MTH5123 Differential Equations

Lecture Notes

Week 1

School of Mathematical Sciences
Queen Mary University of London

Introduction

An **ordinary differential equation (ODE) of $n-$th order** is a relation between an **unknown** function $y = y(x)$ of a **single independent** real variable $x \in \mathbb{R}$, and the derivatives

$$y' \equiv \frac{dy}{dx}, \ldots, y^{(n)} \equiv \frac{d^n y}{dx^n}.$$

Symbolically we can write any ODE in the form

$$F(x, y, y', \ldots, y^{(n)}) = 0 \tag{1}$$

The highest derivative entering (1) defines the **order** of the ODE. Examples are $y' + y = 0$, $y'' - x^2 y' + \sin y = 0$, etc.

Definition: Any function $y = f(x)$ defined in some interval $x \in (A, B)$, which when substituted to eq.(1) reduces it to an identity, is called a **solution** of eq.(1), and (A, B) is called its interval of definition.

The majority of interesting differential equations (not only ordinary ones!) comes from modelling problems in various branches of physics, such as classical mechanics (Newton's equations of motion), quantum mechanics (Schrödinger's eqn.), the theory of electricity and magnetism (Maxwell's equations), hydrodynamics (Navier-Stokes eqn.), etc. They also play important roles in ecological and biological problems (logistic equation for population growth and extinction), engineering (e.g. launching and control of aircrafts and missiles, problems of combustion, satellite navigation), economics and finances (resource optimization; dynamics of stock exchange indices, etc.).

Note: The role of the independent variable x in applications is most frequently played by the **time variable** t, and we are then interested in functions $y(t)$. In that case the standard notations for derivatives are: $\dot{y} \equiv \frac{dy}{dt}, \ddot{y} \equiv \frac{d^2y}{dt^2}$, etc.

Example:
Newton's Second Law for a point mass m moving along a single (say, vertical) coordinate y under the influence of a force f reads **mass** $\times$ **acceleration** $=$ **force**. By definition, velocity is given by the first derivative $v = \dot{y}$ and acceleration is given by second derivative $a = \ddot{y}$ of the coordinate $y(t)$. Hence Newton's Second Law takes the form of the second-order differential equation

$$m\ddot{y} = f(t, y, \dot{y}) \ , \tag{2}$$

where the force f may in general be time-dependent and velocity-dependent. According to Newtonian mechanics, all complex mechanical motion in the world is governed by second order differential equations, hence their importance. One of the simplest systems of that sort is represented by a point mass m attached to the loose end of a massless elastic spring of length l, with the other end of the spring being fixed to a ceiling (see Fig. 1).

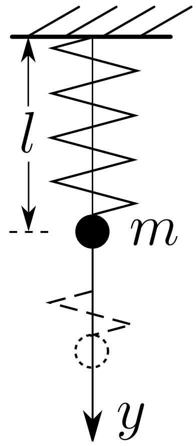

Figure 1: Elastic spring of equilibrium length l with an attached mass m.

Measuring the coordinate y from the ceiling downwards, the mass is subject to a force equal to the sum of three contributions: the position-independent **gravity force** $f_g = mg$, the position dependent **elastic force** $f_{el} = -k(y-l)$ (Hook's law of elasticity), and the **friction force** $f_a = -\gamma\,\dot{y}$ which is proportional to the velocity and is directed against the actual motion. Then (2) takes the form

$$m\ddot{y} = mg - k(y-l) - \gamma\dot{y}\,. \tag{3}$$

Here g is the gravity acceleration, k is the spring constant depending on the spring's material, and γ is the friction coeffcient. We will be able to analyze this equation and the resulting motion in due time, after we learn the methods allowing one to solve such equations.

Example:
Another example in biology is the logistic equation, which is also called the Verhulst model. The logistic equation describes a model of population growth introduced by Pierre Verhulst (1845, 1847). In this model, the initial stage of population growth is approximately exponential when the population size N is small; then the growth slows when population size increases, and stops when the population size reaches the maximum capacity of the environment K. The change of population size over time is governed by a first order non-linear differential equation.

$$\frac{dN(t)}{dt} = rN(t)(1 - \frac{N(t)}{K}). \tag{4}$$

Here, r is the per capita growth rate of a population in the time interval dt. We can see that when $N(t) = K$, $\frac{dN(t)}{dt} = 0$, the population stops growing and its size does not change further. In biology, the maximum population size K is called as the carrying capacity of a population under a certain environment. For some species, e.g. elephants in tropic forests, the carrying capacity K can be very small such as hundreds, as elephants need a lot of food and large space and thus limited number of individuals can be afforded by a natural habitat. However, if we think of colon cancer population in our body, the carrying capacity K can be as large as more than 10^{11} cells, as our body cell is very small, a detectable nail size tumour has more than 10^9 cells. In this case, the initial growth of a tumour from a single cell can be approximately as exponential growth, where $N(t)/K \to 0$ when t is small and $N(0) = 1$.

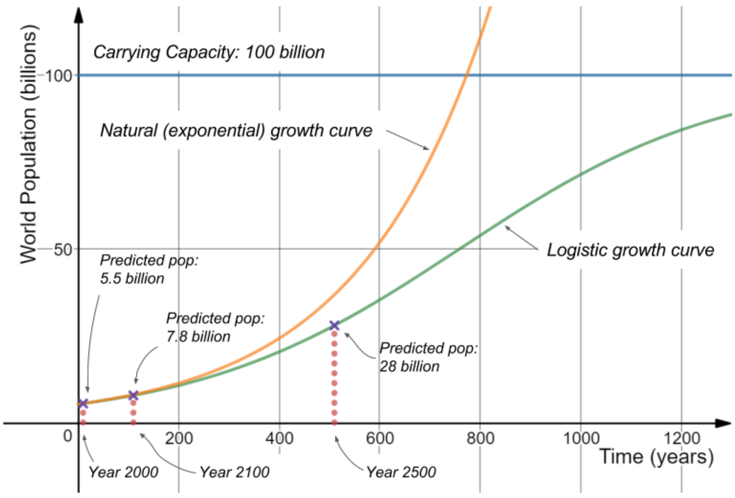

Figure 2: A illustrative figure of world population growth (from *steemiteducation* website)

