Notes for a Course on
Numerics of Dynamical Systems

W. Rheinboldt

These notes are in support of the course "Zur Numerik dynamischer Systeme" (TU Muenchen, WS 2008). They were compiled from parts of lecture notes prepared for several other courses taught in earlier years. Some explanatory material has been added to bridge larger gaps, but otherwise, only minor editorial changes were made. Accordingly, there are overlaps in material and discrepancies in notation, and the assumed prerequisites are not always identical.
Chapter 1

Continuous Model Problems

A "dynamical system" is, in essence, a mathematical model of the time-evolution of some process, such as, for example,

- the motion of a pendulum under the influence of gravity,
- the motion of the planets around the sun,
- flow of a fluid or gas in some pipe,
- the variation of populations under given conditions,
- the behavior of an economic system with time,
- the evolution of a climatic system.

The topic of dynamical systems is very extensive and has a huge literature. The aim here will be to develop a basic understanding of the area with emphasis on numerical aspects.

A basic group of dynamical systems is modelled by an initial value problem of the form

\[ \dot{x} = G(x), \quad x \in \mathbb{R}^n, \quad x(0) = x^0, \]  \tag{1.1}

where \( G \) is a sufficiently smooth function on \( \mathbb{R}^n \). Many of the above mentioned problems have this form. But there is reason not to restrict the attention to problems with a continuous time variable. Indeed, it is often more natural to measure time in discrete steps, and hence to consider dynamical systems in the form of recursions:

\[ x^{k+1} = G(x^k), \quad x^k \in \mathbb{R}^n, \quad k = 0, 1, 2, \ldots \]  \tag{1.2}

We begin with a simple example of a continuous time problem.
1.1 The Planar Pendulum

A venerable example of a dynamical system of the form (1.1) is the planar pendulum.

A story goes that Galileo Galilei (1564 - 1642), as a young man, sat in the cathedral of Pisa and watched an overhead candelabra swing at the end of a long chain suspended from the high ceiling. He appears not have been attentive to the service, but wondered

• what is the period of the oscillation; i.e., how fast does the lamp swing,
• how does this period depend on the length of the chain
• how does the period depend on the weight of the lamp.

In his later career Galilei investigated many phenomena related to the pull of gravity, including, besides the pendulum, the free fall of a body, the trajectory of cannon balls, and the timing of bodies rolling down an inclined plane. The mathematical tools available at the time of Galilei were still inadequate for problems of this type. But his superb experimental skills led him to profound insights, such as, for instance, the observation, that in all the indicated problems, the velocity is independent of the mass of the body.

We consider a mathematical model of a planar pendulum, as sketched in Figure 1.1.

![Planar Pendulum Diagram](image)

The arclength (displacement) between the current position and the rest position ($\theta = 0$) equals $s = \ell \theta$. Let $m$ be the mass of the bob and $F$ the force acting on it, then Newtons second law provides that

$$m \ddot{s} = m \ell \ddot{\theta} = F.$$
With the acceleration of gravity $g$, the force of gravity equals $mg$. As sketched, this force can be decomposed into a tangential component $-mg \sin \theta$ pointing in the direction opposite to $\text{sign}(\theta)$, and a component $mg \cos \theta$ in the direction of the rigid, massless ‘support-string’. This component does not contribute to the motion. Thus, we have

$$m \ell \ddot{\theta} = -mg \sin \theta$$

As Galilei observed, the mass $m$ cancels and we obtain the differential equation

$$\frac{d^2 \theta}{dt^2} + \frac{g}{l} \sin \theta = 0.$$  \hspace{1cm} (1.3)

Obviously, our pendulum model is idealized: in particular, we assumed (i) a point mass, (ii) a fixed geometry, (iii) a massless, rigid ‘string’, and most importantly, (iv) no friction.

In general, for dynamical systems Willard Gibbs introduced in 1901 the concept of a phase space. This is the set of all possible states of the system. Thus, informally, we may characterize a dynamical system as a ”rule for state-transitions” in phase space. In many cases, the phase space is a subset of a finite dimensional linear space $\mathbb{R}^n$ and the components of the state $x \in \mathbb{R}^n$ are certain variables and/or parameters characterizing the system.

In the case of the pendulum, (1.3) shows, that at any time $t$ the state of the pendulum is determined by the vector $(\theta, \dot{\theta})$ consisting of the position angle $\theta$ and the velocity $\dot{\theta}$. Thus, for the pendulum the phase space is the plane $\mathbb{R}^2$. For an initial state $x^0 = (\theta_0, \dot{\theta}_0)^T \in \mathbb{R}^2$ at time $t = 0$, the motion $x = x(t) \in \mathbb{R}^2$ is the solution of the initial value problem (1.1):

$$\begin{align*}
\frac{dx_1}{dt} &= x_2, \\
\frac{dx_2}{dt} &= -\frac{g}{l} \sin x_1,
\end{align*}$$

$x(0) = x^0 = (\theta_0, \dot{\theta}_0)^T$. \hspace{1cm} (1.4)

For a trace of these solutions in phase space, we can use a simple Matlab code, such as, Algorithm 1.1, where we set $g/\ell = 1$.

**Listing 1.1: Plot solution of pendulum equation**

```matlab
function pendul(x0, T)
    [t, x] = ode45(@pend, [0 T], x0);
    plot(x(:,1), x(:,2));
    end

function yp = pend(t, y)
    yp = [y(2); -g/l * sin(y(1))];
    end
```

For different choices of the initial state $x^0$ we obtain a ‘phase-portrait’ of the form shown in Figure 1.2.
1.2 Properties of Initial-value Problems

A basic background of the theory of initial value problems for ODEs is assumed. The standard existence theory for initial value problems (1.1) works with locally Lipschitz continuous mappings. Recall that this includes all mappings

$$G : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad G(x) := (g_1(x), \ldots, g_n(x))^\top, \quad x = (x_1, \ldots, x_n)^\top,$$

which are continuously differentiable on an open set \(\Omega\); i.e., for which all first order partial derivatives of each component \(g_j\) exist and, together with \(g_j\), are continuous on \(\Omega\). As usual, we denote the first derivative of \(G\) by \(DG(x)\). In terms of the bases used in (1.5), \(DG(x)\) has the matrix representation

$$
\begin{pmatrix}
\frac{\partial}{\partial x_1} g_1(x) & \cdots & \frac{\partial}{\partial x_n} g_1(x) \\
\vdots & \ddots & \vdots \\
\frac{\partial}{\partial x_1} g_n(x) & \cdots & \frac{\partial}{\partial x_n} g_n(x)
\end{pmatrix}
$$

A mapping

$$x : \mathcal{J} \subset \mathbb{R}^1 \rightarrow \mathbb{R}^1, \quad 0 \in \mathcal{J}, \quad \mathcal{J} \text{ open interval}$$

is a solution of (1.1), if \(x\) is continuously differentiable on \(\mathcal{J}\) and satisfies the differential equation and the initial condition. A solution \(x\) on some interval \(\mathcal{J}\) has an extension, if there is another solution \(\hat{x}\) on a larger interval \(\hat{\mathcal{J}} \supset \mathcal{J}\), such that \(x(t) = \hat{x}(t)\) for \(t \in \mathcal{J}\).

