_1 2^t - \frac{1}{2}C_2(-1)^t - \frac{1}{2}t - \frac{1}{4} \end{aligned}$$

The constants C_1 and C_2 may be determined if we know the initial values x_0 and y_0. For example, if $x_0 = y_0 = 1$ we have from the general solution above that

$$\begin{aligned} 1 &= C_1 + C_2 - 1/2 \\ 1 &= C_1 - \frac{1}{2}C_2 - 1/4 \end{aligned}$$

which implies that $C_1 = 4/3$ and $C_2 = 1/6$ so the solution becomes

$$x_t = \frac{4}{3} \cdot 2^t + \frac{1}{6}(-1)^t - \frac{1}{2} \qquad y_t = \frac{4}{3} \cdot 2^t + \frac{1}{12}(-1)^t - \frac{1}{2}t - \frac{1}{4}$$

□

Exercise 2.2.1. Find the general solution of the systems

$$\text{a) } \begin{aligned} x_{t+1} &= 2y_t + t \\ y_{t+1} &= -x_t + 3y_t \end{aligned} \qquad \text{b) } \begin{aligned} x_{t+1} &= x_t + 2y_t \\ y_{t+1} &= 3x_t \end{aligned}$$

□

Another way to find the solution of a system is to use the matrix formulation (2.2.3). Indeed, suppose that the initial vector $\mathbf{x}_0$ is known. Then:

$$\mathbf{x}_1 = A\mathbf{x}_0 + \mathbf{b}(0)$$
$$\mathbf{x}_2 = A\mathbf{x}_1 + \mathbf{b}(1) = A(A\mathbf{x}_0 + b(0)) + \mathbf{b}(1) = A^2\mathbf{x}_0 + A\mathbf{b}(0) + \mathbf{b}(1)$$

and by induction (we leave the details to the reader)

$$\mathbf{x}_t = A^t\mathbf{x}_0 + A^{t-1}\mathbf{b}(0) + A^{t-2}\mathbf{b}(1) + \cdots + \mathbf{b}(t-1) \qquad (2.2.5)$$

In the important special case $\mathbf{b} = \mathbf{0}$ we have the result:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t \Leftrightarrow \mathbf{x}_t = A^t\mathbf{x}_0 \qquad (2.2.6)$$

where A^0 is equal to the identity matrix I.

Exercise 2.2.2. Consider the matrix

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

a) Compute A^2 and A^3.

b) Let t be a positive integer and use induction to find a formulae for A^t.

c) Let $\mathbf{x} = (x_1, x_2)^T$ and solve the difference equation $\mathbf{x}_{t+1} = A\mathbf{x}_t$ where $\mathbf{x}_0 = (a, b)^T$. □

Our next goal is to solve the linear system

$$\mathbf{x}_{t+1} = A\mathbf{x}_t \qquad (2.2.7)$$

in terms of eigenvalues and eigenvectors. Thus, consider (2.2.7) and assume a solution of the form $\mathbf{x}_t = \lambda^t\mathbf{u}$. Then

$$\lambda^{t+1}\mathbf{u} - A\lambda^t\mathbf{u} = 0$$
$$\Leftrightarrow \qquad (2.2.8)$$
$$(A - \lambda I)\mathbf{u} = 0$$

so λ is nothing but an eigenvalue belonging to A and $\mathbf{u}$ is the associated eigenvector. As is well known, the eigenvalues may be computed from the relation

$$|A - \lambda I| = 0 \tag{2.2.9}$$

There are two cases to consider.

(A) If the $n \times n$ matrix A is diagonalizable over the complex numbers, then A has n distinct eigenvalues $\lambda_1, ..., \lambda_n$ and moreover, the associated eigenvectors $\mathbf{u}_1, ..., \mathbf{u}_n$ are linear independent. Consequently, the general solution of the linear system (2.2.7) may be cast in the form

$$\mathbf{x}_t = C_1\lambda_1^t\mathbf{u}_1 + C_2\lambda_2^t\mathbf{u}_2 + \cdots + C_n\lambda_n^t\mathbf{u}_n \tag{2.2.10}$$

(B) If A is not diagonalizable (i.e. A has multiple eigenvalues) we may proceed in much of the same way as in the corresponding theory for continuous systems, see Grimshaw (1990). Suppose that λ is an eigenvalue with multiplicity m and let $\mathbf{u}_1, ..., \mathbf{u}_p$ be a basis for the eigenspace of λ if $p < m$. Then we seek a solution of the form $\mathbf{x}_t = \lambda^t(\mathbf{v} + t\mathbf{u})$ where $\mathbf{u}$ is one of the $\mathbf{u}_i$'s.

Then from (2.2.7) one easily obtains

$$(A - \lambda I)\mathbf{v} = \mathbf{u} \tag{2.2.11a}$$

$$(A - \lambda I)\mathbf{u} = 0 \tag{2.2.11b}$$

and after multiplying (2.2.11a) with $(A - \lambda I)$ from the left we arrive at

$$(A - \lambda I)^2\mathbf{v} = 0 \tag{2.2.12}$$

Now suppose that we can find $\mathbf{v}_1, ..., \mathbf{v}_q$ such that $\mathbf{v}_1, ..., \mathbf{v}_q$, $\mathbf{u}_1, ..., \mathbf{u}_p$ are linear independent. Let $\hat{\mathbf{u}}_j = (A - \lambda I)\mathbf{v}_j$, $j = 1, ..., q$. We claim that $\hat{\mathbf{u}}_1, ..., \hat{\mathbf{u}}_q$ are linear independent. Indeed, suppose the opposite. Then $d_1\hat{\mathbf{u}}_1 + \cdots + d_q\hat{\mathbf{u}}_q = 0$ where not all constants $d_1 ... d_n = 0$. This implies (see (2.2.11a)) that $d_1\mathbf{v}_1 + \cdots + d_q\mathbf{v}_q = \mathbf{u}$

which contradicts the assumption that $\mathbf{v}_1, ..., \mathbf{v}_q, \mathbf{u}_1, ..., \mathbf{u}_p$ are linear independent. Now, if $p + q = m$ we are done. If $p + q < m$ we continue in the same fashion by seeking a solution of the form $\mathbf{x}_t = \lambda^t(\mathbf{w} + t\mathbf{v} + (1/2)t^2\mathbf{u})$. In this case (2.2.7) implies

$$(A - \lambda I)\mathbf{w} = \lambda\left(\mathbf{v} + \frac{1}{2}\mathbf{u}\right) \tag{2.2.13a}$$

$$(A - \lambda I)\mathbf{v} = \mathbf{u} \tag{2.2.13b}$$

$$(A - \lambda I)\mathbf{u} = 0 \tag{2.2.13c}$$

which again leads to

$$(A - \lambda I)^3\mathbf{w} = 0 \tag{2.2.14}$$

and we proceed in the same way as before. Either we are done or we keep on seeking solutions where cubic terms of t are included. Sooner or later we will obtain the necessary number of linear independent eigenvectors. □

Let us now illustrate the theory presented above through three examples. In Example 2.2.2 we deal with the easiest case where the coefficient matrix A has distinct real eigenvalues. In Example 2.2.3 we consider eigenvalues with multiplicity larger than one, and finally, in Example 2.2.4, we analyse the case where the eigenvalues are complex conjugated.

Example 2.2.2. Let

$$\mathbf{x} = (x_1, x_2)^T, \quad A = \begin{pmatrix} 2 & 1 \\ -3 & 6 \end{pmatrix}$$

and solve $\mathbf{x}_{t+1} = A\mathbf{x}_t$.

Assuming $\mathbf{x} = \lambda^t\mathbf{u}$ the eigenvalue equation (2.2.9) becomes

$$\begin{vmatrix} 2-\lambda & 1 \\ -3 & 6-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^2 - 8\lambda + 15 = 0 \Leftrightarrow \lambda_1 = 5\,, \quad \lambda_2 = 3$$

The eigenvector $\mathbf{u}_1 = (u_1, u_2)^T$ belonging to $\lambda_1 = 5$ satisfies (cf. (2.2.8))

$$\begin{pmatrix} 2-5 & 1 \\ -3 & 6-5 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence, we choose $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

In the same way, the eigenvector $\mathbf{u}_2 = (u_1, u_2)^T$ belonging to $\lambda_2 = 3$ satisfies

$$\begin{pmatrix} -1 & 1 \\ -3 & 3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Thus $\mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Therefore, according to (2.2.10), the general solution is

$$\mathbf{x}_t = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t = C_1 5^t \begin{pmatrix} 1 \\ 3 \end{pmatrix} + C_2 3^t \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

□

Example 2.2.3. Let

$$\mathbf{x} = (x_1, x_2, x_3)^T, \quad A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

and solve $\mathbf{x}_{t+1} = A\mathbf{x}_t$.

Assuming $\mathbf{x}_t = \lambda^t \mathbf{u}$, we arrive at the eigenvalue equation

$$\begin{vmatrix} 1-\lambda & 1 & 1 \\ 0 & 1-\lambda & 2 \\ 0 & 0 & 1-\lambda \end{vmatrix} = 0 \Leftrightarrow (1-\lambda)^3 = 0$$

so we conclude that $\lambda = 1$ is the only eigenvalue and that it has multiplicity 3. Therefore, according to (B) the general solution of the problem is

$$\mathbf{x}_t = C_1 \lambda^t \mathbf{u} + C_2 \lambda^t (\mathbf{v} + t\mathbf{u}) + C_3 \lambda^t \left(\mathbf{w} + t\mathbf{v} + \frac{1}{2} t^2 \mathbf{u} \right)$$

where $\lambda = 1$ and $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ must be found from (2.2.13a,b,c). Let $\mathbf{u} = (u_1, u_2, u_3)^T$, $\mathbf{v} = (v_1, v_2, v_3)^T$ and $\mathbf{w} = (w_1, w_2, w_3)^T$. (2.2.13c) implies

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} v_2 + v_3 = 0 \\ 2v_3 = 0 \end{matrix}$$

so $u_3 = 0 \Rightarrow u_2 = 0$ and u_1 is arbitrary so let $u_1 = 1$. Therefore $\mathbf{u} = (1, 0, 0)^T$. (2.2.13b) implies

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} v_2 + v_3 = 1 \\ 2v_3 = 0 \end{matrix}$$

thus, $v_3 = 0$, $v_2 = 1$ and v_1 may be chosen arbitrary so we let $v_1 = 0$. This yields $\mathbf{v} = (0,1,0)^T$.

Finally, from (2.2.13a):

$$\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = 1\left(\mathbf{v} + \frac{1}{2}\mathbf{u}\right) = \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix} \Leftrightarrow \begin{array}{c} w_2 + w_3 = \frac{1}{2} \\ 2w_3 = 1 \end{array}$$

Hence, $w_3 = 1/2$, $w_2 = 0$ and we may choose $w_1 = 0$ so $\mathbf{w} = (0,0,1/2)^T$. Consequently, the general solution may be written as

$$\mathbf{x}_t = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}_t = C_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + C_2\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)$$
$$+ C_3\left(\begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \end{pmatrix} + t\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \frac{1}{2}t^2\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right)$$

□

Example 2.2.4. Let

$$\mathbf{x} = (x_1, x_2)^T, \quad A = \begin{pmatrix} -2 & 1 \\ -1 & -2 \end{pmatrix}$$

and solve $\mathbf{x}_{t+1} = A\mathbf{x}_t$.

Suppose $\mathbf{x}_t = \lambda^t \mathbf{v}$. (2.2.9) implies

$$\begin{vmatrix} -2-\lambda & 1 \\ -1 & -2-\lambda \end{vmatrix} = 0 \Leftrightarrow \lambda^2 + 4\lambda + 5 = 0$$

$\Leftrightarrow \lambda_1 = -2 + i$, $\lambda_2 = -2 - i$ (distinct complex eigenvalues).

Further: $|\lambda_1| = \sqrt{(-2)^2 + 1^2} = \sqrt{5}$ $\cos\theta = (-2)/\sqrt{5}$ $\sin\theta = 1/\sqrt{5}$ so $\lambda_1 = \sqrt{5}(\cos\theta + i\sin\theta)$.

The eigenvector $\mathbf{u} = (u_1, u_2)^T$ corresponding to λ_1 may be found from

$$\begin{pmatrix} -2-(-2+i) & 1 \\ -1 & -2-(-2+i) \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{array}{c} -iu_1 + u_2 = 0 \\ -u_1 - iu_2 = 0 \end{array}$$

Let $u_2 = t$, $u_1 = -it$ so $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = t\begin{pmatrix} -i \\ 1 \end{pmatrix}$, so we choose $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} -i \\ 1 \end{pmatrix}$ as eigenvector. Therefore (by use of Moivre's formulae), the solution in complex form becomes

$$\mathbf{x}_t = \sqrt{5}^t(\cos\theta t + i\sin\theta t)\begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} \sqrt{5}^t\{-i\cos\theta t + \sin\theta t\} \\ \sqrt{5}^t\{\cos\theta t + i\sin\theta t\} \end{pmatrix}$$

Two linear independent real solutions are found by taking the real and imaginary parts respectively:

$$\text{Real part} \begin{pmatrix} x_{1r} \\ x_{2r} \end{pmatrix}_t = \sqrt{5}^t\begin{pmatrix} \sin\theta t \\ \cos\theta t \end{pmatrix}$$

$$\text{Imaginary part} \begin{pmatrix} x_{1i} \\ x_{2i} \end{pmatrix}_t = \sqrt{5}^t\begin{pmatrix} -\cos\theta t \\ \sin\theta t \end{pmatrix}$$

Thus, the general solution may be written as

$$\mathbf{x}_t = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_t = C_1\begin{pmatrix} x_{1r} \\ x_{2r} \end{pmatrix}_t + C_2\begin{pmatrix} x_{1i} \\ x_{2i} \end{pmatrix}_t = \begin{pmatrix} \sqrt{5}^t\{C_1\sin\theta t - C_2\cos\theta t\} \\ \sqrt{5}^t\{C_1\cos\theta t + C_2\sin\theta t\} \end{pmatrix}$$

□

Exercise 2.2.3. Let $\mathbf{x} = (x_1, x_2)^T$, $A = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}$ and find the general solution of

a) $\mathbf{x}_{t+1} = A\mathbf{x}_t$,

b) $\mathbf{x}_{t+1} = B\mathbf{x}_t$,

c) Let $\mathbf{x} = (x_1, x_2, x_3)^T$ and find the general solution of $\mathbf{x}_{t+1} = C\mathbf{x}_t$ where

$$C = \begin{pmatrix} -3 & 1 & -1 \\ -7 & 5 & -1 \\ -6 & 6 & -2 \end{pmatrix}$$

□

We close this section by a definition and an important theorem about stability of linear systems.

Definition 2.2.1. The linear system (2.2.7) is globally asymptotic stable if $\lim_{t\to\infty} \mathbf{x}_t = 0$. □

Theorem 2.2.1. The linear system (2.2.7) is globally asymptotic stable if and only if all the eigenvalues λ of A are located inside the unit circle $|z| = 1$ in the complex plane.

Proof: In case of distinct eigenvalues the result follows immediately from (2.2.10).

Eigenvalues with multiplicity m lead according to our previous discussion to terms in the solution of form $t^q\lambda^t$ where $q \leq m - 1$.

Now, if $|\lambda| < 1$, let $|\lambda| = 1/s$ where $s > 1$. Then by L'Hopital's rule: $\lim_{t\to\infty}(t^q/s^t) = 0$ so the result follows here too. □

2.3 The Leslie matrix

In Part I of this book we illustrated many aspects of the theory which we established by use of the quadratic map. Here in Part II we will use Leslie matrix models in a similar fashion.

Leslie matrix models are age-structured population models. They were independently developed in the 1940's by Bernardelli (1941), Lewis (1942) and Leslie (1945, 1948) but were not widely adopted by human demographers until the late 1960's and by ecologists until the 1970's. The ultimate book on matrix population models which we refer to is "Matrix population models" by Hal Caswell (2001). Here we will deal with only a limited number of aspects of these models.

Let $\mathbf{x}_t = (x_{0,t}, ..., x_{n,t})^T$ be a population with $n+1$ nonoverlapping age classes at time t. $x = x_0 + \cdots + x_n$ is the total population.

Next, introduce the Leslie matrix

$$A = \begin{pmatrix} f_0 & f_1 & \cdots & f_n \\ p_0 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & \\ 0 & \cdots & p_{n-1} & 0 \end{pmatrix} \tag{2.3.1}$$

The meaning of the entries in (2.3.1) is as follows: f_i is the average fecundity (the average number of daughters born per female) of a member located in the i'th age class. p_i may be interpreted as the survival probability from age class i to age class $i+1$ and clearly $0 \leq p_i \leq 1$. The relation between $\mathbf{x}$ at two consecutive time steps (years) may then be expressed as

$$\mathbf{x}_{t+1} = A\mathbf{x}_t \tag{2.3.2}$$

or in map notation

$$f : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \qquad \mathbf{x} \to A\mathbf{x} \tag{2.3.3}$$

Depending on the species under consideration, nonlinearities may show up on different entries in the matrix. For example, in fishery models it is often assumed that density effects occur mainly through the first year of life so one may assume $f_i = f_i(x)$. It is also customary to write $f(x)$ as a product of a density independent part F and a density dependent part $\hat{f}(x)$ so $f(x) = F\hat{f}(x)$. Frequently used fecundity functions are:

$$f(x) = F\,e^{-\alpha x} \tag{2.3.4}$$

which is often referred to as the overcompensatory Ricker relation and

$$f(x) = \frac{F}{1+\alpha x} \tag{2.3.5}$$

the compensatory Beverton and Holt relation.

Instead of assuming $f = f(x)$ one may alternatively suppose $f = f(y)$ where $y = \alpha_0 x_0 + \cdots + \alpha_n x_n$ is the weighted sum of the age classes. If only one age class, say x_i, contributes to density effects one writes $f = f(x_i)$. In the case where an age class x_i is not fertile we simply write $F_i = 0$. (Species where most age classes are fertile are called iteroparous. Species where fecundity is restricted to the last age class only are called semelparous.)

The survival probabilities may of course also be density dependent so in such cases we adopt the same strategy as in the fecundity case and write $p(\cdot) = P\hat{p}(\cdot)$ where P is a constant.

A final but important comment is that one in most biological relevant situations supposes $p'(\cdot) \leq 0$ and $f'(\cdot) \leq 0$. The standard counter example is when the Allé effect is modelled. Then one may use $f'(x) \geq 0$ and/or $p'(x) \geq 0$ in case of small populations x. (Allé effects will not be considered here.)

In the subsequent sections we shall analyse nonlinear maps and as already mentioned the theory will be illustrated by use of (2.3.2), (2.3.3). However, if both $f_i = F_i$ and $p_i = P_i$ the Leslie matrix is linear and we let

$$M = \begin{pmatrix} F_0 & \cdots & & & F_n \\ P_0 & 0 & \cdots & & 0 \\ 0 & & & & \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & P_{n-1} & 0 \end{pmatrix} \tag{2.3.6}$$

We close this section by a study of the linear case

$$f : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}, \qquad \mathbf{x} \to M\mathbf{x} \tag{2.3.7}$$

The eigenvalues of M may be obtained from $|M - \lambda I| = 0$.

Exercise 2.3.1.

a) Assume that M is a 2×2 matrix. Show that $|M - \lambda I| = 0 \Leftrightarrow \lambda^2 - F_0\lambda - P_0F_1 = 0$.

b) Suppose that M is 3×3 and show that the eigenvalue equation becomes

$$\lambda^3 - F_0\lambda^2 - P_0F_1\lambda - P_0P_1F_2 = 0$$

c) Generalize and show that if M is a $(n+1) \times (n+1)$ matrix then the eigenvalue equation may be written

$$\lambda^{n+1} - F_0\lambda^n - P_0F_1\lambda^{n-1} - \cdots - P_0P_1 \cdots P_{n-1}F_n = 0 \tag{2.3.8}$$

□

Next, we need some definitions:

Definition 2.3.1. A matrix A is nonnegative if all its elements are greater or equal to zero. It is positive if all elements are positive.

Clearly, the Leslie matrix is nonnegative. □

Definition 2.3.2. A nonnegative matrix A and its associated life cycle graph is irreducible if its life cycle graph is strongly connected (i.e. if between every pair of distinct nodes Ni, N_j in the graph there is a directed path of finite length that begins at N_i and ends at N_j). □

Definition 2.3.3. A reducible life cycle graph contains at least one age group that cannot contribute by any developmental path to some other age group. □

Examples of two irreducible Leslie matrices and one reducible one with associated life cycle graphs are given in Figure 9.

Definition 2.3.4. An irreducible matrix A is said to be primitive if it becomes positive when raised to sufficiently high powers. Otherwise A is imprimitive (cyclic) with index of imprimity equal to the greatest common divisor of the loop lengths in the life cycle graph. □

Exercise 2.3.2. Show by direct calculation that the first irreducible Leslie matrix in Figure 9 is primitive and that the second one is imprimitive (cyclic) with index of imprimity equal to 3. □

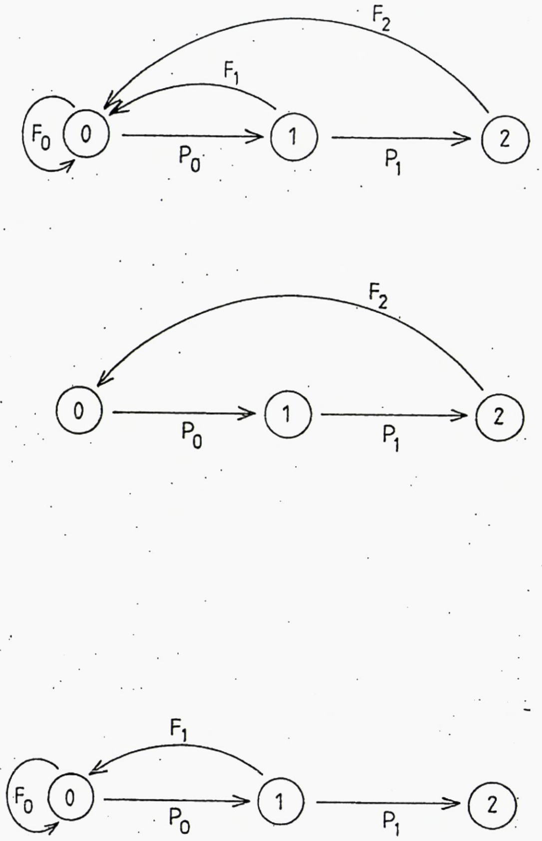


Figure 9: Two irreducible and one reducible matrices with corresponding life cycle graphs.