1 Properties of first-order ODEs

Explicit Solutions of Simple Types of First Order ODEs

In this chapter we familiarize ourselves with a few simple **first-order differential equations**, always written in the **normal form** $y' = f(x, y)$ or $\dot{y} = f(t, y)$, which allow for a complete analytical solution. The simplest case is one when the right-hand side is independent of y, that is

$$y' = f(x)$$

Solutions to such an equation amount to finding the antiderivative for the right-hand side, that is, to a simple integration: $y = \int f(x)dx + C$, where C is an arbitrary constant. We will see that general solutions of first order ODEs will always contain an arbitrary constant. The above type belongs in fact to a more general class of explicitly solvable first order ODEs as discussed in the next section.

1.1 Separable First Order ODEs

These are equations with the **right-hand side being a product of two factors**, one depending only on the variable x and another one depending only on the unknown function y, that is

$$\frac{dy}{dx} = f(x)g(y)\,, \tag{1.1}$$

where both $f(x)$ and $g(y)$ are assumed to be continuous. The first observation is that if $y_1, \ldots, y_k$ are roots of the algebraic equation $g(y) = 0$, then the constant functions

$$y(x) = y_1,\; y(x) = y_2, \ldots,\; y(x) = y_k$$

are solutions of the ODE (1.1).
To find non-constant solutions scientists and engineers usually employ the following *heuristic* method (i.e mathematically ill-defined, but producing sensible results) of **separation of variables**, which allows one to solve (1.1) by the following steps: One starts with treating the derivative $\frac{dy}{dx}$ *as if* it was a ratio of two algebraic quantities dy and dx. That way one formally *separates variables* as $\frac{dy}{g(y)} = f(x)dx$ and by integrating both sides arrives at the relation

$$\int \frac{dy}{g(y)} = \int f(x)dx + C\,, \tag{1.2}$$

where C is an arbitrary constant. Denoting the result of integration on the left-hand side as $\int \frac{dy}{g(y)} \equiv H(y)$ (1.2) takes the form $H(y) = \int f(x)dx + C$. At the final step we may try to express $y(x)$ by formally defining the **inverse function** $H^{-1}(u)$ in such a way that $H\left(H^{-1}(u)\right) = u$, which is equivalent to

$$H(y) = u \quad \Leftrightarrow \quad y = H^{-1}(u)\,.$$

This allows one to write a one-parameter family of solutions to (1.1) as

$$y = H^{-1}\left(\int f(x)dx + C\right)\,.$$

Note: There may exist more than one function inverse to a given function. For example, suppose $H(y) = y^2$. Solving $y^2 = u$ we find $y = \pm\sqrt{u}$ for any $u > 0$. Hence there exist two different inverse functions $H^{-1}(u > 0) = \sqrt{u}$ and $H^{-1}(u > 0) = -\sqrt{u}$. To find *all* solutions to an ODE by separation of variables we need to use *all* possible inverse functions $H^{-1}(u)$.

Example:
Find non-constant solutions of the ODEs

$$(\mathbf{a})\; y' = xy^2, \qquad (\mathbf{b})\; y' = \frac{2xy}{1+y} \qquad (\mathbf{c})\; y' = 3y^{2/3}$$

Solution:
(**a**) Separating the variables we have

$$H(y) = \int \frac{dy}{y^2} = \int x dx + C$$

which gives $H(y) = -\frac{1}{y}$ on the left-hand side so that the equation $H(y) = -\frac{1}{y} = u$ is solved by $y = -1/u$. This defines the inverse function $H^{-1}(u) = -1/u$. On the right-hand side the integration gives $\frac{1}{2}x^2 + C$.

The general solution to the ODE is then given by applying the function H^{-1} to the right-hand side

$$y = H^{-1}\left(\frac{x^2}{2} + C\right) = -\frac{1}{\frac{x^2}{2} + C}$$

for any value of the constant C.

(**b**) In this case

$$H(y) = \int dy\, \frac{y+1}{2y} = \int x dx = \frac{1}{2}x^2 + C$$

We further write on the left-hand side

$$H(y) = \int dy\, \frac{y+1}{2y} = \int dy \left[\frac{1}{2} + \frac{1}{2y}\right] = \frac{y}{2} + \frac{1}{2}\ln|y|$$

However, in this case it is not possible to solve $\frac{y}{2} + \frac{1}{2}\ln|y| = u$ explicitly, so we neither can write an explicit formula for the inverse function $H^{-1}(u)$, nor find the general solution $y(x)$ explicitly. In such a case it is conventional to say that the general solution to the ODE is given in *implicit form* by the relation $\frac{y}{2} + \frac{1}{2}\ln|y| = \frac{1}{2}x^2 + C$.

(**c**). To find the general solution valid for $y \neq 0$, we define $H(y) = \int \frac{dy}{3y^{2/3}} = y^{1/3}$, so that solving $H(y) = y^{1/3} = u$ defines the inverse function $H^{-1}(u) = u^3$. As $f(x) = 1$ we have on the right-hand side $\int f(x)\, dx = x + C$. Finally, the general solution is given by applying the inverse H^{-1} to the right-hand side: $y = H^{-1}(x + C) = (x + C)^3$.

