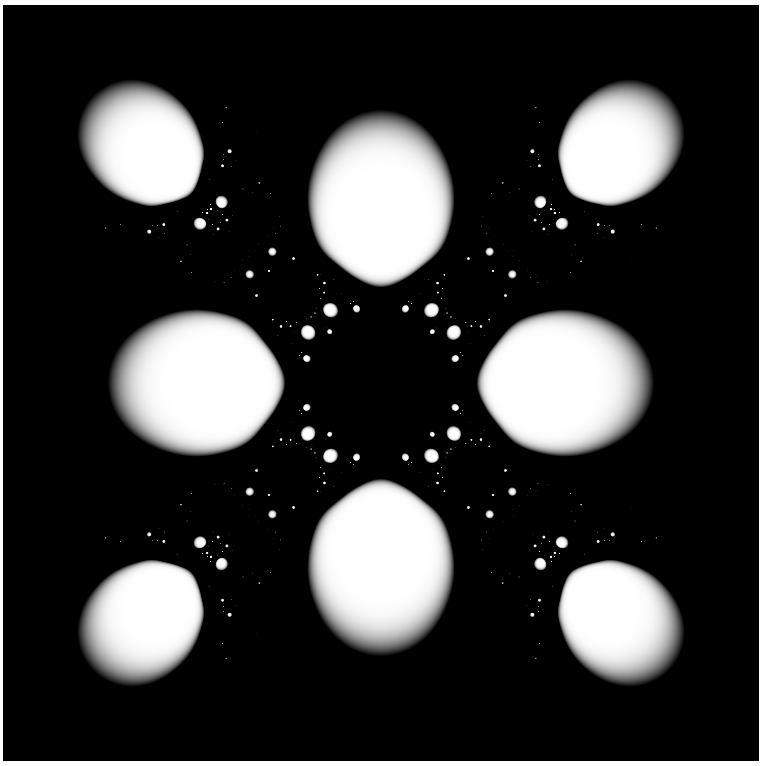
ALESSANDRO COLOMBO

NOTES OF NONLINEAR DYNAMICS



UPDATED ON 21/5/2025

these are the lecture notes from the course on nonlinear dynamics that i teach at politecnico di milano since 2019. these notes are continuously updated, and may contain errors. feel free to contact me (alessandro.colombo@polimi.it) if you notice any, or if you have suggestions to improve the contents of these notes.

Contents

Mathematics is the part of physics where experiments are cheap

(Arnold, 1997)

About these notes

These notes are an introduction to nonlinear dynamics, chaos and bifurcation theory, tailored for master students in Engineering with some background in linear systems theory. This document is not intended to be a substitute for a textbook on nonlinear dynamics; many excellent books already exist, and some are mentioned throughout these notes. Rather, these notes constitute a miniature list of topics that anybody interested in nonlinear dynamics should at least superficially understand, with just enough detail to give the reader the tools to construct a coherent picture and decide which topics to investigate further from standard textbooks.

The course from which these notes are taken is meant to give an introductory view of the main features of nonlinear systems and dynamics from the point of view of systems *analysis*, which is the qualitative and quantitative study of a system's behaviour, and to complement the notions of nonlinear control *synthesis* taught in other master courses. Along the way, we learn some notions of chaos and fractal theory, which are at once a fascinating and essential element of nonlinear dynamics.

The choice of topics touched on in this course is general enough that it should be of interest to students in most domains of Engineering and Physics, as well as any reader interested in understanding the mechanisms that decide the complex behaviour of natural and human-designed processes.

To help the reader through these notes, I have highlighted theorems, definitions, and remarks:

♦ **Theorem**
Theorems are highlighted like this. They form the rigorous structure of the theory we discuss in class.

★ Remarks are highlighted like this. They are not formal statements, but they capture details that are worth reading twice.

■ **Definition**
Definitions are highlighted like this. They set the stage for the subsequent discussion. Highlighting should make it easier to find them when you need to interpret a theorem or other concept.

Throughout sections you will find a set of examples that should help you motivate and understand the theory; at the end of each section, a set of exercises is given to test if you did. The answers to most exercises are not printed, so you will have to reason and convince yourself of the correctness of your answer. A printed answer is typically too strong a temptation. There are some exceptions to this rule, when exercises are especially challenging, or when I want to suggest a particular way to answer a question. Even in these cases, you should look at the answer only after you have spent some time trying to answer the question yourself. Challenging exercises are also marked by one or more stars (* **Exercise**).

On the right margin of some pages, you will find notes or figures. These are comments and, sometimes, technicalities that were left out of the main text and should not be necessary to follow it.

Introduction to dynamical systems

Keywords: **Dynamical systems, existence and uniqueness theorems, trajectory or orbit.**

Examples of nonlinear dynamical system

Example 1 (Pendulum). *Consider the following model of a pendulum with friction:*

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -\frac{g}{l}\sin(x_1) - \frac{k}{ml^2}x_2. \end{aligned}$$

The state $x := (x_1, x_2)$ *consists of the angular position* x_1*, and the angular velocity* x_2*, while the parameters* m*,* g*, and* l *are mass, gravity, and length of the arm.* $x_1 = 0$ *represents the pendulum at rest in the downward position, and* $x_1 = \pi$ *is the rest position in the upward position.* $\dot{x}$ *represents the derivative of* x *with respect to time.*

From the model equations, we easily see that both the state $(x_1 = 0, x_2 = 0)$ *and the state* $(x_1 = \pi, x_2 = 0)$ *imply* $\dot{x} = 0$*, that is, the system has two equilibria. Since linear systems can only have zero, one, or infinitely many equilibria, the full behaviour of this system and the relation between its two equilibria can only be studied in a nonlinear model.*

Example 2 (Repressilator). *The repressilator is an oscillating gene regulatory network that was first realised and discussed by Elowitz and Leibler, 2000. The network consists of three genes, each encoding a protein, each protein repressing the expression of the next gene in a loop.*

The action of the three genes, plus a fourth (GFP) that was used to express a fluorescent protein to measure oscillations, is represented by the following scheme, where the $\perp$ *symbol means 'represses'. The network was*

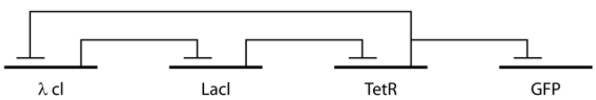


Figure 1: Image from Wikipedia

modelled by the following system of differential equations

$$\begin{aligned}
\dot{m}_{lacl} &= -m_{lacl} + \frac{\alpha}{1+p_{cl}} + \alpha_0, \\
\dot{p}_{lacl} &= -\beta(p_{lacl} - m_{lacl}), \\
\dot{m}_{tetR} &= -m_{tetR} + \frac{\alpha}{1+p_{lacl}} + \alpha_0, \\
\dot{p}_{tetR} &= -\beta(p_{tetR} - m_{tetR}), \\
\dot{m}_{cl} &= -m_{cl} + \frac{\alpha}{1+p_{tetR}} + \alpha_0, \\
\dot{p}_{cl} &= -\beta(p_{cl} - m_{cl}),
\end{aligned}$$

where m_i are concentrations of mRNA and p are concentrations of protein. A numerical simulation of the system shows that the concentrations of the proteins, for plausible values of the parameters, oscillate in time. This was

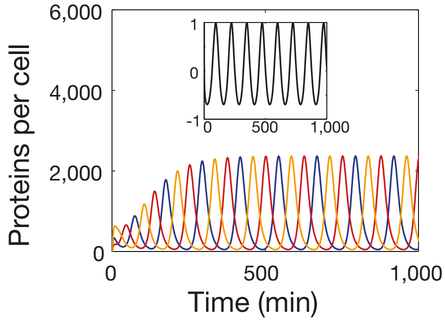


Figure 2: Image from Elowitz and Leibler, 2000

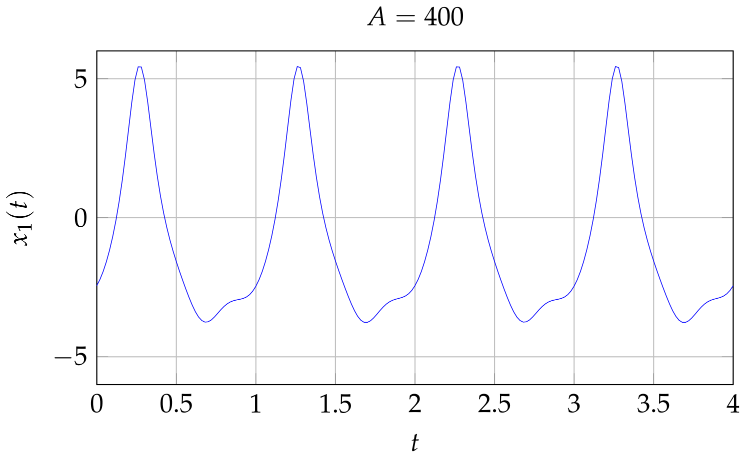
the first reported oscillatory network obtained by genetic engineering, designed from scratch from the above model and implanted in E. Coli bacteria. Unlike linear systems, this model oscillates on a unique, isolated limit cycle, regardless of the initial conditions. Therefore, after some transient, its output (measured as the density of the fluorescent protein) oscillates with a set frequency and amplitude, which is always the same and only depends on the model parameters.

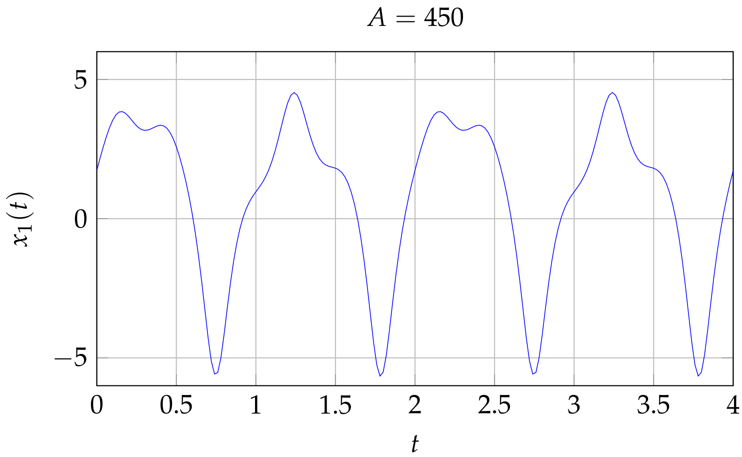
Example 3 (Mass-spring-damper with nonlinear spring)**.** *The following equations model a spring-mass-damper system with a nonlinear spring, whose elastic coefficient increases with increasing displacement:*

$$\begin{aligned}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= -x_1(1 + 10x_1^2) - x_2 + A\sin(2\pi t).
\end{aligned}$$

If the system were fully linear, we would expect to observe an oscillation of the position x_1 that is sinusoidal of frequency 1, just like the input. If, on top of the linear dynamics, we assume some nonlinear output function, we can expect to see this sinusoidal signal nonlinearly deformed at the output. We will therefore see an oscillation of frequency 1, together with its harmonics: oscillations with frequencies that are integer multiples of 1.

The system in this example however is not just a linear system with a nonlinear output function, it is a fully nonlinear dynamical system. Let us see what happens to $x_1(t)$ as A changes from 400 to 450.





The subharmonic, visible in this picture, is the result of a *period-doubling* bifurcation.

When $A = 400$, x_1 *is a periodic signal of period* 1, *as expected. When* $A = 450$, *however,* x_1 *has period* 2*! The nonlinear dynamics have generated a subharmonic of the periodic input.*

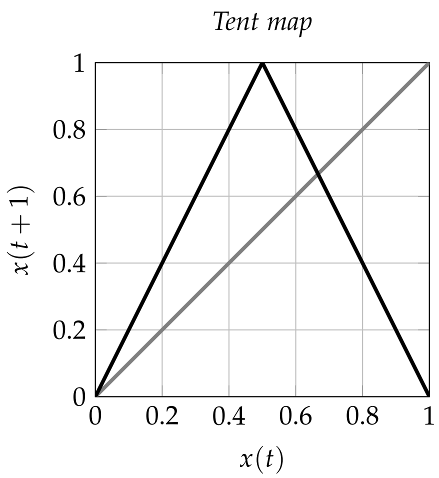
Example 4 (Tent map)**.** *Consider the following 1-dimensional discrete-time system:*

$$x(t+1) = \begin{cases} 2x(t), & x(t) \leq 0.5, \\ 2 - 2x(t), & x(t) > 0.5, \end{cases}$$

with $x \in [0, 1]$. *This is known as the tent map.*

As with all 1-dimensional discrete-time systems, we can represent its dynamics by plotting $f(x)$ *and the bisectrix of the first quadrant in the plane* $(x(t), x(t+1))$.

The tent map is, more precisely, the parametric family of systems $x(t+1) = \begin{cases} px(t,) & x(t) \leq 0.5, \\ p(1 - x(t)), & x(t) > 0.5. \end{cases}$



This kind of plots are also known as Lamerey diagram (Arnold, 1992)

The equilibria of the system are the points of intersection of $f(x)$ and the bisectrix. In this case, the system has two equilibria, and $f(x)$ is continuous but not everywhere differentiable.

Later on, we will learn that these two equilibria are both unstable, that is, they tend to repel nearby orbits. So where do solutions of this system go? We will see that this is a chaotic system: it has infinitely many unstable limit cycles, and a non-periodic orbit that is dense in $[0, 1]$.

Background

■ **Definition: Dynamical system (continuous time)**

A continuous time dynamical system is a system of equations

$$\dot{x} = f(x, t, u, p)$$

where

- $x \in \mathbb{R}^n$ is the state,
- $\dot{x}$ is the time-derivative of the state,
- $t \in \mathbb{R}$ is the time,
- $u \in \mathbb{R}^m$ is the input,
- $p \in \mathbb{R}^q$ are the parameters,
- f is the vector field.

n is called the *order* of the dynamical system

★ We will usually denote by x_i the i-th component of x, and by f_i the i-th equation of the system f.

Definition: Dynamical system (discrete time)
A discrete-time dynamical system is a system of equations

$$x(t+1) = f(x(t), t, u(t), p),$$

where

- $x \in \mathbb{R}^n$ is the state,
- $t \in \mathbb{Z}$ is the time,
- $u \in \mathbb{R}^m$ is the input,
- $p \in \mathbb{R}^q$ are the parameters,
- f is the vector field.

Definition: Classes of systems
A system where

- $f := f(x, t, p)$ is called **autonomous**,
- $f := f(x, u, p)$ is called **time-invariant**,
- $f := A(t, p)x + B(t, p)u$ is called **linear**,
- $x \in \mathbb{R}^n_+$, $\forall\, t \geq 0$ is called **positive**,
- solutions $x(t)$ are defined for both $t \geq 0$ and $t \leq 0$ is called **reversible**.

Unless otherwise stated, in this course we will deal with autonomous, time-invariant, continuous-time nonlinear systems. Often, when dependence on parameters is not relevant, we will also omit p among the arguments.

Definition: Phase space
The phase space is the set of all possible states of a dynamical system

In most cases, we will assume that the phase space is $\mathbb{R}^n$.

Definition: Flow of an autonomous system
The map $\phi_t(x) : \mathbb{R}^n \to \mathbb{R}^n$, seen as a map from the phase space to itself in the parameter t, which maps an initial state x to a state $x(t)$, is called the flow of the system.

This definition applies both to continuous and discrete-time systems.

We will use the same notation for the flow of discrete and continuous-time systems. Depending on the case, the value of t is a real or an integer number.

Definition: Trajectory or Orbit
A trajectory or orbit of x, which we call $\phi_\cdot(x)$, is the set of states spanned by $\phi_t(x)$ as t changes.

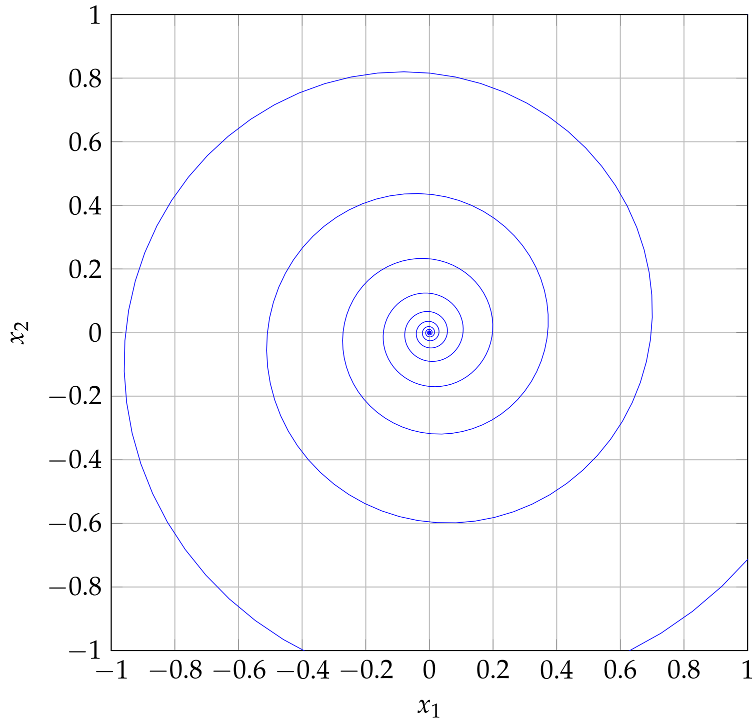
The term 'trajectory', with this meaning, is more often used in the control theoretic literature. In the nonlinear dynamics community the term 'orbit' is more commonly used.

Example 5 (examples of orbit)**.** *Here are some examples of orbits.*

In the following figure, we see an orbit converging to a stable focus in the system

$$\dot{x} = \begin{pmatrix} -1 & 10 \\ -10 & -1 \end{pmatrix} x.$$

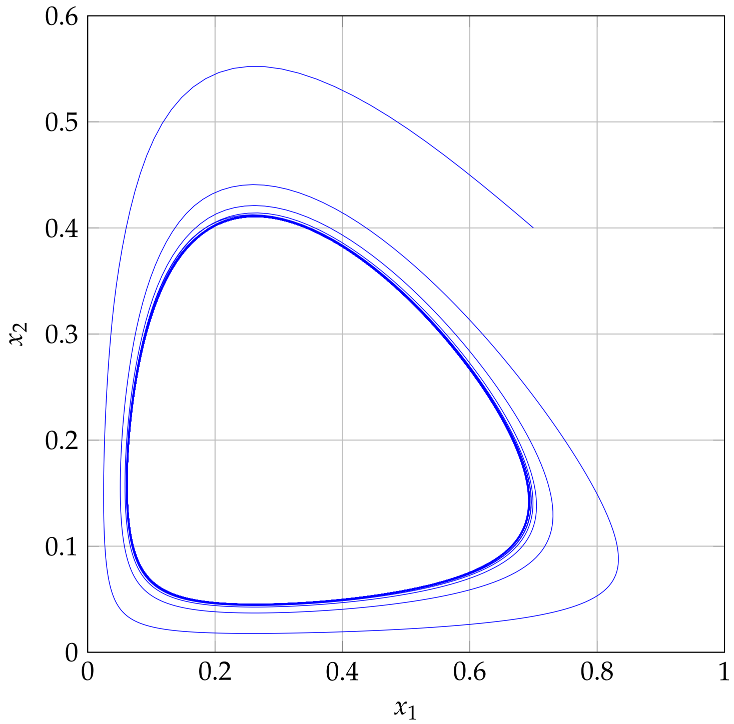
Notice that the orbit consists of all points of the spiral except the origin, which is only reached asymptotically. The origin is an orbit by itself.



In the next figure, we see an orbit converging to a stable cycle in the system

$$\dot{x}_1 = 7x_1(1 - x_1) - \frac{50x_1x_2}{1 + \frac{50}{15}x_1},$$
$$\dot{x}_2 = -7x_2 + \frac{50x_1x_2}{1 + \frac{50}{15}x_1}.$$

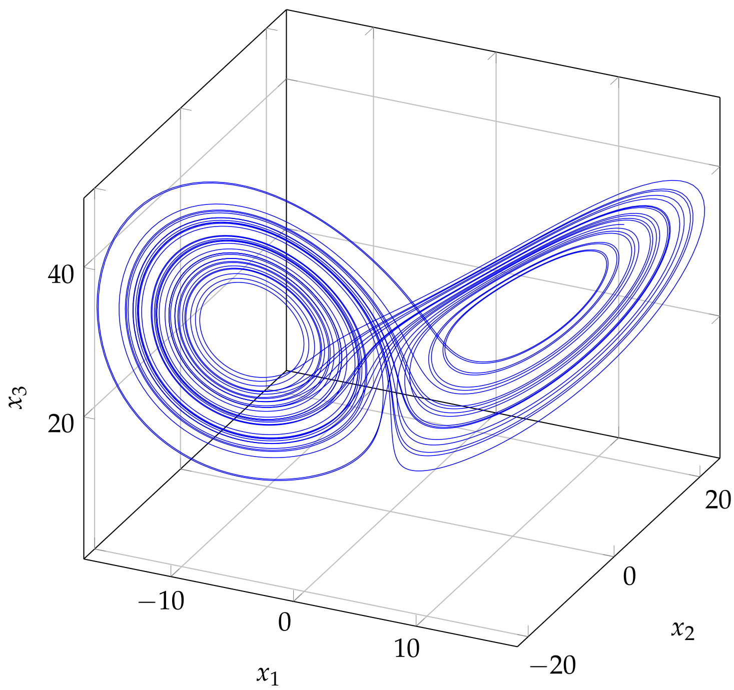
This is known as the Rosenzweig-MacArthur model (Rosenzweig and MacArthur, 1963)



Once again, the cycle is not part of this orbit and is an orbit by itself.

Finally, the following figure shows the chaotic orbit in Lorenz's system

$$\begin{aligned}\dot{x}_1 &= 10(x_2 - x_1),\\ \dot{x}_2 &= x_1(28 - x_3) - x_2,\\ \dot{x}_3 &= x_1 x_2 - \frac{8}{3}x_3.\end{aligned}$$



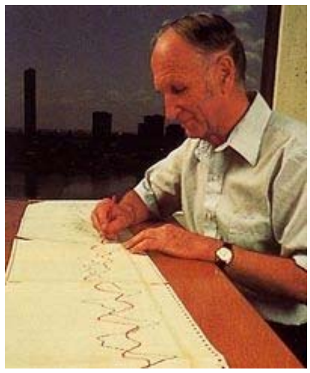


Figure 3: Edward Lorenz, by the window at MIT's Green Building (then dept. of meteorology), studying the first plots of the attractor.

Existence and uniqueness of solutions

In the case of continuous-time systems, determining the set of values of t where $\phi_t(x)$ is defined is not trivial. In particular, depending on the continuity properties of the vector field, we can ensure the existence of $\phi_t(x)$ in an interval around $t = 0$, or globally for all t. These conditions are specified in the theorems that follow.

Note that the same problem is much simpler for discrete-time systems, as long as we only care about $t \geq 0$.

■ **Definition: Lipschitz function**
$f(x)$ is Lipschitz if there is a K such that, for all $x, y \in \mathbb{R}^n$,

$$\|f(x) - f(y)\| \leq K\|x - y\|;$$

$f(x)$ is locally Lipschitz if for any $x \in \mathbb{R}^n$ there exists a neighbourhood of x where $f(x)$ is Lipschitz (K can be different in different neighbourhoods).

In the formula, the norm $\|\cdot\|$ is any vector norm.

We will sometimes use the term *globally Lipschitz* to distinguish functions that not just locally Lipschitz.

We can extend the above definition to time-dependent vector fields as follows.

■ **Definition: Uniformly Lipschitz function**
$f(x, t)$ is uniformly Lipschitz [resp. uniformly locally Lipschitz], if it is Lipschitz [resp. locally Lipschitz] for all fixed t

Consider now a continuous-time system.

♦ **Picard-Lindelöf existence and uniqueness theorem**
Let $f(x)$ be locally Lipschitz in $\mathbb{R}^n$. Then the flow $\phi_t(x)$ exists and is unique in a compact neighbourhood of $t = 0$.

The complete theorem actually provides an estimate of the size of the t-neighbourhood, based on the size of the x-neighbourhood. See e.g. (Meiss, 2007)

In other words, the above theorem is stating that, if $f(x)$ is locally Lipschitz, then a single orbit passes through $\bar{x}$: different orbits don't intersect. Moreover, given that a neighbourhood must (by definition!) contain x in its interior, the orbit exists both for $t > 0$ and $t < 0$, provided that $|t|$ is small enough. Note that functions that are continuously differentiable on a compact and convex set are also Lipschitz in the set [1], therefore continuous differentiability can be substituted to Lipschitz continuity to obtain a slightly more restrictive version of the result. Finally, note that the above theorem can be extended to time-variant vector fields $f(x, t)$, by asking that the vector field be uniformly locally Lipschitz, and continuous in t in a compact neighbourhood of $t = 0$.

[1] see e.g., Arnold, 1992, p273

Most of the tools that we study in this course refer to the behaviour of a system in some neighbourhood of one of its orbits, for instance, an equilibrium or a periodic orbit. In all these cases, the Picard-Lindelöf theorem is all we need to ensure that orbits are well-behaved in such a neighbourhood. In some cases, we may want to prove that $\phi_t(x)$ exists for all $t \in \mathbb{R}$. In these cases, there are different ways forward, the two following theorems discuss in particular the case of a time-invariant system.

Any vector field f that is locally Lipschitz *unbounded* can be made bounded by the time-reparametrization

$$\frac{dy}{d\tau} = \frac{f(y)}{1+|f(y)|}.$$

♦ **Bounded global existence**
If $f(x)$ is locally Lipschitz and bounded, then $\phi_t(x)$ exists for all $t \in \mathbb{R}$.

♦ **Global existence**
If $f(x)$ is (globally) Lipschitz, then $\phi_t(x)$ exists for all $t \in \mathbb{R}$.

Given all of the above results, one may be interested in understanding how, in the lack of the conditions for global existence, the flow may fail to exist for some finite t. We will see in Exercise 2 an example of how this happens. In general, it is useful to know the following.

♦ **Theorem**
If $\phi_t(x) \in D$ does not exist for all $t \geq 0$, then $\phi_t(x)$ must escape any compact neighbourhood of x for sufficiently large t.

See e.g. Theorem 3.3 in Khalil, 2002.

In other words, if the flow does not exist for all t, this is caused by the solution going to infinity.

All the above theorems provide conditions for the existence of solutions in a neighbourhood of $t = 0$, that is, the solution is well-defined both forward and backwards in time. When any one of these results holds, we can say that the system is reversible:

■ **Definition: Reversibility**
A system is *reversible* if, for all x, $\phi_t(x)$ exists and is unique in an open neighbourhood of $t = 0$.

★ A continuous-time system that satisfies the conditions of any one of the above theorems is reversible.

As we said before, the existence and uniqueness of the solutions of a discrete-time system forward in time is not an issue. However, in general, discrete-time systems do not have unique solutions backwards in time. This is unless $f(x)$ is injective. We thus have the following fact

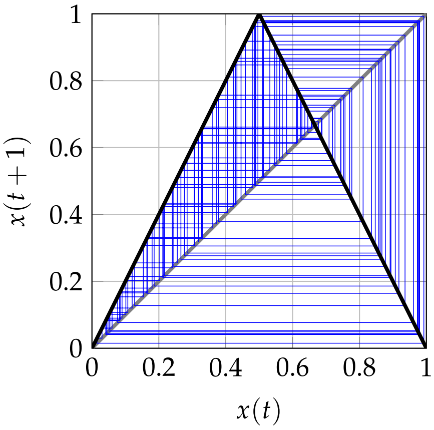
★ Discrete-time systems are, in general, not reversible.

Given our definition of orbit and the concept of reversibility, we can observe that

★ If a system is reversible, its orbits form a partition of the phase space.

Example 6. *The tent map:*

$$x(t+1) = \begin{cases} 2x(t), & x(t) \leq 0.5, \\ 2 - 2x(t, & x(t) > 0.5, \end{cases}$$



which we have seen before, is a 1-dimensional not reversible discrete-time system: states that are symmetric with respect to $x = 0.5$ *are mapped to the same state.*

Exercises

Exercise 1

Consider a stable linear system with a sinusoidal input of period 1, and an output function $y = g(x)$, with g possibly nonlinear. Explain why y can only have frequency components that are integer multiples of 1.

Exercise 2

Discuss the existence and uniqueness of the flow of

$$\dot{x} = x^2.$$

Answer of exercise 2

Take any fixed $\bar{x}$ and a compact neighbourhood N of $\bar{x}$. We have

$$|f(y) - f(x)| \leq \max_{x \in N} \frac{\partial f}{\partial x} |y - x|.$$

Therefore, $f(x)$ is Lipschitz on any compact neighbourhood, so by the Picard-Lindelöf theorem $\phi_t(x)$ exists and is unique in a neighbourhood of $t = 0$.

However, $f(x)$ is not globally Lipschitz and is not bounded, so $\phi_t(x)$ does not exist for all $t \in \mathbb{R}$.

In this simple case, we can see what happens by explicitly solving the differential equation. Integrating by parts we have

$$\frac{dx}{x^2} = dt,$$

which gives

$$-\frac{1}{x} = t + c \Rightarrow x = \frac{1}{\frac{1}{x(0)} - t}.$$

If $x(0)$ is negative, $x(t)$ tends to 0 as $\frac{1}{t}$. If, however, $x(0)$ is positive, $x(t)$ goes to ∞ as $t \to \frac{1}{x(0)}$. That is, the orbit is defined on the full real line, but the flow is only defined over a bounded time interval. A similar reasoning holds for $t \to -\infty$.

Exercise 3

Consider the system

$$\dot{x} = x^2 - 1,$$

with state defined in the closed interval $[-1, 1]$. How many orbits does this system have?

Exercise 4

Prove that the flow of the linear system

$$\dot{x} = Ax$$

exists for all t.

Hint: use the induced matrix norm $\|A\|$.
Note that, if $\|\cdot\|$ is the Euclidean norm, then $\|A\|$ is the largest singular value of A, that is, the square root of the largest eigenvalue of $A^T A$.

Exercise 5

Prove that the flow of

$$\dot{x} = |x|$$

exists for all t.

Exercise 6

Determine which of the following continuous and discrete-time systems are reversible:

- $\dot{x} = x^2$,
- $\dot{x} = \cos(x) - x$,
- $x(t+1) = x^2$,
- $x(t+1) = x^3$,
- $x(t+1) = x^3 - x$.

Exercise 7

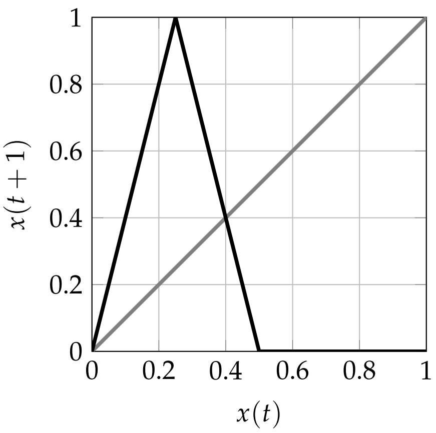
Consider the control system

$$\begin{aligned}
\dot{x}_1 &= x_1 + u, \\
\dot{x}_2 &= -2x_1 + 2x_2 + x_1^3 + u, \\
u &= -18x_1 + 12x_2.
\end{aligned}$$

1. Is it a continuous or discrete time system?
2. Is it autonomous or non-autonomous?
3. Is it time-invariant?
4. Is it linear or nonlinear?
5. Which of these sets could be a phase space?
   - $x_1 \in \mathbb{R}$,
   - $(x_1, x_2) \in \mathbb{R}^2$,
   - $(x_1, x_2, u) \in \mathbb{R}^3$.

Exercise 8

Compute, graphically, the evolution of this map starting from each of the equilibria



Exercise 9

Let $F(t)$ be the t-th term of the Fibonacci sequence: $1, 1, 2, 3, 5, 8, 13 \ldots$ The sequence is an orbit, for suitable initial conditions, of the x_2 variable of the discrete-time dynamical system

$$x_1(t+1) = x_2(t)$$
$$x_2(t+1) = x_1(t) + x_2(t)$$

Discuss the truth of the following statement:

$$\lim_{t \to \infty} \frac{c^t}{F(t)} = 0, \ \forall c \in \mathbb{R},$$

i.e., the Fibonacci sequence grows faster than any exponential.

Equilibria and fixed points

Keywords: **Equilibrium, fixed point, Lyapunov stability, asymptotic stability, GAS, positively invariant set, Lyapunov theorem, LaSalle invariance principle.**

Definition of equilibrium and fixed points

We have discussed, in the previous chapter, the conditions that guarantee the existence of the flow and the orbit of a dynamical system. It is now time to learn to characterize the behaviour of these solutions, and we start by investigating systems near the simplest kind of orbits, those where the flow is constant. We call these equilibria or fixed points, reserving the use of these two names to continuous and discrete-time systems, respectively.

■ **Definition: Equilibrium**
An equilibrium of an autonomous continuous-time system is a state x where $f(x) = 0$.

■ **Definition: Fixed point**
A fixed point of an autonomous discrete-time system is a state x where $f(x) = x$.

■ **Definition: Equilibrium or fixed point of non-autonomous systems**
An equilibrium [resp., fixed point] of a non-autonomous system is a pair (x, u), with u constant, where $f(x, u) = 0$ [resp., where $f(x, u) = x$].

In the linear world, the Rouchè Capelli theorem limits most systems to have a unique equilibrium.

★ **Equilibria and fixed points of a linear system**
The continuous-time linear system

$$\dot{x} = Ax + Bu,$$

with $u(t) = \bar{u}$ has a unique equilibrium in $-A^{-1}B\bar{u}$, provided that A is non-singular, otherwise it has infinitely many equilibria.
The discrete-time linear system

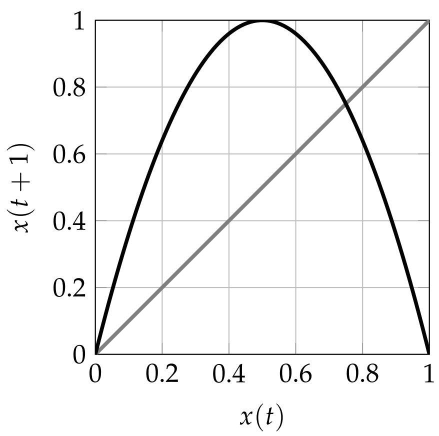
$$x(t+1) = Ax(t) + Bu(t),$$

with $u(t) = \bar{u}$ has a unique fixed point in $-(A-I)^{-1}B\bar{u}$, provided that $A - I$ is non-singular, otherwise it has infinitely many fixed points.

The analysis of equilibria and fixed points of nonlinear systems is, however, significantly more complex. Take for instance the following models:

Example 7. *Consider the logistic map*

$$x(t+1) = 4x(t)(1 - x(t))$$



Its fixed points are the solutions of

$$x = 4x - 4x^2.$$

Point $x = 0$ is obviously a fixed point, while a second fixed point is $x = \frac{3}{4}$.

Example 8. *The Rosenzweig-MacArthur model of a prey-predator food chain is*

$$\dot{x}_1 = rx_1\left(1 - \frac{x_1}{K}\right) - \frac{ax_1x_2}{1 + ahx_1},$$
$$\dot{x}_2 = -dx_2 + \frac{eax_1x_2}{1 + ahx_1},$$

where x_1 is the total mass of prey and x_2 is the total mass of predators.

The equilibria of the system are all pairs (x_1, x_2) where both $f_1 = 0$ and $f_2 = 0$. Predictably, $x_1 = 0$ and $x_2 = 0$ are an equilibrium (a food chain with no animals remains such). Also, $x_2 = 0$ and $x_1 = K$ is an equilibrium *(a food chain with no predators, where the prey reaches its carrying capacity K).*

A third equilibrium is a little more challenging to compute. From the law of $\dot{x}_2$, assuming $x_2 \neq 0$, we have

$$\bar{x}_1 = \frac{d}{ea - dah}.$$

Moreover, from this same equation, we see that $\frac{ax_1}{1+ahx_1} = \frac{d}{e}$. Substituting this in the law of $\dot{x}_1$ we have

$$rx_1\left(1 - \frac{x_1}{K}\right) - \frac{d}{e}x_2 = 0,$$

so

$$\bar{x}_2 = \frac{e}{d}r\bar{x}_1\left(1 - \frac{\bar{x}_1}{K}\right).$$

When this equilibrium is positive, it represents a state where prey and predators coexist.

In the above example, we have seen how the computation of the third equilibrium requires significantly more effort than in the case of a linear system; things are even worse than it may look:

★ There is no explicit formula for the computation of all equilibria of an arbitrary nonlinear system.

Therefore, even determining the number and location of equilibria can be a challenging task.

Stability of equilibria and fixed points

When we study the equilibria or the fixed points of a dynamical system, one of the first properties we typically want to ascertain is its *stability*. Loosely speaking, this indicates whether a small perturbation from the equilibrium should evolve towards the equilibrium, or away from it. The practical implication of stability is obvious: a system's state will tend to stay close to its stable equilibria or fixed points, even in the presence of a little bit of noise, but will diverge from its unstable equilibria or fixed points as soon as the state is affected by an arbitrarily small perturbation.

■ **Definition: Lyapunov stability**

An equilibrium or fixed point $\bar{x}$ is Lyapunov stable if, for every neighbourhood N of $\bar{x}$, there exists a neighbourhood $M \subset N$ such that

$$x \in M \Rightarrow \phi_t(x) \in N, \forall\, t \geq 0.$$

Probably the first mathematical model of population dynamics was Fibonacci's model of a population of rabbits (Fibonacci, 1202)(chapter XII, part 7.30), which in modern discrete-time systems theory is described by the system

$$x_1(t+1) = x_2(t)$$
$$x_2(t+1) = x_1(t) + x_2(t).$$

Here, $x_1(t)$ represents the number of pairs of juvenile rabbits at month t, and $x_2(t)$ is the number of pairs of adult rabbits.

The next relevant step was taken by Thomas Robert Malthus, who wrote an influential essay on population dynamics (Malthus, 1798), stating that when free of the constraints of limited resources, populations tend to grow exponentially. Nowadays we would formalize this statement as

$$\dot{x} = rx,$$

which is indeed known as the Malthusian growth model.

An improved mathematical model, accounting for the effects of limited resources on population growth, was then discussed by Verhulst (1845), who formulated the logistic growth model

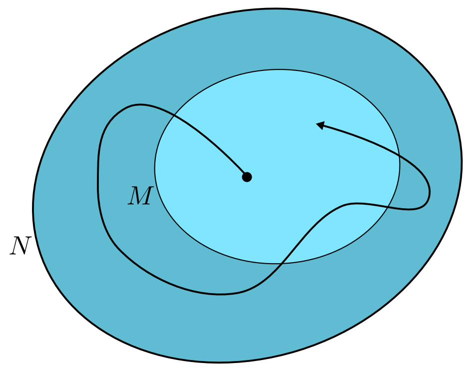
$$\dot{x} = rx\left(1 - \frac{x}{k}\right).$$

The interaction of a prey and a predator species was first modelled with differential equations by Volterra (1926), and the same equations were studied by Lotka (1920) in the context of chemical reactions. The model was

$$\dot{x}_1 = ax_1 - bx_1x_2$$
$$\dot{x}_2 = cx_1x_2 - dx_2.$$

The model by Rosenzweig and MacArthur (1963) is an evolution of the Lotka-Volterra, that utilizes a logistic growth model for the prey x_1, and an improved model for the prey-predator interaction, known as the Holling type-II functional response. Note that, when $x_2 = 0$, the dynamics of the prey is the logistic equation $\dot{x}_1 = rx_1\left(1 - \frac{x_1}{K}\right)$.

Aleksandr Lyapunov [1857-1918]



Definition: Unstable equilibrium or fixed point
An equilibrium or fixed point that is not Lyapunov stable is unstable.

Definition: Asymptotic stability
An equilibrium or fixed point $\bar{x}$ is asymptotically stable if it is Lyapunov stable, and there exists a neighbourhood N of $\bar{x}$ such that

$$\lim_{t \to \infty} \phi_t(x) = \bar{x}, \; \forall \, x \in N.$$

Definition: Global asymptotic stability (GAS)
An equilibrium or fixed point $\bar{x}$ is globally asymptotically stable (GAS) if it is asymptotically stable with neighbourhood $N = \mathbb{R}^n$.

There is yet another definition of stability, sometimes used in control theory.

Definition: Global exponential stability
A GAS equilibrium or fixed point $\bar{x}$ is globally exponentially stable if there exist constants $a > 0$, $b > 0$ such that

$$\|x(t) - \bar{x}\| \leq ae^{-bt}\|x(0) - \bar{x}\|$$

for all $t > 0$.

Global exponential stability is a stronger form of GAS, as it guarantees a bound on the speed of convergence of x to $\bar{x}$. However, while all other forms of stability that we have seen so far characterize a given equilibrium regardless of a choice of variables (i.e., regardless of how we choose to represent a system), global exponential stability does instead depend on the choice of variables, as the following example shows.

Example 9. *Consider the system*

$$\dot{x} = -x,$$

which obviously has a single equilibrium in 0 *that is GAS and exponentially stable over the whole real line, therefore is globally exponentially stable. Take the new variable*

$$y = \begin{cases} \ln(x+1), & x \geq 0, \\ -\ln(-x+1), & x < 0. \end{cases}$$

The inverse change of variables is

$$x = \begin{cases} e^y - 1, & y \geq 0, \\ 1 - e^{-y}, & y < 0. \end{cases}$$

This gives

$$\dot{y} = \begin{cases} \frac{\dot{x}}{x+1} = e^{-y} - 1, & y > 0, \\ \frac{\dot{x}}{1-x} = 1 - e^{y}, & y < 0. \end{cases}$$

For large enough $|y(0)|$ *we have* $|y(t)| \simeq |y(0)| - t$*, while global exponential stability would imply that, given any constant* c*, there exists a time* $\bar{t}$ *such that*

$$|y(\bar{t})| \leq c|y(0)|, \forall y(0).$$

The equilibrium in the new variables is therefore no longer globally exponentially stable.

Stability criteria for continuous-time systems

Let us now look at the tools that we have in our toolbox to prove that an equilibrium is stable. We start from the well-known eigenvalue criterion for linear systems. Call λ_1 the dominant eigenvalue of the A matrix of the linear system, that is, the rightmost eigenvalue in the complex plane (or one of the rightmost ones, if there are more than one).

♦ **Eigenvalue criterion for linear systems**

- If $\Re(\lambda_1) < 0$, the linear system is asymptotically stable;
- if $\Re(\lambda_1) > 0$, the linear system is unstable;
- if $\Re(\lambda_1) = 0$ and the geometric multiplicity of the eigenvalues with $\Re(\lambda_i) = 0$ equals their algebraic multiplicity, the linear system is Lyapunov stable but not asymptotically stable;
- if $\Re(\lambda_1) = 0$ and the geometric multiplicity of the eigenvalues with $\Re(\lambda_i) = 0$ is less than their algebraic multiplicity, the linear system is unstable.

Notice how, in the above theorems, we have talked about stability *of a system*, which is of course a meaningless concept in nonlinear systems, which can have an arbitrary number of equilibria with different stability. We may however expect that, besides this difference, the above theorems could translate into equivalent criteria for nonlinear systems, at least as long as we consider dynamics sufficiently close to the equilibrium. This is true, but only to an extent, as we will soon learn that in some circumstances the eigenvalue criterion cannot be applied to equilibria of a nonlinear system. To start looking into this issue, let us review the relation between linear and nonlinear equilibria: the linearization process. We consider in the following the case of an autonomous vector field $f(x)$, but the extension to the case of a nonautonomous vector field $f(x, u)$ subject to a constant input $u = \bar{u}$ is rather simple.

■ **Definition: Jacobian matrix**
Given a differentiable vector field $f(x)$, its Jacobian $J_f(x)$ is the matrix of partial derivatives

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \\ \vdots & \vdots & \end{pmatrix}.$$

■ **Definition: Linearisation**
Given a nonlinear vector field $f(x)$ and one of its equilibria $\bar{x}$, the linearisation at the equilibrium is the linear system

$$\dot{\xi} = J_f(\bar{x})\xi,$$

where $J_f(\bar{x})$ is the Jacobian of f evaluated at the equilibrium $\bar{x}$, and ξ is a perturbation to x.