The following existence theorem holds:

**1.2.1.** Let (1.5) be continuously differentiable on the open set \(\Omega\). Then for any \(x^0 \in \Omega\) the initial value problem (1.1) has a unique solution (1.6) on an interval \(\mathcal{J}\), which is maximal in the sense, that the solution has no extension to any (strictly) larger interval. The maximal interval has one of the following four forms

$$-\infty < t < -\infty, \quad -\infty < t < t_+, \quad t_- < t < \infty, \quad t_- < t < t_+,$$

where, as \(t\) approaches \(t_-\) or \(t_+\) the point \(x(t)\) approaches the boundary of \(\Omega\).
For example, \( \dot{x} = 1 + x^2, \ x(0) = 0, \) has the solution \( x = \tan(t) \) on the maximal interval \((-\pi/2, \pi/2)\). Here \( x(t) \) tends to \( \pm \infty \) for \( t \to \pm \pi/2 \). For \( \dot{x} = -x^2, \ x(0) = 1, \) the solution \( x(t) = 1/(1+x) \) is defined on the maximal interval \((-1, \infty)\) and satisfies \( \lim_{t \to -1} x(t) = \infty \).

We will see shortly, that for a linear system \( \dot{x} = Ax, \ x(0) = x^0, \) with \( A \in \mathbb{R}^{n \times n} \) all solutions are defined for \( t \in (-\infty, \infty) \).

For an initial value problem (1.1) the zeroes of the function \( G \) play a special role:

**1.2.2 Definition.** Consider the ODE \( \dot{x} = G(x) \) defined by some continuously differentiable mapping (1.5) on the open set \( \Omega \).

(a) A point \( x^* \in \Omega \) is a stationary point (or equilibrium) of the ODE if \( G(x^*) = 0 \).

(b) A stationary point \( x^* \) is stable if it has a neighborhood \( U \) of \( x^* \), which contains another neighborhood \( U_1 \), such that every solution \( x = x(t) \) started with \( x(0) \in U_1 \) is well-defined and satisfies \( x(t) \in U \) for all \( t > 0 \).

(c) A stationary point \( x^* \) is asymptotically stable, if it is stable and \( \lim_{t \to \infty} x(t) = x^* \) for any solution with \( x(0) \in U_1 \).

(d) A stationary point, that is not stable, is called unstable.

If \( x^* \in \Omega \) is a stationary point, then the initial value problem \( \dot{x} = G(x), \ x(0) = x^* \), has the constant solution \( x(t) = x^* \ \forall t \in (-\infty, \infty) \).

The equilibria of the pendulum equation (1.4) are the points \( x = (\pm k\pi, 0)^T, \ k = 0, 1, 2, \ldots \). For even \( k \), the pendulum is hanging straight down. These equilibria are obviously stable (but not asymptotically stable), since – in phase space – any motion started near one of these points will circle around it and hence stay in its neighborhood. On the other hand, the states

\[
(\pm(2k+1)\pi, 0)^T, \quad k = 0, 1, 2, \ldots,
\]

mean that the pendulum is ‘standing’ straight up. Already physical intuition suggests that these states should be unstable. As we saw, if \( x^* \) is one of these points, then there are four solutions of which two appear to end at \( x^* \) and two start at the point. But this picture is misleading, since \( x^* \) belongs to none of these solutions. Each of the four solutions is defined for all \( t \in (-\infty, \infty) \) and satisfies either \( \lim_{t \to \infty} x(t) = x^* \) or \( \lim_{t \to -\infty} x(t) = x^* \). In particular, we see that arbitrarily close to \( x^* \) we can start solutions, which move away from the point. Therefore, by definition, the point is indeed unstable. The solutions between two successive unstable equilibria; e.g., between \(-\pi \) and \( \pi \), are called heterocyclic orbits.
1.3 Predator-Prey Problems

1.3.1 The Volterra-Lotka Model

Mathematical models of predation are amongst the oldest in ecology. In 1926 the Italian biologist Humberto D’Ancona completed a statistical study of the numbers of species of fish sold on the fish markets of three Italian ports. When fishing was good, the number of fishermen increased, drawn by the success of others. After a time, the fish declined, perhaps due to over-harvest, and then the number of fishermen also declined. After some time, the cycle repeated. D’Ancona asked his father-in-law, the famous Italian mathematician Vito Volterra (1860-1940) to come up with a mathematical model that might explain the observations. Volterra soon developed a series of models for the interactions between any two or more species.

Independently an American mathematical biologist Alfred J. Lotka (1880-1949) formulated many of the same models as Volterra. His primary example of a predator-prey system comprised a population of plants and an herbivorous animal dependent on these plants for food.

The basic models of Volterra and Lotka use the following assumptions, that might be unrealistic in most practical situations:

- the predators are totally dependent on a single prey species as its only food supply,
- the food supply of the prey species is unlimited, and
- there is no threat to the prey other than the specific predator.

Let \( x(t) \) denote the size of the prey population at time \( t \). If no predators are present, the assumptions require, that the prey population increases at a constant rate; i.e., that \( \dot{x} = ax \), for some \( a > 1 \). Correspondingly, without prey the predator population is decreasing at some constant rate: \( \dot{y} = -by \), \( b > 0 \). The interaction between the species is assumed to be proportional to the product of both populations. Hence, altogether, we arrive at the so called Volterra-Lotka equations:

\[
\begin{align*}
\dot{x} &= ax - cxy \\
\dot{y} &= -by + dxy
\end{align*}
\] (1.7)

The population of prey is decreased by the predators and a greater number of predators will cause a greater decrease in the prey population. In turn, this will lead to a decrease of the predator population, and thus we should expect a cyclic increase and decrease in the populations.
In addition to \((x, y)^\top = (0, 0)^\top\), when there are neither predators nor prey, the system (1.7) has an equilibrium at \(x^* = (-b/d, a/c)^\top\). For \(a = 2\), \(b = 9\), \(c = 1\), \(d = 3\), and hence \(x^* = (3, 2)^\top\), Figure 1.3 shows a phase portrait of the system. It exhibits the expected periodic behavior around the stable (but not asymptotically stable) stationary point \(x^*\).

![Phase portrait of the Volterra-Lotka equations](image)

Figure 1.3: Phase portrait of the Volterra-Lotka equations

1.3.2 The Logistic Equation

The Volterra-Lotka equations are very simplistic and have a number of shortcomings. In particular, the specific periodic orbit around the equilibrium is entirely determined by the choice of the initial state. More realistically, such systems should be expected to have a so called limit-cycle; that is, they should approach, after some time, a specific periodic behavior for all reasonably close initial states.

An obvious problem is the assumption that, in the absence of predators, the prey population increases with a fixed rate. Such an exponential growth rate is unrealistic, since natural limits of the living space and available food resources always enforce some maximal population size \(x_{\text{max}}\).