It is easy to check by direct substitution that the heuristic "separation of variables" method indeed works perfectly, but a mathematician must be concerned with finding a justification of the correct results obtained by an ill-defined method. A mathematically legitimate way of solving (1.1) goes as follows. Let the equation $g(y) = 0$ have distinct real roots $y = y_1 < y_2 < y_3 \ldots$ so that $y(x) = y_1$, $y(x) = y_2$, etc. are solutions to (1.1) (which are called in this case *special solutions*). Consider now any open interval (A, B) which contains none of the roots $y_1, y_2, \ldots$. Then $g(y) \neq 0$ for any $y \in (A, B)$ (that is $g(y)$ retains its sign inside the interval). Then inside the interval we can rewrite (1.1) as

$$\frac{1}{g(y)}\, y' = f(x) \tag{1.3}$$

Define the function $H(y)$ via the indefinite integral:

$$H(y) = \int \frac{1}{g(y)} dy.$$

and consider a function of variable x defined as $H(y(x))$. Then using the chain rule of differentiation we have

$$\frac{d}{dx} H(y(x)) = \frac{dH}{dy}\frac{dy}{dx} = \frac{1}{g(y)}\, y' = f(x) \tag{1.4}$$

so that we conclude that $H(y(x))$ is an antiderivative of $f(x)$, hence

$$H(y(x)) = \int f(x)\,dx + C\,.$$

Since $g(y)$ retains its sign in (A, B) the derivative $\frac{dH}{dy} = \frac{1}{g(y)}$ is of the same sign in the interval. Therefore the function $H(y)$ is either strictly increasing, or strictly decreasing in (A, B), hence it has a unique functional inverse H^{-1} inside that interval. The general solution $y(x)$ of (1.1) is therefore given by

$$y(x) = H^{-1}\left(\int f(x)\,dx + C\right) \tag{1.5}$$

and indeed coincides with one predicted by the heuristic method.

1.2 First order ODEs which can be reduced to be separable

1. Consider equations of the type

$$y' = f(ax + by + c), \quad \text{where } a, b, c \text{ are real constants} \tag{1.6}$$

Introducing a new function $z(x) = ax + by + c$ we see that this equation can be rewritten as $y' = f(z)$. Then (1.6) becomes equivalent to

$$z' = a + by' = a + bf(z), \tag{1.7}$$

which is a particular type of separable equation (1.1).

Example:

Solve the equation

$$y' = (4y - x - 6)^2$$

Solution:

We introduce $z = 4y - x - 6$ so that the equation can be written as $y' = z^2$. Then $z' = 4y' - 1 = 4z^2 - 1$ which is a separable ODE. Separating variables we get

$$\int \frac{dz}{4z^2 - 1} = \int dx + C$$

and performing the integration in the left-hand side as:

$$H(z) = \frac{1}{2}\int \left(\frac{1}{2z - 1} - \frac{1}{2z + 1}\right) dz$$

we see that

$$H(z) = \frac{1}{4}\left(\ln|2z-1| - \ln|2z+1|\right) = \frac{1}{4}\ln\left|\frac{2z-1}{2z+1}\right|$$

The inverse function $H^{-1}(u)$ is obtained by solving $H(z) = u$, that is

$$\frac{1}{4}\ln\left|\frac{2z-1}{2z+1}\right| = u \qquad \Leftrightarrow \qquad \left|\frac{2z-1}{2z+1}\right| = e^{4u}$$

Solving for z (exercise for yourself!) gives explicitly two possible solutions

$$z(u) = \frac{1}{2}\frac{1+e^{4u}}{1-e^{4u}} \quad \text{or} \quad z(u) = \frac{1}{2}\frac{1-e^{4u}}{1+e^{4u}}.$$

Denoting the functions on the right-hand side as $z = H^{-1}(u)$ we see that the solution $z(x)$ is given in either case by

$$z(x) = H^{-1}(x+C) = \frac{1}{2}\frac{1 \pm e^{4(x+C)}}{1 \mp e^{4(x+C)}}$$

It is convenient to write $\pm e^{4C} = A$, where the constant A may have an arbitrary sign. Finally, using the definition of z we see that y is expressed in terms of the above z via

$$y = \frac{1}{4}(z(x) + x + 6) = \frac{1}{4}\left(x + 6 + \frac{1}{2}\frac{1+Ae^{4x}}{1-Ae^{4x}}\right)$$

which gives the general solution to the original ODE.

Queen Mary
University of London

MTH5123 Differential Equations

Lecture Notes

Week 2

School of Mathematical Sciences
Queen Mary University of London

2. Consider equations of the type

$$y' = F\left(\frac{y}{x}\right) \tag{1.8}$$

Such ODEs do not change if we rescale $x \to kx$ and $y \to ky$ for any real constant factor $k \neq 0$, hence they are known under the name **scale-invariant** first order ODEs. To reduce them to separable equations one introduces a new function $z(x) = y(x)/x$ which implies $y(x) = xz(x)$. Differentiating this equation gives $y' = z(x) + xz'(x)$, and (1.8) can be rewritten in the form $z + xz' = F(z)$ or equivalently

$$z' = \frac{1}{x}\left[F(z) - z\right] \tag{1.9}$$

which is indeed separable.

Example:

Solve the equation

$$xy' = y - xe^{y/x}.$$

Solution:

After dividing both sides by x we see that the equation is of the form (1.8) with the right-hand side $F(z) = z - e^z$. Therefore it is equivalent to the separable equation

$$z' = -\frac{1}{x}e^z\,.$$

Solving it by standard means leads to $e^{-z} = \ln|x| + C$ or

$$z = -\ln\left(\ln|x| + C\right)\,.$$

Finally the general solution to the original ODE is

$$y(x) = -x\ln\left(\ln|x| + C\right)\,.$$

1.3 First order linear ODEs

This class of equations is given by

$$y' = A(x)\,y + B(x)\,, \tag{1.10}$$

where the two functions $A(x) \neq 0$ and $B(x)$ are known. These equations are called **linear** (in y), because y and its derivative y' occur only to the first power, they are not multiplied together, nor do they appear as the argument of a function (such as $\sin y, \exp(y)$, etc.). If $B(x) = 0$, the equation is called **homogeneous**, if $B(x) \neq 0$ it is called **inhomogeneous**.