Note that, in this formula, the Jacobians are numerical matrices.

Through these tools, we can formulate the eigenvalue criterion for nonlinear systems.

♦ **Eigenvalue criterion for nonlinear systems**
Let $J_f(x)$ be the Jacobian of a vector field $f(x)$ with respect to the state x, and let $\bar{x}$ be an equilibrium. If all eigenvalues of $J_f(\bar{x})$ have a negative real part, then the equilibrium $\bar{x}$ of the nonlinear vector field $f(x)$ is asymptotically stable. If at least one of the eigenvalues of $J_f(\bar{x})$ has a positive real part, then $\bar{x}$ is unstable.

This theorem was first proved by Lyapunov in 1892. Notice how it does not say anything about the case of a dominant eigenvalue on the imaginary axis.

In cases where an equilibrium with no positive eigenvalues has some on the imaginary axis, the analysis of the linearisation does not tell us anything about stability. To discuss more thoroughly this issue we will need the Hartman-Großman theorem, which we will see later. In the meantime, we can reformulate the above criterion using the concept of hyperbolic equilibrium:

■ **Definition: Hyperbolic equilibrium or fixed point**
An equilibrium [A fixed point] is hyperbolic if it does not have eigenvalues on the imaginary axis [on the unit circle].

We can then say that

⧫ **Eigenvalue criterion for hyperbolic equilibria**
The stability of a hyperbolic equilibrium $\bar{x}$ of a nonlinear vector field $f(x)$ coincides with the stability of the linearisation $\dot{\xi} = J_f(\bar{x})\xi$.

Example 10. *Consider again the logistic model*

$$\dot{x} = x(1-x)$$

and its two equilibria $x = 0$ and $x = 1$. In 0 we have

$$J_f(0) = [1-2x]_{x=0} = 1,$$

while in 1 we have

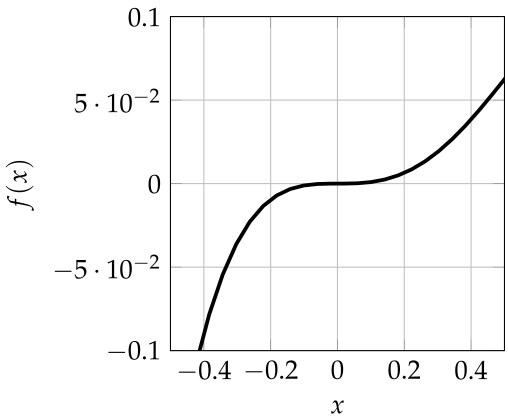
$$J_f(1) = [1-2x]_{x=1} = -1.$$

This confirms what we proved earlier, that $x = 0$ is unstable, while $x = 1$ is asymptotically stable.

Consider instead the modified model

$$\dot{x} = x^3(1-x),$$

which still has two equilibria, $x = 0$ and $x = 1$. We have $J_f(1) = -1$, as before, but $J_f(0) = 0$. The linearisation would classify the equilibrium in 0 as Lyapunov (but not asymptotically) stable. Let us look, however, at the plot of $f(x)$ near $x = 0$:



We have that $f(x) < 0$ for $x < 0$, and $f(x) > 0$ for $x > 0$, for x small enough. $x = 0$ must therefore be unstable.

The case of purely imaginary eigenvalues is not the only limitation of the eigenvalue criterion. A second, important limitation is that it provides no means to understand how far from the equilibrium the stability property holds. If multiple, stable equilibria are present, we should in fact expect different initial conditions to converge to different equilibria. Given these limitations, we should move on to see some more stability criteria, typically less simple but more informative in the context of nonlinear dynamics.

■ **Definition: Positive definite function**
A function $V(x) : \mathbb{R}^n \to \mathbb{R}$ is positive definite if

$$V(\bar{x}) = 0$$

for some $\bar{x}$, and

$$V(\bar{x}) > 0$$

for all $x \neq \bar{x}$.

The definition of *negative* definite is simply obtained be reversing the inequality.

For example,

$$x^2$$

and

$$|x|$$

are positive definite, but also

$$x^T P x,$$

where P is any symmetric matrix with positive eigenvalues, is positive definite.

■ **Definition: Positive semidefinite function**
A function $V(x) : \mathbb{R}^n \to \mathbb{R}$ is positive semidefinite if

$$V(\bar{x}) = 0$$

for some $\bar{x}$, and

$$V(\bar{x}) \geq 0$$

for all $x \neq \bar{x}$.

♦ **Lyapunov theorem**
Let x be the state of a dynamical system, $\bar{x}$ an equilibrium, and consider a continuously differentiable and positive definite function $V(x)$, defined in a neighbourhood of $\bar{x}$, with $V(\bar{x}) = 0$. Then:

- if $\dot{V}(x)$ is negative semidefinite, then $\bar{x}$ is Lyapunov stable;
- if $\dot{V}(x)$ is negative definite, then $\bar{x}$ is asymptotically stable;
- if $\dot{V}(x)$ is positive definite, then $\bar{x}$ is unstable.

For a proof see, e.g. (Khalil, 2002)

Continuously differentiable = continuous with continuous derivative

A function $V(x)$ satisfying the above statement is called a *Lyapunov function*. Note that nothing is required (not even existence) from the Lyapunov function far away from the equilibrium. We can however obtain a stronger result if $V(x)$ exists in $\mathbb{R}^n$ and is unbounded, in the sense specified below.

♦ **Lyapunov theorem and GAS (Barbashin-Krasovskii theorem)**
If the above theorem holds with $\dot{V}$ negative definite and defined over all $\mathbb{R}^n$, and V is radially unbounded, that is $V(x) \to +\infty$ as $\|x\| \to \infty$, then $\bar{x}$ is GAS.

Example 11. *The scalar continuous-time system*

$$\dot{x} = -x^3$$

has an equilibrium in 0 *with eigenvalue* 0*. Linear system theory would classify it as Lyapunov stable, but not asymptotically. However, using the Lyapunov function* $V(x) := x^2$*, we see that*

$$\dot{V}(x) = -2x^4,$$

which is negative definite and radially unbounded. The origin is therefore GAS, though its linearisation is not.

Lyapunov theorem is used to discuss the stability of nonhyperbolic equilibria, but it does not provide a means to establish which orbits converge to the equilibrium. This problem is addressed by another theorem, proved by Lasalle and Krasovskii, starting from the concept of *invariant set*.

■ **Definition: Positively-invariant set**
A compact set S is positively invariant (with respect to a vector field $f(x)$) if, for all $x \in S$, $\phi_t(x) \in S$, for all $t \geq 0$.

Note that, so long as $\phi_t(x) \in S$ and S is compact, the flow $\phi_t(x)$ exists for all $t \geq 0$. We have seen this together with the Picard-Lindelöf theorem. Therefore in this definition we can safely assume existence of $\phi_t(x)$ for all $t \geq 0$.

■ **Definition: Invariant set**
A compact set S is invariant (with respect to a vector field $f(x)$) if, for all $x \in S$, $\phi_t(x) \in S$, for all $t \in \mathbb{R}$.

We can alternatively define an invariant set S as a set with the following property:

$$x \in S \Rightarrow \phi.(x) \in S,$$

that is, an invariant set is a set that contains the entire orbit through any of its points.

♦ **LaSalle (Krasovskii) invariance principle**
Let M be a positively invariant subset of R^n, and consider a continuously differentiable function $V(x)$, such that $\dot{V}(x)$ is negative semidefinite in M. Let N be the set where $\dot{V}(x) = 0$, and let P be the largest invariant set within N. Then $\phi_t(x)$ approaches P as $t \to \infty$, for all $x \in M$.

This result was developed independently by LaSalle (who published it in 1960) and Krasovskii (in 1959), though it is now more commonly known as the LaSalle principle.

The hypothesis on V, of being continuously differentiable on a compact set M, implies that it is bounded in M.

★ Contrary to the Lyapunov theorem, the above principle does not require that V be positive definite. Moreover, on top of determining the asymptotic behaviour of solutions, it allows us to determine a compact set M from which all initial conditions eventually converge to P.

Here a short remark is due on the ways to prove that a given set M is positively invariant. While the general theory of positively invariant sets stands on an important result known as Nagumo's theorem, in our notes here we will always deal with fairly well-behaved sets, which essentially fall under one of these two families

Nagumo's theorem provides conditions for the positive invariance of a set using the concept of $T_M(x)$, the *tangent cone* to M at x: it is the set of all $z \in \mathbb{R}^n$ for which there exists a sequence $\{x'_1, x'_2, \dots\} \subset M$, with $x_k \overset{k\to\infty}{\to} x$, such that
$$\lim_{k\to\infty} \frac{x_k}{\|x_k\|} = \frac{z}{\|z\|}.$$
Then, assuming M is compact, we have that M is positively invariant provided that $f(x) \in T_M(x)$, for all $x \in M$. See for example (Aubin, 1991) or (Blanchini and Miani, 2008).

1. The compact set M is the sublevel set $\{x : V(x) \leq c\}$ of some differentiable function V, and it is such that $\frac{\partial V}{\partial x} \neq 0$ for all $x \in \partial M$.
2. The compact set M is a polytope with a nonempty interior, defined as the intersection of a finite set of half-spaces $\{x : g_i(x) \leq c_i\}$, where $g_i(x) : \mathbb{R}^n \to \mathbb{R}$, $i \in \{1, \dots m\}$ are m different linear functions.

★ **Proof of invariance of M, case 1**
If M is as in case 1, invariance of M is guaranteed provided that $\dot{V}(x) \leq 0$, for all $x \in \partial M$.

The above condition can be read as stating that, starting from any point x on the boundary of M, the value $V(x)$ that is found by following the flow $\phi_t(x)$ does not increase (at least for sufficiently small t). This means that the flow is not crossing the boundaries outwards from M. Similar reasoning supports the proof for case 2:

★ **Proof of invariance of M, case 2**
If M is as in case 2, invariance of M is guaranteed provided that $\dot{g}_i(x) \leq 0$, for all $x \in \partial M$ such that $g_i(x) = c_i$.

Example 12. *Let us take the linear system*

$$\dot{x} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x,$$

and show that the region defined by the set of inequalities

$$\begin{aligned} x_1 &\geq 0, \\ x_2 &\geq 0, \\ x_1 + 2x_2 &\leq 4, \end{aligned}$$

is positively invariant.

To proceed, we can write each of the inequalities in terms of a function $g_i(x)$ that is negative within the region, and positive outside. Then, all *that is left to show is that the derivatives of these functions along the flow* $\frac{\partial}{\partial t}g_i(x)$ *are negative on all of the region sides. Of course, at corners defined by the intersection between two* $g_i(x)$ *and* $g_j(x)$*, we must require both* $\frac{\partial}{\partial t}g_i(x) \leq 0$ *and* $\frac{\partial}{\partial t}g_j(x) \leq 0$*.*

In our case, we can take

$$\begin{aligned} g_1 &= -x_1, \\ g_2 &= -x_2, \\ g_3 &= x_1 + 2x_2 - 4. \end{aligned}$$

We have

$$\frac{\partial}{\partial x}g_1\dot{x} = \begin{pmatrix} -1 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x = x_1 = 0, \forall \{x_1 = 0, x_2 \in [0,2]\};$$

$$\frac{\partial}{\partial x}g_2\dot{x} = \begin{pmatrix} 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x = x_2 = 0, \forall \{x_2 = 0, x_1 \in [0,4]\};$$

$$\frac{\partial}{\partial x}g_3\dot{x} = \begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} x = \\ - x_1 - 2x_2 \leq 0, \forall \{x_1 \in [0,4], x_2 \in [0,2]\}.$$

The region is therefore positively invariant.

Example 13. *Consider the logistic growth model*

$$\dot{x} = x(1-x),$$

which has fixed points in 0 *and* 1*.*

Take the function $V(x) = x^2$*, defined in a neighbourhood of* 0*. It is positive definite, and*

$$\dot{V}(x) = \frac{\partial V}{\partial x}\dot{x} = 2x^2(1-x)$$

is positive definite near $x = 0$*. The fixed point in* 0 *is unstable.*

Now consider the function $V(x) = (1-x)^2$ *in a neighbourhood of* 1*. It is again positive definite, and*

$$\dot{V}(x) = \frac{\partial V}{\partial x}\dot{x} = -2x(1-x)^2,$$

which is negative definite around $x = 1$*. This is therefore an asymptotically stable fixed point.*

Finally, consider an interval $M := [0, b]$ *for some* $b > 1$*. We know that* $\dot{x} = 0$ *if* $x = 0$*, and* $\dot{x} < 0$ *if* $x = b$*, therefore the set* M *is positively invariant (solutions cannot escape it). Consider* $V(x) = (1-x)^2$ *with* $x \in M$*. We have* $\dot{V}(x) = -2x(1-x)^2$*, which is equal to* 0 *in* $x = 0$ *and* $x = 1$*. These are both fixed points and therefore are both positively invariant. By the LaSalle invariance principle, we know that all* $x \in M$ *converge to one of these two fixed points for* $t \to \infty$*.*

In fact, since we know that $x = 0$ *is unstable, we can further specify that all* $x \in M$*, except* $x = 0$*, converge to* $x = 1$ *for* $t \to \infty$*.*

LaSalle invariance principle can be further adapted to prove the global asymptotic stability of an equilibrium as follows (and, in this form, it goes once again under the name of Barbashin-Krasovskii theorem)

Theorem
Consider a continuously differentiable function $V(x)$ that is positive definite and radially unbounded in $\mathbb{R}^n$, such that $\dot{V}(x)$ is negative semidefinite in $\mathbb{R}^n$. Let N be the set where $\dot{V}(x) = 0$, and suppose that the only positively invariant set in N is the isolated equilibrium $\bar{x}$. Then $\bar{x}$ is GAS.

Stability criteria for discrete-time systems

The results of the previous section hold, with minimal changes, for discrete-time systems. The main change is that the time-derivative must be substituted with a finite time difference. The only remarkable difference (remarkable at least in that we will meet it again in the future) is in the eigenvalue criterion, which is modified as follows.

Eigenvalue criterion for discrete-time nonlinear systems
A discrete-time linear system is asymptotically stable provided that all eigenvalues of A are strictly within the unit circle, unstable if at least one eigenvalue lies outside of the unit circle.

Example 14. *Take the logistic map*

$$x(t+1) = 4x(1-x)$$

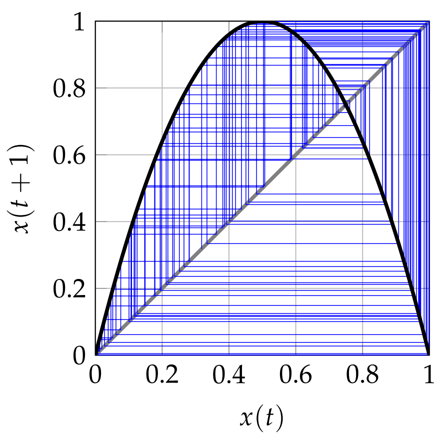
with equilibria in 0 *and* $\frac{3}{4}$. *We have*

$$J_f(0) = [4-8x]_{x=0} = 4,$$

and

$$J_f\left(\frac{3}{4}\right) = [4-8x]_{x=\frac{3}{4}} = -2.$$

By the eigenvalue criterion, both equilibria are unstable. Notice, however, that the system's orbits are bounded within the interval $[0,1]$.



Exercises

Exercise 10

Check that the Rosenzweig-MacArthur model

$$\begin{aligned}\dot{x}_1 &= x_1\left(1 - \frac{x_1}{8}\right) - \frac{x_1 x_2}{1 + x_1},\\ \dot{x}_2 &= -x_2 + 2\frac{x_1 x_2}{1 + x_1},\end{aligned}$$

has equilibria in $(0,0)$, $(8,0)$, and $\left(1, \frac{7}{4}\right)$, and discuss the stability of the equilibria.

Exercise 11

Use the Lyapunov theorem to prove that the equilibrium of

$$\dot{x} = \begin{pmatrix} -2 & 0 \\ 1 & -1 \end{pmatrix}$$

is GAS

Exercise 12

Find the equilibria of

$$\dot{x} = -x^3 + x^4$$

and discuss their stability using any one of the criteria seen in this chapter.

*** Exercise 13**

Use the LaSalle invariance principle and Lyapunov theorem to determine the stability of the origin in the following system, and to determine a region of initial conditions that converge to the origin:

$$\begin{aligned}\dot{x}_1 &= x_1(x_1^2 - x_2^2 - 1),\\ \dot{x}_2 &= x_2(x_2^2 + 3x_1^2 - 1).\end{aligned}$$

Answer of exercise 13

Let us start by studying the stability of the origin, using the Lyapunov theorem. As usual, lacking any better clue, we should start looking for $V(x)$ from the simplest quadratic functions. Let us take for instance the candidate function $V(x) = x_1^2 + x_2^2$. We have

$$\begin{aligned}\dot{V} &= 2x_1^2(x_1^2 - x_2^2 - 1) + 2x_2^2(x_2^2 + 3x_1^2 - 1)\\ &= 2x_1^4 + 2x_2^4 + 4x_1^2x_2^2 - 2x_1^2 - 2x_2^2\\ &= 2((x_1^2 + x_2^2)^2 - (x_1^2 + x_2^2)).\end{aligned}$$

This is a negative definite function near the origin since the linear term dominates for any x such that $x_1^2 + x_2^2 < 1$. The equilibrium is therefore asymptotically stable.

LaSalle invariance principle now requires finding a positively invariant set M, and given that we have already determined a quadratic $V(x)$ whose time derivative is negative definite near the origin, we may try to construct M as a sublevel set of $V(x)$. Our choice of M can be any circle

$$M := \{(x_1, x_2) : V(x) \leq \rho < 1\}$$

for some $\rho > 0$. We have that M is compact and $\dot{V}(x) \leq 0$ in all of M, so M is positively invariant. Furthermore, the only point where $\dot{V} = 0$ is the origin. Therefore, by the LaSalle invariance principle, all $x \in M$ converge to 0.

* Exercise 14

Consider once again the Rosenzweig-MacArthur model

$$\dot{x}_1 = x_1\left(1 - \frac{x_1}{8}\right) - \frac{x_1 x_2}{1 + x_1},$$
$$\dot{x}_2 = -x_2 + 2\frac{x_1 x_2}{1 + x_1},$$

and the equilibria that we found in Exercise 10. Prove that the three equilibria are enclosed in a compact, positively invariant region.

Exercise 15

Consider the system

$$\dot{x}_1 = (x_1^2 - 1)(2x_1 + x_2),$$
$$\dot{x}_2 = -x_2.$$

1. Prove that the set $\{x : -1 \leq x_1 \leq 1, -c \leq x_2 \leq c\}$ for arbitrary $c > 0$ is positively invariant.
2. Compute all equilibria of the system
3. Using the LaSalle invariance principle and the function $V = x_2^2$, discuss the behaviour of $x(t)$ for $t \to \infty$ for all $x(0) \in M$.

Answer of exercise 15

1. Set M is the polytope $\{x : -1 \leq x_1 \leq 1, -c \leq x_2 \leq c\}$. To prove that it is invariant, we should show that on all boundaries the vector field points within M. In general, we can do this by defining each portion of the boundary of M as the level set of a linear function $g(x) : \mathbb{R}^n \to \mathbb{R}$ whose gradient points outside of M. Then the invariance condition becomes $\dot{g}(x) = \frac{\partial g}{\partial x} f(x) \leq 0$.

   Take for instance $g(x) := -x_1 - 1$, for the left boundary. We have

   $$\frac{\partial g}{\partial x} f(x) = (-1, 0) f(x)|_{x_1=-1} = 0.$$

   The same happens on the boundary $x_1 = -1$, *mutatis mutandis*. If we now take $g := x_2 - c$ (for the boundary $x_2 = c$), we have

   $$\frac{\partial g}{\partial x} f(x) = (0, 1) f(x)|_{x_2=c} = -c < 0,$$

and similarly for the fourth boundary segment. Therefore, M is invariant.

2. $f(x) = 0$ for $x_1 \in \{-1, 0, 1\}, x_2 = 0$.

3. We already proved that M is positively invariant, and we have $\dot{V} = \frac{\partial V}{\partial x} f(x) = 2x_2\dot{x}_2 = -2x_2^2$. It is negative semidefinite, with $\dot{V} = 0$ in $\{x : x_1 \in [-1, 1], x_2 = 0\}$ (remember that we are only considering the set M now.)

   By the invariance principle, all $x(0) \in M$ converge to this set as $t \to \infty$. If we further investigate the stability of the 3 equilibria contained in this set, we find that the two equilibria in $x_1 = \pm 1$ are unstable, while the one in $x_1 = 0$ is stable. Therefore, all $x(0) \in M$ except those with $x_1(0) = \pm 1$ converge to $(0, 0)$.

Exercise 16

Consider the system

$$\begin{aligned}
\dot{x}_1 &= -2x_1 + x_2 - x_1^2 x_2^2, \\
\dot{x}_2 &= x_1 - x_2.
\end{aligned}$$

1. Prove that all initial conditions in the box $x_1 \in [0, 10]$, $x_2 \in [0, 10]$ converge to the origin.
2. Prove that the origin is not GAS.

Exercise 17

Prove that the continuous-time system

$$\begin{aligned}
\dot{x}_1 &= x_1 x_2 - x_1 x_2^2 \\
\dot{x}_2 &= -x_1^2 - x_2
\end{aligned}$$

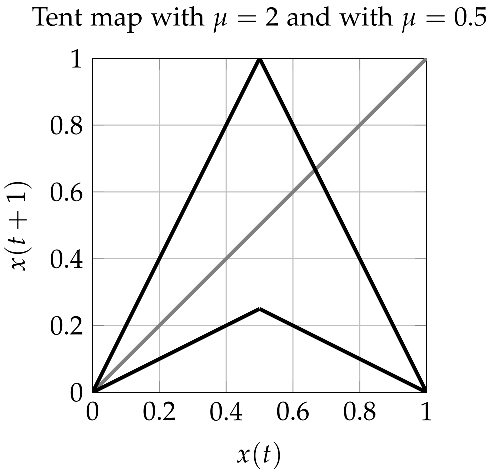
has a globally asymptotically stable equilibrium at the origin.

Exercise 18

By graphical inspection, determine the equilibria of the tent map

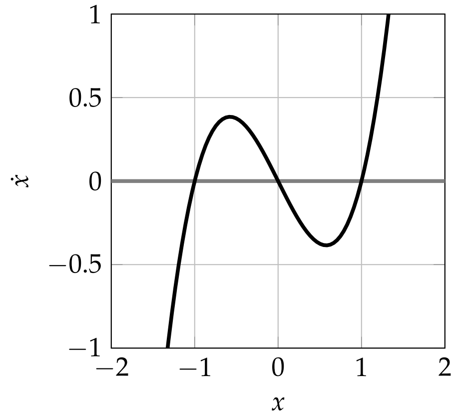
$$x(t+1) = \begin{cases} \mu x(t), & x(t) \leq 0.5, \\ \mu(1 - x(t)), & x(t) > 0.5, \end{cases}$$

and their stability, as μ changes between 0.5 and 2



Exercise 19

By graphical inspection, determine the equilibria of the following function, seen as the vector field of a continuous-time system, and their stability. Discuss how the phase portrait changes as $f(x)$ is shifted up or down.



Phase portrait near an equilibrium or a fixed point

Keywords: **Hyperbolic equilibria, node, focus, saddle, nullclines, Poincaré index.**

Geometry of the orbits near an equilibrium

Much of what we learn in this course regards the representation and the analysis of the qualitative features of orbits in the state space, that is, the qualitative study of the *phase portrait:*

■ **Definition: phase portrait**
The phase portrait is the set of all the orbits of a dynamical system

Let us consider, for the moment, the neighbourhood of an equilibrium. Our attention so far has been directed towards the asymptotic behaviour of orbits near an equilibrium: does it approach the equilibrium or not? For a more complete understanding of the behaviour of a system near its equilibria, we may want to investigate not just the limit $t \to \infty$, but the behaviour, that is, the shape of the orbits for all $t \geq 0$. This means studying the phase portrait near the equilibrium.

We can start our investigation by looking at some particular orbits, which are defined by the eigenvectors of the equilibrium.

★ Every real eigenvector of a linear continuous-time [discrete-time] system defines a one-dimensional invariant space through the equilibrium [fixed point].

An invariant space is simply an invariant set that is also a vector space, that is, it is closed with respect to addition and multiplication by a scalar.

The above property is easily proved: take an initial condition x that is an eigenvector. In this case

$$\dot{x} = Ax = \lambda x$$

specifies a straight orbit moving in the direction of x, and therefore entirely contained in the the one-dimensional space defined by the eigenvector. Obviously enough, initial conditions in this space approach the equilibrium along a straight line if the eigenvalue is negative or depart from it if the eigenvalue is positive.

Note that the same conclusion holds for a discrete-time system:

$$x(t+1) = Ax(t) = \lambda x(t).$$

★ The real and imaginary parts of a pair of complex eigenvectors of a linear continuous or discrete-time system define a two-dimensional invariant space through the equilibrium

The above property is slightly more complex to deduce. We can first notice that, given eigenvectors $\alpha \pm i\beta$ and the corresponding eigenvalues $a \pm ib$, we have

$$A((\alpha + i\beta) + (\alpha - i\beta)) = 2\alpha a - 2\beta b$$

and

$$A((\alpha + i\beta) - (\alpha - i\beta)) = i(2\alpha b + 2\beta a),$$

which we can simplify into

$$A\alpha = \alpha a - \beta b$$

and

$$A\beta = \alpha b + \beta a.$$

This means that, if an initial condition x can be written as a weighted sum of α and β, then $\dot{x} = Ax$ is itself a weighted sum of α and β. The space defined by the vectors α and β is therefore invariant, and orbits approach or depart from the equilibrium depending on the stability of the two eigenvalues. Once again, the same holds for discrete-time dynamics.

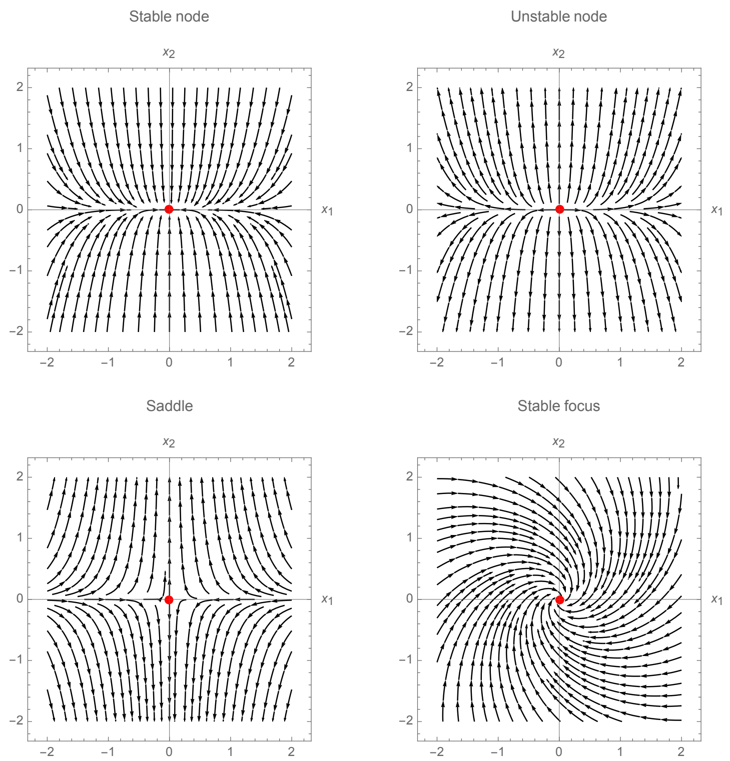
We have thus seen how eigenvalues of linear systems identify invariant subspaces through equilibria and fixed points. It would be interesting if a similar property held at equilibria and fixed points of nonlinear systems. This is true, up to a point: in a nonlinear system, the mentioned spaces become manifolds (curved lines or surfaces) and only exist locally around the equilibrium. These limitations notwithstanding, much of the geometry of the orbits near an equilibrium or fixed point of a nonlinear system can be described through a careful analysis of its eigenvalues and eigenvectors.

Besides the above classification, in 2D systems, we have a richer terminology to define equilibria with different-looking phase portraits.

■ **Definition: Types of continuous-time hyperbolic equilibria**

A hyperbolic equilibrium of a planar system with

- two real negative eigenvalues is called a *stable node*,
- two real positive eigenvalues is called an *unstable node*,
- one positive and one negative eigenvalue is called an *saddle*,
- a pair of complex eigenvalues is called a *focus* (stable or unstable, depending on the real parts).



Can you tell the eigenvectors in each of these phase portraits? And can you sketch the phase portrait near the stable node with Jacobian $\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$? (note that it has only one eigenvector)

The above classification only considers hyperbolic equilibria, and diagonalizable matrices (that is, matrices whose eigenvalues have the same geometric and algebraic multiplicity). In the cases of nondiagonalizable matrices, the portraits are slightly different due to the lack of one of the two invariant subspaces. The cases of nonhyperbolic equilibria instead are more fundamentally different and require more careful analysis to decide the geometry of nearby orbits in the invariant set correspondent to the nonhyperbolic eigenvalues. We will understand more of these cases in the coming chapters.

In dimensions higher than 2 the same terminology is sometimes used to describe the local appearance of the phase portrait, but terms are somewhat mixed when subspaces of different nature coexist. The most common notation in this case is to call

- node: a hyperbolic equilibrium whose eigenvalues are all real and of the same sign

- focus: a hyperbolic equilibrium with at least two complex conjugate eigenvalues, and whose eigenvalues are all in the same side of the imaginary axis

- saddle-focus: a hyperbolic equilibrium with at least two complex-conjugate eigenvalues, and with eigenvalues on both sides of the imaginary axis

The equivalent definitions in discrete time are the following.

Definition: Types of discrete-time hyperbolic fixed points
A hyperbolic fixed point of a planar system with

- real eigenvalues in the interval $[0, 1)$ is called a *stable node*,
- real eigenvalues in the interval $(1, \infty)$ is called an *unstable node*,
- one eigenvalue in $[0, 1)$ and one in $(1, \infty)$ is called a *saddle*,
- eigenvalues anywhere else in the complex plane, except on the unit circle, is called a *focus*.

With the above terminology, we can now rigorously classify the types of hyperbolic equilibria and fixed points in planar continuous and discrete-time systems: given their type and eigenvectors, we can easily sketch the geometry of the orbits in their neighbourhood.

Example 15. *Let us sketch the phase portrait of*

$$\dot{x} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} x.$$

Its eigenvalues solve

$$\lambda^2 - 2 = 0 \Rightarrow \lambda = \pm\sqrt{2}.$$

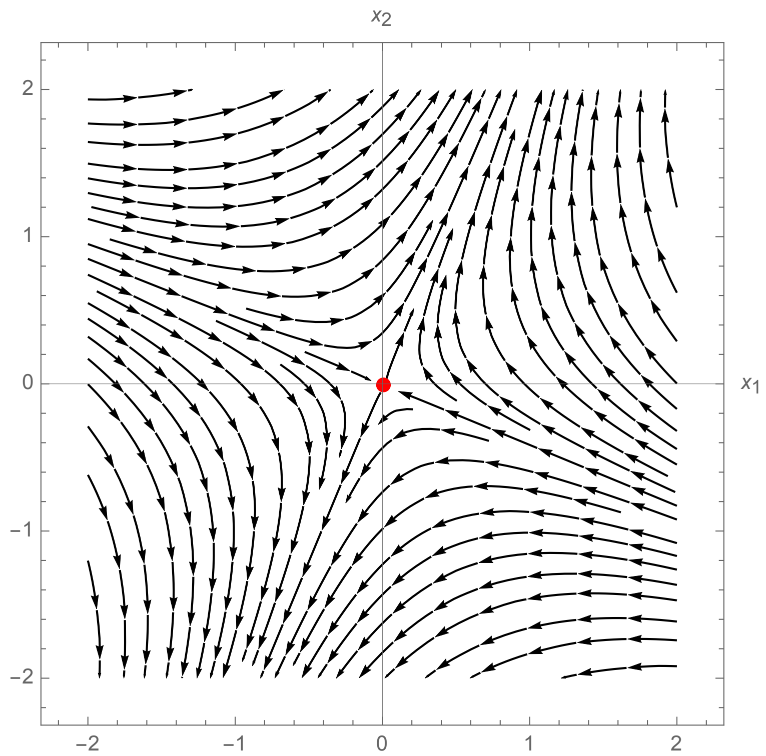
The equilibrium is therefore a saddle. We can find the eigenvectors by solving

$$\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \sqrt{2} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \Rightarrow v_2 = v_1 \left(1 + \sqrt{2}\right) \Rightarrow v = \begin{pmatrix} 1 \\ 1 + \sqrt{2} \end{pmatrix}$$

and

$$\begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = -\sqrt{2} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \Rightarrow v_1 = -v_2 \left(1 + \sqrt{2}\right) \Rightarrow v = \begin{pmatrix} -1 - \sqrt{2} \\ 1 \end{pmatrix}.$$

The system has two orthogonal eigenvectors, one, unstable, in the first and third quadrants, and the other one, stable, in the second and fourth quadrants.



Planar systems present an additional advantage, other than having a small and well-defined set of types of equilibria or fixed points. Since the vector field $f(x)$ has values in the plane, the solutions of $f_i(x) = 0$ (or $f_i(x) = x_i$, for discrete-time systems) form a set of curves in the plane, which one can easily sketch and use to understand a systems's behaviour. These curves are known as *nullclines*. They prove particularly useful in continuous-time systems, where the fact that orbits cannot cross each other makes the analysis through nullclines particularly powerful. For this reason, we discuss here their usage only in the context of continuous-time systems, even though an extension to the discrete-time domain would be technically simple.

■ **Definition: Nullclines**

The nullclines of a 2-dimensional vector field $f(x)$ are the curves defined by the two equations

$$f_i(x) = 0$$

Example 16. *Let us compute the nullclines of the linear system we studied before:*

$$\dot{x} = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} x.$$

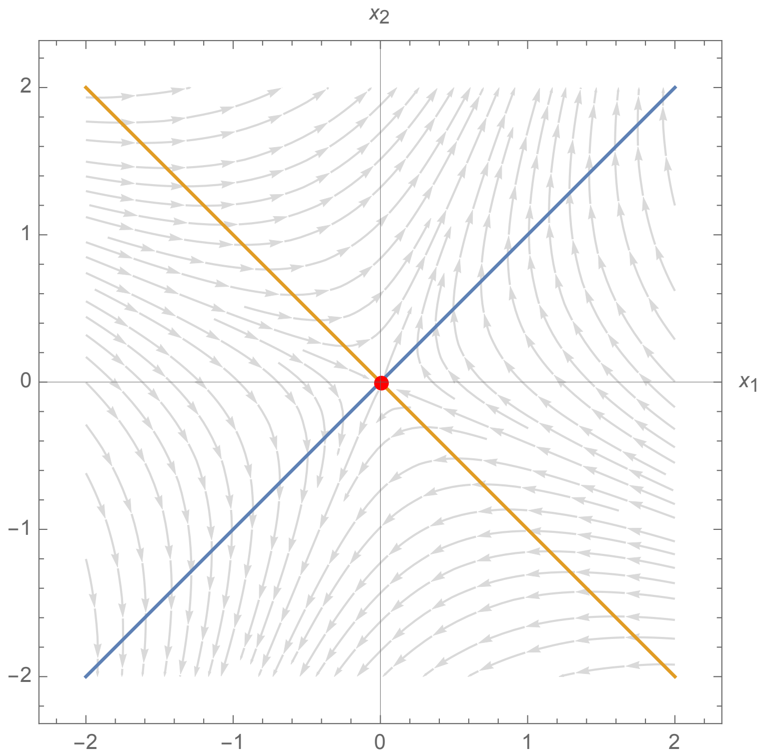
The x_1 nullcline solves

$$-x_1 + x_2 = 0 \Rightarrow x_1 = x_2.$$

The x_2 nullcline solves

$$x_1 + x_2 = 0 \Rightarrow x_1 = -x_2.$$

We can use these to help us sketch the phase portrait. Along the x_1 nullcline the flow lines have a vertical tangent, along the x_2 one they have a horizontal tangent.



Example 17 (Phase portrait of a prey-predator model)**.** *Let us find the equilibria of the prey-predator model*

$$\dot{x}_1 = x_1(1 - \frac{x_1}{4}) - x_1 x_2,$$
$$\dot{x}_2 = x_1 x_2 - x_2.$$

using nullclines.

The x_1 nullclines are $x_1 = 0$ and $x_1 = 4 - 4x_2$; the x_2 nullclines are $x_2 = 0$ and $x_1 = 1$. We have equilibria at each intersection of an x_1 nullcline with an x_2 nullcline, therefore in $(0,0)$, $(4,0)$, and $\left(1, \frac{3}{4}\right)$.

To sketch the phase portrait near the three equilibria we can study their linearization. We have

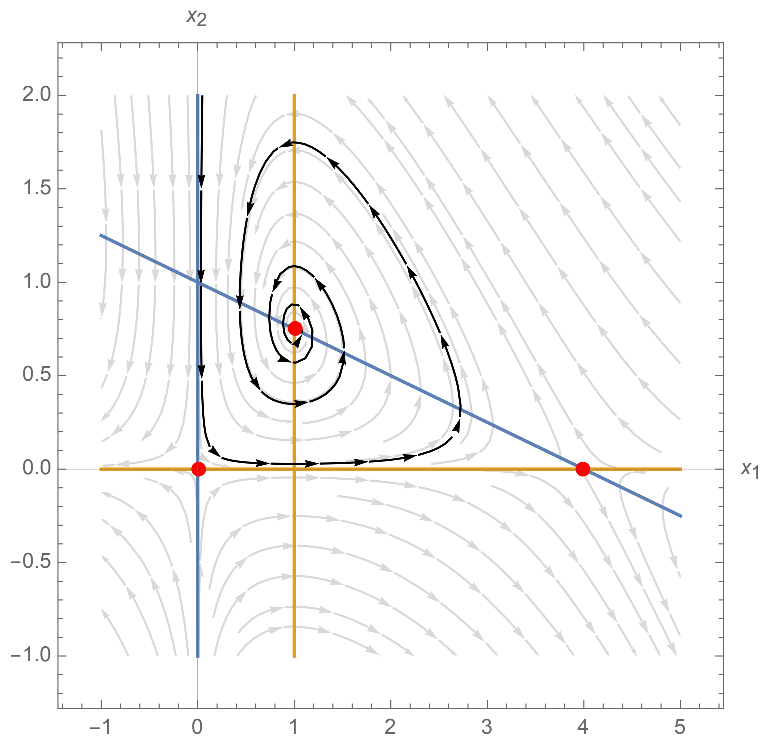
$$J_f(x) = \begin{pmatrix} 1 - \frac{x_1}{2} - x_2 & -x_1 \\ x_2 & x_1 - 1 \end{pmatrix},$$

therefore

$$J_f(0,0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J_f(4,0) = \begin{pmatrix} -1 & -4 \\ 0 & 3 \end{pmatrix}, \quad J_f\left(1, \frac{3}{4}\right) = \begin{pmatrix} -\frac{1}{4} & -1 \\ \frac{3}{4} & 0 \end{pmatrix}.$$

We see that $(0,0)$ and $(4,0)$ are saddles, while $\left(1, \frac{3}{4}\right)$ is a stable focus(it has negative trace and positive determinant, and $\mathrm{tr}^2 - 4\det < 0$, that is, the discriminant is negative).

Let us now sketch the global phase portrait of the system



Reasoning on the nullclines, we can now imagine what would happen if, for example, the carrying capacity, equal to 4 in the equation of $\dot{x}_1$, was to decrease. The only nullcline affected by this change would be

$$x_1 = k - kx_2$$

(the slanted blue line in the plot). As k is reduced, the slope of the line increases (while the line keeps crossing the x_2 axis at 1), and the focus equilibrium approaches the rightmost saddle. When k becomes less than 1, we expect the two equilibria to collide. At this value, the strictly positive equilibrium crosses to the fourth quadrant, and prey and predator can no longer coexist.

Poincaré index theory

We have just seen how the uniqueness of orbits (the Picard-Lindelöf theorem) in a planar continuous time system bears as a consequence a large set of constraints on how orbits can move about the phase space. This is why so much about the phase portrait of a continuous-time planar system can be understood by studying its nullclines. An even stronger impression of the consequences of uniqueness comes from the Poincaré index theory, which is based on the following observations.

■ **Definition: Poincaré index**
Given a planar non-intersecting, oriented, and closed curve L and a planar continuous vector field f, the index $I_L(f)$ is the integer number of rotations of the vector field $f(x)$ as x traverses the curve, counted positive for rotations in the same direction of rotation as the curve L.

A planar non-intersecting curve is sometimes called a *Jordan curve*

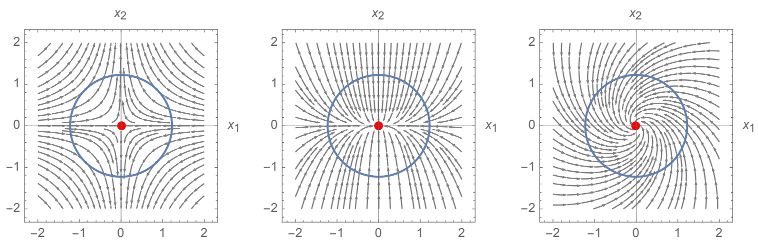
Here rotations are counted with a sign, positive in the direction of rotation of the curve L.

♦ **Invariance of the Poincaré index**
If the curve is deformed without crossing equilibria, its Poincaré index does not change.

♦ **Index of a sum of curves**
The index of a sum of curves is the sum of the indices of the curves

Based on the above theorems, we can now imagine taking a curve that encloses exactly one equilibrium, and shrinking it until it encircles the equilibrium tightly, that is, until it lies within a vector field that is well approximated by the linearisation around the equilibrium.

Example 18. *Let us look at the indices of a Jordan curve surrounding a saddle, a node, and a focus*



It is now easy to conclude that

♦ **Theorem**
The index of a curve that encloses

- exactly one node or focus is 1,
- exactly one saddle is -1,
- no equilibria is 0.

The index is independent of the stability of the node!

We can, with little abuse of notation, say that the index of a node or focus is 1, and that of a saddle is -1. We can now complete the above list as follows

♦ **Theorem**
The index of a curve that encloses a set of nodes, foci, and saddles, is $N - M$, where N is the number of nodes and foci, and M is the number of saddles.

Considering that the index of a periodic orbit is $+1$ (easy to check), we can also state the following

♦ **Theorem**
Every periodic orbit in the plane must enclose a set of equilibria whose indices sum exactly to $+1$.

Example 19. *A control system is designed so that*

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

with f and g continuous, has a GAS equilibrium at $x = 0$. The system is then perturbed, becoming

$$\dot{x} = f(x, u) + h(x),$$
$$u = g(x),$$

with $h(x)$ continuous and $\frac{\|h(x)\|}{\|f(x,u(x))\|} < \rho < 1$ for $\|x\| > C$, for some constant C. Let us discuss what are some possible and impossible scenarios for the perturbed system.

First, let us take a closed curve L around the origin, so that $\|x\| > C$ for all $x \in L$. By our assumption on the relative sizes of $h(x)$ and $f(x, u(x))$, the curve must have the same index with and without the perturbation. Therefore the curve must have an index 1, since it encloses exactly 1 equilibrium in the unperturbed case. This imposes some surprisingly tight restrictions on the effects of $h(x)$, even though we are not assuming anything about the size of $h(x)$ near the origin.

A new stable equilibrium appears near $x = 0$. *This is not possible, since the total index of the equilibria within ρ would become 2.*

A new pair of equilibria, a saddle and a stable or unstable node, appear near $x = 0$. *This is possible since the total index of the equilibria within ρ would remain unchanged.*

The equilibrium changes stability *This is possible since its index is not a function of its stability.*

The equilibrium disappears *This is not possible, since the index of ρ would become 0.*

Exercises

Exercise 20

A planar, parameter-dependent system

$$\dot{x} = f(x, p)$$

has a single stable equilibrium $\bar{x}$ and no periodic orbits in a region enclosed by the curve L. Suppose that, as f changes continuously, $I_L(f)$ remains constant and no other equilibrium crosses L. Can we say that

- the equilibrium $\bar{x}$ may disappear,
- a new saddle equilibrium may appear inside L,
- a new focus equilibrium may appear inside L,
- a new pair of equilibria, a node and a saddle, may appear inside L,
- a periodic orbit may appear inside L,
- a pair of periodic orbits may appear inside L.