Regarding nonnegative matrices the main results may be summarized in the following theorem which is often referred to as the Perron–Frobenius theorem.

Theorem 2.3.1 (Perron–Frobenius).

1) If A is positive or nonnegative and primitive, then there exists a real eigenvalue $\lambda_0 > 0$ which is a simple root of the characteristic equation $|A - \lambda I| = 0$. Moreover, the eigenvalue is strictly greater than the magnitude of any other eigenvalue, $\lambda_0 > |\lambda_i|$ for $i \neq 0$. The eigenvector $\mathbf{u}_0$ corresponding to λ_0 is real and strictly positive. λ_0 may not be the only positive eigenvalue but if there are others they do not have nonnegative eigenvectors.

2) If A is irreducible but imprimitive (cyclic) with index of imprivity $d+1$ there exists a real eigenvalue $\lambda_0 > 0$ which is a simple root of $|A-\lambda I| = 0$ with associated eigenvector $\mathbf{u}_0 > \mathbf{0}$. The eigenvalues λ_i satisfy $\lambda_0 \geq |\lambda_i|$ for $i \neq 0$ but there are d complex eigenvalues equal in magnitude to λ_0 whose values are $\lambda_0 \exp(2k\pi i/(d+1))$, $k = 1, 2, ..., d$.

For a general proof of Theorem 2.3.1 we refer to the literature. See for example Horn and Johnson (1985).

Concerning the Leslie matrix M (2.3.6) we are particularly interested in two cases: (I) the case where all fecundities $F_i > 0$, and (II) the semelparous case where $F_i = 0$, $i = 0, ..., n-1$ but $F_n > 0$. In both cases it is assumed that $0 < P_i \leq 1$ for all i.

Let us prove Theorem 2.3.1 assuming (I):

Since $F_n > 0$ and $0 < P_i \leq 1$ it follows directly from (2.3.8) that $\lambda = 0$ is impossible. Therefore, we may divide (2.3.8) by λ^{n+1} to obtain

$$f(\lambda) = \frac{F_0}{\lambda} + \frac{P_0 F_1}{\lambda^2} + \cdots + \frac{P_0 P_1 \cdots P_{n-1} F_n}{\lambda^{n+1}} = 0 \qquad (2.3.9)$$

Clearly, $\lim_{\lambda \to 0} f(\lambda) = \infty$, $\lim_{\lambda \to \infty} f(\lambda) = 0$, and since $f'(\lambda) < 0$ for $\lambda > 0$ it follows that there exists a unique positive λ_0 which satisfies $f(\lambda_0) = 1$. Therefore, assume $\lambda_0^{-1} = e^{\gamma}$ and rewrite (2.3.9) as

$$f(\lambda) = F_0 e^{\gamma} + P_0 F_1 e^{2\gamma} + \cdots + P_0 P_1 \cdots P_{n-1} F_n e^{n\gamma} = 1 \qquad (2.3.10)$$

Next, let $\lambda_j^{-1} = \exp(\alpha + \beta i) = e^{\alpha}(\cos\beta + i \sin\beta)$ for $j = 1, ..., n$ and assume β real and positive and $\beta \neq 2k\pi$, $k = 1, 2, \ldots$. Then $\lambda_j^{-p} = e^{\alpha p}(\cos p\beta + i \sin p\beta)$ which inserted into $f(\lambda)$, considering the real part only, gives

$$F_0 e^{\alpha} \cos\beta + P_0 F_1 e^{2\alpha} \cos 2\beta + \cdots + P_0 P_1 \cdots P_{n-1} F_n e^{n\alpha} \cos n\beta = 1 \quad (2.3.11)$$

Now, since there are at least two consecutive positive fecundity values F_j and F_{j+1} it follows that $\cos j\beta$ and $\cos(j+1)\beta$ cannot both be equal to unity since $\beta \neq 2k\pi$. Consequently, by comparing (2.3.10) and (2.3.11), we have $e^{\alpha} > e^{\gamma} \Leftrightarrow |\lambda_j| < \lambda_0$ for $j = 1, ..., n$.

Finally, in order to see that the eigenvector $\mathbf{u}_0$ corresponding to λ_0 has only positive elements, recall that $\mathbf{u}_0$ must be computed from $M\mathbf{u}_0 = \lambda_0 \mathbf{u}_0$, and in order to avoid $\mathbf{u}_0 = \mathbf{0}$ we must choose one of the components of $\mathbf{u}_0 = (u_{00}, ..., u_{n0})^T$ free, so let $u_{00} = 1$. Then from $M\mathbf{u}_0 = \lambda_0 \mathbf{u}_0$: $P_0 \cdot 1 = \lambda_0 u_{10}$, $P_1 u_{10} = \lambda_0 u_{20}, ..., P_{n-1} u_{n-10} = \lambda_0 u_{n0}$ which implies

$$u_{10} = \frac{P_0}{\lambda_0}, \quad u_{20} = \frac{P_1 u_{10}}{\lambda_0} = \frac{P_0 P_1}{\lambda_0^2} \cdots u_{n0} = \frac{P_0 \cdots P_{n-1}}{\lambda_0^n}$$

which proves what we want. □

(This proof is based upon Frauenthal (1986).) The proof of Theorem 2.3.1 under the assumption (II) is left to the reader.

—

Let us now turn to the asymptotic behaviour of the linear map (2.3.7) in light of the results of Theorem 2.3.1.

In the case where all $F_i > 0$ we may express the solution of (2.3.7) (cf. (2.2.10)) as

$$\mathbf{x}_t = c_o \lambda_0^t \mathbf{u}_0 + c_1 \lambda_1^t \mathbf{u}_1 + \cdots + c_n \lambda_n^t \mathbf{u}_n \tag{2.3.12}$$

where λ_i (real or complex, λ_0 real) are the eigenvalues of M numbered in order of decreasing magnitude and $\mathbf{u}_i$ are the corresponding eigenvectors. Further,

$$\frac{\mathbf{x}_t}{\lambda_0^t} = c_0 \mathbf{u}_0 + c_1 \left(\frac{\lambda_1}{\lambda_0}\right)^t \mathbf{u}_1 + \cdots + c_n \left(\frac{\lambda_n}{\lambda_0}\right)^t \mathbf{u}_n$$

and since $\lambda_0 > |\lambda_i|$, $i \neq 0$

$$\lim_{t \to \infty} \frac{\mathbf{x}_t}{\lambda_0^t} = c_0 \mathbf{u}_0 \tag{2.3.13}$$

Consequently, if M is nonnegative and primitive, the long term dynamics of the population are described by the growth rate λ_0 and the stable population structure $\mathbf{u}_0$. Thus $\lambda_0 > 1$ implies an exponential increasing population, $0 < \lambda_0 < 1$ an exponential decreasing population, where we in all cases have the stable age distribution $\mathbf{u}_0$.

If M is irreducible but imprimitive with index of imprimity $d+1$ it follows from part 2 of the Perron–Frobenius theorem that the limit (2.3.13) may be expressed as

$$\lim_{t\to\infty} \frac{\mathbf{x}_t}{\lambda_0^t} = c_0\mathbf{u}_0 + \sum_{k=1}^{d} c_k e^{(2k\pi/(d+1))it}\mathbf{u}_i \tag{2.3.14}$$

As opposed to the dynamical consequences of 1) in the Perron–Frobenius theorem we now conclude from (2.3.14) that $\mathbf{u}_0$ is not stable in the sense that an initial population not proportional to $\mathbf{u}_0$ will converge to it. Instead, the limit (2.3.14) is periodic with period $d+1$.

Example 2.3.1 (Bernardelli 1941). The first paper where the matrix M was considered came in 1941. There, Bernardelli considered a hypothetic beetle population obeying the equation

$$\mathbf{x}_{t+1} = B\mathbf{x}_t \qquad \text{where} \qquad B = \begin{pmatrix} 0 & 0 & 6 \\ 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{pmatrix}$$

Clearly, B is irreducible and imprimitive with index of imprimity equal to 3 (cf. Exercise 2.3.2). Moreover, the eigenvalues of B are easily found to be $\lambda_1 = 1$ and $\lambda_{2,3} = \exp(\pm 2\pi i/3)$ and it is straightforward to show that $B^3 = I$ so each initial age distribution will repeat itself in a regular manner every third year as predicted by (2.3.14). In Figure 10 we show the total hypothetic beetle population together with the three age classes as function of time, and clearly there is no stable age distribution. □

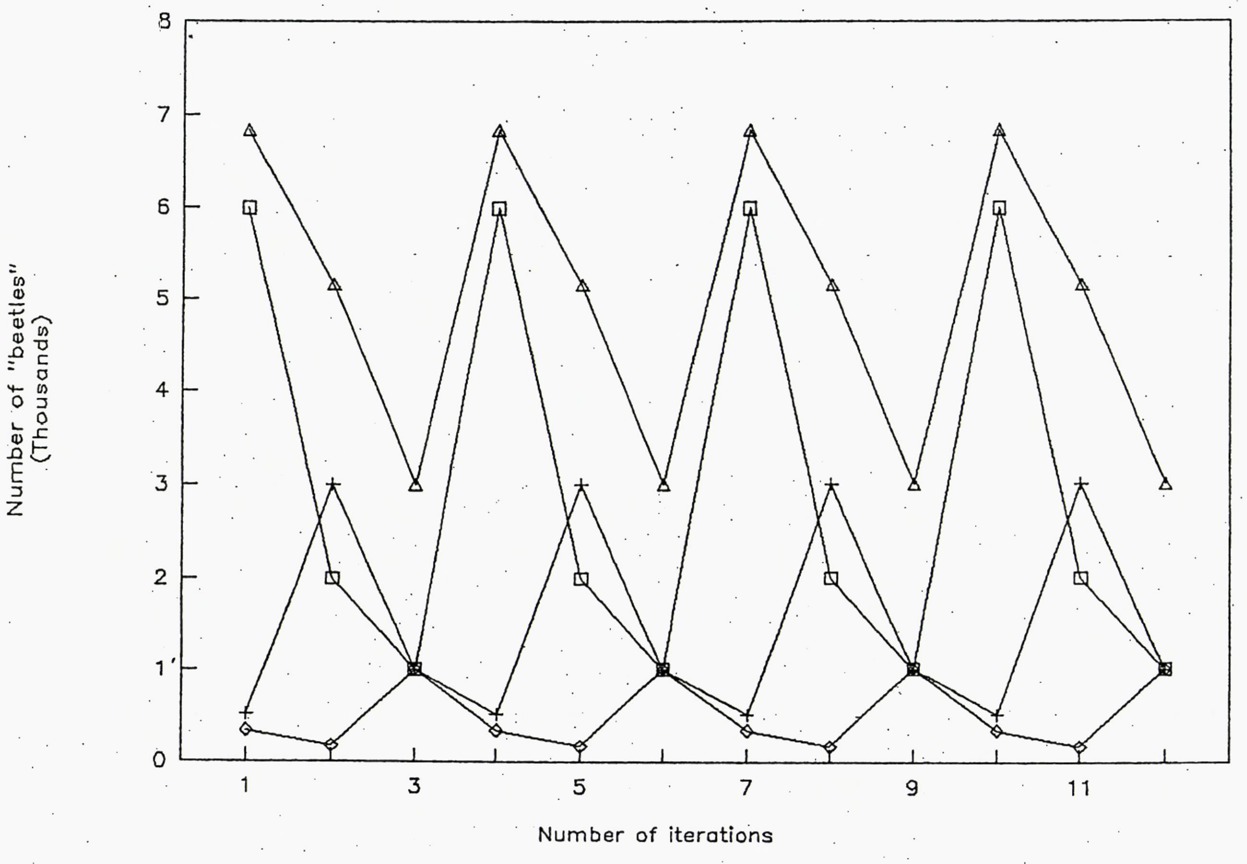


Figure 10: The hypothetical "beetle" population of Bernardelli as function of time. △ is the total population. □, + and ◇ correspond to the zeroth, first and second age classes respectively. Clearly, there is no stable age distribution.

2.4 Fixed points and stability of nonlinear systems

In this section we turn to the nonlinear case $\mathbf{x} \rightarrow f(\mathbf{x})$ which in difference equation notation may be cast in the form

$$\begin{aligned} x_{1,t+1} &= f_1(x_{1,t}, ..., x_{n,t}) \\ &\vdots \\ x_{n,t+1} &= f_n(x_{1,t}, ..., x_{n,t}) \end{aligned} \tag{2.4.1}$$

Definition 2.4.1. A point $\mathbf{x}^* = (x_1^*, ..., x_n^*)$ which satisfies $\mathbf{x}^* = f(\mathbf{x}^*)$ is called a fixed point for (2.4.1). □

Exercise 2.4.1. Assume that $F_0 + P_0F_1 > 1$, $x = x_0 + x_1$ and find the nontrivial fixed point (x_0^*, x_1^*) of the two-dimensional Leslie matrix model (the Ricker model)

$$\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \longrightarrow \begin{pmatrix} F_0e^{-x} & F_1e^{-x} \\ P_0 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \tag{2.4.2}$$

According to Definition 2.4.1 the fixed point satisfies

$$x_0^* = F_0e^{-x^*}x_0^* + F_1e^{-x^*}x_1^* \tag{2.4.3a}$$

$$x_1^* = P_0x_0^* \tag{2.4.3b}$$

and if we insert (2.4.3b) into (2.4.3a) we obtain $1 = e^{-x^*}(F_0 + P_0F_1)$, hence the total equilibrium population becomes $x^* = \ln(F_0 + P_0F_1)$. Further, since $x^* = x_0^* + x_1^*$ and $x_1^* = P_0x_0^*$ we easily find

$$(x_0^*, x_1^*) = \left(\frac{1}{1+P_0}x^*, \frac{P_0}{1+P_0}x^*\right) \tag{2.4.4}$$

(Note that $F_0+P_0F_1 > 1$ is necessary in order to obtain a biological acceptable solution.) □

Exercise 2.4.1. Still assuming $F_0 + P_0F_1 > 1$, show that the fixed point (x_0^*, x_1^*) of the two-dimensional Beverton and Holt model

$$\begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{F_0}{1+x} & \frac{F_1}{1+x} \\ P_0 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \end{pmatrix} \tag{2.4.5}$$

becomes

$$(x_0^*, x_1^*) = \left(\frac{1}{1+P_0}x^*, \frac{P_0}{1+P_0}x^*\right) \tag{2.4.6}$$

where $x^* = F_0 + P_0F_1 - 1$. □

Example 2.4.2. Find the nontrivial fixed point of the general Ricker model:

$$\begin{pmatrix} x_0 \\ \vdots \\ x_1 \end{pmatrix} \longrightarrow \begin{pmatrix} F_0e^{-x} & \cdots & & & F_ne^{-x} \\ P_0 & 0 & \cdots & & 0 \\ \vdots & & & & \vdots \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & 0 & P_{n-1} & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ \vdots \\ x_n \end{pmatrix} \tag{2.4.7}$$

The fixed point $\mathbf{x}^* = (x_0^*, ..., x_n^*)$ obeys

$$x_0^* = e^{-x^*}(F_0 x_0^* + \cdots + F_n x_n^*)$$
$$x_1^* = P_0 x_0^*$$
$$\vdots$$
$$x_n^* = P_{n-1} x_{n-1}^*$$

From the last n equations we have $x_1^* = P_0 x_0^*$, $x_2^* = P_1 x_1^* = P_0 P_1 x_0^*$, $x_n^* = P_0 \cdots P_{n-1} x_0^*$ which inserted into the first equation give

$$1 = e^{-x^*}(F_0 + P_0 F_1 + P_0 P_1 F_2 + \cdots + P_0 \cdots P_{n-1} F_n) \tag{2.4.8}$$

Hence,

$$x^* = \ln(F_0 + P_0 F_1 + \cdots + P_0 \cdots P_{n-1} F_n) = \ln\left(\sum_{i=0}^{n} F_i L_i\right)$$

where $L_i = P_0 P_1 \cdots P_{i-1}$ and by convention $L_0 = 1$. From $\sum x_i^* = x^*$ and $x_1^* = P_0 x_0^* = L_1 x_0^*$, $x_2^* = P_0 P_1 x_0^* = L_2 x_0^*$ and $x_i^* = L_i x_0^*$ we obtain

$$(x_0^*, ..., x_n^*) = \left(\frac{L_1}{\sum_{i=0}^n L_i} x^*, \cdots \frac{L_i}{\sum_{i=0}^n L_i} x^*, \cdots \frac{L_n}{\sum_{i=0}^n L_i} x^*\right) \tag{2.4.9}$$

Again, $\sum_{i=0}^n F_i L_i > 1$ is required in order to have an acceptable biological equilibrium. □

Exercise 2.4.2. Generalize Exercise 2.4.1 in the same way as in Example 2.4.2 and obtain a formulae for the fixed point of the $n+1$ dimensional Beverton and Holt model. □

In order to reveal the stability properties of the fixed point $\mathbf{x}^*$ of (2.4.1) we follow the same pattern as we did in Section 1.4. Let $\mathbf{x} = \mathbf{x}^* + \boldsymbol{\xi}$, then expand $f_i(\mathbf{x})$ in its Taylor series about x^*, taking the linear terms only in order to obtain

$$x_{1,t+1}^* + \xi_{1,t+1} \approx f_i(\mathbf{x}_t^*) + \frac{\partial f_1}{\partial x_1}\xi_{1,t} + \cdots + \frac{\partial f_1}{\partial x_n}\xi_{n,t}$$
$$\vdots$$
$$x_{n,t+1}^* + \xi_{n,t+1} \approx f_n(\mathbf{x}_t^*) + \frac{\partial f_n}{\partial x_1}\xi_{1,t} + \cdots + \frac{\partial f_n}{\partial x_n}\xi_{n,t}$$

where all derivatives are evaluated at $\mathbf{x}^*$. Moreover, $x^*_{i,t+1} = f_i(\mathbf{x}^*_t)$. Consequently, the linearized map (or linearization) of (2.4.1) becomes

$$\begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x^*) & \cdots & \frac{\partial f_1}{\partial x_n}(x^*) \\ \vdots & & \\ \frac{\partial f_n}{\partial x_1}(x^*) & \cdots & \frac{\partial f_n}{\partial x_n}(x^*) \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} \tag{2.4.10}$$

where the matrix is called the Jacobian.

If the fixed point $\mathbf{x}^*$ of (2.4.1) shall be locally asymptotic stable we clearly must have

$$\lim_{t\to\infty} \boldsymbol{\xi}_t \to 0 \tag{2.4.11}$$

and according to Theorem 2.2.1 this is equivalent to say:

Theorem 2.4.1. The fixed point $\mathbf{x}^*$ of the nonlinear system (2.4.1) is locally asymptotic stable if and only if all the eigenvalues λ of the Jacobian matrix are located inside the unit circle $|z| = 1$ in the complex plane. □

Example 2.4.3.

a) Define $\hat{F}\hat{x} = F_0 x_0^* + F_1 x_1^*$ and show that the fixed point (2.4.4) of the Ricker map (2.4.2) is locally asymptotic stable provided

$$\hat{F}\hat{x}(1 + P_0) > 0 \tag{2.4.12a}$$

$$2F_0 + \hat{F}\hat{x}(P_0 - 1) > 0 \tag{2.4.12b}$$

$$2P_0F_1 + F_0 - P_0\hat{F}\hat{x} > 0 \tag{2.4.12c}$$

b) Assume that $F_0 = F_1 = F$ (same fecundity in both age classes) and show that (2.4.12b), (2.4.12c) may be expressed as

$$F < \frac{1}{1 + P_0} e^{2/(1-P_0)} \tag{2.4.13b}$$

$$F < \frac{1}{1 + P_0} e^{(1+2P_0)/P_0} \tag{2.4.13c}$$

Solution:

a) Rewrite (2.4.2) as

$$x_0 \to f_1(x_0, x_1) = F_0 e^{-x} x_0 + F_1 e^{-x} x_1$$

$$x_1 \to f_2(x_0, x_1) = P_0 x_0$$

Then the Jacobian becomes

$$J = \begin{pmatrix} e^{-x^*}(F_0 - \hat{F}\hat{x}) & e^{-x^*}(F_1 - \hat{F}\hat{x}) \\ P_0 & 0 \end{pmatrix} \tag{2.4.14}$$

and the eigenvalue equation $|J - \lambda I| = 0$ may be cast in the form

$$\lambda^2 - \frac{F_0 - \hat{F}\hat{x}}{F_0 + P_0 F_1}\lambda - P_0 \frac{F_1 - \hat{F}\hat{x}}{F_0 + P_0 F_1} = 0 \tag{2.4.15}$$

where we have used $e^{-x^*} = (F_0 + P_0 F_1)^{-1}$.

(2.4.15) is a second order polynomial and $|\lambda| < 1$ if the corresponding Jury criteria (2.1.14) are satisfied. Therefore, by defining

$$a_1 = -\frac{F_0 - \hat{F}\hat{x}}{F_0 + P_0 F_1} \qquad a_2 = -P_0 \frac{F_1 - \hat{F}\hat{x}}{F_0 + P_0 F_1}$$

we easily obtain from (2.1.14) that the fixed point is locally asymptotic stable provided the inequalities (2.4.12a)–(2.4.12c) hold.