Example:

$y' = \sin(x)y$ homogeneous
$y' = e^x y + x$ inhomogeneous
$y' = 1 - y^2 + x$ nonlinear

The method of solution of such equations proceeds in two steps:

Step 1: Solve the *homogeneous* equation $y' = A(x)\, y$, which is separable. The general solution is found to be

$$\int \frac{dy}{y} = \int A(x)\, dx + C \Rightarrow \ln|y| = \int A(x)\, dx + C \tag{1.11}$$

and finally

$$y = De^{\int A(x)\, dx}, \tag{1.12}$$

where $D \in \mathbb{R}$ is an arbitrary real constant (also called a free parameter).

Step 2 is known as the **variation of parameter** method. It amounts to looking for the solution of (1.10) in the form

$$y = D(x)\, e^{\int A(x)\, dx}, \tag{1.13}$$

where $D(x)$ is now an unknown function to be determined by substituting (1.13) to (1.10). This gives

$$y' = D'(x)\, e^{\int A(x)\, dx} + A(x)D(x)\, e^{\int A(x)\, dx} = A(x)D(x)\, e^{\int A(x)\, dx} + B(x)\,,$$

which after cancelling equal terms on both sides is equivalent to

$$D'(x)\, e^{\int A(x)\, dx} = B(x). \tag{1.14}$$

This allows us to write $D'(x) = e^{-\int A(x)\, dx}\, B(x)$ and to recover $D(x)$ by simple integration

$$D(x) = \int e^{-\int A(x)\, dx}\, B(x)\, dx + C \tag{1.15}$$

finally yielding the general solution of (1.10) in the form

$$y(x) = e^{\int A(x)\, dx} \left(\int e^{-\int A(x)\, dx}\, B(x)\, dx + C \right) \quad \forall C \in \mathbb{R} \tag{1.16}$$

Note: An alternative method to derive the same result is the *integrating factor method*, as you have seen in Calculus 2.

Example:
Solve the equation

$$y' + 2xy = x.$$

Solution:
First we solve $y' + 2xy = 0$ by separation of variables obtaining $y = De^{-x^2}$, where D is an arbitrary constant. Now we assume $D = D(x)$ and substitute $y = D(x)e^{-x^2}$ to the full non-homogeneous equation:

$$y' = D'(x)e^{-x^2} + D(x)(-2x)e^{-x^2}.$$

Thus, we have

$$D'(x)e^{-x^2} + D(x)(-2x)e^{-x^2} + 2xD(x)e^{-x^2} = x$$

which implies $D'(x) = xe^{x^2}$, hence $D(x) = \int xe^{x^2}\, dx = \frac{1}{2}e^{x^2} + C$. Finally, the general solution to the original ODE is given by

$$y(x) = \left(\frac{1}{2}e^{x^2} + C \right) e^{-x^2} = \frac{1}{2} + Ce^{-x^2}.$$

1.4 Exact first order ODEs.

Exact ODEs are of the form

$$P(x,y) + Q(x,y)\,\frac{dy}{dx} = 0\,. \tag{1.17}$$

We would like to find solutions of this class of ODEs in *implicit form* $F(x,y) = C\,,\ y = y(x)$, for a constant C. Using the chain rule we observe that

$$\frac{dF(x,y(x))}{dx} = \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y}\,\frac{dy}{dx} = 0\,, \tag{1.18}$$

which coincides with (1.17) if we define

$$P(x,y) = \frac{\partial F}{\partial x}, \quad Q(x,y) = \frac{\partial F}{\partial y}\,. \tag{1.19}$$

Using these definitions we have

$$\frac{\partial}{\partial y}P(x,y) = \frac{\partial^2 F}{\partial y \partial x}, \quad \frac{\partial}{\partial x}Q(x,y) = \frac{\partial^2 F}{\partial x \partial y}. \tag{1.20}$$

If F is twice differentiable in both x and y with continuous second order partial derivatives, we have (according to the *mixed derivatives theorem* in Calculus 2)

$$\frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y}\,,$$

and we conclude that the equation

$$\frac{\partial}{\partial y}P(x,y) = \frac{\partial}{\partial x}Q(x,y) \tag{1.21}$$

must hold. Equation (1.21) is the crucial condition for (1.17) to be **exact**. For any exact ODE the general solution can always be written in the implicit form $F(x,y) = C$.
To determine the form of the function $F(x,y)$, one may start with the first equation in (1.19) by integrating it over the variable x to

$$P(x,y) = \frac{\partial F}{\partial x} \quad \Rightarrow F(x,y) = \int P(x,y)dx + g(y)\,, \tag{1.22}$$

where the function $g(y)$ is an arbitrary function of the variable y, yet to be determined. To find $g(y)$ we use the second equation in (1.19)

$$Q(x,y) = \frac{\partial F}{\partial y} = \frac{\partial}{\partial y}\int P(x,y)dx + g'(y)\,, \tag{1.23}$$

which gives

$$g'(y) = Q(x,y) - \frac{\partial}{\partial y}\int P(x,y)dx \tag{1.24}$$

The missing function $g(y)$ can then be found by straightforward integration of this equation.

Example:
Show that the equation

$$3x^2 + y - (3y^2 - x)\frac{dy}{dx} = 0$$

is exact and find its general solution in implicit form.