Exercise 21

Sketch the phase portrait of

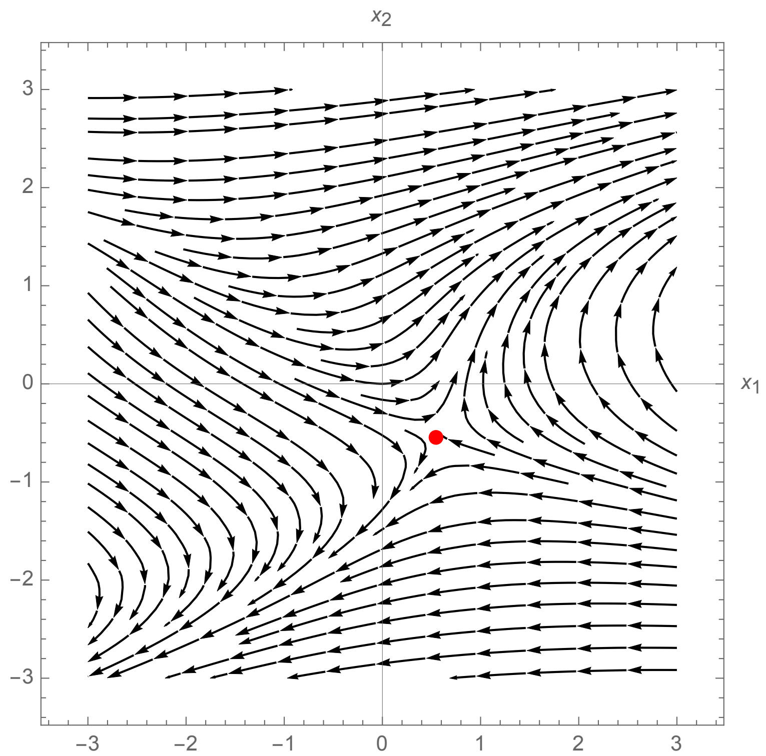
$$\dot{x} = \begin{pmatrix} -1 & 1 \\ 3 & 1 \end{pmatrix} x,$$

highlighting the eigenvectors of the equilibrium and the direction of motion of the vector field in each of the regions separated by the eigenvectors.

Exercise 22

The phase portrait below shows the neighbourhood of the only equilibrium in the system

$$\begin{aligned} \dot{x}_1 &= 1 - x_1 + x_1 x_2 + x_2^3, \\ \dot{x}_2 &= x_1 + x_2. \end{aligned}$$



Explain why the system cannot have a periodic orbit.

Exercise 23

Sketch the nullclines of

$$\dot{x}_1 = x_1 x_2 + x_1^2,$$
$$\dot{x}_2 = -x_1 + x_2 + 1,$$

and study the type and stability of the equilibria.

Exercise 24

Sketch the nullclines of

$$\dot{x}_1 = x_2(x_1^3 - x_1),$$
$$\dot{x}_2 = -x_1 + x_2^2,$$

identify the equilibria and their type, and sketch qualitatively the direction of motion of the orbits in each of the regions of the phase space separated by the nullclines.

Answer of exercise 24

The nullclines solve $f_i(x) = 0$. From f_1 we obtain the straight lines $x_1 = 0$, $x_1 = \pm 1$, $x_2 = 0$. From f_2 we have $x_1 = x_2^2$. Equilibria lie at the intersections of a nullcline of f_1 with one of f_2, at $(0, 0)$, $(1, 1)$, and $(1, -1)$. The Jacobians are

$$J_f(0,0) = \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix},$$

$$J_f(1,1) = \begin{pmatrix} 2 & 0 \\ -1 & 2 \end{pmatrix},$$

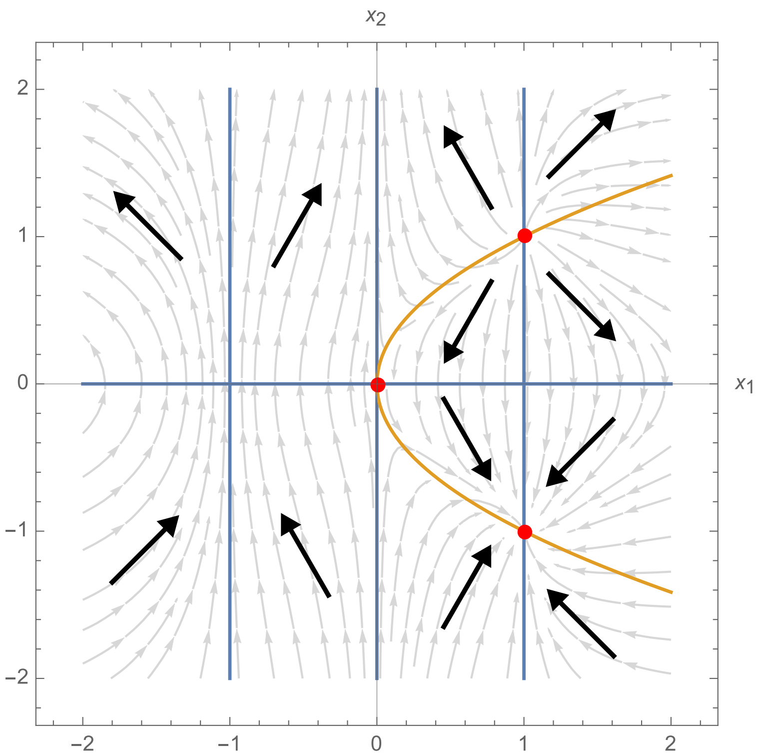
and

$$J_f(1,-1) = \begin{pmatrix} -2 & 0 \\ -1 & -2 \end{pmatrix}.$$

Therefore the first equilibrium is nonhyperbolic (we could tell by the fact that it lies at a nontransversal intersection of two nullclines), while the other two are respectively an unstable node and a stable node.

To sketch qualitatively the direction of motion of the orbits, it is sufficient to choose any one region and determine the sign of $\dot{x}_1$ and $\dot{x}_2$ within the region. Then, whenever we cross a nullcline that is a simple root of f_1 or f_2 the corresponding component of $\dot{x}$ changes sign.

The figure below portrays the nullclines, the qualitative direction of motion (black arrows), and a computer-generated representation of the phase portrait (grey arrows).



Exercise 25

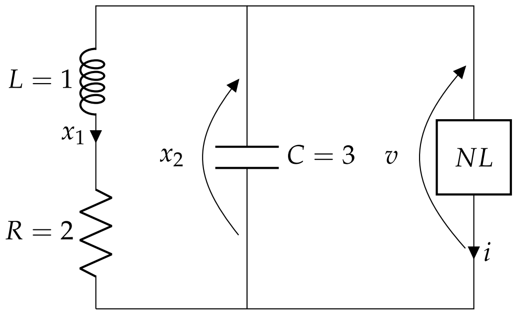
Sketch the nullclines of

$$\dot{x}_1 = x_1^2 - x_2 - 1,$$
$$\dot{x}_2 = x_1^2 + x_2,$$

identify the equilibria and their type, and sketch qualitatively the direction of motion of the orbits in each of the regions of the phase space separated by the nullclines.

Exercise 26

Sketch the phase portrait (nullclines, equilibria, qualitative direction of motion) of the following circuit



where the nonlinear element NL has characteristic

$$i = -v + v^3.$$

Homeomorphisms and diffeomorphisms

Keywords: **Homeomorphism, diffeomorphism, change of variables, topological conjugacy, inverse function theorem, topological equivalence.**

Mappings between flows

The functional form of a dynamical system is usually the result of multiple modelling decisions. We know very well however that, for example, the asymptotic stability of a linear equilibrium is a function of its eigenvalues, and these are independent of the choice of variables with which we decide to describe the system. In other words, many interesting properties of a dynamical system do not depend on these modelling choices and may be easier to investigate in some functional representations than in others. The tools to transform a system from one representation to another, preserving these properties, are called homeomorphisms and diffeomorphisms. We now learn what they are, but we will elaborate on how to use them, to transform continuous-time or discrete-time systems, in the next chapter.

■ **Definition: Homeomorphism**
A homeomorphism is a map that is bijective, continuous, and with continuous inverse.

A bijective map is a one-to-one map, that is, exactly one element of the codomain corresponds to each element of the domain.

■ **Definition: Topological conjugacy**
Two flows $\phi_t(x)$ and $\psi_t(y)$ are topologically conjugate if there exists a homeomorphism $h : y = h(x)$ such that, for all x and for all t,

$$h(\phi_t(x)) = \psi_t(h(x)).$$

■ **Definition: Local topological conjugacy**
Two flows $\phi_t(x)$ and $\psi_t(y)$ are locally topologically conjugate near the states $\bar{x}$ and $\bar{y}$ if the flows, restricted to two neighbourhoods of $\bar{x}$ and $\bar{y}$ respectively, are topologically conjugate.

We can see the conjugacy relation as the following commuting diagram: it states that the flow of x, which is $\phi_t(x)$, is mapped through

h onto the flow of y with $y = h(x)$, which is $\psi_t(h(x)) = h(\phi_t(x))$.

Example 20. *Consider the 1-dimensional system*

$$\dot{x} = -x,$$

whose flow is

$$x(t) = e^{-t}x(0),$$

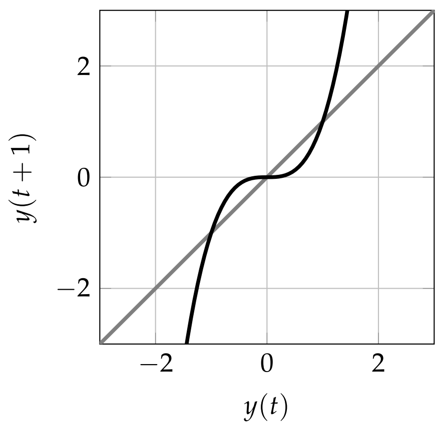
and take $y = h(x) := x^3$. This is a homeomorphism: it is bijective, continuous, and its inverse $x = \sqrt[3]{y}$ is continuous, though it is not continuously differentiable. If we map x onto y as $y = h(x)$, and $\phi_t(x)$ onto $\psi_t(y)$ as $\psi_t(y) = h(\phi_t(x))$, we obtain

$$\begin{array}{ccc} \dot{x} = -x & \xrightarrow{\phi_t} & e^{-t}x(0) \\ \downarrow h & & \downarrow h \\ \dot{y} = -3y & \xrightarrow{\psi_t} & e^{-3t}y(0) \end{array}$$

The diagram commutes, and the two flows, of $\dot{x} = -x$ and of $\dot{y} = -3y$, are topologically conjugate. Notice that the equilibrium (which is unique, since the systems are linear with a nonsingular matrix) has different eigenvalues in the two systems.

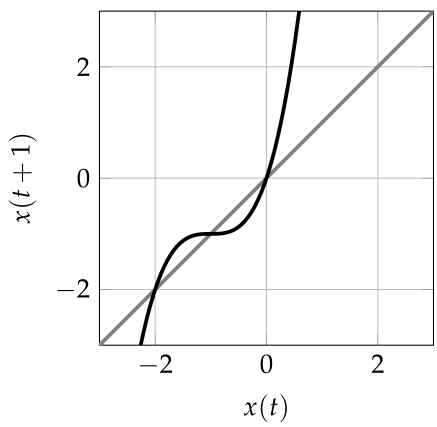
Example 21. *Take the discrete-time systems*

$$y(t+1) = \psi_1(y) := y^3(t)$$



and

$$x(t+1) = \phi_1(x) := x^3(t) + 3x^2(t) + 3x(t),$$



and the homeomorphism

$$y = h(x) = x + 1.$$

We have

$$\psi_1(h(x)) = (x+1)^3 = (x^3(t) + 3x^2(t) + 3x(t)) + 1 = h(\phi_1(x)).$$

This implies that $\psi_t(h(x)) = h(\phi_t(x))$, the two flows are therefore topologically conjugate.

We can easily deduce two interesting results from the definition of topological conjugacy:

♦ **Theorem**
Topologically conjugate flows have the same number of equilibria or fixed points, with the same stability.

♦ **Theorem**
Invariant sets are preserved through topological conjugacy.

This means that topologically conjugate flows behave essentially the same. This is, clearly, a very strong requirement. We will see in the next chapter that local topological conjugacy is a more useful definition since by only focussing on a small neighbourhood of a state (typically, of an equilibrium) we can prove much more general results about the properties of the dynamics in that neighbourhood.

From the above examples, we may also notice that homeomorphisms $h(x)$ between discrete-time systems with vector fields $f_1(y)$ and $f_2(x)$ provide an explicit mapping between the vector fields, as long as we can compute h^{-1}:

$$f_2(x) = h^{-1}(f_1(h(x))).$$

This means that f_2 can be obtained explicitly if h and its inverse h^{-1} are known. It also implies as we see in the next example, that homeomorphisms preserve the eigenvalues of the fixed points of discrete-time flows.

Example 22. *(Homeomorphisms and eigenvalues of fixed points) Consider a system*

$$x(t+1) = f(x(t)),$$

and the homeomorphism $x = h(y)$. We have

$$y(t+1) = h^{-1}\left(f(h(y))\right).$$

Assume that $\bar{x} = h(\bar{y})$ is a fixed point. The Jacobian at the fixed point is

$$\left[\frac{\partial}{\partial y} h^{-1}(f(h(y)))\right]_{y=\bar{y}} = \left[J_{h^{-1}}(f(h(y)))J_f(h(y))J_h(y)\right]_{y=\bar{y}}$$
$$\overset{f(h(\bar{y}))=h(\bar{y})}{=} J_{h^{-1}}(h(\bar{y}))J_f(h(\bar{y}))J_h(\bar{y}).$$

Now notice that matrices $J_{h^{-1}}(h(\bar{y}))$ and $J_h(\bar{y})$ are inverses of each other, since $h^{-1}(h(y)) = y$ and $\frac{\partial}{\partial y}\left(h^{-1}(h(y))\right) = J_{h^{-1}}(h(y))J_h(y) = \frac{\partial}{\partial y}y = I$. Therefore, the matrix $J_{h^{-1}}(h(\bar{y}))J_f(h(\bar{y}))J_h(\bar{y})$ is similar to $J_f(\bar{x})$, and they share the same eigenvalues.

We had seen previously that the same did not hold at the equilibrium of a continuous-time system.

♦ **Theorem**
Homeomorphisms preserve the eigenvalues of fixed points, but not those of equilibria.

In other words, homeomorphisms appear to preserve more structure when applied to discrete-time flows, than when they are applied to continuous time flows. We recover, on continuous-time systems, the same behaviour that homeomorphisms have on discrete-time systems if we slightly tighten the requirements on the map h.

■ **Definition: Diffeomorphism**
A diffeomorphism is a map that is bijective, continuously differentiable, with a continuously differentiable inverse.

Example 23 (Linear change of variables are diffeomorphisms)**.** *Consider a linear system*

$$\dot{x} = Ax + Bu,$$

and a similar system

$$\dot{y} = M^{-1}AMy + M^{-1}Bu,$$

obtained through the change of variables

$$x = My,$$

with M nonsingular. The map My is continuous, its differential M is continuous, and it has a continuously differentiable inverse. Therefore, it is a diffeomorphism.

♦ **Theorem**
Diffeomorphisms preserve the eigenvalues of equilibria and fixed points.

In other words, flows that are topologically conjugate through a diffeomorphism share equilibria (or fixed points) with the same sets of eigenvalues.

Example 24. *(Diffeomorphisms and eigenvalues of equilibria) Consider a system*

$$\dot{x} = f(x),$$

and the diffeomorphism $x = h(y)$. We can write the system in the new variables as

$$\dot{x} = J_h(y)\dot{y} = f(h(y)),$$

therefore

$$\dot{y} = (J_h(y))^{-1} f(h(y))$$

The reason why a homeomorphism does not provide an explicit mapping between continuous-time flows is in this equation: in general, $J_h(y)$ and $(J_h(y))^{-1}$ are not everywhere defined unless we deal with diffeomorphisms.

Assume that $\bar{x}$ is an equilibrium, and $\bar{y}$ is the corresponding equilibrium in the y variables. The Jacobian at the equilibrium is

$$\left[\frac{\partial}{\partial y}\left((J_h(y))^{-1} f(h(y))\right)\right]_{y=\bar{y}}$$
$$= \left[\frac{\partial}{\partial y}(J_h(y))^{-1} f(h(y)) + (J_h(y))^{-1} J_f(h(y))J_h(y)\right]_{y=\bar{y}}$$
$$\overset{f(h(\bar{y}))=0}{=} (J_h(\bar{y}))^{-1} J_f(h(\bar{y}))J_h(\bar{y}).$$

The matrix $(J_h(\bar{y}))^{-1} J_f(h(\bar{y}))J_h(\bar{y})$ is similar to $J_f(\bar{x})$, therefore they share the same eigenvalues.

◆ **Inverse function theorem**
Consider the continuously differentiable map $y = h(x)$. If $J_h(\bar{x})$ is nonsingular then, in a neighbourhood of $\bar{y} = h(\bar{x})$, h is invertible and $h^{-1}(y)$ is continuously differentiable.

Even though it is tempting to use this result to prove that f is globally a diffeomorphism simply by showing that $J_h(x)$ is nonsingular everywhere, a few more assumptions are needed for this. One form of the Hadamard-Levy theorem, for instance, states that a continuously differentiable function h is a diffeomorphism from $\mathbb{R}^n \to \mathbb{R}^n$ provided that $J_h(x)$ is nonsingular for all x, *and* $\|J_h^{-1}(x)\| < a + b\|x\|$, for some $a, b > 0$. See (Plastock, 1974) for more equivalent conditions.

The above theorem states that a continuously differentiable function $g(x)$ is a diffeomorphism in a neighbourhood of $\bar{x}$ provided that the Jacobian $J_g(\bar{x})$ is nonsingular. While constructing a global diffeomorphism is typically challenging, we will more often be satisfied with diffeomorphisms defined in a neighbourhood of some special state $\bar{x}$, hence nonsingularity of the Jacobian is all we will need.

Definition of topological equivalence

Topological conjugacy, which we defined in the previous section, is a very demanding relation: it means that every orbit of ϕ_t is mapped, at each instant of time, onto an orbit of ψ_t. The following relation is a little looser, in that it requires mapping of orbits onto orbits, but it allows for time to be stretched at will, as long as direction is preserved.

■ **Definition: Topological equivalence**
Two flows are topologically equivalent if there exists a homeomorphism mapping orbits of one onto orbits of the other, preserving the direction of time.

Example 25. *Consider the two scalar systems*

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

and

$$\dot{y} = |y|.$$

We have already encountered the first, and we know that its flow is

$$x(t) = \phi_t(x) = \begin{cases} \frac{1}{\frac{1}{x(0)} - t}, & x(0) \neq 0, \\ 0, & x(0) = 0. \end{cases}$$

Solutions converge to 0 *from the left, if* $x(0) < 0$*, while they diverge in finite time if* $x(0) > 0$*, or stay at* 0 *if* $x(0) = 0$*.*

The flow of the second system, on the other hand, can be easily computed by looking separately at the domains $y < 0$*,* $y = 0$*, and* $y > 0$*:*

$$y(t) = \psi_t(y) = \begin{cases} e^{-t} y(0), & y(0) < 0, \\ 0, & y(0) = 0, \\ e^{t} y(0), & y(0) > 0. \end{cases}$$

Since the first flow is defined only on a time domain $t \in \left[0, \frac{1}{x(0)}\right)$ *when* $y(0) > 0$*, there exists no homeomorphism such that*

$$h(\phi_t(x)) = \psi_t(h(x)), \forall t.$$

The two flows are not topologically conjugate. They are, however, topologically equivalent, and the homeomorphism mapping orbits to orbits is simply the identity: $y = x$*.*

★ Discrete-time flows that are topologically equivalent are also topologically conjugate.

This follows from the fact that time is discrete, so when mapping a discrete-time system we don't have the option of stretching it or shrinking it.

■ **Definition: Local topological equivalence**
Two flows $\phi_t(x)$ and $\psi_t(y)$ are locally topologically equivalent near the states $\bar{x}$ and $\bar{y}$ if the flows, restricted to two neighbourhoods of $\bar{x}$ and $\bar{y}$ respectively, are topologically equivalent.

We need one more extension of the definition of topological equivalence, that applies to flows that depend on parameters. Its use will become apparent once we will meet bifurcations and their unfoldings as parametric systems.

■ **Definition: Topological equivalence for parametric flows**
Two flows $\phi_t(x, \alpha)$ and $\psi_t(y, \beta)$ in the parameters α and β are topologically equivalent if there exists a homeomorphism $\beta = b(\alpha)$, and $\phi_t(x, \alpha)$ and $\psi_t(y, b(\alpha))$ are topologically equivalent for all α.

In a nutshell, topological equivalence is extended to parametric families of flows by asking that parameters be mapped through a homeomorphism between the two families and that the two flows be topologically equivalent for each parameter pair.

Exercises

Exercise 27

Decide which of these maps are homeomorphisms or diffeomorphisms:

$$y = x^2, \quad x \in \mathbb{R};$$
$$y = \begin{pmatrix} 1 & 2 \\ -1 & 1 \end{pmatrix} x, \quad x \in \mathbb{R}^2;$$
$$y = \begin{cases} \sqrt{x}, & x \geq 0, \\ -\sqrt{-x}, & x < 0, \end{cases}, \quad x \in \mathbb{R};$$
$$y = x + e^x, \quad x \in \mathbb{R} \text{ (* this is a bit harder do discuss).}$$

Exercise 28

Prove that the maps

$$x(t+1) = \begin{cases} 2x(t), & x \leq 0.5, \\ 2(1 - x(t)), & x > 0.5 \end{cases}$$

and

$$x(t+1) = 4x(t)(1 - x(t)),$$

defined for $x \in [0,1]$, are topologically conjugate with homeomorphism

$$h(x) := \sin^2\left(\frac{\pi x}{2}\right), x \in [0,1].$$

Exercise 29

Consider the system

$$\dot{x} = \begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} x + \begin{pmatrix} x_1 x_2 \\ 0 \end{pmatrix}.$$

Define a diffeomorphism that diagonalizes the Jacobian of the equilibrium in 0, and compute the vector field in the new variables.

Answer of exercise 29

The linear change of variables diagonalizing A will work. We start by finding the eigenvectors of A:

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} p = 1p \Rightarrow p = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and

$$\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} p = 2p \Rightarrow p = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

We can thus diagonalize the system through the change of variables

$$y = Tx,$$

with

$$T^{-1} = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

We have

$$T = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix},$$

therefore

$$\begin{aligned} y_1 &= x_1 - 2x_2, \\ y_2 &= x_2. \end{aligned}$$

and

$$\begin{aligned} x_1 &= y_1 + 2y_2, \\ x_2 &= y_2. \end{aligned}$$

Using these in the original system, we obtain

$$\begin{aligned} \dot{y} &= J_h(y)^{-1} f(h(y)) = Tf(T^{-1}y) \\ &= \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \left(\begin{pmatrix} 1 & 2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} y + \begin{pmatrix} (y_1 + 2y_2)y_2 \\ 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} y + \begin{pmatrix} (y_1 + 2y_2)y_2 \\ 0 \end{pmatrix} \end{aligned}$$

** Exercise 30 Node to focus homeomorphism

Prove that the linear node

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

and the linear focus

$$\dot{x} = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix} x$$

are topologically conjugate.

Answer of exercise 30

The orbits of the linear node are straight lines approaching the origin as e^{-t}. The linear focus instead has eigenvalues $-1 \pm i$; its orbits spiral towards the origin, with $\rho(t) = e^{-t}\rho(0)$, $\theta(t) = t + \theta(0)$.

We can map the orbits of the node onto those of the focus by twisting them around the origin. Let us call (ρ_n, θ_n) the coordinates of the node system, and (ρ_f, θ_f) those of the focus system, and let us fix the circle $\rho = 1$ as the set of initial conditions for the orbits so that $\rho(0) = 1$. We have

$$\begin{aligned} \rho_n(t) &= e^{-t}, \\ \theta_n(t) &= \theta_n(0), \end{aligned}$$

for the node system, and

$$\rho_f(t) = e^{-t},$$
$$\theta_f(t) = \theta_f(0) + t,$$

for the focus system. To twist the orbit of the first system onto those of the second, it is sufficient to rotate any state (ρ, θ) by an angle $-\ln(\rho)$, thus the map

$$\begin{pmatrix} \rho_f \\ \theta_f \end{pmatrix} = h \begin{pmatrix} \rho_n \\ \theta_n \end{pmatrix} := \begin{pmatrix} \rho_n, \\ \theta_n - \ln(\rho_n) \end{pmatrix}$$

takes orbits of one system onto orbits of the other. The map is invertible and can be made continuous in a neighbourhood of the origin by defining

$$\theta_f = \theta_c \quad \text{when } \rho_c = 0.$$

It is therefore a homeomorphism.

Having constructed a homeomorphism between the orbits of the two systems, we may conclude that the flows are topologically equivalent, while the exercise asked for topological conjugacy. This is, however, guaranteed by the fact that the time domains over which the flows of the two systems are defined are the same. Let us verify this by explicitly testing the relation $h(\phi_t(x_n)) = \psi_t(h(x_n))$. Let us define $x_n := \begin{pmatrix} \rho_n \\ \theta_n \end{pmatrix}$, $x_f := \begin{pmatrix} \rho_f \\ \theta_f \end{pmatrix}$, $\phi_t(x_n) := \begin{pmatrix} e^{-t}\rho_n \\ \theta_n \end{pmatrix}$, $\psi_t(x_f) := \begin{pmatrix} e^{-t}\rho_f \\ \theta_f + t \end{pmatrix}$.

We have

$$h(\phi_t(x_n)) = \begin{pmatrix} e^{-t}\rho_n \\ \theta_n - \ln\left(e^{-t}\rho_n\right) \end{pmatrix} = \begin{pmatrix} e^{-t}\rho_n \\ \theta_n + t - \ln\left(\rho_n\right) \end{pmatrix}$$

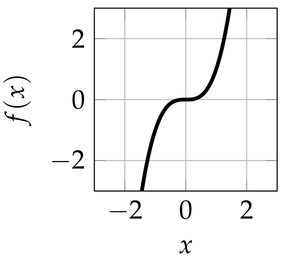
and

$$\psi_t(h(x_n)) = \begin{pmatrix} e^{-t}\rho_n \\ \theta_n - \ln\left(\rho_n\right) + t \end{pmatrix}.$$

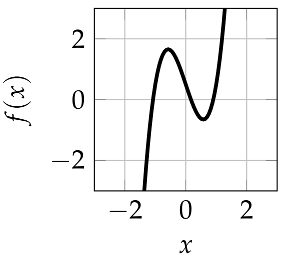
This proves that the homeomorphism h indeed establishes a topological conjugacy between the flows.

Exercise 31

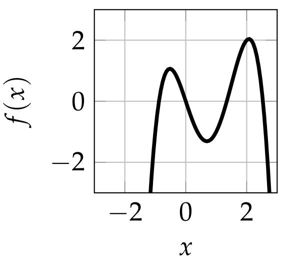
Decide which of the following scalar continuous-time systems have topologically equivalent flows:


f(x)
2
0
−2
−2 0 2
x

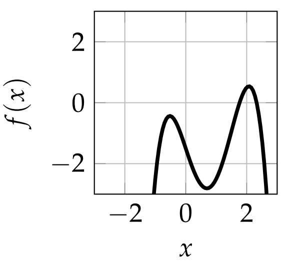
1.


f(x)
2
0
−2
−2 0 2
x

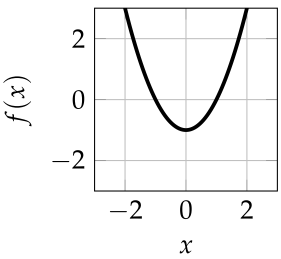
2.


f(x)
2
0
−2
−2 0 2
x

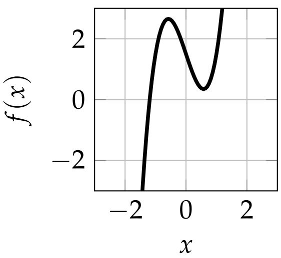
3.


f(x)
2
0
−2
−2 0 2
x

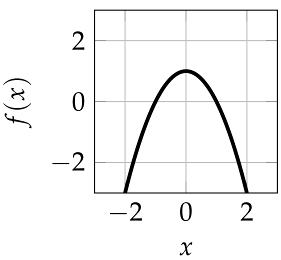
4.


f(x)
2
0
−2
−2 0 2
x

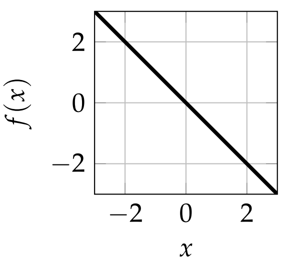
5.


f(x)
2
0
−2
−2 0 2
x

6.


f(x)
2
0
−2
−2 0 2
x

7.


f(x)
2
0
−2
−2 0 2
x

8.

Normal forms

Keywords: **Hartman-Großman theorems, centre manifold theorem, stable unstable and centre manifolds, normal form, Poincaré normal form, feedback linearisation normal form**

Invariant manifolds near an equilibrium

In the previous chapter, we established relations between flows of discrete or continuous-time dynamical systems. The concepts of topological conjugacy and topological equivalence are the bases on which the theory of normal forms is constructed. This is the theory dealing with how a system can be transformed into an equivalent but simpler form (in some sense that, of course, depends on the application), and with how systems can be classified according to their topological properties.

In this chapter, to keep the discussion simple, we break the symmetry that we have preserved so far between discrete and continuous time, and present the theory only for continuous-time systems, for the sake of simplicity. A similar theory can of course be constructed for discrete-time systems, through predictable changes in the algebra.

♦ **Hartman-Großman theorem**
If $\bar{x}$ is a hyperbolic equilibrium of a continuously differentiable vector field with flow $\phi_t(x)$, then $\phi_t(x)$ is locally topologically conjugate to its linearisation at $\bar{x}$.

The theorem states that ϕ_t is topologically conjugate to its linearisation, but in general it is not diffeomorphic to it.

Example 26. *Consider the system*

$$\dot{x} = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} x + \begin{pmatrix} 1 + x_2 \\ -1 \end{pmatrix} u + \begin{pmatrix} x_1^2 \\ 2x_1x_2 + x_2^2 \end{pmatrix}.$$

When $u = 0$, the system has an equilibrium in 0 with eigenvalues $\pm\sqrt{2}$ (it is a saddle). We stabilize the system at $u = 0$, $x = 0$ with a static state feedback

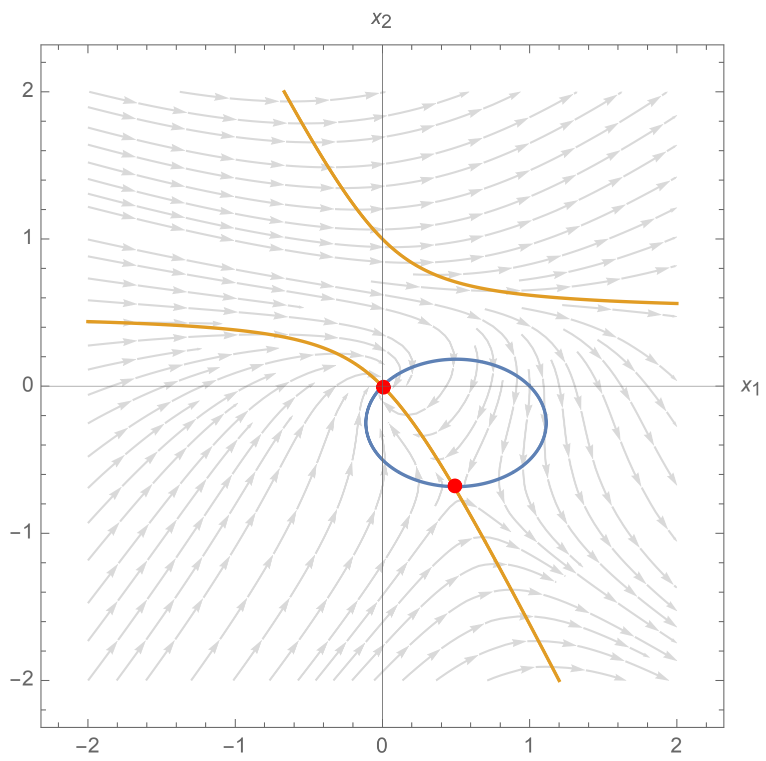
$$u = 2x_2.$$

The linearised controlled system is

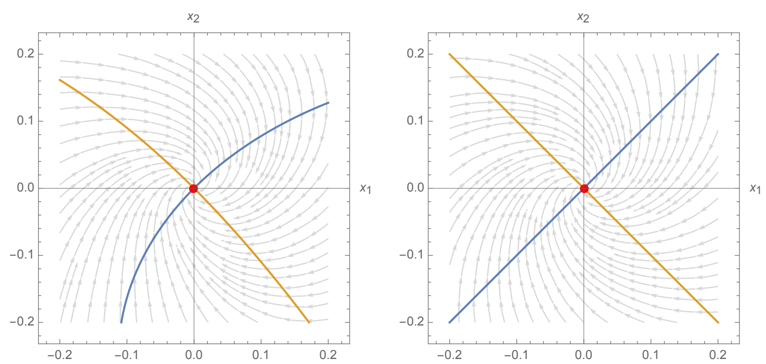
$$\dot{x} = \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 0 & 2 \end{pmatrix} x = \begin{pmatrix} -1 & 1 \\ -1 & -1 \end{pmatrix} x,$$

which has eigenvalues $-1 \pm i$*. The Hartman-Großman theorem guarantees that the nonlinear system is topologically conjugate to the linear one near the origin, therefore there exists a neighbourhood of* 0 *where orbits approach* 0*. Outside of this neighbourhood, however, dynamics may be very different from what we would expect.*

Here is the phase portrait of the controlled nonlinear system around the origin,



and a zoom-in of the nonlinear flow (on the left), and of the linearised flow (on the right)



The Hartman-Großman theorem thus justifies the use of linearisation to study the dynamics of hyperbolic equilibria and to design controllers. However, it does not give any information regarding non-hyperbolic equilibria. For this kind of information, we must use a different tool.

♦ **Centre manifold theorem**
Consider a continuously differentiable vector field $f(x)$ near the equilibrium $\bar{x}$. Consider the eigenspaces E^s, E^u, E^c spanned by the stable, unstable, and zero eigenvectors of $J_f(\bar{x})$. In a neighbourhood of $\bar{x}$ there exist three continuously differentiable invariant manifolds, respectively tangent to E^s, E^u, and E^c at $\bar{x}$.

In this statement, a continuously differentiable invariant manifold is an invariant subset of $\mathbb{R}^n$ with the shape of a surface whose tangent space changes continuously.

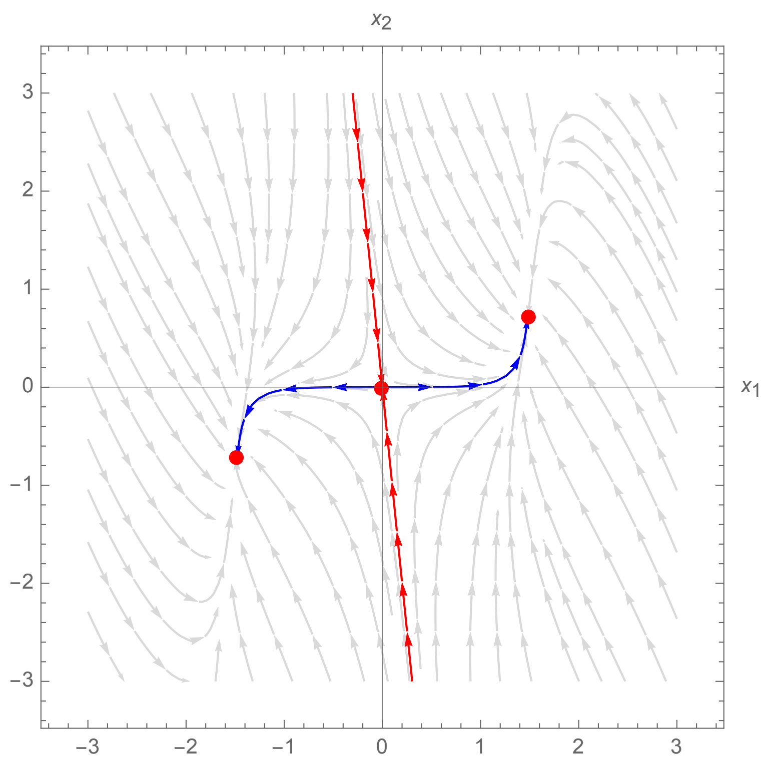
By analogy with their linear counterpart, let us denote by E^s, E^u, E_c the above-mentioned manifolds.

■ **Definition: Stable, unstable, and centre manifold**
The three manifolds E^s, E^u, and E^c are called the stable, unstable, and centre manifold, respectively.

Example 27. *The phase portrait below is of the system*

$$\dot{x}_1 = x_1 - \frac{x_1^3}{2} + \frac{x_2}{5},$$
$$\dot{x}_2 = \frac{x_1^5}{10} - x_2.$$

It has a saddle in $(0,0)$*, and two stable nodes. The unstable manifold of the saddle, in blue, converges to the two nodes. The stable manifold, in red, marks the boundary between the orbit converging to one or the other node.*



Stable manifolds of saddles are sometimes called *separatrices*, as they separate regions of the state space whose orbit have different long-term behaviour, as in this case.

♦ **Nonhyperbolic Hartman-Großman**
Let $\bar{x}$ be a nonhyperbolic equilibrium of a continuously differentiable vector field $f(x)$. The flow of $f(x)$ is locally topologically conjugate to the flow of a system

$$\begin{aligned}\dot{x} &= Cx + F(x, y, z)\\ \dot{y} &= Sy\\ \dot{z} &= Uz\end{aligned}$$

where C, S and U are matrices whose eigenvalues have zero, negative and positive real parts, respectively, the block-diagonal matrix with blocks C, S and U is similar to $J_f(\bar{x})$, and F is a nonlinear function.
The three sets of equations define the dynamics in the manifolds E^c, E^s, and E^u, respectively

Note that, unlike what is stated in the (hyperbolic) Hartman-Großman theorem, here the system is *not* conjugate to a linear system. The dynamics in the centre manifold is not equivalent to a linear dynamics.

Example 28. *Consider the control system*

$$\begin{aligned}\dot{x} &= \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \end{pmatrix} u + \begin{pmatrix} 0 \\ x_2^3 \end{pmatrix},\\ y &= x_1.\end{aligned}$$

We design a feedback control $u = -2y = -2x_1$, obtaining the closed loop dynamics

$$\dot{x} = \begin{pmatrix} -1 & 0 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ x_2^3 \end{pmatrix}.$$

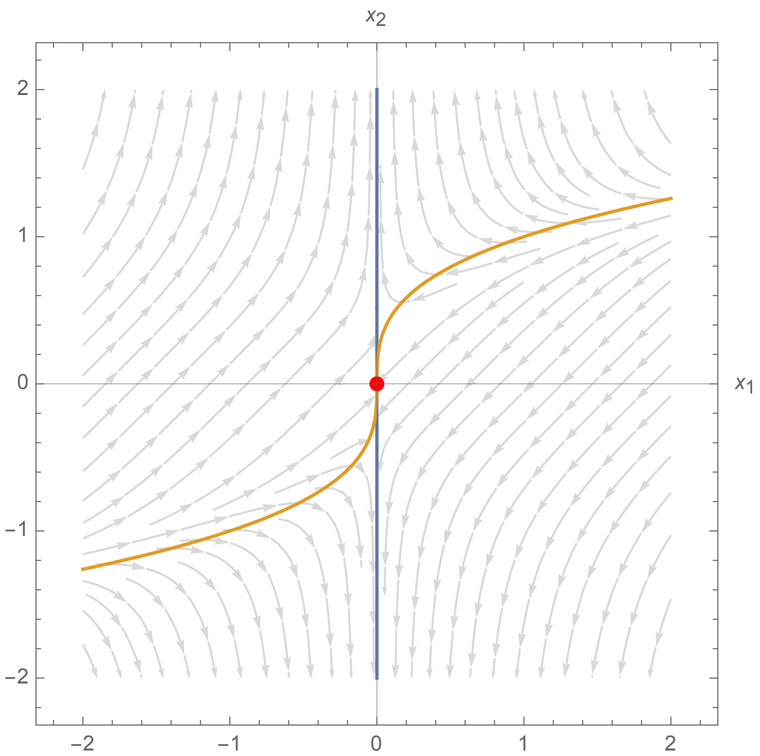
The system has an equilibrium at $x = 0$, with an eigenvalue -1, corresponding to eigenvector $(1, 1)$, and an eigenvalue 0 with eigenvector $(0, 1)$. From the analysis of the linearised dynamics, we would expect the origin to be Lyapunov stable, with small perturbations converging to the null eigenvector, which coincides with the x_2 axis. This may be good enough for our purpose... if it were true.

The Hartman-Großman theorem tells us that the dynamics along the eigenvector $(1, 1)$ will indeed resemble those of the linearisation, at least near the origin. This is not true, however, along the null eigenvector.

The nonlinear term x_2^3 in the equation of $\dot{x}_2$ destabilizes the x_2 axis:

$$|\dot{x}_2| > 0, \forall x_1 = 0, x_2 \neq 0.$$

In the nonlinear system, the origin therefore behaves like a saddle. We see here the phase portrait of the nonlinear system.



An equilibrium such as this one, which has a stable and an unstable direction, despite not having both a stable and an unstable manifold, is called a *topological saddle*

Example 29 (The SIR model and the epidemic threshold). *The following model, known as the SIR model (Kermack and McKendrick, 1927), is a simplistic representation of a rapidly spreading infectious disease. It is frequently used as the basis to develop more articulate, predictive models of diseases such as influenza or SARS-COV-2.*

$$\dot{S} = -aIS,$$
$$\dot{I} = aIS - bI,$$
$$\dot{R} = bI.$$

The variables S, I, and R represent the fraction of individuals in a given population that are susceptible, infected, or removed (recovered or dead) at a given time instant, while the positive parameters a and b represent the transmission and removal rates. The model ignores the flows of newborns and dead by causes other than the disease, which is the reason why it is only used in this form to study diseases that spread over a faster time scale than natural birth and death. In the model, $S(0)$, $I(0)$ and $R(0)$ are assumed to sum to 1 (to make up the entirety of the population), and it is easy to verify that the simplex

$$\{(S, I, R) : S + I + R = 1, S \geq 0, I \geq 0, R \geq 0\} \tag{1}$$

is positively invariant.

All equilibria of the model lie in the set $\{(S, I, R) : I = 0\}$. Let us study the stability of these equilibria, and to simplify our task let us ignore the third equation (note how the first two equations are independent of R, so they form a closed system).

We have

$$J_f(S, I = 0) = \begin{pmatrix} 0 & -aS \\ 0 & aS - b \end{pmatrix}.$$

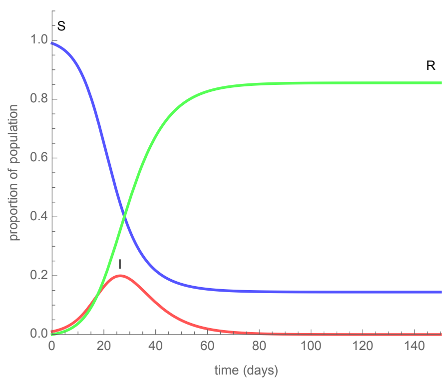


Figure 4: Evolution of S, I, and R with $S(0) = 0.99$, $I(0) = 0.01$, $R(0) = 0$, $a = 2.25/7$, $b = 1/7$.