Remark: Scrutinizing the criteria, it is obvious that (2.4.12a) holds for any (positive) equilibrium population x^*. It is also clear that in case of $\hat{F}\hat{x}$ sufficiently small the same is true for both (2.4.12b,c) as well which allow us to conclude that (x_0^*, x_1^*) is stable in case of "small" equilibrium population x^*. However, if $\hat{F}\hat{x}$ becomes large, both (2.4.12b) and (2.4.12c) contain a large negative term so evidently there are regions in parameter space where (2.4.12b) or (2.4.12c) or both are violated and consequently regions where (x_0^*, x_1^*) is no longer stable.

b) If $F_0 = F_1 = F$, then $\hat{F}\hat{x} = Fx^*$, thus (2.4.15) may be expressed as

$$\lambda^2 - \frac{1 - x^*}{1 + P_0}\lambda - P_0 \frac{1 - x^*}{1 + P_0} = 0 \tag{2.4.16}$$

and the criteria (2.4.12b), (2.4.12c) simplify to

$$2 + x^*(P_0 - 1) > 0$$

$$2P_0 + 1 - Px^* > 0$$

(2.4.13b) and (2.4.13c) are now established by use of $x^* = \ln[F(1 + P_0)]$. A final but important observation is that whenever $0 < P_0 < 1/2$, (2.4.13b) will be violated prior to (2.4.13c) if F is increased. On the other hand, if $1/2 < P_0 \leq 1$, (2.4.13c) will be violated first through an increase of F. (As we shall see later, this fact has a crucial impact of the possible dynamics in the unstable parameter region.) □

Example 2.4.4 (Example 2.4.2 continued). Let the fecundities be equal (i.e. $F_0 = \cdots = F_n = F$) in the general $n + 1$ dimensional Ricker model that we considered in Example 2.4.2. Then, $x^* = \ln(FD)$, $D = \sum_{i=0}^{n} L_i$ and the fixed point $\mathbf{x}^*$ may be written as $\mathbf{x}^* = (x_0^*, ..., x_i^*, ..., x_n^*)$ where $x_i^* = (L_i/D)x^*$. The eigenvalue equation (cf. (2.4.16)) may be cast in the form

$$\lambda^{n+1} - \frac{1}{D}(1 - x^*) \sum_{i=0}^{n} L_i \lambda^{n-i} = 0 \qquad (2.4.17)$$

Our goal is to show that the fixed point $\mathbf{x}^*$ is locally asymptotic stable whenever $x^* < 2$ (i.e. that all the eigenvalues λ of (2.4.17) are located inside the unit circle.)

In contrast to Example 2.4.3, Theorem 2.1.8 obviously does not work here so instead we appeal to Theorem 2.1.9 (Rouché's theorem). Therefore, assume $|1 - x^*| < 1$, let $f(\lambda) = \lambda^{n+1}$, $g(\lambda) = -(1/D)(1 - x^*) \sum_{i=0}^{n} L_i \lambda^{n-i}$ and rewirite (2.4.17) as $f(\lambda) + g(\lambda) = 0$. Clearly, f and g are analytic functions on and inside the unit circle C and the equation $f(\lambda) = 0$ has $n + 1$ roots inside C.

On the boundary we have

$$
\begin{aligned}
|g(\lambda)| &= \left| -\frac{1}{D}(1-x^*)\sum_{i=0}^{n} L_i\lambda^{n-i} \right| \\
&\leq \left| \frac{L_0}{D}(1-x^*)\lambda^n \right| + \left| \frac{L_1}{D}(1-x^*)\lambda^{n-1} \right| + \cdots + \left| \frac{L_n}{D}(1-x^*) \right| \\
&\leq |1-x^*| < |f(\lambda)|
\end{aligned}
$$

Thus, according to Theorem 2.1.9, $f(\lambda)+g(\lambda)$ and $f(\lambda)$ have the same number of zeros inside C, hence (2.4.17) has $n+1$ zeros inside the unit circle which proves that $x^* < 2$ is sufficient to guarantee a stable fixed point. □

Exercise 2.4.2 (Exercise 2.4.1 continued).

a) Consider the two-dimensional Beverton and Holt model (see Exercise 2.4.1) and show that the fixed point (x_0^*, x_1^*) is always stable. ($F_0 = F_1 = F$.)

b) Generalize to $n+1$ age classes. ($F_0 = \cdots = F_n = F$.) □

Exercise 2.4.3: Assume $P_0 < 1$ and consider the two-dimensional semelparous Ricker model:

$$
\begin{aligned}
x_{0,t+1} &= F_1 e^{-x_t} x_1 \\
x_{1,t+1} &= P_0 x_0
\end{aligned}
\tag{2.4.18}
$$

a) Compute the nontrivial fixed point (x_0^*, x_1^*).

b) Show that the eigenvalue equation may be written as

$$
\lambda^2 + \frac{x_1^*}{P_0}\lambda - (1 - x_1^*) = 0
$$

and use the Jury criteria to conclude that (x_0^*, x_1^*) is always unstable.

c) Show that

$$
\begin{aligned}
x_{0,t+2} &= (P_0F_1)e^{-x_{t+1}} x_{0,t} \\
x_{1,t+2} &= (P_0F_1)e^{-x_t} x_{1,t}
\end{aligned}
\tag{2.4.19}
$$

d) Assume that there exists a two-cycle where the points in the cycle are on the form $(A, 0)$, $(0, B)$ and show that the cycle is $((1/P_0)\ln(P_0F_1), 0)$, $(0, \ln(P_0F_1))$.

e) Show that the two cycle in d) is stable provided $0 < P_0F_1 < e^2$. □

—

Next, consider the general system (2.4.1) and its linearization (2.4.10) and let λ be the eigenvalues of the Jacobian. We now define the following decompositions of $\mathbb{R}^n$.

Definition 2.4.2.

E^s is the subspace which is generated by the (generalized) eigenvectors whose eigenvalues satisfy $|\lambda| < 1$.

E^c is the subspace which is generated by the (generalized) eigenvectors whose eigenvalues satisfy $|\lambda| = 1$.

E^u is the subspace which is generated by the (generalized) eigenvectors whose corresponding eigenvalues satisfy $|\lambda| > 1$.

$\mathbb{R}^n = E^s \oplus E^c \oplus E^u$ and the subspaces E^s, E^c and E^u are called the stable, the center and the unstable subspace respectively. □

By use of the definition above, the stability result stated in Theorem 2.4.1 may be reformulated as follows:

$$\mathbf{x}^* = (x_0^*, ..., x_n^*) \text{ is locally asymptotic stable if } E^u = \{\mathbf{0}\} \text{ and } E^c = \{\mathbf{0}\}.$$
$$\mathbf{x}^* \text{ is unstable if } E^u \neq \{\mathbf{0}\}.$$

$\mathbf{x}^* = (x_0^*, ..., x_n^*)$ is called a hyperbolic fixed point if $E^c = \{\mathbf{0}\}$ (cf. Section 1.4). ($\mathbf{x}^*$ is attracting if $|\lambda| < 1$, repelling if $|\lambda| > 1$.)

We close this section by stating two general theorems which link the nonlinear behaviour close to a fixed point to the linear behaviour.

Theorem 2.4.2 (Hartman–Grobman). Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a C' diffeomorphism where $\mathbf{x}^*$ is a hyperbolic fixed point and let Df be the linearization. Then there exists a homeomorphism h defined in a domain U about $\mathbf{x}^*$ such that

$$(h \circ f)(\xi) = Df(x^*) \circ h(\xi) \tag{2.4.20}$$

for $\xi \in U$. □

Theorem 2.4.3. There exists a stable manifold $W^s_{\text{loc}}(\mathbf{x}^*)$ and an unstable manifold $W^u_{\text{loc}}(\mathbf{x}^*)$ which are a) invariant, and b) is tangent to E^s and E^u at $\mathbf{x}^*$ and have the same dimension as E^s and E^u. □

2.5 The Hopf bifurcation

There are three ways in which the fixed point $\mathbf{x}^* = (x_0^*, ..., x_1^*)$ of a nonlinear map, $f_\mu : \mathbb{R}^n \to \mathbb{R}^n$ may fail to be hyperbolic. One way is that an eigenvalue λ of the linearization crosses the unit circle (sphere) through 1. Then a saddle-node bifurcation occurs. Another possibility is that λ crosses the unit circle at -1 which in turn leads to a flip bifurcation. (In section 1.5 we analysed the generic properties of both the saddle node and the flip.) The third possibility is that a pair of complex eigenvalues $\lambda, \overline{\lambda}$ cross the unit circle. In this case the fixed point will undergo a Hopf bifurcation which we will now describe. Note that the saddle-node and the flip bifurcations may occur in one-dimensional maps, $f_\mu : \mathbb{R} \to \mathbb{R}$. The Hopf bifurcation may take place when the dimension n of the map is equal or larger than two. In this section we will restrict the analysis to the case $n = 2$ only. Later on in section 2.6 we will show how both the flip and the Hopf bifurcation may be analysed in case of $n > 2$.

Theorem 2.5.1 (Hopf). Let $f_\mu : \mathbb{R}^2 \to \mathbb{R}^2$ be a two-dimensional one-parameter family of maps whose fixed point is $\mathbf{x}^* = (x_0^*, x_1^*)$. Moreover, assume that the eigenvalues $\lambda(\mu)$, $\overline{\lambda}(\mu)$ of the linearization are complex conjugates. Suppose that

$$|\lambda(\mu_0)| = 1 \qquad \text{but } \lambda^i(\mu_0) \neq 1 \text{ for } i = 1, 2, 3, 4 \tag{2.5.1}$$

and

$$\frac{d|\lambda(\mu_0)|}{d\mu} = d \neq 0 \tag{2.5.2}$$

Then, there exists a series of near identity transformations h such that $hf_\mu h^{-1}$ in polar coordinates may be written as

$$hf_\mu h^{-1}(r, \varphi) = ((1 + d\mu)r + ar^3, \ \varphi + c + br^2) + \text{ higher order terms} \tag{2.5.3}$$

Moreover, if $a \neq 0$ there is an $\varepsilon > 0$ and a closed curve ξ_μ of the form $r = r_\mu(\varphi)$ for $0 < \mu < \varepsilon$ which is invariant under f_μ. □

Before we prove the theorem let us give a few remarks.

Remark 1. Performing near identity transformations as stated in the theorem is also called normal form calculations. Hence, formulae (2.5.3) is nothing but the original map written in normal form.

Remark 2. If $d > 0$ (cf. (2.5.2)) then the complex conjugated eigenvalues cross the unit circle outwards which of course means that (x_0^*, x_1^*) loses its stability at bifurcation threshold $\mu = \mu_0$. If $d < 0$ the eigenvalues move inside the unit circle. □

Remark 3. $\lambda(\mu_0) = 1$ or $\lambda^2(\mu_0) = 1$ (cf. 2.5.1)) correspond to the well known saddle-node or flip bifurcations respectively. $\lambda^3(\mu_0) = 1$ and $\lambda^4(\mu_0) = 1$ are special and are referred to as the strong resonant cases. If λ is third or fourth root of unity there will be additional resonant terms in formulae (2.5.3). □

Remark 4. As is well known, if a saddle node bifurcation occurs at $\mu = \mu_0$ it means that in case of $\mu < \mu_0$ there are no fixed points but when μ passes through μ_0 two branches of fixed points are born, one branch of stable points, one branch of unstable points.

If the fixed point undergoes a flip bifurcation at $\mu = \mu_0$ we have (in the supercritical case) that the fixed point loses its stability at $\mu = \mu_0$ and that a stable period 2 orbit is created.

Theorem 2.5.1 says that when (x_0^*, x_1^*) undergoes a Hopf bifurcation at $\mu = \mu_0$ a closed invariant curve surrounding (x_0^*, x_1^*) is established whenever $\mu > \mu_0$, $|\mu - \mu_0|$ small. □

Proof of Theorem 2.5.1. Let (x_0^*, x_1^*) be the fixed point of the two-dimensional map $\mathbf{x} \to f(\mathbf{x})$ $(\mathbf{x} = (x_0, x_1)^T)$ and assume that the eigenvalues of the Jacobian $Df(x_0^*, x_1^*)$ are $\lambda, \overline{\lambda} = a_1 + a_2 i$. Next, define the 2×2 matrix T which columns are the real and imaginary parts of the eigenvectors corresponding to the eigenvalues at the bifurcation. Then, after expanding the right-hand side of the map in a Taylor series, applying the change of coordinates $(\hat{x}_0, \hat{x}_1) = (x_0 - x_0^*, x_1 - x_1^*)$ (in order to bring the bifurcation to the origin) together with the transformations

$$\begin{pmatrix} \hat{x}_0 \\ \hat{x}_1 \end{pmatrix} = T \begin{pmatrix} x \\ y \end{pmatrix} \qquad \begin{pmatrix} x \\ y \end{pmatrix} = T^{-1} \begin{pmatrix} \hat{x}_0 \\ \hat{x}_1 \end{pmatrix}$$

our original map may be cast into standard form at the bifurcation as

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \cos 2\pi\theta & -\sin 2\pi\theta \\ \sin 2\pi\theta & \cos 2\pi\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} R_1(x, y)) \\ R_2(x, y) \end{pmatrix} \tag{2.5.4}$$

Our next goal is to simplify the higher order terms R_1 and R_2. This will be done by use of normal form calculations (near identity transformations). The calculations are simplified if they first are complexified. Thus we introduce

$$x' = \cos 2\pi\theta x - \sin 2\pi\theta y + R_1(x, y)$$

$$y' = \sin 2\pi\theta x + \cos 2\pi\theta y + R_2(x, y)$$

$$z = x + yi \qquad z' = x' + y'i \qquad R = R_1 + R_2 i$$

and rewrite (2.5.4) as

$$f : \mathbb{C} \to \mathbb{C}, \qquad z \to f(z) = e^{2\pi\theta i} z + R(z) \tag{2.5.5}$$

where the remainder is on the form

$$R(z) = R^{(k)}(z) + O(|z|^{k+1})$$

Here, $R^{(k)} = r_1^{(k)} z^k + r_2^{(k)} z^{k-1}\overline{z} + \cdots + r_{k+1}^{(k)} \overline{z}^k$.

Next, define

$$z = Z(w) \qquad w = W(z) = Z^{-1}(z) \tag{2.5.6}$$

Then

$$z' = f(z))f(Z(w)) \tag{2.5.7}$$

which in turn implies

$$w' = \hat{f}(w) = Z^{-1}(z') = (Z^{-1} \circ f \circ Z)(w) \tag{2.5.8}$$

Now, we introduce the near identity transformation

$$z = Z(w) = w + P^{(k)}(w) \tag{2.5.9}$$

and claim that

$$w = z - P^{(k)}(z) + O(|z|^{k+1}) = W(z) \tag{2.5.10}$$

This is nothing but a consequence of (2.5.9). Indeed we have

$$\begin{aligned} w &= z - P^{(k)}(w) = z - P^{(k)}(W(z)) \\ &= w + P^{(k)}(w) - P^{(k)}(w + P^{(k)}(w)) \\ &= w + \text{terms of order higher than } k \end{aligned}$$

Thus, we may now by use of the relations

$$\begin{aligned} f(z) &= e^{2\pi\theta i} z + R^{(k)}(z) + \text{h.o.} \\ Z(w) &= w + P^{(k)}(w) \\ Z^{-1}(z') &= z' - P^{(k)}(z') + \text{h.o.} \end{aligned}$$

(where h.o. means higher order) compute $\hat{f}(w)$. This is done in two steps. First,

$$z' = (f \circ Z)(w) = e^{2\pi\theta i}w + e^{2\pi\theta i}P^{(k)}(w) + R^{(k)}(w + ...)$$

Then

$$\begin{aligned}\hat{f}(w) &= (Z^{-1} \circ f \circ Z)(w) = z' - P^{(k)}(z') + \text{h.o.} \qquad (2.5.11)\\ &= e^{2\pi\theta i}w + e^{2\pi\theta i}P^{(k)}(w) + R^{(k)}(w + ...) - P^{(k)}(e^{2\pi\theta i}w) + \text{h.o.}\end{aligned}$$

Next, we want to choose constants in order to remove as many terms in $R^{(k)}(w)$ as possible. To this end let H^k be the space of polynomials of degree k and consider the map

$$K : H^k \to H^k \qquad K(P) = e^{2\pi\theta i}P(w) - P(e^{2\pi\theta i}w) \qquad (2.5.12)$$

Clearly, $w^l\overline{w}^{k-l}$ is a basis for H^k and we have

$$\begin{aligned}K(w^l\overline{w}^{k-l}) &= e^{2\pi\theta i}w^l\overline{w}^{k-l} - e^{2\pi\theta il}w^l\, e^{-2\pi\theta i(k-l)}\overline{w}^{k-l}\\ &= \left[e^{2\pi\theta i} - e^{2\pi\theta i(2l-k)}\right] w^l\overline{w}^{k-l}\\ &= \lambda w^l\overline{w}^{k-l}\end{aligned}$$

where $k = 2, 3, 4, ..., 0 \leq l \leq k$.

From this we conclude that terms in $R^{(k)}(w)$ of the form $w^l\overline{w}^{k-l}$ such that $\lambda(\theta, k, l) = 0$ cannot be removed by near identity transformations. There are two cases to consider: (A) θ irrational, and (B) θ rational.

(A) Assume θ irrational. Then $\lambda = 0 \Leftrightarrow 2l = k + 1$ thus k is an odd number. Here $k = 1$ corresponds to the linear term and the next unremoval terms are proportional to $w^2\overline{w}$ and $w|w|^4$ (i.e. third and fifth order terms).

(B) Supppose $\theta = \mu/r$ rational, $\mu, r \in \mathbb{N}$, μ/r. Then $\lambda = 0 \Leftrightarrow (2l - (k + 1))\mu/r = m$ where $m \in Z$. This implies $(2l - (k + 1))\mu = mr$. Therefore r must be a factor in $(2l - (k + 1))$. Thus the smallest k $(l = 0)$, equals $r - 1$ which means that the first unremoval terms are proportional to $\overline{w}^{r-1}$. When

$r = 2$ the flip occurs. The cases $r = 3, 4$ which corresponds to eigenvalues of third and fourth root of unity repectively are special (cf. Remark 3 after Theorem 2.5.1.)

Now, considering the generic case, θ irrational, we may through normal form calculations remove all terms in $R^{(k)}$ except from those which are proportional to $w^2\overline{w}$ and $w|\overline{w}|^4$, hence (2.5.5) may be cast into normal form as

$$z' = f(z) = e^{2\pi\theta i}z(1 + \alpha\mu + \beta|z|^2) \tag{2.5.13}$$

where α and β are given complex numbers. Now introducing polar coordinates (r, φ), (2.5.13) may be expressed as

$$r' = r(1 + d\mu + ar^2) \tag{2.5.14a}$$

$$\varphi' = \varphi + c + br^2 \tag{2.5.14b}$$

which is nothing but formulae (2.5.3) in the theorem.

Finally, observe that the fixed point r^* of (2.5.14a) is

$$r^* = \sqrt{-\frac{d\mu}{a}} \tag{2.5.15}$$

Considering the case where the eigenvalues $\lambda, \overline{\lambda}$ leave the unit circle at bifurcation threshold as μ is increased (i.e. $d > 0$, $\mu > 0$), $a < 0$ is necessary in order for r^* to exist. The eigenvalue of the linearization of (2.5.14a) is $\sigma = 1 - 2d\mu$, hence r^* is stable whenever $d\mu$ small. (Another way to see that $a < 0$ is necessary for r^* to be an attractor is simply to solve (2.5.14a) by graphical analysis.) Therefore we conclude that $(a < 0, \mu > 0, d > 0)$ the outcome of a supercritical Hopf bifurcation is an invariant attracting circle (curve) as displayed in Figure 11. □

Referring to section 1.5 where we treated the flip bifurcation we stated and proved a theorem (Theorem 1.5.1) where we gave conditions for the flip to be supercritical. Regarding the Hopf bifurcation there exists a similar theorem which was first proved by Wan (1978).

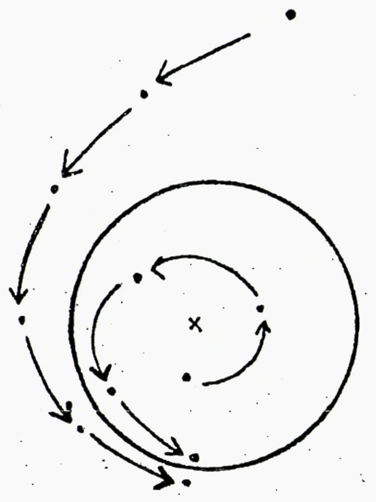


Figure 11: The outcome of a supercritical Hopf bifurcation. A point close to the unstable fixed point x moves away from x and approaches the attracting curve (indicated by a solid line). In the same way an initial point located outside the curve is also attracted.

Theorem 2.5.2 (Wan). Consider the map $K : \mathbb{R}^2 \to \mathbb{R}^2$ on standard form

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} f(x,y) \\ g(x,y) \end{pmatrix} \tag{2.5.16}$$

with eigenvalues $\lambda, \overline{\lambda} = e^{\pm i\theta}$. Then the Hopf bifurcation is supercritical whenever the quantity d (cf. (2.5.2)) in Theorem 2.5.2 is positive and the quantity a (cf. (2.5.14a)) is negative. a may be expressed as

$$a = -Re\left[\frac{(1-2\lambda)\overline{\lambda}^2}{1-\lambda}\xi_{11}\xi_{20}\right] - \frac{1}{2}|\xi_{11}|^2 - |\xi_{02}|^2 + Re(\overline{\lambda}\xi_{21}) \tag{2.5.17}$$

where

$$\xi_{20} = \frac{1}{8}\left[(f_{xx} - f_{yy} + 2g_{xy}) + i(g_{xx} - g_{yy} - 2f_{xy})\right]$$
$$\xi_{11} = \frac{1}{4}\left[(f_{xx} + f_{yy}) + i(g_{xx} + g_{yy})\right]$$
$$\xi_{02} = \frac{1}{8}\left[(f_{xx} - f_{yy} - 2g_{xy}) + i(g_{xx} - g_{yy} + 2f_{xy})\right]$$
$$\xi_{21} = \frac{1}{16}\left[(f_{xxx} + f_{xyy} + g_{xxy} + g_{yyy}) + i(g_{xxx} + g_{xyy} - f_{xxy} - f_{yyy})\right]$$

□

For a formal proof we refer to Wan's original paper (Wan, 1978).