Solution:
We identify $P(x, y) = 3x^2 + y$, hence $\frac{\partial P}{\partial y} = 1$. Similarly, $Q(x, y) = -(3y^2 - x)$, hence $\frac{\partial Q}{\partial x} = 1$. Since $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$ the equation is exact.
We find its implicit solution in the form $F(x, y) = C$ by

$$F(x, y) = \int P(x, y)dx + g(y) = \int (3x^2 + y)\, dx + g(y) = x^3 + xy + g(y)\ ,$$

where $g(y)$ is yet undetermined. We further have

$$\frac{\partial F}{\partial y} = x + g'(y) = Q(x, y) = -(3y^2 - x), \quad \Rightarrow g'(y) = -3y^2\ .$$

This allows us to find

$$g(y) = \int (-3y^2)\, dy = -y^3 + C_1\ ,$$

where C_1 is an arbitrary constant. There is no need to keep C_1, as it can always be absorbed into the constant C. The general solution of the original equation in implicit form is obtained as

$$F(x, y) = x^3 + xy - y^3 = C.$$

Note:
The same ODE can be presented in a different form, for example:

$$\frac{dy}{dx} = \frac{3x^2 + y}{3y^2 - x}$$

One needs to recognize the equivalence of this equation to the form of an exact ODE by then applying the same procedure for a solution.

MTH5123 Differential Equations

Lecture Notes

Week 3

School of Mathematical Sciences
Queen Mary University of London

2 Initial value problem for first-order ODEs: existence and uniqueness of solutions

In the previous chapter we have studied a few types of ODEs whose solutions $y(x)$ or $y(t)$ we were able either to write down explicitly, or to characterize implicitly in terms of a given function of two variables, e.g. as $F(x, y) = C$. The class of ODEs for which this can be done is however rather small. Most ODEs which are encountered in practice can not be solved in this or similar ways. If, however, we have an ODE for which we know that a solution **exists**, we may proceed to investigate its properties (e.g., the behaviour of $y(t)$ for large $t \to \infty$) *regardless of whether we know the explicit form for the solution.* The corresponding methods are known as a *qualitative study* of ODEs, and some of them will be briefly discussed later on in this module.
Another important aspect is the **uniqueness** of a solution. For this we first need to define the *initial value problem* for a first-order ODE. Such a problem asks to find a function $y(x)$ solving a given ODE by taking a specific value $y = b$ at the argument $x = a$. That is, the solution must satisfy the *initial value* $y(a) = b$. We have seen for first-order ODEs like $\frac{dy}{dx} = f(y, x)$ that the general solution always has a *single* free parameter, which is the constant of integration. Solving an initial value problem requires to determine this free parameter by using the given initial condition $y(a) = b$.

2.0 Initial Value Problem

Example:
Consider the first order separable ODE

$$\frac{dy}{dx} = \frac{1}{2y} \, .$$

Its solution is found to be $y(x) = \pm\sqrt{x + C}$ with arbitrary C. For this example, the initial condition $y(a) = b$ yields $b = \pm\sqrt{a + C}$ which is satisfied if $C = b^2 - a$. Hence the solution to the initial value problem is given by

$$y(x) = \begin{cases} \sqrt{x + b^2 - a}, & \text{if} \quad b > 0 \\ -\sqrt{x + b^2 - a}, & \text{if} \quad b < 0 \\ \text{both } \sqrt{x - a} \text{ and } -\sqrt{x - a}, & \text{if} \quad b = 0 \end{cases} .$$

We see that for $b = 0$ there are **two** solutions of the initial value problem, whereas for all other values of $b \neq 0$ there is **only one** solution. In the latter case we say that the solution is **unique**.
By a **unique solution** we mean the following:

Definition:
An initial value problem for an ODE with initial condition $y(a) = b$ has a **unique** solution if for any two solutions $y_1(x)$ and $y_2(x)$ satisfying the same initial condition $y_1(a) = y_2(a) = b$, there exists positive number $A > 0$ and a positive $B > 0$ such that

$$y_1(x) = y_2(x), \quad \forall x \in (a - A, a + A), y \in (b - B, b + B).$$

In other words, uniqueness implies that two such solutions are **identical** for all (x, y) in the region $\mathcal{D}$ $|x - a| < A$ and $|y - b| < B$ for some $A > 0, B > 0$.

We have already seen that for some initial conditions solutions to initial value problems may not be unique.

Example:
Consider the first order ODE

$$\frac{dy}{dx} = 3y^{2/3}$$

with initial condition $y(0) = 0$. Obviously, $y(x) = 0$ is a solution to this initial value problem. On the other hand, this is a separable ODE, which in fact we have solved before: By separating the variables we found $y^{1/3} = x + C$, hence we have a family $\mathcal{F}$ of solutions $y = (x + C)^3$. Obviously, $y = x^3$ belongs to this family and *also* satisfies the above initial condition. Consequently, the solution to our initial value problem $y(0) = 0$ is *not unique.*
Similarly, for the initial value problem $y(a) = 0$ there are two solutions passing through the point $(a, 0)$: $y(x) = 0$ and $y(x) = (x - a)^3$; see Fig. 2.1. Moreover, the curve $M_1N_1 \bigcup N_1N_2 \bigcup N_2K_2$, which is a mix between both these basic solutions, is another solution, as the left and the right derivatives at N_1 and N_2 are equal.

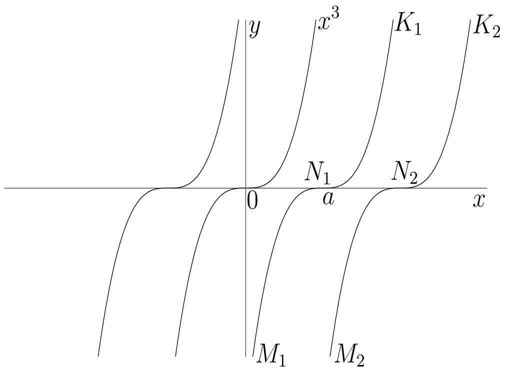

Figure 2.1: Sketch of the solutions $y = (x + C)^3$ to the ODE $y' = 3y^{2/3}$.