The matrix has a null eigenvalue, with eigenvector $(1, 0)$*, and a second eigenvalue* $aS - b$*, with eigenvector* $(-aS, aS - b)$*. We can thus divide equilibria into* 3 *groups:*

1. *equilibria with* $aS - b > 0$*, which have an unstable eigenvalue,*
2. *one equilibrium with* $aS - b = 0$*, which has a second null eigenvalue,*
3. *equilibria with* $aS - b < 0$*, which have a stable eigenvalue.*

Let us study the nonlinear dynamics of equilibria in each group. Equilibria in group 1 have a centre manifold, locally contained in the set $\{(S, I) : I = 0\}$*, and an unstable manifold. No orbit can reach these equilibria forward in time. The equilibrium in group 2 has a centre manifold and no stable or unstable manifold. Notice also how the matrix in this case is not diagonalizable (the centre eigenspace is 1-dimensional). In this case, the Hartman-Großman theorem does not provide any additional information, and the full* 2*-dimensional dynamics describes the behaviour in the centre manifold. We can however note that the equilibrium (call it* (S_∞, I_∞)*) can only be reached from states* $S : S > S_\infty$*, since* S *is strictly decreasing. The dynamics of* I *near the equilibrium is thus equal to*

$$\dot{I} = aIS - bI = aI(S - S_\infty) + \underbrace{(aS_\infty - b)}_{=0} I = aI(S - S_\infty) > 0.$$

Orbits within the centre manifold cannot reach the equilibrium $(S_\infty, I_\infty = 0)$*. Finally, equilibria in group* 3 *have a stable manifold, and can therefore be reached by orbits within the simplex (1). All in all, we see that all orbits of the system converge to equilibria with state* $S : aS - b < 0$*, that is, with* $S < \frac{b}{a}$*. The value* $\frac{a}{b}$*, also known as* $\mathcal{R}_0$*, is called the basic reproduction number, and* $1/\mathcal{R}_0$ *is called the epidemic threshold. If* $\mathcal{R}_0 > 1$ *then the state* $(S = 1, I = R = 0)$ *is unstable, a small perturbation of the infected compartment will trigger an epidemic which will not stop before* S *has reached a limiting value* $S_\infty \leq \frac{1}{\mathcal{R}_0}$*. If* $\mathcal{R}_0 < 1$ *then the state* $(S = 1, I = R = 0)$ *is stable: a small number of infected individuals cannot trigger an epidemic.*

The nonhyperbolic Hartman-Großman theorem states that, near any nonhyperbolic equilibrium, dynamics can be partitioned into three sets corresponding to the three manifolds. Dynamics in the centre manifold, unlike dynamics in the other two manifolds, are completely decided by higher-order terms. However, in most cases, we will still observe convergence or divergence from the equilibrium, even in the centre manifold. That is, in most cases if we ignore the time-dependence of the orbits (and only focus on topological equivalence) we can expect to see dynamics qualitatively similar to those that we know for hyperbolic equilibria.

★ Planar nonhyperbolic equilibria that are locally topologically equivalent to a node, focus, or saddle, are called topological nodes, foci, and saddles.

An important consequence of the centre manifold and the Hartman-Großman theorem is that, when we need to study a nonhyperbolic equilibrium, we can focus on the analysis of the reduced model

$$\dot{x} = Cx + F(x, y, z),$$

since the dynamics in the stable and unstable manifolds are fully defined by the corresponding, hyperbolic eigenvalues and eigenvectors. This is good news, since in most cases of interest the centre manifold is of dimension 1 or 2, depending on whether a single real eigenvalue, or a pair of complex conjugate eigenvalues, are crossing the imaginary axis, and independently of the order of the system we are studying.

In light of this observation, one may be interested in seeing whether the dynamics of most nonhyperbolic equilibria in their centre manifold can be captured and classified by some sort of standardized equations. To some extent, this is indeed possible, through the use of *normal forms*, which we start to acquaint with in the next section.

Normal forms

A normal form of a vector field is a special functional form, in which the vector field can be transformed through a suitable diffeomorphism (that is, by a nonlinear change of variables). This is of course a very loose definition, but it is meant to be so, as normal forms are many and designed for different purposes. We see next two families of normal forms that are frequently encountered in nonlinear dynamics and control, and we will encounter a third family –topological normal forms– when we will talk about bifurcations.

Refer to (Murdock, 2003) for a thorough discussion of normal forms theory

To simplify the writing of our normal forms, we will use two common tools of differential geometry, the Lie derivative of scalar and vector functions.

■ **Definition: Lie derivative of a scalar function**
The Lie derivative of a scalar function $g(x)$ with respect to vector field $f(x)$ is

$$L_f g := \frac{\partial g}{\partial x} f(x) = J_g(x) f(x).$$

Note how this is the derivative of the function $g(x)$ in the direction of the flow of $f(x)$. We already encountered it, for example in Lyapunov theorem, where $\dot{V}(x) = L_f V(x) = \frac{\partial V}{\partial x} f(x)$.

The Lie derivative of a scalar function g with respect to vector field f is the directional derivative of the function, with respect to the flow of the vector field.

■ **Definition: Lie bracket of vector functions**
The Lie bracket of a vector functions $g(x)$ and $f(x)$ is

$$[f, g] := \frac{\partial g}{\partial x} f(x) - \frac{\partial f}{\partial x} g(x) = J_g(x) f(x) - J_f(x) g(x).$$

With these tools we can start to address the first family of normal forms: the Poincaré normal form. This is, loosely speaking, a vector field representation that makes the system as close as possible to its linearisation in a neighbourhood of one of its equilibria.

Consider a sufficiently differentiable vector field $f(x)$ with an equilibrium at 0. We want to find a coordinate change (hence a diffeomorphism) that transforms it as much as possible into its linearisation $J_f(0)x$.

We can start by writing the Taylor expansion of $f(x)$:

$$\dot{x} = J_f(0)x + f^{[2]}(x) + f^{[3]}(x) + \dots,$$

where $f^{[i]}$ are homogeneous polynomials of degree i. We aim at constructing the change of variables by eliminating the terms $f^{[i]}$ one degree at a time. Let us consider, to begin with, a near-identity change of variables

$$x = h(y) = y + \psi^{[2]}(y),$$

where $\psi^{[2]}(y)$ is an arbitrary second-degree polynomial. Its inverse is

$$y = x - \psi^{[2]}(x) + O(\|x\|^3).$$

To see this, in the scalar case, given $x = y + py^2$ consider the Taylor expansion of the unknown inverse function

$$h^{-1}(x) = c_1 y + c_2 y^2 + c_3 y^3 + \dots.$$

Iteratively solving $h(h^{-1}(x)) = x$ we find the coefficients of h^{-1}: we have $c_1 = 1$ and $c_2 = -p$, $c_3 = 2p^2 \dots$

Differentiating both sides of the above equation in time, we obtain

$$\begin{aligned}
\dot{y} &= \dot{x} - J_{\psi^{[2]}}(x)\dot{x} + O(\|x\|^3) \\
&= J_f(0)x + f^{[2]}(x) - J_{\psi^{[2]}}(x)J_f(0)x + O(\|x\|^3) \\
&= J_f(0)y + J_f(0)\psi^{[2]}(y) + f^{[2]}(y) - J_{\psi^{[2]}}(y)J_f(0)y + O(\|y\|^3) \\
&= J_f(0)y + f^{[2]}(y) - [J_f(0)y, \psi^{[2]}(y)] + O(\|y\|^3).
\end{aligned}$$

Thus, we can eliminate all second-degree terms by choosing $\psi^{[2]}$ such that

$$[J_f(0)y, \psi^{[2]}(y)] = f^{[2]}(y).$$

This is known as the *homological equation...*

The same formal computation can be repeated for higher-order terms.

We can see the above as a linear operator

$$[J_f(0)y, \cdot],$$

...while this is the *homological operator.*

acting on the (finite-dimensional) space of second-degree homogeneous polynomials. In the case that $J_f(0)$ is diagonalizable, it can be shown that the eigenvalues of the homological equation, seen as a linear operator on the polynomials ψ, are integer multiples of the quantity $m_1\lambda_1 + \dots + m_n\lambda_n - \lambda_i$, where n is the dimension of the system, λ_i are the eigenvalues of $J_f(0)$, m_i are nonnegative integers, and $\sum m_i = 2$. It follows that the homological equation is a nonsingular operator provided that the equation

$$m_1\lambda_1 + \dots + m_n\lambda_n - \lambda_i = 0$$

has no solution in the variables m_i. This is called a *non-resonance* condition. Combinations of $m_1, \dots, m_n$ that satisfy the equation corresponds monomials, of the form $y_1^{m_1} y_2^{m_2} \dots y_n^{m_n}$, that cannot be eliminated through the Poincaré normal form reduction. These are called *resonant terms* in the normal form. The condition for the existence of resonant terms, which we have seen for the second-order terms, is similar for terms of higher order.

■ **Definition: Poincaré normal form**
The Poincaré normal form of a vector field near an equilibrium is the functional form of $f(x)$ once all non-resonant terms have been eliminated, by iterating the above procedure.

Example 30 (Poincaré normal form of a 1D hyperbolic equilibrium)**.** *Take the scalar system*

$$\dot{x} = -x + 2x^2 - x^3,$$

with an equilibrium in $x = 0$ *and* $J_f(0) = -1$. *Let us construct a diffeomorphism to remove the second-order term. We must find* $\psi^{[2]}$ *such that*

$$[-y, \psi^{[2]}] = -\frac{\partial}{\partial y}\psi^{[2]} y + \psi^{[2]} = 2y^2.$$

The function $\psi^{[2]}$*, combined with its derivatives, must be equal to a quadratic term, so it is fair to assume* $\psi^{[2]} = ay^2$. *We obtain*

$$-2ay^2 + ay^2 = 2y^2 \Rightarrow a = -2,$$

We should thus use the local diffeomorphism $x = y - 2y^2$ *(and* $y = x + 2x^2 + O(|x|^3)$*) obtaining*

$$\begin{aligned}\dot{y} &= J_h^{-1}(y) f(h(y)) \\ &= \frac{1}{1-4y}(-y + 2y^2 + 2y^2 + 8y^4 - 8y^3 - (y - 2y^2)^3) \\ &= -y\frac{(1+4y)}{1+4y} + O(|y|^3) = -y + O(|y|^3)\end{aligned}$$

We have eliminated the second-order term. We could proceed to eliminate higher order terms in the same manner, though computation of the coefficients becomes rapidly cumbersome.

Example 31 (Poincaré normal form of a saddle-node)**.** *Consider the system*

$$\dot{x} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} x + \begin{pmatrix} 2x_1^2 \\ x_1 x_2 + x_2^2 \end{pmatrix}.$$

Let us solve the homological equation for the second-order terms

$$\frac{\partial}{\partial y}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} y - \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} 2y_1^2 \\ y_1 y_2 + y_2^2 \end{pmatrix}.$$

The equation above, assuming $\psi_i(y) = a_i y_1^2 + b_i y_1 y_2 + c_i y_2^2$*, can be written as*

$$\begin{pmatrix} 2a_1 y_1 + b_1 y_2 & b_1 y_1 + 2c_1 y_2 \\ 2a_2 y_1 + b_2 y_2 & b_2 y_1 + 2c_2 y_2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} y - \begin{pmatrix} a_1 y_1^2 + b_1 y_1 y_2 + c_1 y_2^2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2y_1^2 \\ y_1 y_2 + y_2^2 \end{pmatrix},$$

We can easily prove the resonance condition at second order for 1 and 2-dimensional systems. Start by considering a 1-dimensional system. The unknown polynomials $\psi(y) = ay^2$ can be represented as the space of 1 dimensional vectors $a \in \mathbb{R}$. Assuming an equilibrium with Jacobian $J_f(0) = \lambda$, the homological operator is

$$(2\lambda - \lambda) a y^2.$$

This can be seen as the linear operator $(2\lambda - \lambda)$ applied to the 1-dimensional vector a, and its eigenvalue is nonzero provided that $2\lambda - \lambda \neq 0$.

In a 2-dimensional system, the unknown polynomials $\psi_1(y) = a_1 y_1^2 + b_1 y_1 y_2 + c_1 y_2^2$ and $\psi_2(y) = a_2 y_1^2 + b_2 y_1 y_2 + c_2 y_2^2$ can be represented by the 6-dimensional vector $(a_1, b_1, c_1, a_2, b_2, c_2)$. If we assume that the Jacobian $J_f(0)$ is diagonal with eigenvalues λ_i on the diagonal, then the matrix representation of the homological operator is itself diagonal, and equal to

$$\begin{pmatrix} 4\lambda_1 - 2\lambda_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 + \lambda_2 - \lambda_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4\lambda_2 - 2\lambda_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4\lambda_1 - 2\lambda_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 + \lambda_2 - \lambda_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4\lambda_2 - 2\lambda_2 \end{pmatrix}$$

Once again, its eigenvalues are nonzero provided that $m_1\lambda_1 + m_2\lambda_2 - \lambda_i \neq 0$, for all $i \in \{1, 2\}$ and for all $m_1, m_2 \geq 0$ such that $m_1 + m_2 = 2$.

that is,

$$\begin{pmatrix} a_1 y_1^2 - c_1 y_2^2 \\ 2a_2 y_1^2 + b_2 y_1 y_2 \end{pmatrix} = \begin{pmatrix} 2y_1^2 \\ y_1 y_2 + y_2^2 \end{pmatrix}.$$

Comparing term by term we obtain the set of conditions

$$\begin{cases} a_1 = 2, \\ c_1 = 0, \\ a_2 = 0, \\ b_2 = 1. \end{cases}$$

We do not have conditions on b_1 and c_2, so we can put them at 0*, but we also do not have the means to eliminate the y_2^2 term in the second equation. This is a resonant term, and we should have expected to find one, given that the eigenvalues $\lambda_1 = 1$ and $\lambda_2 = 0$ of $J_f(0)$ satisfy $0\lambda_1 + 2\lambda_2 - \lambda_2 = 0$. Note that, in general, resonant terms are not unique and depend on the system's representation.*

The above coefficients give

$$[J_f(0)y, \psi^{[2]}(y)] = \begin{pmatrix} 2y_1^2 \\ y_1 y_2 \end{pmatrix},$$

therefore, in the new variables, the system becomes

$$\dot{y} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} y + \begin{pmatrix} 0 \\ y_2^2 \end{pmatrix} + O(\|y\|^3).$$

Normal form reduction, to degree 2*, has given us a definition (up to degree* 2*) of the dynamics within the stable manifold of the saddle ($\dot{y}_1 = y_1$) and within the centre manifold ($\dot{y}_2 = y_2^2$).*

The fact that y_2^2 is resonant means that this term, or an equivalent one, persists in any possible representation of the system, among all representations that can be obtained by polynomial change of variables! It is, in some sense, a more relevant nonlinear term than the others. We can understand why this is so also by looking at what would happen if we perturbed the equation of $\dot{y}_2$ by an arbitrary small constant:

$$\dot{y}_2 = \epsilon + y_2^2 + O(\|y\|^3).$$

Near 0 *we would now have* 0 *or* 2 *equilibria, depending on the value of ϵ, and this is due to the resonant nonlinear term.*

In the above discussion, we have seen how to change variables to make a system as similar as possible to a linear system, without touching its linear part. This is, of course, not the only normal form of interest. Another one, particularly interesting in nonlinear control, is the feedback linearisation normal form.

We begin by considering a nonautonomous nonlinear system with an affine dependence on the input, that is, a system of the form

$$\dot{x} = f(x) + g(x)u.$$

Let us assume for simplicity that u is a scalar signal. The question is whether, by exploiting the additional degree of freedom provided by the input u, we can find a change of variable *and a static feedback input* such that the feedback system is linear.

To do so, we use the Lie derivative notation, which we remember is the directional derivative along the flow of a scalar function. Assume that we have a scalar function $h(x)$ such that the system satisfies the following condition

Definition: Feedback linearisation conditions
Let $x \in \mathbb{R}^n$, and N be an open subset of $\mathbb{R}^n$. There exists $h(x) : N \to \mathbb{R}$ such that, for all $x \in N$,

$$\begin{aligned} L_g h(x) &= 0, \\ L_g L_f h(x) &= 0, \\ &\vdots \\ L_g L_f^{n-2} h(x) &= 0, \\ L_g L_f^{n-1} h(x) &\neq 0. \end{aligned}$$

In the control literature, this condition implies that the system has *relative degree* n, but the notion of relative degree and its implications are beyond the scope of this discussion.

Then, we can prove the following result.

Theorem
If the feedback linearization conditions are satisfied, the mapping

$$z = \psi(x)$$

with

$$\begin{aligned} \psi_1 &:= h(x), \\ \psi_2 &:= L_f h(x), \\ &\vdots \\ \psi_n &:= L_f^{n-1} h(x), \end{aligned}$$

is a diffeomorphism in N.

A proof is found in (Isidori, 1995)

In the coordinates z we have

$$\dot{z}_i = \frac{\partial \psi_i(x)}{\partial x} \dot{x} = L_f \psi_i(x) + u L_g \psi_i(x).$$

This gives

$$\begin{aligned} \dot{z}_1 = \frac{\partial h}{\partial x} \dot{x} = \frac{\partial h}{\partial x} f(x) + \overbrace{\frac{\partial h}{\partial x} u g(x)}^{=u L_g h = 0} &= L_f h = z_2, \\ \dot{z}_2 &= z_3, \\ &\vdots \\ \dot{z}_{n-1} &= z_n, \\ \dot{z}_n &= L_f^n h(x) + L_g L_f^{n-1} h(x) u. \end{aligned}$$

Defining

$$\begin{aligned} a(z) &:= L_f^n h(\psi^{-1}(z)), \\ b(z) &:= L_g L_f^{n-1} h(\psi^{-1}(z)), \qquad (\neq 0 \text{ by assumption}), \end{aligned}$$

we obtain the new system

$$\begin{aligned} \dot{z}_1 &= z_2, \\ &\vdots \\ \dot{z}_n &= a(z) + b(z)u, \end{aligned}$$

which is finally linearised by setting

$$u(z) := \frac{v - a(z)}{b(z)}.$$

■ **Definition: Feedback linearisation normal form**
The feedback linearisation normal form is the system

$$\begin{aligned} \dot{z}_1 &= z_2, \\ &\vdots \\ \dot{z}_n &= a(z) + b(z)u. \end{aligned}$$

★ Note that the above system with

$$u(z) := \frac{v - a(z)}{b(z)}$$

is linear and controllable (reachable).

Example 32. *Consider the system*

$$\dot{x} = \begin{pmatrix} x_2 \\ x_1^2 + x_2^2 \\ x_1 - x_2 \end{pmatrix} + \begin{pmatrix} x_3 \\ x_3 \\ 0 \end{pmatrix} u, \quad h(x) = x_3.$$

We have

$$\begin{aligned} L_g h(x) &= \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} g & &= 0, \\ L_f h(x) &= \begin{pmatrix} 0 & 0 & 1 \end{pmatrix} f & &= x_1 - x_2, \\ L_g L_f h(x) &= \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} g & &= 0, \\ L_f^2 h(x) &= \begin{pmatrix} 1 & -1 & 0 \end{pmatrix} f & &= x_2 - x_1^2 - x_2^2, \\ L_g L_f^2 h(x) &= \begin{pmatrix} -2x_1 & 1 - 2x_2 & 0 \end{pmatrix} g & &= x_3(-2x_1 + 1 - 2x_2). \end{aligned}$$

For all $x_3 \neq 0$ and $(1 - 2x_1 - 2x_2) \neq 0$ the system is fully feedback linearisable, so the mapping

$$z = \begin{pmatrix} x_3 \\ x_1 - x_2 \\ x_2 - x_1^2 - x_2^2 \end{pmatrix}$$

is a local diffeomorphism. We can verify this by checking that

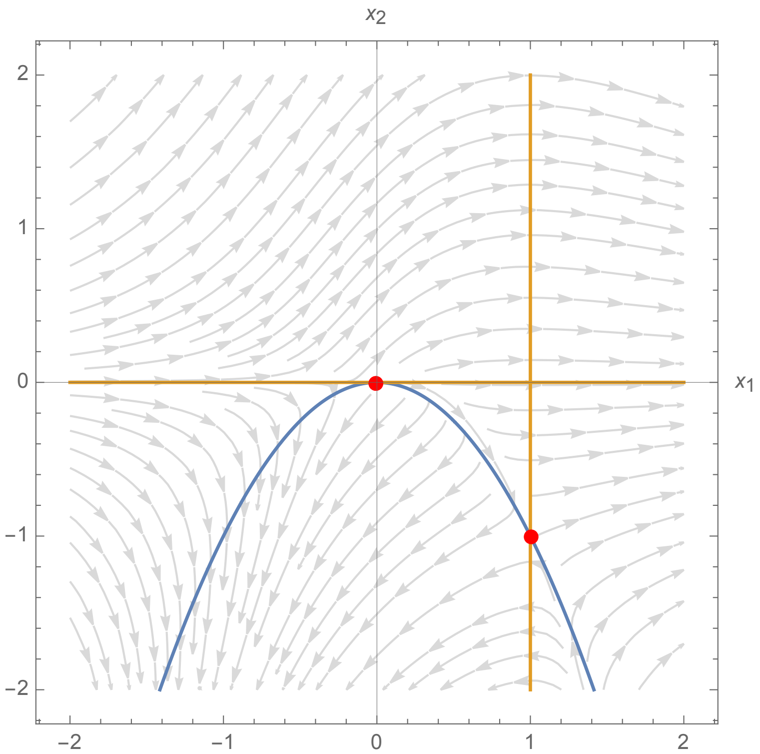
$$\det J_\psi(x) = \det \begin{pmatrix} 0 & 0 & 1 \\ 1 & -1 & 0 \\ -2x_1 & 1-2x_2 & 0 \end{pmatrix} \neq 0.$$

The condition $x_3 \neq 0$ ensures that a linearising input exists.

Example 33 (Limitations on the stability of the controlled equilibrium)**.** *The system*

$$\begin{aligned} \dot{x}_1 &= x_1^2 + x_2 + ux_1, \\ \dot{x}_2 &= -x_1x_2 + x_2 - ux_2, \end{aligned}$$

has an unstable equilibrium in $\bar{x} = (1, -1)$ for $\bar{u} = 0$.



We are going to stabilize it through feedback linearisation, with $h(x) = x_1x_2$. We have

$$\begin{aligned} L_g h(x) &= \begin{pmatrix} x_2 & x_1 \end{pmatrix} g & &= 0, \\ L_f h(x) &= \begin{pmatrix} x_2 & x_1 \end{pmatrix} f & &= x_1x_2 + x_2^2, \\ L_g L_f h(x) &= \begin{pmatrix} x_2 & x_1 + 2x_2 \end{pmatrix} g & &= -x_2^2, \\ L_f^2 h(x) &= \begin{pmatrix} x_2 & x_1 + 2x_2 \end{pmatrix} f & &= 3x_2^2 - 2x_1x_2^2 + x_1x_2; \end{aligned}$$

the system is, therefore, feedback linearisable in the domain $x_2 \neq 0$, through the change of variables

$$z = \psi(x) = \begin{pmatrix} x_1x_2 \\ x_1x_2 + x_2^2 \end{pmatrix}, \tag{2}$$

which has inverse

$$x = \psi^{-1}(z) = \begin{pmatrix} \pm\frac{z_1}{\sqrt{z_2 - z_1}} \\ \pm\sqrt{z_2 - z_1} \end{pmatrix}.$$

We already know that the largest region containing $\bar{x}$ where the diffeomorphism is defined is the set $\{x : x_2 < 0\}$, and using (2) we see that the image of $\{x : x_2 < 0\}$ through the diffeomorphism is the region $\{z : z_2 - z_1 > 0\}$. Therefore, ψ and ψ^{-1} define a diffeomorphism between compact subsets of the sets $\{x : x_2 < 0\}$ and $\{z : z_2 - z_1 > 0\}$.

The coordinates of $\bar{x}$ in the z variables are

$$\bar{z} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Setting

$$u = \frac{v - L_f^2 h(x)}{L_g L_f h(x)} = \frac{v - (3x_2^2 - 2x_1 x_2^2 + x_1 x_2)}{-x_2^2}, \tag{3}$$

we obtain the closed-loop dynamics

$$\begin{aligned} \dot{z}_1 &= z_2, \\ \dot{z}_2 &= v, \end{aligned}$$

and we want to stabilize the equilibrium $\bar{z}$ of the above system. Let us call A and B the dynamic and input matrix of the above linear system. We can assign the eigenvalues -1 and -2 to the dynamic matrix by setting

$$v = Kz + w = \begin{pmatrix} -2 & -3 \end{pmatrix} z + w,$$

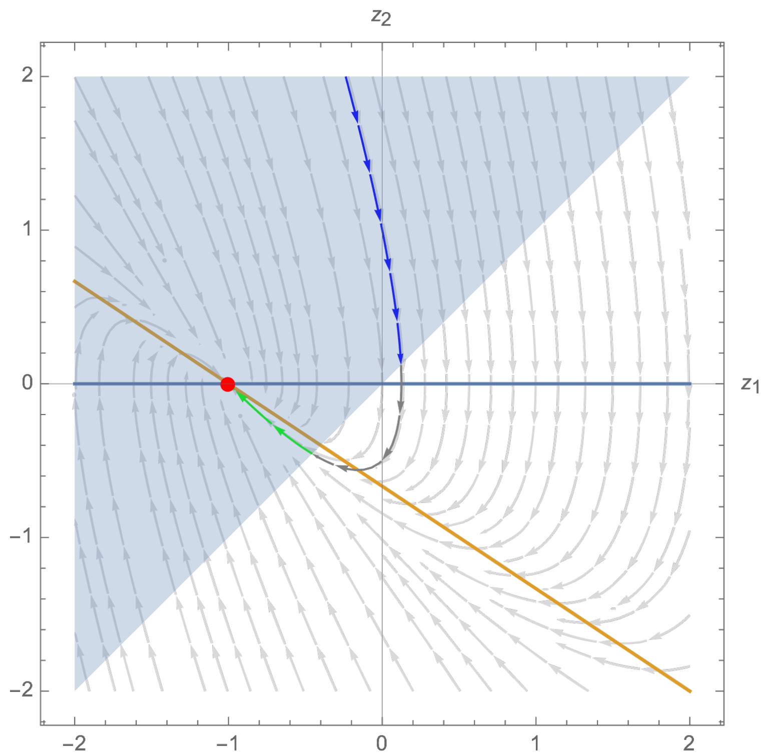
and then translate the equilibrium to $\bar{z}$ by choosing $w = \bar{w}$, where $\bar{w}$ is such that

$$-(A + BK)^{-1} B\bar{w} = \bar{z}.$$

This gives

$$B\bar{w} = -(A + BK)\bar{z},$$

that is, $\bar{w} = -2$.



Remember that the diffeomorphism maps the state space $\{x : x_2 \neq 0\}$ into the shaded region $\{z : z_2 - z_1 > 0\}$. States z out of this region are not related to any state in the x coordinates. For this reason, the orbit $\phi.(0,1)$, in this figure, is mapped onto two separate orbits in the x space, corresponding to the blue and the green segments respectively. The grey segment has no representation in the x space.

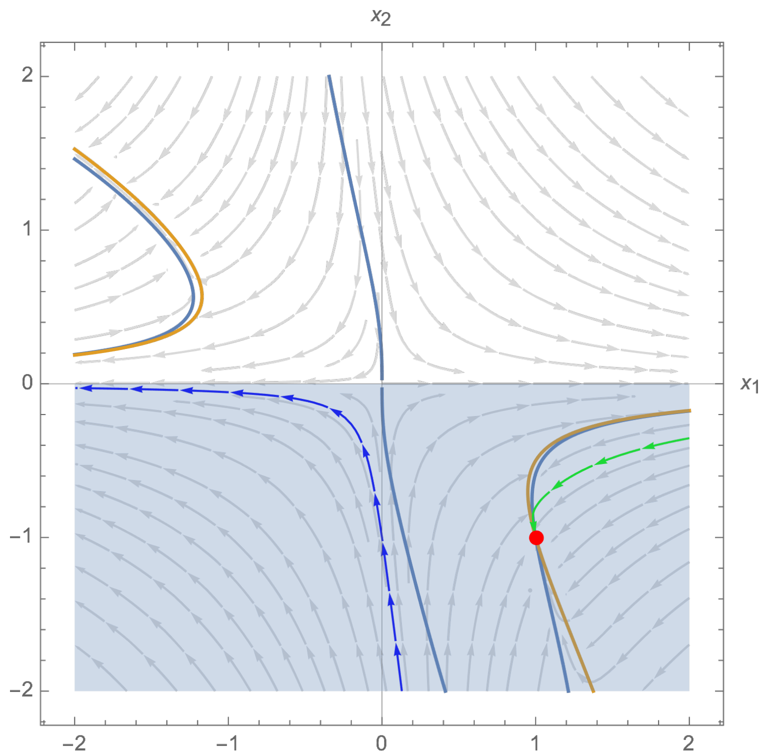
We can now use the expression of v in (2) and (3) to compute the input u:

$$u = \frac{\overbrace{-2x_1x_2 - 3(x_1x_2 + x_2^2)}^{Kz}\overbrace{-2}^{\bar{w}} - (3x_2^2 - 2x_1x_2^2 + x_1x_2)}{-x_2^2}$$
$$= \frac{-2 - 6x_1x_2 - 6x_2^2 + 2x_1x_2^2}{-x_2^2}.$$

This input stabilizes the equilibrium $\bar{x}$, assigning it the eigenvalues -1 and -2. We have the phase portrait above, in the z coordinates.

Now, we have seen how the diffeomorphism ψ maps compact subsets of the sets $\{x : x_2 < 0\}$ and $\{z : z_2 - z_1 > 0\}$, and we know that $\bar{z}$ is globally asymptotically stable. Does this mean that $\bar{x}$ is globally asymptotically stable? Obviously not, since the diffeomorphism that relates the two systems is only defined in $\{x : x_2 < 0\}$.

Then, does this at least imply that the equilibrium $\bar{x}$ attracts any initial condition in $\{x : x_2 < 0\}$? Unfortunately, even this is not true, because the set $\{z : z_2 - z_1 > 0\}$ is not positively invariant. We see that clearly in the phase portrait above, where the orbit $\phi.(0,1)$ (the orbit that crosses coordinates $z = (0,1)$) leaves the set before converging to the equilibrium. The image of this orbit through the diffeomorphism, $\psi^{-1}(\phi.(0,1))$, is split in two by the diffeomorphism (it is, effectively, two different orbits, represented in blue and green in the figures above and below). In particular, the blue orbit, despite belonging entirely to the set $\{x : x_2 < 0\}$, does not converge to $\bar{x}$.



This is the phase portrait of

$$\dot{x} = f(x) + \frac{Kz + \bar{w} - L_f^n h(\psi^{-1}(z))}{L_g L_f^{n-1} h(\psi^{-1}(z))}.$$

The equilibrium in $(1, -1)$ is now a stable node, with eigenvalues -1 and -2. However, it only attracts initial conditions in the subsets of $\psi^{-1}(\{z : z_2 - z_1 > 0\})$ that are positively invariant under the flow of $\dot{z} = (A + BK)z + B\bar{w}$. The blue and green orbits in this portrait are the images of the blue and green segments of the orbit in the previous figure. Notice how they correspond, here, to different orbits.

Exercises

Exercise 32

Discuss if the following statements are true, or false, or if their truth depends on further assumptions.

1. If an equilibrium $\bar{x}$ of $\dot{x} = f(x)$ is hyperbolic, then near $\bar{x}$ the flow is topologically equivalent to its linearization at $\bar{x}$.
2. If an equilibrium $\bar{x}$ of $\dot{x} = f(x)$ is hyperbolic, then the nonlinear terms of $f(x)$ near $\bar{x}$ can be removed up to an arbitrary order by a suitable change of variables.
3. System $\dot{x} = f(x, u)$ has feedback linearisation normal form

$$\dot{x} = \begin{pmatrix} x_2 \\ x_3 \\ x_4 \\ 1 + x_1^2 - x_4 x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 - x_2 \end{pmatrix} u.$$

It is feedback linearisable in the region

$$\|x\| < 2.$$

Exercise 33

Verify that the feedback linearisation condition holds for $h(x) = x_1$ on the following system, and compute the diffeomorphism that changes the system in the feedback-linearisation normal form:

$$\dot{x} = \begin{pmatrix} x_3 \\ x_1^2 + x_3 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 + x_1 \\ 0 \end{pmatrix} u.$$

Exercise 34

Compute the Poincaré normal form to the second order of

$$\dot{x} = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} x + \begin{pmatrix} 2x_1 x_2 \\ x_1^2 - x_1 x_2 \end{pmatrix}.$$

Exercise 35

A reaction wheel pendulum is modelled by the equations

See (Spong, Corke, and Lozano, 2001) for the model derivation

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -a\sin(x_1) - bu, \\ \dot{x}_3 &= a\sin(x_1) + cu, \end{aligned}$$

with $a = 78$, $b = 206$, $c = 31161$. Here x_1 is the angle w.r.t the upward equilibrium, x_2 the angular velocity of the pendulum, x_3 the angular velocity of the wheel. Prove that the model is feedback linearisable around 0, and determine in what region the diffeomorphism is well defined.

Attractors

Keywords: **ω-limit set, stability of sets, attractor, limit cycle, torus, chaotic attractor, sensitive dependence on initial conditions**

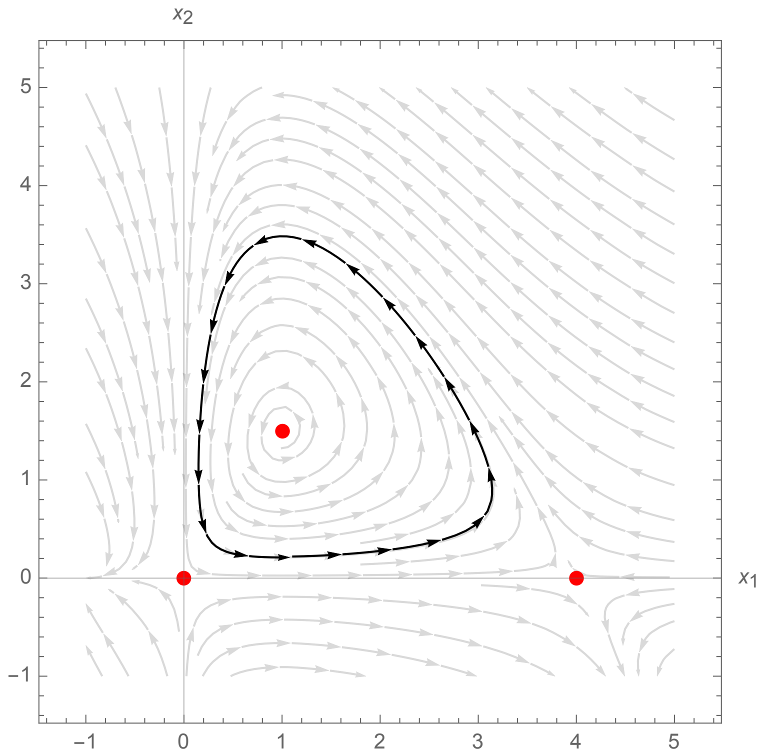
Definition of attractor

The previous chapter gave us the tools to classify systems according to their topology, that is, their qualitative behaviour. If the concept of topological equivalence is intuitively quite simple, the complexity of normal form reduction and analysis should make it apparent that there is very little hope of studying systems through homeomorphic or diffeomorphic transformations of their full phase portrait. A much more reasonable endeavour is to analyse a system near its asymptotic behaviour, that is, near the set of states where we expect the system to be most of the time. For a stable linear system, this would of course be the equilibrium. We can extend this idea, formally, by defining a more general set of asymptotic behaviours, which we call the *attractors*.

In this chapter, we bring discrete-time systems back to the game: all definitions, unless explicitly stated, apply both to continuous as well as to discrete-time systems.

Example 34. *The following figure shows the phase portrait of a Rosenzweig MacArthur prey-predator model, with equations*

$$\begin{aligned}
\dot{x}_1 &= x_1\left(1 - \frac{x_1}{4}\right) - \frac{x_1 x_2}{1 + x_1}, \\
\dot{x}_2 &= -x_2 + 2\frac{x_1 x_2}{1 + x_1}.
\end{aligned}$$



How does the system behave for $t \to \infty$, for initial conditions in the positive quadrant?

As we have seen for many other prey-predator models, the system has three equilibria: two saddles and a focus, which in this case is unstable. None of the equilibria is asymptotically stable so, even though they shape the geometry of the phase portrait, none of them describes the asymptotic behaviour of the system. This instead appears to settle onto a periodic orbit: a limit cycle.

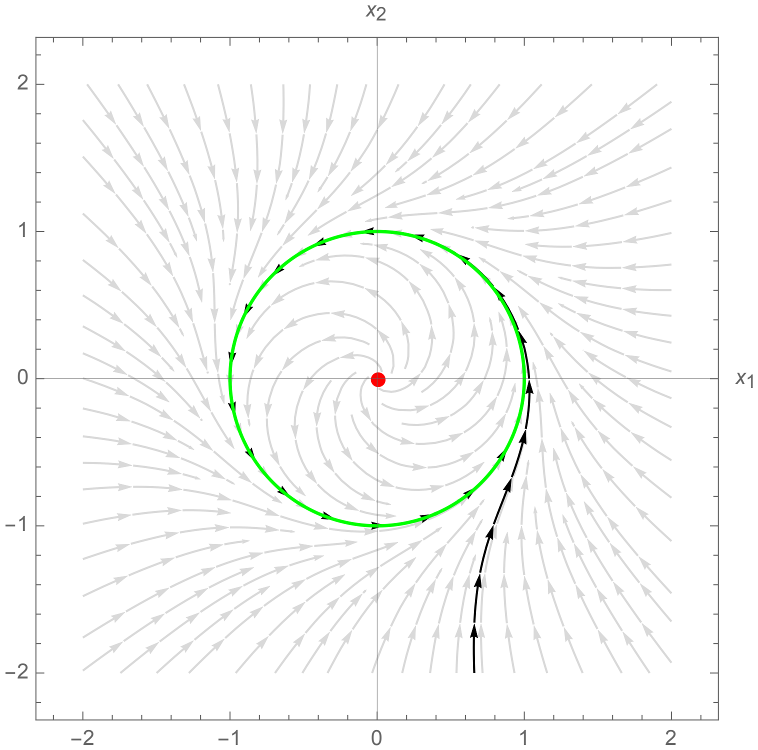
What we observe in the example is a typical feature of nonlinear systems: the asymptotic behaviour is often described by more complex sets than an equilibrium.

Definition: ω-limit point
A point $\bar{x}$ is an ω-limit point of the orbit $\phi_\cdot(x)$ if there exists a sequence $t_1, t_2, \ldots$, tending to $+\infty$, such that $\lim_{k\to\infty} \phi_{t_k}(x) \to \bar{x}$.

Definition: ω-limit set
The ω-limit set of x, denoted $\omega(x)$, is the set of all limit points of x.

Definition: α-limit set
The α-limit set of x, denoted $\alpha(x)$, is defined as the ω-limit set, but with $t \to -\infty$.

Example 35. *The red equilibrium in the following figure is an ω-limit set of itself and nothing else and is the α-limit set of all orbits within the* *green circle. The green circle is the α-limit set of itself, and the ω-limit set of $\mathbb{R}^2\backslash 0$, since the flow of any points in $\mathbb{R}^2\backslash 0$ converges to the green circle and visits a neighbourhood of any point of the circle infinitely often as $t \to \infty$.*



■ **Definition: Lyapunov stability of a set**
A positively invariant set S of a discrete or continuous-time system is Lyapunov stable if, for every neighbourhood N of S, there exists a neighbourhood $M \subset N$ such that

$$x \in M \Rightarrow \phi_t(x) \in N, \forall t \geq 0.$$

Now, to define asymptotic stability for sets, let us first define the distance between a point x and a set S as

$$d(x, S) := \inf_{y \in S} \|y - x\|.$$

■ **Definition: Asymptotic stability**
A positively invariant set S of a discrete or continuous-time system is asymptotically stable if it is Lyapunov stable, and there exists a neighbourhood N of S such that

$$\lim_{t \to \infty} d(\phi_t(x), Ss) = 0,$$

for all $x \in N$.

Definition: Attractor (Meiss, 2007)
An attractor is a set that is

1. compact,
2. asymptotically stable,
3. and the ω-limit set of some x.

There are multiple slightly different definitions of attractor, see (Milnor, 1985). Wiggins (2003) for example, requires *topological transitivity* instead of the third condition.

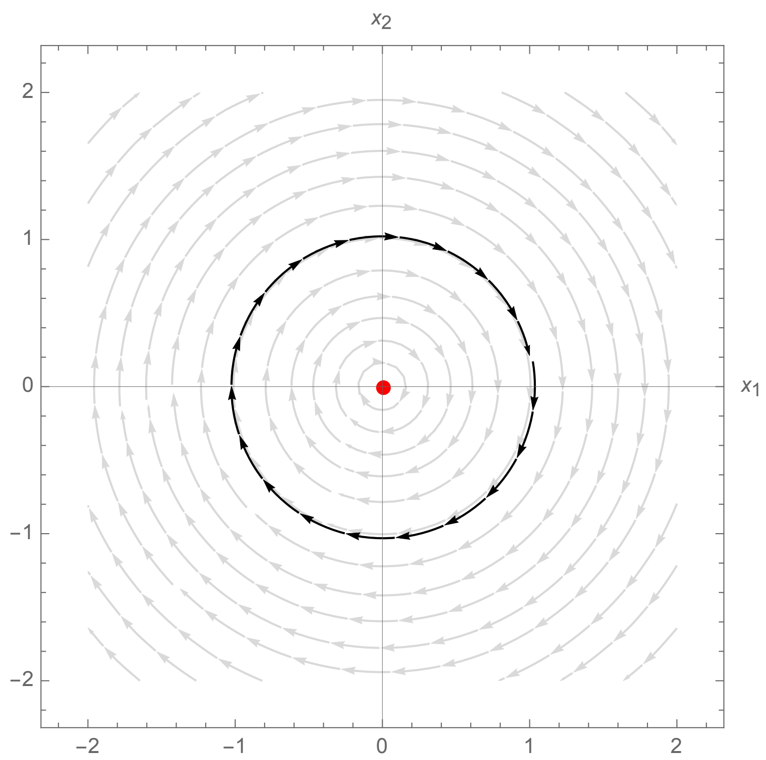
In the above definition, the third statement requires that *the whole* attractor be the ω-limit set of *the same* x. This allows us to distinguish an attractor from the union of separate attractors.

Note that conditions 2 and 3 in the definition of attractor are both necessary.

★ There exist Lyapunov-stable sets that are not attractors.

Example 36. *Consider the linear system*

$$\dot{x} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x.$$



A planar, linear equilibrium with imaginary eigenvalues is called a *centre*.

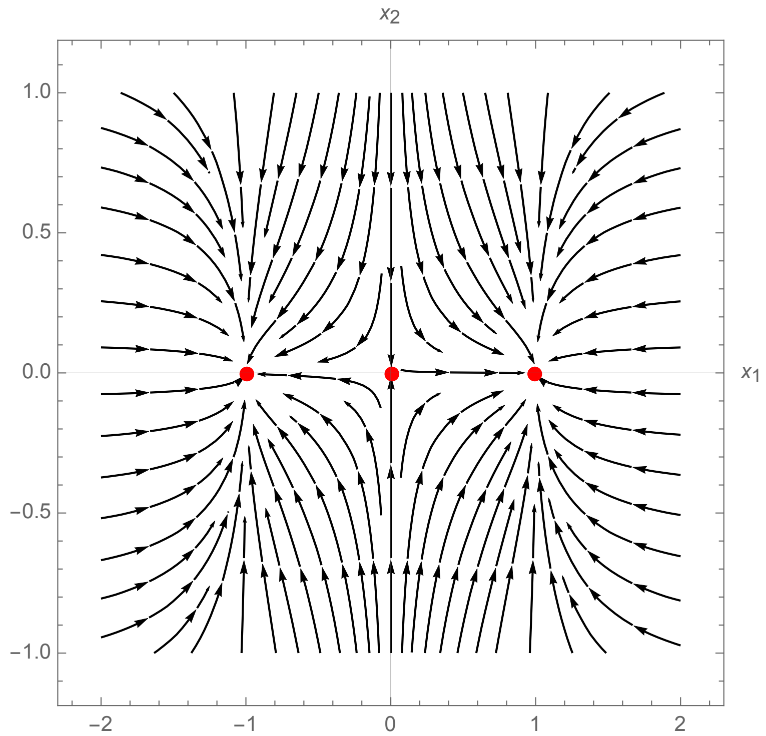
Its orbits are concentric circles around the origin. The origin is therefore Lyapunov stable (check!), but it is not asymptotically stable.