In that case, we should expect that for population sizes near the maximum, the growth rate decreases in proportion to the remaining 'carrying capacity'. In other words, a model of the growth of a single population might have the form

\[
\dot{x} = g(x), \quad x(0) = x^0, \quad g(x) := \mu x (1 - x / x_{\text{max}}). \tag{1.8}
\]

Here \(g\) is the so called logistic or Verhulst map. It was first considered by P. F. Verhulst in 1838 after he had read Thomas Malthus' famous essay on the limits of growth of biological populations. A. J. Lotka rediscovered the logistic equation in 1925, calling it the law of population growth. It has found also applications in medicine as a model of tumor-growth.
It is easily checked, that (1.8) has the explicit solution

\[ x(t) = \frac{x_{\text{max}}}{1 + \left(\frac{x_{\text{max}}}{x^0} - 1\right) \exp(-\mu t)}. \]  

(1.9)

This are the sigmoid curves, as shown in Figure 1.4 for \( x_{\text{max}} = 1, x^0 = 0.1, \mu = 0.2. \)

\[ \begin{align*}
\text{Figure 1.4: The sigmoid curve}
\end{align*} \]

### 1.3.3 The Holling-Tanner Model

The logistic map has its own interesting features, and we will return to that shortly. In connection with predator-prey problems, it is natural to assume, that, without interference, the population growth of both the prey and the predators will be governed by the logistic map. But then the interaction of the species requires a closer look.

In 1959 the biologist Holling suggested, that the factor \( c \) in (1.7) measuring the impact of the predators on the prey should decrease with increasing prey population. Later Tanner proposed, that when the growth of the predator population follows the logistic map, then there is no need for the interference term \( dxy \); instead, it is more natural to let the maximum carrying capacity of predators to be inverse proportional to the size of the prey population. Together this led to the socalled Holling-Tanner equations

\[ \begin{align*}
\dot{x} &= ax(1 - \frac{x}{x_{\text{max}}}) - \frac{c}{\mu + x} xy \\
\dot{y} &= by(1 - \frac{\nu}{x} y)
\end{align*} \]  

(1.10)

Here \( c \) measures the rate of predation, \( \mu \) characterizes the time required for a predator to search and find a prey, and \( \nu \) represents the number of prey required to support one predator at equilibrium. Studies of several pairs of interacting species have shown, that the theoretical predictions of this model are broadly in line with practical reality when based on reasonable parameters. Without entering into further studies about this model, we use (1.10) with the following parameter values \( a = 1, x_{\text{max}} = 7, b = 0.2, c = 6/7, \mu = 1 \)
and $\nu = 0.5$. As shown in Figure 1.5, the system indeed has a limit cycle, which is reached for different initial states. Moreover Figure 1.6 indicates that the resulting periods of the oscillations are relatively constant.

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{holling-tanner-cycle.png}
\caption{Holling-Tanner Cycle}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.4\textwidth]{holling-tanner-popul.png}
\caption{Holling-Tanner Popul.}
\end{figure}

1.4 The Lorenz Attractor

In the two example we can identify several tasks, such as

- the computation of parts of the phase portrait,
- the evaluation of all equilibria and their stability properties,
- the determination of periodic solutions.

There are many other features that can occur in a phase portrait of a dynamical system. Around 1963, a meteorologist, Ed. N. Lorenz, developed a very simplified model of certain convection phenomena in the earth’s atmosphere. His system is most commonly expressed as three coupled non-linear differential equations

\[
\begin{align*}
\frac{dx_1}{dt} &= a(x_2 - x_1) \\
\frac{dx_2}{dt} &= x_1(b - x_3) - x_2 \\
\frac{dx_3}{dt} &= x_1x_2 - cx_3
\end{align*}
\]

Similar system also arise naturally in models of lasers and dynamos. The system involves three parameters. One commonly used choice is $a = 10$, $b = 28$, $c = 8/3$. Another is $a = 28$, $b = 46.92$, $c = 4$. 

9
In his work on celestial mechanics the great French mathematician Henri Poincare (1854 - 1912) proved, that there are dynamical systems, for which – in his words – "... small differences in the initial state of the system can lead to very large differences in its final state. A small error in the former could then produce an enormous one in the latter. Prediction becomes impossible, and the system appears to behave randomly.” There were no doubts about the correctness of these profound results, but it was thought, that they concerned rare systems, which do not occur in practical applications.

Thus it was a surprise, that the solutions of the small, simple system of Lorenz equations show an extreme sensitivity to the choice of the initial conditions, and, moreover, that they do not reach a steady state, nor tend to a limit cycle; that is, a periodic set. Instead, they approach more and more closely a bizarre looking set, now often called a "strange" attractor. Some typical trajectories of the Lorenz system are shown in Figure 1.7). The picture becomes much more instructive, if a code is used, that shows the dynamics of the process; i.e, how the state $x(t)$ moves along the solution curves with time $t$. It is intuitively obvious, that the solutions are "attracted" to a certain set. We will see later an example of a discrete-time system, where such an attractor becomes more visible.

Thus Poincare’s work achieved a new and profound importance. It had to be accepted, that already simple, small systems can have a strong sensitivity to small changes in the initial conditions, and that attractors can be very strange indeed. Today, there is general agreement, that these phenomena are very frequent and occur in many important practical problems. Some examples of such systems include the solar system (Poincare), the weather (Lorenz), turbulence in fluids (Libchaber), solar activity (Parker), population growth (May), and many others. In fact, sensitivity to initial data is one of the reasons,
why reliable weather and climate predictions can only be given for relatively short time intervals. Even larger and faster computers cannot overcome the inherent difficulties in the equations.

We arrive here at a very fundamental observation with profound philosophic implications. A widespread opinion is that every event – including human cognition and behaviour, decision and action – is causally determined by an unbroken chain of prior occurrences. This is the basis of the philosophic proposition called determinism, which was given a major boost by the development of Newtonian mechanics. It had led to initial value problems for systems of ODEs, and as noted, typically a solution of such a problem is uniquely determined when an initial state is known at some time. Many other mathematical models are equally deterministic.

But here we come to a limit of our cognition. All our computations are, by necessity, finite. We can only work with finite length numbers and we can only compute for a finite length of time. In order to compute the trajectory of a dynamical system for an infinite length of time, we need the initial state with an infinite precision. It is a fallacy to expect, that small changes have only small effects.
Chapter 2

Discrete Dynamical Systems

So far we assumed that the time $t$ is a continuous variable, and this led us to dynamical systems in the form of ODEs. As noted, often it is more appropriate to consider time as a discrete variable, which then leads to a dynamical system in the recursive form (1.2).

2.1 Orbits and Attractors

A discrete dynamical system uses a fixed discrete sequence $\{t_k\}$ of times and generates a sequence $x^0, x^1, x^2, \ldots$ of states from some phase space, say, $\mathbb{R}^n$, where $x^k$ corresponds to the time $t_k$. Frequently one assumes, that $t_k = k\Delta$ with a fixed $\Delta > 0$.

As indicated before, many discrete dynamical systems in $\mathbb{R}^n$ are defined in the form

$$x^{k+1} = G(x^k), \quad k = 0, 1, 2, \ldots$$  \hspace{1cm} (2.1)

where $G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a continuous mapping. This includes many iterative processes of numerical analysis, as, for example Newton’s method for the solution of a system of equations $F(x) = 0$. In that case, we have $G(x) := x - DF(x)^{-1}F(x)$, where $DF(x)$ denotes the Jacobian matrix of the nonlinear mapping $F$.

The orbit of a point $x \in \mathbb{R}^n$ is the sequence $\{x^k\}$ generated by (2.1) for the starting point $x^0 = x$. With the iterated maps defined by

$$G^k : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad G^0 := Id_n, \quad G^{k+1} := G \circ G^k, \quad k = 0, 1, \ldots$$  \hspace{1cm} (2.2)

the orbit of $x \in \mathbb{R}^n$ is the sequence $\{G^k(x)\}$.

