Chapter 1

Linear equations

1.1 Objects of study

Many problems in economics, biology, physics and engineering involve rate of change dependent on the interaction of the basic elements–assets, population, charges, forces, etc.–on each other. This interaction is frequently expressed as a system of ordinary differential equations, a system of the form

$$x_1'(t) = f_1(t, x_1(t), x_2(t), \dots, x_n(t)), \tag{1.1}$$
$$x_2'(t) = f_2(t, x_1(t), x_2(t), \dots, x_n(t)), \tag{1.2}$$
$$\vdots$$
$$x_n'(t) = f_n(t, x_1(t), x_2(t), \dots, x_n(t)). \tag{1.3}$$

Here the (known) functions $(\tau, \xi_1, \dots, \xi_n) \mapsto f_i(\tau, \xi_1, \dots, \xi_n)$ take values in $\mathbb{R}$ (the real numbers) and are defined on a set in $\mathbb{R}^{n+1}$ ($\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$, $n+1$ times).

We seek a set of n unknown functions $x_1, \dots, x_n$ defined on a real interval I such that when the values of these functions are inserted into the equations above, the equality holds for every $t \in I$.

Introducing the vector notation

$$x(t) := \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad x'(t) := \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \quad \text{and} \quad f(\tau, \xi) := \begin{bmatrix} f_1(\tau, \xi_1, \dots, \xi_n) \\ \vdots \\ f_n(\tau, \xi_1, \dots, \xi_n) \end{bmatrix} \quad \xi := \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix},$$

the system of differential equations can be abbreviated simply as

$$x'(t) = f(t, x(t)).$$

Definition. A function $x : [t_0, t_1] \to \mathbb{R}^n$ is said to be a *solution* of (1.1)-(1.3) if x is differentiable on $[t_0, t_1]$ and it satisfies (1.1)-(1.3) for each $t \in [t_0, t_1]$.

In addition, an initial condition may also need to be satisfied: $x(0) = x_0 \in \mathbb{R}^n$, and a corresponding solution is said to satisfy the *initial value problem*

$$x'(t) = f(t, x(t)), \quad x(t_0) = x_0.$$

In this course we will mainly consider the case when the functions $f_1, \dots, f_n$ do not depend on t (that is, they take the same value for all t).

In most of this course, we will consider autonomous systems, which are defined as follows.

Definition. If f does not depend on t, that is, it is simply a function defined on some subset of $\mathbb{R}^n$, taking values in $\mathbb{R}^n$, the differential equation

$$x'(t) = f(x(t)),$$

is called *autonomous*.

But we begin our study with an even simpler case, namely when these functions are *linear*, that is,

$$f(\xi) = A\xi,$$

where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Then we obtain the 'vector' differential equation

$$x'(t) = Ax(t),$$

which is really the system of scalar differential equations given by

$$x_1' = a_{11}x_1 + \cdots + a_{1n}x_n, \tag{1.4}$$

$$\vdots$$

$$x_n' = a_{n1}x_1 + \cdots + a_{nn}x_n. \tag{1.5}$$

In many applications, the equations occur naturally in this form, or it may be an approximation to a nonlinear system.

Exercises.

1. Classify the following differential equations as autonomous/nonautonomous. In each autonomous case, also identify if the system is linear or nonlinear.
   (a) $x'(t) = e^t$.
   (b) $x'(t) = e^{x(t)}$.
   (c) $x'(t) = e^t y(t)$, $y'(t) = x(t) + y(t)$.
   (d) $x'(t) = y(t)$, $y'(t) = x(t)y(t)$.
   (e) $x'(t) = y(t)$, $y'(t) = x(t) + y(t)$.

2. Verify that the differential equation has the given function or functions as solutions.

   (a) $x'(t) = e^{\sin x(t)} + \cos(x(t))$; $x(t) \equiv \pi$.

   (b) $x'(t) = ax(t)$, $x(0) = x_0$; $x(t) = e^{ta}x_0$.

   (c) Let p be a polynomial function:

   $$p(\lambda) := a_n\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0.$$

   Let λ_0 be root of this polynomial, that is, λ_0 is a number satisfying $p(\lambda_0) = 0$. The differential equation under consideration is:

   $$a_n\frac{d^n x}{dt^n}(t) + a_{n-1}\frac{d^{n-1}x}{dt^{n-1}}(t) + \cdots + a_1\frac{dx}{dt}(t) + a_0x(t) = 0.$$

   The claimed solution is $x(t) = Ce^{\lambda_0 t}$, where C is a constant.

   (d) $\begin{bmatrix} x_1'(t) \\ x_2'(t) \end{bmatrix} = \begin{bmatrix} 2x_2(t) \\ -2x_1(t) \end{bmatrix}$; $\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} \sin(2t) \\ \cos(2t) \end{bmatrix}$.

   (e) $x'(t) = 2t(x(t))^2$; $x_1(t) = \frac{1}{1-t^2}$ for $t \in (-1, 1)$, $x_2(t) \equiv 0$.

   (f) $x'(t) = 3(x(t))^{\frac{2}{3}}$, $x(0) = 0$; $x_1(t) = 0$ and $x_2(t) = t^3$ for $t \geq 0$.

   (g) $\frac{\partial x}{\partial t}(s,t) - \frac{\partial^2 x}{\partial s^2}(s,t) = 0$; $x(s,t) = Ce^{t\tau_0 + s\sigma_0}$, where C is a constant and τ_0, σ_0 are fixed numbers satisfying $\tau_0 - \sigma_0^2 = 0$.

3. Find value(s) of m such that $x(t) = t^m$ is a solution to $2tx'(t) = x(t)$ for $t \geq 1$.

4. Show that every solution of $x'(t) = (x(t))^2 + 1$ is an increasing function.

1.2 Using Maple to investigate differential equations

1.2.1 Getting started

Maple should be installed on all public school computers; with the newest version being "Maple 2016". (Some computers will also have "Maple 17", which, confusingly, is an older version than "Maple 2016".) You can obtain a free copy of Maple 2016 from the IMT Walk In Centre on the 1st floor of the LSE Library.

Background material about Maple can be found at:

1. the *MapleSoft Student Help Center*;

2. Maple's own tutorials and help, which can be found under *Maple Help* in Maple itself;

3. MA100 Maple tutorials, which can be found on the *MA209 Moodle page*.

1.2.2 Differential equations in Maple

Here we describe some main Maple commands related to differential equations.

1. *Defining differential equations.* For instance, to define the differential equation $x' = x + t$, we give the following command.

```
[> ode1 := diff(x(t),t) = t + x(t);
```

Here, `ode1` is the label or name given to the equation, `diff(x(t),t)` means that the function $t \mapsto x(t)$ is differentiated with respect to t, and the last semicolon indicates that we want Maple to display the answer upon execution of the command. Indeed, on hitting the enter-key, we obtain the following output.

$$ode1 \ := \frac{d}{dt}x(t) = t + x(t)$$

The differentiation of x can also be expressed in another equivalent manner as shown below.

```
[> ode1 := D(x)(t) = t + x(t);
```

A second order differential equation, for instance $x'' = x' + x + \sin t$ can be specified by

```
[> ode2 := diff(x(t),t,t) = diff(x(t),t) + x(t) + sin(t);
```

or equivalently by the following command.

```
[> ode2 := D(D(x))(t) = D(x)(t) + x(t) + sin(t);
```

A system of ODEs can be specified in a similar manner. For example, if we have the system

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -x_1, \end{aligned}$$

then we can specify this as follows:

```
[> ode3a := diff(x1(t),t) = x2(t);  ode3b := diff(x2(t),t) = -x1(t);
```

2. *Solving differential equations.* To solve say the equation `ode1` from above, we give the command

```
[> dsolve(ode1);
```

which gives the following output:

$$x(t) = -t - 1 + e^t\_C1$$

The strange "$\_C1$" is a constant. The strange name is used by Maple as an indication that the constant is generated by Maple (and has not been introduced by the user).

To solve the equation with a given initial value, say with $x(0) = 1$, we use the command:

```
[> dsolve({ode1, x(0)=1});
```

If our initial condition is itself a parameter α, then we can write

```
[> dsolve({ode1, x(0)=alpha});
```

which gives:

$$x(t) = -t - 1 + e^t(1 + \alpha)$$

We can also give a name to the equation specifying the initial condition as follows

```
[> ic1 := x(0)=2;
```

and then solve the initial value problem by writing:

```
[> dsolve({ode1, ic1});
```

Systems of differential equations can be handled similarly. For example, the ODE system `ode3a, ode3b` can be solved by

```
[> dsolve({ode3a, ode3b});
```

and if we have the initial conditions $x_1(0) = 1$, $x_2(0) = 1$, then we give the following command:

```
[> dsolve({ode3a, ode3b, x1(0)=1, x2(0)=1});
```

3. *Plotting solutions of differential equations.* The tool one has to use is called `DEplot`, and so one has to activate this at the outset using the command:

```
[> with(DEtools):
```

Once this is done, one can use for instance the command `DEplot`, which can be used to plot solutions. This command is quite complicated with many options, but one can get help from Maple by using:

```
[> ?DEplot;
```

For the equation `ode1` above, the command

```
[> DEplot(ode1, x(t), t=-2..2, [[x(0)=0]]);
```

will give a nice picture of a solution to the associated initial value problem, but it contains some other information as well. The various elements in the above command are: `ode1` is the label specifying which differential equation we are solving, `x(t)` indicated the dependent variable, `t=-2..2` indicates the independent variable and its range, and `[[x(0)=0]]` gives the initial value.

One can also give more than one initial value, for instance:

```
[> DEplot(ode1, x(t), t=-2..2, [[x(0)=-1], [x(0)=0], [x(0)=1]]);
```

The colour of the plot can also be changed:

```
[> DEplot(ode1, x(t), t=-2..2, [[x(0)=-1], [x(0)=0], [x(0)=1]],
   linecolour=blue);
```

The arrows one sees in the picture show the *direction field*, a concept we will discuss in Chapter 2. One can hide these arrows:

```
[> DEplot(ode1, x(t), t=-2..2, [[x(0)=-1], [x(0)=0], [x(0)=1]],
   arrows=NONE);
```

To make plots for higher order ODEs, one must give the right number of initial values. We consider an example for `ode2` below:

```
[> DEplot(ode2, x(t), t=-2..2, [[x(0)=0, D(x)(0)=0], [x(0)=0, D(x)(0)=2]]);
```

One can also handle systems of ODEs using DEplot, and we give an example below.

```
[> DEplot({ode3a,ode3b}, {x1(t),x2(t)}, t=0..10, [[x1(0)=1, x2(0)=0]],
   scene=[t,x1(t)]);
```

If one wants x_1 and x_2 to be displayed in the same plot, then we can use the command display as demonstrated in the following example.

```
[> with(plots):
[> plot1 := DEplot({ode3a,ode3b}, {x1(t),x2(t)}, t=0..10,
   [[x1(0)=1, x2(0)=0]], scene=[t,x1(t)]):
[> plot2 := DEplot({ode3a,ode3b}, {x1(t),x2(t)}, t=0..10,
   [[x1(0)=1, x2(0)=0]], scene=[t,x2(t)], linecolour=red):
[> display(plot1,plot2);
```

In Chapter 2, we will learn about 'phase portraits' which are plots in which we plot one solution against the other (with time giving this parametric representation) when one has a 2D system. We will revisit this subsection in order to learn how we can make phase portraits using Maple.

Exercises.

1. In each of the following initial-value problems, find a solution using Maple. Verify that the solution exists for some $t \in I$, where I is an interval containing 0.

   (a) $x' = x + x^3$ with $x(0) = 1$.

   (b) $x'' + x = \frac{1}{2}\cos t$ with $x(0) = 1$ and $x'(0) = 1$.

   (c) $\left.\begin{array}{rcl} x_1' &=& -x_1 + x_2 \\ x_2' &=& x_1 + x_2 + t \end{array}\right\}$ with $x_1(0) = 0$ and $x_2(0) = 0$.

2. In forestry, there is interest in the evolution of the population x of a pest called 'spruce budworm', which is modelled by the following equation:

$$x' = x\left(2 - \frac{1}{5}x - \frac{5x}{2+x^2}\right). \tag{1.6}$$

   The solutions of this differential equation show radically different behaviour depending on what initial condition $x(0) = x_0$ one has in the range $0 \le x_0 \le 10$.

   (a) Use Maple to plot solutions for several initial values in the range $[0, 10]$.

   (b) Use the plots to describe the different types of behaviour, and also give an interval for the initial value in which the behaviour occurs. (For instance: For $x_0 \in [0, 8)$, the solutions $x(t)$ go to 0 as t increases. For $x_0 \in [8, 10]$ the solutions $x(t)$ go to infinity as t increases.)

   (c) Use Maple to plot the function

$$f(x) = x\left(2 - \frac{1}{5}x - \frac{5x}{2+x^2}\right),$$

   in the range $x \in [0, 10]$. Can the differential equation plots be explained theoretically? HINT: See Figure 1.2.

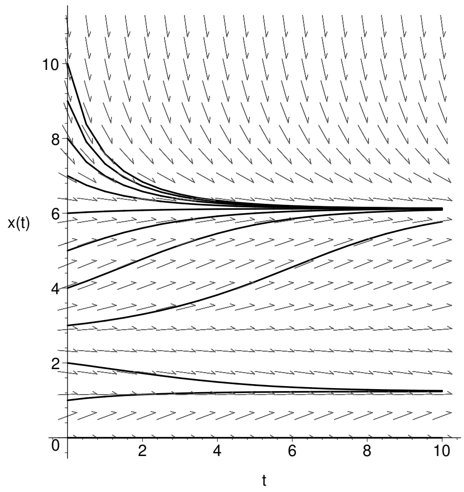

Figure 1.1: Population evolution of the budworm for various initial conditions.

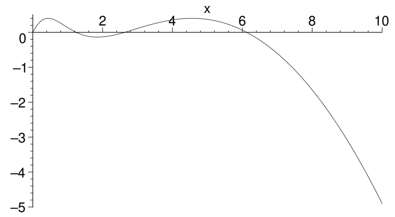

Figure 1.2: Graph of the function f.

1.3 High order ODE to a first order ODE. State vector.

Note that the system of equations (1.1)-(1.3) are *first order*, in that the derivatives occurring are of order at most 1. However, in applications, one may end with a model described by a set of high order equations. So why restrict our study only to first order systems? In this section we learn that such high order equations can be expressed as a system of first order equations, by introducing a 'state vector'. So throughout the sequel we will consider only a system of first order equations.

Let us consider the second order differential equation

$$y''(t) + a(t)y(t) + b(t)y(t) = u(t). \tag{1.7}$$

If we introduce the new functions x_1, x_2 defined by

$$x_1 = y \text{ and } x_2 = y',$$

then we observe that

$$\begin{aligned} x_1'(t) &= y'(t) = x_2(t), \\ x_2'(t) &= y''(t) = -a(t)y'(t) - b(t)y(t) + u(t) = -a(t)x_2(t) - b(t)x_1(t) + u(t), \end{aligned}$$

and so we obtain the system of first order equations

$$x_1'(t) = x_2(t), \tag{1.8}$$
$$x_2'(t) = -a(t)x_2(t) - b(t)x_1(t) + u(t), \tag{1.9}$$

which is of the form (1.1)-(1.3).