★ There exist asymptotically stable sets that are not attractors.

Example 37. *The system*

$$\dot{x}_1 = x_1 - x_1^3,$$
$$\dot{x}_2 = -x_2,$$

has the following phase portrait



It has a saddle in 0, *and stable nodes in* $(-1, 0)$ *and* $(1, 0)$. *The interval* $[-1, 1]$ *on the* x_1 *axis is a compact set and is asymptotically stable. However, it is not an attractor. It contains three distinct* ω*-limit sets, but it is not itself an* ω*-limit set of any state* x.

★ There exist ω-limit sets that are not attractors.

Example 38. *An unstable equilibrium is the* ω*-limit set of itself. A saddle is the* ω*-limit set of its stable manifold, but it is clearly not asymptotically stable.*

An attractor, by definition, must be asymptotically stable, therefore it must attract initial conditions in its neighbourhood. The union of these initial conditions forms its *basin of attraction*.

■ **Definition: Basin of attraction**

The basin of attraction of an attractor O is the set

$$\left\{ x : \lim_{t\to\infty} d(\phi_t(x), O) = 0 \right\}.$$

Example 39. *Consider again the system in Example 37. The basins of attraction of the two nodes are the positive and negative half-plane, respectively. They are separated by the stable manifold of the saddle in* 0*, which, in this case, coincides with the* x_2 *axis. In general, basins of attractions of different attractors are separated by invariant surfaces which are either sets of orbits going to infinity, or stable manifolds of saddle sets (saddle equilibria, cycles, etc.) For this reason, these manifolds are also called separatrices.*

Consequently, the analysis of saddle sets can be as important as the analysis of attractors, to determine the global behaviour of a nonlinear system.

Types of attractors

According to the definition given before, a stable node or focus is an attractor. There are, of course, more complex types of attractors. Here we see some of the most common ones.

■ **Definition: Limit cycle**
A limit cycle is an isolated periodic orbit.

The above definition identifies a limit cycle as a periodic orbit that has no other periodic orbit in a neighbourhood. This means that limit cycles are not necessarily attractors, and indeed non-attracting limit cycles play an important role in organizing the dynamics of many nonlinear systems. However, we will most often deal with limit cycles that are also attractors, which we simply call *stable limit cycles*.

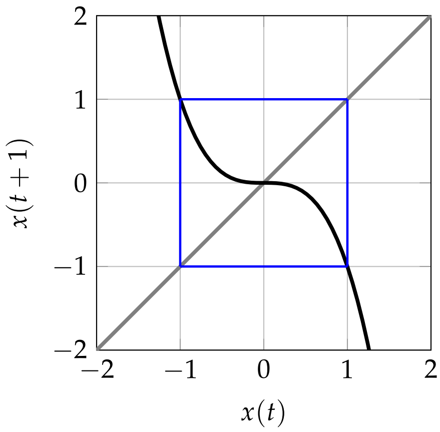
■ **Definition: Stable limit cycle**
A stable limit cycle is a limit cycle that is asymptotically stable and is therefore an attractor.

The examples of limit cycles that we have seen before were mostly from continuous-time systems. In this case, the limit cycle is a closed curve. A limit cycle of a discrete-time system has a markedly different appearance: it is a closed but non-connected set.

Example 40. *Consider the discrete-time systems*

$$x(t+1) = -x^3(t).$$

It has a fixed point in $x = 0$*, but we can also easily see that* $f(1) = -1$ *and* $f(-1) = 1$*. The two points* -1 *and* 1 *form a periodic orbit, and with a little more effort we could prove that it is isolated: it is a limit cycle.*



Note that the blue line is not an orbit! It is a graphical representation of the iteration between the two states of the limit cycle: $x = -1$ and $x = 1$.

Besides limit cycles, nonlinear systems can exhibit many other forms of asymptotic dynamics that are not possible in linear systems, and yet are very relevant in applications. A major one, that has had a deep impact on our understanding of many physical phenomena, is chaotic dynamics.

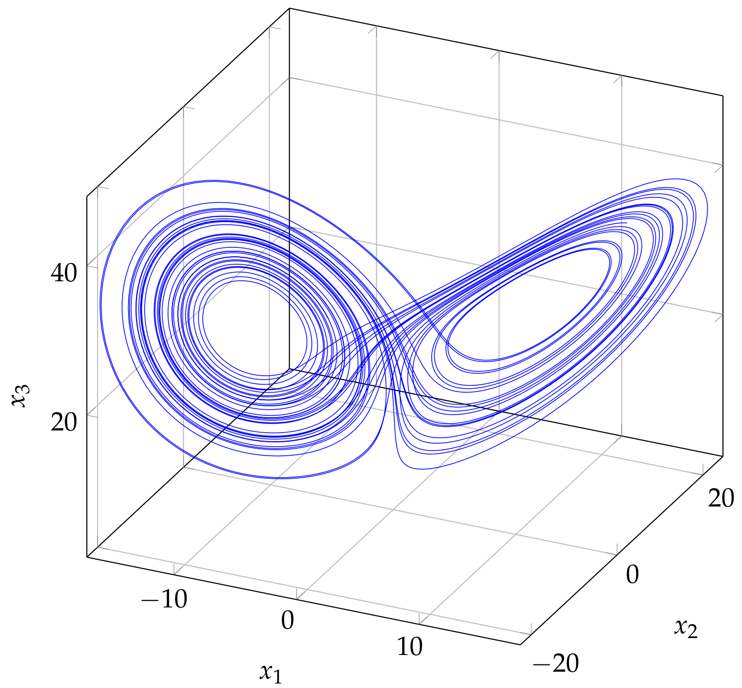
■ **Definition: Sensitive dependence on initial conditions**
A positively invariant set N exhibits sensitive dependence on initial conditions if there exists a constant $r > 0$ such that, for all $\epsilon > 0$ and x in N, there exist $y \in N$ with $\|x - y\| < \epsilon$ such that $\|\phi_t(x) - \phi_t(y)\| > r$ for some $t \geq 0$.

■ **Definition: Chaotic attractor**
An attractor exhibiting sensitive dependence on initial conditions is a chaotic attractor

Example 41 (Lorenz attractor)**.** *Orbits of the Lorenz system*

$$\begin{aligned}
\dot{x}_1 &= 10(x_2 - x_1), \\
\dot{x}_2 &= x_1(28 - x_3) - x_2 \\
\dot{x}_3 &= x_1 x_2 - \frac{8}{3}x_3,,
\end{aligned}$$

which models the amplitudes of the fundamental modes of oscillation of a 2D fluid with a temperature gradient, converge to a chaotic attractor. The attractor can be shown to be a fractal set.

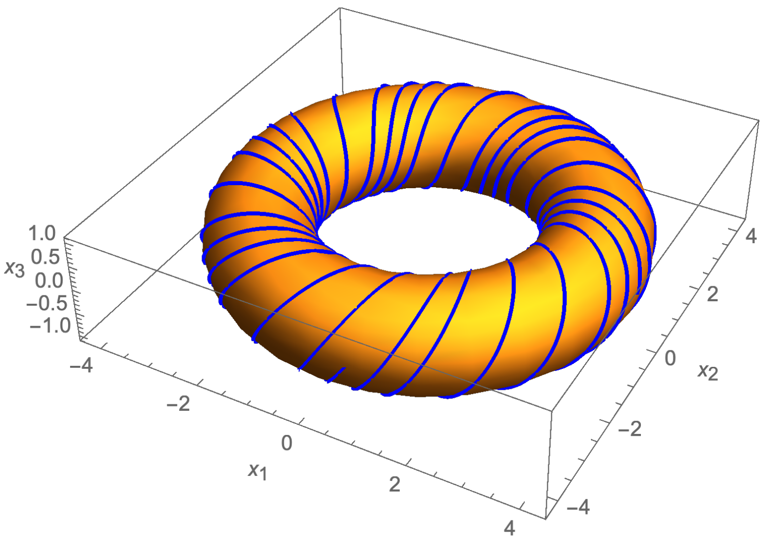


In between the relative dullness of periodic dynamics, and the complexity of chaotic dynamics, lies the set of those behaviours obtained from the composition of multiple periodic functions, with frequencies arranged in a way that makes the overall dynamics nonperiodic. This happens when two or more periodic components combine, with frequencies that are not in rational relation (e.g., frequencies θ_1 and θ_2 such that $k_1\theta_1 + k_2\theta_2 \neq 0$ for any nonzero integer k_1 and k_2.)

■ **Definition: Quasiperiodic attractor**

A quasiperiodic attractor is an attractor where the states $x(t)$ oscillate according to a quasiperiodic function $F(t\theta)$, where $\theta := (\theta_1, \ldots, \theta_d)$ and $k\theta \neq 0, \forall k \in \mathbb{Z}^d \backslash 0$.

Example 42. *A quasiperiodic orbit with* 2 *frequencies in non-rational relation can typically be represented, through a suitable change of variables, as an infinitely long curve wrapped around a torus, as in the following figure (where the two angular frequencies are* 1 *and* 2π*).*

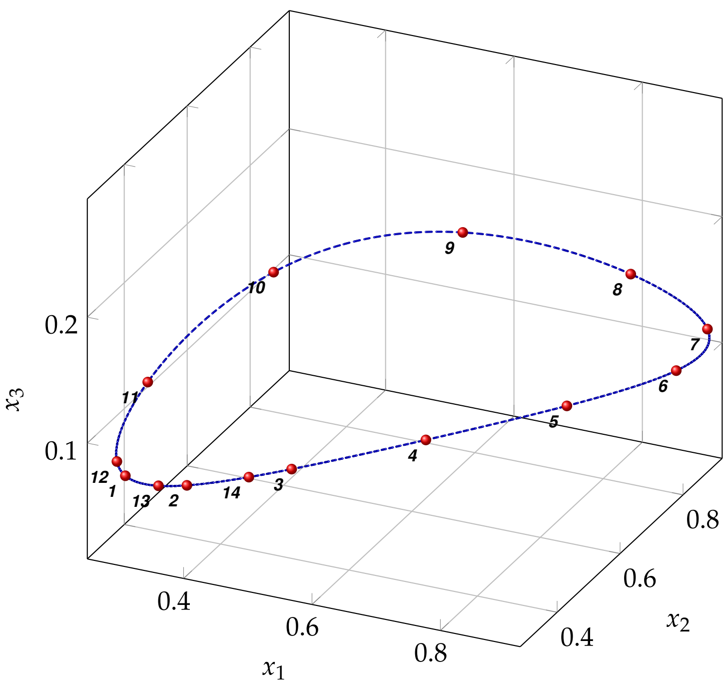


Note that the blue line represents only a small portion of the orbit, which would densely cover the whole torus surface. You can see the interrupted orbit in the lower left end of the figure.

Example 43 (A discrete-time quasiperiodic attractor)**.** *The discrete-time system*

$$x_1(t+1) = x_2(t),$$
$$x_2(t+1) = (1+r)x_2(t) - rx_1(t)x_2(t) - cx_2(t)x_3(t),$$
$$x_3(t+1) = cx_2(t)x_3(t),$$

which is an embedding in 3 *dimensions of a discrete-time prey-predator model from (Kot, 2005), has, for* $r = 0.58$ *and* $c = 1.85$*, the following quasiperiodic attractor.*

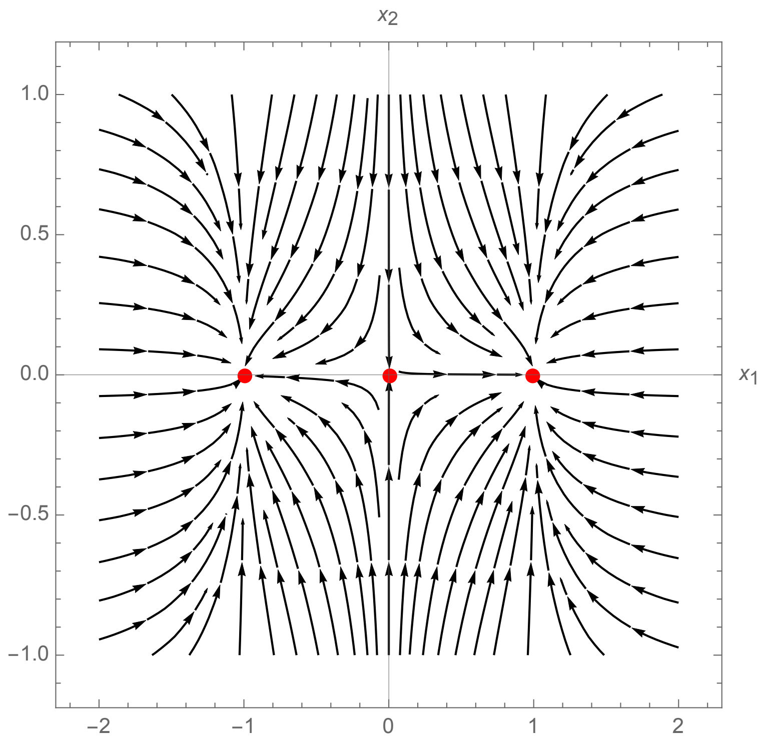


The apparently continuous curve consists of a multitude of nearby iterations of the map arranged along the torus; to give an idea of how states within the *attractor are mapped, 14 subsequent iterations of the map are plotted in red and numbered sequentially.*

Exercises

Exercise 36

Identify the stable and unstable manifolds of the three equilibria in the following phase portrait and the basin of attraction of the stable equilibria.



Exercise 37

Propose a nonlinear dynamical system with a periodic orbit that is not a limit cycle.

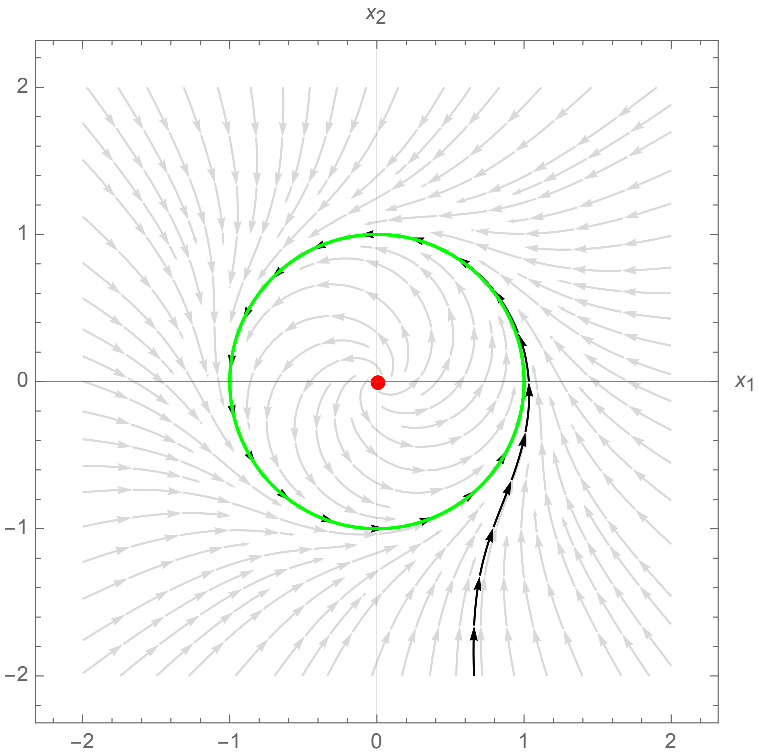
Exercise 38

The system

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

which we will eventually meet as Hopf normal form, has a stable limit cycle (hence, an attractor) of the equation

$$x_1^2 + x_2^2 = 1 :$$



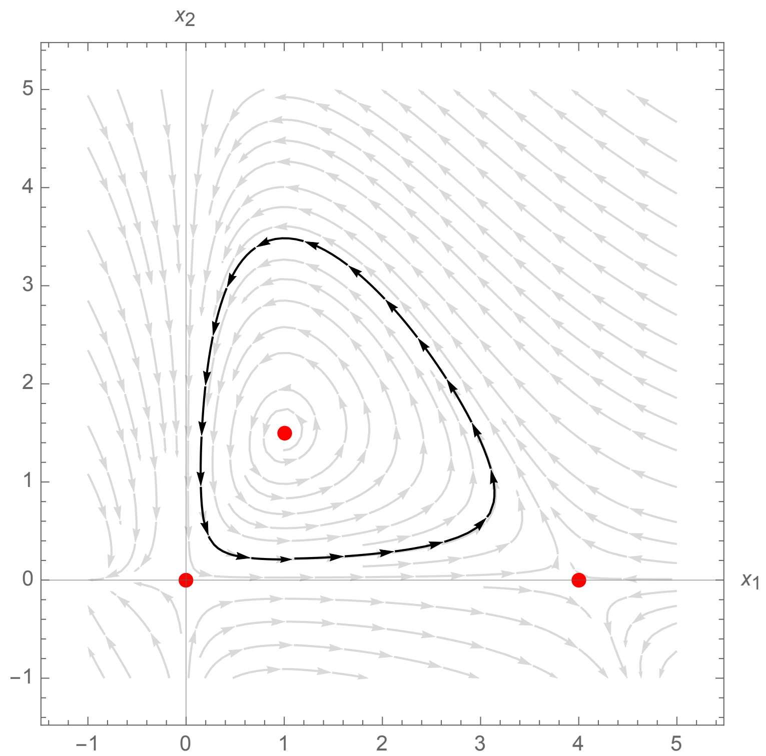
If we modify it by a time rescaling, as follows:

$$\dot{x}_1 = \left(x_1 - x_2 - x_1(x_1^2 + x_2^2)\right)(1 - (x_1^2 + x_2^2))^2,$$
$$\dot{x}_2 = \left(x_1 + x_2 - x_2(x_1^2 + x_2^2)\right)(1 - (x_1^2 + x_2^2))^2,$$

is the circle $x_1^2 + x_2^2 = 1$ still an attractor?

Exercise 39

What is the basin of attraction of the stable limit cycle in this Rosenzweig-MacArthur model?

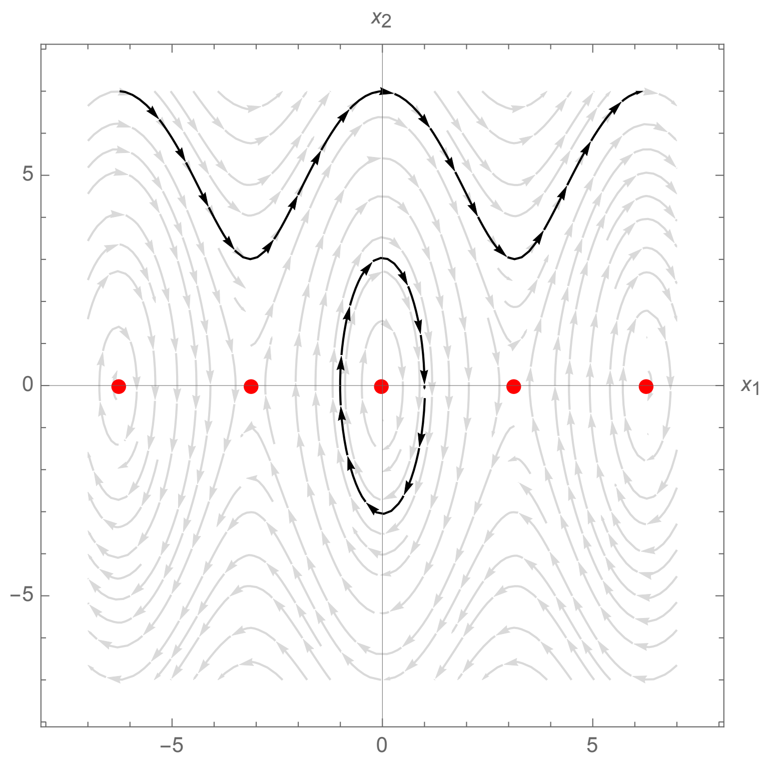


Exercise 40

The frictionless pendulum of equations

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -10\sin(x_1),$$

with $x \in \mathbb{R}^2$, has phase portrait



Discuss the existence of periodic orbits and limit cycles in the model.

Limit cycles

Keywords: **Poincaré-Bendixon theorem, Dulac's criterion, Poncaré map, stable unstable and centre manifold of a cycle, monodromy matrix, Floquet multipliers, multipliers stability criterion.**

In this chapter, we focus on periodic orbits and limit cycles of contiuous-time systems, and in paticular we learn about some of the tools that can be used to prove the existence or nonexistence of limit cycles, or to study their stability and characterize the neighbouring phase portrait. A good part of the chapter is devoted to planar systems, where more powerful tools are available.

Example 44. *Let us consider a spring-mass model with a nonlinear friction characteristic:*

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1 - x_2(x_1^2 - 1).$$

This is known as the Van der Pol oscillator and was originally proposed as a model for a vacuum tube circuit.

The system has a single equilibrium in the origin, and

$$J_f(0) = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix}$$

has eigenvalues $\frac{1}{2} \pm i\frac{\sqrt{3}}{2}$. *It is therefore an unstable focus. From physical reasoning, however, we can see that for large displacements of the spring (large* $|x_1|$*) the system has positive friction coefficient* $(x_1^2 - 1)$ *and should therefore behave more or less like a regular, dissipative damped spring. We can thus expect the system to settle on a bounded asymptotic behaviour, which must however not be an equilibrium. Can we prove that it is a limit cycle, and how?*

In general, proving or disproving the existence of limit cycles is a tough problem. It is somewhat simpler in planar continuous-time systems, where a set of rather powerful tools can be constructed based on the fact that periodic orbits divide the plane in two disjoint sets.

For similar results in higher-dimensional systems, we will have to use ingeniously some results from bifurcation theory and normal form reduction, which we will learn later on.

Existence theorems for continuous-time systems in $\mathbb{R}^2$

The first (non)-existence theorem is a simple consequence of the Poincaré index of periodic orbits, which we already met a few chapters ago.

Theorem
A limit cycle in the plane must enclose a set of equilibria of total index 1.

Therefore, for example, regions of the plane without equilibria cannot have limit cycles. The following theorem adds some more structure to this basic observation.

Poincaré Bendixon theorem
Let $f(x)$ be a continuously differentiable continuous-time vector field in $\mathbb{R}^2$, and M be a positively invariant compact subset of $\mathbb{R}^2$ containing a finite number of equilibria of $f(x)$. For all $x \in M$, one of the three following statements holds:

1. $\omega(x)$ is an equilibrium,
2. $\omega(x)$ is a periodic orbit,
3. $\omega(x)$ consists of a finite number of equilibria $p_1, \ldots, p_n$ and orbits γ with $\alpha(\gamma) = p_i$ and $\omega(\gamma) = p_j$.

For a proof see e.g. (Wiggins, 2003)

This theorem is important for what it proves to exist, as well as for what it proves cannot exist:

Corollary
A continuous-time system in $\mathbb{R}^2$ cannot have quasiperiodic or chaotic attractors.

Here the fact that we are considering continuous-time systems is essential: even 1D discrete-time systems can have chaotic attractors, see the tent map. We will later see that, even for continuous-time systems, continuous differentiability is essential. Discontinuous or hybrid systems can behave like discrete-time ones in this respect.

Corollary
If M does not contain stable equilibria or saddles, then every $x \in M$ converges to a periodic orbit.

Example 45. *The following equations, from (Strogatz, 1994), describe a kinetic model of glycolysis in yeast cells*

$$\begin{aligned} \dot{x}_1 &= -x_1 + a x_2 + x_1^2 x_2, \\ \dot{x}_2 &= b - a x_2 - x_1^2 x_2, \end{aligned}$$

with $a, b > 0$.

The nullclines are

$$\dot{x}_1 = 0 \Rightarrow x_2 = \frac{x_1}{a + x_1^2}$$

and

$$\dot{x}_2 = 0 \Rightarrow x_2 = \frac{b}{a + x_1^2}.$$

and intersect at $x_1 = b$, $x_2 = \frac{b}{a+b^2}$.

Consider now the subset M (green in the figure below) of the positive quadrant enclosed by the lines

$$\begin{aligned} x_1 &= 0, \\ x_2 &= 0, \\ x_2 &= \frac{b}{a}, \\ x_1 + x_2 &= b + \frac{b}{a}. \end{aligned}$$

We can see that M is positively invariant:

$$\begin{aligned} \frac{\partial}{\partial x}(-x_1)f(x)\Big|_{x_1=0, x_2\in[0,b/a]} &= -\, ax_2 < 0, \\ \frac{\partial}{\partial x}(-x_2)f(x)\Big|_{x_1\in[0,b+b/a], x_2=0} &= -\, b < 0, \\ \frac{\partial}{\partial x}(x_2 - b/a)f(x)\Big|_{x_1\in[0,b], x_2=b/a} &= -\, \frac{b}{a}x_1^2 \leq 0, \\ \frac{\partial}{\partial x}(x_1 + x_2 - b - b/a)f(x)\Big|_{x_1\in[b,b+b/a], x_2\in[0,b/a]} &= b - x_1 < 0. \end{aligned}$$

By the Poincaré Bendixon theorem, all initial conditions in M converge to the equilibrium, to a homoclinic connection of the stable and unstable manifold of the equilibrium, or to a periodic orbit.

Now, note that

$$J_f(b, b/(a+b^2)) = \begin{pmatrix} \frac{-a+b^2}{a+b^2} & a+b^2 \\ \frac{-2b^2}{a+b^2} & -a-b^2 \end{pmatrix}.$$

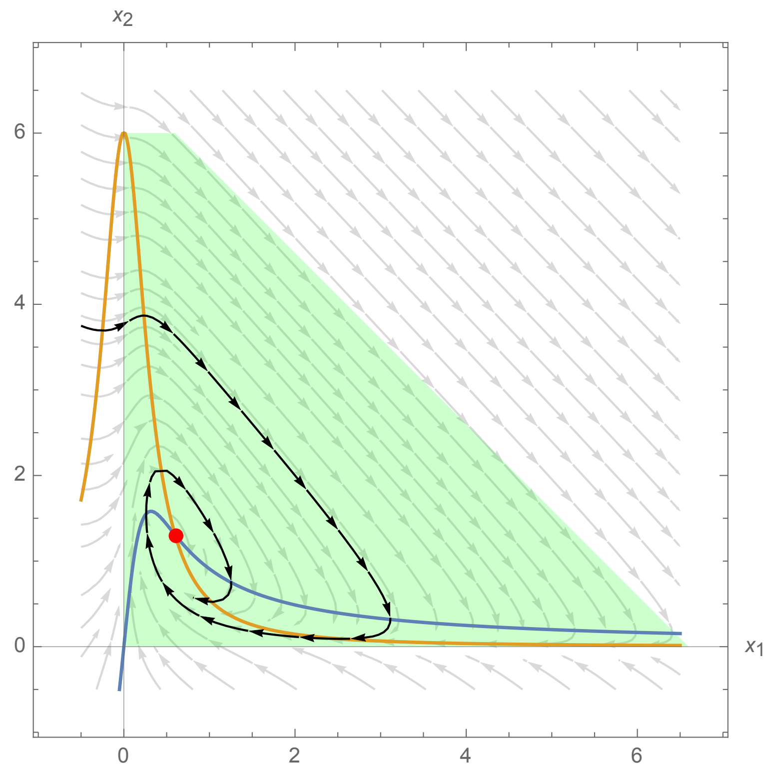
We have

$$\det(J_f) = a + b^2 > 0,$$

so the equilibrium is not a saddle, and we can rule out convergence to a homoclinic connection. Then, we have

$$\mathrm{tr}(J_f) = \frac{-a + b^2 - (a+b^2)^2}{a+b^2}.$$

For values where $\mathrm{tr}(J_f) > 0$ *the equilibrium is unstable. In this case, M must contain a periodic orbit.*



While the Poincaré-Bendixon theorem is typically used to prove the existence of limit cycles, the following theorem is used to exclude the existence of cycles in a given region.

Dulac's criterion
Let $f(x)$ be a continuously differentiable continuous-time vector field defined in a simply connected subset N of $\mathbb{R}^2$. If there exists a continuously differentiable function $g(x) : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\nabla \cdot (g(x)f(x)) := \frac{\partial}{\partial x_1} g(x) f_1(x) + \frac{\partial}{\partial x_2} g(x) f_2(x)$$

is everywhere nonzero and has a constant sign, then there are no periodic orbits lying entirely in N.

A subset of $\mathbb{R}^2$ is simply connected if any closed curve in it can be shrunk to a point without leaving the set.

Here $\nabla\cdot$ is the divergence.

Example 46. *Consider*

$$\dot{x}_1 = -1 - x_1 + x_2^2,$$
$$\dot{x}_2 = x_2(1 + x_1^2 - x_2^2).$$

Note that the x_1 axis is an invariant set, therefore periodic orbits can only exist if entirely contained in the regions above or below the x_1 axis.

Now, consider $g(x) = \frac{1}{x_2}$. We have

$$\nabla \cdot g(x)f(x) = \frac{\partial}{\partial x_1}\left(-\frac{1}{x_2} - \frac{x_1}{x_2} + x_2\right) + \frac{\partial}{\partial x_2}\left(1 + 2x_1^2 - x_2^2\right)$$
$$= -\frac{1}{x_2} - 2x_2.$$

This has a constant sign for $x_2 > 0$ or $x_2 < 0$, therefore the system cannot have limit cycles.

The next approach to proving existence of periodic orbits applies only to continuous-time dynamical systems with a very special structure: conservative systems.

■ **Definition**
A continuous-time system with vector field $f(x)$ is conservative if there exists a scalar function $V(x) : \mathbb{R}^n \to \mathbb{R}$ such that $\dot{V}(x) = 0$, for all $x \in \mathbb{R}^n$, and such that $V(x)$ is nonconstant on every open set.

This means that the function $V(x)$ is constant along any orbit of the system. While the above property may seem excessively constraining at first sight, it is in fact relatively common. Consider, for example, the law of motion of a 1-degree of freedom object subject to Newton's law, without friction:

$$m\ddot{x} = F(x).$$

Let us define the potential energy $E(x)$ as a function such that $\frac{\partial E(x)}{\partial x} = -F(x)$. Using the above equations, and multiplying by $\dot{x}$, we obtain

$$m\dot{x}\ddot{x} + \frac{\partial E(x)}{\partial x}\dot{x} = 0,$$

which can be written as

$$\frac{d}{dt}\left[\frac{m}{2}\dot{x}^2 + E(x)\right]$$

Thus, the function $V(x) = \frac{m}{2}\dot{x}^2 + E(x)$ (the total energy) is constant along orbits. We can use this fact to prove existence of periodic orbits near an equilibrium, as follows.

⧫ **Theorem on periodic orbits of conservative systems.**
Consider a planar, continuously differentiable, and conservative vector field $f(x)$, with conserved quantity $V(x)$. If an isolated equilibrium $\bar{x}$ is a local maximum or minimum of $V(x)$, then all the orbits in a sufficiently small neighbourhood of $\bar{x}$ are periodic.

See Strogatz (1994), Theorem 6.5.1

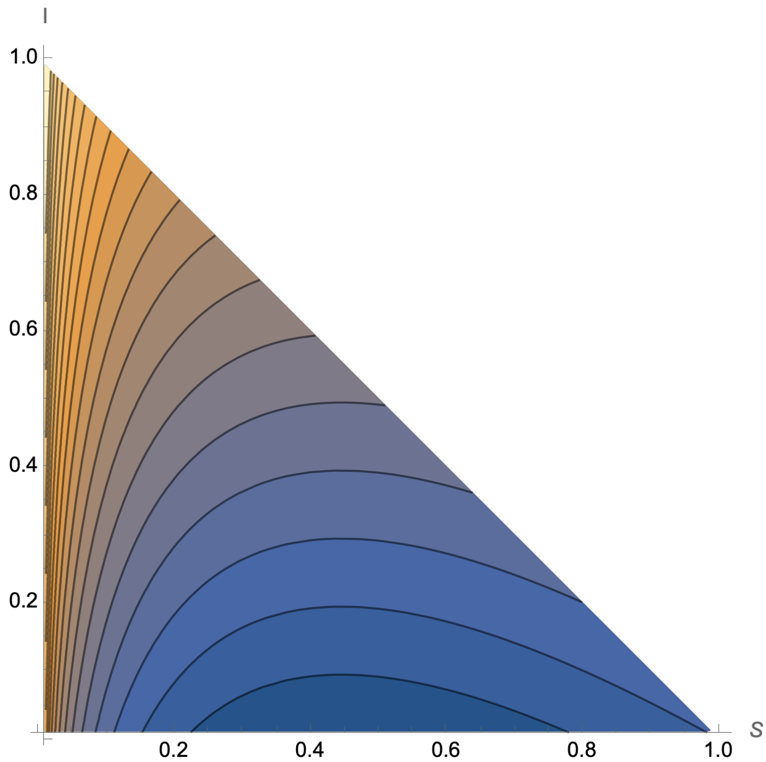
★ A consequence of the above theorem is that the periodic orbits are not limit cycles, since they are not isolated.

Notice that the existence of a conserved quantity alone is not sufficient to prove existence of periodic orbits, since the level curves of $V(x)$ are not necessarily closed. Notice also that, while it is always true that any orbit of a planar conservative system belongs to a unique level cure, the converse is not true: level curves may include multiple orbits. As an example of this, consider the S and I variables of the SIR model:

$$\dot{S} = -aIS,$$
$$\dot{I}_2 = aIS - bI.$$

The quantity $V(x) = -\frac{b}{a}\log(S) + S + I$ is a conserved quantity, hence the above system is conservative. Its level curves are plotted in the figure below. We see that each level curve includes both an orbit corresponding to an epidemic wave, and up to two equilibria, which lie at the intersection of the level curve with the axis $I = 0$.

For conserved quantities in more general SIR models see, e.g., (Mestres and Cortes, 2022)



Finally, the last existence theorem that we see is very specific, in terms of the assumptions, but applies to a family of systems that is relatively common in electrical and mechanical engineering. It was indeed mainly used in the first half of the 20th century, to prove oscillations in certain models of nonlinear electrical circuits.

♦ **Lienard theorem**

Consider a system of the form

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -g(x_1) - f(x_1)x_2.$$

If

1. $f(x_1)$ and $g(x_1)$ are continuously differentiable for all x_1,
2. $g(-x_1) = -g(x_1)$ for all x_1 (g is an odd function),
3. $g(x_1) > 0$ for $x_1 > 0$,
4. $f(-x_1) = f(x_1)$ for all x_1 (f is an even function),
5. $F(x_1) := \int_0^{x_1} f(u)du$ has exactly one positive root a, is negative for $0 < x_1 < a$ positive nondecreasing for $x_1 > a$, and tends to ∞ as $x \to \infty$,

then the system has a unique stable limit cycle surrounding the origin.

Example 47. *Consider again the Van der Pol oscillator*

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1 - x_2(x_1^2 - 1).$$

We have

$$g(x) = x_1$$

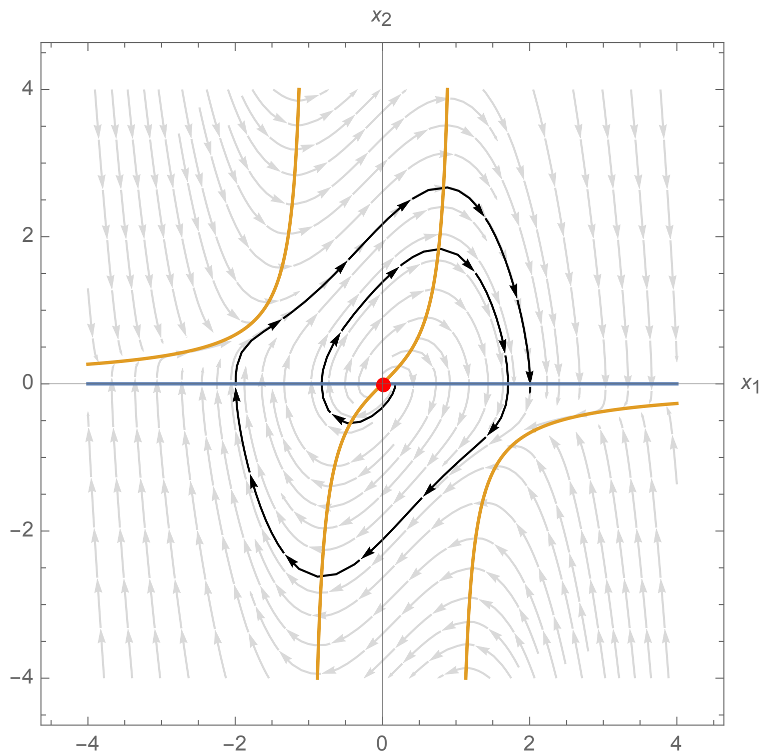
and

$$f(x) = x_1^2 - 1.$$

The function g is odd and positive for positive x_1, while f is even and

$$\int_0^{x_1} f(u)du = \frac{1}{3}x_1^3 - x_1,$$

which has exactly one positive root $x_1 = \sqrt{3}$, and is negative between 0 and $\sqrt{3}$. The system is therefore guaranteed to have a unique stable limit cycle. The phase portrait is given below.



Poincaré maps

Let us now address the problem of determining the stability of a limit cycle. If we are dealing with a discrete-time system, then once we have proved that a given state x is a fixed point of the m-times iterated map $f^{(m)} = f \circ f \circ \ldots f(x)$ while it is not a fixed point for any map iterated less than m times (i.e, it is a periodic orbit of period m) its stability is easily discussed by studying the Jacobian of $f^{(m)}$ at x. All that we know about invariant manifolds translates as well to this map. Things might seem less straightforward for a limit cycle in a continuous-time system: how can we determine the stability of the cycle? Does it have stable and unstable manifolds, like equilibria do? We can answer these questions quite easily by reducing the analysis to that of a discrete-time system.

■ **Definition: Poincaré section**
Given a continuous-time dynamical system of order n and a point $\bar{x} \in \mathbb{R}^n$, a Poincaré section in a neighbourhood of $\bar{x}$ is a surface S of dimension $n - 1$ such that $f(x)$ is not tangent to S at any point in the neighbourhood.

■ **Definition: Poincaré map**
The Poincaré map on a section S is the map $S \to S$ obtained by taking any point $x \in S$ and following $\phi_t(x)$ until its first return on S

Of course, the Poincaré map may in general not exist, for an arbitrary section S. It is however defined in a neighbourhood of any periodic orbit, and this is the context where we use it most often.

Example 48. *The following system, represented in polar coordinates and defined in $\mathbb{R}^2\backslash 0$, has a limit cycle of radius 1 around the origin:*

$$\dot{\rho} = 1 - \rho,$$
$$\dot{\theta} = 2\pi.$$

Let us take as Poincaré section a segment of constant θ, $\rho \in (0, \infty]$. We can compute explicitly the Poincaré map by solving the equation of ρ for $t \in [0, 1]$. We obtain

$$\rho(1) = e^{-1}\rho(0) + e^{-1}\int_0^1 e^{\tau} d\tau = e^{-1}(\rho(0) - 1) + 1.$$

We see that 1 is a fixed point of the Poincaré map. We should have expected this since we have a periodic circular orbit of radius 1. We also see that the

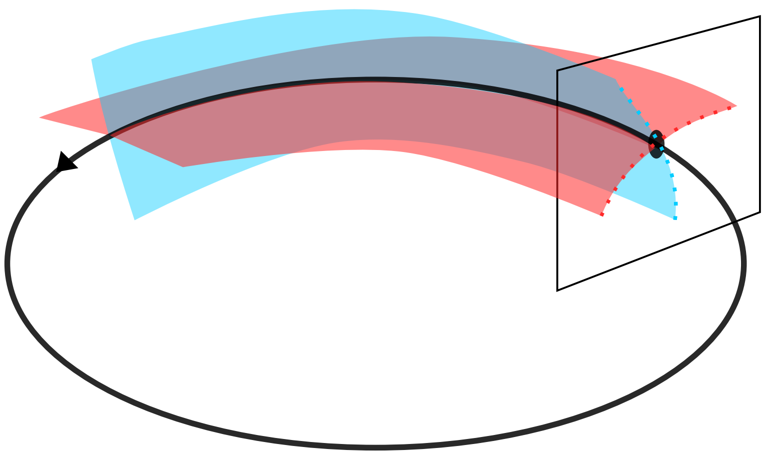
$$J_f(1) = e^{-1},$$

the eigenvalue of the Jacobian is within the unit circle. The fixed point of the map is therefore asymptotically stable.

A continuous-time periodic orbit is therefore a fixed point of some Poincaré map. We have also seen that fixed points have stable, unstable, and centre manifolds, by the nonhyperbolic Hartman-Großman theorem. We can conclude that any continuous-time periodic orbit, just like any equilibrium, has a stable, an unstable, and a centre manifold.

Definition: Stable, unstable, and centre manifold of cycles
The stable, unstable, and centre manifold of a cycle are manifolds formed by orbits that intersect the stable, unstable, and centre manifold of the corresponding fixed point on the Poincaré map.

Example 49. *A saddle cycle in 3 dimensions, that is, a cycle whose Poincaré map is a saddle fixed point, has a 2-dimensional stable and a 2-dimensional unstable manifold.*



Stability and the monodromy matrix

■ **Definition: Fundamental matrix**
The fundamental matrix of a time-dependent linear system

$$\dot{x} = A(t)x$$

is the solution of the matrix differential equation

$$\dot{X} = A(t)X, \quad X \in \mathbb{R}^{n \times n}, \quad X(0) = I.$$

Using the superposition principle (which holds for time-dependent linear systems), one can see how the fundamental matrix $X(t)$ provides a means to determine $x(t)$ as

$$x(t) = X(t)x(0).$$

In the context of limit cycle analysis, we can use the above idea to study the behaviour of the nonlinear vector field in the vicinity of the periodic orbit $\phi_{\cdot}(x)$, by studying the dynamics of the linearised vector field

$$\dot{\xi} = J_f(\phi_t(x))\xi$$

over one period T of the orbit.

■ **Definition: Monodromy matrix**
Given a periodic orbit $\phi_{\cdot}(x)$ of period T of a continuously differentiable vector field $f(x)$, its monodromy matrix is the fundamental matrix $X(T)$ of

$$\dot{\xi} = J_f(\phi_t(x))\xi$$

at time T.

Notice that the eigenvalues of $X(T)$ are in general not related to those of J_f, so, for example, one can be unstable while the other is uniformly stable.

■ **Definition: Floquet multipliers**
The eigenvalues of the monodromy matrix are called Floquet multipliers of the periodic orbit.

Periodicity of the orbit puts a constraint on the multipliers and eigenvectors of the matrix: perturbations in the direction of the flow must be mapped onto themselves after one period. This implies that there always exists one direction, parallel to the flow, along which the matrix has unitary eigenvalue.

♦ **Trivial multiplier**
The monodromy matrix always has a multiplier equal to 1, called the trivial multiplier, and corresponding to an eigenvector parallel to $f(x(0)) = f(x(T))$.

♦ **Theorem**
The n multipliers of the monodromy matrix are the $n-1$ eigenvalues of the Poincaré map linearised at the fixed point, plus the trivial multiplier 1.

Sketch of proof. The linearised Poincaré map is obtained as the composition of the monodromy matrix, and a projection of the resulting vectors back onto the Poincaré section along the flow. The flow is, however, parallel to the leading order to the direction of the eigenvector corresponding to the trivial multiplier. The composition of the two linear transformations therefore has a zero eigenvalue, which takes the place of the trivial multiplier, plus all the eigenvalues of the original monodromy matrix. □

Given the monodromy matrix $X(t)$, an infinitesimal perturbation ξ from the periodic orbit on the Poincaré section is mapped back onto the section by the mapping $PX\xi$, with

$$P := I - \frac{f(0)n^\top}{n^\top f(0)},$$

where $f(0)$ is the vector field where the cycle intersects the Poincaré section, and n is a normal vector to the section. It is easy to see how $f(0)$ is an eigenvector of P with eigenvalue 0.