For the discrete system (2.1) the concept of a stationary point becomes that of a fixed point; that is a point $x^* \in \mathbb{R}^n$ for which $x^* = G(x^*)$. The stability concepts carry over analogously:
A fixed point $x^*$ is stable if for any small $\epsilon > 0$ there exists a $\delta > 0$ such that $\|x^k - x^*\| < \epsilon$, $\forall k > 0$, for every orbit with starting point $\|x_0 - x^*\| < \delta$. A stable fixed point $x^*$ is asymptotically stable if, in addition, $\lim_{k \to \infty} x^k = x^*$. In other words, $x^*$ has an open neighborhood $U \subset \mathbb{R}^n$, such that the orbit of any point $x_0 \in U$ converges to $x^*$. In the discrete-time case, asymptotically stable fixed points are also called *attractive* fixed points.

Clearly, the orbit of a fixed point consists solely of the point itself. More generally, a periodic point of the system is any point $x \in \mathbb{R}^n$, such that $G^m(x) = x$ for some integer $m \geq 1$, called a period of $x$. Then the orbit of $x$ consists of the finite point set

$$\{x, G^1(x), \ldots, G^m(x)\}, \quad G^m(x) = x.$$  \tag{2.3}

An invariant set of the system (2.2) is a subset $S \subset \mathbb{R}^n$, such that $G(S) \subset S$ and hence also $G^k(S) \subset S$ for all $k \geq 0$. In particular, any stationary point is an invariant set, and more generally, the orbit (2.3) of a periodic point is a finite invariant set. Note that if $S$ is invariant, then the closure of $S$ is invariant as well. Accordingly, we restrict ourselves to closed invariant sets.

The concept of an attractive fixed point can be generalized to any invariant set. Roughly speaking, an invariant set is attractive if all orbits started in some neighborhood of the set tend to it. More precisely, a (closed) invariant set $S$ is an *attractor*, if there exists an open set $U \subset \mathbb{R}^n$, such that

$$S = \bigcap_{k=0,1,\ldots} G^k(U)$$  \tag{2.4}

Since $G^0$ is the identity on $\mathbb{R}^n$, we certainly have $S \subset U$; that is, $U$ is an open neighborhood of $S$, often call a fundamental neighborhood of the invariant set. If the fundamental neighborhood is bounded, then the invariant set is necessarily compact.

It is very common for dynamical systems to have more than one attractor.

### 2.2 The Discrete Logistic System

A widely studied example of a scalar discrete dynamical system uses the scaled logistic map $g(x) = \mu x(1 - x)$. The corresponding recursion

$$x^{k+1} = \mu x^k (1 - x^k), \quad k = 0, 1, 2, \ldots$$  \tag{2.5}

was popularized in a seminal paper in 1976 by the biologist Robert May. Its relative simplicity has made it an excellent entry-point into considerations of the concept of chaos and of systems exhibiting high sensitivity to the choice of initial conditions.
The graph of $g$ is a parabola, which passes through the $x$-axis at $x = 0$ and $x = 1$ and has its maximum $\mu/4$ at $x = 1/2$. The fixed points (stationary points) of (2.5) are $x = 0$ and $x^*(\mu) = 1 - 1/\mu$. Since 

$$0 \leq g(x) = \mu x (1 - x) \leq g(\frac{1}{2}) = \frac{\mu}{4} \quad \forall \ x \in [0, 1],$$

we see that $g$ maps $[0, 1]$ into itself as long as $0 \leq \mu \leq 4$.

We begin with some illustrations. In this case, the iteration can be followed graphically. The figures show how this is done: Go from $x^k$ to the graph, from the graph to the diagonal, from the diagonal to the graph, etc.

For $0 < \mu < 1$ we observe that the sequence \( \{x^k\} \) tends to zero, while for $1 < \mu < 3$ it converges to the fixed point $x^*(\mu) = 1 - 1/\mu$, which is therefore an attractive stationary point. For values between 2 and 3 this convergence becomes oscillatory. Then for $\mu > 3$ we
observe that the oscillations become 2-cycles. Thereafter things become more complicated, and after \( \mu = 4 \) the iteration will eventually leave the interval.

In order to understand the initial convergence, we prove a simple convergence theorem for scalar iterations.

### 2.2.1. Let \( g : \mathbb{R}^1 \rightarrow \mathbb{R}^1 \) be continuously differentiable and consider the iteration \( x^{k+1} = g(x^k) \), \( k = 0, 1, \ldots \). If \( x^* = g(x^*) \) is a stationary point, where \( |g'(x^*)| < 1 \), then there exists a \( \delta > 0 \), such that for any \( x^0 \) satisfying \( |x^0 - x^*| \leq \delta \), the iteration converges to \( x^* \). On the other hand, if \( |g'(x^*)| > 1 \), then the sequence diverges for any \( x^0 \).

**Proof.** By the differentiability assumption we know that for any \( \epsilon > 0 \) there exists a \( \delta > 0 \) such that
\[
|g(x) - g(x^*) - g'(x^*)(x - x^*)| \leq \epsilon|x - x^*|, \quad \forall |x - x^*| \leq \delta.
\]
If \( |g'(x^*)| < 1 \), then we can choose \( \epsilon > 0 \), such that \( \alpha = |g'(x^*)| + \epsilon < 1 \). Then, for \( |x^0 - x^*| \leq \delta \), we have
\[
|x^{k+1} - x^*| = |g(x^k) - g(x^*)| \\
\leq |g'(x^*)(x^k - x^*)| + |g(x^k) - g(x^*) - g'(x^*)(x^k - x^*)| \\
\leq \alpha|x^k - x^*| \leq \alpha^{k+1}|x^0 - x^*|
\]
which proves the convergence. On the other hand, if \( |g'(x^*)| > 1 \), then for \( \epsilon > 0 \) such that \( \alpha = |g'(x^*)| - \epsilon > 1 \) we obtain
\[
|x^{k+1} - x^*| = |g(x^k) - g(x^*)| \\
\geq |g'(x^*)(x - x^*)| - |g(x^k) - g(x^*) - g'(x^*)(x^k - x^*)| \\
\geq \alpha|x^k - x^*| \geq \alpha^{k+1}|x^0 - x^*|
\]
which implies the divergence. \( \Box \)
For $g(x) = \mu x(1-x)$ we have $g'(x) = \mu (1-2x)$ and hence $g'(0) = \mu < 1$ for $0 < \mu < 1$, and $|g'(x^*(\mu))| = |2-\mu| < 1$ for $1 < \mu < 3$. In other words, for $0 < \mu < 1$ we have $\lim_{k \to \infty} x^k = 0$; i.e., the population will die out. In the case $\mu = 1$, it is readily seen, that for small positive $x^0$ the iterates converge monotonically to zero. However, for small negative $x^0$ they will diverge.

For $1 < \mu < 3$ it follows that $\lim_{k \to \infty} x^k = x^*(\mu)$. More specifically, for $1 < \mu < 2$ the convergence is very rapid and independent of the initial point. For $2 \leq \mu < 3$ it becomes slower and, as we saw, oscillates around $x^*(\mu)$ for some time. For $\mu = 3$ there is still convergence, but it is excruciatingly slow.

For $\mu > 3$ we find that $|g'(x^*(\mu))| = |2-\mu| > 1$ and hence the sequence no longer converges to $x^*(\mu)$ – the point has become an unstable stationary point. At the point $\mu = 3$ the 1-cycles at $x^*(\mu)$ change into two 2-cycles. Accordingly, this is called a period-doubling bifurcation point.