(The idea of the proof is simple enough: we start with the original map, write it on standard form (i.e. (2.5.16)) and for each of the near identity transformations we then perform we express the new variables in terms of the original ones, thereby obtaining a in (2.5.14a) expressed in terms of the original quantities. The problem of course is that the calculations involved are indeed cumbersome and time-consuming as formulae (2.5.17) suggests.)

Example 2.5.1. Consider the stage-structured cod model proposed by Wikan and Eide (2004).

$$\begin{aligned} x_{1,t+1} &= F\,e^{-\beta x_{2,t}} x_{2,t} + (1-\mu_1)x_{1,t} \\ x_{2,t+1} &= P x_{1,t} + (1-\mu_2)x_{2,t} \end{aligned} \tag{2.5.18}$$

Here the cod stock x is split into one immature part x_1 and one mature part x_2. F is the density independent fecundity of the mature part while β measures the "strength" of cannibalism from the mature population upon the immature population. P is the survival probability from the immature stage to the mature stage and μ_1, μ_2 are natural death rates. We further assume: $0 < P \leq 1$, $0 < \mu_1, \mu_2 < 1$, $\beta > 0$, $F > 0$ and $FP > \mu_1, \mu_2$.

Assuming $x_1^* = x_{1,t+1} = x_{1,t}$ and $x_2^* = x_{2,t+1} = x_{2,t}$ the fixed point of (2.5.18) is found to be

$$(x_1^*, x_2^*) = \left[\frac{\mu_2}{\beta P}\ln\left(\frac{FP}{\mu_1\mu_2}\right), \frac{1}{\beta}\ln\left(\frac{FP}{\mu_1\mu_2}\right)\right] \tag{2.5.19}$$

The eigenvalue equation of the linearized map becomes (we urge the reader to work through the details)

$$\lambda^2 - (2-\mu_1-\mu_2)\lambda + (1-\mu_1)(1-\mu_2) - \mu_1\mu_2(1-\beta x_2^*) = 0 \tag{2.5.20}$$

Now, defining $a_1 = -(2-\mu_1-\mu_2)$, $a_2 = (1-\mu_1)(1-\mu_2) - \mu_1\mu_2(1-\beta x_2^*)$ and appealing to the Jury criteria (2.1.14) it is straightforward to show that the fixed point is stable as long as the inequalities

$$\beta\mu_1\mu_2 x_2^* > 0 \tag{2.5.21a}$$

$$2(2-\mu_1-\mu_2)+\mu_1\mu_2\beta x_2^* > 0 \tag{2.5.21b}$$

$$\mu_1+\mu_2-\beta\mu_1\mu_2 x_2^* > 0 \tag{2.5.21c}$$

hold. Clearly, (2.5.21a) and (2.5.21b) hold for any positive x_2^*. Thus, there will never be a transfer from stability to instability through a saddle-node or a flip bifurcation. (2.5.21c) is valid in case of x_2^* sufficiently small. Hence, the fixed point is stable in case of small equilibrium populations. However, if x_2^* is increased, as a result of increasing F which we from now on will use as our bifurcation parameter, it is clear that (x_1^*, x_2^*) will lose its stability at the threshold

$$x_2^* = \frac{\mu_1+\mu_2}{\beta\mu_1\mu_2} \tag{2.5.22a}$$

or alternatively when

$$F = \frac{\mu_1\mu_2}{P}\, e^{(\mu_1+\mu_2)/\mu_1\mu_2} \tag{2.5.22b}$$

Consequently, the fixed point will undergo a Hopf bifurcation at instability threshold and the complex modulus 1 eigenvalues become

$$\lambda, \overline{\lambda} = \frac{2-\mu_1-\mu_2}{2} \pm \frac{b}{2} i \tag{2.5.23}$$

where $b = \sqrt{4(\mu_1+\mu_2)-(\mu_1+\mu_2)^2}$.

In order to show that the Hopf bifurcation is supercritical we have to compute d (defined through (2.5.2)) and a (defined through (2.5.17)) and verify that $d > 0$ and $a < 0$.

By first computing λ from (2.5.20) we find

$$|\lambda| = \sqrt{(1-\mu_1)(1-\mu_2)-\mu_1\mu_2(1-\beta x_2^*)} \tag{2.5.24}$$

which implies

$$\frac{d}{dF}|\lambda| = \frac{1}{2\sqrt{(1-\mu_1)(1-\mu_2)-\mu_1\mu_2(1-\beta x_2^*)}} \cdot \frac{\mu_1\mu_2}{F}$$

and since the square root is equal to 1 at bifurcation and F is given by (2.5.22b) we obtain

$$\frac{d}{dF}|\lambda| = \frac{1}{2} P e^{-(\mu_1+\mu_2)/\mu_1\mu_2} = d > 0 \tag{2.5.25}$$

which proves that the eigenvalues leave the unit circle through an enlargement of the bifurcation parameter F.

In order to compute a we first have to express (2.5.18) on standard form (2.5.16). At bifurcation the Jacobian may be written as

$$J = \begin{pmatrix} 1-\mu_1 & \frac{1}{P}[(\mu_1\mu_2 - (\mu_1+\mu_2)] \\ P & 1-\mu_1 \end{pmatrix} \tag{2.5.26}$$

so by use of standard techniques the eigenvector $(Z_1, Z_2)^T$ belonging to λ is found to be

$$(z_1, z_2)^T = \left(\frac{\mu_1 - \mu_2}{2P} + \frac{b}{2P} i, \; 1 + 0i\right)^T \tag{2.5.27}$$

and the transformation matrix T and its inverse may be cast in the form

$$T = \begin{pmatrix} \frac{\mu_2-\mu_1}{2P} & -\frac{b}{2P} \\ 1 & 0 \end{pmatrix} \qquad T^{-1} = \begin{pmatrix} 0 & 1 \\ -\frac{2P}{b} & \frac{\mu_2-\mu_1}{b} \end{pmatrix} \tag{2.5.28}$$

The next step is to expand $f(x_2) = Fe^{-\beta x_2}$ up to third order. Then (2.5.18) becomes

$$\begin{aligned} x_{1,t+1} = &\Big\{ f(x_2^*) + f'(x_2^*)(x_{2,t} - x_2^*) + \frac{1}{2} f''(x_2^*)(x_{2,t} - x_2^*)^2 \\ &+ \frac{1}{6} f'''(x_2^*)(x_{2,t} - x_2^*)^3 \Big\} x_{2,t} + (1-\mu_1)x_{1,t} \\ x_{2,t+1} = &P_{x,t} + (1-\mu_2)x_{2,t} \end{aligned}$$

and by introducing the change of coordinates $(\hat{x}_1, \hat{x}_2) = (x_1 - x_1^*, x_2 - x_2^*)$, in order to bring the bifurcation to the origin, the result is

$$\begin{aligned} \hat{x}_{1,t+1} = &(1-\mu_1)\hat{x}_{1,t} + \frac{1}{P}\mu_1\mu_2(1 - \beta x_2^*)\hat{x}_{2,t} - \frac{\beta}{P}\mu_1\mu_2\left(1 - \frac{\beta}{2}x_2^*\right)\hat{x}_{2,t}^2 \\ &+ \frac{\beta^2}{P}\mu_1\mu_2\left(\frac{1}{2} - \frac{\beta}{6}x_2^*\right)\hat{x}_{2,t}^3 \end{aligned} \tag{2.5.29a}$$

$$\hat{x}_{2,t+1} = P\hat{x}_{1,t} + (1-\mu_2)\hat{x}_{2,t} \tag{2.5.29b}$$

where all terms of higher order than three have been neglected.

Finally, by applying the transformations

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = T \begin{pmatrix} u \\ v \end{pmatrix} \quad \begin{pmatrix} u \\ v \end{pmatrix} = T^{-1} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} \tag{2.5.30}$$

we obtain after some algebra that the original map (2.5.18) may be cast into standard form as

$$\begin{aligned} u_{t+1} &= \frac{2-\mu_1-\mu_2}{2}\, u_t - \frac{b}{2}\, v_t \\ u_{t+1} &= \frac{b}{2}\, u_t + \frac{2-\mu_1-\mu_2}{2}\, v_t + g(u_t, v_t) \end{aligned} \tag{2.5.31}$$

where

$$g(u,v) = \frac{2\beta}{b}\,\mu_1\mu_2 \left(1 - \frac{\beta}{2}x_2^*\right) u^2 - \frac{2\beta^2}{b}\,\mu_1\mu_2 \left(\frac{1}{2} - \frac{\beta}{6}x_2^*\right) u^3$$

Now at last, we are ready to compute the terms in formulae (2.5.17)

$$g_{uu} = \frac{4\beta}{b}\,\mu_1\mu_2 A \qquad g_{uuu} = -\frac{12\beta^2}{b}\,\mu_1\mu_2 B$$

where $A = 1 - (\beta/2)x_2^*$, $B = (1/2) - (\beta/6)x_2^*$. This yields:

$$\xi_{20} = \frac{1}{8}\, i g_{uu} \quad \xi_{11} = \frac{1}{4}\, i g_{uu} \quad \xi_{02} = \frac{1}{8}\, i g_{uu} \quad \xi_{21} = \frac{1}{16}\, i g_{uuu}$$

and

$$Re\left[\frac{(1-2\lambda)\overline{\lambda}^2}{1-\lambda}\,\xi_{11}\xi_{20}\right] = -\frac{g_{uu}^2}{256(\mu_1+\mu_2)} \times \left[3(\mu_1+\mu_2)\left[(2-u_1-u_2)^2 - b^2\right] - 2(2-\mu_1-\mu_2)b^2\right]$$

so finally, by computing $|\xi_{11}|^2 = (1/16)g_{uu}^2$, $|\xi_{02}|^2 = (1/64)g_{uu}^2$, $Re(\overline{\lambda}\xi_{21}) = (1/32)bg_{uuu}$ and inserting into (2.5.17) we eventually arrive at

$$a = -\frac{\beta^2}{16(\mu_1+\mu_2)}\left\{(2\mu_1\mu_2)^2 + (\mu_1+\mu_2)\left[(2\mu_1\mu_2 - (\mu_1+\mu_2))^2 - \mu_1\mu_2\right]\right\} \tag{2.5.32}$$

which is negative for all $0 < \mu_1, \mu_2 < 1$. Consequently, the fixed point (2.5.19) undergoes a supercritical Hopf bifurcation at the threshold (2.5.22a,b) (i.e. when (x_1^*, x_2^*) fails to be stable through an increase of F, a closed invariant attracting curve surrounding (x_1^*, x_2^*) is established). □

In the next example, which we will mainly present as an exercise, most of the cumbersome and time-consuming calculations we had to perform in Example 1.5.1 are avoided.

Example 2.5.2. Assume that the parameter $\mu > 1$ and consider the map

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} y \\ \mu y(1-x) \end{pmatrix} \tag{2.5.33}$$

a) Show that the nontrivial fixed point

$$(x^*, y^*) = \left(\frac{\mu-1}{\mu}, \frac{\mu-1}{\mu}\right)$$

b) Compute the Jacobian and show that the eigenvalue equation may be expressed as

$$\lambda^2 - \lambda + \mu - 1 = 0$$

c) Use the Jury criteria (2.1.14) and show that the fixed point is stable whenever $1 < \mu < 2$ and that a Hopf bifurcation occurs at the threshold $\mu = 2$.

d) Show that $|\lambda| = \sqrt{\mu - 1}$ and moreover that

$$\frac{d}{d\mu}|\lambda|_{\mu=2} > 0$$

which proves that the eigenvalues leave the unit circle at bifurcation threshold.

e) Assuming $\mu = 2$, apply the change of coordinates $(\hat{x}, \hat{y}) = (x - (1/2), y - (1/2))$ together with the transformations

$$\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix} = T\begin{pmatrix} u \\ v \end{pmatrix} \quad \begin{pmatrix} u \\ v \end{pmatrix} = T^{-1}\begin{pmatrix} \hat{x} \\ \hat{y} \end{pmatrix}$$

where

$$T = \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 1 & 0 \end{pmatrix}$$

(verify that the columns in T are the real and imaginary parts of the eigenvectors belonging to the eigenvalues of the Jacobian respectively) and show that (2.5.33) may be written on standard form at bifurcation threshold as

$$\begin{pmatrix} u \\ v \end{pmatrix} \to \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} f(u,v) \\ g(u,v) \end{pmatrix} \tag{2.5.34}$$

where $f(u,v) = -u^2 - \sqrt{3}\,uv$ and $g(u,v) = (1/\sqrt{3})u^2 + uv$.

f) Referring to Theorem 2.5.2 show that the quantity a defined in (2.5.17) is negative, hence that in case of $\mu > 2$, $|\mu - 2|$ small, there exists an attracting curve surrounding the unstable fixed point (x^*, y^*). □

Exercise 2.5.1 (Strong resonant case I). Consider the two-age structured population model

$$(x_1, x_2) \to (F_2 x_2, P\,e^{-x_1} x_1) \tag{2.5.35}$$

where $0 < P \leq 1$, $F_2 > 0$ and $PF_2 > 1$.

a) Show that the fixed point $(x_1^*, x_2^*) = (\ln(PF_2), (1/F_2)\ln(PF_2))$.

b) Show that the eigenvalue equation may be cast in the form $\lambda^2 + x_1^* - 1 = 0$ and further that a Hopf bifurcation takes place at the threshold $x_1^* = 2$ (or equivalently when $F_2 = (1/P)\exp(2)$).

c) Show that λ equals fourth root of unity at bifurcation threshold.

Note that the result obtained in c) violates assumption (2.5.1) in Theorem 2.5.1 which of course means that neither Theorem 2.5.1 nor Theorem 2.5.2 applies on map (2.5.35). We urge the reader to perform numerical experiments where $F_2 > (1/P)\exp(2)$ in order to show that when (x_1^*, x_2^*) fails to be stable, an exact 4-periodic orbit with small amplitude is established. (For further reading, cf. Wikan (1997.) □

Exercise 2.5.2 (Strong resonant case II). Repeat the analysis from the previous exercise on the map

$$(x_1, x_2) \to \left(F\, e^{-(x_1+x_2)}(x_1 + x_2), x_1\right) \tag{2.5.36}$$

Hint: λ equals third root of unity at bifurcation threshold. □

We close this section by once again emphasizing that the outcome of a supercritical Hopf bifurcation is that when the fixed point fails to be stable an attracting invariant curve which surrounds the fixed point is established. In section 2.7 we shall focus on the nonstationary dynamics on such a curve as well as possible routes to chaos. However, before we turn to those questions we shall in the next section present the center manifold theorem which plays a key role in order to analyse the nature of bifurcations in higher dimensional problems.

2.6 The center manifold theorem

Recall that in our treatment of the flip bifurcation (cf. section 1.5) we considered one-dimensional maps of the form $f : \mathbb{R} \to \mathbb{R}$ and when we studied the Hopf bifurcation in the previous section the main theorems were stated for two-dimensional maps $f : \mathbb{R}^2 \to \mathbb{R}^2$. Let us now turn to higher-dimensional maps, $f : \mathbb{R}^n \to \mathbb{R}^n$. Of course, $|\lambda| = 1$ at bifurcation in these cases too but how do we determine the nature of the bifurcation involved when the fixed point fails to be hyperbolic?

The main conclusion is that there exists a method which applied to a map on the form $f : \mathbb{R}^n \to \mathbb{R}^n$ reduces the bifurcation problem to a study of a map $g : \mathbb{R}^2 \to \mathbb{R}^2$ (Hopf), or $g : \mathbb{R} \to \mathbb{R}$ (flip). The cornerstone in the theory which allows this conclusion is the center manifold theorem for maps which we now state.

Theorem 2.6.1 (Center manifold theorem). Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be a C^k, $k \geq 2$ map and assume that the Jacobian $Df(0)$ has a modulus 1 eigenvalue and, moreover, that all eigenvalues of $Df(0)$ splits into two parts α_c, α_s such that

$$|\lambda| = \begin{cases} 1 & \text{if } \lambda \in \alpha_c \\ <1 & \text{if } \lambda \in \alpha_s \end{cases}$$

Further, let E_c be the (generalized) eigenspace of α_c, $\dim E_c = d < \infty$. Then there exists a domain V about 0 in $\mathbb{R}^n$ and a C^k submanifold W of V of dimension d passing through 0 which is tangent to E_c at 0 which satisfies:

I) If $x \in W$ and $f(x) \in V$ then $f(x) \in W$.

II) If $f^{(n)}(x) \in V$ for all $n = 0, 1, 2, \ldots$ then the distance from $f^{(n)}(x)$ to W approaches zero as $n \to \infty$. □

For a proof of Theorem 2.6.1, cf. Marsden and McCracken (1976, p. 28–43).

Concerning Hopf bifurcation problems, the essence of Theorem 2.6.1 is that there exists an invariant manifold of dimension $2 \subset \mathbb{R}^n$ which has the eigenspace belonging to the complex eigenvalues as tangent space at the bifurcating nonhyperbolic fixed points. In case of flip bifurcation problems, $\dim W = 1$. Thus close to the bifurcation, our goal is to restrict the original map to the invariant center manifold W and then proceed with the analysis by using the results in Theorems 2.5.1 and 2.5.2 in case of Hopf bifurcation problems and Theorem 1.5.1 in the flip case.

Let us now in general terms describe how such a restriction may be carried out. To this end, consider our discrete system written in the form

$$\begin{aligned} \mathbf{x}_{t+1} &= A\mathbf{x}_t + F(\mathbf{x}_t, \mathbf{y}_t) \\ \mathbf{y}_{t+1} &= B\mathbf{y}_t + G(\mathbf{x}_t, \mathbf{y}_t) \end{aligned} \tag{2.6.1}$$

where all the eigenvalues of A are on the boundary of the unit circle and those of B within the unit circle. (If the system we want to study is not on the form as in (2.6.1) we first apply the procedure in Example 2.5.1, see also the proof of Theorem 2.5.1.)

Now, since the center manifold W is tangent to the (generalized) eigenspace E_c, we may represent it as a local graph

$$W = \{(\mathbf{x}, \mathbf{y})/\mathbf{y} = h(\mathbf{x})\} \qquad h(0) = Dh(0) = 0 \tag{2.6.2}$$

and by substituting (2.6.2) into (2.6.1) we have

$$\mathbf{y}_{t+1} = h(\mathbf{x}_{t+1}) = h(A\mathbf{x}_t + F(\mathbf{x}_t, h(\mathbf{x}_t)) = Bh(\mathbf{x}_t) + G(\mathbf{x}_t, h(\mathbf{x}_t))$$

or equivalently

$$h(A\mathbf{x} + F(\mathbf{x}, h(\mathbf{x}))) - Bh(\mathbf{x}) - G(\mathbf{x}, h(\mathbf{x})) = 0 \tag{2.6.3}$$

An explicit expression of $h(\mathbf{x})$ is out of reach in most cases, but one can approximate h by its Taylor series at the bifurcation as

$$h(x) = ax^2 + bx^3 + O(x^4) \tag{2.6.4}$$

where the coefficients a, b are determined through (2.6.3), and finally the restricted map is obtained by inserting the series of h into (2.6.1).

Example 2.6.1. Consider the Leslie matrix model

$$f : \mathbb{R}^2 \to \mathbb{R}^2 \qquad \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} F(1-\gamma x)^{1/\gamma} & F(1-\gamma x)^{1/\gamma} \\ P & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{2.6.5}$$

where $x = x_1 + x_2$ is the total population. □

(2.6.5) is often referred to as the Deriso–Schnute population model. Note that if $\gamma \to 0$, (2.6.5) is nothing but the Ricker model (see (2.3.4) and Examples 2.4.1 and 2.4.3). If $\gamma = -1$ we are left with the Beverton and Holt model (see (2.3.5) and Exercise 2.4.1).

Our goal is to show that under the assumptions $F(1+P) > 1$, $0 < P < 1/2$, $\gamma > -(1-P)/2$ the fixed point (x_1^*, x_2^*) of (2.6.5) will undergo a supercritical flip bifurcation at instability threshold.

We urge the reader to verify the following properties:

$$(x_1^*, x_2^*) = \left(\frac{1}{1+P}x^*, \frac{1}{1+P}x^*\right) \tag{2.6.6}$$

where $x^* = (1/\gamma)[1-(P+PF)^{-\gamma}]$. Defining $f(x) = F(1-\gamma x)^{1/\gamma}$ the Jacobian becomes

$$\begin{pmatrix} f'x^*+f & f'x^*+f \\ P & 0 \end{pmatrix}$$

where $f = f(x^*) = 1/(1+P)$ and $f' = f'(x^*)$.

Show by use of the Jury criteria (2.1.14) that whenever $0 < P < 1/2$, $\gamma > -(1-P)/2$ the fixed point (2.6.6) will undergo a flip bifurcation when $f'x^* = -2/(1-P^2)$ and that the Jacobian at bifurcation threshold equals

$$\begin{pmatrix} -\frac{1}{1-P} & -\frac{1}{1-P} \\ P & 0 \end{pmatrix} \tag{2.6.7}$$

and moreover, that the eigenvalues of (2.6.7) are $\lambda_1 = -1$ and $\lambda_2 = -P/(1-P)$.

Now, in order to show that the flip bifurcation is of supercritical nature we must appeal to Theorem 1.5.1 but since that theorem deals with one-dimensional maps, we first have to express (2.6.5) on the appropriate form (2.6.1) and then perform a center manifold restriction as explained through (2.6.2)–(2.6.4).