We can now translate, with minimal changes, all that we know about the stability of fixed points in discrete-time systems to the study of the stability of periodic orbits.

♦ **Multipliers criterion for cycles**
If all the Floquet multipliers of a periodic orbit except the trivial one are strictly inside the unit circle, then the orbit is an asymptotically stable limit cycle. If at least one multiplier is outside of the unit circle, then the orbit is an unstable set.

This is, of course, equivalent to the eigenvalue criterion for discrete-time systems. The fact that the periodic orbit is a *limit* cycle follows from the Hartman-Großman theorem for discrete-time hyperbolic equilibria.

The Poincaré map is therefore a fundamental tool to study the stability of limit cycles. The Jacobian of this map around the fixed point corresponding to the periodic orbit has a close relative in a matrix, known as the fundamental matrix, used in the analysis of the stability of a time-dependent linear system.

Exercises

Exercise 41

A continuous time system has a periodic orbit with a monodromy matrix

$$M = \begin{pmatrix} -1 & 2 & 0 \\ 2 & -1 & 0 \\ -2 & -1 & -2 \end{pmatrix}.$$

1. Is this a limit cycle, and what are its Floquet multipliers?
2. Is the orbit an attractor?
3. Consider now the system obtained by reversing time. Is the periodic orbit an attractor in this system?

Exercise 42

Which of these can be monodromy matrices of a continuous-time, continuously differentiable vector field around a periodic orbit?

$$\begin{pmatrix} 1 & 0 \\ -3 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 3 & -2 \\ 0 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}.$$

[Hint, by Picard-Lindelöf the orbits of the systems are unique in the neighbourhood of any state, hence they cannot merge.]

**** Exercise 43**

Can the following matrix be the Jacobian of the Poincaré map of a limit cycle in 3 dimensions?

$$\begin{pmatrix} -0.5 & 0 \\ 0 & 2 \end{pmatrix}$$

Hint: think about the geometry of the stable and unstable manifolds.

**** Exercise 44**

Consider a fixed point of a Poincaré map of a 3-dimensional system, with a stable and an unstable manifold. Assume that the dynamics in the stable manifold has an equation

$$x(t+1) = -0.5x(t),$$

while the dynamics in the unstable manifold is given by

$$x(t+1) = \frac{-3\epsilon x}{x^2 + 2\epsilon}$$

for some $\epsilon << 1$.

1. Identify the fixed points of the Poincaré map and of the second-iterated Poincaré map in the unstable manifold.
2. In light of the above analysis, sketch the 3D phase portrait near the limit cycle.

Exercise 45

Prove that the system

$$\begin{aligned} \dot{x}_1 &= -x_1^3 + 2x_1x_2 + x_1, \\ \dot{x}_2 &= -x_2^3 - x_2^2 - 2x_2, \end{aligned}$$

cannot have periodic orbits.

Exercise 46

Prove that the SIR model

$$\begin{aligned}\dot{S} &= -aIS,\\ \dot{I} &= aIS - bI,\\ \dot{R} &= bI,\end{aligned}$$

cannot have periodic orbits.

Exercise 47

Prove that the system

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= -x_1 + x_2(1 - x_1^2 - x_2^2),\end{aligned}$$

has a periodic orbit.

Exercise 48

Prove that the following model has a limit cycle:

$$\begin{aligned}\dot{x}_1 &= x_2,\\ \dot{x}_2 &= -x_1^3 - x_2(3x_1^2 - \cos(x_1)).\end{aligned}$$

** **Exercise 49**

Prove that the Lotka Volterra model

$$\begin{aligned}\dot{x}_1 &= x_1(a - bx_2),\\ \dot{x}_2 &= x_2(cx_1 - d).\end{aligned}$$

has infinitely many periodic orbits around the equilibrium $\bar{x} = (d/c, a/b)$, and that these are not limit cycles.

Chaotic attractors

Keywords: **Lyapunov exponents, properties of the Lyapunov exponents, fractal set, strange attractor.**

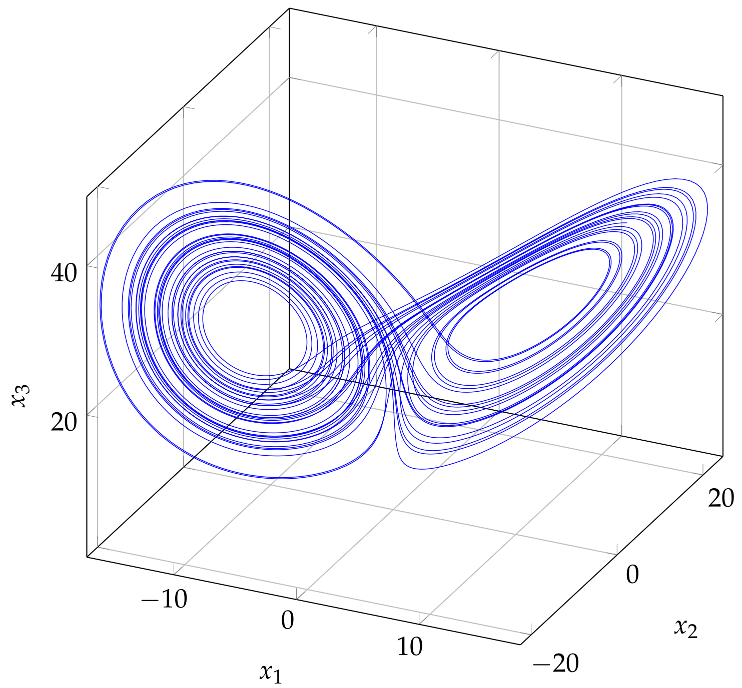
Chaotic dynamics

According to the Poincaré Bendixon theorem, the only attractors of a planar continuous-time system can be equilibria or limit cycles. As soon as we add a third dimension or we move to discrete time, however, things can get a lot more complicated: the menu of possible attractors now includes tori and chaotic attractors. We will learn later on how tori are rather fragile objects, in the sense that they tend to disappear under very small perturbations of the vector field, unless a special structure is assumed (e.g., Hamiltonian systems). Chaotic attractors, on the other hand, are surprisingly common.

We have already seen one in the Lorenz system, which was originally derived as a simplified model of atmospheric convection:

Example 50 (Lorenz system)**.**

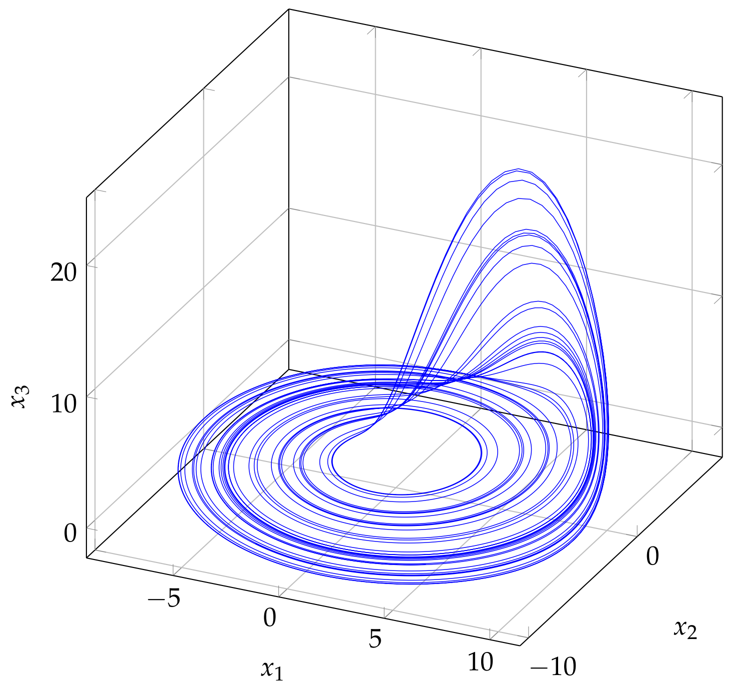
$$\begin{aligned}
\dot{x}_1 &= 10(x_2 - x_1), \\
\dot{x}_2 &= x_1(28 - x_3) - x_2, \\
\dot{x}_3 &= x_1 x_2 - \frac{8}{3} x_3.
\end{aligned}$$



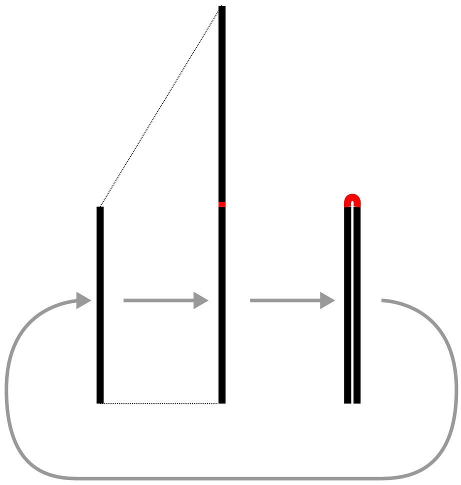
Another famous one is found in the Rössler system. Initially constructed as an academic example of a simple system exhibiting chaos (it is a linear system with the addition of a single quadratic term), the Rössler model was later found to be a good prototype for chaotic dynamics in chemical reactions (Scott, 1991).

Example 51 (Rössler system)**.**

$$\begin{aligned}\dot{x}_1 &= -x_2 - x_3,\\ \dot{x}_2 &= x_1 + 0.2x_2,\\ \dot{x}_3 &= 0.2 + x_3(x_1 - 5.7).\end{aligned}$$



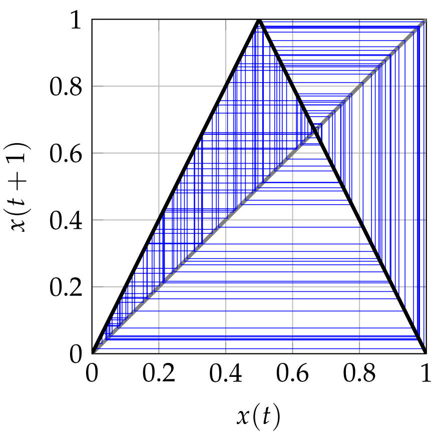
Unfortunately, if chaotic attractors are common, characterizing an attractor as chaotic, that is, distinguishing it from a very long limit cycle, is not trivial. The essence of chaos, sensitive dependence on initial within a bounded positively invariant set, is generated in the majority of chaotic attractors by a common topological principle, which is elegantly described by a mechanism known as Smale's horseshoe:



A set of states within the attractor is stretched (this introduces the sensitive dependence on initial conditions), then folded back to itself (this ensures positive invariance). The simplest prototype of this mechanism is provided by the tent map,

$$x(t+1) = \begin{cases} 2x(t), & x(t) \leq 0.5, \\ 2 - 2x(t), & x(t) > 0.5. \end{cases}$$

where the set is the real interval $[0, 1]$, the stretching is given by the multiplication by 2, and the folding by the tent shape.



Consider the binary representation of the points in the interval $[0, 1]$, and notice that the binary number $0.111\ldots$, with infinitely repeating 1s, correspond to unity, since it is $\lim_{i\to\infty} \frac{1}{2^i}$. In this representation, we can easily see that $1 - x$ is obtained by flipping all the fractional digits of x. Let us call this operator $\neg$. Let us also define the left-shift operator $\leftarrow (x)$, which shifts left the bits of a binary representation, and implements the multiplication by 2. Finally, let us call x_1 the first fractional digit of x The tent map can now be written as

$$x(t+1) = \begin{cases} \leftarrow (x(t)) & x_1(t) = 0 \\ \leftarrow (\neg x(t)) & x_1(t) = 1 \end{cases}$$

The structure of the tent map orbits can, in this way, be deduced from the sequences of repeating 0s and 1s in the binary representation of the real numbers in $[0, 1]$.

Sensitive dependence on the initial condition for the dynamics of the $[0, 1]$ interval of this map can be proved as a simple exercise of symbolic dynamics (see note on the right). A ghost of the stretch-and-fold mechanism can also be seen in the single lobe of the Rössler attractor, which stretches the orbits spiralling out and reinjects them within the spiral, and in the connection of the two lobes of the Lorenz attractor, which are connected so as to reinject into each other orbits once they have spiralled sufficiently far within the two wings of the butterfly. The attractors of less trivial systems, however, cannot be treated as easily, and a formal proof of chaoticity by e.g. symbolic dynamics is in general far from trivial. A more common approach is then to extrapolate from the tools we have been using so far to analyse the stability of equilibria and periodic orbits, that is, eigenvalues and the fundamental matrix. The tool that results from this extrapolation are the Lyapunov exponents.

Lyapunov exponents

Formally, an attractor is chaotic if it exhibits sensitive dependence on initial conditions. Notice that sensitive dependence alone is not an indication of chaos; in

$$\dot{x} = x$$

solutions depend sensitively on initial conditions (distance between $\phi_t(x)$ and $\phi_t(x')$ grows exponentially for $x \neq x'$), yet there is nothing chaotic about this linear system. If, however, the orbits *within* the attractor diverge exponentially, this is indeed an indication of chaos. In practice, this condition can be evaluated recurring to the computation of the attractor's *Lyapunov exponents*.

Consider once again the fundamental matrix of the system

$$\dot{\xi} = J_f(\phi_t(x))\xi,$$

which we used to define the monodromy matrix when $\phi_\cdot(x)$ is a periodic orbit.

If we now take a generic orbit $\phi_\cdot(x)$, periodic or not, we may attempt to measure how fast nearby orbits converge to it, or diverge from it, by measuring the rate of divergence

$$\frac{\|X(t)\xi\|}{\|\xi\|},$$

where X is the fundamental matrix and ξ is a small perturbation from $\phi_\cdot(x)$. This rate is of course time-dependent. We may obtain a time-independent quantity by taking its time average:

$$\lim_{T\to\infty} \sup_{t>T} \frac{1}{t} \log \frac{\|X(t)\xi\|}{\|\xi\|}.$$

In the above function, $\lim_{T\to\infty} \sup_{t>T}$ is more commonly denoted $\limsup_{t\to\infty}$, and is needed (as opposed to a simple $\lim_{t\to\infty}$) to handle cases where the limit itself does not exist, though in many cases we may use the simple $\lim_{t\to\infty}$.

More details on when lim can be used are found in (Wiggins, 2003), for example.

The value of the above limit of course depends on the choice of the perturbation ξ, and is somehow related to the different directions in which the matrix $J_f(\phi_\cdot(x))$ may expand or contract vectors. One can however prove that, given an n-dimensional flow, there are at most n different values of this limit.

■ **Definition: Lyapunov exponents**
The Lyapunov exponents of an orbit $\phi_\cdot(x)$ are the (at most) n values of

$$\limsup_{t\to\infty} \frac{1}{t} \log \frac{\|X(t)v\|}{\|v\|},$$

for $v \in \mathbb{R}^n$, where $X(t)$ is the fundamental matrix of

$$\dot{\xi} = J_f(\phi_t(x))\xi.$$

To clarify the meaning of this expression, imagine that vector ξ be stretched exponentially with exponent λ, so that $\xi(t) = e^{\lambda t}\xi$. Then, the 'Lyapunov exponent' of this transformation is

$$\lim_{t\to\infty} \frac{1}{t} \log e^{\lambda t} = \lambda.$$

★ Imagine writing the Lyapunov exponents in decreasing order

$$\lambda_1 > \lambda_2 > \dots,$$

and imagine taking a small m-dimensional cube of initial perturbations, with $1 \le m \le n$, and measuring its m-dimensional volume $V(t)$ as it is transformed by the flow of

$$\dot{\xi} = J_f(\phi_t(x))\xi.$$

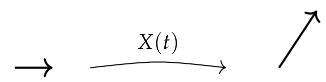
We have

$$\lambda_1 + \lambda_2 + \ldots + \lambda_m = \lim_{t\to\infty} \frac{1}{t} \log \frac{V(t)}{V(0)}.$$

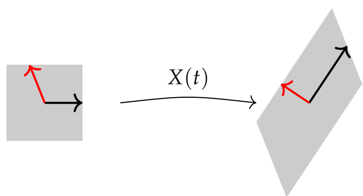
Hence, the largest exponent measures how the main direction of the cube (the base) is transformed by the flow, the other exponents measure how the minor direction (the heights) are transformed, their sum giving the rate of change of the volume.

Why n different exponents, and why this relation between exponents and areas, volumes,...?

Take a one-dimensional straight segment of initial conditions. It is transformed by the linear time-varying vector field in a straight segment (linearity!), of length on average $e^{\lambda_1 t}$ times the length of the original segment, where λ_1 is the (largest) Lyapunov exponent:



Now take a two-dimensional square of initial conditions. It is transformed into a parallelogram:



As $t \to \infty$, the long side of the parallelogram grows faster than the short side, as they grow exponentially with different exponents, and the parallelogram becomes thinner. Asymptotically, most of the vectors of initial conditions within the original square become aligned with the long side of the parallelogram, so it does not matter which initial condition we choose to compute the largest Lyapunov exponent, the asymptotic value will remain almost always the same.

There is, however, one (time-dependent) direction which will give a different value: the red arrow in the figure. This direction transforms according to the height of the parallelogram, proportional to $e^{\lambda_2 t}$ with $\lambda_2 < \lambda_1$.

If we now take a 3-dimensional cube of initial conditions we can repeat the reasoning, finding a third exponent $\lambda_3 < \lambda_2 < \lambda_1$. We can find at most n such exponents, and the sum of the first m gives the rate of growth of an m-dimensional volume.

♦ **Theorem**
The Lyapunov exponents of an orbit $\phi_\cdot(x)$ are independent of the initial condition.

This is a simple consequence of the $\lim_{t\to\infty}$ in the definition. It means that the Lyapunov exponents are a property of the orbit, not of the initial condition.

♦ **Theorem**
Almost all orbits converging to the same attractor have the same Lyapunov exponents

We can therefore expect to compute the Lyapunov exponents of an attractor by computing those of a generic orbit converging to it. Moreover, Lyapunov exponents characterize the attractor in a coordinate-free way, that is, they do not change through a change of variables:

♦ **Theorem**
Corresponding attractors in flows that are topologically conjugated through a diffeomorphism have the same Lyapunov exponents.

Note that the above theorem implicitly requires the diffeomorphism to be defined over the whole attractor.

For the simple attractors seen before – equilibria and limit cycles – we have a trivial relation between the Lyapunov exponents and the eigenvalues or multipliers.

♦ **Theorem**
The Lyapunov exponents of an equilibrium are

$$\Re(\lambda_i)$$

where λ_i are the eigenvalues of the Jacobian matrix at the equilibrium.

♦ **Theorem**
The Lyapunov exponents of a periodic orbit are $\frac{1}{T}\log\|\lambda_i\|$, where λ_i are the Floquet multipliers of the orbit.

The above theorem in particular implies that 0 is always a Lyapunov exponent of a periodic orbit, given that 1 is always a Floquet multiplier. This is actually a more general property:

♦ **Theorem**
If $\phi.(x)$ is a bounded orbit of a continuous-time system, that does not tend to an equilibrium, then it has 0 as a Lyapunov exponent.

Intuitively, this property follows from the fact that two points on a bounded orbit not converging to an equilibrium cannot tend to each other (or the orbit would be converging to an equilibrium), but also cannot diverge (since the orbit is bounded and therefore $f(x)$ is bounded), hence on average they maintain a constant distance.

We are finally ready to state a criterion to determine if an attractor is chaotic.

♦ **Lyapunov exponents criterion for chaos**
If an attractor has a positive Lyapunov exponent then it is chaotic, with *exponential* sensitive dependence on initial conditions.

In the theorem, exponential sensitive dependence on initial conditions means that orbits separate at an exponential rate. This is not always the case, though there are no general-purpose results to determine sensitive dependence on initial conditions in attractors with no positive Lyapunov exponents.

We have seen how Lyapunov exponents measure how fast nearby orbits diverge. The largest among these exponents should therefore give us an estimate of how far ahead in time the numerical integration of an orbit can be considered reliable, provided that any numerical integration scheme is bound to introduce some error, albeit small.

■ **Definition: Lyapunov time**

Given an orbit with positive largest Lyapunov exponent λ_1, its Lyapunov time is

$$\frac{1}{\lambda_1}.$$

The Lyapunov time should be considered an estimate of the order of magnitude of the horizon of predictability of an orbit, definitely not an accurate measure of it.

We easily conclude from the above definitions that, after n Lyapunov times, the distance between nearby initial conditions is expected to grow by a factor e^n.

Lyapunov exponents computation

The estimation of the largest Lyapunov exponent is relatively straightforward: given an arbitrary vector v of norm 1,

$$L \simeq \frac{1}{T} \log \|X(T)v\|, T >> 0.$$

This is because an arbitrary vector will typically contain components aligned with the direction of maximum growth, hence in the long term it will align along this direction.

The computation of the full Lyapunov spectrum is a bit more complex and relies on the following fact

♦ **Theorem**

Any real square matrix can be decomposed in the product QR of two matrices, an orthogonal matrix Q and an upper triangular matrix R with positive diagonal elements such that $R_{ii} \geq R_{(i+1)(i+1)}$.

Not all definitions of the QR decomposition enforce positive diagonal elements. Matlab implementation, for example, does not. It is however always possible to choose Q such that R has positive diagonal elements.

♦ **Theorem**

Given the QR decomposition of matrix $X(T)$, the product of the first m elements R_{ii}, $1 \leq m \leq n$ gives the area of a generic m-dimensional unit cube C_m defined by the first m columns of the identity matrix, under the linear transformation $X(T)C_m$.

In other words, the diagonal elements of R represent how fast an m-dimensional cube, $1 \leq m \leq n$, is stretched in different directions by the flow. This is precisely what we need to compute the Lyapunov exponents.

♦ **Theorem**
For almost all coordinate choices, the Lyapunov exponents of $f(x)$ are given by

$$\lim_{t\to\infty} \frac{1}{t} \log R_{ii}.$$

See (Geist, Parlitz, and Lauterborn, 1990) or (Dieci and Vleck, 2002) for more information on the meaning of 'for almost all coordinate choices'.

In practice, the above computational method cannot be applied as it is, because the ratios between the elements R_{ii} diverge in time, causing numerical problems. A more robust approach is to compute the exponents as the average of the expansion rates of multiple boxes of initial conditions, over a sequence of small time intervals, as sketched in the following algorithm.

■ **Definition: numerical computation of the Lyapunov spectrum**

1: choose $N \gg 1$
2: define a small time interval τ
3: define $Q_0 = I$
4: set $t = 0$
5: **for** $k = 1 : N$ **do**
6: compute $Y(\tau)$ as solution of

$$\dot{Y} = J_f(\phi_t(x(k\tau)))Y, \quad Y(0) = Q_{k-1}$$

7: compute R_k and Q_k as QR decomposition of $Y(\tau)$.
8: **end for**
9: **return** $\lambda_i = \frac{1}{N\tau}\sum_k \log R_{k,ii}$.

By the definition of the fundamental matrix $X(T)$, when $T = k\tau$ we have

$$X(T) = X_k X_{k-1} \cdots X_1$$

where X_i is the fundamental matrix of the orbit between times $(i-1)\tau$ and $i\tau$. Remember that each fundamental matrix in the above equation is computed by solving a differential equation with initial condition $X(0) = I$. By using initial condition $Y(0) = Q_{k-1}$ in the algorithm, we obtain $Y_k = X_k Q_{k-1}$, that is, $X_k = Y_k Q_{k-1}^{-1}$. If we now write $Y_k = Q_k R_k$, and put this in the equation of $X(T)$, we obtain

$$\begin{aligned} X(T) &= Q_k R_k Q_{k-1}^{-1} Q_{k-1} R_{k-1} Q_{k-2}^{-1} \cdots R_1 \\ &= Q_k R_k R_{k-1} \cdots R_1. \end{aligned}$$

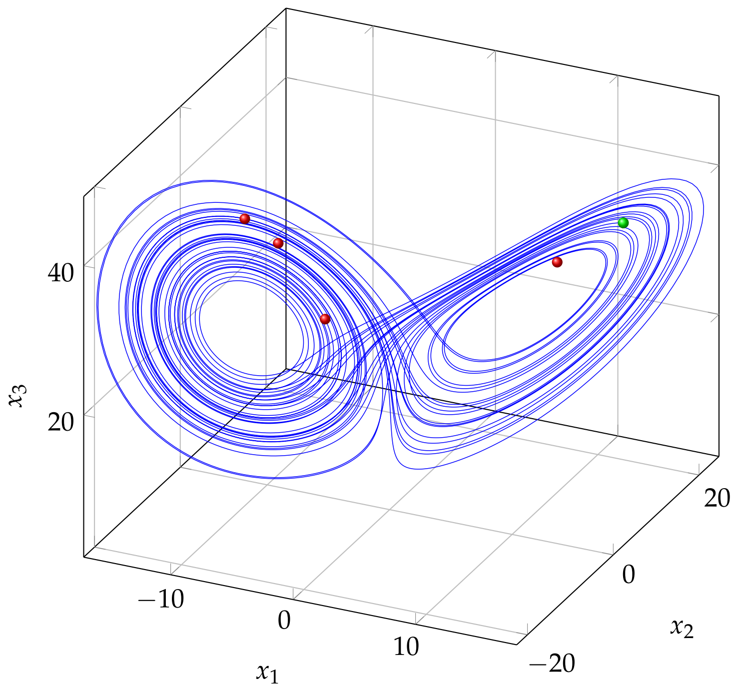
This means that $X(T)$ has for QR factorization a matrix $Q = Q_k$, and a matrix R whose diagonal elements are the products of the diagonal elements of $R_1, \ldots, R_k$. This explains why line 9 in the algorithm produces a correct estimate of the Lyapunov exponents for large enough N.

Example 52. *The Lyapunov exponents of the Lorenz attractor, with standard parameters, are approximately* 0.906, 0, *and* -14.572. *The largest one indicates that generic neighbouring initial conditions diverge as* $e^{0.906t}$, *with Lyapunov time* $\frac{1}{0.906} \simeq 1.1$. *If we take initial conditions that differ by* 10^{-3} *and integrate for* 10 *unit times, that is, about* 9 *Lyapunov times, we expect the initial conditions to separate by a factor of about*

$$e^9 \simeq 10^4.$$

Of course, while the speed of divergence is reasonably approximated by the first Lyapunov exponent when the orbits of the two initial conditions are very close, we are stretching the tool quite far in the above calculation. Nevertheless, our estimate happens to work surprisingly well in the Lorenz system. The following figure shows in green 4 *initial conditions, at a distance of* 10^{-3} *from each other, and in red the corresponding states after numerically integrating for* 10 *units of time.*

Note that, while the Lyapunov time may accurately estimate divergence of extremely near initial conditions, the farther away they split the less accurate is the estimate. For instance, in this example, two solutions within the attractor will never be split by more than about 40 units, which is the diameter of the attractor.



Lyapunov exponents of discrete-time systems

The discussion in the two previous sections holds, with minimal changes, for discrete-time systems. The numerical computation of the exponents is identical since, as we have seen even in continuous-time systems it is reduced to the iteration of discrete-time steps.

Example 53. *The tent map*

$$x(t+1) = \begin{cases} 2x(t), & x(t) \leq 0.5, \\ 2 - 2x(t), & x(t) > 0.5, \end{cases}$$

has

$$\|J_f(x)\| = 2, \forall x \neq 0.5.$$

For an orbit not going through $x = 0.5$, we have

$$\lim_{t\to\infty} \frac{1}{t} \log 2^t = \log 2.$$

This is one of the few cases where the Lyapunov exponent can be computed analytically.

Notice that the interval $[0,1]$ is not an attractor, by our definition. Indeed, if we take $\mathbb{R}$ as the host space, no neighbourhood of $[0,1]$ is attracted to $[0,1]$ asymptotically: the set is invariant but not asymptotically stable. It becomes a (chaotic) attractor if the tent structure is repeated throughout $\mathbb{R}$.

Fractal geometry

We can now characterize the complexity of nonlinear dynamics in terms of the Lyapunov exponents, and the ensuing sensitive dependence on initial conditions of chaotic attractors. Another manifestation of dynamic complexity, possibly even more pervasive than chaos, is the fractal dimension of relevant objects, and namely of attractors and basin boundaries. To discuss this we need to start from a definition of dimension more general, and flexible, than the intuitive definition we have of 0, 1, 2-dimensional sets (points, lines, surfaces,...). We find such a tool in the definition of measure and dimension initially introduced by Felix Hausdorff in 1918.

Definition: Hausdorff measure of a set A
The Hausdorff measure of a set A in parameter d is

$$H^d(A) := \lim_{\epsilon \to 0} \inf \sum_{i=1}^{\infty} \mathrm{diam}(B_i(\epsilon))^d,$$

where $B_i(\epsilon)$ are open sets of a diameter smaller than ϵ forming a countable cover of A, and $\mathrm{diam}(\cdot)$ is the diameter of an open set, that is, the sup of the distance between arbitrary points in the set.

The Hausdorff measure of parameter d thus measures the speed of growth in the number of elements B_i necessary to cover set A, as their diameter, weighed by the power d, shrinks. We can intuitively see how for large d the measure is 0 (the number of elements grows slower than the d-th power of their diameter shrinks), while for small enough d it goes to ∞. There is one special d that marks the transition between these two limits, and that gives the Hausdorff dimension.

The Hausdorff dimension and other definitions of fractal dimension are sometimes used to characterize the complexity of a dataset, receiving rather interesting interpretations. For instance, Fuss and Niegl (2008) used the Hausdorff dimension of the force signal of a climber on a hold to characterize athletic performance, considering it as a proxy of how precisely the climber interacts with the climbing wall. Schubert et al. (2009) and Melillo, Bracale, and Pecchia (2011) used the *correlation dimension* of the ECG signal, an alternative way of measuring fractal dimension, to measure stress.

Definition: Hausdorff dimension of a set A
The Hausdorff dimension of set A is

$$\inf d : H^d(A) = 0.$$

The Hausdorff dimension is, in simpler words, the weight by which we must weigh the diameters of the covering elements in order to balance their increasing number. This generalizes the idea that, for example, to cover a line the number of cover elements must increase inversely to their diameter, while to cover a surface the number must grow as the square of their diameter. The definition is however quite involved, and not easy to apply in general. There exists a less rigorous definition, which provides incorrect results for some sets but has the merit of having a simple implementation in numerous cases of interest.

Definition: Box-counting dimension
Let S be a set in $\mathbb{R}^n$, and $N(\epsilon)$ the number of n-dimensional cubes of side ϵ necessary to cover S. The box-counting dimension of S is

$$\lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}}.$$

The Box-counting dimension can be identically defined using spheres instead of cubes. It is not mathematically sound (*e.g.*, it estimates $d = 1$ for the countable set of the rationals in $[0, 1]$) and it is not the most efficient dimension to compute numerically, but it is the simplest one to explain.

Example 54 (Box-counting dimension of regular sets)**.** *Regular sets have integer box-counting dimensions. Consider, for instance, a planar* *region of area A and a 3-dimensional region of volume V in $\mathbb{R}^3$. The number of cubic boxes necessary to cover the planar region, for small ϵ is equal to $\frac{A}{\epsilon^2}$, hence*

$$\lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}} = \lim_{\epsilon \to 0} \frac{\log \left(\frac{A}{\epsilon^2} \right)}{\log \frac{1}{\epsilon}} = \lim_{\epsilon \to 0} \frac{\log A + 2 \log \frac{1}{\epsilon}}{\log \frac{1}{\epsilon}} = 2.$$

Similarly, The number of cubic boxes necessary to cover the 3-dimensional region is proportional to $\left(\frac{1}{\epsilon}\right)^3$, hence

$$\lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}} = \lim_{\epsilon \to 0} \frac{\log \left(\frac{V}{\epsilon^3} \right)}{\log \frac{1}{\epsilon}} = \lim_{\epsilon \to 0} \frac{\log V + 3 \log \frac{1}{\epsilon}}{\log \frac{1}{\epsilon}} = 3.$$

■ **Definition: Cantor set**

The Cantor set is the subset of $\mathbb{R}$ obtained by taking the unit interval, removing its central third, and then iteratively repeating the operation on each of the obtained segments.

Many continuous-time chaotic attractors have Poincaré maps that are homeomorphic to a Cantor set.

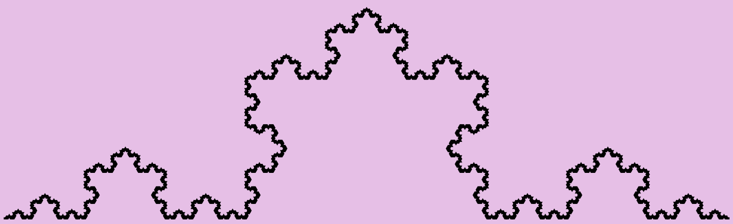
Example 55 (Box-counting dimension of the Cantor set)**.** *Let S_k be the approximation of the Cantor set obtained by iterating k times the above procedure so that S_0 is the unit interval, S_1 the union of $[0, 1/3]$ and $[2/3, 1]$, and so on.*

The set S_k is covered by exactly 2^k boxes of side $\epsilon = \frac{1}{3^k}$. Taking $\epsilon \to 0$, and simultaneously taking $k \to \infty$, we have

$$\lim_{\epsilon \to 0} \frac{\log N(\epsilon)}{\log \frac{1}{\epsilon}} = \lim_{k \to \infty} \frac{\log 2^k}{\log 3^k} = \frac{\log 2}{\log 3} \simeq 0.6309.$$

■ **Definition: Koch curve**

The Koch curve is obtained by taking the unit interval, substituting its central third with the two upper sides of an equilateral triangle, and iterating the procedure on each of the sides of the curve.



In the figure, the 6-th iteration in the computation of the Koch curve.

Example 56 (Box-counting dimension of the Koch curve)**.** *Consider the Koch curve in $\mathbb{R}^2$. For this example, it is easier to take a spherical cover.*

Call K_k the approximation of the Koch curve obtained after k iterations of the generating algorithm, with K_0 being the unit segment. The set is minimally covered by 4^k spheres of diameter $\frac{1}{3^k}$.

Defining $\epsilon = \frac{1}{3^k}$, and taking $k \to \infty$, we have

$$d = \lim_{k\to\infty} \frac{\log 4^k}{\log 3^k} = \frac{\log 4}{\log 3} \simeq 1.2619.$$

■ **Definition: Fractal set**
A fractal set is a set with non-integer dimension

In this definition, box-counting, or other dimensions can be used indifferently.

■ **Definition: Strange attractor**
Attractors with fractal geometry are called strange attractors.

This definition follows the convention in (Meiss, 2007). Other references use the term *strange* with slightly different meanings.

The Lorenz attractor, for instance, is believed to be fractal, and therefore strange.

★ Notice that there exists attractors that are strange but not chaotic, and chaotic attractors that are not strange.

Fractal geometry, however, does not come in the game of nonlinear dynamics only when dealing with chaotic attractors. The boundary of a basin of attraction can in fact be fractal, and this is possible even for the simplest type of attractor: the stable equilibrium.

Example 57. *Newton's method for root finding consists of the iteration of the discrete-time, nonlinear system*

$$x(t+1) = x(t) - J_f^{-1}(x(t))f(x(t)).$$

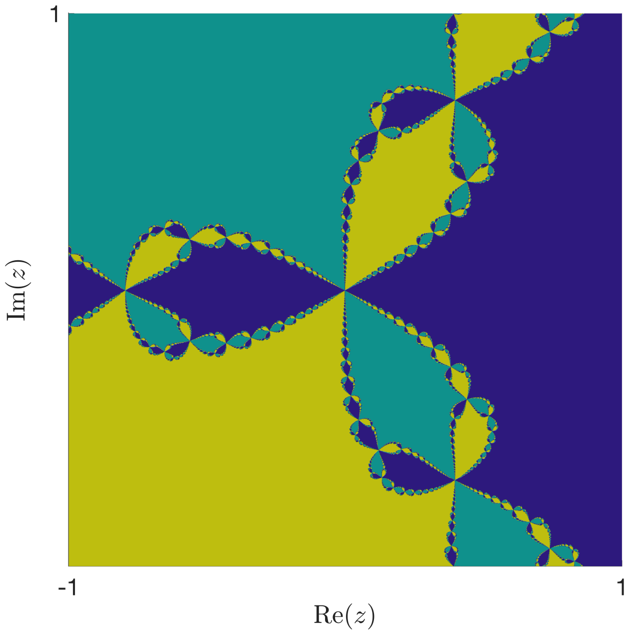
Consider the complex function $f(z) = z^3 - 1$, which can be written as a real vector-valued function

$$f(x) = \begin{pmatrix} x_1^3 - 3x_1x_2^2 - 1 \\ 3x_1^2x_2 - x_2^3 \end{pmatrix}.$$

The function has, of course, three roots equal to

$$(1,0), (\cos(2\pi/3), \sin(2\pi/3)), (\cos(4\pi/3), \sin(4\pi/3)),$$

and each of these roots is an asymptotically stable fixed point of Newton's algorithm. If we attempt to find them using Newton's method, however, the one we converge to depends on our choice of initial conditions in a very complex way: the boundary of the basins of attraction of the three fixed points, corresponding to the three roots, are in fact fractal. This is evident in the picture below, which represents in three different colours the root towards which the initial conditions is attracted.



An obvious consequence of the above observation is that deciding precisely what perturbations in the state of a multi-stable system may lead it to settle on a different attractor is, in general, a very hard problem. Clearly, in the nonlinear setting, a Lyapunov function and LaSalle's invariance theorem can only provide a very rough under-approximation of a generic basin of attraction.

In closing, note how the definitions of dimension introduced before are exclusively geometric. In applying them to the study of an attractor, the attractor set must first be computed, and then its dimension must be calculated based on one or the other definition. There exist however other definitions, more closely related to the dynamic nature of attractors. One that is not rigorously justified, but very simple to implement, is the following.

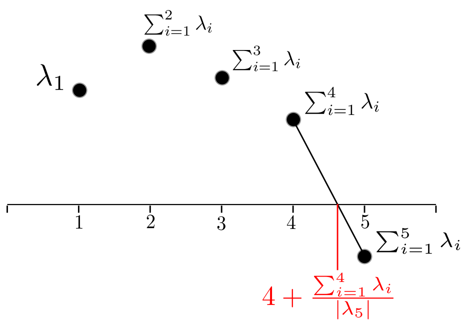
Definition: Lyapunov dimension of an attractor

Consider an attractor with Lyapunov exponents $\lambda_1, \lambda_2, \ldots$, in decreasing order. Its Lyapunov dimension is

$$k + \frac{\sum_{i=1}^{k} \lambda_i}{|\lambda_{k+1}|},$$

with

$$k := \max \left\{ k : \sum_{i=1}^{k} \lambda_i \geq 0 \right\}.$$



The intuition behind this definition is that the (fractal) dimension of the attractor is the dimension of volumes that neither grow nor shrink under the effect of the attractor's flow. In the defining formula, k is the largest integer dimension of volumes that increase under the flow, while the fractional term approximates the fraction of dimension missing to exactly balance growth and contraction.

Example 58. *The tent map has only one Lyapunov exponent equal to* $\log 2$*, therefore it has dimension* 1*. The Lorenz attractor, with Lyapunov exponents* 0.906*,* 0*, and* -14.572*, is equal to* $2 + \frac{0.906}{14.572} \simeq 2.06$*.*

Exercises

Exercise 50

Consider the linear system

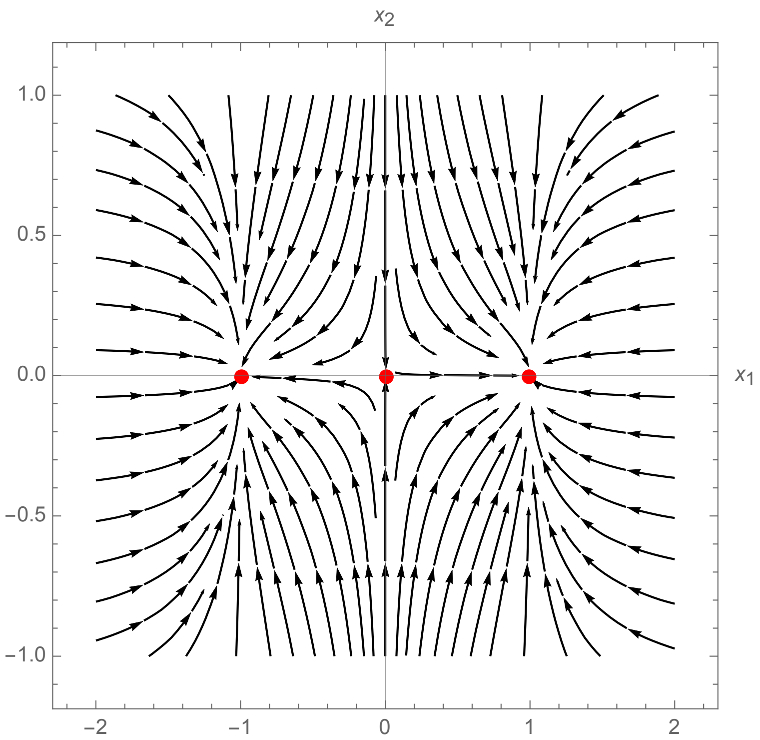
$$\dot{x} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} x.$$

1. Take a square of initial conditions centred around the origin, and draw how the square is transformed by the flow after $t = 1, 10, 100$.
2. Is the area of the parallelogram increasing or decreasing?
3. What are the Lyapunov exponents of the equilibrium?
4. Consider an arbitrary planar polytope of initial conditions, not necessarily including the origin. Does the area of the polytope increase or decrease as t grows?

Exercise 51

Compute the Lyapunov exponents of all the orbits of the system

$$\dot{x}_1 = x_1 - x_1^3,$$
$$\dot{x}_2 = -x_2.$$



Exercise 52

Consider the following model

$$\dot{x}_1 = x_1 + x_1 x_2 - \tanh(x_2^4) + x_1 x_2^3 - 3,$$
$$\dot{x}_2 = x_2 - \cos(x_1 x_2) + x_1^2 x_2^7.$$

Can the model have a chaotic attractor?

Exercise 53

We have seen in an exercise that the Lotka Volterra model admits an infinite number of periodic orbits surrounding the positive equilibrium. What are the Lyapunov exponents of these orbits?

Answer of exercise 53

The periodic orbits must have two 0 exponents: one coming from the trivial Floquet multiplier, the second due to the fact that nearby orbits are also periodic, therefore neighbouring initial conditions diverge at most polynomially in t (due to the growing phase difference).

*** Exercise 54**

Let

$$\dot{x} = \Sigma(x, p)$$

be a continuous-time model of a food chain in the parameter p. Let $p = p_0 + g(w)$ with w solution of

$$\dot{w} = W(w),$$

where $g(w)$ is the output of a nonlinear system evolving on a chaotic attractor, modelling the chaotic variation of the parameter p due to the weather. The Lyapunov exponents of the only attractor of $\Sigma(x, p_0)$ (the unperturbed food chain) are $-2, -1, 0$, while those of the only attractor of W are $-4, 0, 1$. What can we say about the exponents of the seasonally perturbed system

$$\dot{x} = \Sigma(x, p_0 + g(w)),$$
$$\dot{w} = W(w)?$$

Answer of exercise 54

The dynamics of W does not depend on that of Σ, therefore the set of exponents of any attractor of the aggregate system must contain the exponents $-4, 0, 1$. The remaining exponents instead depend on the coupling between W and Σ, but nothing can be said in general about their value.

Exercise 55

A chaotic attractor has Lyapunov exponents $\{-4, -1, 0, 3\}$. What is the Lyapunov time of the system?

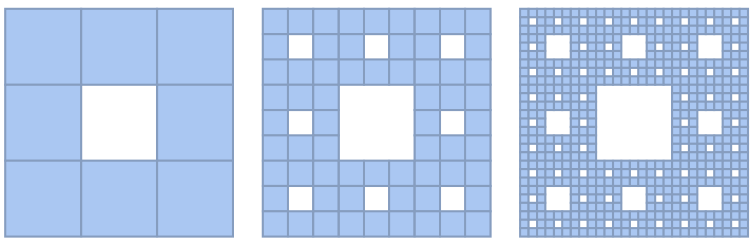
Exercise 56

A nonlinear continuous-time system has a quasi-periodic attractor on an invariant 2-torus, that is, an invariant torus that is a 2-dimensional manifold.