2.0.1 Picard-Lindelöf Existence and Uniqueness Theorem for I.V.P

Uniqueness of solutions is important for a number of reasons. Suppose we are able to find by some technique a one-parameter family $\mathcal{F}$ of solutions $y = y_{\mathcal{F}}(x)$ of a first-order ODE. Furthermore, suppose for any point (a, b) in some domain $\mathcal{D}$ of the xy plane we can always find a solution in our family $\mathcal{F}$ which satisfies $y_{\mathcal{F}}(a) = b$. If we know that uniqueness holds, then our family $\mathcal{F}$ of solutions must contain **all** desired solutions to our ODE, and we need look no further for other solutions.
Uniqueness will also be of importance if, for instance, we wanted to approximate a solution numerically. If two *different* solutions passed through a point, then successive approximations could very well jump from one solution to the other - with misleading consequences. It is therefore important to know under which conditions one can expect an ODE to have a unique solution for a specified initial condition. In case of first-order ODEs the answer is largely given by the **Picard-Lindelöf Existence and Uniqueness Theorem**, which we state without proof:

Theorem: Picard-Lindelöf Existence and Uniqueness Theorem
Consider the initial value problem

$$\frac{dy}{dx} = f(x, y) \quad \text{with} \quad y(a) = b\,. \tag{2.1}$$

Consider the IVP (2.1) on a rectangular domain D of the form $|x - a| \le A$ and $|y - b| \le B$ (see Fig. 2.2), then it has **one and only one solution** in $\mathcal{D}$ provided the following two conditions are satisfied:

- The function $f(x, y)$ is *continuous* in $\mathcal{D}$ and therefore bounded: $|f(x, y)| \le M\ \forall (x, y) \in \mathcal{D}$ for some positive constant $M > 0$. We also have to impose the restriction $A \le B/M$ on the width of $\mathcal{D}$.
- It has *bounded derivative* $\frac{\partial f}{\partial y}$ everywhere in $\mathcal{D}$, that is, the value $K = \max_{(x,y)\in\mathcal{D}} \left|\frac{\partial f}{\partial y}\right|$ is finite: $0 < K < \infty$.

Note:

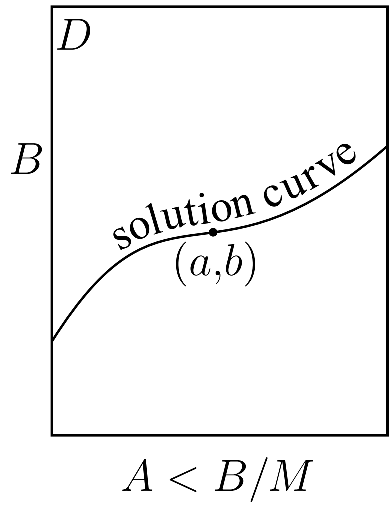

Figure 2.2: Sketch of a a rectangular space $\mathcal{D} : \{a - A \leq x \leq a + A, \quad b - B \leq y \leq b + B\}$, where uniqueness is ensured due to the Picard-Lindelöf theorem.

To see where the condition $A \leq B/M$ comes from one needs to go through the proof.
If the derivative $\frac{\partial f}{\partial y}$ is *continuous* everywhere in $\mathcal{D}$ then it is necessarily bounded.
The second condition is known as the *Lipschitz condition* and the constant K as the *Lipschitz constant*.

Example:
For the example of non-uniqueness given above, with $f(y) = 3y^{2/3}$ we have $\frac{\partial f}{\partial y} = 2y^{-1/3}$ which *diverges* close to $y = b = 0$ thus violating the Lipschitz condition and hence invalidating the Picard-Lindelöf theorem.

Note:
It is important to understand that the existence and uniqueness properties of the solution are guaranteed by the Picard–Lindelöf theorem only *locally*, i.e., sufficiently close to the point (a, b). In no way it implies existence and uniqueness everywhere.

Example:
Consider the initial value problem

$$\frac{dy}{dx} = f(x, y), \quad y(0) = 1\,,$$

where $f(x, y)$ is defined as

$$f(x, y) = x^2 |y|^{\frac{1}{3}},$$

i) Explain why the Picard-Lindelöf Theorem guarantees the existence and uniqueness of the solution to the above I.V.P on a rectangular domain $\mathcal{D} = (|x| \leq A,\ |y - 1| \leq B)$ in the xy plane only for heights B satisfying $0 < B < 1$. Find the value of the Lipschitz constant K for the above problem for given A and B.

ii) Suppose that the height B of $\mathcal{D}$ is given and satisfies $0 < B < 1$. Show that the width A must then satisfy the inequality

$$0 < A \leq \frac{B^{1/3}}{(1+B)^{1/9}}$$

Solution:

i) The right-hand side of $f(x, y)$ depends on $|y|$, hence formally it is defined separately for positive and negative values of y:

$$f(x,y) = \begin{cases} x^2 y^{1/3}, & y > 0 \\ x^2(-y)^{1/3}, & y < 0 \end{cases}$$

This polynomial expression is continuous everywhere in $\mathcal{D}$, as nothing goes wrong at $y = 0$:

$$\lim_{y\to 0^+} f(x,y) = \lim_{y\to 0^-} f(x,y) = 0\,.$$

On the other hand, the derivative

$$\frac{\partial f}{\partial y} = \begin{cases} \frac{1}{3}x^2 y^{-2/3} & \text{for} \quad y > 0 \\ -\frac{1}{3}x^2(-y)^{-2/3} & \text{for} \quad y < 0 \end{cases}$$

is defined and finite everywhere *except* at $y = 0$ where it diverges. Therefore this point has to be *excluded* from $\mathcal{D}$ to ensure the conditions of the Picard-Lindelöf Theorem. This means the interval $|y - 1| \leq B$ (or, equivalently, $y \in [1 - B, 1 + B]$ with $B > 0$) cannot contain $y = 0$, which is only possible for $0 < B < 1$.