Solving (1.7) is equivalent to solving the system (1.8)-(1.9). To see the equivalence, suppose that (x_1, x_2) satisfies the system (1.8)-(1.9). Then x_1 is a solution to (1.7), since

$$(x_1'(t))' = x_2'(t) = -b(t)x_1(t) - a(t)x_1'(t) + u(t),$$

which is (1.7). On the other hand, if y is a solution to (1.7), then define $x_1 = y$ and $x_2 = y'$, and proceeding as in the preceding paragraph, this yields a solution of (1.8)-(1.9).

More generally, if we have an nth order scalar equation

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y'(t) + a_0(t)y(t) = u(t),$$

then by introducing the vector of functions

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} := \begin{bmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{bmatrix}, \tag{1.10}$$

we arrive at the equivalent first order system of equations

$$\begin{aligned} x_1'(t) &= x_2(t), \\ x_2'(t) &= x_3(t), \\ &\vdots \\ x_{n-1}'(t) &= x_n(t), \\ x_n'(t) &= -a_1(t)x_1(t) - \cdots - a_{n-1}(t)x_{n-1}(t) + u(t). \end{aligned}$$

The auxiliary vector in (1.10) comprising successive derivatives of the unknown function in the high order differential equation, is called a *state*, and the resulting system of first order differential equations is called a *state equation*.

Exercises. By introducing appropriate state variables, write a state equation for the following (systems of) differential equations:

1. $x'' + \omega^2 x = 0$.
2. $x'' + x = 0$, $y'' + y' + y = 0$.
3. $x'' + t \sin x = 0$.

1.4 The simplest example

The differential equation

$$x'(t) = ax(t) \tag{1.11}$$

is the simplest differential equation. It is also one of the most important. First, what does it mean? Here $x : \mathbb{R} \to \mathbb{R}$ is an unknown real-valued function (of a real variable t), and $x'(t)$ is its derivative at t. The equation (1.11) holds for every value of t, and a denotes a constant.

The solutions to (1.11) are obtained from calculus: if C is any constant, then the function f given by $f(t) = Ce^{ta}$ is a solution, since

$$f'(t) = Cae^{ta} = a(Ce^{ta}) = af(t).$$

Moreover, there are no other solutions. To see this, let u be any solution and compute the derivative of v given by $v(t) = e^{-ta}u(t)$:

$$\begin{aligned} v'(t) &= -ae^{-ta}u(t) + e^{-ta}u'(t) \\ &= -ae^{-ta}u(t) + e^{-ta}au(t) \quad (\text{since } u'(t) = au(t)) \\ &= 0. \end{aligned}$$

Therefore by the fundamental theorem of calculus,

$$v(t) - v(0) = \int_0^t v'(t)dt = \int_0^t 0dt = 0,$$

and so $v(t) = v(0)$ for all t, that is, $e^{-ta}u(t) = u(0)$. Consequently $u(t) = e^{ta}u(0)$ for all t.

So we see that the *initial value problem*

$$x'(t) = ax(t), \quad x(0) = x_0$$

has the unique solution

$$x(t) = e^{ta}x_0, \quad t \in \mathbb{R}.$$

As the constant a changes, the nature of the solutions changes. Can we describe qualitatively the way the solutions change? We have the following cases:

1° $a < 0$. In this case,

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} e^{ta}x(0) = 0x(0) = 0.$$

Thus the solutions all converge to zero, and moreover they converge to zero exponentially, that is, they there exist constants $M > 0$ and $\epsilon > 0$ such that the solutions satisfy an inequality of the type $|x(t)| \leq Me^{-\epsilon t}$ for all $t \geq 0$. (Note that not every decaying solution of an ODE has to converge exponentially fast–see the example on page 56). See Figure 1.3.

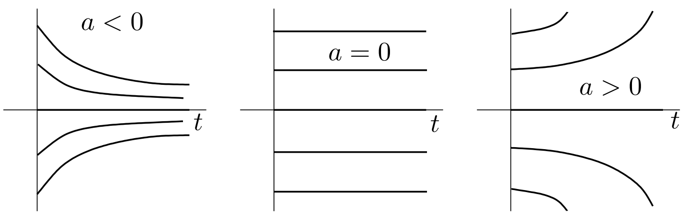

Figure 1.3: Exponential solutions $e^{ta}x_0$.

2° $a = 0$. In this case,

$$x(t) = e^{t0}x(0) = 1x(0) = x(0) \text{ for all } t \geq 0.$$

Thus the solutions are constants, the constant value being the initial value it starts from. See Figure 1.3.

3° $a > 0$. In this case, if the initial condition is zero, the solution is the constant function taking value 0 everywhere. If the initial condition is nonzero, then the solutions 'blow up'. See Figure 1.3.

We would like to have a similar idea about the qualitative behaviour of solutions, but when we have a system of linear differential equations. It turns out that for the system

$$x'(t) = Ax(t),$$

the behaviour of the solutions depends on the eigenvalues of the matrix A. In order to find out why this is so, we first give an expression for the solution of such a linear ODE in the next two sections. We find that the solution is notationally the same as the scalar case discussed in this section: $x(t) = e^{tA}x(0)$, with the little 'a' now replaced by the matrix 'A'! But what do we mean by the exponential of a matrix, e^{tA}? We first introduce this concept in the next section, and subsequently, we will show how it enables us to solve the system $x' = Ax$.

1.5 The matrix exponential

In this section we introduce the exponential of a square matrix A, which is useful for obtaining explicit solutions to the linear system $x'(t) = Ax(t)$. We begin with a few preliminaries concerning vector-valued functions.

A *vector-valued function* $t \mapsto x(t)$ is a vector whose entries are functions of t. Similarly, a *matrix-valued function* $t \mapsto A(t)$ is a matrix whose entries are functions:

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}, \quad A(t) = \begin{bmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{m1}(t) & \dots & a_{mn}(t) \end{bmatrix}.$$

The calculus operations of taking limits, differentiating, and so on are extended to vector-valued and matrix-valued functions by performing the operations on each entry separately. Thus by definition,

$$\lim_{t \to t_0} x(t) = \begin{bmatrix} \lim\limits_{t \to t_0} x_1(t) \\ \vdots \\ \lim\limits_{t \to t_0} x_n(t) \end{bmatrix}.$$

So this limit exists iff $\lim\limits_{t \to t_0} x_i(t)$ exists for all $i \in \{1, \dots, n\}$. Similarly, the derivative of a vector-valued or matrix-valued function is the function obtained by differentiating each entry separately:

$$\frac{dx}{dt}(t) = \begin{bmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{bmatrix}, \quad \frac{dA}{dt}(t) = \begin{bmatrix} a_{11}'(t) & \dots & a_{1n}'(t) \\ \vdots & & \vdots \\ a_{m1}'(t) & \dots & a_{mn}'(t) \end{bmatrix},$$

where $x_i'(t)$ is the derivative of $x_i(t)$, and so on. So $\frac{dx}{dt}$ is defined iff each of the functions $x_i(t)$ is differentiable. The derivative can also be described in vector notation, as

$$\frac{dx}{dt}(t) = \lim_{h \to 0} \frac{x(t+h) - x(t)}{h}. \tag{1.12}$$

Here $x(t+h) - x(t)$ is computed by vector addition and the h in the denominator stands for scalar multiplication by h^{-1}. The limit is obtained by evaluating the limit of each entry separately, as above. So the entries of (1.12) are the derivatives $x_i(t)$. The same is true for matrix-valued functions.

Suppose that analogous to

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \dots, \quad a \in \mathbb{R},$$

we define

$$e^A = I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots, \quad A \in \mathbb{R}^{n\times n}. \tag{1.13}$$

In this section, we will study this matrix exponential, and show that the matrix-valued function

$$e^{tA} = I + tA + \frac{t^2}{2!}A^2 + \frac{t^3}{3!}A^2 + \dots$$

(where t is a variable scalar) can be used to solve the system $x'(t) = Ax(t)$, $x(0) = x_0$: indeed, the solution is given by $x(t) = e^{tA}x_0$.

We begin by stating the following result, which shows that the series in (1.13) converges for any given square matrix A.

Theorem 1.5.1 *The series (1.13) converges for any given square matrix A.*

We have collected the proofs together at the end of this section in order to not break up the discussion.

Since matrix multiplication is relatively complicated, it isn't easy to write down the matrix entries of e^A directly. In particular, the entries of e^A are usually *not* obtained by exponentiating the entries of A. However, one case in which the exponential is easily computed, is when A is a diagonal matrix, say with diagonal entries λ_i. Inspection of the series shows that e^A is also diagonal in this case and that its diagonal entries are e^{λ_i}.

The exponential of a matrix A can also be determined when A is *diagonalizable* , that is, whenever we know a matrix P such that $P^{-1}AP$ is a diagonal matrix D. Then $A = PDP^{-1}$, and using $(PDP^{-1})^k = PD^kP^{-1}$, we obtain

$$\begin{aligned}
e^A &= I + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots \\
&= I + PDP^{-1} + \frac{1}{2!}PD^2P^{-1} + \frac{1}{3!}PD^3P^{-1} + \dots \\
&= PIP^{-1} + PDP^{-1} + \frac{1}{2!}PD^2P^{-1} + \frac{1}{3!}PD^3P^{-1} + \dots \\
&= P\left(I + D + \frac{1}{2!}D^2 + \frac{1}{3!}D^3 + \dots\right)P^{-1} \\
&= Pe^DP^{-1} \\
&= P\begin{bmatrix} e^{\lambda_1} & & \\ 0 & \ddots & \\ & & e^{\lambda_n} \end{bmatrix}P^{-1},
\end{aligned}$$

where $\lambda_1, \dots, \lambda_n$ denote the eigenvalues of A.

Exercise. $(**)$ The set of diagonalizable $n \times n$ complex matrices is *dense* in the set of all $n \times n$ complex matrices, that is, given any $A \in \mathbb{C}^{n\times n}$, there exists a $B \in \mathbb{C}^{n\times n}$ arbitrarily close to A (meaning that $|b_{ij} - a_{ij}|$ can be made arbitrarily small for all $i, j \in \{1, \dots, n\}$) such that B has n distinct eigenvalues.

Hint: Use the fact that every complex $n \times n$ matrix A can be 'upper-triangularised': that is, there exists an invertible complex matrix P such that PAP^{-1} is upper triangular. Clearly the diagonal entries of this new upper triangular matrix are the eigenvalues of A.

In order to use the matrix exponential to solve systems of differential equations, we need to extend some of the properties of the ordinary exponential to it. The most fundamental property is $e^{a+b} = e^a e^b$. This property can be expressed as a formal identity between the two infinite series which are obtained by expanding

$$\begin{aligned} e^{a+b} &= 1 + \tfrac{(a+b)}{1!} + \tfrac{(a+b)^2}{2!} + \dots \quad \text{and} \\ e^a e^b &= \left(1 + \tfrac{a}{1!} + \tfrac{a^2}{2!} + \dots\right)\left(1 + \tfrac{b}{1!} + \tfrac{b^2}{2!} + \dots\right). \end{aligned} \tag{1.14}$$

We cannot substitute matrices into this identity because the commutative law is needed to obtain equality of the two series. For instance, the quadratic terms of (1.14), computed without the commutative law, are $\frac{1}{2}(a^2 + ab + ba + b^2)$ and $\frac{1}{2}a^2 + ab + \frac{1}{2}b^2$. They are not equal unless $ab = ba$. So there is no reason to expect e^{A+B} to equal $e^A e^B$ in general. However, if two matrices A and B happen to commute, the formal identity can be applied.

Theorem 1.5.2 *If $A, B \in \mathbb{R}^{n \times n}$ commute (that is $AB = BA$), then $e^{A+B} = e^A e^B$.*

The proof is at the end of this section. Note that the above implies that e^A is always invertible and in fact its inverse is e^{-A}: Indeed $I = e^{A-A} = e^A e^{-A}$.

Exercises.

1. Give an example of 2×2 matrices A and B such that $e^{A+B} \neq e^A e^B$.
2. Compute e^A, where A is given by $A = \begin{bmatrix} 2 & 3 \\ 0 & 2 \end{bmatrix}$.

   Hint: $A = 2I + \begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$.

We now come to the main result relating the matrix exponential to differential equations. Given an $n \times n$ matrix, we consider the exponential e^{tA}, t being a variable scalar, as a matrix-valued function:

$$e^{tA} = I + tA + \frac{t^2}{2!}A^2 + \frac{t^3}{3!}A^3 + \dots.$$

Theorem 1.5.3 *e^{tA} is a differentiable matrix-valued function of t, and its derivative is Ae^{tA}.*

The proof is at the end of the section.

Theorem 1.5.4 (Product rule.) *Let $A(t)$ and $B(t)$ be differentiable matrix-valued functions of t, of suitable sizes so that their product is defined. Then the matrix product $A(t)B(t)$ is differentiable, and its derivative is*

$$\frac{d}{dt}(A(t)B(t)) = \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt}.$$

The proof is left as an exercise.

Theorem 1.5.5 *The first-order linear differential equation*

$$\frac{dx}{dt}(t) = Ax(t), \quad t \in \mathbb{R}, \quad x(0) = x_0 \tag{1.15}$$

has the unique solution $x(t) = e^{tA}x_0$.

Proof We have

$$\frac{d}{dt}(e^{tA}x_0) = Ae^{tA}x_0,$$

and so $t \mapsto e^{tA}x_0$ solves $\frac{dx}{dt}(t) = Ax(t)$. Furthermore, $x(0) = e^{0A}x_0 = Ix_0 = x_0$.

Finally we show that the solution is unique. Let x be a solution to (1.15). Using the product rule, we differentiate the matrix product $e^{-tA}x(t)$:

$$\frac{d}{dt}(e^{-tA}x(t)) = -Ae^{-tA}x(t) + e^{-tA}Ax(t).$$

From the definition of the exponential, it can be seen that A and e^{-tA} commute, and so the derivative of $e^{-tA}x(t)$ is zero. Therefore, $e^{-tA}x(t)$ is a constant column vector, say C, and $x(t) = e^{tA}C$. As $x(0) = x_0$, we obtain that $x_0 = e^{0A}C$, that is, $C = x_0$. Consequently, $x(t) = e^{tA}x_0$. ■

Thus the matrix exponential enables us to solve the differential equation (1.15). Since direct computation of the exponential can be quite difficult, the above theorem may not be easy to apply in a concrete situation. But if A is a diagonalizable matrix, then the exponential can be computed: $e^A = Pe^DP^{-1}$. To compute the exponential explicitly in all cases requires putting the matrix into Jordan form. But in the next section, we will learn yet another way of computing e^{tA} by using Laplace transforms.

We now go back to prove Theorems 1.5.1, 1.5.2, and 1.5.3.