1. How many null Lyapunov exponents should we expect to see for an orbit converging to the attractor?
2. What is the largest Lyapunov exponents of the attractor?
3. For a small parameter change the attractor becomes periodic (it phase locks). How many null Lyapunov exponents should we expect to see now?

Exercise 57

The Sierpinski carpet is obtained by taking the unit square, removing a square of side $\frac{1}{3}$ from its centre, then removing a square of side $\frac{1}{3^2}$ from the centre of the 8 remaining squares, and iterating. Compute the box-counting dimension of the carpet.



Exercise 58

Consider the model $\dot{x} = 4x^2 - 10x + 6$, defined in the interval $x \in [1, \infty)$, and the mapping $y = \sqrt{x - 1}$.

1. Determine if the mapping is a homeomorphism or a diffeomorphism.
2. Write the model in the variable y.
3. Compute the Lyapunov exponent of the orbit with initial condition $x = 1.2$, and that of the corresponding orbit in the y variable.

Structural stability and bifurcations

Keywords: **Structural stability, bifurcation, bifurcation diagram, saddle-node, fold, Hopf, period-doubling, Neimark-Sacker**

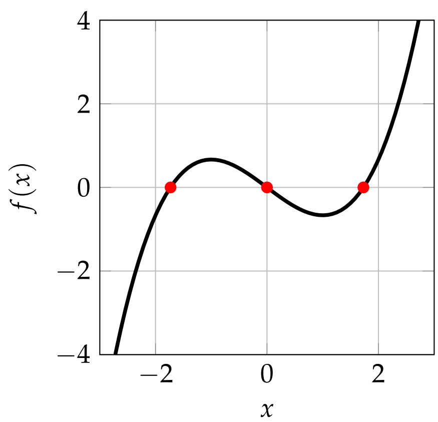
Structural stability

Definition: Structurally stable system
A system $f(x, p)$, with parameters $p \in \mathbb{R}^q$ is structurally stable if its flow is topologically equivalent to the flow of all systems $f(x, p')$ with p' in a neighbourhood of p.

Example 59. *Consider the one-dimensional model*

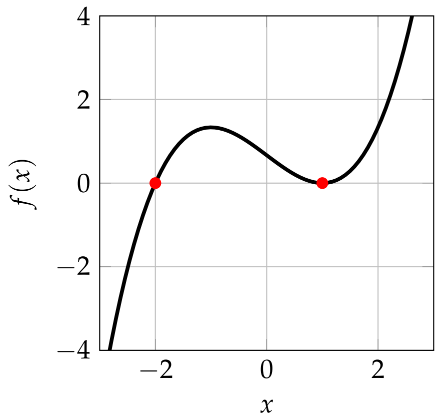
$$\dot{x} = -x + \frac{x^3}{3} + p,$$

whose vector field is depicted below for $p = 0$*:*



The model has 3 equilibria when $p = 0$*: two unstable ones in* $\pm\sqrt{3}$*, and a stable one in* 0*. By changing* p *we can move* $f(x)$ *up and down, moving its equilibria. If* p *is only slightly changed, the equilibria remain the same, and we could formally prove that the flow remains topologically equivalent to that with* $p = 0$*. It is structurally stable.*

If we set $p = \frac{2}{3}$*, however, we obtain the following vector field.*



Now we only have 2 *equilibria, an unstable one in* -2*, and a nonhyperbolic one in* 1*. A small perturbation of* p *on either side will change the system into one with* 3 *or* 1 *equilibria. The vector field* $f(x, 2/3)$ *therefore is not structurally stable.*

Example 60 (Structural stability and robustness of a feedback control)**.** *Let us attempt to design a feedback control to stabilize the system*

$$\dot{x} = x + u(1 + x) + d$$

around the origin, assuming a constant but unknown disturbance d*. Proceeding as by standard linear systems theory, we can linearise at* $x = 0, u = 0$ *obtaining*

$$\dot{x} = x + u,$$

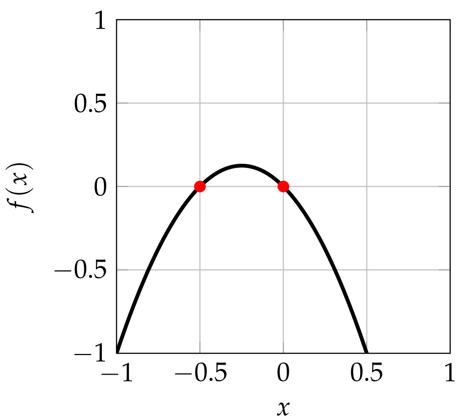
which can be stabilized at the origin by choosing

$$u = -kx, k > 1.$$

Let us take, for example, $k = 2$*.*

$$\dot{x} = -x - 2x^2 + d.$$

When $d = 0$*,* $f(x)$ *is*



We see that the origin is indeed a stable equilibrium, as expected. This equilibrium persists but is shifted right if the disturbance d *has a positive value. If however* $d < -\frac{1}{8}$*, the equilibrium collides with another, unstable equilibrium, and disappears! The unlucky control engineer would end with a system state drifting to* $-\infty$*, for any initial condition.*

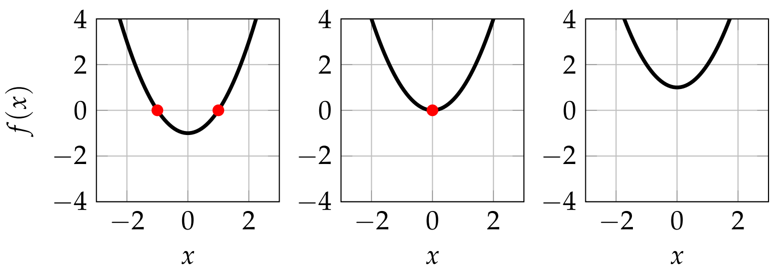
■ **Definition: Bifurcation point**
A bifurcation point is a point $\bar{p}$ in parameter space where $f(x, p)$ is not structurally stable.

A bifurcation, according to the above definition, is a phenomenon related to the whole phase space. In practice, however, we are frequently interested in the behaviour of a system near one of its attractors. In this case, we look for bifurcations only within its neighbourhood. In Example 59, for instance, we can observe that the bifurcations involve only the two equilibria at a time. We can effectively reduce the problem of studying each one of the two bifurcations in Example 59 to that of studying the behaviour of a simpler second-order system.

Example 61 (Saddle-node bifurcation of an equilibrium)**.** *For parameter values close to the bifurcation point, the flow in Example 59 near the rightmost or the leftmost pair of equilibria is locally topologically equivalent to that of the family*

$$\dot{x} = x^2 + p,$$

for p near 0.



■ **Definition: Bifurcation set**
The bifurcation set is the set of bifurcation points in parameter space.

■ **Definition: Codimension of the bifurcation set**
A bifurcation set is of codimension k if it is locally an m-dimensional manifold in n parameters, with $n - m = k$.

★ From a different point of view, we may view the codimension of the bifurcation set as the number of equations in the n parameters that are needed to define the set.

Example 62. *Consider the 2-parameter family of systems*

$$\dot{x} = p_1 x^2 + p_2.$$

If $p_1 > 0$, the system has 2 equilibria when $p_2 < 0$, and no equilibria when $p_2 > 0$. If $p_1 < 0$ the two equilibria exist when $p_2 > 0$ and no *equilibria exist for* $p_2 < 0$*. When* $p_1 = 0$ *the system has infinitely many or no equilibria, depending on* p_2*. Overall, the curves* $p_1 = 0$ *and* $p_2 = 0$ *in the* (p_1, p_2) *plane are codimension-1 bifurcation sets.*

Example 63. *Consider the 2-parameter family of systems*

$$\dot{x} = \frac{x^3}{3} - p_1 x - p_2.$$

For small p_2 *and* $p_1 > 0$ *it has 3 equilibria, two of which collide and then disappear as* p_2 *changes. This happens at parameter values for which* $f(x)$ *is tangent to the horizontal axis. We can identify the bifurcation set by solving*

$$\frac{\partial}{\partial x}\left(\frac{x^3}{3} - p_1 x - p_2\right) = x^2 - p_1 = 0 \Rightarrow x = \pm\sqrt{p_1},$$

and then solving

$$\left(\frac{x^3}{3} - p_1 x - p_2\right)\bigg|_{x=\pm\sqrt{p_1}} = \mp\frac{2p_1^{\frac{3}{2}}}{3} - p_2 = 0 \Rightarrow p_2 = \mp\frac{2p_1^{\frac{3}{2}}}{3}.$$

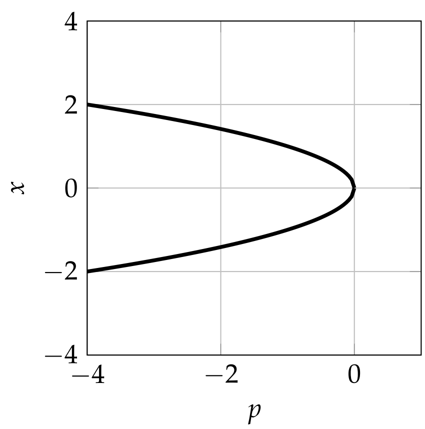
This is a codimension-1 bifurcation set. Notice that it contains, as a subset, the codimension-2 bifurcation set $p_1 = p_2 = 0$*, where all three equilibria coincide.*

■ **Definition: 1-parameter bifurcation diagram**
A 1-parameter bifurcation diagram is a plot of a system's equilibria [or limit cycles] against a parameter p in the (x_i, p) plane [or (x_i, x_j, p) space, for cycles].

Example 64. *Consider again the family of systems*

$$\dot{x} = x^2 + p,$$

Its 1-parameter bifurcation diagram is



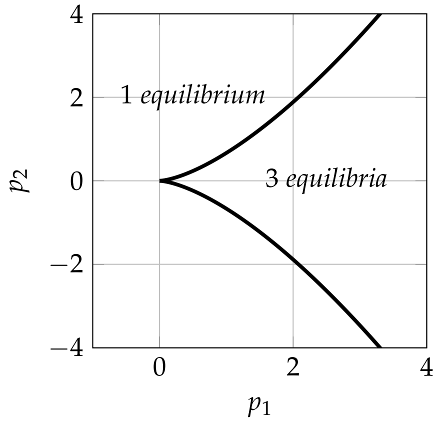
A codimension-1 bifurcation set in the plane is, for example, a curve.

■ **Definition: 2-parameter bifurcation diagram**
A 2-parameter bifurcation diagram is a plot of the bifurcation set of a system in a parameter plane.

Example 65. *Consider again the* 2*-parameter family of systems*

$$\dot{x} = \frac{x^3}{3} - p_1 x - p_2.$$

Its 2*-parameter bifurcation diagram is*



*Notice how the bifurcation set is a smooth manifold, except at the codimension-*2 *bifurcation in* $(0, 0)$.

The definition of bifurcation, and the fact that topologically equivalent flows have corresponding invariant sets, implies that a bifurcation must occur any time two invariant sets of a system collide at the change of a parameter. This is by no means a rigorous definition, but many of the most common and interesting bifurcations can indeed be thought of as collisions of equilibria, cycles, and other familiar objects.

★ **Bifurcations as collisions**
A collision of cycles or equilibria when a parameter changes is a bifurcation.

A different set of bifurcations, called *global* bifurcations, happen when the stable and unstable manifolds of saddle equilibria or cycles collide. These are called homoclinic bifurcations (connections between invariant manifolds of the same equilibrium/cycle) or heteroclinic bifurcations (connections between invariant manifolds of different equilibria/cycles)

While the above remark helps us visualize bifurcations, it does not provide a very useful means for their numerical study. A more practical set of tools comes as a consequence of the Hartman-Großman theorem, and the observation we did in an earlier exercise, that flows near linear equilibria with the same number of stable and unstable eigenvalues are locally topologically equivalent, irrespective of the type of the equilibria. These two results, together, tell us that nothing special can happen near an equilibrium, as long as it remains hyperbolic. We must look for bifurcations in nonhyperbolic equilibria.

★ **Bifurcations as loss of hyperbolicity**
Bifurcations of a small neighbourhood of an equilibrium occur at parameter values where the equilibrium becomes non-hyperbolic.

Homoclinic and heteroclinic bifurcations escape this rule, as nothing special happens to the equilibrium or cycle itself.

A second important tool then comes from the centre manifold theorem, which tells us that the hyperbolic eigenvectors of a nonhyperbolic equilibrium identify manifolds (the stable and unstable ones) where, again, nothing special can happen. Therefore, all the interesting phenomena we can expect from a bifurcation must happen within the centre manifold. This is very good news since the centre manifold has a dimension equal to the number of nonhyperbolic eigenvalues. In the most common cases, this is 1 or 2.

Topological normal forms

We are now ready to start our systematic study of the most common codimension-1 local bifurcations. We have already seen what codimension-1 means. The adjective 'local' means that the bifurcation is related to a loss of structural stability in the neighbourhood of an equilibrium or a fixed point. In other words, it is a loss of local topological equivalence between the flow of $f(\bar{x}, \bar{p})$, and the flow of $f(\bar{x}, p)$, for all p near $\bar{p}$.

When possible, we will discuss these bifurcations through the analysis of their *topological normal forms*. These are a third type of normal form, on top of the two seen previously. Let us introduce topological normal forms for continuous-time systems, leaving the obvious extension to discrete-time systems to the reader.

■ **Definition: Topological normal form**
System

$$\dot{x} = f(x, p) \text{ or } x(t+1) = f(x(t), p)$$

is a topological normal form for a bifurcation of an equilibrium or fixed point $\bar{x}$ satisfying a set of bifurcation conditions at $p = 0$ if all the systems that satisfy the same set of conditions have flows that are locally topologically equivalent, near $\bar{x}$, to the flow of $f(x, p)$.

Constructing topological normal forms is not an easy task, in general, and is an open problem even in one of the cases we see in our very limited foray into the topic. The procedure is however quite straightforward in principle:

1. A set of conditions on the eigenvalues of the equilibrium or fixed point $\bar{x}$ and on the coefficients of the vector field at $(x = \bar{x}, p = 0)$ is defined (the bifurcation conditions).

2. The Poincaré normal form of $f(x, p)$ near $x = \bar{x}$ is computed up to a suitable order

3. It is proved that the flow of the truncated normal form is locally topologically equivalent to the flow of any non-truncated normal form with the same coefficients.

The normal forms we will see are therefore truncated Poincaré normal forms, with the added property that, in a neighbourhood of $\bar{x}$, the tail of the expansion of the Poincaré normal form is irrelevant for the topology of the flow. The terms that appear in the normal forms are, consequently, the resonant terms of all flows with the same bifurcation conditions.

The saddle-node bifurcation

Consider an equilibrium $\bar{x}$ of a continuous-time system $\dot{x} = f(x, p)$.

■ **Definition: Saddle-node bifurcation**
A saddle-node of $f(x, p)$ at the equilibrium $\bar{x}$ occurs at parameter values where the Jacobian $J_f(\bar{x}, p)$ has a single 0 eigenvalue.

The corresponding centre manifold is therefore 1-dimensional. Assume, without loss of generality, that the saddle-node occurs at $\bar{x} = 0$, $p = 0$. Call $f_1(x_1, p)$, with $x_1 \in \mathbb{R}$, $p \in \mathbb{R}$, the 1-dimensional dynamics in the centre manifold. We have the following.

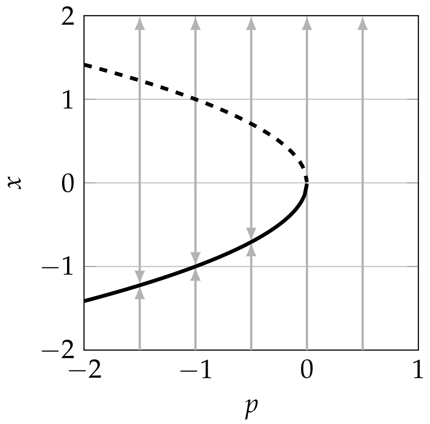
⧫ **Theorem: the normal form of a generic saddle-node**
Assume that $\frac{\partial^2}{\partial x_1^2} f_1(0,0) \neq 0$ and $\frac{\partial}{\partial p} f_1(0,0) \neq 0$. The flow in the centre manifold is locally topologically equivalent to the flow of

$$\dot{x}_1 = p \pm x_1^2.$$

For (many!) more details on the normal forms of this and other bifurcations, see (Kuznetsov, 2004)

Let us look at the dynamics of this normal form, for instance for the system $\dot{x}_1 = p + x_1^2$.

1-parameter bifurcation diagram of the saddle-node



In the diagram, the dashed line indicates the unstable equilibrium, and the solid line the stable one.

As p crosses 0 a stable node (below) and an unstable node (above) collide, and disappear for $p > 0$. Remember that, in general, this is what happens in the 1-dimensional centre manifold of a higher-dimensional equilibrium. Therefore in general we should expect to see an equilibrium with n_s stable eigenvalues and n_u unstable eigenvalues collide with one that has $n_s - 1$ stable eigenvalues and $n_u + 1$ unstable ones (this justifies the name saddle-node).

The theorem above tells us that, under some *genericity* conditions, the equation

$$\det J_f(\bar{x}, p) = 0$$

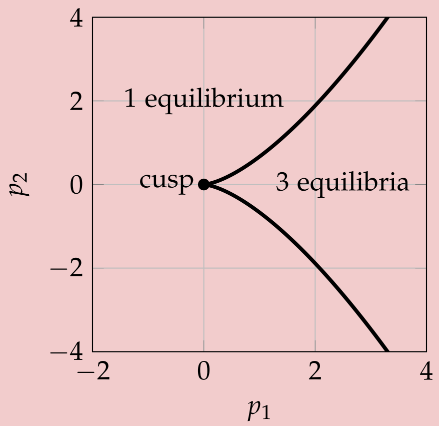
defines a saddle-node bifurcation set. Since this is a single scalar equation, the corresponding set is typically a codimension 1 set (an $(n-1)$-dimensional surface in the space of n parameters).

If we imagine moving along the codimension-1 bifurcation sets, we should expect to cross, occasionally, codimension-2 bifurcation points where one of the genericity conditions is violated. The corresponding codimension-2 bifurcations can themselves be unfolded, providing a normal form that is valid assuming further genericity conditions. Without going too deep into the details, at least for the case of the saddle-node it is worth looking at the two classes of codimension-2 bifurcations that happen to violate the two genericity conditions seen above.

Let us start by observing what happens when the first genericity condition, on the quadratic tangency of $f_1(x, 0)$ with the horizontal axis, is violated.

★ **The cusp bifurcation**

If the condition $\frac{\partial^2}{\partial x_1^2} f_1(0,0) \neq 0$ is violated, then $f_1(x_1, 0)$ is cubically tangent to $x_1 = 0$. The corresponding bifurcation is a codimension-2 bifurcation, called *cusp*, where two different saddle-node curves merge tangentially. We have seen the bifurcation diagram before:

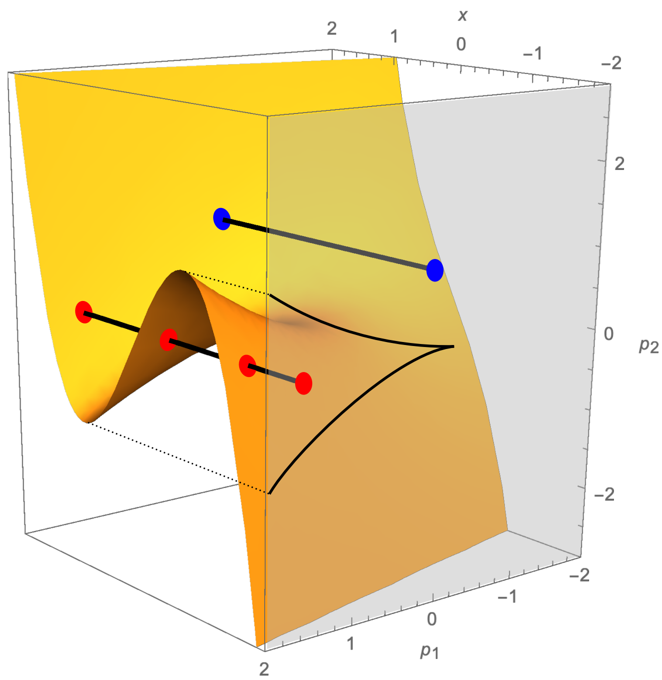


In some neighbourhood of the codimension-2 point, there must exist not 2 but 3 equilibria!

We may gain some further intuition about the geometry of the above bifurcation diagram if we imagine plotting the locus of the equilibria of a system of the form

$$\dot{x} = x^3 - p_1 x - p_2$$

in the space (x, p_1, p_2):

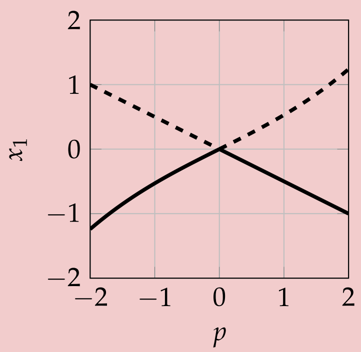


The locus is the yellow surface. If we imagine tracing a line in the x direction from the parameter pair p corresponding to the red dot, we cross the surface thrice: we have 3 equilibria. If instead, we start from the blue dot, we only cross the surface once: 1 equilibrium. The transition between these two scenarios, which is caused by the two *folds* of the surface, appears in the (p_1, p_2) plane as the cusp bifurcation diagram above.

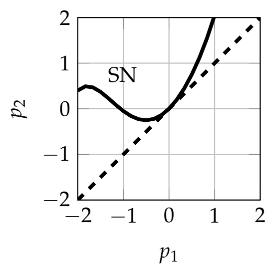
Let us now observe what happens when the second genericity condition, on the dependence of $f_1(0, p)$ on p, is violated.

★ **The transcritical bifurcation**

If the condition $\frac{\partial}{\partial p} f_1(\bar{x}_1(p), p) \neq 0$ is violated, then $f_1(x_1, p)$ defines a parabola that touches the horizontal axis for $p = 0$, but bounces back as p crosses 0. The corresponding bifurcation, called *transcritical* bifurcation, involves two equilibria that exist on either side of the bifurcation point and collide at the bifurcation point, exchanging stability through the bifurcation:



The transcritical appears to be a codimension-2 bifurcation since it occurs as a consequence of two equality conditions: $J_{f_1}(0,0) = 0$ and $\partial f_1(0,0)/\partial p = 0$. This is, however, not entirely correct. Given an arbitrary one-dimensional system $\dot{x} = f(x, p)$, consider arbitrary polynomial perturbations to the system, expressed as countably many parameters $p_1, p_2, \ldots$. Assume that an equilibrium of the system satisfies the conditions for the transcritical bifurcation $\partial f / \partial p_i = 0$ for a given p_i. This does not mean that the same condition will be satisfied for a different choice of parameters. In other words, from this point of view, the transcritical bifurcation is nothing more than the portrait seen while following a path in parameter plane (p_1, p_2) which happens to touch a generic saddle-node curve tangentially, for example following the path $\{p_1 = p_2\}$ in the following diagram.



A different path in parameter space would classify the same bifurcation point as a regular codimension-1 saddle-node (SN, solid black curve).

We may see an example of a transcritical bifurcation if we lower the carrying capacity (the coefficient 1/4 in the logistic model of the prey) in Example 17. This shifts the diagonal nullclines down, and eventually, the strictly positive equilibrium collides with the equilibrium on the x_1 axis, in a transcritical bifurcation.

Notice that, in this case, the bifurcation happens as we change a single parameter, and will keep happening (maybe for slightly different parameter values) even if we change all other parameters in the system. In other words, it appears to be a codimension-1 bifurcation. This has do to the special structure of the prey-predator model, and is not entirely unexpected: the particular algebraic structure of many mathematical models makes generic (i.e., codimension-1) bifurcations which are non generic (i.e., with codimension higher than 1) in the world of arbitrary dynamical systems.

Saddle-nodes, bistability, and hysteresis

The cusp structure seen above is just one of many, common examples of multistability in nonlinear systems. A multistable system has more than one attractor, and its asymptotic behaviour depends on the initial conditions. What makes the cusp structure (and other similar ones) particularly interesting is that the system may lose one or the other of the two attractors when parameters change. It can exhibit, in other words, a *hysteretic* behaviour. With specific reference to the cusp in the above figure, if p_1 is positive, and p_2 represents a slowly changing quantity that oscillates around 0, we may observe the system settling onto one of the two stable equilibria, and the suddenly jumping onto the other equilibrium when parameters cross both of the saddle-node curves, only to come back to the first equilibrium as parameters cross the two curves the other way.

Hysteresis is sometimes an unwanted, even tragic phenomenon in natural systems. For example, in some models describing the interaction between ocean water temperature and polar ice cover the state can tip from an equilibrium where temperature is on average colder and it is kept cool by albedo due to polar ice, to a much warmer one, where polar ice essentially disappears. Due to hysteresis, once the system is on the 'warmer' equilibrium it can only transition back to the first equilibrium through a large perturbation of parameters. In other systems, hysteresis has a welcome effect. In thermostats, for instance, it is often introduced on purpose in order to avoid the thermal controller switching on and off too quickly.

Let us have a more formal look at hysteresis:

■ **Definition: Bistability**
A bistable system is a system with two attractors.

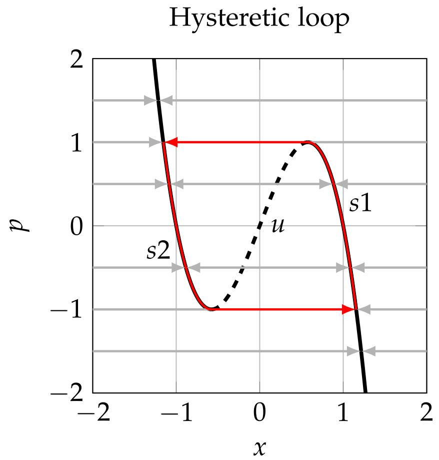
We may imagine a system with two equilibria, even though, in principle, the two attractors can be anything.

■ **Definition: Hysteresis**
A system whose parameters change slowly in time is said to exhibit hysteresis if its steady state depends on the history of its parameters, and not just their value.

Imagine, at first, that we have a system with three equilibria, two of which are stable (call them $s1$ and $s2$), and one is unstable (call it u), and the system depends on a parameter p. When p crosses 1 from below, the unstable equilibrium collides with $s1$ in a saddle-node, and after the saddle-node, the remaining equilibrium ($s2$) becomes GAS.

For $p > 1$, we expect the system to settle more or less quickly onto $s2$. If we now decrease p below 1, however, the system will stay on $s2$, *even if it was on* $s1$, *before!*.

Imagine now that the same structure repeats, symmetrically, when p crosses -1 from above: now $s2$ collides with u. If we follow p in and out of the interval $[-1, 1]$, we observe the system suddenly transitioning between the equilibria $s1$ and $s2$, in a hysteretic loop.



The cusp structure is of course not the only possible explanation of a hysteretic behaviour. Many of the bifurcations that we will encounter in the following can be combined in a multistable system to give rise to hysteretic phenomena. In some systems, like models of ferromagnetic materials, hysteresis appears as an even more complex phenomenon. In many cases, however, a hysteretic behaviour hints to the existence of a bifurcation structure such as the cusp, and knowledge of this can help in decoding the system dynamics.

The fold bifurcation

The saddle-node bifurcation of continuous-time equilibria has a close relative in the fold bifurcation of fixed points.

Along the lines of what we before, consider a fixed point $\bar{x}$ of a discrete-time system $x(t+1) = f(x, p)$.

Here a note on terminology is due. The fold and the saddle-node are so tightly linked that often the two names are used interchangeably. Here, we are using the name saddle-node to refer to the bifurcations of equilibria and fold to refer to the bifurcation of fixed points and limit cycles.

Definition: Fold bifurcation
A fold of $f(x, p)$ at the fixed point $\bar{x}$ occurs at parameter values where the Jacobian $J_f(\bar{x}, p)$ has a single 1 eigenvalue.

The corresponding centre manifold is therefore 1-dimensional. Assume, without loss of generality, that the fold occurs at $\bar{x} = 0$, $p = 0$, and call $f_1(x_1, p)$, with $x_1 \in \mathbb{R}$, $p \in \mathbb{R}$, the 1-dimensional dynamics in the centre manifold. We have the following.

Theorem: the normal form of a generic fold
Assume that $\frac{\partial^2}{\partial x_1^2} f_1(0, 0) \neq 0$ and $\frac{\partial}{\partial p} f_1(0, 0) \neq 0$. The flow in the centre manifold is locally topologically equivalent to the flow of

$$x_1(t+1) = p + x_1(t) \pm x_1^2(t).$$

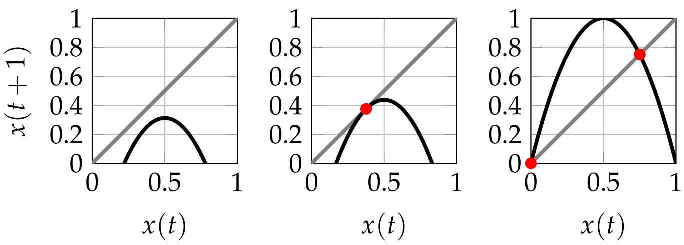
We have seen before that any continuous-time system induces, in the neighbourhood of a limit cycle, a discrete-time system with an equilibrium corresponding to the limit cycle. The above theorem gives us conditions on the Poincaré map of the limit cycle: the cycle undergoes a generic fold bifurcation whenever one of its nontrivial Floquet multipliers crosses the unit circle at 1.

Example 66. *Let us consider a translated logistic map*

$$x(t+1) = 4x(t)(1 - x(t)) + p.$$

For $p < -\frac{9}{16}$, the system has no fixed points, while for $p > -\frac{9}{16}$ it has two fixed points, in

$$x = \frac{3 \pm \sqrt{9 + 16p}}{8}.$$



At the bifurcation point $p = -\frac{9}{16}$ the equilibrium $x = \frac{3}{8}$ has Jacobian

$$J_f\left(\frac{3}{8}\right) = 4 - 8\frac{3}{8} = 1,$$

with

$$\frac{\partial^2}{\partial x^2} f = -8 \neq 0,$$

and

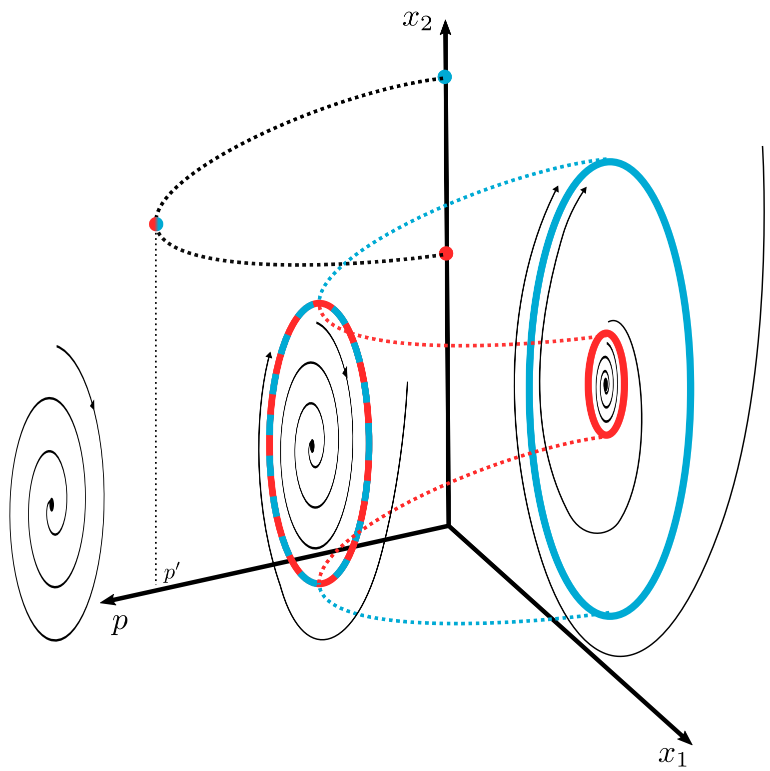
$$\frac{\partial}{\partial p} f = 1 \neq 0.$$

The system has a generic fold bifurcation at $p = -\frac{9}{16}$, and near $p = -\frac{9}{16}$ its flow is locally topologically equivalent to that of

$$x_1(t+1) = p + x_1(t) - x_1^2(t).$$

Notice that the eigenvalue of the rightmost equilibrium, which is initially $\simeq 1$, crosses through -1 as p increases. This is a second bifurcation, called period-doubling. We will see it later in this chapter.

Example 67. *Let us try to imagine how the fold bifurcation appears when it happens to the fixed points of the Poincaré map of a cycle. For some parameter value $p < p'$, a stable limit cycle (blue) and an unstable limit cycle (red) coexist. As $p \to p'$, the corresponding fixed points on the Poincaré map approach each other quadratically, and so do the cycles. At p' the two cycles collide and disappear as $p > p'$.*



The Hopf bifurcation

Consider a system $f(x, p)$, having an equilibrium $\bar{x}$ with two complex conjugate eigenvalues $a(p) \pm ib(p)$.

■ **Definition: Hopf bifurcation**
A Hopf of $f(x, p)$ at the equilibrium point $\bar{x}$ occurs at parameter values where the Jacobian $J_f(\bar{x}, p)$ has two purely imaginary eigenvalues.

Assume, without loss of generality, that the Hopf occurs at $\bar{x} = 0$, $p = 0$.

♦ **Theorem: the normal form of a generic Hopf**
Assume that, when $p = 0$, $l_1(p) \neq 0$ and $\frac{\partial}{\partial p} a(p) \neq 0$. The flow in the centre manifold is locally topologically equivalent to the flow of

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} \beta & -1 \\ 1 & \beta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \sigma(x_1^2 + x_2^2) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

where $\beta := \frac{a(p)}{b(p)}$, and $\sigma = \operatorname{sgn}(l_1(0)) = \pm 1$.

The function $l_1(p)$, which appears in the theorem, is a function of the system's parameters known as the *first Lyapunov coefficient*. After writing the planar system as a complex system $\dot{z} = \lambda z + g(z, z^*, p), z \in \mathbb{C}$, where z^* is the complex conjugate of z, we have

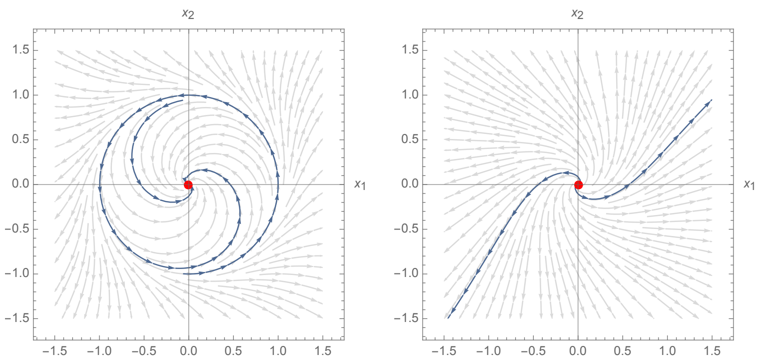
$$l_1(0) := \frac{1}{2\omega_0^2} \Re\left(i g_{20} g_{11} + \omega_0 g_{21}\right)$$

where $\omega_0 = \Im(\lambda)$ and g_{ij} are the coefficients of the terms $\frac{g_{ij}}{i!j!} z^i (z^*)^j$ in the Taylor expansion of g. See (Kuznetsov, 2004) for more details.

Let us see how this normal form behaves near $p = 0$. First, notice that the eigenvalues of the linearisation of the above normal form are

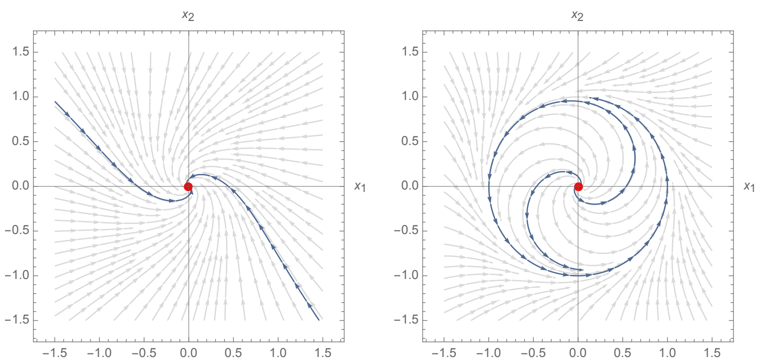
$$\beta \pm i,$$

so they cross the imaginary axis when those of J_f do. If we now take $\sigma = +1$, and let β change between -1 and $+1$, we obtain the following phase portrait



For $\beta = -1$, the stable focus is surrounded by an unstable limit cycle. This shrinks around the focus, with a radius equal to $\sqrt{-\beta}$, leaving an unstable focus for $\beta > 0$.

If instead we take $\sigma = -1$ we have the following portrait.



Now the limit cycle, with radius $\sqrt{\beta}$, exists for $\beta > 0$, and it is stable. The first case ($\sigma = +1$) is called the *subcritical* Hopf bifurcation. The second case is the *supercritical* Hopf bifurcation.

The value of sigma is determined by the sign of $l_1(p)$:

$$\sigma := \operatorname{sgn}(l_1(p)).$$

Example 68 (Catastrophic bifurcation)**.** *There is a significant difference in the behaviour we can observe from a system undergoing a subcritical or a supercritical Hopf bifurcation, as a parameter is slowly changed.*

Imagine to be observing a system $f(x, p)$ whose state has settled on one of its stable equilibria, and that this equilibrium, due to a slow change of a parameter, approaches a Hopf bifurcation. If this bifurcation is supercritical, we should expect to see the asymptotic behaviour of the system turning from a steady state to a small oscillation (the stable limit cycle). Should the parameter then reverse its drift, the state will go back to a steady value.

If, on the contrary, the bifurcation is subcritical, after the bifurcation no attractor is left in a neighbourhood of the equilibrium. The state will likely drift away to some other region of the phase space, possibly far away, in what is sometimes termed a catastrophe, that is, a sudden large change in the state due to an infinitesimal change in a parameter. Restoring the original value of the parameter will not, in general, bring the state back to its original value.

The period-doubling bifurcation

Having seen what happens in discrete-time systems as a single eigenvalue crosses the unit circle at 1 (the fold bifurcation), we now investigate the effect of a single eigenvalue crossing the unit circle at -1.

■ **Definition: Period-doubling bifurcation**
A period-doubling of $f(x, p)$ at the fixed point $\bar{x}$ occurs at parameter values where the Jacobian $J_f(\bar{x}, p)$ has an eigenvalue equal to -1.

Assume, without loss of generality, that the period-doubling occurs at $\bar{x} = 0$, $p = 0$, and let $f_1(x_1, p)$, with $x_1 \in \mathbb{R}$, $p \in \mathbb{R}$, be the 1-dimensional dynamics in the centre manifold. We have the following.

♦ **Theorem: the normal form of a generic period-doubling**
Assume that

$$\frac{1}{2}\left(\frac{\partial^2}{\partial x_1^2} f_1(0,0)\right)^2 + \frac{1}{3}\frac{\partial^3}{\partial x_1^3} f_1(0,0) \neq 0$$

and

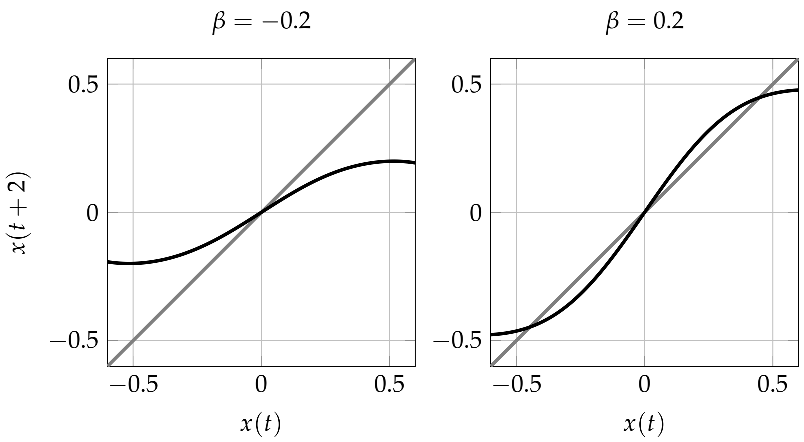
$$\frac{\partial^2}{\partial x_1 \partial p} f_1(0,0) \neq 0.$$

The flow in the centre manifold is locally topologically equivalent to the flow of

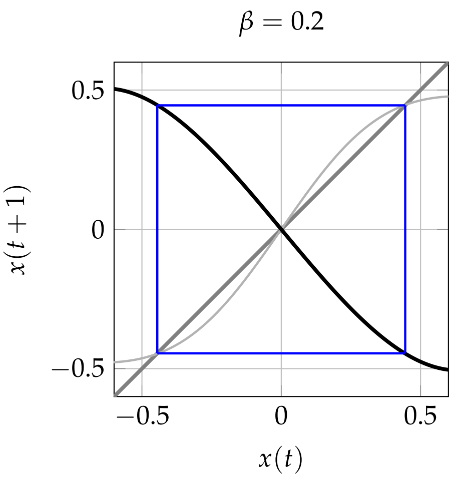
$$x_1(t+1) = -(1+\beta)x_1 + c x_1^3,$$

with $\beta(p) = 0$ for $p = 0$, $\text{sgn}(\beta) = \text{sgn}(p)$, and $c = \pm 1$.

Let us see how this map behaves near $p = 0$ with $c = +1$. We easily see that the map has a fixed point at $x_1 = 0$, for all values of p, and that this point becomes unstable when $p > 0$, while it is stable for $p < 0$. Let us now look at a plot of the second iterate of the map:



The map has a single fixed point for $\beta < 0$, which corresponds to the fixed point of the first iterate. However, for $\beta > 0$, two more fixed points appear, symmetrically around $x = 0$. We can understand the behaviour at these points by looking at an iteration of the first iterate map (the second iterate is reported in grey for reference).

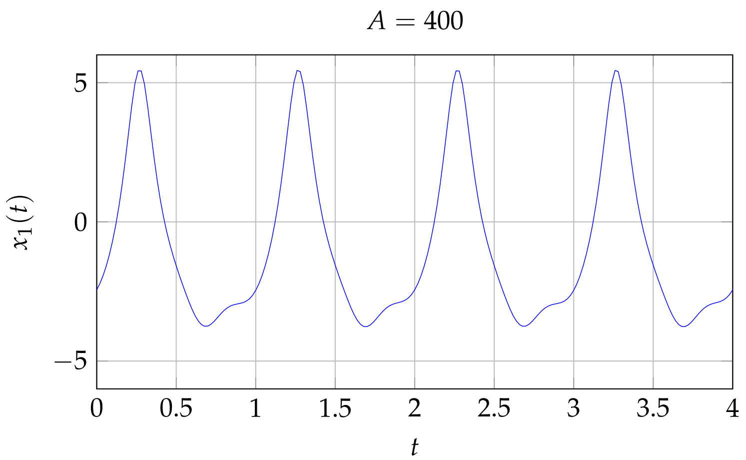


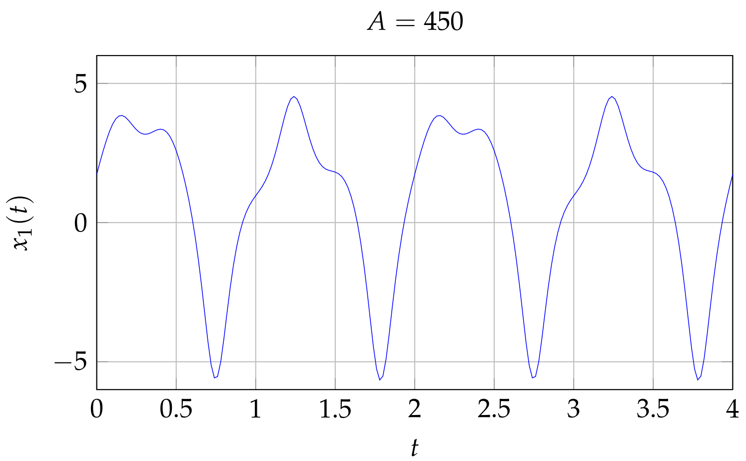
We would have found a similar scenario if we considered the case $c = -1$, with the only difference that the period-2 fixed point would have been unstable, coexisting with the stable period-1 fixed point. The period-doubling bifurcation is of course found also when a multiplier of a limit cycle crosses the unit circle at -1. If the bifurcating cycle is a stable one, at the bifurcation we expect the cycle to change stability, while it collides with a stable or unstable cycle twice its period. The stability of the double-period cycle depends on the sign of parameter c in the normal form.