Our figures illustrate this. We see, how, for $\mu < 3$ the iteration converges to a single point and at $\mu = 3$ switches to a 2-cycle. If one is careful, one detects the next period-doubling bifurcation point, where 4-cycles begin. This pattern continues and the period-doubling bifurcations come faster and faster and change into chaos. This can be seen in Figure 2.7.

![Figure 2.7: Period Doubling of the Logistic Curve](image)

An analytical study of this quickly becomes tedious. But it is instructive to look at least at the case of the 2-cycles:

A short calculation shows that

$$g^2(x) = \mu^2 x \left[ -\mu x^3 + 2\mu x^2 - (1 + \mu) x + 1 \right]$$
and hence that the stationary points of $g^2$ are the roots of

$$x - g^2(x) = \mu^2(x - (1 - \frac{1}{\mu})) q_2(x) = 0, \quad q_2(x) := x^2 - (1 + \frac{1}{\mu})x + \frac{1}{\mu}(1 + \frac{1}{\mu})$$

where we used that the stationary points of $g^2$ include those of $g$. Thus the true 2-cycles are the roots of $q_2$; that is,

$$z_{\pm}(\mu) := \frac{1}{2} \left( (1 + \frac{1}{\mu}) \pm \frac{1}{\mu} \sqrt{(\mu - 3)(\mu + 1)} \right).$$

These roots are only real for $\mu \geq 3$. Moreover for $\mu = 3$ we have the double root $z_{\pm}(3) = 2/3$. The two roots $z_-(\mu)$ and $z_+(\mu)$ form a 2-cycle. In fact, because of $z_-(\mu) + z_+(\mu) = 1 + 1/\mu$ a short calculation shows that

$$g(z_-(\mu)) = g(1 + \frac{1}{\mu} - z_+(\mu)) = z_+(\mu) - \mu q_2(z_+(\mu)) = z_+(\mu),$$

and analogously that $g(z_+(\mu)) = z_-(\mu)$.

From $g^2(x) = \mu g(x)(1 - g(x))$ it follows that $(g^2)'(x) = \mu g'(x)(1 - 2g(x))$. Hence, in view of $z_-(\mu) z_+(\mu) = (1/\mu)(1 + 1/\mu)$ we obtain that

$$(g^2)'(z_-(\mu)) = \mu^2(1 - 2z_-(\mu))(1 + 2z_+(\mu))$$

$$= \mu^2[1 - 2(z_-(\mu) + z_+(\mu)) + 4z_-(\mu) z_+(\mu)]$$

$$= \mu^2[1 - 2(1 + \frac{1}{\mu}) + 4 \frac{1}{\mu}(1 + \frac{1}{\mu})]$$

$$= -\mu^2 + 2\mu + 4$$

and the first line also implies that $(g^2)'(z_-(\mu)) = (g^2)'(z_+(\mu))$. For $\mu = 3$ the quadratic polynomial in the last line has the value 1 and for $\mu \geq 3$ it decreases monotonically and reaches the value $-1$ for $\mu = 1 + \sqrt{6}$. This implies that

$$|(g^2)'(z_-(\mu))| = |(g^2)'(z_+(\mu))| < 1 \quad \forall 3 < \mu < 1 + \sqrt{6},$$

Thus, by the theorem the iteration with $g^2$ converges in this $\mu$-interval to the calculated 2-cycles.

At $\mu = 1 + \sqrt{6}$ we have the next period-doubling bifurcation point, where now 4-cycles begin. The computation with $g^4$ is more tedious, but shows that indeed for a small interval beyond $1 + \sqrt{6}$ the iterates converge to stable 4-cycles.

The following table shows the intervals of the $2^k$-cycles for the system:
<table>
<thead>
<tr>
<th>cycle</th>
<th>( \mu )</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>3.449490</td>
</tr>
<tr>
<td>8</td>
<td>3.544090</td>
</tr>
<tr>
<td>16</td>
<td>3.564407</td>
</tr>
<tr>
<td>32</td>
<td>3.568750</td>
</tr>
<tr>
<td>64</td>
<td>3.56969</td>
</tr>
<tr>
<td>128</td>
<td>3.56989</td>
</tr>
<tr>
<td>256</td>
<td>3.569934</td>
</tr>
<tr>
<td>512</td>
<td>3.569943</td>
</tr>
<tr>
<td>1024</td>
<td>3.5699451</td>
</tr>
<tr>
<td>2048</td>
<td>3.56994557</td>
</tr>
<tr>
<td>acc. pnt.</td>
<td>3.569945672</td>
</tr>
</tbody>
</table>

Obviously, the period doubling bifurcations come faster and faster and tend to the point \( \mu_\infty \approx 3.569945672 \ldots \). It can be shown that

\[
\delta = \lim_{k \to \infty} \frac{\mu_{k+1} - \mu_k}{\mu_{k+2} - \mu_{k+1}} = 4.669201609102990\ldots
\]

is a constant, now known as the Feigenbaum constant. In fact, this constant is universal for all one-dimensional maps on a bounded interval that satisfy a certain general condition. We will not go into details.

As we observed, beyond \( \mu_\infty \) there is chaos. We can no longer see any oscillations, and slight variations in the initial point yield dramatically different results over time.

However, there are still certain isolated values of \( \mu \), that appear to show non-chaotic behavior; these are sometimes called islands of stability. For instance, beginning at \( 1 + 2\sqrt{2} \approx 3.83 \ldots \) there is a range of parameters, which show a three-cycle, and, for slightly higher \( \mu \)-values, a 6 and then a 12 cycle, etc. There are other ranges which yield 5-cycles, etc. In fact, periods of any length occur. Beyond \( \mu = 4 \), the \( x^k \) eventually leave the interval \([0, 1]\) and diverge for almost all initial values.

### 2.3 Periodic Behavior in Standard Iterations

The solution behavior we observed for the discrete logistic equation is by no means atypical. In fact, a similar behavior arises in standard iterative methods for solving nonlinear equations.

As a very simple example suppose that we apply the standard Newton method

\[
x^{k+1} = x^k - \frac{f(x^k)}{f'(x^k)} \quad \forall k = 0, 1, 2, \ldots
\]
to the cubic equation

\[ g(x) := x^3 - x = 0, \quad x \in \mathbb{R}^1, \quad (2.7) \]

with the roots \( x_1^* = -1, \ x_2^* = 0, \ \text{and} \ x_3^* = 1 \). We consider the sets \( A_j \) of all startpoints \( x^0 \) for which \( (2.6) \) converges to \( x_j^* \), \( j = 1, 2, 3 \). These sets are the attraction basins of the roots; they are shown in Figure 2.8.

\[
\begin{array}{c|c|c}
 x^0 & (a) & (b) \\
\hline
(-\infty, -1/\sqrt{3}) & \text{Limit } = x_1^* = -1 & \text{Limit } = x_2^* = 0 \\
(-1/\sqrt{5}, 1/\sqrt{5}) & \text{Limit } = x_2^* = 0 & \text{Limit } = x_3^* = 1 \\
(1/\sqrt{3}, \infty) & \text{Limit } = x_3^* = 1 & \text{Limit } = x_3^* = 1 \\
\pm(1/\sqrt{5}) & \text{period 2} & \end{array}
\]

For \( x^0 \in (-1/\sqrt{3}, -1/\sqrt{5}) \) and \( x^0(1/\sqrt{5}, 1/\sqrt{3}) \) the behaviour is less predictable. There are points from where the method converges to some root and there are also points where the iterates cycle. In other words, the attraction domains of the three roots consists of intervals between which there are intervals with a chaotic convergence behavior.