The form (2.6.1) is achieved by performing the same kind of calculations as in Example 2.5.1. The eigenvectors belonging to λ_1 and λ_2 are easily found to be $(-1/P, 1)^T$ and $(-1/(1-P), 1)^T$ respectively so the transformation matrix T and its inverse become

$$T = \begin{pmatrix} -\frac{1}{P} & -\frac{1}{1-P} \\ 1 & 1 \end{pmatrix} \quad T^{-1} = \begin{pmatrix} \frac{P(1-P)}{2P-1} & \frac{P}{2P-1} \\ -\frac{P(1-P)}{2P-1} & -\frac{1-P}{2P-1} \end{pmatrix} \tag{2.6.8}$$

Further, by expanding f up to third order, i.e.

$$f(x) \approx f(x^*) + f'(x*)(x-x^*) + \frac{1}{2}f''(x^*)(x-x^*)^2 + \frac{1}{6}f'''(x^*)(x-x^*)^3$$

and applying the change of coordinates $(\hat{x}_1, \hat{x}_2) = (x_1 - x_1^*, x_2 - x_2^*)$, using the fact that $f'x^* = -2/(1-P^2)$ at bifurcation threshold gives

$$\begin{aligned} \hat{x}_{1,t+1} &= -\frac{1}{1-P}\hat{x}_{1,t} - \frac{1}{1-P}\hat{x}_{2,t} + \{1\}\hat{x}_t^2 + \{2\}\hat{x}_t^3 \\ \hat{x}_{2,t+1} &= P\hat{x}_{1,t} \end{aligned} \tag{2.6.9}$$

where all terms of higher order than 3 have been neglected and $\{1\}$ and $\{2\}$ are defined through

$$\{1\} = f' + \frac{1}{2} f'' x^* \qquad \{2\} = \frac{1}{2} f'' + \frac{1}{6} f''' x^*$$

Now, performing the transformations

$$\begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix} = T \begin{pmatrix} u \\ v \end{pmatrix} \qquad \begin{pmatrix} u \\ v \end{pmatrix} = T^{-1} \begin{pmatrix} \hat{x}_1 \\ \hat{x}_2 \end{pmatrix}$$

on (2.6.9) we arrive at

$$\begin{aligned} u_{t+1} &= -u_t + g(u_t, v_t) \\ v_{t+1} &= -\frac{P}{1-P} v_t - g(u_t, v_t) \end{aligned} \tag{2.6.10}$$

where $g(u, v) = A[(1-P)^2 u + P^2 v]^2 + B[(1-P)^2 + P^2 v]^3$

$$A = \frac{1}{P(2P-1)(1-P)} \{1\} \qquad B = -\frac{1}{P^2(2P-1)(1-P)^2} \{2\}$$

and we observe that (2.6.10) is nothing but the original map (2.6.5) written on the desired form (2.6.1).

The next step is to restrict (2.6.10) to the center manifold. Thus, assume

$$v = i(u) = Ku^2 + Lu^3 \tag{2.6.11}$$

By use of (2.6.3) we now have

$$i(-u_t + g(u_t, v_t)) + \frac{P}{1-P} i(u_t) + g(u_t, i(u_t)) = 0$$

which is equivalent to

$$\left[K + \frac{PK}{1-P} + (1-P)^4 A\right] u^2 + \left[\frac{PL}{1-P} - 2KA(1-P)^4 - L + 2AP^2(1-P)^2 K + B(1-P)^6\right] u^3 = 0$$

from which we obtain

$$K = -(1-P)^5 A \qquad L = (1-P)^7 [B + 2A^2(1-P)(1-2P)]$$

Finally, by inserting $v = Ku^2 + Lu^3$ into the first component of (2.6.10) the restricted map may be cast in the form

$$u_{t+1} = h(u_t) = -u_t + A(1-P)^4 u_t^2 + (1-P)^6[B - 2A^2P^2(1-P)]u_t^3 + O(u^4) \tag{2.6.12}$$

Since $u \to h(u)$ is a one-dimensional map we may now proceed by using Theorem 1.5.1 in order to show that the flip bifurcation is supercritical. A time-consuming but straightforward calculation now yields that the quantity b defined in Theorem 1.5.1 becomes

$$\begin{aligned} b &= \frac{1}{2}\left(\frac{\partial^2 h}{\partial u^2}\right)^2 + \frac{1}{3}\frac{\partial^3 h}{\partial u^3} \\ &= \left[\frac{2\gamma}{1-P}+1\right]^2 \frac{2(1-P)^3}{P^2(1+P)(1-2P)}\left\{(P-\gamma)^2 + \frac{1}{6}(1-\gamma)(4\gamma - 3P + 1)\right\} \end{aligned} \tag{2.6.13}$$

at bifurcation. Here we may observe that $W(\gamma) = \{\ \}$ attains its minimum when $\gamma = (9/4)P - 3/4$ and that $W((9/4)P - 3/4) > 0$ whenever $0 < P < 1/2$. Hence $b > 0$.

Regarding the nondegeneracy condition a defined in Theorem 1.5.1, it may be expressed as

$$a = \frac{\partial h}{\partial F}\frac{\partial^2 h}{\partial u^2} - \left(\frac{\partial h}{\partial u} - 1\right)\frac{\partial^2 h}{\partial u \partial F} \neq 0 \text{ at } (u,v) = (0,0)$$

Now, since the bifurcation is transformed to the origin it follows that $\partial h/\partial u = -1$ and $\partial h/\partial F = 0$. Therefore the condition $a \neq 0$ simplifies to

$$a = 2\frac{\partial^2 h}{\partial u \partial F} \neq 0 \Leftrightarrow 2\frac{\partial \lambda}{\partial F} \neq 0$$

since in general $\partial h/\partial u = \lambda$. From the Jacobian:

$$\lambda = \frac{1}{2}\left(w - \sqrt{w^2 + 4Pw}\right)$$

where

$$w = f'x^* + f = \frac{1}{1+P}\left\{-\frac{1}{\gamma}[(F+FP)^\gamma - 1] + 1\right\}$$

it follows that

$$\begin{aligned} 2\frac{\partial \lambda}{\partial F} &= \frac{dw}{dF} - \frac{1}{2\sqrt{w^2+4Pw}}\left(2w\frac{dw}{dF} + 4P\frac{dw}{dF}\right) \\ &= \frac{dw}{dF}\left[1 - \frac{1}{\sqrt{w^2+4Pw}}(w+2P)\right] \end{aligned}$$

At bifurcation, $w = -(1-P)^{-1}$ which inserted into the expression above gives

$$2\frac{\partial \lambda}{\partial F} = -2\left[\frac{2\gamma}{1-P}+1\right]^{1-(1/\gamma)} \frac{(1-P)^2}{1-2P} \tag{2.6.14}$$

and clearly, (2.6.14) is nonzero whenever $0 < P < 1/2$. Consequently, the flip bifurcation is supercritical, which means that when the fixed point fails to be stable, a stable two-periodic orbit is established.

—

We close this section by showing the dynamics beyond the flip bifurcation threshold for the Ricker map

$$(x_0, x_1) \to (F\,e^{-x}(x_0+x_1), Px_0) \tag{2.6.15}$$

which is a special case of map (2.6.5) (the case $\gamma \to 0$). Assuming $F(1+P) > 1$ the nontrivial fixed point of (2.6.15) is

$$(x_0^*, x_1^*) = \left(\frac{1}{1+P}\ln(F(1+P)), \frac{P}{1+P}\ln(F(1+P))\right)$$

and whenever $0 < P < 1/2$ we have according to the preceding example that the fixed point undergoes a supercritical flip bifurcation at the threshold $F = (1/(1+P))\exp(2/(1-P))$.

Now, consider the value $P = 0.2$. Under this choice the fixed point is stable in the F interval $0.834 < F < 10.152$ and in Figure 12 we have plotted the bifurcation diagram of (2.6.15) in the range $5 < F < 80$. We clearly identify the supercritical flip at the threshold $F = 10.152$ and beyond that stable periodic orbits of period 2^k are established through further increase of F so what we recognize is essentially the same kind of dynamical behaviour as we found when we considered one-dimensional maps. Beyond the point of accumulation for the flip bifurcation sequence the dynamics becomes chaotic as displayed in Figure 13. Note that the chaotic attractor consists of 4 disjoint subsets (branches) that are visited once every fourth iteration so a certain kind of four periodicity is preserved in the chaotic regime. In case of higher F values the branches merge together.

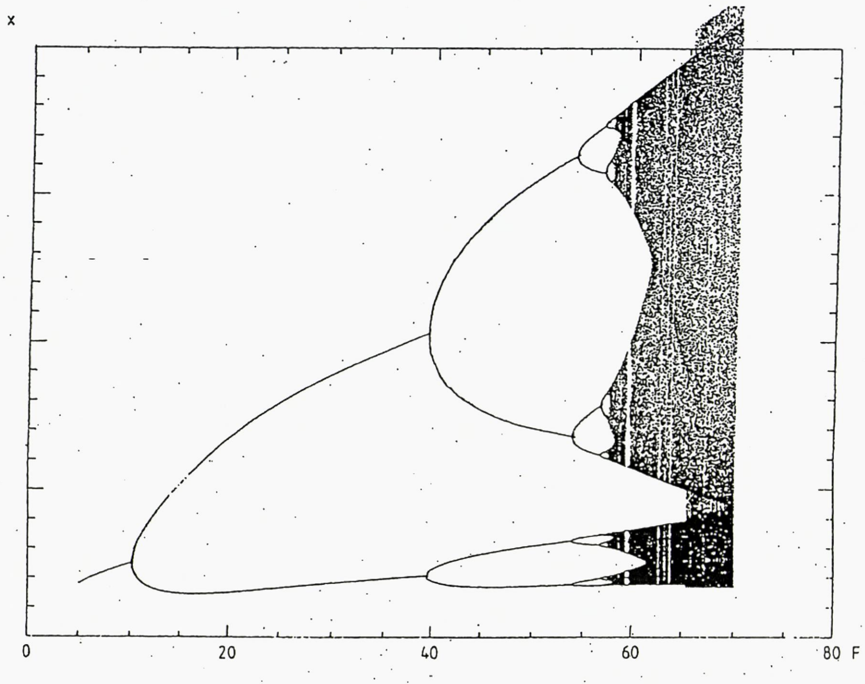


Figure 12: The bifurcation diagram of map (2.6.14) in the case $P = 0.2$. For small F values we see the stable fixed point of (2.6.14) which undergoes a supercritical flip bifurcation when $F = 10.152$. Through further increase of F stable orbits of period 2^k are created until an accumulation value F_a for the flip bifurcations is reached. Beyond F_a the dynamics is chaotic.

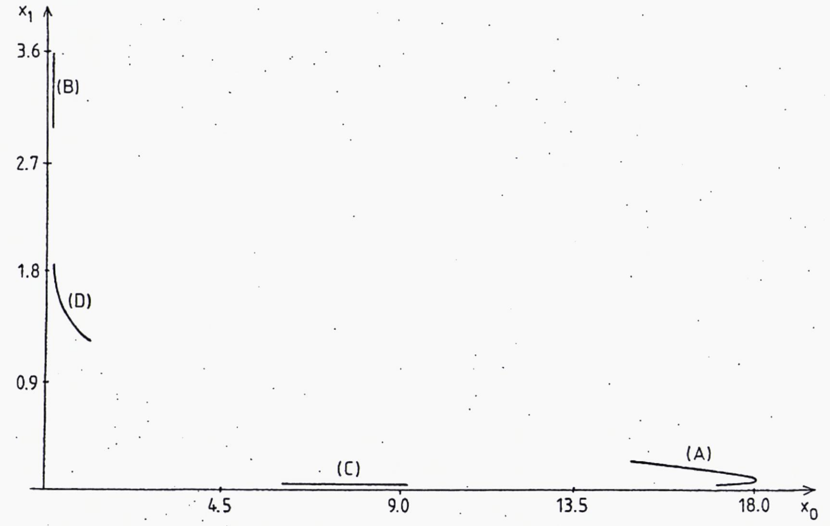


Figure 13: The chaotic attractor consisting of four separate branches just beyond the point of accumulation for the flip bifurcations in the case $P = 0.2$, $F = 58.5$. The dynamics goes in the direction $A \to B \to C \to D$.

2.7 Beyond the Hopf bifurcation, possible routes to chaos

As we proved in section 2.5, the outcome of a supercritical Hopf bifurcation is that when the fixed point of a discrete map fails to be stable, an attracting invariant curve which surrounds the fixed point is created. Our goal in this section is to describe the dynamics on such an invariant curve. We will also discuss possible routes to chaos and as it will become clear, the dynamics may be much richer than in the one-dimensional cases discussed in Part I.

In general terms, the dynamics on an invariant curve (circle) created by a Hopf bifurcation may be analysed by use of equation (2.5.14b). Indeed, if we substitute the fixed point r^* of (2.5.14a) into (2.5.14b) we arrive at

$$\varphi \to \varphi + c - \frac{bd}{a}\mu = \varphi + \sigma(\mu) \tag{2.7.1}$$

where $c = \arg\lambda$. Also recall that when we derived (2.5.14a,b) we first transformed the bifurcation to the origin. If the Hopf bifurcation occurs at a threshold $\mu_0 \neq 0$, $\sigma(\mu) = c + (bd/a)(\mu_0 - \mu)$.

Now, the essential feature is that successive iterations of (2.7.1) simply "move" or rotate points from one location to another on the invariant curve. Hence, the original map $f_\mu : \mathbb{R}^2 \to \mathbb{R}^2$ may be regarded as being topological equivalent to a circle map $g : S' \to S'$ once the invariant curve is established. Moreover, considering g, one may define its rotation number as the average amount that points are rotated by an iteration of the map. Therefore, we may (to leading order, recall that (2.7.1) is a truncated map) regard (2.7.1) as a circle map with rotation number $\sigma(\mu)$.

Remark 1. A more precise definition of the rotation number may be achieved along the following line: Given a circle map $g : S \to S$ we first "lift" g to the real line $\mathbb{R}$ by use of $\pi : \mathbb{R} \to S$, $\pi(x) = \cos(2\pi x) + i\sin(2\pi x)$ and then define the lift F as $F : \mathbb{R} \to \mathbb{R}$, $\pi \circ F = g \circ \pi$. Next, let $\sigma_0(F) = \lim_{n\to\infty} F^n(x)/x$ and finally define the rotation number of $g, \sigma(g)$, as the unique number in $[0, 1\rangle$

such that $\sigma_0(F) - \sigma(g)$ is an integer. In Devaney's book there is an excellent introduction to circle maps. □

Returning to map (2.7.1) the rotation number may be irrational or rational. In the former case this means that as the number of iterations of the map tends to infinity, the invariant curve will be filled with points. Whenever σ irrational, an orbit of a point is often referred to as a quasistationary orbit. If $\sigma = 1/n$, rational, the dynamic outcome is an n-period orbit. It is of great importance to realize that whenever the rotation number is rational for a given parameter value $\mu = \mu_r$, it follows from the implicit function theorem that there exists an open interval about μ_r where the periodicity is maintained. This phenomenon is known as frequency locking of periodic orbits. Consequently, periodic dynamics will occur in parameter regions, not at isolated parameter values only. As we shall see, such regions (or intervals) may in fact be large. So in order to summarize: beyond the Hopf bifurcation (and outside the strongly resonant cases where λ is third or fourth root of unity) there are quasistationary orbits restricted to an invariant curve and there may also be orbits of finite period established through frequency locking as the value of the parameter μ in the model is increased.

Our next goal is by way of examples to study in more detail the interplay between these cases as well as studying possible routes to chaos.

—

Example 2.7.1. First, consider the two-age class population model

$$(x_0, x_1) \to (Fx, P\,e^{-\alpha x} x_0) \tag{2.7.2}$$

which is a semelparous species model where the fecundity F is constant while the survival probability $p(x) = P\exp(-\alpha x)$ is density dependent. α is a positive number (scaling constant) and we assume that $PF > 1$.

It is easy to verify that (2.7.2) possesses the following properties: The fixed point may be expressed as

$$(x_0^*, x_1^*) = \left(\frac{F}{1+F}x^*, \frac{1}{1+F}x^*\right) \tag{2.7.3}$$

where $x^* = x_0^* + x_1^* = \alpha^{-1} \ln(PF)$. Moreover, the eigenvalue equation may be cast in the form

$$\lambda^2 + \frac{\ln(PF)}{1+F}\lambda + \frac{F\ln(PF)}{1+F} - 1 = 0 \tag{2.7.4}$$

and from the Jury criteria one obtains that the fixed point is stable in case of PF small but undergoes a Hopf bifurcation at the threshold

$$P = P_c = \frac{1}{F}e^{2(1+F)/F} \tag{2.7.5}$$

Note that α drops out of (2.7.4), (2.7.5) which simply means that stability properties are independent of α. At bifurcation threshold (2.7.5) the solution of the eigenvalue equation becomes

$$\lambda = -\frac{1}{F} \pm \sqrt{1 - \frac{1}{F^2}}\, i \tag{2.7.6}$$

A final observation is that by rewriting (2.7.2) on standard form (as in Example 2.5.1) and then apply Theorem 2.5.2, it is possible to prove that the bifurcation is supercritical.

Now, let us scrutinize a numerical example somewhat closer. Assume $P = 0.6$. Then from (2.7.5) the F value at bifurcation threshold is numerically found to be $F = F_c = 14.1805$. We want to investigate the dynamics when $F > F_c$. In Figure 14 we show the dynamics just beyond the instability threshold in the case $(\alpha, P, F) = (0.02, 0.6, 15)$. From an initial state (x_{00}, x_{10}) 500 iterations have been computed and the last 20 together with the (unstable!) fixed point are plotted. The invariant curve is indicated by the dashed line so clearly the original map (2.7.2) does nothing but rotate points around that curve, i.e. (2.7.2) acts as a circle map.

Moreover, Figure 14 demonstrates a clear tendency towards 4-periodic dynamics. This is as expected due to the location of the eigenvalues. Indeed, when $F_c = 14.1805$ it follows from (2.7.6) that the eigenvalues are located very close to the imaginary axis ($\lambda_{1,2} = -0.0750 \pm \sqrt{0.9975}\, i$), and since the rotation

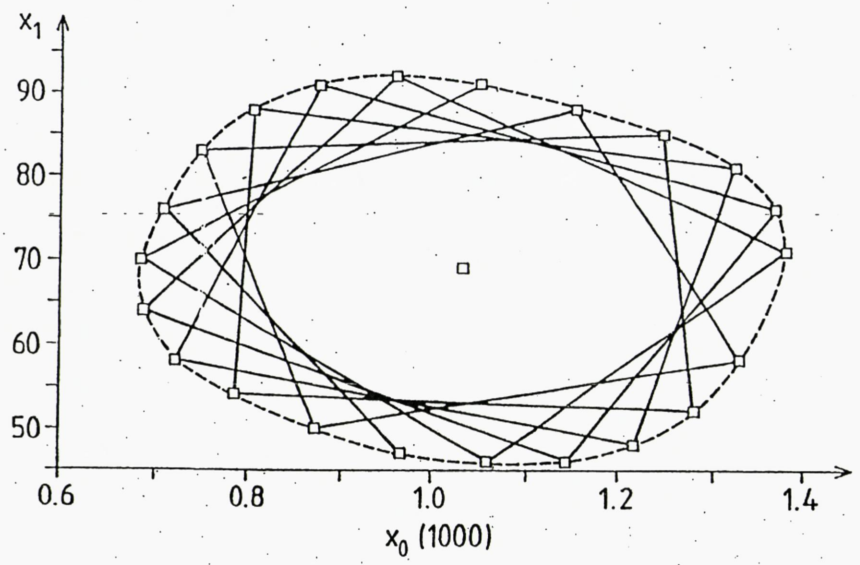


Figure 14: The dynamics of map (2.7.2) (a quasistationary orbit), just beyond the Hopf bifurcation threshold.

number (up to leading order!) has the form $\sigma(F) = c + (bd/a)(F_c - F)$ where $c = \arg\lambda$ it follows that σ must be close to $1/4$ in case of F close to F_c.

If we increase F beyond 15 we observe (due to frequency locking!) that an exact 4-periodic orbit is established. This is shown in Figure 15 in the case $(\alpha, P, F) = (0.02, 0.6, 20)$ and further, it is possible to verify numerically that the exact 4-periodicity is maintained as long as F does not superceede the value 21.190.

At $F = 21.190$ the fourth iterate of (2.7.2) undergoes a flip bifurcation, thus an 8-periodic orbit is established, and through further enlargement of F we find that new flip bifurcations take place at the parameter values 24.232 and 24.883 which again result in orbits of period 16 and 32 respectively. Hence we oberve nothing but the flip bifurcation sequence which we discussed in Part I. The point of accumulation for the flip bifurcation is found to be $F_a \approx 25.07$ and in case of $F > F_a$ the dynamics becomes chaotic.

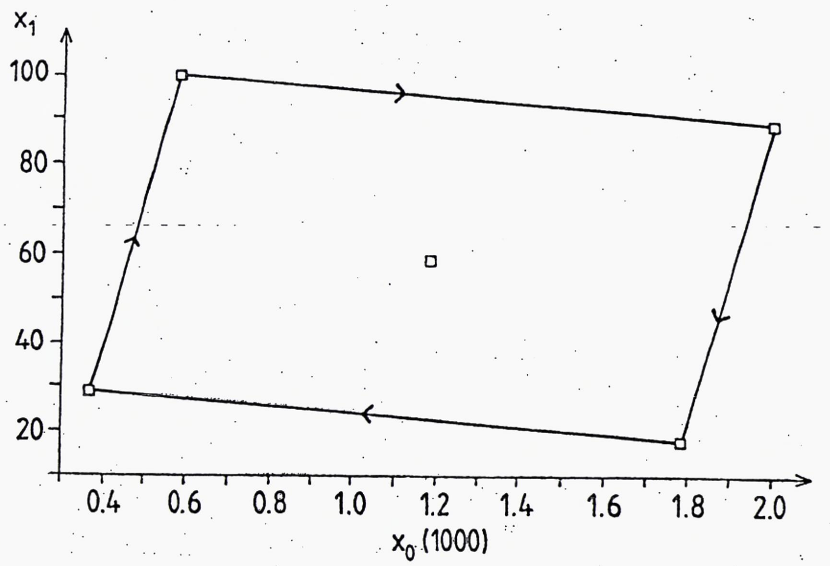


Figure 15: A 4-periodic orbit generated by map (2.7.2).