For want of a more compact notation, we will denote the i,j-entry of a matrix A by A_{ij} here. So $(AB)_{ij}$ will stand for the entry of the matrix product matrix AB, and $(A^k)_{ij}$ for the entry of A^k. With this notation, the i,j-entry of e^A is the sum of the series

$$(e^A)_{ij} = I_{ij} + A_{ij} + \frac{1}{2!}(A^2)_{ij} + \frac{1}{3!}(A^3)_{ij} + \dots.$$

In order to prove that the series for the exponential converges, we need to show that the entries of the powers A^k of a given matrix do not grow too fast, so that the absolute values of the i,j-entries form a bounded (and hence convergent) series. Consider the following function $\|\cdot\|$ on $\mathbb{R}^{n\times n}$:

$$\|A\| := \max\{|A_{ij}| \mid 1 \le i,j \le n\}. \tag{1.16}$$

Thus $|A_{ij}| \le \|A\|$ for all i,j. This is one of several possible "norms" on $\mathbb{R}^{n\times n}$, and it has the following property.

Lemma 1.5.6 *If* $A, B \in \mathbb{R}^{n\times n}$, *then* $\|AB\| \le n\|A\|\|B\|$, *and for all* $k \in \mathbb{N}$, $\|A^k\| \le n^{k-1}\|A\|^k$.

Proof We estimate the size of the i,j-entry of AB:

$$|(AB)_{ij}| = \left|\sum_{k=1}^{n} A_{ik}B_{kj}\right| \le \sum_{k=1}^{n} |A_{ik}||B_{kj}| \le n\|A\|\|B\|.$$

Thus $\|AB\| \le n\|A\|\|B\|$. The second inequality follows from the first inequality by induction. ■

Proof (of Theorem 1.5.1:) To prove that the matrix exponential converges, we show that the series

$$I_{ij} + A_{ij} + \frac{1}{2!}(A^2)_{ij} + \frac{1}{3!}(A^3)_{ij} + \dots$$

is absolutely convergent, and hence convergent. Let $a = n\|A\|$. Then

$$\begin{aligned} |I_{ij}| + |A_{ij}| + \frac{1}{2!}|(A^2)_{ij}| + \frac{1}{3!}|(A^3)_{ij}| + \dots &\le 1 + \|A\| + \frac{1}{2!}n\|A\|^2 + \frac{1}{3!}n^2\|A\|^3 + \dots \\ &= 1 + \frac{1}{n}\left(a + \frac{1}{2!}a^2 + \frac{1}{3!}a^3 + \dots\right) = 1 + \frac{e^a - 1}{n}. \end{aligned}$$

■

Proof (of Theorem 1.5.2:) The terms of degree k in the expansions of (1.14) are

$$\frac{1}{k!}(A+B)^k = \frac{1}{k!}\sum_{r+s=k}\binom{k}{r}A^rB^s \quad \text{and} \quad \sum_{r+s=k}\frac{1}{r!}A^r\frac{1}{s!}B^s.$$

These terms are equal since for all k, and all r, s such that $r + s = k$,

$$\frac{1}{k!}\binom{k}{r} = \frac{1}{r!s!}.$$

Define

$$S_n(A) = I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \dots + \frac{1}{n!}A^n.$$

Then

$$\begin{aligned} S_n(A)S_n(B) &= \left(I + \frac{1}{1!}A + \frac{1}{2!}A^2 + \dots + \frac{1}{n!}A^n\right)\left(I + \frac{1}{1!}B + \frac{1}{2!}B^2 + \dots + \frac{1}{n!}B^n\right) \\ &= \sum_{r,s=0}^{n}\frac{1}{r!}A^r\frac{1}{s!}B^s, \end{aligned}$$

while

$$\begin{aligned} S_n(A+B) &= I + \frac{1}{1!}(A+B) + \frac{1}{2!}(A+B)^2 + \dots + \frac{1}{n!}(A+B)^n \\ &= \sum_{k=0}^{n}\sum_{r+s=k}\frac{1}{k!}\binom{k}{r}A^rB^s = \sum_{k=0}^{n}\sum_{r+s=k}\frac{1}{r!}A^r\frac{1}{s!}B^s. \end{aligned}$$

Comparing terms, we find that the expansion of the partial sum $S_n(A+B)$ consists of the terms in $S_n(A)S_n(B)$ such that $r + s \le n$. We must show that the sum of the remaining terms tends to zero as k tends to ∞.

Lemma 1.5.7 *The series*

$$\sum_k \sum_{r+s=k}\left|\left(\frac{1}{r!}A^r\frac{1}{s!}B^s\right)_{ij}\right|$$

converges for all i, j.

Proof Let $a = n\|A\|$ and $b = n\|B\|$. We estimate the terms in the sum using Lemma 1.5.6:

$$|(A^r B^s)_{ij}| \leq \|A^r B^s\| \leq n\|A^r\|\|B^s\| \leq n(n^{r-1}\|A\|^r)(n^{s-1}\|B\|^s) \leq a^r b^s.$$

Therefore

$$\sum_k \sum_{r+s=k} \left| \left(\frac{1}{r!} A^r \frac{1}{s!} B^s \right)_{ij} \right| \leq \sum_k \sum_{r+s=k} \frac{a^r}{r!} \frac{b^s}{s!} = e^{a+b}.$$

The theorem follows from this lemma because, on the one hand, the i,j-entry of $(S_k(A)S_k(B) - S_k(A+B))_{ij}$ is bounded by

$$\sum_{r+s>k} \left| \left(\frac{1}{r!} A^r \frac{1}{s!} B^s \right)_{ij} \right|.$$

According to the lemma, this sum tends to 0 as k tends to ∞. And on the other hand, $S_k(A)S_k(B) - S_k(A+B)$ tends to $e^A e^B - e^{A+B}$. ■

This completes the proof of Theorem 1.5.2. ■

Proof (of Theorem 1.5.3:) By definition,

$$\frac{d}{dt}e^{tA} = \lim_{h\to 0} \frac{1}{h}(e^{(t+h)A} - e^{tA}).$$

Since the matrices tA and hA commute, we have

$$\frac{1}{h}(e^{(t+h)A} - e^{tA}) = \left(\frac{1}{h}(e^{hA} - I)\right) e^{tA}.$$

So our theorem follows from this lemma:

Lemma 1.5.8 $\displaystyle\lim_{h\to 0} \frac{1}{h}(e^{hA} - I) = A.$

Proof The series expansion for the exponential shows that

$$\frac{1}{h}(e^{hA} - I) - A = \frac{h}{2!}A^2 + \frac{h^2}{3!}A^3 + \dots . \tag{1.17}$$

We estimate this series. Let $a = |h|n\|A\|$. Then

$$\begin{aligned}\left|\left(\frac{h}{2!}A^2 + \frac{h^2}{3!}A^3 + \dots\right)_{ij}\right| &\le \left|\frac{h}{2!}(A^2)_{ij}\right| + \left|\frac{h^2}{3!}(A^3)_{ij}\right| + \dots \\ &\le \frac{1}{2!}|h|n\|A\|^2 + \frac{1}{3!}|h|^2n^2\|A\|^3 + \dots \\ &= \|A\|\left(\frac{1}{2!}a + \frac{1}{3!}a^2 + \dots\right) = \frac{\|A\|}{a}(e^a - 1 - a) = \|A\|\left(\frac{e^a-1}{a} - 1\right).\end{aligned}$$

Note that $a \to 0$ as $h \to 0$. Since the derivative of e^x is e^x, $\displaystyle\lim_{a\to 0}\frac{e^a - 1}{a} = \left.\frac{d}{dx}e^x\right|_{x=0} = e^0 = 1.$

So (1.17) tends to 0 with $h \to 0$. ■

This completes the proof of Theorem 1.5.3. ■

Exercises.

1. $(*)$ If $A \in \mathbb{R}^{n\times n}$, then show that $\|e^A\| \le e^{n\|A\|}$. (In particular, for all $t \ge 0$, $\|e^{tA}\| \le e^{tn\|A\|}$.)
2. (a) Let $n \in \mathbb{N}$. Show that there exists a constant C (depending only on n) such that if $A \in \mathbb{R}^{n\times n}$, then for all $v \in \mathbb{R}^n$, $\|Av\| \le C\|A\|\|v\|$. (Throughout this course, by $\|v\|$, where v is a vector in $\mathbb{R}^n$, we mean its *Euclidean norm*, that is, the square root of the sum of the squares of the n components of v.)

   (b) Show that if λ is an eigenvalue of A, then $|\lambda| \le n\|A\|$.
3. (a) $(*)$ Show that if λ is an eigenvalue of A and v is a corresponding eigenvector, then v is also an eigenvector of e^A corresponding to the eigenvalue e^λ of e^A.

   (b) Solve $x'(t) = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix} x(t)$, $x(0) = x_0$, when

   (i) $x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, (ii) $x_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$, and (iii) $x_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

   HINT: In parts (i) and (ii), observe that the initial condition is an eigenvector of the square matrix in question, and in part (iii), express the initial condition as a combination of the initial conditions from the previous two parts.

4. (∗) Prove that $e^{tA^\top} = (e^{tA})^\top$. (Here $M^\top$ denotes the transpose of the matrix M.)

5. (∗) Let $A \in \mathbb{R}^{n\times n}$, and let $\mathscr{S} = \{x : \mathbb{R} \to \mathbb{R}^n \mid \forall t \in \mathbb{R},\ x'(t) = Ax(t)\}$. In this exercise we will show that $\mathscr{S}$ is a finite dimensional vector space with dimension n.

   (a) Let $C(\mathbb{R};\mathbb{R}^n)$ denote the vector space of all functions $f : \mathbb{R} \to \mathbb{R}^n$ with pointwise addition and scalar multiplication. Show that $\mathscr{S}$ is a subspace of $C(\mathbb{R};\mathbb{R}^n)$.

   (b) Let $e_1, \dots, e_n$ denote the standard basis vectors in $\mathbb{R}^n$. By Theorem 1.5.5, we know that for each $k \in \{1, \dots, n\}$, there exists a unique solution to the initial value problem $x'(t) = Ax(t)$, $t \in \mathbb{R}$, $x(0) = e_k$. Denote this unique solution by f_k. Thus we obtain the set of functions $f_1, \dots, f_n \in \mathscr{S}$. Prove that $\{f_1, \dots, f_n\}$ is linearly independent.

   Hint: Set $t = 0$ in $\alpha_1 f_1 + \cdots + \alpha_n f_n = 0$.

   (c) Show that $\mathscr{S} = \operatorname{span}\{f_1, \dots, f_n\}$, and conclude that $\mathscr{S}$ is a finite dimensional vector space of dimension n.

1.6 Computation of e^{tA}

In the previous section, we saw that the computation of e^{tA} is easy if the matrix A is diagonalizable. However, not all matrices are diagonalizable. For example, consider the matrix

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The eigenvalues of this matrix are both 0, and so if it were diagonalizable, then the diagonal form will be the zero matrix, but then if there did exist an invertible P such that $P^{-1}AP$ is this zero matrix, then clearly A should be zero, which it is not!

In general, however, every matrix has what is called a *Jordan canonical form*, that is, there exists an invertible P such that $P^{-1}AP = D + N$, where D is diagonal, N is *nilpotent* (that is, there exists a $n \geq 0$ such that $N^{n+1} = 0$), and D and N commute. Then one can compute the exponential of A:

$$e^{tA} = Pe^{tD}\left(I + tN + \frac{1}{2!}t^2N^2 + \cdots + \frac{1}{n!}t^nN^n\right)P^{-1}.$$

However, the algorithm for computing the P taking A to the Jordan form requires some sophisticated linear algebra. So we give a different procedure for calculating e^{tA} below, using Laplace transforms. First we will prove the following theorem.

Theorem 1.6.1 *For large enough s,* $\displaystyle\int_0^\infty e^{-st}e^{tA}dt = (sI - A)^{-1}.$

Proof First choose a s_0 large enough so that $s_0 > n\|A\|$. Then for all $s > s_0$, we have

$$\begin{aligned}
\int_0^\infty e^{-ts}e^{tA}dt &= \int_0^\infty e^{-t(sI-A)}dt \\
&= \int_0^\infty (sI-A)^{-1}(sI-A)e^{-t(sI-A)}dt \\
&= (sI-A)^{-1}\int_0^\infty (sI-A)e^{-t(sI-A)}dt \\
&= (sI-A)^{-1}\int_0^\infty -\frac{d}{dt}e^{-t(sI-A)}dt \\
&= (sI-A)^{-1}\left(-e^{-ts}e^{tA}\Big|_{t=0}^{t=\infty}\right) \\
&= (sI-A)^{-1}(0+I) \\
&= (sI-A)^{-1}.
\end{aligned}$$

(In the above, we used the Exercise 1 on page 16, which gives $\|e^{tA}\| \le e^{tn\|A\|} \le e^{ts_0} = e^{ts}e^{t(s_0-s)}$, and so $\|e^{-ts}e^{tA}\| \le e^{t(s_0-s)}$. Also, we have used Exercise 2b, which gives invertibility of $sI - A$.) ■

If s is not an eigenvalue of A, then $sI - A$ is invertible, and *Cramer's rule*[1] says that,

$$(sI-A)^{-1} = \frac{1}{\det(sI-A)}\text{adj}(sI-A).$$

Here $\text{adj}(sI - A)$ denotes the *adjoint* of the matrix $sI - A$, which is defined as follows: its (i, j)th entry is obtained by multiplying $(-1)^{i+j}$ and the determinant of the matrix obtained by deleting the jth row and ith column of $sI - A$. Thus we see that each entry of $\text{adj}(sI - A)$ is a polynomial in s whose degree is at most $n - 1$. (Here n denotes the size of A–that is, A is a $n \times n$ matrix.)

Consequently, each entry m_{ij} of $(sI - A)^{-1}$ is a rational function, in other words, it is ratio of two polynomials (in s) p_{ij} and $q := \det(sI - A)$:

$$m_{ij} = \frac{p_{ij}(s)}{q(s)}$$

Also from the above, we see that $\deg(p_{ij}) \le \deg(q) - 1$. From the fundamental theorem of algebra, we know that the monic polynomial q can be factored as

$$q(s) = (s-\lambda_1)^{m_1}\dots(s-\lambda_k)^{m_k},$$

where $\lambda_1, \dots, \lambda_k$ are the distinct eigenvalues of $q(s) = \det(sI - A)$, with the algebraic multiplicities $m_1, \dots, m_k$.

By the "partial fraction expansion" one learns in calculus, it follows that one can find suitable coefficients for a decomposition of each rational entry of $(sI - A)^{-1}$ as follows:

$$m_{ij} = \sum_{l=1}^{k}\sum_{r=1}^{m_k}\frac{C_{l,r}}{(s-\lambda_l)^r}.$$

Thus if $f_{ij}(t)$ denotes the (i, j)th entry of e^{tA}, then its Laplace transform will be an expression of the type m_{ij} given above. Now it turns out that this determines the f_{ij}, and this is the content of the following result.

[1]For a proof, see for instance Artin [3].

Theorem 1.6.2 *Let $a \in \mathbb{C}$ and $n \in \mathbb{N}$. If f is a continuous function defined on $[0, \infty)$, and if there exists a s_0 such that for all $s > s_0$,*

$$F(s) := \int_0^\infty e^{-st} f(t)dt = \frac{1}{(s-a)^n},$$

then

$$f(t) = \frac{1}{(n-1)!} t^{n-1} e^{ta} \quad \text{for all } t \geq 0.$$

Proof The proof is beyond the scope of this course, but we refer the interested reader to Exercise 11.38 on page 342 of Apostol [1]. ■

So we have a procedure for computing e^{tA}: form the matrix $sI - A$, compute its inverse (as a rational matrix), perform a partial fraction expansion of each of its entry, and take the inverse Laplace transform of each elementary fraction. Sometimes, the partial fraction expansion may be avoided, by making use of the following corollary (which can be obtained from Theorem 1.6.2, by a partial fraction expansion!).