The situation becomes more evident for the complex cubic

\[ z^3 - z = 0, \quad z \in \mathbb{C}^1. \quad (2.8) \]

In terms of the real and complex parts the complex Newton method can be written as a real iterative

\[ x^{k+1} = G(x^k), \quad k = 0, 1, \ldots \quad G(x) = \begin{pmatrix} x(1)^3 - 3 \cdot x(1) \cdot x(2)^2 - 1 \\ x(2)^3 - 3 \cdot x(1)^2 \cdot x(2) \end{pmatrix}, \quad x \in \mathbb{R}^2. \]
Figure 2.9 shows the attraction basins of the three roots $x^*_r = 1$, $x^*_\pm = -(1 \pm i\sqrt{3})/2$ in different colors.

We see again that each of the roots is inside an open subset of its attraction basin. But between these open sets there are so called fractal sets, where the convergence is very irregular.

A fractal is generally "a rough or fragmented geometric shape, that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole." This property is called self-similarity. The term "fractal" was coined by Benot Mandelbrot around 1975. Because they appear to be similar at all levels of magnification, fractals are informally considered to be infinitely complex. Natural objects that approximate fractals include clouds, mountain ranges, lightning bolts, coastlines, and snow flakes. A classical example is the well known Sierpinski triangle shown in Figure 2.10.

The literature on fractals is large. It should be noted, that fractals arise not only in discrete, but also in continuous dynamical systems.

2.4 The Henon Attractor

The Lorenz system showed, that simple dynamical systems (with continuous time) can show a strange behavior. In the study of that behavior, the astronomer Michel Henon was led to look for simpler systems, which were easier to analyze. In 1976, he introduced the
following discrete dynamical system

\[ x^{k+1} = G(x^k), \quad k = 0, 1, \ldots, \quad G(x) := \begin{pmatrix} 1 + x_2 - ax_1^2 \\ bx_1 \end{pmatrix}, \quad x \in \mathbb{R}^2. \] (2.9)

which indeed does have a strange attractor. Here \(a\) and \(b\) are parameters, which, in the initial configuration, were set to \(a = 1.4\) and \(b = 0.3\).

Generally, the determinant of the Jacobian matrix of the map denotes the expansion rate of the volume (area) per iteration in the phase space. For the Henon map, we have

\[ DG(x) := \begin{pmatrix} -2ax_1 & 1 \\ b & 0 \end{pmatrix} \]

and hence \(|\det DG(x)| = |-b| = 0.3\); i.e., the map contracts areas. By taking the limit of the contraction rate, i.e., \(b = 0\), we obtain the one-dimensional system \(x^{k+1} = 1 - a(x^k)^2\), which can be transformed into the logistic map.

The Henon map may be decomposed into three mappings \(G = G_3 \circ G_2 \circ G_1\), where

\[ G_1(x) := \begin{pmatrix} x_1 \\ 1 + x_2 - ax_1^2 \end{pmatrix} \]

bends – but does not contract – an area in the \(x_2\)-direction,

\[ G_2(x) := \begin{pmatrix} bx_1 \\ x_2 \end{pmatrix} \]

is a contraction in the \(x_1\) direction, and, finally,

\[ G_3(x) := \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \]
is a reflection across the line \(x_2 = x_1\).

We see this with a simple Matlab code applied to the rectangle \([-1, 1] \times [-1, 1]\) (see Figure 2.11). After the second step, we observe already the emergence of a sichel-shaped figure. This is indeed an indication of what we obtain after many iterations. For this we start at an arbitrary initial point \(x^0\) in the vicinity of the sichel and iterate for several thousand steps. If \(x^0\) is not too far away, the result will always be the same figure shown in Figure 2.12. It turns out to be an attractive invariant set of \(G\) in the earlier defined sense. Although composed of lines, the orbits of points on this set do not flow continuously, but jump from one location to another.

![Figure 2.11: Henon Map Steps 1,2](image1)

![Figure 2.12: Henon Attractor](image2)

The figure is the same for all initial points, for which the iteration does not diverge, but the resulting sequences \(\{x^k\}\) are very different. In fact, even for small changes of the initial point, the sequences will eventually diverge from each other and evolve separately. The Henon attractor also has a great amount of fine structure and successive magnifications show an ever increasing degree of detail. On closer inspection, what looks like a line turns out to be a set of lines, and when these lines are magnified further, they also turn out to be sets of lines, etc.
Chapter 3

Planar Linear Systems

In this chapter we begin with an elementary analysis of a simple class of dynamical systems, namely two-dimensional real, linear systems

\[ \dot{x} = Ax, \quad x(0) = x^0, \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \]  

(3.1)

Planar linear systems (3.1) are not entirely artificial problems as the three simple examples in the next section will show.

3.1 Examples of Planar Systems

3.1.1 A Shock-absorber

As a model of a shock absorber, consider a mass \( m \) supported on a vertically mounted spring, which is constrained to move only along the axis of the spring. The motion of the mass is hydraulically damped. For this it is attached to a piston, that moves in a cylinder filled with hydraulic fluid; see Figure 3.1.

Figure 3.1: A Shock Absorber
Let $x$ denote the displacement of the mass below its equilibrium position and assume that (i) the spring is linear with restoring force $-rx$, $r > 0$, and (ii) the force exerted by the piston opposing the motion equals $2cm\dot{x}$ where $c \geq 0$, and $m\dot{x}$ is the momentum of the mass. By Newton’s law, the equation of motion is then $m\ddot{x} = -rx - 2cm\dot{x}$, which in first order form becomes the linear system

$$\dot{x} = Ax, \quad A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -\beta \end{pmatrix}, \quad \omega = \sqrt{\frac{r}{m}} > 0, \quad \beta = 2c > 0,$$

(3.2)

where $\omega$ will turn out to be the natural frequency of the undamped system.

### 3.1.2 An LRC Circuit

Generally, an LRC circuit is an electrical circuits built only with (linear) resistors, inductors, and capacitors. For these three types of components the following laws hold:

1. The current $i_R$ flowing through a resistor is – by Ohm’s law – proportional to the potential difference $u_R$ across it:

$$u_R = Ri_R, \quad R > 0 \text{ resistance.}$$

2. The potential difference $u_L$ across an inductor is proportional to the time derivative of the current $i_L$ passing through it:

$$u_L = L\frac{di_L}{dt}, \quad L > 0 \text{ inductance.}$$

3. The current $i_C$ through a capacitor is proportional to the time derivative of the potential difference $u_C$ across it:

$$i_C = C\frac{du_C}{dt}, \quad C > 0 \text{ capacitance.}$$

The circuit equations can now be built-up by applying the Kirchhoff laws:

- **Current law:** The sum of the currents flowing into a node is equal to the sum of the currents flowing out of it,

- **Voltage law:** The sum of the potential differences around any closed loop in a circuit is zero.