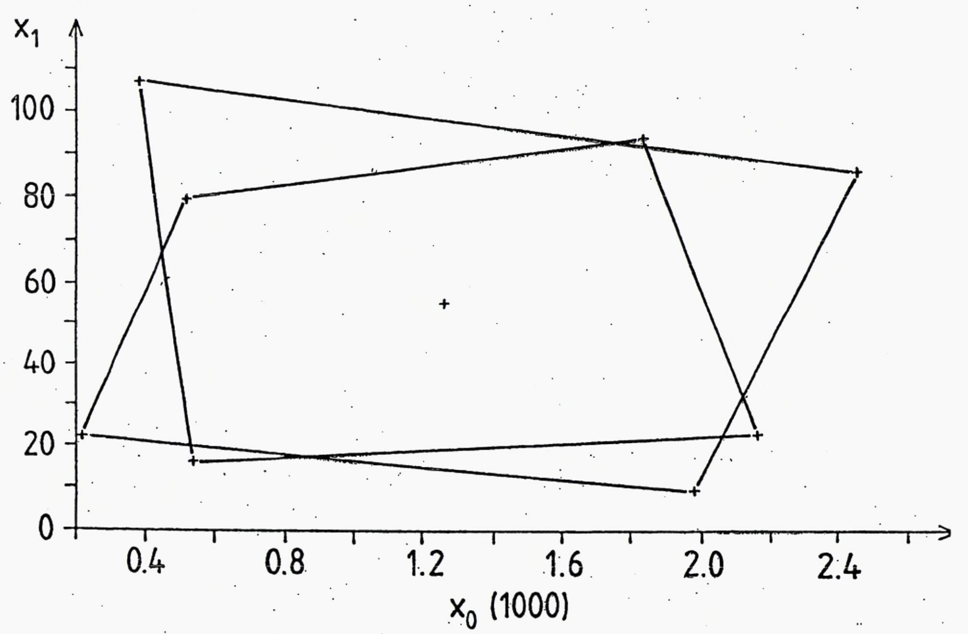


Figure 16: An 8-periodic orbit generated by map (2.7.2).

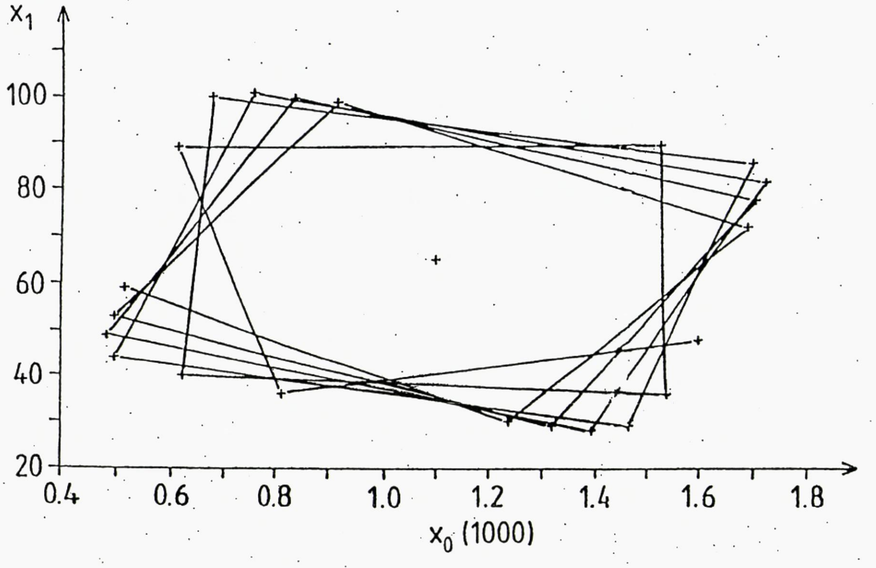


Figure 17: A 32-periodic orbit generated by map (2.7.2).

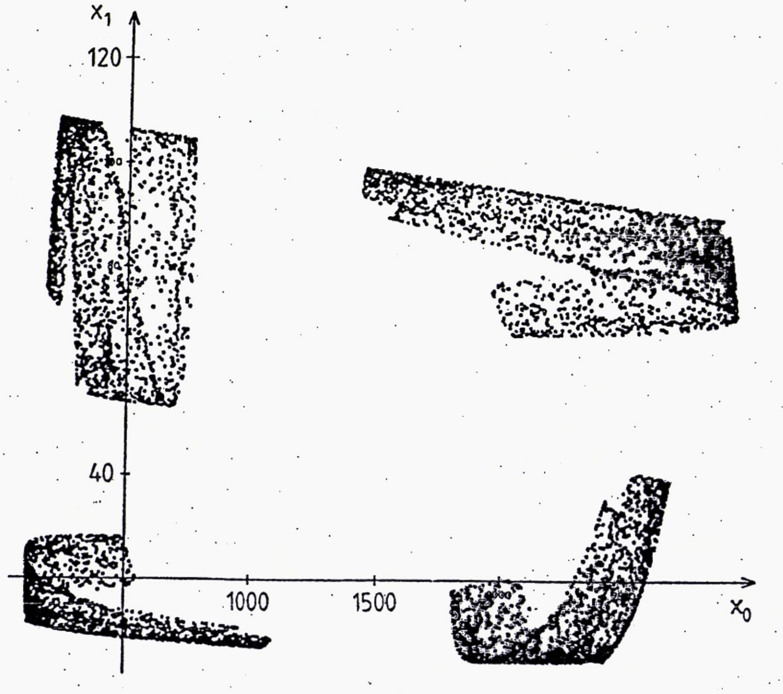


Figure 18: Map (2.7.2) in the chaotic regime.

These findings are shown in Figures 16, 17 and 18. In Figures 16 and 17 periodic orbits of period 8 and 32 are displayed. In Figure 18 we show the chaotic attractor. Note that the attractor is divided in 4 disjoint subsets and that each of the subsets are visited once every fourth iteration so there is a kind of 4-periodicity preserved, even in the chaotic regime. □

Example 2.7.2. The next example is basically the same as the previous one but the dimension of the map has been extended by 1 and we consider a general survival probability $p(x)$, $0 < p(x) \leq 1$, $p'(x) \leq 0$, instead of $p(x) = P\exp(-x)$. Hence we consider the problem

$$(x_1, x_2, x_3) \to (F_3 x_3, p(x)x_1, p(x)x_2) \tag{2.7.7}$$

Skipping computational details (which are much more cumbersome here than in our previous example) we find that the nontrivial fixed point is

$$(x_1^*, x_2^*, x_3^*) = \left(\frac{x^*}{K}, p(x^*)\frac{x^*}{K}, p^2(x^*)\frac{x^*}{K}\right) \tag{2.7.8}$$

where $K = \sum_{i=1}^{3} p^{i-1}(x^*)$ and $x^* = p^{-1}(F_3^{-1/(n-1)})$. ($p^{-1}$ denotes the inverse of p.)

Moreover, by first computing the Jacobian and then use the Jury criteria, it is possible to show that (2.7.8) is stable as long as

$$-p'(x^*)\frac{x^*}{K} < p(x^*)\,\frac{1 + p(x^*) - 2p^2(x^*)}{(1 + p(x^*))(1 - p^2(x^*))} \tag{2.7.9}$$

(2.7.8) becomes unstable when F_3 is increased to a level F_{H1} where (2.7.9) becomes an equality. At that level a (supercritical) Hopf bifurcation occurs and the complex modulus 1 eigenvalues may be expressed as

$$\lambda_{1,2} = -\frac{p^2(x^*)}{1 + p(x^*)} \pm \sqrt{1 - \frac{p^4(x^*)}{(1 + p(x^*))^2}}\, i \tag{2.7.10}$$

Now, for comparison reasons, assume that $p(x) = P\exp(-x)$ just as in Example 2.7.1. Then it easily follows that F_3 is a "large" number at bifurcation threshold F_{H1} and further that $p(x^*) \ll 1$. Consequently, $\lambda_{1,2}$ are located very close to the imaginary axis, in fact even closer than the eigenvalues from Example 2.7.1. When we increase F_3 beyond F_{H1} we observe the following dynamics: In case of $F_3 - F_{H1}$ small we find an almost 4-periodic orbit restricted on an invariant curve and through further enlargement of F_3 we once again find (through frequency locking) that an exact 4-periodic orbit is the outcome. Thus the dynamics is qualitatively similar to what we found in Example 2.7.1. However, if we continue to increase F_3 we do not experience the flip bifurcation sequence. Instead we find that the fourth iterate of map (2.7.7) undergoes a (supercritical) Hopf bifurcation at a threshold $F_3 = F_{H2}$. Therefore, beyond that threshold, and in case of $F_3 - F_{H2}$ small, the dynamics is restricted on 4 disjoint invariant attracting curves which are visited once every fourth iteration. This is displayed in Figure 19. At an even higher value, $F_3 = F_s$, map (2.7.7) undergoes a subcritical bifurcation which implies that whenever $F_3 > F_s$ there is no attractor at all so in this part of parameter space we simply find that points (x_1, x_2, x_3) are randomly distributed in state space. □

So far we have demonstrated that although the dynamics is a quasistationary orbit just beyond the original Hopf bifurcation threshold, the dynamical outcome may be a periodic orbit as we penetrate deeper into the unstable parameter region. Such a phenomenon may happen when $|\arg\lambda|$ is close to $\pi/4$ at bifurcation threshold (4-periodicity). Another possibility (among others!) is that $|\arg\lambda|$ is close to $2\pi/3$ (3-periocidity).

Note, however, that if $\arg\lambda$ is close to a "critical" value, say $\pi/2$, at bifurcation it does not necessarily imply that a periodic orbit is created when we continue to increase the bifurcation parameter. In fact, when the parameter is enlarged the location of the eigenvalues may move away from the imaginary axis, hence the periodicity will be less pro-

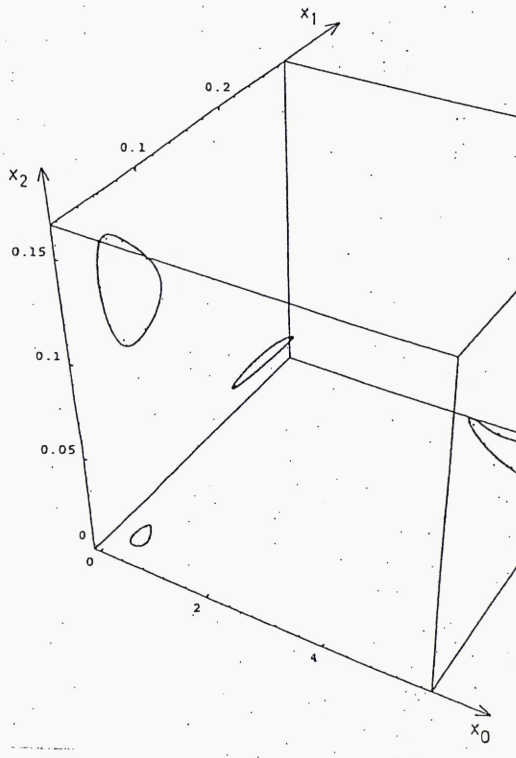


Figure 19: Map (2.7.7) after the secondary Hopf bifurcation.

nounced as the bifurcation parameter growths. In our next example there is no periodicity at all.

Example 2.7.3. Consider two-dimensional population map

$$(x_1, x_2) \rightarrow (F\,e^{-x_2}x_1 + F\,e^{-x_2}x_2, Px_1) \tag{2.7.11}$$

Hence, only the second age class x_2 contributes to density effects. As before, $F > 0$, $0 < P \leq 1$ and $F(1+P) > 1$.

We urge the reader to verify that the fixed point (x_1^*, x_2^*) may be written as

$$(x_1^*, x_2^*) = \left(\frac{1}{P}x_2^*, \ln[(1+P)F]\right) \tag{2.7.12}$$

and further that a (supercritical) Hopf bifurcation occurs at the threshold

$$F = F_H = \frac{1}{1+P}e^{(1+2P)/(1+P)} \tag{2.7.13}$$

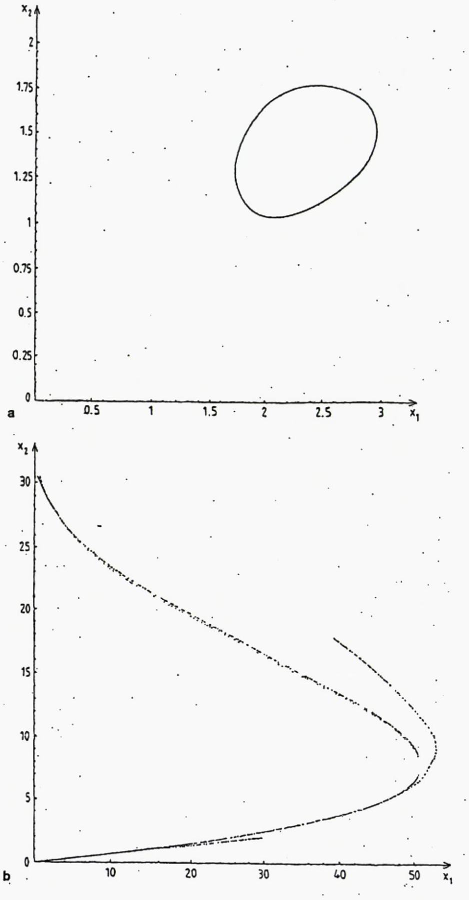


Figure 20: Dynamics generates by map (2.7.11). Parameter values: (a) $(P, F) = (0.6, 2.5)$; (b) $(P, F) = (0.6, 5.0)$.

and finally that the solution of the eigenvalue equation at threshold (2.7.13) becomes

$$\lambda = \frac{1}{2(1+P)} \left\{ 1 \pm \sqrt{4(1+P)^2 - 1}\, i \right\} \tag{2.7.14}$$

Now, assume that P is not close to zero. Then, the location of λ clearly suggests that frequency locking into an orbit of finite period will not take place. In Figure 20a we show the invariant curve just beyond the bifurcation threshold $(P, F) = (0.6, 2.5)$ and on that curve we find no tendency towards periodic dynamics.

As we continue to increase F (P fixed) the "radius" of the invariant curve becomes larger. Eventually, the invariant curve becomes kinked and signals that the attractor is not topological equivalent to a circle anymore and finally the curve breaks up and a chaotic attractor is born. This is exemplified in Figure 20b. □

In our final example all bifurcations that we have previously discussed are present.

Example 2.7.4. Referring to section 2.4, Examples 2.4.1 and 2.4.3 we showed that the fixed point (x_0^*, x_1^*) of map (2.4.2), i.e.

$$(x_0, x_1) \to (F_0\, e^{-\alpha x} x_0 + F_1\, e^{-\alpha x} x_1, P_0 x_0)$$

is stable in case of small equilibrium populations $x^* = x_0^* + x_1^*$ but eventually will undergo a supercritical Hopf bifurcation at the threshold

$$F = F_H = \frac{1}{1+P_0}\, e^{(1+2P_0)/P_0}$$

provided $1/2 < P_0 < 1$ and equal fecundities $F_0 = F_1 = F$. In Figure 21 we have generated the bifurcation of the map in the case $P_0 = 0.9$, $\alpha = 0.01$. The bifurcation parameter F is along the horizontal axis, the total population x along the vertical. Omitting computational details (which may be obtained in Wikan and Mjølhus (1996)) we shall now use Figure 21 in order to reveal the dynamics of (2.4.2).

In case of $5.263 < F < 10.036$ there is one attractor, namely the stable fixed point (x_0^*, x_1^*). (The lower limit 5.263 is a result of the requirement $F(1+P) > 1$.) At the threshold $F_s = 10.036$ a 3-cyclic attractor with large amplitude is created. Thus beyond F_s there exists a parameter (F) interval where there are two coexisting attractors and the ultimate fate of an orbit depends on the initial condition. It is a well known fact that multiple attractors indeed may occur in nonlinear systems. What happens in our case is that the third iterate of the original map (2.4.2) undergoes a saddle-node bifurcation at

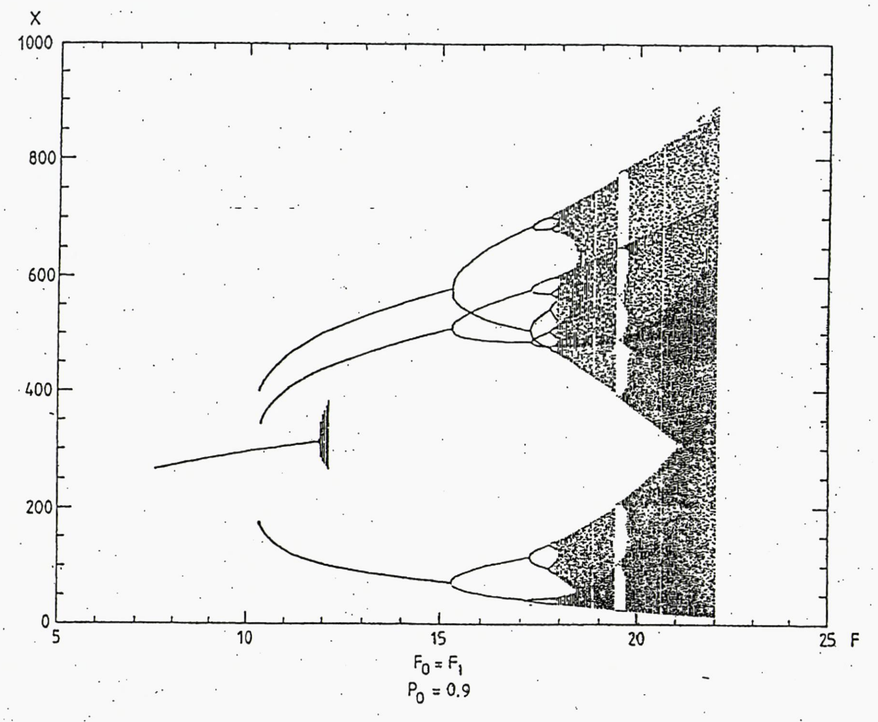


Figure 21: The bifurcation diagram generated by map (2.4.2).

F_s. This may be verified numerically by computing the Jacobian of the third iterate and show that the dominant eigenvalue of the Jacobian equals unity. Moreover (referring to section 1.5, see also Exercise 1.4.2 in section 1.4), a 3-cycle consisting of unstable points is also created through the saddle node at threshold F_s. This repelling 3-cycle is of course invisible to the computer.

In the interval $10.036 < F < 11.81$ the large amplitude 3-cycle and the fixed point are coexisting attractors. At $F_H = 11.81$ the fixed point undergoes a supercritical Hopf bifurcation (for a proof, cf. Wikan and Mjølhus (1996)), thus in case of $F > F_H$, $F - F_H$ small, there is coexistence between the 3-cylic attractor and a quasistationary orbit restricted to an invariant curve. The coexistence takes place in the interval $11.81 < F < 12.20$. In somewhat more detail we also find that since $\arg \lambda$ (where λ is the eigenvalue of the Jacobian of (2.4.2)) is close to $2\pi/3$ at F_H there is a clear tendency towards 3-periodic dynamics on the invariant curve but there is no frequency locking into an exact 3-periodic orbit.

At $F_K = 12.20$ the invariant curve disappears. Consequently, in case of $F > F_K$, there is again only one attractor, namely the attracting 3-cycle. The reason that the invariant curve disappears at threshold F_K is that it is "hit" by the three branches of the repelling 3-cycle. This phenomenon is somewhat akin to what is called a crisis in the chaos literature.

As we continue to increase F successive flip bifurcations occur, creating orbits of period $3 \cdot 2^k$, $k = 1, 2, ...$, in much of the same way as we have seen in earlier examples. Eventually an accumulation value F_a for the flip bifurcations is reached, and beyond that value the dynamics becomes chaotic. At first the chaotic attractor consists of three separate branches which are visited once every third iteration. When F is even more increased the branches merge together. □

Through our previous examples, which all share the common feature that the original (first) bifurcation is a Hopf bifurcation, we have experienced that the nonstationary dynamics beyond the instability threshold may indeed be different from map to map. In the following exercises even more possible dynamical outcomes are demonstrated.

Exercise 2.7.1. Consider the map

$$(x_0, x_1) \to (F_1 x_1, P_0(1 - \gamma\beta x)^{1/\gamma} x_0)$$

where $\beta > 0$, $\gamma \leq 0$.

a) Compute the nontrivial fixed point (x_0^*, x_1^*).

b) Assume that $\gamma > \gamma_c = -(F_1/2(1 + F_1)$ and show that the fixed point undergoes a Hopf bifurcation at the threshold

$$P_0 = \frac{1}{F_1}\left[1 + \gamma\frac{2(1 + F_1)}{F_1}\right]^{1/\gamma}$$

c) Assume that $\gamma > \gamma_c$ but $\gamma - \gamma_c$ small. Investigate numerically the dynamical outcomes when P_0 is fixed and F_1 is increased beyond the bifurcation threshold.

d) (difficult!) Show that the Hopf bifurcation is supercritical. □

Exercise 2.7.2. Consider the semelparous population model

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}_{t+1} = \begin{pmatrix} 0 & 0 & F_2 e^{-x} \\ P_0 & 0 & 0 \\ 0 & P_1 & 0 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix}_t$$

a) Show that the fixed point is

$$(x_0^*, x_1^*, x_2^*) = \left(\frac{1}{1+P_0+P_0P_1} x^*, \frac{P_0}{1+P_0+P_0P_1} x^*, \frac{P_0P_1}{1+P_0+P_0P_1} x^* \right)$$

where $x^* = \ln(P_0P_1F_2)$.

b) Compute the Jacobian and show that the eigenvalue equation may be cast in the form

$$\lambda^3 + \varepsilon\lambda^2 + P_0\varepsilon\lambda + P_0P_1\varepsilon - 1 = 0$$

where $\varepsilon = x^*/(1+P_0+P_0P_1)$.

c) Use the Jury criteria (2.1.16) and show that the fixed point is stable whenever

$$\varepsilon_4 < \varepsilon < \varepsilon_2$$

where

$$\varepsilon_4 = \frac{1+P_0-2P_0P_1}{P_0P_1(1-P_0P_1)} \quad \text{and} \quad \varepsilon_2 = \frac{2}{1-P_0+P_0P_1}$$

d) Use the result in c) and show that the fixed point is stable provided

$$\frac{1}{2} < P_0 < 1 \qquad P_1 > \frac{1+P_0}{3P_0}$$

e) The results from c) and d) are special in the sense that they imply that the fixed point is unstable in case of x^* (or F_2) small, becomes stable for larger values of x^* (or F_2) and then becomes unstable again through further enlargement of x^* (or F_2). Note that ε_4 and ε_2 are Hopf and flip bifurcation thresholds respectively. Investigate (numerically the dynamics in case of $\varepsilon < \varepsilon_4$ (i.e. x^* small) and $\varepsilon < \varepsilon_2$ (i.e. x^* large). (Hint: cf. Exercise 2.4.3.) □

Part III
Related Topics

3.1 The fundamental equation of discrete dynamic programming

In this section and in the next we shall give a brief introduction to discrete dynamic optimization. When one wants to solve problems within this field there are mainly two methods (together with several numerical alternatives which we will not treat here) available. Here, in section 3.1, we shall state and prove the fundamental equation of discrete dynamic programming which perhaps is the most frequently used method. In section 3.2 we shall solve optimization problems by use of a discrete version of the maximum principle.