Corollary 1.6.3 *Let f be a continuous function defined on $[0, \infty)$, and let there exist a s_0 such that for all $s > s_0$, F defined by*

$$F(s) := \int_0^\infty e^{-st} f(t)dt,$$

is one of the functions given in the first column below. Then f is given by the corresponding entry in the second column.

F	f
$\frac{b}{(s-a)^2 + b^2}$	$e^{ta} \sin(bt)$
$\frac{s-a}{(s-a)^2 + b^2}$	$e^{ta} \cos(bt)$
$\frac{b}{(s-a)^2 - b^2}$	$e^{ta} \sinh(bt)$
$\frac{s-a}{(s-a)^2 - b^2}$	$e^{ta} \cosh(bt)$

Example. If $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $sI - A = \begin{bmatrix} s & -1 \\ 0 & s \end{bmatrix}$, and so

$$(sI - A)^{-1} = \frac{1}{s^2} \begin{bmatrix} s & 1 \\ 0 & s \end{bmatrix} = \begin{bmatrix} \frac{1}{s} & \frac{1}{s^2} \\ 0 & \frac{1}{s} \end{bmatrix}.$$

By using Theorem 1.6.2 ('taking the inverse Laplace transform'), we obtain $e^{tA} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix}$. ◊

Exercises.

1. Compute e^{tA}, when $A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$.

2. Compute e^{tA}, for the 'Jordan block' $A = \begin{bmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{bmatrix}$.

   Remark: In general, if

$$\text{if } A = \begin{bmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \lambda \end{bmatrix}, \text{ then } e^{tA} = e^{\lambda t} \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \dots & \frac{t^{n-1}}{(n-1)!} \\ & \ddots & \ddots & & \vdots \\ & & & & \vdots \\ & & & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{bmatrix}.$$

3. (a) Compute e^{tA}, when $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

   (b) Find the solution to

$$x'' + kx = 0, \quad x(0) = 1, \quad x'(0) = 0.$$

   (Here k is a fixed positive constant.)
   HINT: Introduce the state variables $x_1 = \sqrt{k}x$ and $x_2 = x'$.
   Suppose that $k = 1$, and find $(x(t))^2 + (x'(t))^2$. What do you observe? If one identifies $(x(t), x'(t))$ with a point in the plane at time t, then how does this point move with time?

4. Suppose that A is a 2×2 matrix such that

$$e^{tA} = \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}, \quad t \in \mathbb{R}.$$

   Find A.

1.7 Stability considerations

Just as in the scalar example $x' = ax$, where we saw that the sign of (the real part of) a allows us to conclude the behaviour of the solution as $x \to \infty$, it turns out that by looking at the real parts of the eigenvalues of the matrix A one can say similar things in the case of the system $x' = Ax$. We will study this in this section.

We begin by proving the following result.

Lemma 1.7.1 *Suppose that $\lambda \in \mathbb{C}$ and k is a nonnegative integer. For every $\omega > \text{Re}(\lambda)$, there exists a $M_\omega > 0$ such that for all $t \geq 0$, $|t^k e^{\lambda t}| \leq M_\omega e^{\omega t}$.*

Proof We have

$$e^{(\omega-\mathrm{Re}(\lambda))t} = \sum_{n=0}^{\infty} \frac{(\omega-\mathrm{Re}(\lambda))^n t^n}{n!} \geq \frac{(\omega-\mathrm{Re}(\lambda))^k t^k}{k!},$$

and so $t^k e^{(\mathrm{Re}(\lambda)-\omega)t} \leq M_\omega$, where

$$M_\omega := \frac{k!}{(\omega-\mathrm{Re}(\lambda))^k} > 0.$$

Consequently, for $t \geq 0$, $|t^k e^{\lambda t}| = t^k e^{\mathrm{Re}(\lambda)t} = t^k e^{(\mathrm{Re}(\lambda)-\omega)t} e^{\omega t} \leq M_\omega e^{\omega t}$. ■

In the sequel, we denote the set of eigenvalues of A by $\sigma(A)$, sometimes referred to as the *spectrum* of A.

Theorem 1.7.2 *Let $A \in \mathbb{R}^{n\times n}$.*

1. *Every solution of $x' = Ax$ tends to zero as $t \to \infty$ iff for all $\lambda \in \sigma(A)$, $\mathrm{Re}(\lambda) < 0$. Moreover, in this case, the solutions converge exponentially to 0: there exist $\epsilon > 0$ and $M > 0$ such that for all $t \geq 0$, $\|x(t)\| \leq Me^{-\epsilon t}\|x(0)\|$.*
2. *If there exists a $\lambda \in \sigma(A)$ such that $\mathrm{Re}(\lambda) > 0$, then for every $\delta > 0$, there exists a $x_0 \in \mathbb{R}^n$ with $\|x_0\| < \delta$, such that the unique solution to $x' = Ax$ with initial condition $x(0) = x_0$ satisfies $\|x(t)\| \to \infty$ as $t \to \infty$.*

Proof 1. We use Theorem 1.6.1. From Cramer's rule, it follows that each entry in $(sI - A)^{-1}$ is a rational function with the denominator equal to the characteristic polynomial of A, and then by using a partial fraction expansion and Theorem 1.6.2, it follows that each entry in e^{tA} is a linear combination of terms of the form $t^k e^{\lambda t}$, where k is a nonnegative integer and $\lambda \in \sigma(A)$. By Lemma 1.7.1, we conclude that if each eigenvalue of A has real part < 0, then there exist positive constants M and ϵ such that for all $t \geq 0$, $\|e^{tA}\| < Me^{-\epsilon t}$.

On the other hand, if each solution tends to 0 as $t \to \infty$, then in particular, if $v \in \mathbb{R}^n$ is an eigenvector[2] corresponding to eigenvalue λ, then with initial condition $x(0) = v$, we have $x(t) = e^{tA}v = e^{\lambda t}v$, and so $\|x(t)\| = e^{\mathrm{Re}(\lambda)t}\|v\| \stackrel{t\to\infty}{\longrightarrow} 0$, and so it must be the case that $\mathrm{Re}(\lambda) < 0$.

2. Let $\lambda \in \sigma(A)$ be such that $\mathrm{Re}(\lambda) > 0$, and let $v \in \mathbb{R}^n$ be a corresponding eigenvector[3]. Given $\delta > 0$, define $x_0 = \frac{\delta}{2\|v\|}v \in \mathbb{R}^n$. Then $\|x_0\| = \frac{\delta}{2} < \delta$, and the unique solution x to $x' = Ax$ with initial condition $x(0) = x_0$ satisfies $\|x(t)\| = \frac{\delta}{2}e^{\mathrm{Re}(\lambda)t} \to \infty$ as $t \to \infty$. ■

In the case when we have eigenvalues with real parts equal to zero, then a more careful analysis is required and the boundedness of solutions depends on the algebraic/geometric multiplicity of the eigenvalues with zero real parts. We will not give a detailed analysis, but consider two examples which demonstrate that the solutions may or may not remain bounded.

[2]With a complex eigenvalue, this vector is not in $\mathbb{R}^n$! But the proof can be modified so as to still yield the desired conclusion.

[3]See the previous footnote!

Examples. Consider the system $x' = Ax$, where

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then the system trajectories are constants $x(t) \equiv x(0)$, and so they are bounded.

On the other hand if

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

then

$$e^{tA} = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix},$$

and so with the initial condition $x(0) = \delta \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, we have $\|x(t)\| = \delta\sqrt{1+t^2} \to \infty$ as $t \to \infty$ for all $\delta > 0$. So even if one starts arbitrarily close to the origin, the solution can become unbounded. $\Diamond$

Exercises.

1. Determine if all solutions of $x' = Ax$ are bounded, and if so if all solutions tend to 0 as $t \to \infty$.

   (a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$

   (b) $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$

   (c) $\begin{bmatrix} 1 & 1 & 0 \\ 0 & -2 & -1 \\ 0 & 0 & -1 \end{bmatrix}$

   (d) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & -2 \end{bmatrix}$

   (e) $\begin{bmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 1 & 0 & -1 \end{bmatrix}$

   (f) $\begin{bmatrix} -1 & 0 & -1 \\ 0 & -2 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

2. For what values of $\alpha \in \mathbb{R}$ can we conclude that all solutions of the system $x' = Ax$ will be bounded for $t \geq 0$, if $A = \begin{bmatrix} \alpha & 1+\alpha \\ -(1+\alpha) & \alpha \end{bmatrix}$?

Chapter 2

Phase plane analysis

2.1 Introduction

In the preceding chapter, we learnt how one can solve a system of linear differential equations. However, the equations that arise in most practical situations are inherently nonlinear, and typically it is impossible to solve these explicitly. Nevertheless, sometimes it is possible to obtain an idea of what its solutions look like (the "qualitative behaviour"), and we learn one such method in this chapter, called *phase plane analysis*.

Phase plane analysis is a graphical method for studying 2D autonomous systems. This method was introduced by mathematicians (among others, Henri Poincaré) in the 1890s.

The basic idea of the method is to generate in the state space motion trajectories corresponding to various initial conditions, and then to examine the qualitative features of the trajectories. As a graphical method, it allows us to visualise what goes on in a nonlinear system starting from various initial conditions, without having to solve the nonlinear equations analytically. Thus, information concerning stability and other motion patterns of the system can be obtained. In this chapter, we learn the basic tools of the phase plane analysis.

2.2 Concepts of phase plane analysis

2.2.1 Phase portraits

The phase plane method is concerned with the graphical study of 2 dimensional autonomous systems:

$$\begin{aligned} x_1'(t) &= f_1(x_1(t), x_2(t)), \\ x_2'(t) &= f_2(x_1(t), x_2(t)), \end{aligned} \tag{2.1}$$

where x_1 and x_2 are the states of the system, and f_1 and f_2 are nonlinear functions from $\mathbb{R}^2$ to $\mathbb{R}$. Geometrically, the state space is a plane, and we call this plane the *phase plane*.

Given a set of initial conditions $x(0) = x_0$, we denote by x the solution to the equation (2.1). (We assume throughout this chapter that given an initial condition there exists a *unique* solution *for all* $t \geq 0$: this is guaranteed under mild assumptions on f_1, f_2, and we will learn more about this in Chapter 4.) With time t varied from 0 to ∞, the solution $t \mapsto x(t)$ can be represented geometrically as a curve in the phase plane. Such a curve is called a *(phase plane) trajectory.* A family of phase plane trajectories corresponding to various initial conditions is called a *phase portrait* of the system. From the assumption about the existence of solution, we know that from each point in the phase plane there passes a curve, and from the uniqueness, we know that there can be only one such curve. Thus no two trajectories in the phase plane can intersect, for if they did intersect at a point, then with that point as the initial condition, we would have two solutions[1], which is a contradiction!

To illustrate the concept of a phase portrait, let us consider the following simple system.

Example. Consider the system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -x_1. \end{aligned}$$

Thus the system is a linear ODE $x' = Ax$ with $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. Then $e^{tA} = \begin{bmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{bmatrix}$, and so if the initial condition expressed in polar coordinates is

$$\begin{bmatrix} x_{10} \\ x_{20} \end{bmatrix} = \begin{bmatrix} r_0 \cos\theta_0 \\ r_0 \sin\theta_0 \end{bmatrix},$$

then it can be seen that the solution is

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} r_0 \cos(\theta_0 - t) \\ r_0 \sin(\theta_0 - t) \end{bmatrix}, \quad t \geq 0. \tag{2.2}$$

We note that

$$(x_1(t))^2 + (x_2(t))^2 = r_0^2,$$

which represents a circle in the phase plane. Corresponding to different initial conditions, circles of different radii can be obtained, and from (2.2), it is easy to see that the motion is clockwise. Plotting these circles on the phase plane, we obtain a phase portrait as shown in the Figure 2.1.

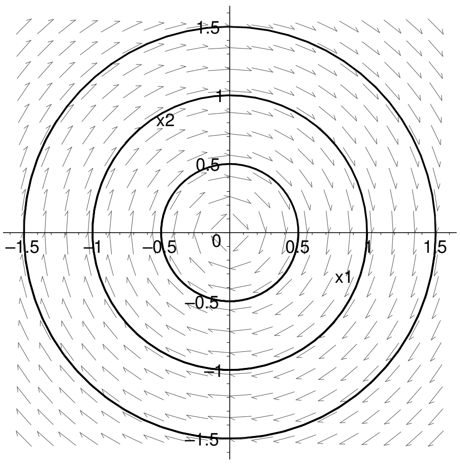

Figure 2.1: Phase portrait.

We see that the trajectories neither converge to the origin nor diverge to infinity. They simply circle around the origin. ◊

[1]Here we are really running the differential equation *backwards* in time, but then we can make the change of variables $\tau = -t$.

2.2.2 Singular points

An important concept in phase plane analysis is that of a singular point.

Definition. A *singular point* of the system $\begin{cases} x_1' &= f_1(x_1, x_2) \\ x_2' &= f_2(x_1, x_2), \end{cases}$ is a point (x_{1*}, x_{2*}) in the phase plane such that $f_1(x_{1*}, x_{2*}) = 0$ and $f_2(x_{1*}, x_{2*}) = 0$.

Such a point is also sometimes called an *equilibrium point* (see Chapter 3), that is, a point where the system states can stay forever: if we start with this initial condition, then the unique solution is $x_1(t) = x_{1*}$ and $x_2(t) = x_{2*}$. So through that point in the phase plane, only the 'trivial curve' comprising just that point passes.

For a linear system $x' = Ax$, if A is invertible, then the only singular point is $(0, 0)$, and if A is not invertible, then all the points from the kernel of A are singular points. So in the case of linear systems, either there is only one equilibrium point, or infinitely many singular points, none of which is then isolated. But in the case of nonlinear systems, there can be more than one isolated singular point, as demonstrated in the following example.

Example. Consider the system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\frac{1}{2}x_2 - 2x_1 - x_1^2, \end{aligned}$$

whose phase portrait is shown in Figure 2.2. The system has two singular points, one at $(0, 0)$,

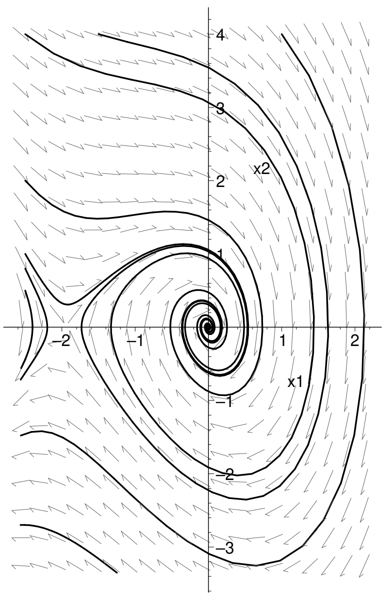

Figure 2.2: Phase portrait.

and the other at $(-2, 0)$. The motion patterns of the system trajectories starting in the vicinity of the two singular points have different natures. The trajectories move towards the point $(0, 0)$, while they move away from $(-2, 0)$. ♢

One may wonder why an equilibrium point of a 2D system is called a singular point. To answer this, let us examine the slope of the phase trajectories. The slope of the phase trajectory at time t is given by

$$\frac{dx_2}{dx_1} = \frac{\frac{dx_2}{dt}}{\frac{dx_1}{dt}} = \frac{f_2(x_1, x_2)}{f_1(x_1, x_2)}.$$

When both f_1 and f_2 are zero at a point, this slope is undetermined, and this accounts for the adjective 'singular'.