As an example, we consider a simple LRC circuit with a resistor, inductor, and capacitor in series in a circular loop, see Figure 3.2.
In this case, we obtain the equations
\[ i_R = i_C = i_L, \quad \text{current law} \]
\[ u_R + u_L + u_C = 0, \quad \text{voltage law} \]
\[ u_R = R \, i_R, \quad \text{resistor} \]
\[ L \frac{di_L}{dt} = u_L, \quad \text{inductance} \]
\[ C \frac{du_C}{dt} = i_C, \quad \text{capacitor} \]

With \( u = u_C \) and \( i = i_R \), simplification yields
\[ \frac{du}{dt} = \frac{1}{C} \, i, \quad \frac{di}{dt} = -\frac{1}{L} \left( u_R + u_C \right) = -\frac{R}{L} \, i - \frac{1}{L} \, u. \]

This is a planar linear system (3.1) with the matrix
\[ A = \begin{pmatrix} 0 & 1 \\ -\alpha & -\beta \end{pmatrix}, \quad \alpha = \frac{1}{LC} > 0, \quad \beta = \frac{2R}{L} > 0, \quad (3.3) \]
which has exactly the same form as (3.2) with coefficients of the same sign.

### 3.1.3 An Economic Model

Consider an economy, where at time \( t \), \( Y = Y(t) \) is the total output, \( C = C(t) \) the consumption, \( I = I(t) \) the investment, and \( G \) the (constant) government expenditure. If there are no further influences, we then require that
\[ Y(t) = C(t) + I(t) + G. \quad (3.4) \]

It is not unreasonable to assume that the consumption \( C(t) \) is proportional to the output \( Y(t - \tau) \), where \( \tau \) denotes a response time. If we use the first order approximation
\[ Y(t - \tau) = Y(t) - \tau \dot{Y}(t), \]
this means that
\[ C(t) = (1 - s)[Y(t) - \tau \dot{Y}(t)]. \quad (3.5) \]
where $s > 0$ is a savings factor.

Similarly, investments may be assumed to be proportional to the change in output but again with a time lag; i.e.,

$$I(t + \sigma) = a\dot{Y}(t).$$

Again in first order approximation this gives

$$I(t) + \sigma\dot{I}(t) = a\dot{Y}(t). \quad (3.6)$$

By substituting (3.5) into (3.4) we obtain

$$(1 - s)\tau\dot{Y}(t) = -sY(t) + I(t) + G \quad (3.7)$$

and hence after differentiation and substitution of (3.6)

$$(1 - s)\tau\dot{Y}(t) = -s\dot{Y}(t) + \frac{1}{\sigma}[a\dot{Y}(t) - I(t)] \quad (3.8)$$

Here $I(t)$ can be eliminated by means of (3.7). At the same time it is useful to introduce the translated variable

$$y = Y - \frac{G}{s}.$$  

We then obtain the second order ODE

$$\sigma\tau(1 - s) \ddot{y} + [s\sigma + (1 - s)\tau - a] \dot{y} + s \ y = 0 \quad (3.9)$$

In first order form this is a planar linear system (3.1) with the matrix

$$A = \begin{pmatrix} 0 & 1 \\ -\alpha & -\beta \end{pmatrix}, \quad \alpha = \frac{s}{\sigma\tau(1 - s)} > 0, \quad \beta = \frac{s\sigma + (1 - s)\tau - a}{2\sigma\tau(1 - s)}. \quad (3.10)$$

Again, this matrix has an analogous form as (3.2) and (3.3), but here $\beta$ need not be positive. We will see shortly, that this has a profound effect.

### 3.2 Transformations of Planar Systems

For the analysis of any linear system

$$\dot{x} = Ax$$

it is natural to look for a simplifying coordinate transformation $x = Uy$. Here $U$ is some nonsingular matrix and the transformed system becomes

$$\dot{y} = Qy, \quad Q = U^{-1}AU.$$
A principal tool for constructing transformations, which produce matrices $Q$ of a simpler form, is the use of the eigenvalues and eigenvectors of $A$. We will discuss later the computation of eigenvalues and eigenvectors for general matrices and concentrate here on our simple case of $2 \times 2$ matrices

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$ 

Recall, that a complex number $\lambda$ is an eigenvalue of $A$ if there exists a nonzero, complex vector $u \in \mathbb{C}^2$, such that $Au = \lambda u$; that is, $(A - \lambda I)u = 0$. Hence, $A - \lambda I$ must be a singular matrix and this holds exactly if $\det(A - \lambda I) = 0$.

In our case, we obtain the characteristic polynomial

$$p_A(\lambda) = \det(A - \lambda I) = \lambda^2 - \text{tr}(A)\lambda + \det(A)$$

(3.11)

where $\text{tr}(A) = a + d$ is the trace of $A$ and $\det(A) = ad - bc$ the determinant. Obviously the roots of $p_A$ are

$$\lambda_{1,2} = \frac{1}{2} \left( \text{tr}(A) \pm \sqrt{\Delta} \right), \quad \Delta := \text{tr}(A)^2 - 4 \det(A).$$

(3.12)

Hence we have three principal cases:

(a) $\Delta > 0$: Two distinct real eigenvalues $\lambda_1 > \lambda_2$,

(b) $\Delta < 0$: Two conjugate complex eigenvalues $\lambda_{1,2} = \alpha \pm i\beta$,

(c) $\Delta = 0$: A double root $\lambda_1 = \lambda_2 = \frac{1}{2} \text{tr}(A)$.

There is no need for developing formulas for the corresponding eigenvectors. Instead, we concentrate on illustrative examples for the three cases.

**Case (a):** We will prove later, that when the real eigenvalues $\lambda_1$ and $\lambda_2$, are distinct, then the corresponding eigenvectors $u^1, u^2$ are linearly independent. Clearly, $Au^1 = \lambda_1 u^1, Au^2 = \lambda_2 u^2$, is equivalent with

$$A(u^1, u^2) = (u^1, u^2)D, \quad D = \text{diag}(\lambda_1, \lambda_2),$$

and for $U = (u^1, u^2)$ the transformed matrix equals $U^{-1}AU = D$. In the new coordinate system the dynamical system

$$\dot{y}_1 = \lambda_1 y_1, \quad \dot{y}_2 = \lambda_2 y_2, \quad y_1(0) = y_1^0, \quad y_2(0) = y_2^0,$$

has the solution

$$y_1(t) = \exp(t\lambda_1) \ y_1^0, \quad y_2(t) = \exp(t\lambda_2) \ y_2^0.$$

**Case (b):** For distinct conjugate complex eigenvalues $\lambda_{1,2} = \alpha \pm i\gamma$ and

$$A(v + iw) = (\alpha + i\gamma)(v + iw),$$

27
we have
\[ A(v - iw) = (\alpha - i\gamma)(v - iw). \]

Then
\[ A(v, w) = (v, w) T, \quad T := \begin{pmatrix} \alpha & \gamma \\ -\gamma & \alpha \end{pmatrix}, \]
and \( v \) \( w \) are linearly independent. Hence with \( U = (v, w) \) we obtain the transformed system \( \dot{y} = Ty, U^{-1}AU = T \), which has the solutions
\[
\begin{align*}
y_1(t) &= c_1 e^{t\alpha} \cos(\gamma t) + c_2 e^{t\alpha} \sin(\gamma t) \\
y_2(t) &= -c_1 e^{t\alpha} \sin(\gamma t) + c_2 e^{t\alpha} \cos(\gamma t)
\end{align*}
\]
where the constants \( c_1 \) \( c_2 \) are chosen so as to satisfy the initial conditions.