The Lipschitz constant K can be found according to

$$K = \max_{(x,y)\in\mathcal{D}} \left|\frac{\partial f}{\partial y}\right|$$

$$= \max\left[x^2\right]_{-A\leq x\leq A} \cdot \max\left[\frac{1}{3}|y|^{-2/3}\right]_{1-B\leq y\leq 1+B} = \frac{A^2}{3(1-B)^{2/3}}$$

ii) The modulus of the function $|f(x, y)| = x^2|y|^{1/3}$ on the right-hand side of the ODE grows with both $|x|$ and $|y|$, hence for a given B its maximum M in $\mathcal{D}$ is achieved for $x = \pm A$ and $y = 1 + B$:

$$M = \max_{(x,y)\in D}|f(x, y)| = A^2(1 + B)^{1/3}$$

This in turn implies that the width $A > 0$ should satisfy $A \leq B/M = \frac{B}{A^2(1+B)^{1/3}}$. Rearranging gives

$$A^3 \leq \frac{B}{(1+B)^{1/3}} \quad \text{or} \quad 0 < A \leq \frac{B^{1/3}}{(1+B)^{1/9}}\,.$$

2.1 Systems of first-order ODEs

Our main goal in this section is to extend the previous theory of existence and uniqueness to systems of n first-order ODEs for n unknown functions $y_1(t), y_2(t), \ldots, y_n(t)$ written in the **normal form**

$$\begin{cases} \dot{y}_1 = f_1(t, y_1, \ldots, y_n) \\ \dot{y}_2 = f_2(t, y_1, \ldots, y_n) \\ \qquad \ldots \\ \qquad \ldots \\ \dot{y}_n = f_n(t, y_1, \ldots, y_n) \end{cases} \tag{2.2}$$

For sake of brevity we consider only the case $n = 2$. Using vector notation we can write down any such system in normal form as

$$\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}), \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} f_1(t, y_1, y_2) \\ f_2(t, y_1, y_2) \end{pmatrix} . \tag{2.3}$$

By imposing the **initial conditions** at $t = a$

$$y_1(a) = b_1 \ , \ \ y_2(a) = b_2 \ , \tag{2.4}$$

where b_1 and b_2 are given constants, we can formulate the **initial value problem** for the system (2.3).

The equivalence of n first-order ODEs to 1 n-th order ODE

The system (2.2) is *very fundamental*, as *any* n-th order ODE of the form

$$\frac{d^n y}{dx^n} = F\left(x, y, \frac{dy}{dx}, \frac{d^2 y}{dx^2}, \ldots, \frac{d^{n-1} y}{dx^{n-1}}\right) \tag{2.5}$$

is in fact equivalent to a system of the form (2.2). To see this, introduce the notation $t \equiv x$ and further denote

$$y_1(t) \equiv y(x), \ y_2(t) \equiv \frac{dy}{dx}, \ y_3(t) \equiv \frac{d^2 y}{dx^2}, \ \ldots, y_n(t) \equiv \frac{d^{n-1} y}{dx^{n-1}} \ . \tag{2.6}$$

The functions $y_1(t), \ldots, y_n(t)$ then satisfy the $(n-1)$ relations

$$\dot{y}_1(t) = y_2(t), \ \dot{y}_2(t) = y_3(t), \ \ldots, \dot{y}_{n-1}(t) = y_n(t) \tag{2.7}$$

such that (2.5) takes the form, in the new notation,

$$\dot{y}_n(t) = F\left(t, y_1, y_2, y_3, \ldots, y_n\right) \ . \tag{2.8}$$

Equations (2.7)-(2.8) are exactly of the form of (2.2) with

$$f_1(t, \mathbf{y}) = y_2, \quad f_2(t, \mathbf{y}) = y_3, \quad \ldots \tag{2.9}$$

$$\ldots, \quad f_{n-1}(t, \mathbf{y}) = y_n \ , \ f_n(t, \mathbf{y}) = F\left(t, y_1, y_2, y_3, \ldots, y_n\right) \ .$$

Example:
Transform the second order ODE

$$\frac{d^2y}{dx^2} = 6y - 4\frac{dy}{dx}$$

to a system of first-order ODEs.

Solution:
By using $t \equiv x$, identify according to (2.6)

$$y_1(t) \equiv y(x) \ , \ y_2(t) \equiv \frac{dy}{dx} \ .$$

Differentiate these equations, cf. (2.7):

$$\dot{y}_1(t) = y' = y_2(t) \ , \ \dot{y}_2(t) = y'' \ .$$

According to (2.8) we then obtain the system of two first-order ODEs

$$\dot{y}_1(t) = y_2(t) \ , \ \dot{y}_2(t) = 6y - 4y' = 6y_1 - 4y_2 \ .$$

This example belongs to an important special class of n-th order ODEs, linear ODEs of second order (here with constant coefficients), which we will discuss in more detail in the next chapter.

MTH5123 Differential Equations

Lecture Notes

Week 4

School of Mathematical Sciences
Queen Mary University of London

2.2 Linear ODEs of second order: general facts

2.2.1 General form and Initial Value Problem of Linear second-order ODEs

Linear ODEs of second order are equations of the form

$$a_2(x)\frac{d^2y}{dx^2} + a_1(x)\frac{dy}{dx} + a_0(x)y = f(x)\,, \tag{2.5}$$

where $a_0(x), a_1(x), a_2(x)$ as well as $f(x)$ are assumed to be continuous real functions in some interval $A < x < B$, and $a_2(x) \neq 0$. A *general solution* of this ODE is the set of functions containing *all possible solutions* of that equation.

The Initial value problem for this ODE is composed by the ODE itself and initial conditions $y(a) = b_1$ and $y'(a) = b_2$, which refer to the value of $y(x)$ and $y'(x)$ when the independent variable $x = a$. Rewriting this 2nd-order ODE as a system of two 1st-order ODEs in normal form (see equation (2.3) before and the corresponding example in week 3), it follows from the generalised Picard-Lindelöf Theorem that for *any* $a \in (A, B)$ and *any* initial conditions

$$y(a) = b_1, \quad y'(a) = b_2 \tag{2.6}$$

the ODE (2.5) has *one and only one* solution.