Singular points are important features in the phase plane, since they reveal some information about the system. For nonlinear systems, besides singular points, there may be more complex features, such as limit cycles. These will be discussed later in this chapter.

Note that although the phase plane method is developed primarily for 2D systems, it can be also applied to the analysis of nD systems in general, but the graphical study of higher order systems is computationally and geometrically complex. On the other hand with 1D systems, the phase "plane" is reduced to the real line. We consider an example of a 1D system below.

Example. Consider the system $x' = -x + x^3$.

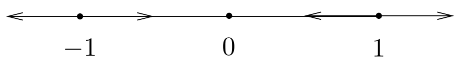

Figure 2.3: Phase portrait of the system $x' = -x + x^3$.

The singular points are determined by the equation

$$-x + x^3 = 0,$$

which has three real solutions, namely -1, 0 and 1. The phase portrait is shown in Figure 2.3. Indeed, for example if we consider the solution to

$$x'(t) = -x(t) - (x(t))^3, \quad t \geq t_0, \quad x(t_0) = x_0,$$

with $0 < x_0 < 1$, then we observe that

$$x'(t_0) = -x_0 - x_0^3 = -\underbrace{x_0}_{>0}\underbrace{(1 - x_0^2)}_{>0} < 0,$$

and this means that $t \mapsto x(t)$ is decreasing, and so the "motion" starting from x_0 is towards the left. This explains the direction of the arrow for the region $0 < x < 1$ in Figure 2.3. $\Diamond$

Exercises.

1. Locate the singular points of the following systems.

(a) $\begin{cases} x_1' &= x_2, \\ x_2' &= \sin x_1. \end{cases}$

(b) $\begin{cases} x_1' &= x_1 - x_2, \\ x_2' &= x_2^2 - x_1. \end{cases}$

(c) $\begin{cases} x_1' &= x_1^2(x_2 - 1), \\ x_2' &= x_1 x_2. \end{cases}$

(d) $\begin{cases} x_1' &= x_1^2(x_2-1), \\ x_2' &= x_1^2 - 2x_1x_2 - x_2^2. \end{cases}$

(e) $\begin{cases} x_1' &= \sin x_2, \\ x_2' &= \cos x_1. \end{cases}$

2. Sketch the following parameterised curves in the phase plane.

   (a) $(x_1, x_2) = (a\cos t, b\sin t)$, where $a > 0$, $b > 0$.

   (b) $(x_1, x_2) = (ae^t, be^{-2t})$, where $a > 0$, $b > 0$.

3. Draw phase portraits of the following 1D systems.

   (a) $x' = x^2$.

   (b) $x' = e^x$.

   (c) $x' = \cosh x$.

   (d) $x' = \sin x$.

   (e) $x' = \cos x - 1$.

   (f) $x' = \sin(2x)$.

4. Consider a 2D autonomous system for which there exists a unique solution for every initial condition in $\mathbb{R}^2$ for all $t \in \mathbb{R}$.

   (a) Show that if (x_1, x_2) is a solution, then for any $T \in \mathbb{R}$, the shifted functions (y_1, y_2) given by

   $$\begin{aligned} y_1(t) &= x_1(t+T), \\ y_2(t) &= x_2(t+T), \end{aligned}$$

   ($t \in \mathbb{R}$) is also a solution.

   (b) (∗) Can $x(t) = \left(\frac{2\cos t}{1+(\sin t)^2}, \frac{\sin(2t)}{1+(\sin t)^2}\right)0$, $t \in \mathbb{R}$, be the solution of such a 2D autonomous system?

   HINT: By the first part, we know that $t \mapsto y_1(t) := x(t+\pi/2)$ and $t \mapsto y_2(t) := x(t+3\pi/2)$ are also solutions. Check that $y_1(0) = y_2(0)$. Is $y_1 \equiv y_2$? What does this say about uniqueness of solutions starting from a given initial condition?

   (c) Using Maple, sketch the curve

   $$t \mapsto \left(\frac{2\cos t}{1+(\sin t)^2}, \frac{\sin(2t)}{1+(\sin t)^2}\right).$$

   (This curve is called the *lemniscate*.)

5. Consider the ODE (1.6) from Exercise 2 on page 6.

   (a) Using Maple, find the singular points (approximately).

   (b) Draw a phase portrait in the region $x \geq 0$.

6. A simple model for a national economy is given by

$$\begin{aligned} I' &= I - \alpha C \\ C' &= \beta(I - C - G), \end{aligned}$$

where

I	denotes the national income,
C	denotes the rate of consumer spending, and
G	denotes the rate of government expenditure.

The model is restricted to $I, C, G \geq 0$, and the constants α, β satisfy $\alpha > 1$, $\beta \geq 1$.

(a) Suppose that the government expenditure is related to the national income according to $G = G_0 + kI$, where G_0 and k are positive constants. Find the range of positive k's for which there exists an equilibrium point such that I, C, G are nonnegative.

(b) Let $k = 0$, and let (I_0, C_0) denote the equilibrium point. Introduce the new variables $I_1 = I - I_0$ and $C_1 = C - C_0$. Show that (I_1, C_1) satisfy a linear system of equations:

$$\begin{bmatrix} I_1' \\ C_1' \end{bmatrix} = \begin{bmatrix} 1 & -\alpha \\ \beta & -\beta \end{bmatrix} \begin{bmatrix} I_1 \\ C_1 \end{bmatrix}.$$

If $\beta = 1$ and $\alpha = 2$, then conclude that in fact the economy oscillates.

2.3 Constructing phase portraits

Phase portraits can be routinely generated using computers, and this has spurred many advances in the study of complex nonlinear dynamic behaviour. Nevertheless, in this section, we learn a few techniques in order to be able to roughly sketch the phase portraits. This is useful for instance in order to verify the plausibility of computer generated outputs. We describe two methods: one involves the analytic solution of differential equations. If an analytic solution is not available, the other tool, called the method of isoclines, is useful.

2.3.1 Analytic method

There are two techniques for generating phase portraits analytically. One is to first solve for x_1 and x_2 explicitly as functions of t, and then to eliminate t, as we had done in the example on page 26.

The other analytic method does not involve an explicit computation of the solutions as functions of time, but instead, one solves the differential equation

$$\frac{dx_2}{dx_1} = \frac{f_1(x_1, x_2)}{f_2(x_1, x_2)}.$$

Thus given a trajectory $t \mapsto (x_1(t), x_2(t))$, we eliminate the t by setting up a differential equation for the derivative of the second function 'with respect to the first one', not involving the t, and by solving this differential equation. We illustrate this in the same example.

Example. Consider the system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -x_1. \end{aligned}$$

We have $\frac{dx_2}{dx_1} = \frac{-x_1}{x_2}$, and so $x_2 \frac{dx_2}{dx_1} = -x_1$. Thus $\frac{d}{dx_1}\left(\frac{1}{2}x_2^2\right) = x_2 \frac{dx_2}{dx_1} = -x_1$.

Integrating with respect to x_1, and using the fundamental theorem of calculus, we obtain $x_2^2 + x_1^2 = C$. This equation describes a circle in the (x_1, x_2)-plane. Thus the trajectories satisfy

$$(x_1(t))^2 + (x_2(t))^2 = C = (x_1(0))^2 + (x_2(0))^2, \quad t \geq 0,$$

and they are circles. We note that when $x_1(0)$ belongs to the right half plane, then $x_2'(0) = -x_1(0) < 0$, and so $t \mapsto x_2(t)$ should be decreasing. Thus we see that the motion is clockwise, as shown in Figure 2.1. ♢

Exercises.

1. Sketch the phase portrait of $\left\{ \begin{array}{rcl} x_1' &=& x_2, \\ x_2' &=& x_1. \end{array} \right.$
2. Sketch the phase portrait of $\left\{ \begin{array}{rcl} x_1' &=& -2x_2, \\ x_2' &=& x_1. \end{array} \right.$
3. (a) Sketch the curve curve $y(x) = x(A + B \log|x|)$ where A, B are constants and $B > 0$.
   (b) (∗) Sketch the phase portrait of $\left\{ \begin{array}{rcl} x_1' &=& x_1 + x_2, \\ x_2' &=& x_2. \end{array} \right.$
   HINT: Solve the system, and try eliminating t.

2.3.2 The method of isoclines

At a point (x_1, x_2) in the phase plane, the slope of the tangent to the trajectory is $\frac{dx_2}{dx_1} = \frac{f_2(x_1,x_2)}{f_1(x_1,x_2)}$. An *isocline* is a curve in $\mathbb{R}^2$ defined by $\frac{f_2(x_1,x_2)}{f_1(x_1,x_2)} = \alpha$, where α is a real number. This means that if we look at all the trajectories that pass through various points on the same isocline, then all of these trajectories have the same slope (equal to α) on the points of this isocline. To obtain trajectories from the isoclines, we assume that the tangent slopes are locally constant. The method of constructing the phase portrait using isoclines is thus the following:

STEP 1. For various values of α, construct the corresponding isoclines. Along an isocline, draw small line segments with slope α. In this manner a field of directions is obtained.

STEP 2. Since the tangent slopes are locally constant, we can construct a phase plane trajectory by connecting a sequence line segments.

We illustrate the method by means of two examples.

Example. Consider the system $\left\{ \begin{array}{rcl} x_1' &=& x_2, \\ x_2' &=& -x_1. \end{array} \right.$ The slope is given by $\frac{dx_2}{dx_1} = \frac{-x_1}{x_2}$, and so the isocline corresponding to slope α is $x_1 + \alpha x_2 = 0$, and these points lie on a straight line. By taking different values for α, a set of isoclines can be drawn, and in this manner a field of directions of

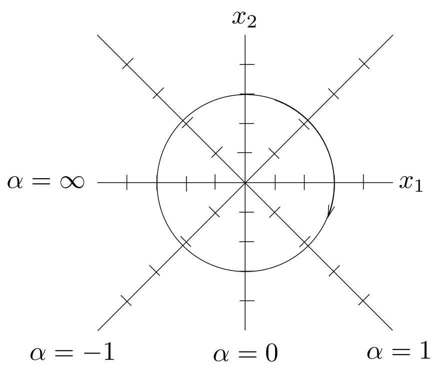

Figure 2.4: Method of isoclines.

tangents to the trajectories are generated, as shown in Figure 2.4, and the trajectories in the phase portrait are circles. If $x_1 > 0$, then $x_2'(0) = -x_1(0) < 0$, and so the motion is clockwise. ♢

Let us now use the method of isoclines to study a nonlinear equation.

Example. *(van der Pol equation)* Consider the differential equation

$$y'' + \mu(y^2 - 1)y' + y = 0. \tag{2.3}$$

By introducing the variables $x_1 = y$ and $x_2 = y'$, we obtain the following 2D system:

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= -\mu(x_1^2 - 1)x_2 - x_1. \end{aligned}$$

An isocline of slope α is defined by $\dfrac{-\mu(x_1^2 - 1)x_2 - x_1}{x_2} = \alpha$, that is, the points on the curve $x_2 = \dfrac{x_1}{(\mu - \mu x_1^2) - \alpha}$ all correspond to the same slope α of tangents to trajectories.

We take the value of $\mu = \frac{1}{2}$. By taking different values for α, different isoclines can be obtained, and short line segments can be drawn on the isoclines to generate a field of directions, as shown in Figure 2.5. The phase portrait can then be obtained, as shown.

It is interesting to note that from the phase portrait, one is able to guess that there exists a closed curve in the phase portrait, and the trajectories starting from both outside and inside seem to converge to this curve[2]. ♢

Exercise. Using the method of isoclines, sketch a phase portrait of the system $\left\{ \begin{aligned} x_1' &= x_2, \\ x_2' &= x_1. \end{aligned} \right.$

2.3.3 Phase portraits using Maple

We consider a few examples in order to illustrate how one can make phase portraits using Maple.

[2]This is also expected based on physical considerations: the van der Pol equation arises from electric circuits containing vacuum tubes, where for small oscillations, energy is fed into the system, while for large oscillations, energy is taken out of the system–in other words, large oscillations will be damped, while for small oscillations, there is 'negative damping' (that is energy is fed into the system). So one can expect that such a system will approach some periodic behaviour, which will appear as a closed curve in the phase portrait.

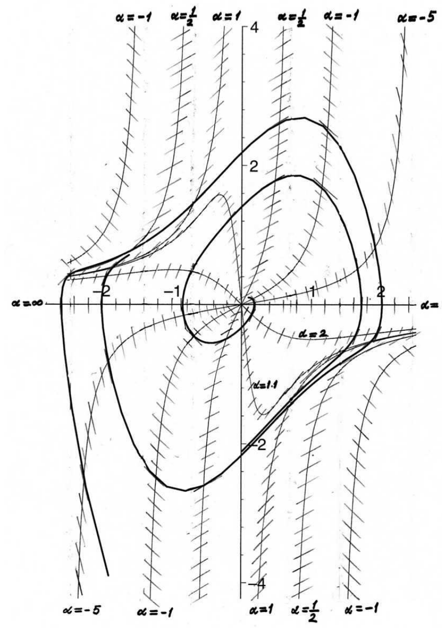

Figure 2.5: Method of isoclines.

Consider the ODE system: $x_1'(t) = x_2(t)$ and $x_2(t) = -x_2(t)$. We can plot x_1 against x_2 by using `DEplot`. Consider for example:

```
> with(DEtools) :
> ode3a := diff(x1(t), t) = x2(t); ode3b := diff(x2(t), t) = −x1(t);
> DEplot({ode3a, ode3b}, {x1(t), x2(t)}, t = 0..10, x1 = −2..2, x2 = −2..2,
  [[x1(0) = 1, x2(0) = 0]], stepsize = 0.01, linecolour = black);
```

The resulting plot is shown in Figure 2.6. The arrows show the direction field.

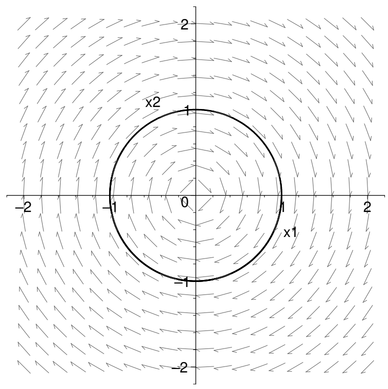

Figure 2.6: Phase portrait for the ODE system $x_1' = x_2$ and $x_2' = -x_1$.

By including some more trajectories, we can construct a phase portrait in a given region, as shown in the following example.