**Case (c):** There are two distinct subcases. If \( \lambda_1, \lambda_2 = a \) is a double eigenvalue and there exist two linearly independent eigenvectors \( u^1 \) and \( u^2 \), then we have
\[ AU = U \text{diag}(\lambda, \lambda) = \lambda U, \quad U = (u^1, u^2). \]

Hence \( U^{-1}AU = \lambda I_2 \) and the transformed system \( \dot{y} = \lambda y \) has the solution
\[
\begin{align*}
y_1(t) &= \exp(t\lambda) \ y_1^0, \\
y_2(t) &= \exp(t\lambda) \ y_2^0.
\end{align*}
\]

On the other hand, we will see, that there may well be only one linearly independent eigenvector \( u \). Then it turns out, that there exists a principal vector
\[ v \in \mathbb{R}^2, \quad (A - \lambda I_2)^2 v = 0, \quad (A - \lambda I_2)v \neq 0, \]
for which necessarily \( (A - \lambda I_2)v = \nu u \). We can scale \( v \) such that \( \nu = 1 \), and therefore obtain
\[ A(u, v) = (u, v) T, \quad T := \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}; \]
which implies that \( u \), \( v \) are linearly independent. Thus with \( U = (u, v) \) the transformed system equals \( \dot{y} = Ty, \quad T = U^{-1}AU \), and has the solutions
\[
\begin{align*}
y_1(t) &= (y_1^0 + y_2^0 \ t) e^{t\lambda}, \\
y_2(t) &= y_2^0 \ e^{t\lambda}.
\end{align*}
\]

This, indeed, covers all likely cases and we obtained the following result:

**3.2.1. Let** \( A \) **be a real** \( 2 \times 2 \) **matrix with eigenvalues** \( \lambda_1, \lambda_2 \). **Then there exists a real, non-singular matrix** \( U \) **such that** \( U^{-1}AU = Q \), **where** \( Q \) **is one of the four matrices:**
\[
\begin{align*}
(a) & \quad \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}, \quad \lambda_1 > \lambda_2, &
(b) & \quad \begin{pmatrix} \alpha & \gamma \\ -\gamma & \alpha \end{pmatrix}, \quad \lambda_{1,2} = \alpha \pm i\beta. \\
(c1) & \quad \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{pmatrix}, \quad \lambda_1 = \lambda_2, &
(c2) & \quad \begin{pmatrix} \lambda_1 & 1 \\ 0 & \lambda_1 \end{pmatrix}, \quad \lambda_1 = \lambda_2.
\end{align*}
\]
3.3 Phase Portraits of Planar Systems

The phase portraits of the four cases in 3.2.1 are visibly different. Clearly $x = 0$ is the only stationary point of the given system $\dot{x} = Ax$. Depending on the signs of the determinant $\det(A)$, the trace $\text{tr}(A)$, and the discriminant $\Delta$, we have the categorization given in the following table:

<table>
<thead>
<tr>
<th>Type</th>
<th>tr</th>
<th>det</th>
<th>$\Delta$</th>
<th>Eigenvalues</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saddles</td>
<td>$&lt; 0$</td>
<td></td>
<td></td>
<td>$\lambda_1 &gt; 0 &gt; \lambda_2$</td>
</tr>
<tr>
<td>Sources</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>node</td>
</tr>
<tr>
<td></td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td>improper node</td>
</tr>
<tr>
<td></td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>$&lt; 0$</td>
<td>spiral (focus)</td>
</tr>
<tr>
<td>Sinks</td>
<td>$&lt; 0$</td>
<td>$&gt; 0$</td>
<td>$&gt; 0$</td>
<td>node</td>
</tr>
<tr>
<td></td>
<td>$&lt; 0$</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td>improper node</td>
</tr>
<tr>
<td></td>
<td>$&lt; 0$</td>
<td>$&gt; 0$</td>
<td>$&lt; 0$</td>
<td>spiral (focus)</td>
</tr>
<tr>
<td>Centers</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>$&lt; 0$</td>
<td>$\lambda_{1,2} = \pm \beta i$, $\alpha &gt; 0$</td>
</tr>
<tr>
<td>Singular</td>
<td>$&gt; 0$</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>line-source</td>
</tr>
<tr>
<td></td>
<td>$= 0$</td>
<td>$= 0$</td>
<td>$= 0$</td>
<td>singular-saddle</td>
</tr>
<tr>
<td></td>
<td>$&lt; 0$</td>
<td>$= 0$</td>
<td>$&gt; 0$</td>
<td>line-sink</td>
</tr>
</tbody>
</table>

In brief, a primary classification is:
- stable points or sinks (Senken),
- unstable points or sources (Quellen),
- saddle points (Sattel).

The stable and unstable cases have the secondary subclasses
- nodes (Knoten)
- centers (Wirbel),
- spiral or foci (Spirale, Brennpunkt)

Saddles have both dynamically contracting and expanding directions and we cannot speak of either stability or unstability. Figures Figure 3.3 – 3.10 show examples for the various cases, including the last two with singular matrices.

We look also at our three examples in section 3.1. In each case the matrix had the form

$$A = \begin{pmatrix} 0 & 1 \\ -\alpha & -\beta \end{pmatrix}.$$  \hspace{1cm} (3.14)

where, in all cases, $\det(A) = \alpha > 0$. In the first two examples we also had $\text{tr}(A) = -\beta \leq 0$ and hence, by our table, the origin is either a stable or asymptotically stable stationary point. But in the economic example we cannot exclude the case $\beta < 0$, in which case the point is unstable.
In the shock-absorber example we had
\[ A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -2c \end{pmatrix}, \quad \omega = \sqrt{\frac{r}{m}} > 0, \quad c > 0. \] \hfill (3.15)

where \( m \) is the mass, \( r \) the spring constant, and \( c \) the hydraulic damping factor. Engineers distinguish the following four cases and shown in Figures 3.11 – 3.14:

1. \( c = 0 \), undamped system with natural frequency \( \omega \). Phase portrait is a center.
2. \( 0 < c < \omega \), underdamped system. Phase portrait a spiral.
3. \( c = \omega \), critically damped system. Phase portrait an improper node.
4. \( c > \omega \), overdamped system. Phase portrait a (proper) node.

The behavior of the LRC circuit is entire analogous, except that then \( \omega = \frac{1}{\sqrt{LC}} > 0 \) and \( c = \frac{R}{L} \). For the economic problem the phase portrait will again be similar to those of the other two example as long as in (3.14) the coefficient \( \beta \) is positive. Then the trace is negative and the origin is stable. But in this case there are situation, where \( \beta < 0 \). Since
\[
\beta = \frac{s\sigma + (1 - s)\tau - a}{2\sigma\tau(1 - s)},
\]
there may well be negative values when the investment rate is high and the lag-times are short. As noted, the system then becomes unstable.
Figure 3.3: Source – Node

Figure 3.4: Source – Spiral

Figure 3.5: Center

Figure 3.6: Saddle

Figure 3.7: Source – Improper Node 2

Figure 3.8: Source – Improper Node 1
Figure 3.9: Singular saddle

Figure 3.10: Line Source

Figure 3.11: Undamped Shockabs.

Figure 3.12: Under-damped Shockabs.

Figure 3.13: Crit. damped Shockabs.

Figure 3.14: Over damped Shockabs.