Example 69 (Mass-spring-damper with nonlinear spring)**.** *Let us look once again at the periodically forced mass-spring-damper model that we saw* *at the beginning of the course.*

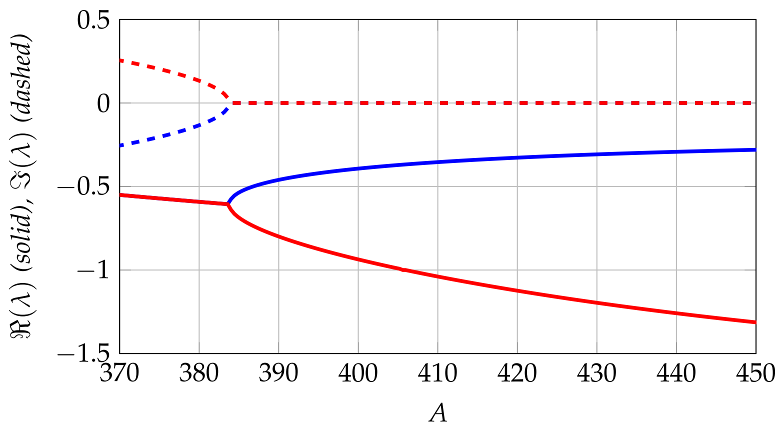
$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1(1 + 10x_1^2) - x_2 + A\sin(2\pi t).$$

We had already observed how, for A going from 400 *to* 450*, the attractor exhibits a subharmonic of the forcing function.*



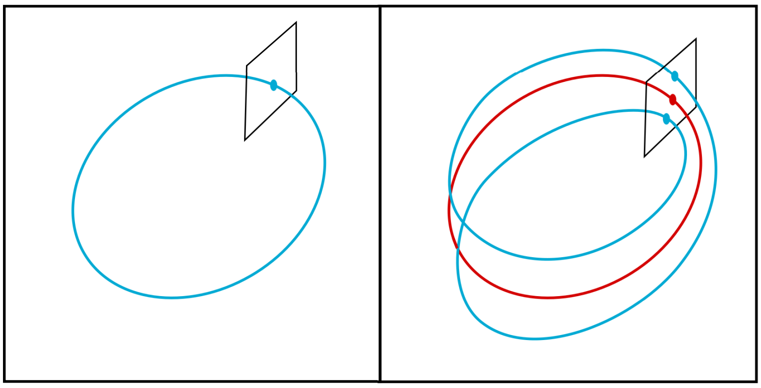


The following plot shows the nontrivial multipliers of the stable limit cycle found at $A = 400$, as a function of A.



We can see that, at $A = 405.93$, one of the multipliers crosses the unit circle at -1, triggering a period-doubling bifurcation. The limit cycle is unstable *for larger values of A: the object that we observed in the simulation above for $A = 450$ is not this cycle, but the double-period one that is born at the bifurcation.*

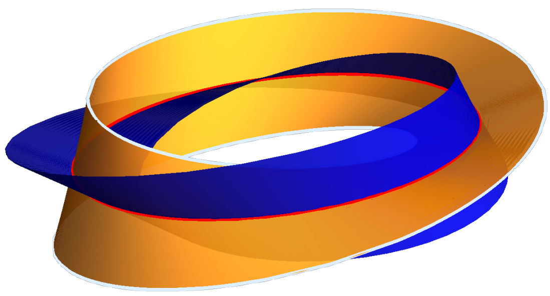
Example 70 (period-doubling in 3D)**.** *Let us see what a period-doubling of a limit cycle looks like in a 3-dimensional continuous-time time-invariant system. In the Poincaré map a fixed point, which we may assume stable, becomes a saddle as a period-2 discrete-time stable cycle appears. Each one of these objects corresponds to a limit cycle in the continuous-time system:*



The unstable manifold of the saddle cycle in this example is a Möbius strip, can you explain why?

Before the bifurcation, we have the stable cycle (blue, on the left). After the bifurcation, the cycle has become saddle (red), while a stable cycle twice the original length is born.

After the period-doubling, stable and unstable manifolds of the saddle cycle appear as follows (saddle cycle in red, stable cycle in light blue, stable and unstable manifolds of the saddle in blue and orange, respectively)



The Neimark-Sacker bifurcation

We already discussed what happens when, in a map, the eigenvalues of a fixed point cross the unit circle at $+1$ (the fold) or at -1 (the period-doubling). The obvious next step is to see what happens when a pair of eigenvalues crosses the unit circle anywhere except at these two points.

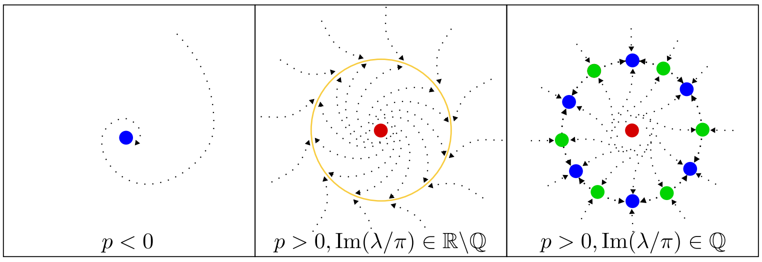
This is the last bifurcation scenario that we will see in this course and belongs by right of its codimension to the set of the most common codimension-1 bifurcations that we have seen so far. However, the complexity of its analysis and bifurcation diagram puts it on a completely different level, so much so that a normal form, in its precise acceptation, is not known, because no known change of variables can cancel the higher-order terms, and even a careful study of the truncated normal form is beyond the scope of our discussion.

Rather than going into the details of the dynamics near the Neimark-Sacker bifurcation, we will here give its definition, and discuss some of the more apparent qualitative features that are typically found near this bifurcation.

■ **Definition: Neimark-Sacker bifurcation**
A Neimark-Sacker of $f(x, p)$ at the fixed point $\bar{x}$ occurs at parameter values where the Jacobian $J_f(\bar{x}, p)$ has two eigenvalues on the unit circle, not in ± 1.

Just like at a Hopf bifurcation, a periodic orbit bifurcates from the equilibrium, at a Neimark-Sacker bifurcation an invariant circle springs out of the fixed point.

Let us consider the *supercritical* scenario, where a stable fixed point transforms into an unstable fixed point plus a stable invariant circle. The *subcritical* scenario is obtained as usual by swapping stability. Let p be the bifurcation parameter and $p = 0$ the bifurcation point, and assume that we have the stable fixed point for $p < 0$ with two eigenvalues approaching the unit circle as $p \to 0$, as in the left panel in the figure below.



As p crosses 0, the fixed point becomes unstable and, if a set of genericity conditions are satisfied, the stable invariant circle emerges around the fixed point.

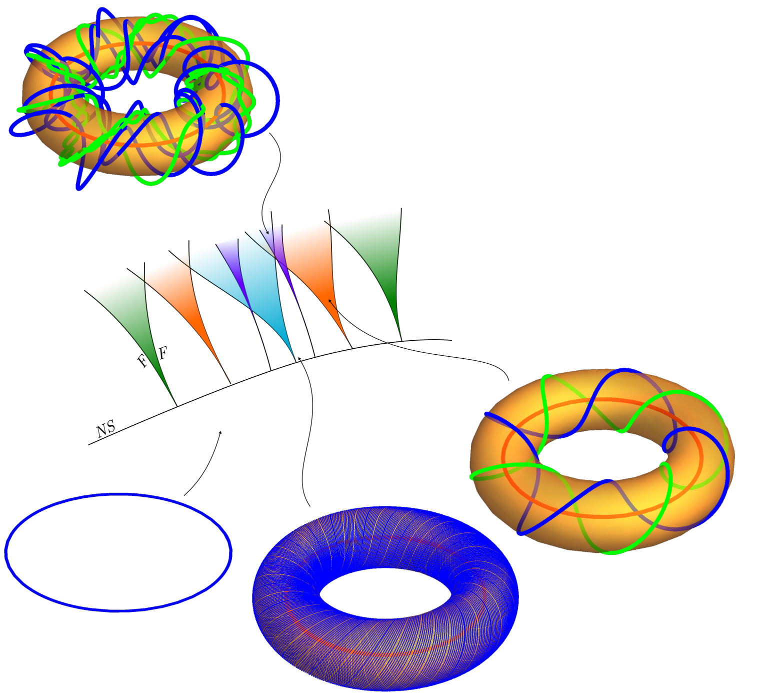
The dynamics within this circle is decided by the imaginary part of the eigenvalues at $p = 0$. If it is an irrational multiple of π, as in the centre panel in the figure below, then the circle (yellow in the figure) is densely covered by an infinitely long orbit. If instead, it is a rational multiple of π, as in the right panel in the figure, then the circle contains a stable and an unstable periodic orbit (the blue and green sets of points), and the rest of the circle consists of the unstable manifold of the unstable periodic orbit.

Now let us imagine how the above picture translates if the discrete-time system is the Poincaré map of a continuous-time system around a cycle. Then, each fixed point becomes a cycle, while the invariant circle corresponds to an invariant *torus* densely covered by a quasiperiodic orbit or by the unstable manifold of the unstable limit cycle.

The dynamics past the bifurcation is however much more complex than one may think from the simplistic description above. Let us get a glance at this complexity while remaining in the continuous-time framework.

We have seen how a pair of limit cycles exist on the torus near bifurcation points where eigenvalues have rational imaginary parts. If we assume, as is usually the case, that the eigenvalues change continuously along the Neimark-Sacker bifurcation, we should expect to encounter infinitely many points along the bifurcation curve where such a condition is satisfied, one for each rational number along the real line. These cycles typically persist beyond the Neimark-Sacker, and exist within a region shaped like a thin wedge, bounded on two sides by two-fold bifurcations.

As rational numbers are dense along the real line, these wedges, known as *Arnold tongues*, are typically dense and overlap. Beyond the Neimark-Sacker, therefore, there are typically multiple coexisting stable and unstable cycles, that interact in extremely complex ways through multiple further bifurcations and may, and typically do, eventually lead to the birth of a chaotic attractor.



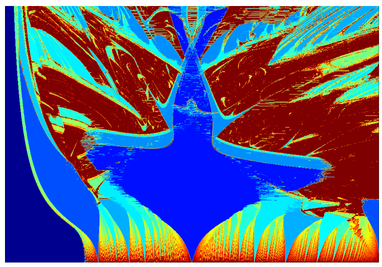


Figure 5: A 2-parameter bifurcation diagram from (Colombo and Rinaldi, 2008). At the lower end of the diagram, we can clearly see a set of Arnold tongues, emerging from the horizontal axis. In the upper part of the diagram, the dynamics is chaotic. For the sake of completeness, we should note that, even though this figure displays the typical characteristics of a bifurcation diagram near a Neimark-Sacker, it comes from a *nonsmooth* model.

In the picture above, for example, the stable blue cycle (lower left) has a Neimark-Sacker (NS) bifurcation.

Shortly after the bifurcation the cycle has changed stability (it is now red) and is surrounded by an invariant torus (yellow) densely covered by a stable orbit (blue). Within an Arnold Tongue (the orange one, for example), the invariant torus still exists, but is wrapped within a stable (blue) and unstable (green) cycle. The torus itself consists of the unstable manifold of the green unstable cycle. At a point where multiple Arnold tongues overlap (top left) there is no longer any invariant torus (you can see orbits crossing through the former invariant torus), and multiple cycles coexist, typically within a chaotic attractor.

Numerical continuation

The normal forms that we have seen in the sections above give us a very precise, yet general picture of the unfolding of all the most common bifurcation, that is, of the phase portrait of a system near the bifurcating equilibrium, fixed point, or limit cycle, for parameters close to the bifurcation parameters. They would however be of limited use in practice, without efficient computational tools to find the bifurcations and analyse the normal form conditions in real-world models. Luckily, such tools exist. The main ingredient in most cases is a procedure to follow, as parameters are changed, a given geometric object (equilibrium, fixed point, cycle), with given additional conditions (e.g. an equilibrium *with one eigenvalue equal to 0*): this is called *numerical continuation*.

To understand what it is, let us consider the simple case of the 2-parameter unfolding of a saddle-node.

Example 71 (Numerical continuation of a saddle-node bifurcation)**.** *Consider an arbitrary system*

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

The equations

$$\begin{aligned} f(x,p) &= 0, \\ \det J_f(x) &= 0, \end{aligned}$$

define the saddle-node bifurcation set, in $\mathbb{R}^{n+2}$. *Assume that* (x_0, p_0) *is one such bifurcation points, and let us rewrite the above system as*

$$F(y) = 0, \quad y := (x, p), \quad y_0 := (x_0, p_0).$$

The Jacobian $J_F(x_0, p_0)$ *has* $n+1$ *rows (*n *from* f, 1 *from* $\det J_f$*) and* $n+2$ *columns, therefore there exists a vector* v *such that*

$$J_F(x_0, p_0)v_0 = 0.$$

This vector is tangent to the bifurcation set at (x_0, p_0), *and is typically unique up to rescaling (in fact, it is unique as long as the conditions for the normal form of a generic saddle-node are satisfied). We can continue the bifurcation set by taking our next guess*

$$\tilde{y}_1 := y_0 + v_0.$$

If we chose the norm $\|v_0\|$ *sufficiently small,* $\tilde{y}_1$ *should be close to a root of* $F(y) = 0$. *To find the correct value of* y_1 *we can use Newton's method, which however requires* $n+2$ *equations for a system of* $n+2$ *unknowns. The additional equation, which we must add, defines the surface on which*

Newton's method is an iterative method to solve $F(y) = 0$, starting from an initial guess $y(0)$, by iterating

$$y(k+1) = y(k) - J_F^{-1}(y(k))F(y(k)).$$

we want our correction to be. A common choice is the pseudo-arclength continuation, which constrains the Newton correction to be in a plane that is orthogonal to v_0. We therefore seek a zero in the variable y_1 of

$$F(y_1) = 0,$$
$$(y_0 + v_0 - y_1)^\top v_0 = 0.$$

By repeating this procedure, we can construct the saddle-node bifurcation set point by point.

In the above example, we have sketched the simplest implementation of a numerical continuation scheme to reconstruct the saddle-node bifurcation set. In practice, for improved numerical efficiency and stability, slightly different schemes are typically used. The strategy is however the same. Also, it is worth noticing that the steps we have followed are not specific to the continuation of the saddle-node. I we substitute $F(y)$ with a system of equations defining a hyperbolic equilibrium, or an equilibrium undergoing a different type of bifurcation, we can follow the exact same steps to follow that object through the parameter space.

■ **Definition: Numerical continuation**
Given a system

$$F(y) = 0, F : \mathbb{R}^{n+1} \to \mathbb{R}^n,$$

the set

$$y : F(y) = 0$$

is one-dimensional, and can be numerically computed in $\mathbb{R}^{n+1}$ by iteratively solving through Newton's method in the variable y_{k+1} the system

$$F(y_{k+1}) = 0,$$
$$((y_k + v_k) - y_{k+1})^\top v_k = 0,$$

where v_k is a vector such that

$$J_F(y_k)v_k = 0, \quad v_k \neq 0,$$

and the initial guess for y_{k+1} is $y_{k+1} = y_k + v_k$.

The above strategy is typically used to plot 2-parameter bifurcation diagrams, as the projection in the (p_1, p_2)-plane of the set $\{x, p : F(x, p) = 0\}$.

Exercises

Exercise 59

The following logistic model represents the dynamics of an animal species with harvesting

$$\dot{x} = 2x\left(1 - \frac{x}{2}\right) - p$$

as a function of the harvesting effort p. Here, p represents harvesting of the species x, and the model only makes sense for $x > 0$, since once $x = 0$ constant harvesting is no longer meaningful.

Study stability and bifurcations of all equilibria.

Exercise 60

Explain why a limit cycle in a planar continuously differentiable system cannot have a multiplier equal to -1.

Exercise 61

Identify all codimension-1 bifurcations of equilibria in the system

$$\begin{aligned}\dot{x}_1 &= x_1^2 + x_2 - x_1 + d,\\ \dot{x}_2 &= -x_1 - 4x_2,\end{aligned}$$

when parameter d changes in $\mathbb{R}$. (You don't need to check if the bifurcations are degenerate)

Answer of exercise 61

We can start by identifying the equilibria, for instance using nullclines.

The x_1 nullcline is the parabola $x_2 = x_1 - x_1^2 + d$, while the x_2 nullcline is the straight line $4x_2 = -x_1$. Since d shifts the parabola up and down, and the parabola is facing down, we can expect two equilibria when d is sufficiently large, disappearing through a saddle-node as d decreases.

For $d = 0$ we have equilibria in $(0,0)$ and $(5/4, -5/16)$. To find the d-coordinate of the saddle-node, we can look for the parameter value where the two nullclines are tangent, which happens when the equation

$$\left(x_1^2 + x_2 - x_1 + d\right)|_{x_2=-x_1/4} = x_1^2 - \frac{5}{4}x_1 + d = 0$$

has a single solution, that is, when

$$\left(\frac{5}{4}\right)^2 - 4d = 0.$$

This is at

$$d = \frac{25}{64}.$$

Next, we should check whether the two equilibria can have a Hopf bifurcation. The Jacobian at an arbitrary equilibrium is

$$J_f(x) = \begin{pmatrix} 2x_1 - 1 & 1 \\ -1 & -4 \end{pmatrix}.$$

At a Hopf bifurcation we must have $\mathrm{tr}(J_f(x)) = 0$ and $\det(J_f(x)) > 0$, that is,

$$\begin{aligned}2x_1 - 5 &= 0,\\ -8x_1 + 6 &> 0.\end{aligned}$$

The system does not have solutions, therefore there cannot be Hopf bifurcations in the system.

Exercise 62

Study the generic bifurcations of equilibria of the system

$$\dot{x}_1 = px_1x_2 + x_1^2$$
$$\dot{x}_2 = 3x_1x_2^3 + 1$$

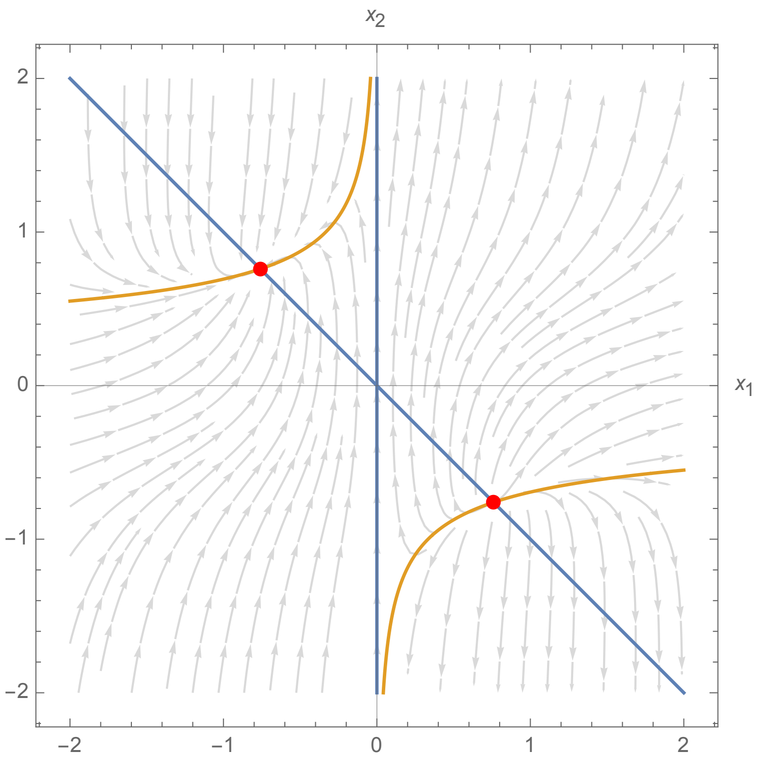
when $p > 0$, and tell if the system can have a limit cycle.

Answer of exercise 62

The equilibria solve the system

$$px_2 = -x_1$$
$$x_1 = -\frac{1}{3x_2^3},$$

therefore they are located at $x_1 = \mp 3^{-\frac{1}{4}} p^{\frac{3}{4}}$, $x_2 = \pm(3p)^{-\frac{1}{4}}$. We may begin studying their bifurcations by sketching the nullclines.



Notice that only the x_1 nullcline (blue) depends on p. The system has exactly two equilibria for all values of $p > 0$, therefore we can exclude the occurrence of a generic saddle-node. We could draw the same conclusion by checking the determinant of the Jacobian at the equilibrium. We have

$$J_f(\hat{x}) = \begin{pmatrix} px_2 + 2x_1 & px_1 \\ 3x_2^3 & 9x_1x_2^2 \end{pmatrix}\Bigg|_{x=\bar{x}} = \begin{pmatrix} \mp 3^{-\frac{1}{4}} p^{\frac{3}{4}} & \pm 3^{-\frac{1}{4}} p^{\frac{7}{4}} \\ \pm 3^{\frac{1}{4}} p^{-\frac{3}{4}} & \mp 3^{\frac{9}{4}} p^{\frac{1}{4}} \end{pmatrix},$$

therefore

$$\det J_f(\bar{x}) = 9p - p \neq 0, \forall\, p > 0.$$

The determinant is never null, therefore the equilibria never have a zero eigenvalue.

To investigate the possibility of a Hopf bifurcation of one of the two equilibria $\bar{x}$, we may check if there exists a value of p such that $J_f(\hat{x})$ has null trace and positive determinant. We have

$$\mathrm{tr} J_f(\bar{x}) = \mp 3^{-\frac{1}{4}} p^{\frac{3}{4}} \mp 3^{\frac{9}{4}} p^{\frac{1}{4}}.$$

The trace has a constant sign for $p > 0$ at both equilibria, therefore we may exclude the occurrence of Hopf bifurcations.

Finally, we may rule out the existence of periodic orbits (and therefore of limit cycles) using Dulac's criterion with $g(x) = \frac{1}{x_1}$:

$$\nabla \cdot g(x) f(x) = \nabla \cdot \begin{pmatrix} p x_2 + x1 \\ 3x_2^3 + \frac{1}{x_1} \end{pmatrix} = 1 + 9x_2^2 > 0, \forall x \in \mathbb{R}, \forall p \in \mathbb{R}.$$

Exercise 63

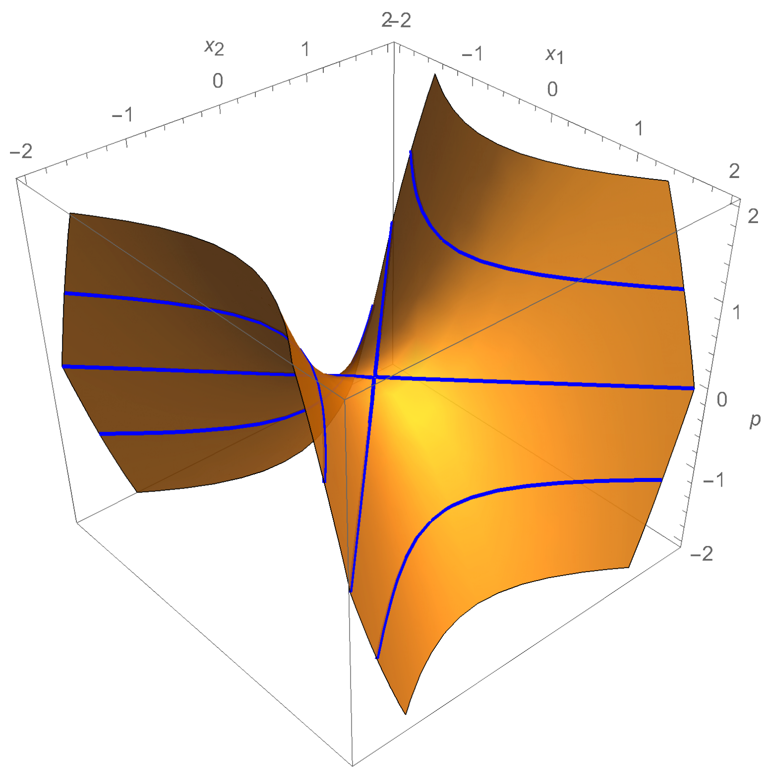
Identify all codimension-1 bifurcations of equilibria in the system

$$\dot{x}_1 = x_1 + 4x_2^2 - 1,$$
$$\dot{x}_2 = x_1^2 - x_2^2 + p,$$

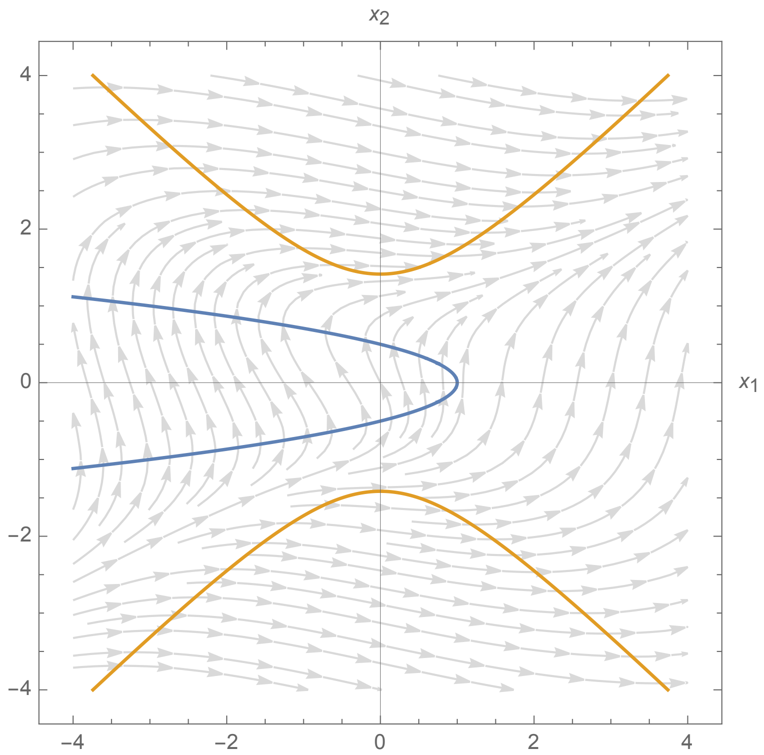
when parameter p changes in $\mathbb{R}$. (You don't need to check if the bifurcations are degenerate)

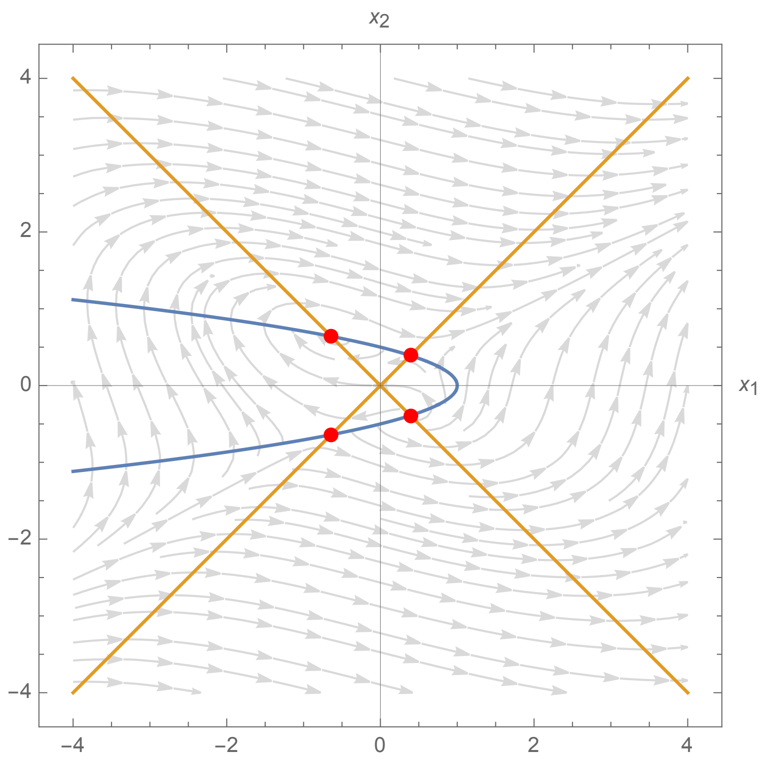
Answer of exercise 63

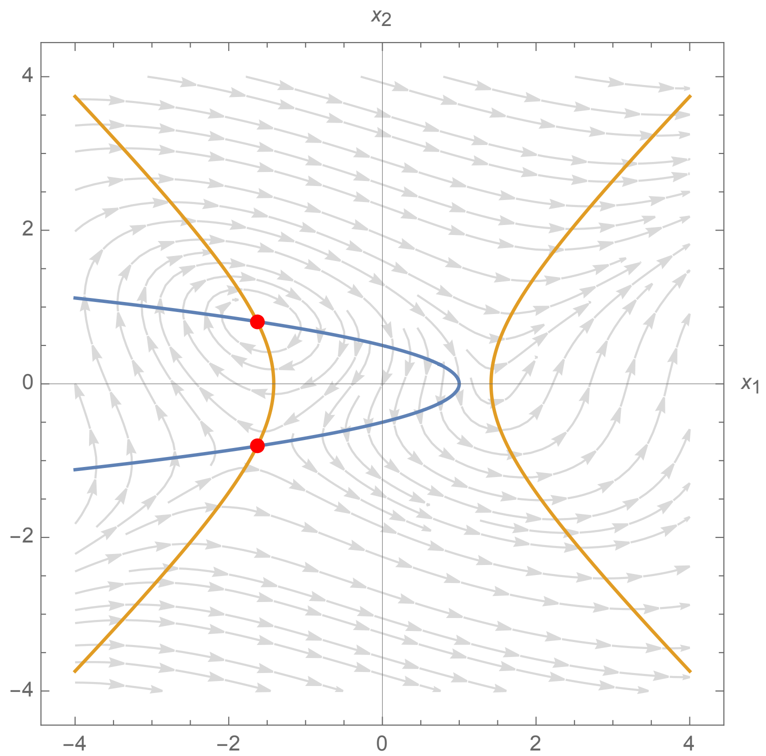
The x_1-nullcline is the parabola $x_1 = 1 - 4x_2^2$, and does not depend on p. To plot the x_2 nullcline we may notice that the function $p = x_2^2 - x_1^2$ defines a *saddle* surface:



For $p = 0$ the x_2 nullcline is composed of a pair of straight lines crossing the origin, while other values of p correspond to two parabolas with axis of symmetry aligned along the x_2 axis for $p > 0$, or the x_1 axis for $p < 0$. The phase portraits for $p = 2$, $p = 0$, and $p = -2$ are as follows







We should expect three saddle-node bifurcations and, possibly, Hopf bifurcations. The Jacobian is

$$J_f(x) = \begin{pmatrix} 1 & 8x_2 \\ 2x_1 & -2x_2 \end{pmatrix}.$$

A saddle-node occurs when $\det(J_f(x)) = 0$, that is

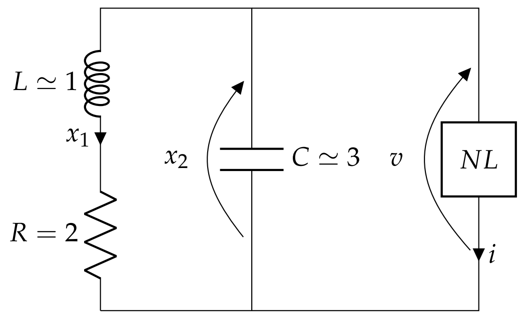
$$-2x_2 - 16x_1x_2 = 0.$$

The above systems has solutions $x_2 = 0$ or $x_1 = -\frac{1}{8}$. When $x_2 = 0$, solving $f(x,p) = 0$ we obtain $x_1 = 1$ and $p = -1$, therefore at this parameter value we have a saddle-node, corresponding to the two equilibria on the right colliding and disappearing. When $x_1 = -1/8$, solving $f(x,p) = 0$ we obtain $x_2 = \pm\frac{3}{4\sqrt{2}}$ and $p = \frac{17}{64}$. At this parameter value, we have two pairs of equilibria simultaneously undergoing saddle-node bifurcations.

A Hopf occurs when $\mathrm{tr}(J_f(x)) = 0$ and $\det(J_f(x)) > 0$. The first condition gives $1 - 2x_2 = 0$, that is, $x_2 = \frac{1}{2}$, and plugging this into $f(x,p) = 0$ we obtain $x_1 = 0$ and $p = \frac{1}{4}$. At these coordinates we have $\det(J_f(x)) = -2x_2 - 16x_1x_2 = -1 < 0$, therefore we do not have Hopf bifurcations.

Exercise 64

Consider the nonlinear circuit



which we saw some chapters ago, where the nonlinear element NL has characteristic

$$i = -v + v^3.$$

The circuit is modelled by the system

$$\dot{x}_1 = \frac{x_2 - Rx_1}{L},$$
$$\dot{x}_2 = \frac{-x_1 + x_2 - x_2^3}{C},$$

and is known to have no limit cycle for the nominal parameter values. While we have a very precise resistor $R = 2$, the capacitor and the inductor have a relatively large tolerance. Determine what are the attractors of the system for its nominal parameters, and what is the effect of the uncertainty in C and L on the attractors, in terms of bifurcations.

Exercise 65

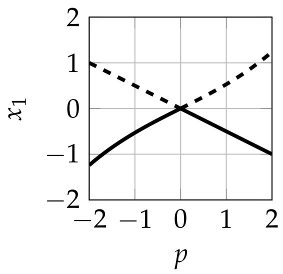
At a transcritical bifurcation, an equilibrium has an eigenvalue equal to 0, with

$$\frac{\partial^2}{\partial x_1^2} f_1(\bar{x}_1(p), p) \neq 0,$$

and

$$\frac{\partial}{\partial p} f_1(\bar{x}_1(p), p) = 0.$$

Perturbing p from the bifurcation value we obtain the following bifurcation diagram.

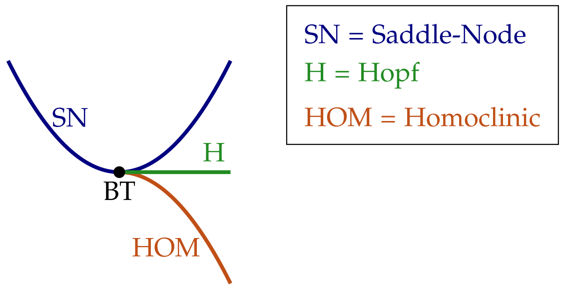


How will the bifurcation diagram change if the equations are slightly perturbed, causing

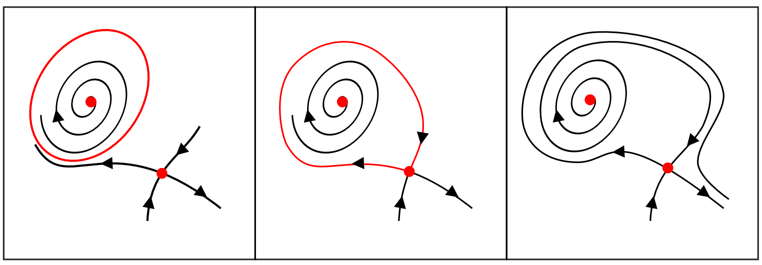
$$\frac{\partial}{\partial p} f_1(\bar{x}_1(p), p) \neq 0?$$

*** Exercise 66**

A Bogdanov-Takens is a codimension-2 bifurcation of continuous-time equilibria, where an equilibrium has two eigenvalues at 0. One of the possible generic 2-parameter bifurcation diagrams is the following



Even though the above diagram is generic for $n-$dimensional systems, let us consider it for a planar system. In that case, the *homoclinic* bifurcation (HOM) implies the structurally unstable connection of the stable and unstable manifolds of a saddle equilibrium, and the simultaneous collision of a limit cycle with the equilibrium, as in the following representation.

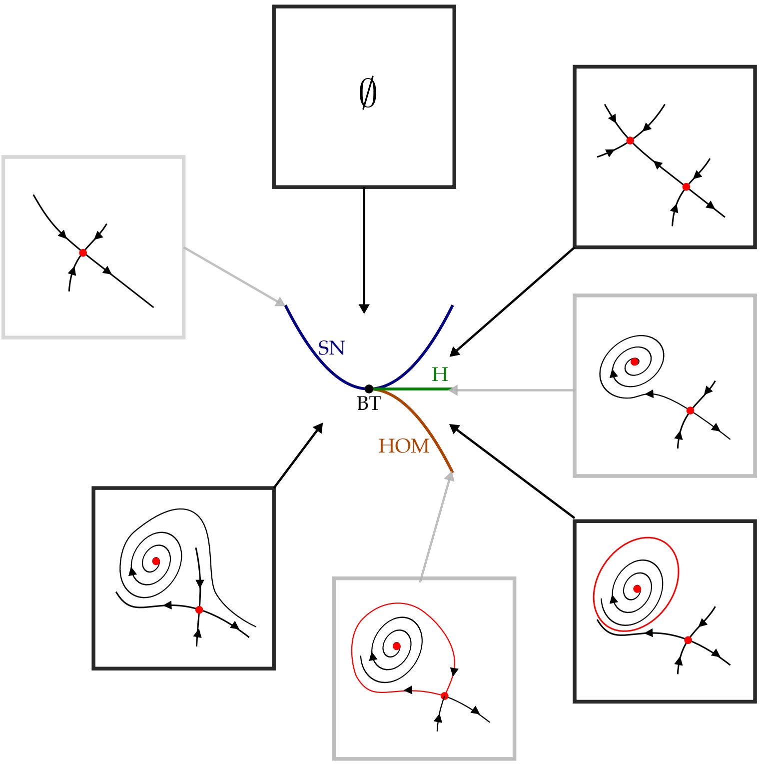


The planar case of the homoclinic bifurcation is described by the Andronov-Leontovich theorem, and can only give rise to scenarios similar to the one depicted here. In more than 2 dimensions, however, more complex scenarios exist, where even an infinite number of cycles (and possibly a chaotic attractor, known as Shilnikov chaos) may be generated at the bifurcation. See (Kuznetsov, 2004) for a discussion.

Using the above information, and everything you know about local codimension-1 bifurcations, sketch the phase portrait you expect to find in each of the four parameter regions surrounding the bifurcation, and on the three boundaries, highlighting the attractors, for a planar system.

Answer of exercise 66

Two equilibria should appear at the saddle-node, and exist below the curve, otherwise, the other bifurcation curves could not be there. Between the SN and the H, one of the two equilibria becomes a focus and undergoes a Hopf bifurcation at H. The cycle disappears through the homoclinic bifurcation.



Notice that at the BT point, the nonhyperbolic equilibrium has 2 null eigenvalues. Moving away from the BT, one of the two equilibria gains two real eigenvalues and becomes a saddle, setting the structure for the homoclinic. The other equilibrium gains a complex conjugate pair, setting the structure for the Hopf. At the BT, the local phase portrait shows on one side the two halves of the stable and unstable manifold of the saddle, and on the other side the circular orbits of the focus.

Bibliography

Arnold, V. I. (1992). *Ordinary differential equations*. Springer-Verlag.

– (1997). *On teaching mathematics*. Talk given at Palais de la Découverte in Paris on 7 March.

Aubin, J. (1991). *Viability Theory*. Birkhauser.

Blanchini, F. and S. Miani (2008). *Set-Theoretic Methods in Control*. Birkhauser.

Colombo, A. and S. Rinaldi (2008). "Chaos in two-party democracies". In: *Int. J. Bifurcat. Chaos* 18, pp. 2133–2140.

Dieci, L. and E. S. V. Vleck (2002). "Lyapunov spectral intervals: Theory and computation". In: *SIAM J. Numer. Anal.* 40, pp. 516–542.

Elowitz, M. B. and S. Leibler (2000). "A synthetic oscillatory network of transcriptional regulators". In: *Nature* 403.6767, pp. 335–338.

Fibonacci (1202). *Liber Abaci*.

Fuss, F. K. and G. Niegl (2008). "Instrumented climbing holds and performance analysis in sport climbing". In: *Sports Technology* 1.6, pp. 301–313.

Geist, K., U. Parlitz, and W. Lauterborn (1990). "Comparison of Different Methods for Computing Lyapunov Exponents". In: *Progress of Theoretical Physics* 83, pp. 875–893.

Isidori, A. (1995). *Nonlinear Control Systems*. Springer London.

Kermack, W. O. and A. G. McKendrick (1927). "A contribution to the mathematical theory of epidemics". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115.772, pp. 700–721.

Khalil, H. K. (2002). *Nonlinear systems*. Prentice-Hall.

Kot, M. (2005). "Torus bubbling in a discrete-time predator–prey model". In: *Journal of Difference Equations and Applications* 11.4-5, pp. 431–441.

Kuznetsov, Yu. A. (2004). *Elements of Applied Bifurcation Theory*. 3rd. Berlin: Springer-Verlag.

Lotka, A. J. (1920). "Undamped oscillations derived from the law of mass action". In: *J. Am. Chem. Soc.* 42, pp. 1595–1599.

Malthus, T. R. (1798). *An Essay on the Principle of Population*. J. Johnson.

Meiss, J. D. (2007). *Differential Dynamical Systems*. SIAM.

Melillo, P., M. Bracale, and L. Pecchia (2011). "Nonlinear Heart Rate Variability features for real-life stress detection. Case study: students under stress due to university examination". In: *BioMedical Engineering OnLine* 10.1, p. 96.

Mestres, P. and J. Cortes (2022). "Safe Policy Design for Controlling Epidemic Spreading under Heterogeneous Testing Capabilities". In: *2022 American Control Conference (ACC)*. IEEE, pp. 697–702.

Milnor, J. (1985). "On the concept of attractor". In: *Commun. Math. Phys.* 99, pp. 177–195.

Murdock, J. (2003). *Normal forms and unfoldings for local dynamical systems*. Springer-Verlag.

Plastock, R. (1974). "Homeomorphisms between Banach spaces". In: *Transactions of the American Mathematical Society* 200, pp. 169–183.

Rosenzweig, M. L. and R. H. MacArthur (1963). "Graphical representation and stability conditions of predator-prey interactions". In: *Am. Nat.* 97, pp. 209–223.

Schubert, C. et al. (2009). "Effects of stress on heart rate complexity—A comparison between short-term and chronic stress". In: *Biological Psychology* 80.3, pp. 325–332.

Scott, S. K. (1991). *Chemical Chaos*. Oxford University Press.

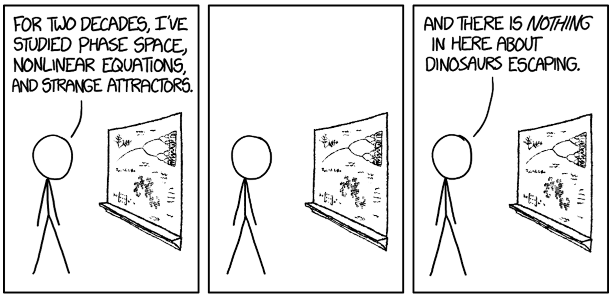
Spong, M. W., P. Corke, and R. Lozano (2001). "Nonlinear control of the Reaction Wheel Pendulum". In: *Automatica* 37.11, pp. 1845–1851.

Strogatz, S. H. (1994). *Nonlinear Dynamics and Chaos*. Reading, MA: Addison-Wesley.

Verhulst, P. (1845). "Recherches mathématiques sur la loi d'accroissement de la population." In: *Nouveaux mémoires de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles* 18, pp. 14–54.

Volterra, V. (1926). "Variazioni e fluttuazioni del numero d'individui in specie animali conviventi". In: *Mem. Accad. Naz. Lincei* 2, pp. 31–113.

Wiggins, S. (2003). *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. 2nd. New York: Springer-Verlag.



thanks, xkcd.com :)