Example. Consider the system $\begin{cases} x_1' &= -x_2 + x_1(1 - x_1^2 - x_2^2), \\ x_2' &= x_1 + x_1(1 - x_1^2 - x_2^2). \end{cases}$ Using the following Maple command, we can obtain the phase portrait shown in Figure 2.7.

```
> with(DEtools) :
> ode1 := diff(x1(t), t) = −x2(t) + x1(t) ∗ (1 − x1(t)^2 − x2(t)^2);
> ode2 := diff(x2(t), t) = x1(t) + x1(t) ∗ (1 − x1(t)^2 − x2(t)^2);
> initvalues := seq(seq([x1(0) = i + 1/2, x2(0) = j + 1/2],
  i = −2..1), j = −2..1) :
> DEplot({ode1, ode2}, [x1(t), x2(t)], t = −4..4, x1 = −2..2, x2 = −2..2, [initvalues],
  stepsize = 0.05, arrows = MEDIUM, colour = black, linecolour = red);
```

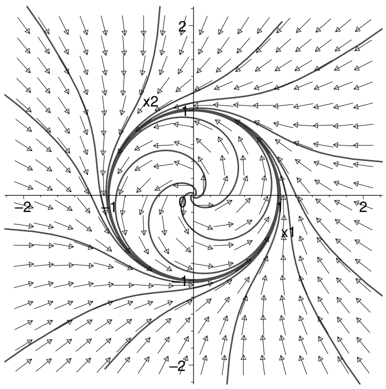

Figure 2.7: Phase portrait.

◊

Exercises.

1. Using Maple, construct phase portraits of the following systems:

   (a) $\begin{cases} x_1' &= x_2, \\ x_2' &= x_1. \end{cases}$

   (b) $\begin{cases} x_1' &= -2x_2, \\ x_2' &= x_1. \end{cases}$

   (c) $\begin{cases} x_1' &= x_1 + x_2, \\ x_2' &= x_2. \end{cases}$

2. Suppose a lake contains two species of fish, which we simply call ‘big fish’ and ‘small fish’. In the absence of big fish, the small fish population x_s evolves according to the law: $x_s' = ax_s$, where $a > 0$ is a constant. Indeed, the more the small fish, the more they reproduce. But big fish eat small fish, and so taking this into account, we have

   $$x_s' = ax_s - bx_s x_b,$$

   where $b > 0$ is a constant. The last term accounts for how often the big fish encounter the small fish–the more the small fish, the easier it becomes for the big fish to catch them, and the faster the population of the small fish decreases.

   On the other hand, the big fish population evolution is given by

   $$x_b' = -cx_b + dx_s x_b,$$

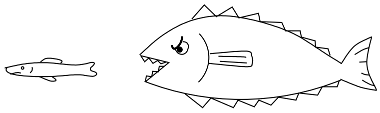

Figure 2.8: Big fish and small fish.

where $c, d > 0$ are constants. The first term has a negative sign which comes from the competition between these predators–the more the big fish, the fiercer the competition for survival. The second term accounts for the fact that the larger the number of small fish, the greater the growth in the numbers of the big fish.

(a) *Singular points.* Show that the (x_s, x_b) ODE system has two singular points $(0, 0)$ and $\left(\frac{c}{d}, \frac{a}{b}\right)$. The point $(0, 0)$ corresponds to the extinction of both species–if both populations are 0, then they continue to remain so. The point $\left(\frac{c}{d}, \frac{a}{b}\right)$ corresponds to population levels at which both species sustain their current nonzero numbers indefinitely.

(b) *Solution to the ODE system.* Use Maple to plot the population levels of the two species on the same plot, with the following data: $a = 2$, $b = 0.002$, $c = 0.5$, $d = 0.0002$, $x_s(0) = 9000$, $x_b(0) = 1000$, $t = 0$ to $t = 100$.

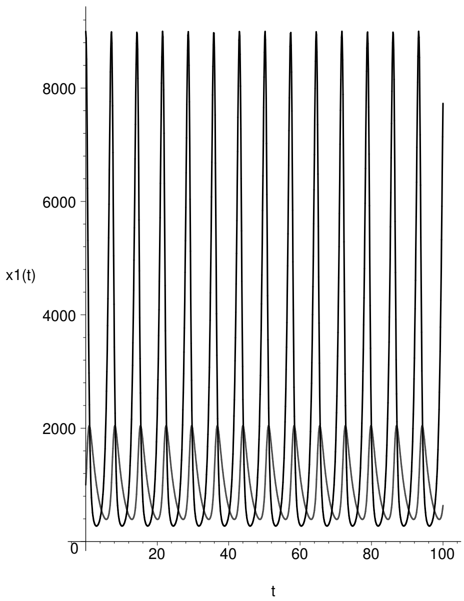

Figure 2.9: Periodic variation of the population levels.

Your plot should show that the population levels vary in a periodic manner, and the population of the big fish lags behind the population of the small fish. This is expected since the big fish thrive when the small fish are plentiful, but ultimately outstrip their food supply and decline. As the big fish population is low, the small fish numbers increase again. So there is a a cycle of growth and decline.

(c) *Phase portrait.* With the same constants as before, plot a phase portrait in the region $x_s = 0$ to $x_s = 10000$ and $x_b = 0$ to 4000.

Also plot in the same phase portrait the solution curves. What do you observe?

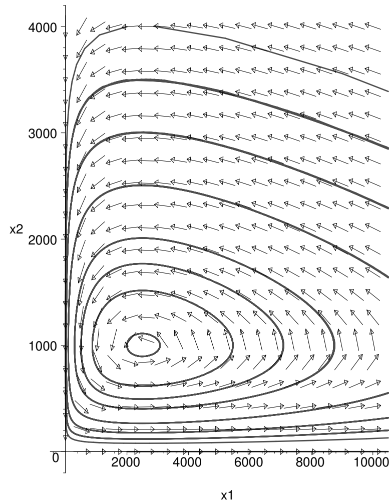

Figure 2.10: Phase portrait for the Lotka-Volterra ODE system.

2.4 Phase plane analysis of linear systems

In this section, we describe the phase plane analysis of *linear* systems. Besides allowing us to visually observe the motion patterns of linear systems, this will also help the development of nonlinear system analysis in the next section, since similar motion patterns can be observed in the local behaviour of nonlinear systems as well.

We will analyse three simple types of matrices. It turns out that it is enough to consider these three types, since every other matrix can be reduced to such a matrix by an appropriate change of basis (in the phase portrait, this corresponds to replacing the usual axes by new ones, which may not be orthogonal). However, in this elementary first course, we will ignore this part of the theory.

2.4.1 Complex eigenvalues

Consider the system $x' = Ax$, where $A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}$. Then $e^{tA} = e^{ta} \begin{bmatrix} \cos(bt) & \sin(bt) \\ -\sin(bt) & \cos(bt) \end{bmatrix}$, and so if the initial condition $x(0)$ has polar coordinates (r_0, θ_0), then the solution is given by

$$x_1(t) = e^{ta} r_0 \cos(\theta_0 - bt) \quad \text{and} \quad x_2(t) = e^{ta} r_0 \sin(\theta_0 - bt), \quad t \geq 0,$$

so that the trajectories are spirals if a is nonzero, moving towards the origin if $a < 0$, and outwards if $a > 0$. If $a = 0$, the trajectories are circles. See Figure 2.11.

2.4.2 Diagonal case with real eigenvalues

Consider the system $x' = Ax$, where

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix},$$

where λ_1 and λ_2 are real numbers. The trajectory starting from the initial condition (x_{10}, x_{20}) is given by $x_1(t) = e^{\lambda_1 t} x_{10}$, $x_2(t) = e^{\lambda_2 t} x_{20}$. We also see that $Ax_1^{\lambda_2} = Bx_2^{\lambda_2}$ with appropriate values

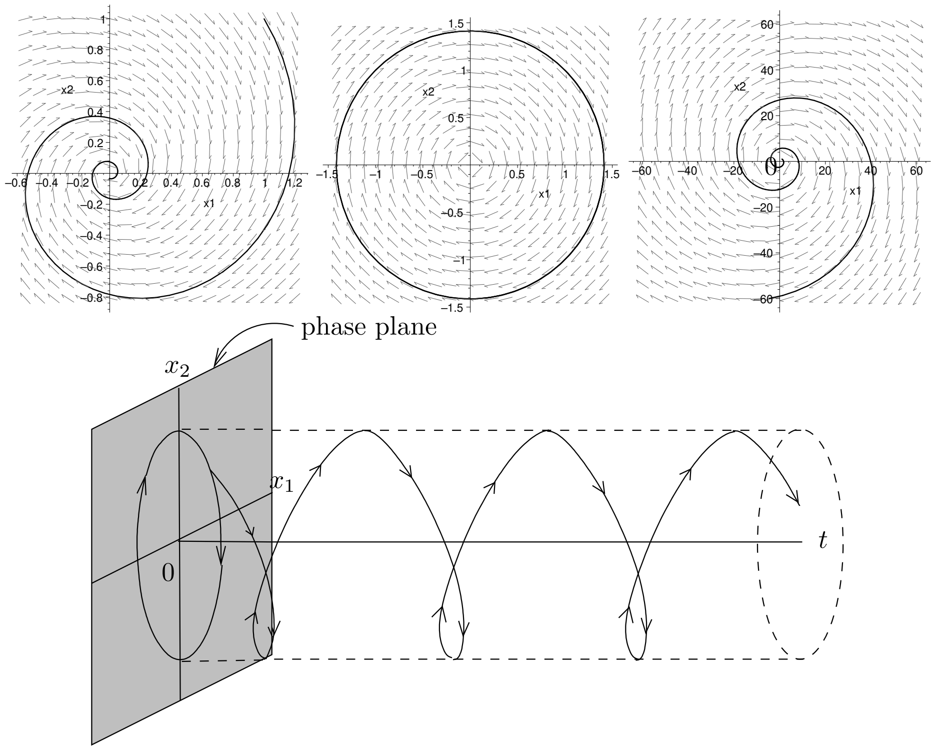

Figure 2.11: Case of complex eigenvalues. The last figure shows the phase plane trajectory as a projection of the curve $(t, x_1(t), x_2(t))$ in $\mathbb{R}^3$: case when $a = 0$, and $b > 0$.

for the constants A and B. See the topmost figure in Figure 2.12 for the case when λ_1, λ_2 are both negative. In general, we obtain the phase portraits shown in Figure 2.12, depending on the signs of λ_1 and λ_2.

2.4.3 Nondiagonal case

Consider the system $x' = Ax$, where

$$A = \begin{bmatrix} \lambda & 1 \\ 0 & \lambda \end{bmatrix},$$

where λ is a real number. It is easy to see that

$$e^{tA} = \begin{bmatrix} e^{\lambda t} & te^{\lambda t} \\ 0 & e^{\lambda t} \end{bmatrix},$$

so that the solution starting from the initial condition (x_{10}, x_{20}) is given by

$$x_1(t) = e^{\lambda t}(x_{10} + tx_{20}), \quad x_2(t) = e^{\lambda t}x_{20}.$$

Figure 2.13 shows the phase portraits for the three cases when $\lambda < 0$, $\lambda = 0$ and $\lambda > 0$.

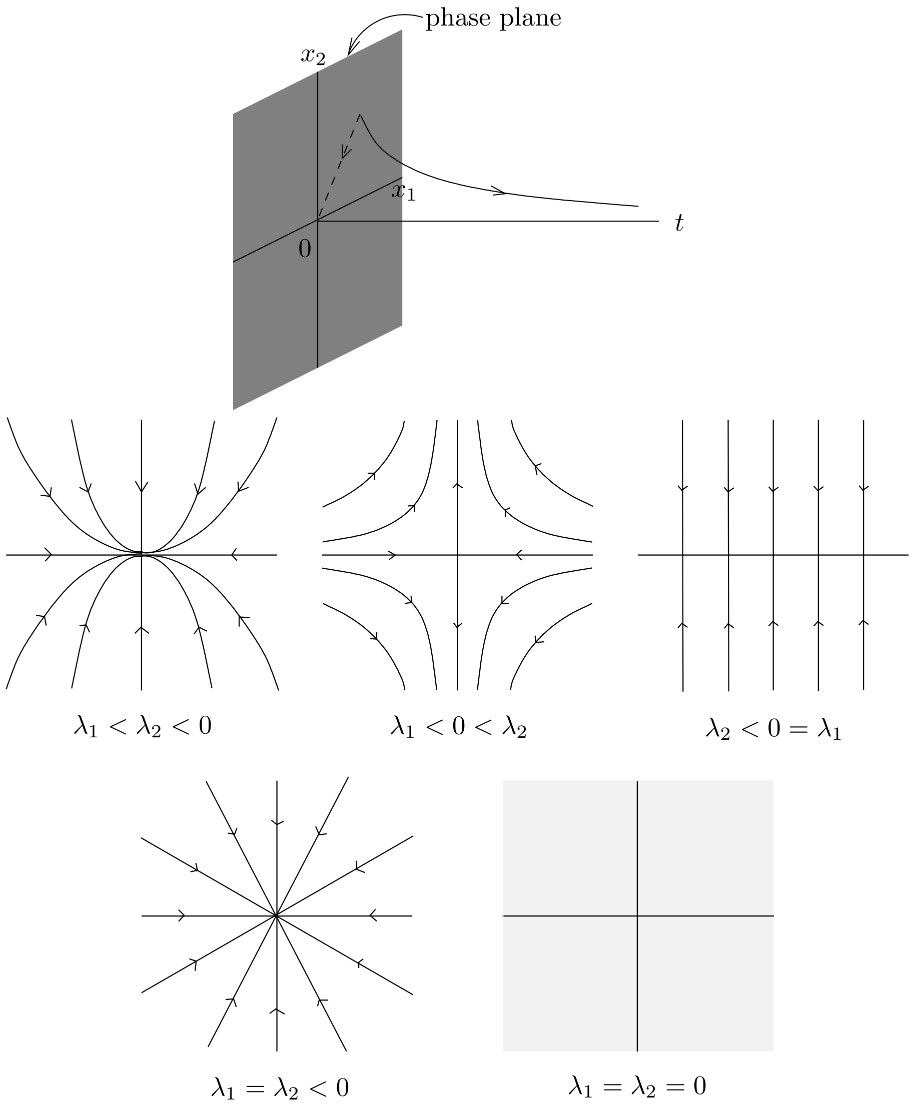

Figure 2.12: The topmost figure shows the phase plane trajectory as a projection of the curve $(t, x_1(t), x_2(t))$ in $\mathbb{R}^3$: diagonal case when both eigenvalues are negative and equal. The other figures are phase portraits in the case when A is diagonal with real eigenvalues. When the eigenvalues have opposite signs, the singular point is called a *saddle point*.

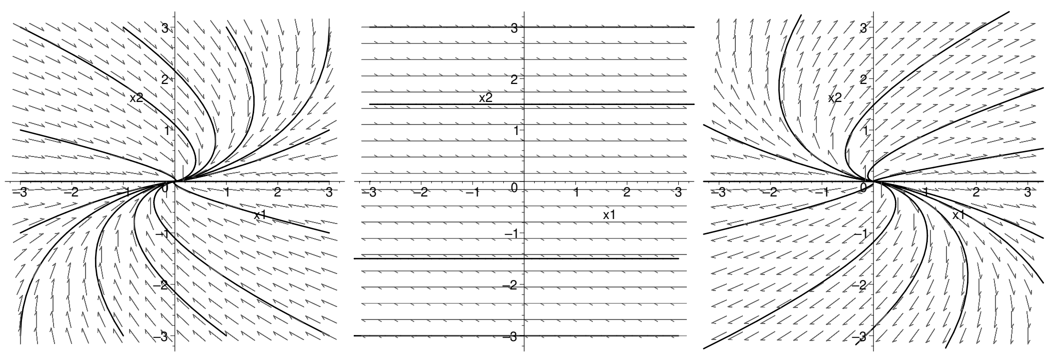

Figure 2.13: Nondiagonal case.