Dynamic optimization is widely used within several scientific branches like economy, physics and biology. As an introduction to the kind of problems that we want to study, let us consider the following example:

Example 3.1.1. Let x_t be the size of a population at time t. Further, assume that x is a species of commercial interest so let $h_t \in [0,1]$ be the fraction of the population that we harvest at each time. Therefore, instead of expressing the relation between x at two consecutive time steps as $x_{t+1} = f(x_t)$ or (if the system is nonautonomous) $x_{t+1} = f(t, x_t)$, we shall from now on assume that

$$x_{t+1} = f(t, x_t, h_t) \tag{3.1.1}$$

If the function f is the quadratic or the Ricker function which we studied in Part I, (3.1.1) may be written as

$$x_{t+1} = r(1-h_t)x_t[1-(1-h_t)x_t] \tag{3.1.2}$$

or

$$x_{t+1} = (1-h_t)x_t \exp[r(1-(1-h_t)x_t)] \tag{3.1.3}$$

respectively. In case of an age-structured population model (cf. the various examples treated in part II) the equation $\mathbf{x}_{t+1} = f(t, \mathbf{x}_t, \mathbf{h}_t)$ may be expressed as

$$\begin{aligned} x_{1,t+1} &= F_1 e^{-x_t} x_{1,t}(1 - h_{1,t}) + F_2 e^{-x_t} x_{2,t}(1 - h_{1,t}) \\ x_{2,t+1} &= P x_{1,t}(1 - h_{2,t}) \end{aligned} \tag{3.1.4}$$

(For simplicity, it is often assumed that $h_t = h$ and $h_{i,t} = h_i$ which means that the population or the age classes are exposed to harvest with constant harvest rate(s).)

Now, returning to equation (3.1.1), assume that $\pi_t = f_0(t, x_t, h_t)$ is the profit we can make of the harvested part of the population at time t. Our ultimate goal is to maximize the profit over a time period from $t = 0$ to $t = T$, i.e. we want to maximize the sum of the profits at times $t = 0, 1, ..., T$. This leads to the problem

$$\underset{h_0, h_1, ..., h_T}{\text{maximize}} \sum_{t=0}^{T} f_0(t, x_t, h_t) \tag{3.1.5}$$

subject to equation (3.1.1) given the initial condition x_0 and $h_t \in [0, 1]$.

To be somewhat more precise, we have arrived at the following situation: Suppose that we at time $t = 0$ apply the harvest rate h_0. Then, according to (3.1.1) $x_1 = f(0, x_0, h_0)$ is known at time $t = 1$. Further, assume that we at time $t = 1$ choose the harvest h_1. Then $x_2 = f(1, x_1, h_1)$ is known and continuing in this fashion, applying (different) harvest rates h_t at each time we also know the value of x_t at each time. Consequently, we also know the profit $\pi_t = f_0(t, x_t, h_t)$ at each time. As stated in (3.1.5) our goal is to choose $h_0, h_1, ..., h_T$ in such a way that $\sum_{t=0}^{T} f_0(t, x_t, h_t)$ is maximized. □

—

Let us now formulate the situation described in Example 3.1.1 in a more general context. Suppose that the *state* variable x evolves according to the equation $x_{t+1} = f(t, x_t, u_t)$ where x_0 is known. At each time t the path that x follows depends on discrete *control* variables $u_0, u_1, ..., u_T$. (In Example 3.1.1 we used harvest rates as control variables.) We assume that $u_t \in U$ where U is called the control region. The sum $\sum_{t=0}^{T} f_0(t, x_t, u_t)$ where f_0 is the quantity we wish to maximize is called the objective function.

Definition 3.1.1. Suppose that $x_s = x$. Then we define the value function as

$$J_s(x) = \underset{u_s, u_{s+1}, ..., u_T}{\text{maximize}} \sum_{t=s}^{T} f_0(t, x_t, u_t) \tag{3.1.6}$$

□

Hence, a more general formulation of the problem we considered in Example 3.1.1 is: maximize $J_s(x)$ subject to $x_{t+1} = f(t, x_t, u_t)$, $x_s = x$ and $u_t \in U$.

We now turn to the question of how to solve the problem.

Suppose that we know the optimal control (optimal with respect to maximizing (3.1.6)) u_s^* at $s = 0$. Then, according to the findings presented in Example 3.1.1, we find the corresponding x_1^* as $x_1^* = f(0, x_0, u_0^*(x_0))$ and if we succeed in finding the optimal control $u_1^*(x_1^*)$ at time $t = 1$ we have $x_2^* = f(1, x_1^*, u_1^*(x_1^*))$ and so on. Thus, suppose that $x_s = x$ at time $t = s$, how shall we choose u_s in the best optimal way? Clearly, if we choose $u_s = u$ as the optimal control we achieve the immediate benefit $f_0(s, x, u)$ and also $x_{s+1} = f(s, x, u)$. This consideration simply means that the highest total benefit which is possible to get from time $s + 1$ to T is $J_{s+1}(x_{s+1}) = J_{s+1}(f(s, x, u))$. Hence, the best choice of $u_s = s$ at time s is the one that maximizes $f_0(s, x, u) + J_{s+1}(f(s, x, u))$. Consequently, we have the following theorem:

Theorem 3.1.1. Let $J_s(x)$ defined through (3.1.6) be the value function for the problem

$$\underset{u}{\text{maximize}} \sum_{t=0}^{T} f_0(t, x_t, u_t) \text{ subject to } x_{t+1} = f(t, x_t, u_t)$$

where $u_t \in U$ and x_0 are given. Then

$$J_s(x) = \max_{u \in U} \left[f_0(s, x, u) + J_{s+1}(f(s, x, u)) \right], \quad s = 0, 1, ..., T - 1 \tag{3.1.7}$$

$$J_T(x) = \max_{u \in U} f_0(T, x, u) \tag{3.1.8}$$

□

Theorem 3.1.1 is often referred to as the fundamental equation(s) of dynamical programming and serves as one of the basic tools for solving the kind of problems that we considered in Example 3.1.1. As we shall demonstrate through several examples, the theorem works "backwards" in the sense that we start to find $u_T^*(x)$ and $J_T(x)$ from (3.1.8). Then we use (3.1.7) in order to find $J_{T-1}(x)$ together with $u_{T-1}^*(x)$ and so on. Hence, all value functions and optimal controls are found recursively.

Example 3.1.2.

$$\underset{u}{\text{maximize}} \sum_{t=0}^{T} (x_t + u_t) \text{ subject to } x_{t+1} = x_t - 2u_t, \quad u_t \in [0,1], \quad x_0 \text{ given}$$

Solution: From (3.1.8), $J_T(x) = \max_u (x + u)$ so clearly, the optimal value of u is $u = 1$. Hence at time $t = T$, $J_T(x) = x + 1$ and $u_T^*(x) = 1$.

Further, from (3.1.7): $J_{T-1}(x) = \max_u [x + u + J_T(x - 2u)] = \max_u [x + u + (x - 2u + 1)] = \max_u [2x - u + 1]$. Consequently, $u = 0$ is the optimal choice, thus at $t = T - 1$ we have $J_{T-1}(x) = 2x + 1$ and $u_{T-1}^*(x) = 0$.

This implies: $J_{T-2}(x) = \max_u [x + u + J_{T-1}(x - 2u)] = \max_u [3x - 3u + 1]$ so again $u = 0$ is the best choice and $J_{T-2}(x) = 3x + 1$ and $u_{T-2}^*(x) = 0$.

From the findings above it is natural to suspect that in general

$$J_{T-k}(x) = (k+1)x + 1, \quad u_{T-k}^*(x) = 0, \quad k = 1, 2, ..., T$$

The formulae is obviously correct in case of $k = 1$ and by induction we have from (3.1.7) that

$$\begin{aligned}
J_{T-(k+1)} &= \max_u [x + u + J_{T-k}(x - 2u)] \\
&= \max_u [x + u + (k+1)(x - 2u) + 1] = \max_u [(k+2)x - 2(k+1)u + 1] \\
&= (k+2)x + 1 = [(k+1) + 1]x + 1
\end{aligned}$$

hence the formulae is correct at time $T-(k+1)$ as well. Therefore

$$J_{T-k}(x) = (k+1)x+1\,, \quad u^*_{T-k}(x) = 0\,, \quad k = 1,2,...,T$$
$$J_T(x) = x+1 \qquad u^*_T(x) = 1$$

□

Example 3.1.3.

$$\text{maximize}\sum_{t=0}^{T}(-u_t^2+u_t-x_t) \text{ subject to } x_{t+1} = x_t+u_t\,, \quad u_t \in \langle -\infty,\infty\rangle\,, \ x_0 \text{ given}$$

Solution: From (3.1.8), $J_T(x) = \max_u(-u^2-x+u)$ and since the function $h(u) = -u^2-x+u$ clearly is concave in u the optimal choice of u must be the solution of $h'(u) = 0$, i.e. $u = 1/2$. Hence, at time $t = T$, $u^*_T(x) = 1/2$ and $J_T(x) = -(1/4)-x+(1/2) = -x+(1/4)$.

Further, (3.1.7) gives $J_{T-1}(x) = \max_u[-u^2-x+u+J_T(x+u)] = \max_u[-u^2-x+u-(x+u)+(1/4)] = \max_u[-u^2-2x+(1/4)]$ and again since $h_1(u) = -u^2-2x+(1/4)$ is concave in u we find that $u=0$ is the optimal choice. Thus $J_{T-1}(x) = -2x+(1/4)$ and $u^*_{T-1}(x) = 0$.

Proceeding in the same way (we urge the reader to work through the details) we find that $J_{T-2}(x) = -3x+(1/2)$, $u^*_{T-2}(x) = -(1/2)$ and $J_{T-3}(x) = -4x+(3/2)$, $u^*_{T-3}(x) = -1$.

Therefore, it is natural to suppose that

$$J_{T-k}(x) = -(k+1)x+b_k$$

where $b_0 = 1/4$ and $u^*_{T-k}(x) = -\frac{k-1}{2}$, $k = 1,2,...,T$. The formulae is obviously correct when $k=0$ and by induction

$$\begin{aligned} J_{T-(k+1)} &= \max_u[-u^2+u-x+J_{T-k}(x+u)] \\ &= \max_u[-(k+2)x-u^2-ku+b_k] \end{aligned}$$

Again, we observe that the function inside the bracket is concave in u so its maximum occurs at $u = -(k/2)$ which means that the corresponding value function becomes

$$J_{T-(k+1)}(x) = -[(k+1)+1]x + b_k + k^2/4 = -[(k+1)+1]x + b_{k+1}$$

It remains to find b_k. The equation $b_{k+1} - b_k = k^2/4$ has the homogeneous solution $C \cdot 1^k = C$. Referring to the remark following Example 3.1.4 we assume a particular solution of the form $p_k = (A + Bk + Dk^2)k$. Hence, after inserting into the equation and equating terms of equal power of k we find that $A = 1/24$, $B = -(1/8)$ and $D = 1/12$ so the general solution becomes $b_k = C + (1/24)k - (1/8)k^2 + (1/12)k^3$. Finally, using the fact that $b_0 = 1/4$ which implies that $C = 1/4$, we obtain

$$J_{T-k}(x) = -(k+1)x + \frac{1}{24}(6 + k - 3k^2 + 2k^3) \text{ and } u^*_{T-k}(x) = -\frac{k-1}{2}$$

□

Example 3.1.4 (Exam exercise, UiO).

$$\underset{u}{\text{maximize}} \sum_{t=0}^{T} (x_t - u_t) \text{ subject to } x_{t+1} = u_t x_t\,, \quad u_t \in [0,2]\,, \ x_0 \text{ given}$$

Solution: $J_T(x) = \max_u (x - u)$. Clearly, $u = 0$ is the optimal choice so $J_T(x) = x$ and $u^*_T(x) = 0$, $J_{T-1}(x) = \max_u[x - u + J_T(ux)] = \max_u[x + (x-1)u]$. Thus, if $x \geq 1$ we choose $u = 2$ and if $x < 1$ we choose $u = 0$. Consequently,

$$J_{T-1}(x) = \begin{cases} x + (x-1)2 = 3x - 2 & \text{if } \ x \geq 1 \ \text{ and } \ u^*_{T-1}(x) = 2 \\ x + (x-1)0 = x & \text{if } \ x < 1 \ \text{ and } \ u^*_{T-1}(x) = 0 \end{cases}$$

(Note that $J_{T-1}(x)$ is a convex function which is continuous at $x = 1$.)

In order to compute $J_{T-2}(x)$ we must consider the cases $J_{T-1}(x) = 3x - 2$ and $J_{T-1}(x) = x$ separately.

Assuming $J_{T-1}(x) = 3x - 2$ we obtain

$$J_{T-2}(x) = \max_u [x - u + 3ux - 2] = \max_u [x + (3x-1)u - 2]$$

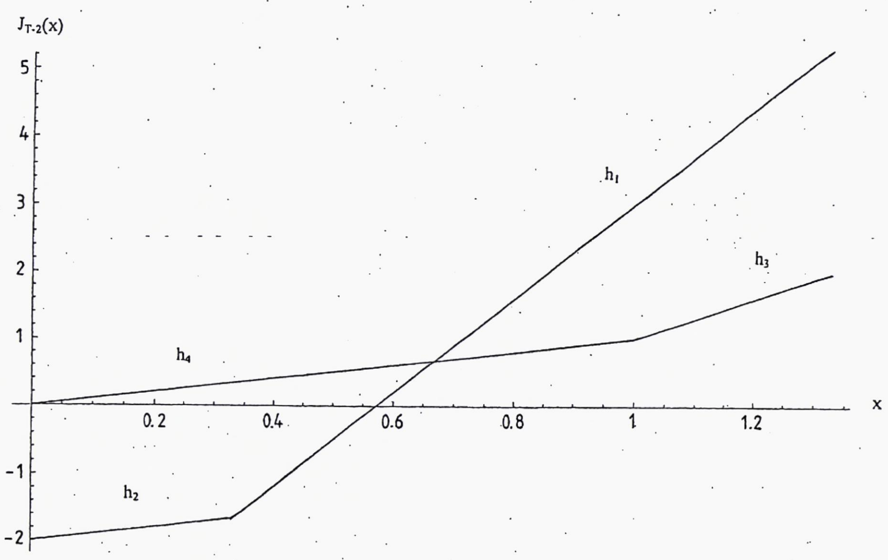


Figure 22: $J_{T-2}(x)$ possibilities.

so if $x \geq 1/3$ our optimal choice is $u = 2$ and if $x < 1/3$ we choose $u = 0$.

In the same way, using $J_{T-1}(x) = x$, we find

$$J_{T-2}(x) = \max_u [x - u + ux] = \max_u [x + (x-1)u]$$

so whenever $x \geq 1$, $u = 2$ and if $x < 1$ our best choice is $u = 0$.

Hence, the possibilities are

$$J_{T-2}(x) = \begin{cases} x + (3x-1)\cdot 2 - 2 = h_1(x) = 7x - 4 & \text{if } \ x \geq 1/3 \\ x + (3x-1)\cdot 0 - 2 = h_2(x) = x - 2 & \text{if } \ x < 1/3 \\ x + (x-1)\cdot 2 = h_3(x) = 3x - 2 & \text{if } \ x \geq 1 \\ x + (x-1)\cdot 0 = h_4(x) = x & \text{if } \ x < 1 \end{cases}$$

In Figure 22 we have drawn the graphs of the h_i functions in their respective domains. The point of intersection between $h_1(x)$ and $h_4(x)$ is $x = 2/3$ so clearly, if $x \geq 2/3$, $h_1(x)$ is the largest function. If $x < 2/3$, $h_4(x)$ is the largest function.

Consequently, we conclude that

$$J_{T-2}(x) = \begin{cases} 7x - 4 & \text{if } \ x \geq 2/3 \ \text{ and } \ u^*_{T-2}(x) = 2 \\ x & \text{if } \ x < 2/3 \ \text{ and } \ u^*_{T-2}(x) = 0 \end{cases}$$

and again we notice that $J_{T-2}(x)$ is a convex function which is continuous at $x = 2/3$.

Now at last, let us try to find the general expression $J_{T-k}(x)$. The formulaes for J_{T-1} and J_{T-2} suggest that our best assumption is

$$J_{T-k}(x) = \begin{cases} a_k x + b_k & x \geq \frac{b_k}{1-a_k} = c \\ x & x < \frac{b_k}{1-a_k} = c \end{cases}$$

$k = 1, 2, ..., T$ and that $u^*_{T-k}(x) = 2$ if $x \geq c$ and $u^*_{T-k}(x) = 0$ if $x < c$.

The formulae is certainly correct in case of $k = 1$. Further, by using the same kind of considerations as in the computation of $J_{T-2}(x)$ and induction there are two separate cases.

$$J_{T-(k+1)}(x) = \max_u [x - u + a_k u x + b_k] = \max_u [x + (a_k x - 1)u + b_k]$$

Hence $x \geq 1/a_k \Rightarrow u = 2$ and $x < 1/a_k \Rightarrow u = 0$, and

$$J_{T-(k+1)}(x) = \max_u [x - u + ux] = \max_u [x + (x-1)u]$$

Thus $x \geq 1 \Rightarrow u = 2$ and $x < 1 \Rightarrow u = 0$.

This yields (just as in the $J_{T-2}(x)$ case) the following

$$J_{T-(k+1)}(x) = \begin{cases} (2a_k + 1)x + b_k - 2 = a_{k+1}x + b_{k+1} = g_1(x) & x \geq 1/a_k \\ x + b_k = g_2(x) & x < 1/a_k \\ 3x - 2 = g_3(x) & x \geq 1 \\ x = g_4(x) & x < 1 \end{cases}$$

and we recognize that the forms of $g_1(x)$ and $g_4(x)$ are in accordance with our assumption and moreover that the point of intersection between $g_1(x)$ and $g_4(x)$ is $b_k(1 - a_k)^{-1}$ which also is consistent with the assumption.

Further, a_k obeys the difference equation $a_{k+1} = 2a_k + 1$. Therefore, the general solution is $a_k = D \cdot 2^k - 1$ and since $a_1 = 3 \Rightarrow D = 2$ we have $a_k = 2^{k+1} - 1$. In the same way, $b_{k+1} = b_k - 2$ (see the remark following this example, see also (1.1.2b)) has the general solution $b_k = K - 2k$ and since $b_1 = -2 \Rightarrow K = 0$ we obtain $b_k = -2k$.

Finally, since (1) $g_1(1) \geq g_3(1)$ and $a_{k+1} > 3$. (2) $g_4(x) > g_2(x)$ and (3) $g_1(x) > g_4(x)$ when $x > b_{k+1}(1 - a_{k+1})^{-1}$ (recall that $a_{k+1} > 3$) we obtain the general solution

$$J_{T-k}(x) = \begin{cases} (2^{k+1} - 1)x - 2k & x \geq \frac{k}{2^k - 1} \quad u^*_{T-k} = 2 \\ x & x < \frac{k}{2^k - 1} \quad u^*_{T-k} = 0 \end{cases}$$

□

Remark: Referring to section 2.1, Exercise 2.1.3, the difference equation $x_{t+2} - 5x_{t+1} - 6x_t = t \cdot 2^t$ has the homogeneous solution $C_1(-1)^t + C_2 6^t$ and since the exponential function 2^t on the right-hand side of the equation is different from both exponential functions contained in the homogeneous solution it suffices to assume a particular solution of the form $(At + B)2^t$ in this case. In Example 3.1.3 we had to solve an equation of the form $x_{t+1} - x_t = at^2$. The homogeneous solution is $C \cdot 1^t = C$ but since $at^2 = at^2 \cdot 1^t$ we have the same exponential function on both sides of the equation. Therefore, we must in this case assume a particular solution of the form $(A + Bt + Dt^2)t$. In the same way, if $x_{t+1} - x_t = bt$ we assume a particular solution $(A + Bt)t$ and finally, in the case $x_{t+1} - x_t = K$, assume a particular solution $A + Bt$ (cf. (1.1.2b)). □

Exercise 3.1.1. Let a be a positive constant and solve the problem

$$\max_u \sum_{t=0}^{T} (x_t + u_t) \text{ subject to } x_{t+1} = x_t - au_t\,, \quad u_t \in [0, 1]\,, \quad x_0 \text{ given}$$

□

Exercise 3.1.2. Solve the problem (Exam Exercise, UiO):

$$\max_u \sum_{t=0}^{T} (x_t - u_t) \text{ subject to } x_{t+1} = x_t + u_t\,, \quad u_t \in [0, 1]\,, \quad x_0 \text{ given}$$

(Hint: Use the remark following Example 3.1.4.) □

Exercise 3.1.3. Solve the problem:

$$\max_u \sum_{t=0}^{T} (x_t + 1) \text{ subject to } x_{t+1} = u_t x_t\,, \quad u_t \in [0,1]\,, \quad x_0 \text{ given}$$

□

3.2 The maximum principle (Discrete version)

When t is a continuous variable, most optimization problems are formulated and solved by use of the maximum principle which was developed by Russian mathematicians about 60 years ago. There is also a great variety of numerical methods within this field. The maximum principle, sometimes referred to as Pontryagin's maximum principle, is the cornerstone in the discipline called optimal control theory which may be regarded as an extension of the classical calculus of variation. An excellent treatment of various aspects of control theory may be found in Seierstad and Sydsæter (1987), see also Sydsæter et al. (2005). In this section we shall briefly discuss a discrete version of the maximum principle which offers an alternative way of dealing with the kind of problems presented in section 3.1.

Consider the problem

$$\text{maximize} \sum_{t=0}^{T} f_0(t, x_t, u_t)\,, \quad u_t \in U\,, \quad U \text{ convex} \tag{3.2.1}$$
$$\text{subject to } x_{t+1} = f(t, x_t, u_t),\ t = 0, 1, ..., T-1,\ x_0 \text{ given.}$$

together with one of the following terminal conditions

$$\text{a) } x_T \text{ free}, \quad \text{b) } x_T \geq X_T, \quad \text{c) } x_T = X_T \tag{3.2.2}$$

Thus, the problem that we consider here is somewhat more general than the one presented in section 3.1 due to the terminal conditions (3.2.2b,c).

Next, define the Hamiltonian by

$$H(t,x,u,p) = \begin{cases} f_0(t,x,u) + pf(t,x,u) & t < T \\ f_0(t,x,u) & t = T \end{cases} \tag{3.2.3}$$

where p is called the adjoint function.

Then we have the following:

Theorem 3.2.1 (The maximum principle, dicrete version). Suppose that (x_t^*, u_t^*) is an optimal sequence for problem (3.2.1), (3.2.2). Then there are numbers $p_0, ..., p_T$ such that

$$u_t^* \text{ maximizes } H_u'(t, x_t^*, u_t^*, p_t)u \quad \text{for} \quad u \in U \tag{3.2.4}$$

Moreover,

$$p_{t-1} = H_x'(t, x_t^*, u_t^*, p_t)\,, \quad t = 1, ..., T-1 \tag{3.2.5a}$$

$$p_{T-1} = f_{0x}'(T, x_T^*, u_T^*) + p_T \tag{3.2.5b}$$

and to each of the terminal conditions (3.2.2) we have the following transversal conditions

a) $p_T = 0$.

b) $p_T \geq 0$ ($= 0$ if $x_T^* > X_T$).

c) p_T no condition. □

Theorem 3.2.1 gives necessary conditions for optimality. Regarding sufficient conditions we have:

Theorem 3.2.2. Suppose that (x_t^*, u_t^*) satisfies all the conditions in Theorem 3.2.1 and in addition that $H(t,x,u,p)$ is concave in (x,u) for every t. Then (x_t^*, u_t^*) is optimal. □

Proof. Our goal is to show that

$$K = \sum_{t=0}^{T} f_0(t, x_t^*, u_t^*) - \sum_{t=0}^{T} f_0(t, x_t, u_t) \geq 0$$

Introducing the notation $f_0 = f_0(t, x, u)$, $f_0^* = f_0(t, x^*, u^*)$ and so on, it follows from (3.2.3) that

$$K = \sum_{t=0}^{T}(H_t^* - H_t) + \sum_{t=0}^{T} p_t(f_t - f_t^*)$$

Now, since H is concave in (x, u) we also have that $H - H^* \leq H_x'^*(x - x^*) + H_u'^*(u - u^*)$. Thus

$$K \geq \sum_{t=0}^{T} H_u'^*(u_t^* - u_t) + \sum_{t=0}^{T} H_x'^*(x_t^* - x_t) + \sum_{t=0}^{T-1} p_t(f_t - f_t^*)$$

Due to (3.2.4) and the concavity of H the first of the three sums above are equal or larger than zero. Indeed, suppose $u_t \in [u_0, u_1]$. If $u_t^* \in (u_0, u_1)$ then $H_u'^* = 0$. If $u_t^* = u_0$, then $H_u'^* \leq 0$ and $u_t^* - u_t \leq 0$ and finally, if $u_t^* = u_1$, $H_u'^* \geq 0$ and $u_t^* - u_t \geq 0$, hence in all cases $H_u'^*(u_t^* - u_t) \geq 0$.

Regarding the second and the third sum they may by use of (3.2.5a), (3.2.5b) and (3.2.1) be written as

$$\sum_{t=0}^{T-1} p_{t-1}(x_t^* - x_t) + (p_{T-1} - p_T)(x_T^* - x_T) + \sum_{t=0}^{T-1} p_t(x_{t+1} - x_{t+1}^*)$$
$$= p_T(x_T - x_T^*) = K1$$

Next, assume x_T free. Then from (3.2.6a), $p_T = 0$ which implies $K1 = 0$. If $x_T \geq X_T$, (3.2.6b) gives $p_T \geq 0$ and since $x_T \geq X_T$ we must have $K1 \geq 0$ if $x_T^* = X_T$. If $x_T^* > X_T$, $p_T = 0$, thus in either case $K1 \geq 0$. Finally, if $x_T = X_T$, $K1 = 0$. Therefore, whatever terminal condition (3.2.2), $K1 \geq 0$ which implies $K \geq 0$ so we are done. □

Example 3.2.1. Solve the problem given in Example 3.1.2 by use of Theorems 3.2.1 and 3.2.2.

Solution: From (3.2.3) it follows

$$H(t, x, u, p) = \begin{cases} x + u + p(x - 2u) & t < T \\ x + u & t = T \end{cases}$$

Consequently, whenever $t < T$, $H'_x = 1 + p$ and $H'_u = -2p$ and if $t = T$, $H'_x = H'_u = 1$.

By use of the results above, (3.2.5a,b) gives

$$p_{t-1} = 1 + p_t \quad t < T\,, \quad p_{T-1} = 1 + p_T$$

and since x_T is free, (3.2.6a) implies that $p_T = 0$ so $p_{T-1} = 1$.

The equation $p_{t-1} = 1 + p_t$ may be rewritten as $p_{t+1} - p_t = -1$ and its general solution is easily found to be $p_t = C - t$. Further, since $P_{T-1} = 1$ it follows that $1 = C - (T - 1)$. Thus $C = T$ so $p_t = T - t$ and we observe that $p_t > 0$ for every $t < T$.

From the preceding findings, (3.2.4) may be formulated as

$$\begin{array}{lll} u = u_t^* & \text{shall maximize} \;\; -2(T-t)u & t < T \\ u = u_T^* & \text{shall maximize} \quad\quad 1u & t = T \end{array}$$

Accordingly, we make the following choices: If $t = T$, choose $u_T^* = 1$. If $t < T$ (recall that $-2(T - t) < 0$), choose $u_t^* = 0$ for every t. Hence, we have arrived at the same conclusion as we did in Example 3.1.2.

A final observation is that the Hamiltonian is linear in (x, u) so H is also concave in (x, u). Consequently, (x_t^*, u_t^*) solves the problem (x_t^* is found at each t from the equation $x_{t+1}^* = x_t^* - 2u_t^*$ and x_0 is given). □

Example 3.2.2. Solve the problem

$$\underset{u}{\text{maximize}} \sum_{t=0}^{T} (x_t - u_t) \text{ subject to } x_{t+1} = x_t + u_t$$

$x_0 = 1$, $x_T = X_T$, $1 < X_T < T + 1$, $u_t \in [0, 1]$.

Solution:

$$H(t, x, u, p) = \begin{cases} x - u + p(x + u) & t < T \\ x - u & t = T \end{cases}$$

Therefore, whenever $t < T$, $H'_x = 1 + p$, $H'_u = -1 + p$ and if $t = T$, $H'_x = 1$ and $H'_u = -1$.

Further, (3.2.5b) gives $p_{T-1} = 1 + p_T$ and (3.2.5a) gives $p_{t-1} = 1 + p_t$ if $t < T$. Clearly (cf. our previous example), the latter difference equation has the general solution $p_t = C - t$ so p_t is a decreasing sequence of points.

From (3.2.4) it follows

$$\begin{array}{lllll} u = u_t^* & \text{shall maximize} & (-1 + p_t)u & t < T \\ u = u_T^* & \text{shall maximize} & -1u & t = T \end{array}$$

Thus at $t = T$ the optimal control is $u_T^* = 0$. In the case $t < T$ we have that if $p_t - 1 \geq 0$, then $u = u_t^* = 1$ and if $p_t - 1 < 0$, we choose $u_t^* = 0$.

First, assume $p_t - 1 \geq 0$ for all $t < T$. Then $u_t^* = 1$ and $x_{t+1}^* = x_t^* + 1$ which has the general solution $x_t^* = K + t$. $x_0^* = 1 \Rightarrow K = 1$, which means that $x_t^* = t + 1$. This implies that $x_T^* = T + 1$ but this is a contradiction since $X_T < T + 1$. Next, assume $p_t - 1 < 0$ for all $t \leq T$. Then $u_t^* = 0$. Thus, $x_{t+1}^* = x_t^*$ which has the constant solution $x_t^* = M$. Again we have reached a contradiction since $1 < X_T$.

Finally, let us suppose that there exists a time t_c such that whenever $t \leq t_c$, then $p_t - 1 \geq 0$ and in case of $t_c < t \leq T$, $p_t - 1 < 0$.

First, consider the case $t \leq t_c$. Then $x_{t+1}^* = x_t^* + 1$ so $x_t^* = K + t$. $x_0 = 1 \Rightarrow K = 1$, hence $x_t^* = t + 1$. If $t > t_c$ we have $x_{t+1}^* = x_t^*$. Hence, x_t^* is a constant, say $x_t^* = M$, and since $x_t^* = X_T$ it follows that $x_t^* = X_T$.

Thus,

$$\begin{array}{lllll} t \leq t_c\,, & p_t - 1 = C - t - 1 \geq 0 & x_t^* = t + 1 & u_t^* = 1 \\ t > t_c\,, & p_t - 1 = C - t - 1 < 0 & x_t^* = X_T & u_t^* = 0 \end{array}$$

It remains to determine t_c and the constant C. At time t_c, $C - t_c - 1 = 0$ so $C = t_c + 1$. Therefore, $p_t = t_c - t$. Further, from $x_{t_c+1} = x_{t_c} + u_{t_c}$ we obtain $X_T = t_c + 1 + 1$ so $t_c = X_T - 2$. Consequently, by use of the conditions in the maximum principle we have arrived at

$$\begin{array}{lll} x_t^* = t + 1 & u_t^* = 1 & 0 \leq t \leq X_T - 2 \\ x_t^* = X_T & u_t^* = 0 & X_T - 2 < t \leq T \end{array}$$

and $p_t = X_T - 2 - t$ for every t. Finally, since H is linear and concave in (x, u) it follows from Theorem 3.2.2 that we have obtained the solution. □

—

We close this section by looking at one extension only.

If we have a problem which involves several state variables $x_1, ..., x_n$ and several controls $u_1, ..., u_m$ we may organize them in vectors, say $\mathbf{x} = (x_1, ..., x_n)$, $\mathbf{u} = (u_1, ..., u_m)$ and reformulate problem (3.2.1), (3.2.2) as

$$\text{maximize} \sum_{t=0}^{T} f_0(t, \mathbf{x}_t, \mathbf{u}_t) \tag{3.2.7}$$

subject to $\mathbf{x}_{t+1} = \mathbf{f}(t, \mathbf{x}_t, \mathbf{u}_t)$, $\mathbf{x}_0$ given, $\mathbf{u}_t \in U$, and terminal conditions on the form

$$\text{a) } x_{i,T} \text{ free,} \quad \text{b) } x_{i,T} \geq X_{i,T}, \quad \text{c) } x_{i,T} = X_{i,T} \tag{3.2.8}$$

The associated Hamiltonian may in case of so-called "normal" problems be defined as

$$H(t, \mathbf{x}, \mathbf{u}, \mathbf{p}) = \begin{cases} f_0(t, \mathbf{x}, \mathbf{u}) + \sum_{i=1}^{n} p_i f_i(t, \mathbf{x}, \mathbf{u}) & t < T \\ f_0(t, \mathbf{x}, \mathbf{u}) & t = T \end{cases} \tag{3.2.9}$$

where $\mathbf{p} = (p_1, ..., p_n)$ is the adjoint function.

Then we may formulate necessary and sufficiently conditions for an optimal solution in the same way as we did in the one-dimensional case.

Theorem 3.2.3. Suppose that $(\mathbf{x}_t^*, \mathbf{u}_t^*)$ is an optimal sequence for problem (3.2.7), (3.2.8) with Hamiltonian defined as in (3.2.9). Then there exists $\mathbf{p}$ such that

$$\mathbf{u} = u_t^* \text{ maximizes} \sum_{i=1}^{m} \frac{\partial H}{\partial u_i}(t, \mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_t) u_i \tag{3.2.10}$$

Moreover

$$p_{i,t-1} = H'_{x_i}(t, \mathbf{x}_t^*, \mathbf{u}_t^*, \mathbf{p}_t)\,, \quad t = 1, ..., T-1 \tag{3.2.11a}$$

$$p_{i,T-1} = \frac{\partial f_0}{\partial x_i}(T, \mathbf{x}_t^*, \mathbf{u}_t^*) + p_{i,T} \tag{3.2.11b}$$

and

a) $p_{i,T} = 0$ if the terminal condition is (3.2.9a).

b) $p_{i,T} \geq 0$ ($= 0$ if $x^*_{i,T} > X_{i,T}$) (3.2.12)

if the condition is (3.2.9b).

c) $p_{i,T}$ free if condition (3.2.9c) applies.

Finally, if H is concave in $(\mathbf{x}, \mathbf{u})$ for each t then $(\mathbf{x}^*_t, \mathbf{u}^*_t)$ solves problem (3.2.7), (3.2.8). □

As usual, we end with an example.

Example 3.2.3. Solve the problem

$$\max \sum_{t=0}^{T} (-u_t^2 - 2x_t) \text{ subject to } x_{t+1} = \frac{1}{2} y_t\,, \; y_{t+1} = u_t + y_t$$

$x_0 = 2$, $y_0 = 1$, $u_t \in \mathbb{R}$, x_T free, y_T free.

Solution: Denoting the adjoint functions by p and q respectively, the Hamiltonian becomes

$$H(t, x, y, u, p, q) = \begin{cases} -u^2 - 2x + \frac{1}{2}py + q(u + y) & t < T \\ -u^2 - 2x & t = T \end{cases}$$

which implies

$$\begin{array}{llll} H'_x = -2 & H'_y = \frac{1}{2}p + q & H'_u = -2u + q & t < T \\ H'_x = -2 & H'_y = 0 & H'_u = -2u & t = T \end{array}$$

Then, from (3.2.11a) it follows that $p_{t-1} = -2$, $q_{t-1} = (1/2)p_t + q_t$ and since x_T, y_T is free, (3.2.12a) implies $p_T = q_T = 0$. Thus (3.2.11b) reduces to $p_{T-1} = -2$ and $q_{T-1} = 0$.

Consequently, $p_t = -2$ for each t and if we insert this result into the difference equation for q we easily obtain the general solution $q_t = C + t$. Moreover, since $q_{T-1} = 0$ it follows that $0 = C + T - 1$ so $C = 1 - T$ which means that $q_t = t - T + 1$.

Now, since the control region is open, it follows from (3.2.10) that $H_u'^* = 0$, thus $-2u_t^* + q_t = 0$ if $t < T$ and $2u_T^* = 0$ whenever $t = T$. Hence at time $t = T$, $u_T^* = 0$ and in case of $t < T$, $u_t^* = (1/2)q_t = 1/2(t - T + 1)$.

Therefore, the problem is in many respects already solved. Indeed, y_t^* is now uniquely determined from the relation $y_{t+1}^* = u_t^* + y_t^*$ (recall that $y_0 = 1$) and x_t^* is subsequently found from $x_{t+1}^* = (1/2)y_t^*$. We leave the details to the reader.

Finally, observe that the Hesse determinant of H $(t < T)$ may be written as

$$\begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{vmatrix}$$

so clearly $(-1)^1\Delta_1 \geq 0$, $(-1)^2\Delta_2 = 0$, $(-1)^3\Delta_3 = 0$ where Δ_i is all possible principal minors of order i respectively. Consequently H is concave in (x, y, u). (At time $t = T$ the result is clear.) □

Exercise 3.2.1. Solve Exercises 3.1.2 and 3.1.3 by use of the maximum principle. □

References

Bernardelli H. (1941) Population waves. *Journal of Burma Research Society* **31**, 1–18.

Caswell H. (2001) *Matrix Population Models.* Sinauer Ass., Inc. Sunderland, Massachusetts.

Collet P. and Eckmann J. P. (1980) Iterated maps on the interval as dynamical systems. *Progress on Physics,* Vol. 1, Birkhäuser – Boston, Boston.

Cushing J. M. (1998) An introduction to structured population dynamics. *SIAM,* Philadelphia.

Devaney R. L. (1989) *An Introduction to Chaotic Dynamical Systems.* Addison–Wesley Publishing Company, Inc.

Feigenbaum M. J. (1978) Quantitative universality for a class of nonlinear transformations. *J. Stat. Phys.* **19,** 25–52.

Frauenthal J. (1986) Analysis of Age-Structure Models. *Biomathematics,* Vol. 17, *Mathematical Ecology,* eds. Hallan T. G. and Levin S. A. Springer Verlag, Berlin, Heidelberg.

Grimshaw R. (1990) *Nonlinear Ordinary Differential Equations.* Blackwell Scientific Publications.

Guckenheimer J. (1977) On the bifurcation of maps of the interval. *Invent. Math.* **39,** 165–178.

Guckenheimer J. and Holmes P. (1983) *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields.* Springer Verlag.

Guckenheimer J., Oster G. F., and Ipaktchi A. (1977) The dynamics of density dependent population models. *J. Math. Biol.* **4,** 101–147.

Hartman P. (1964) *Ordinary Differential Equations.* Wiley, New York.

Horn R. A. and Johnson C. R. (1985) *Matrix Analysis.* Cambrigde University Press.

Katok A. and Hasselblatt B. (1995) *Introduction to the Modern Theory of Dynamical Systems.* Cambridge University Press.

Leslie P H. (1945) On the use of matrices in certain population mathematics. *Biometrika* **33,** 183–212.

Leslie P. H. (1948) Some further notes on the use of matrices in population mathematics. *Biometrika* **35,** 213–245.

Lewis E. G. (1942) On the generation and growth of a population. Sankhya: *The Indian Journal of Statistics* **6,** 93–96.

Marsden J. E. and McCracken M. (1976) *The Hopf Bifurcation and its Applications.* Springer Verlag, New York, Heidelberg.

May R. M. (1976) Simple mathematical models with very complicated dynamics. *Nature* **261,** 459–467.

Mjølhus E., Wikan A., and Solberg, T. (2005) On synchronization in semelparous populations. *J. Math. Biol.* **50,** 1–21.

Murray J. D. (1989) *Mathematical Biology.* Springer, Berlin, Heidelberg.

Seierstad A. and Sydsæter K. (1987) *Optimal Control Theory with Economic Applications.* North-Holland.

Singer D. (1978) Stable orbits and bifurcation of maps of the interval. *SIAM, J. Appl. Math.* **35,** 260–267.

Spiegel M. (1974) *Complex Variables.* Schaum's Outline Series, McGraw–Hill.

Sydsæter K. (2002) *Matematisk analyse,* Bind II, Universitetsforlaget (in Norwegian).

Sydsæter K., Hammond P., Seierstad A., and Strøm A. (2005) *Further Mathematics for Economic Analysis.* Prentice Hall.

Wan Y. H. (1978) Computations of the stability condition for the Hopf bifurcation of diffeomorphisms on $\mathbb{R}^2$. *SIAM, J. Appl. Math.* **34,** 167–175.

Wikan A. (1994) *Bifurcations, Nonlinear Dynamics, and Qualitative Behaviour in a Certain Class of Discrete Age-Structured Population Models.* University of Tromsø, Norway.

Wikan A. and Eide A. (2004) An analysis of a nonlinear stage-structured cannibalism model with application to the northeast Arctic cod stock. *Bull. Math. Biol.* **66,** 1685–1704.

Visste du at...

...Høgskolen i Harstad med sine ca 1.600 studenter er blant landets mindre høgskoler. Et godt studiemiljø og en uformell tone, med nærhet og god oppfølging av den enkelte student, preger høgskolen. Høgskolen ligger idyllisk og sentralt plassert ved byens havnepromenade.

...våre to institutter - økonomi-/samfunnsfag og helse-/sosialfag - gir flere studiemuligheter:

Bachelorutdanninger:

- Handel, service og logistikk
- Økonomi og administrasjon
- Økonomi og administrasjon m/revisjon
- Reiseliv og turisme
- Barnevernpedagog
- Sykepleie
- Vernepleie

Årsstudium:

- Statvitenskap

Videreutdanning/påbygging:

- Revisjon
- Advanced Marketing
- Ledelse
- Psykisk helsearbeid

...Høgskolen i tillegg tilbyr skiftende aktuelle kurs på betalingsvilkår (se www.hih.no). Høgskolen påtar seg også utvikling av skreddersydde kurs, samt utfører forsknings- og utredningsoppdrag for arbeids- og næringsliv. Våre ca 60 fagansatte innehar høy kompetanse og dekker et bredt faglig spekter: Økonomi/administrasjon, sosialøkonomi, ledelse/organisasjon, statistikk, statsvitenskap, sosiologi, spansk, matematikk, pedagogikk, spesialpedagogikk, vernepleie, psykologi, sykepleievitenskap, matematikk, jus.

Besøksadresse: Havnegata 5, 9400 Harstad.
Postadresse: Høgskolen i Harstad, 9480 Harstad
Telefon 77 05 81 00
E-post: postmottak@hih.no
Internett: www.hih.no

HØGSKOLEN I HARSTAD