We note that the angle that a point on the trajectory makes with the x_1 axis is given by

$$\arctan\left(\frac{x_2(t)}{x_1(t)}\right) = \arctan\left(\frac{x_{20}}{x_{10} + tx_{20}}\right),$$

which tends to 0 or π as $t \to \infty$.

Exercises.

1. Draw the phase portrait for the system $\left\{\begin{array}{rcl} x_1' &=& x_1 - 3x_2, \\ x_2' &=& -2x_2, \end{array}\right.$ using the following procedure:

   STEP 1. Find the eigenvectors and eigenvalues: Show that $A := \begin{bmatrix} 1 & -3 \\ 0 & -2 \end{bmatrix}$ has eigenvalues $1, -2$, with eigenvectors $v_1 := \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $v_2 := \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, respectively.

   STEP 2. Set up the coordinate system in terms of the eigenvectors: Since v_1, v_2 form a basis for $\mathbb{R}^2$, the solution vector $x(t)$ can be expressed as a linear combination of v_1, v_2: $x(t) = \alpha(t)v_1 + \beta(t)v_2$. Note that $\alpha(t)$ and $\beta(t)$ are the 'coordinates' of the point $x(t)$ in the directions v_1 and v_2, respectively. In other words, they are the 'projections' of the point $x(t)$ in the directions v_1 and v_2, as shown in the Figure 2.14.

   STEP 3. Eliminate t: Show that $(\alpha(t))^2\beta(t) = (\alpha(0))^2\beta(0)$, and using the 'distorted' coordinate system, draw a phase portrait for the system $x'(t) = Ax(t)$.

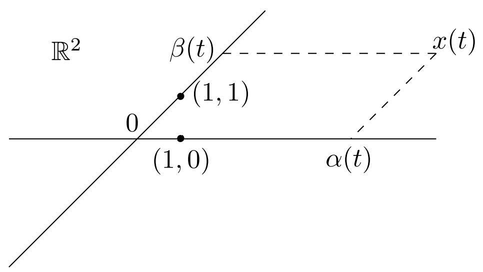

Figure 2.14: The distorted coordinate system.

2. $(*)$ Let $A \in \mathbb{R}^2$ have eigenvalues $a + ib$ and $a - ib$, where $a, b \in \mathbb{R}$, and $b \neq 0$.

   (a) If $v_1 := u + iv$ ($u, v \in \mathbb{R}^2$) is an eigenvector corresponding to the eigenvalue $a + ib$, then show that $v_2 := u - iv$ is an eigenvector corresponding to $a - ib$.

   (b) Using the fact that $b \neq 0$, conclude that v_1, v_2 are linearly independent in $\mathbb{C}^2$.

   (c) Prove that u, v are linearly independent as vectors in $\mathbb{R}^2$. Conclude that the matrix P with the columns u and v is invertible.

   (d) Verify that
   $$P^{-1}AP = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

2.5 Phase plane analysis of nonlinear systems

With the phase plane analysis of nonlinear systems, we should keep two things in mind. One is that the phase plane analysis is related to that of linear systems, because the local behaviour of a nonlinear system can be approximated by a linear system. And the second is that, despite this similarity with linear systems, nonlinear systems can display much more complicated patterns in the phase plane, such as multiple singular points and limit cycles. In this section, we will discuss these aspects. We consider the system

$$\begin{aligned} x_1' &= f_1(x_1, x_2), \\ x_2' &= f_2(x_1, x_2), \end{aligned} \tag{2.4}$$

where we assume that f_1 and f_2 have continuous partial derivatives, and this assumption will be continued for the remainder of this chapter. We will learn later in Chapter 4 that a consequence of this assumption is that for the initial value problem of the system above, there will exist a unique solution. Moreover, we will also make the assumption that solutions exist for all times in $\mathbb{R}$.

2.5.1 Local behaviour of nonlinear systems

In order to see the similarity with linear systems, we decompose the the nonlinear system into a linear part and an 'error' part (which is small close to a singular point), using Taylor's theorem, as follows.

Let (x_{10}, x_{20}) be an isolated singular point of (2.4). Thus $f_1(x_{10}, x_{20}) = 0$ and $f_2(x_{10}, x_{20}) = 0$. Then by Taylor's theorem, we have

$$x_1' = \left[\frac{\partial f_1}{\partial x_1}(x_{10}, x_{20})\right](x_1 - x_{10}) + \left[\frac{\partial f_1}{\partial x_2}(x_{10}, x_{20})\right](x_2 - x_{20}) + e_1(x_1 - x_{10}, x_2 - x_{20}), \tag{2.5}$$

$$x_2' = \left[\frac{\partial f_2}{\partial x_1}(x_{10}, x_{20})\right](x_1 - x_{10}) + \left[\frac{\partial f_2}{\partial x_2}(x_{10}, x_{20})\right](x_2 - x_{20}) + e_2(x_1 - x_{10}, x_2 - x_{20}), \tag{2.6}$$

where e_1 and e_2 are such that $e_1(0,0) = e_2(0,0) = 0$. We translate the singular point (x_{10}, x_{20}) to the origin by introducing the new variables $y_1 = x_1 - x_{10}$ and $y_2 = x_2 - x_{20}$. With

$$\begin{aligned} a &:= \frac{\partial f_1}{\partial x_1}(x_{10}, x_{20}), \qquad & b &:= \frac{\partial f_1}{\partial x_2}(x_{10}, x_{20}), \\ c &:= \frac{\partial f_2}{\partial x_1}(x_{10}, x_{20}), \qquad & d &:= \frac{\partial f_2}{\partial x_2}(x_{10}, x_{20}), \end{aligned}$$

we can rewrite (2.5)-(2.6) as follows:

$$\begin{aligned} y_1' &= ay_1 + by_2 + e_1(y_1, y_2), \\ y_2' &= cy_1 + dy_2 + e_2(y_1, y_2). \end{aligned} \tag{2.7}$$

We note that this new system has $(0,0)$ as a singular point. We will elaborate on the similarity between the phase portrait of the system (2.4) with the phase portrait of its linear part, that is, the system

$$\begin{aligned} z_1' &= az_1 + bz_2, \\ z_2' &= cz_1 + dz_2. \end{aligned} \tag{2.8}$$

Before clarifying the relationship between (2.4) and (2.8), we pause to note some important differences. The system (2.4) may have *many* singular points; *one* of them has been selected and moved to the origin. If a different singular point would have been chosen, then the constants a, b, c, d in (2.8) would have been different. The important point is that any statement relating (2.4) and (2.8), is *local* in nature, in that they apply 'near' the singular point under consideration. By 'near' here, we mean in a sufficiently small neighbourhood or ball around the singular point. Totally different kinds of behaviour may occur in a neighbourhood of other critical points. The transformation above must be made, and the corresponding linear part must be analysed, for each isolated singular point of the nonlinear system.

We now give the main theorem in this section about the local relationship between the nature of phase portraits of (2.4) and (2.8), but we will not prove this theorem.

Theorem 2.5.1 *Let (x_{10}, x_{20}) be an isolated singular point of (2.4), and let*

$$A := \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_{10}, x_{20}) & \frac{\partial f_1}{\partial x_2}(x_{10}, x_{20}) \\ \frac{\partial f_2}{\partial x_1}(x_{10}, x_{20}) & \frac{\partial f_2}{\partial x_2}(x_{10}, x_{20}) \end{bmatrix}.$$

Then we have the following:

1. *If every eigenvalue of A has a negative real part, then all solutions of (2.4) starting in a small enough ball with centre (x_{10}, x_{20}) converge to (x_{10}, x_{20}) as $t \to \infty$. (This situation is abbreviated by saying that the equilibrium point (x_{10}, x_{20}) is 'asymptotically stable'; see §3.3.)*
2. *If the matrix A has an eigenvalue with a positive real part, then there exists a ball B such that for every ball B' of positive radius around (x_{10}, x_{20}), there exists a point in B' such that a solution x of (2.4) starting from that point leaves the ball B. (This situation is abbreviated by saying that the equilibrium point (x_{10}, x_{20}) is 'unstable'; see §3.2.)*

We illustrate the theorem with the following example.

Example. Consider the system

$$\begin{aligned} x_1' &= -x_1 + x_2 - x_1(x_2 - x_1), \\ x_2' &= -x_1 - x_2 + 2x_1^2 x_2. \end{aligned}$$

This nonlinear system has the singular points $(-1, -1)$, $(1, 1)$ and $(0, 0)$. If we linearise around the singular point $(0, 0)$, we obtain the matrix

$$\begin{bmatrix} \left.\frac{\partial}{\partial x_1}(-x_1 + x_2 - x_1(x_2 - x_1))\right|_{(0,0)} & \left.\frac{\partial}{\partial x_2}(-x_1 + x_2 - x_1(x_2 - x_1))\right|_{(0,0)} \\ \left.\frac{\partial}{\partial x_1}(-x_1 - x_2 + 2x_1^2 x_2)\right|_{(0,0)} & \left.\frac{\partial}{\partial x_2}(-x_1 - x_2 + 2x_1^2 x_2)\right|_{(0,0)} \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix},$$

which has eigenvalues $-1 + i$ and $-1 - i$. Thus by Theorem 2.5.1, it follows that for the above nonlinear system, if we start close to $(0, 0)$, then the solutions converge to $(0, 0)$ as $t \to \infty$.

However, not all solutions of this nonlinear system converge to $(0, 0)$. For example, we know that $(1, 1)$ is also a singular point, and so if we start from there, then we stay there. ◊

The above example highlights the *local* nature of Theorem 2.5.1. How close is sufficiently close is generally a difficult question to answer.

Actually more than just similarity of convergence to the singular point can be said. It turns out that if the real parts of the eigenvalues are not equal to zero, then also the 'qualitative' structure is preserved. Roughly speaking, this means that there is a map T mapping a region Ω_1 around (x_{10}, x_{20}) to a region Ω_2 around $(0, 0)$ such that

1. T is one-to-one and onto;
2. both T and T^{-1} are continuous;
3. if two points of Ω_1 lie on the same trajectory of (2.8), then their images under T lie on the same trajectory of (2.4);
4. if two points of Ω_2 lie on the same trajectory of (2.4), then their images under T^{-1} lie on the same trajectory of (2.8).

The mapping is shown schematically in Figure 2.15.

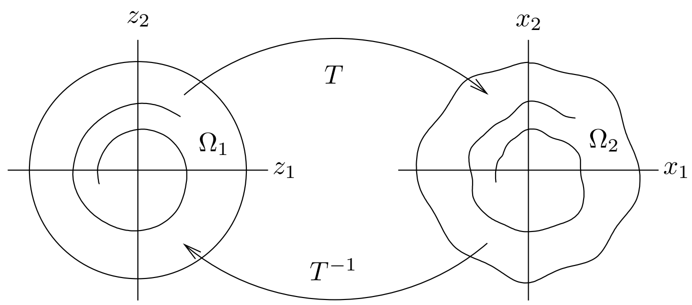

Figure 2.15: The mapping T.

The actual construction of such a mapping is not easy, but we demonstrate the plausibility of its existence by considering an example.

Example. Consider the system

$$\begin{aligned} x_1' &= x_2, \\ x_2' &= x_1 - x_2 + x_1(x_1 - 2x_2). \end{aligned}$$

The singular points are solutions to

$$x_2 = 0 \quad \text{and} \quad x_1 - x_2 + x_1^2 - 2x_1x_2 = 0,$$

and so they are $(0, 0)$ and $(-1, 0)$. Furthermore,

$$\begin{aligned} \frac{\partial f_1}{\partial x_1} &= 0, & \frac{\partial f_1}{\partial x_2} &= 1, \\ \frac{\partial f_2}{\partial x_1} &= 1 + 2x_1 - 2x_2, & \frac{\partial f_1}{\partial x_2} &= -1 - 2x_1. \end{aligned}$$

At $(0, 0)$, the matrix of the linear part is

$$\begin{bmatrix} 0 & 1 \\ 1 & -1 \end{bmatrix},$$

whose eigenvalues satisfy $\lambda^2 + \lambda - 1 = 0$. The roots are real and of opposite signs, and so the origin is a saddle point.

At $(-1, 0)$, the matrix of the linear part is

$$\begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix},$$

whose eigenvalues satisfy $\lambda^2 - \lambda + 1 = 0$. The trajectories are thus outward spirals.

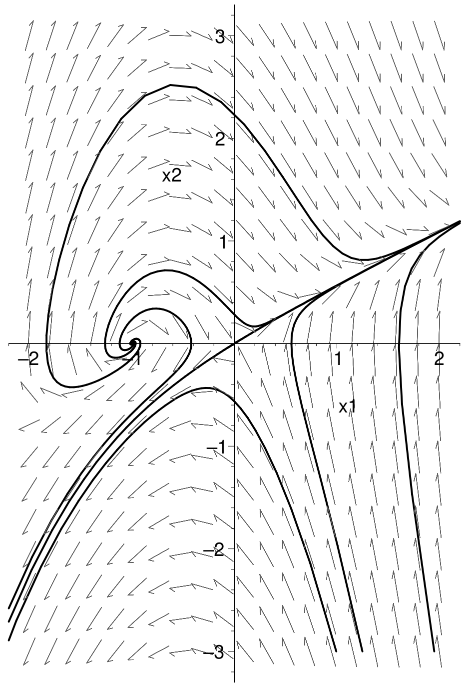

Figure 2.16: Phase portrait.

Figure 2.16 shows the phase portrait of the system. As expected we see that around the points $(0, 0)$ and $(-1, 0)$, the local picture is similar to the corresponding linearisations. $\Diamond$

Finally, we discuss the case when the eigenvalues of the linearisation have real part equal to 0. It turns out that in this case, the behaviour of the linearisation gives no information about the behaviour of the nonlinear system. For example, circles in the phase portrait may be converted into spirals. We illustrate this in the following example.

Example. Consider the system

$$\begin{aligned} x_1' &= -x_2 - x_1(x_1^2 + x_2^2), \\ x_2' &= x_1 - x_2(x_1^2 + x_2^2). \end{aligned}$$

The linearisation about the singular point $(0, 0)$ gives rise to the matrix

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

which has eigenvalues $-i$ and i. Thus the phase portrait of the linear part comprises *circles*.

If we introduce the polar coordinates $r := \sqrt{x_1^2 + x_2^2}$ and $\theta = \arctan(x_2/x_1)$, then we have that

$$\begin{aligned} r' &= -r^3, \\ \theta' &= 1. \end{aligned}$$

Thus we see that the trajectories approach the origin in *spirals*! $\Diamond$

Exercises. Use the linearisation theorem (Theorem 2.5.1), where possible, to describe the behaviour close to the equilibrium points of the following systems:

1. $\left\{ \begin{aligned} x_1' &= e^{x_1+x_2} - x_2, \\ x_2' &= -x_1 + x_1 x_2. \end{aligned} \right.$

2. $\left\{ \begin{aligned} x_1' &= x_1 + 4x_2 + e^{x_1} - 1, \\ x_2' &= -x_2 - x_2 e^{x_1}. \end{aligned} \right.$

3. $\left\{ \begin{aligned} x_1' &= x_2, \\ x_2' &= -x_1^3. \end{aligned} \right.$

4. $\left\{ \begin{aligned} x_1' &= \sin(x_1 + x_2), \\ x_2' &= x_2. \end{aligned} \right.$

5. $\left\{ \begin{aligned} x_1' &= \sin(x_1 + x_2), \\ x_2' &= -x_2. \end{aligned} \right.$

2.5.2 Limit cycles and the Poincaré-Bendixson theorem

In the phase portrait of the van der Pol equation shown in Figure 2.5, we suspected that the system has a closed curve in the phase portrait, and moreover, trajectories starting inside that curve, as well as trajectories starting outside that curve, all tended towards this curve, while a motion starting on that curve would stay on it forever, circling periodically around the origin. Such a curve is called a "limit cycle", and we will study the exact definition later in this section. Limit cycles are a unique feature that can occur only in a nonlinear system. Although in the phase portrait in the middle of the top row of figures in Figure 2.11, we saw that if the real part of the eigenvalues is zero, we have periodic trajectories, these are not limit cycles, since now matter how close we start from such a periodic orbit, we can never approach it. We want to call those closed curves limit cycles such that for all trajectories starting close to it, they converge to it either as the time increases to $+\infty$ or decreases to $-\infty$. In order to explain this further, we consider the following example.

Examples. Consider the system

$$\begin{aligned} x_1' &= x_2 - x_1(x_1^2 + x_2^2 - 1) \\ x_2' &= -x_1 - x_2(x_1^2 + x_2^2 - 1). \end{aligned}$$

By introducing polar coordinates $r = \sqrt{x_1^2 + x_2^2}$ and $\theta = \arctan(x_2/x_1)$, the equations are transformed into

$$\begin{aligned} r' &= -r(r^2 - 1) \\ \theta' &= -1. \end{aligned}$$

When we are on the unit circle, we note that $r' = 0$, and so we stay there. Thus the unit circle is a periodic trajectory. When $r > 1$, then $r' < 0$, and so we see that if we start outside the unit circle, we tend towards the unit circle from the outside. On the other hand, if $r < 1$, then $r' > 0$, and so if we start inside the unit circle, we tend towards it from the inside. This can be made rigorous by examining the analytical solution, given by $r(t) = \left(\sqrt{1 + (1/r_0^2 - 1)e^{-2t}}\right)^{-1}$, $\theta(t) = \theta_0 - t$, where (r_0, θ_0) denotes the initial condition. So we have that all trajectories in the vicinity of the unit circle converge to the unit circle as $t \to \infty$.

Now consider the system

$$\begin{aligned} x_1' &= x_2 + x_1(x_1^2 + x_2^2 - 1) \\ x_2' &= -x_1 + x_2(x_1^2 + x_2^2 - 1). \end{aligned}$$

Again by introducing polar coordinates (r, θ) as before, we now obtain

$$\begin{aligned} r' &= r(r^2 - 1) \\ \theta' &= -1. \end{aligned}$$

When we are on the unit circle, we again have that $r' = 0$, and so we stay there. Thus the unit circle is a periodic trajectory. But now when $r > 1$, then $r' > 0$, and so we see that if we start outside the unit circle, we move away from the unit circle. On the other hand, if $r < 1$, then $r' < 0$, and so if we start inside the unit circle, we again move away from the unit circle. However, if we start with an initial condition (r_0, θ_0) with $r_0 < 1$, and go backwards in time, then we can show that the solution is given by $r(t) = \left(\sqrt{1 + (1/r_0^2 - 1)e^{2t}}\right)^{-1}$, $\theta(t) = \theta_0 - t$, and so we have that all trajectories in the vicinity of the unit circle from inside converge to the unit circle as $t \to -\infty$. ♢

In each of the examples considered above, we would like to call the unit circle a "limit cycle". This motivates the following definitions of ω- and α- limit points of a trajectory, and we will define limit cycles using these notions.

Definition. Let x be a solution of $x' = f(x)$. A point x_* is called an *ω-limit point of x* if there is a sequence of real numbers $(t_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} t_n = +\infty$ and $\lim_{n\to\infty} x(t_n) = x_*$. The set of all ω-limit points of x is denoted by $L_\omega(x)$.

A point x_* is called an *α-limit point of x* if there is a sequence of real numbers $(t_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} t_n = -\infty$ and $\lim_{n\to\infty} x(t_n) = x_*$. The set of all α-limit points of x is denoted by $L_\alpha(x)$.

For example, if x_* is a singular point, clearly $L_\omega(x_*) = L_\alpha(x_*) = x_*$. We consider a few more examples below.

Examples. Consider the system $\left\{\begin{aligned} x_1' &= -x_1 \\ x_2' &= x_2, \end{aligned}\right.$ which has the origin as a saddle singular point. For any trajectory x starting on the x_1-axis (but not at the origin), we have $L_\omega(x) = (0,0)$, while $L_\alpha(x) = \emptyset$. On the other hand, for any trajectory x starting on the x_2-axis (but not at the origin), $L_\omega(x) = \emptyset$, and $L_\alpha(x) = (0,0)$. Finally, for any trajectory x that does not start on the x_1- or the x_2-axis, the sets $L_\omega(x) = L_\alpha(x) = \emptyset$.

Now consider the system $\left\{\begin{aligned} x_1' &= x_2 \\ x_2' &= -x_1, \end{aligned}\right.$ for which all trajectories are periodic and are circles in the phase portrait. For any trajectory starting from a point P, the sets $L_\omega(x)$, $L_\alpha(x)$

are both equal to the circle passing through P. $\Diamond$

We are now ready to define a limit cycle.

Definitions. A *periodic trajectory* is a nonconstant solution such that there exists a $T > 0$ such that $x(t) = x(t + T)$ for all $t \in \mathbb{R}$.

A *limit cycle* is a periodic trajectory that is contained in $L_\omega(x)$ or $L_\alpha(x)$ for some other trajectory x.

Limit cycles represent an important phenomenon in nonlinear systems. They can be found often in engineering and nature. For example, aircraft wing fluttering is an instance of a limit cycle frequently encountered which is sometimes dangerous. In an ecological system where two species share a common resource, the existence of a limit cycle would mean that none of the species becomes extinct. As one can see, limit cycles can be desirable in some cases, and undesirable in other situations. In any case, whether or not limit cycles exist is an important question, and we now study a few results concerning this. In particular, we will study an important result, called the Poincaré-Bendixson theorem, for which we will need a few topological preliminaries, which we list below.

Definitions. Two nonempty sets A and B in the plane $\mathbb{R}^2$ are said to be *separated* if there is no sequence of points $(p_n)_{n\in\mathbb{N}}$ contained in A such that $\lim_{n\to\infty} p_n \in B$, and there is no sequence $(q_n)_{n\in\mathbb{N}}$ contained in B such that $\lim_{n\to\infty} q_n \in A$.

A set that is not the union of two separated sets is said to be *connected*.

A set O is *open* if for every $x \in O$, there exists a $\epsilon > 0$ such that $B(x, \epsilon) \subset O$.

A set Ω is called a *region* if it is an open, connected set.

A set A is said to be bounded if there exists a $R > 0$ large enough so that $A \subset B(0, R)$.

For example, two disjoint circles in the plane are separated, while the quadrant $\{(x, y) \in \mathbb{R}^2 \mid x > 0, y > 0\}$ is not separated from the x-axis.

Although the definition of a connected set seems technical, it turns out that sets that we would intuitively think of as connected in a nontechnical sense are connected. For instance the annulus $\{(x, y) \in \mathbb{R}^2 \mid 1 < x^2 + y^2 < 2\}$ is connected, while the set $\mathbb{Z}^2$ is not.

Roughly speaking, an open set can be thought of as a set without its 'boundary'. For example, the unit disk $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1\}$ is open.

The principal result is the following.

Theorem 2.5.2 (Poincaré-Bendixson) *Let* (x_1, x_2) *be a solution to* $\left\{ \begin{array}{rcl} x_1' &=& f_1(x_1, x_2), \\ x_2' &=& f_2(x_1, x_2), \end{array} \right.$ *such that for all* $t \geq t_0$*, the solution lies in a bounded region of the plane containing no singular points. Then either the solution is a periodic trajectory or its omega limit set is a periodic trajectory.*

The proof requires advanced mathematical techniques to prove, and the proof will be omitted. The Poincaré-Bendixson theorem is false for systems of dimension 3 or more. In the case of 2D systems, the proof depends heavily on a deep mathematical result, known as the Jordan curve theorem, which is valid only in the plane. Although the theorem sounds obvious, its proof is difficult. We state this theorem below, but first we should specify what we mean by a curve.

Definitions. A *curve* is a continuous function $f : [a, b] \to \mathbb{R}^2$. If for every $t_1, t_2 \in (a, b)$ such that $t_1 \neq t_2$, there holds that $f(t_1) \neq f(t_2)$, then the curve is called *simple.* A curve is called *closed* if $f(a) = f(b)$.

Theorem 2.5.3 (Jordan curve theorem) *A simple closed curve divides the plane into two regions, one of which is bounded, and the other is unbounded.*

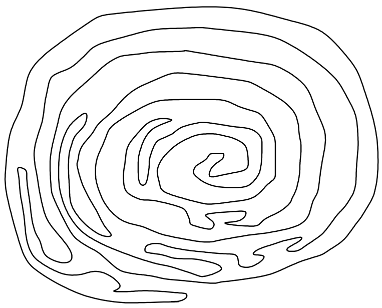

Figure 2.17: Jordan curve theorem.

Now that the inside of a curve is defined, the following result helps to clarify the type of region to seek in order to apply the Poincaré-Bendixson theorem.

Theorem 2.5.4 *Every periodic trajectory of* $\left\{ \begin{array}{rcl} x_1' &=& f_1(x_1, x_2) \\ x_2' &=& f_2(x_1, x_2) \end{array} \right.$ *contains a singular point in its interior.*

This theorem tells us that in order to apply the Poincaré-Bendixson theorem, the singular-point-free-region where the trajectory lies must have at least one hole in it (for the singular point).

We consider a typical application of the Poincaré-Bendixson theorem.

Example.(∗) Consider the system

$$\begin{array}{rcl} x_1' &=& x_1 + x_2 - x_1(x_1^2 + x_2^2)[\cos(x_1^2 + x_2^2)]^2 \\ x_2' &=& -x_1 + x_2 - x_2(x_1^2 + x_2^2)[\cos(x_1^2 + x_2^2)]^2. \end{array}$$

In polar coordinates, the equations are transformed into

$$\begin{array}{rcl} r' &=& r[1 - r^2(\cos r^2)^2] \\ \theta' &=& -1. \end{array}$$

Consider a circle of radius $r_0 < 1$ about the origin. If we start on it, then all trajectories move outward, since

$$r'(t_0) = r_0[1 - r_0^2(\cos r_0^2)^2] > r_0[1 - r_0^2] > 0.$$

Also if $r(t_0) = \sqrt{\pi}$, then

$$r'(t_0) = \sqrt{\pi}[1 - \pi] < 0,$$

and so trajectories starting on the circle with radius $\sqrt{\pi}$ (or close to it) move inwards. Then it can be shown that any trajectory starting inside the annulus $r_0 < r < \sqrt{\pi}$ stays there[3]. But it is clear that there are no singular points inside the annulus $r_1 < r < \sqrt{\pi}$. So this region must contain a periodic trajectory. Moreover, it is also possible to prove that a trajectory starting inside the unit circle is not a periodic trajectory (since it never returns to this circle), and hence its omega limit set is a limit cycle. ♢

It is also important to know when there are no periodic trajectories. The following theorem provides a sufficient condition for the non-existence of periodic trajectories. In order to do that, we will need the following definition.

Definition. A region Ω is said to be *simply connected* if for any simple closed curve C lying entirely within Ω, all points inside C are points of Ω.

For example, the annulus $1 < r < 2$ is not simply connected, while the unit disk $r < 1$ is.

Theorem 2.5.5 *Let Ω be a simply connected set in $\mathbb{R}^2$. If the function $\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2}$ is not identically zero in Ω and does not change sign in Ω, then there are no periodic trajectories of*

$$\begin{aligned} x_1' &= f_1(x_1, x_2) \\ x_2' &= f_2(x_1, x_2) \end{aligned}$$

in the region Ω.

Proof *(Sketch.)* Assume that a periodic trajectory exists with period T, and denote the curve by C. Using Green's theorem, we have

$$\iint \left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} \right) dx_1 dx_2 = \int_C f_2 dx_1 - f_1 dx_2.$$

But

$$\int_C f_2 dx_1 - f_1 dx_2 = \int_0^T (f_2(x_1(t), x_2(t))x_1'(t) - f_1(x_1(t), x_2(t))x_2')dt = 0,$$

a contradiction. ■

Example. Consider the system

$$\begin{aligned} x_1' &= x_2 + x_1 x_2^2, \\ x_2' &= -x_1 + x_1^2 x_2. \end{aligned}$$

Since

$$\left(\frac{\partial f_1}{\partial x_1} + \frac{\partial f_2}{\partial x_2} \right) = x_1^2 + x_2^2,$$

[3]Indeed, for example if the trajectory leaves the annulus, and reaches a point r_* inside the inner circle at time t_*, then go "backwards" along this trajectory and find the first point t_1 such that $r(t_1) = r_0$. We then arrive at the contradiction that $r_* - r_0 = r(t_*) - r(t_1) = \int_{t_1}^{t_*} r'(t)dt > 0$. The case of the trajectory leaving the outer circle of the annulus can be handled similarly.

which is always strictly positive (except at the origin), the system does not have any periodic trajectories in the phase plane. ◊

Exercises.

1. Show that the sets $A = \{(x, y) \in \mathbb{R}^2 \mid xy = 1\}$ and $B = \{(x, y) \in \mathbb{R}^2 \mid xy = 0\}$ are separated.

2. Show that the system

$$\begin{aligned} x_1' &= 1 - x_1 x_2 \\ x_2' &= x_1 \end{aligned}$$

has no limit cycles.

HINT: Are there any singular points?

3. Show that the system

$$\begin{aligned} x_1' &= x_2 \\ x_2' &= -x_1 - (1 + x_1^2) x_2 \end{aligned}$$

has no periodic trajectories in the phase plane.

4. Prove that if the system

$$\begin{aligned} x_1' &= -x_2 + x_1(1 - x_1^2 - x_2^2) \\ x_2' &= x_1 + x_2(1 - x_1^2 - x_2^2) + \frac{3}{4}, \end{aligned}$$

has a periodic trajectory starting inside the circle C $x_1^2 + x_2^2 = \frac{1}{2}$, then it will either intersect C.

HINT: Consider the simply connected region $x_1^2 + x_2^2 < \frac{1}{2}$.

5. Show that $L_\omega(x) = L_\alpha(x) = \{x(t) \mid t \in \mathbb{R}\} = \{x(t) \mid t \in [0, T]\}$ for a periodic trajectory x with period T.