*Reference notes for better understanding, not examinable

Last week (week 3), we showed that using vector notation we can write down a system of 2 first-order ODEs in normal form as $\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y})$, $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $\mathbf{f} = \begin{pmatrix} f_1(t, y_1, y_2) \\ f_2(t, y_1, y_2) \end{pmatrix}$. By imposing the initial conditions at $t = a$ $y_1(a) = b_1$, $y_2(a) = b_2$, where b_1 and b_2 are given constants, we have the initial value problem for the system. In this setting, we have the following generalisation of the **Picard–Lindelöf Theorem:**

Theorem: The initial value problem for the has **one and only one solution** in a cuboid domain $\mathcal{D}$ of the form $|t - a| \leq A$ and $|y_1 - b_1| \leq B_1$, $|y_2 - b_2| \leq B_2$ provided the functions $f_1(t, y_1, y_2), f_2(t, y_1, y_2)$ and all the partial derivatives $\frac{\partial f_i}{\partial y_j}$ for all choices of i, j are continuous in $\mathcal{D}$.

According to the transformation in week 3, we can rewrite the linear 2nd-order ODE as a system of two 1st-order ODEs as

$$\dot{y}_1 = y_2,$$
$$\dot{y}_2 = -\frac{a_0(x)}{a_2(x)}y_1 - \frac{a_1(x)}{a_2(x)})y_2 + \frac{f(x)}{a_2(x)}.$$

Example:

Solving the second order ODE derived from Newton's Second Law $m\ddot{y} = f(t, y, \dot{y})$ for a particle of mass m proceeds by specifying the initial *position* of the particle, $y(t = 0) = y_0$, and the initial *velocity*, $\dot{y}(t = 0) = v_0$. This uniquely determines the evolving dynamics.

Linearity of 2nd-order ODEs

We will use the short-hand notation $\mathcal{L}(y)$ for the left-hand side of (2.5) so that this equation can be written as

$$\mathcal{L}(y) = f \ , \ \mathcal{L}(y) \equiv a_2(x)\frac{d^2y}{dx^2} + a_1(x)\frac{dy}{dx} + a_0(x)y \ . \tag{2.7}$$

If $f(x) = 0$ the equation

$$\mathcal{L}(y) = 0 \tag{2.8}$$

is called **homogeneous**, if $f(x) \neq 0$

$$\mathcal{L}(y) = f \tag{2.9}$$

inhomogeneous. **Linearity** on the left hand side of (2.5) is represented by the property

$$\mathcal{L}(c_1 y_1 + c_2 y_2) = c_1 \mathcal{L}(y_1) + c_2 \mathcal{L}(y_2) \tag{2.10}$$

for any two (two-times differentiable) functions $y_1(x)$ and $y_2(x)$ and arbitrary constant parameters $c_1, c_2 \in \mathbb{R}$. Linearity implies that if $y_1(x), y_2(x), \ldots, y_k(x)$ are any k solutions of the *homogeneous* equation, then *any linear combination*

$$c_1 y_1(x) + c_2 y_2(x) + \ldots + c_k y_k(x)$$

is a solution of the same equation for an arbitrary choice of constant parameters c_k. Using this property one can prove the following

Theorem: Theorem based on the linearity to find the general solution of inhomogeneous ODEs

Suppose that $y_p(x)$ is a *particular solution* of the *inhomogeneous* equation (2.9) and that $y_h(x)$ is the *general solution* of the corresponding *homogeneous* equation (2.8). Then the *general solution* $y_g(x)$ of the *inhomogeneous equation* is given by

$$y_g(x) = y_h(x) + y_p(x)\,. \tag{2.11}$$

Proof:
Suppose $y_g(x)$ solves (2.9). As $y_p(x)$ is also a solution of this equation we have

$$\mathcal{L}\left(y_g\right) = f \quad \text{and} \quad \mathcal{L}\left(y_p\right) = f\,.$$

Taking the difference $y_g(x) - y_p(x)$ we get

$$\mathcal{L}\left(y_g - y_p\right) = \mathcal{L}(y_g) - \mathcal{L}(y_p) = f - f = 0$$

because of linearity. This means that $y_h = y_g - y_p$ must solve the corresponding homogeneous equation (2.8), hence trivially $y_g = y_h + y_p$. But if we take for y_h the *general solution* of the homogeneous equation, accordingly y_g must yield the general solution of the inhomogeneous one.

Note:
This theorem implies that to find the general solution of the inhomogeneous equation (2.9) we first need to find the general solution $y_h(x)$ of the corresponding homogeneous equation (2.8) and then a particular solution $y_p(x)$ of the inhomogeneous equation. The general solution is then obtained by simply adding them together.

2.2.2 Linear 2nd-order ODEs with constant coefficients, $a_2, a_1, a_3 \in \mathbb{R}$

We are considering equation (2.5), where a_0, a_1, a_2 are now given real numbers,

$$a_2 \frac{d^2y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = f(x)\,. \tag{2.12}$$

This equation defines an especially important type of ODE, which allows a full analysis and explicit solutions (not only for $n = 2$ but for similar equations of any order).
A general solution to such second order ODEs must contain two arbitrary constants (or free parameters). The values of these parameters will be fixed if we impose **initial conditions** involving the specification of the value $y(a)$ of the function at some point $x = a$ as well as the first derivative at the same point, $\frac{dy}{dx}|_{x=a}$, see (2.6).
Note that for the ODE (2.12) with constant coefficients the conditions of the Picard-Lindelöf theorem are satisfied *everywhere* (because rewritten as a system of first order ODEs, by definition (2.12) yields functions on the right-hand side of this system of ODEs that fulfill the conditions of this theorem). Hence the solution to any initial value problem is *unique* for $x \in (-\infty, \infty)$. Due to this fact, if one is able to find a family of solutions to such an equation depending on two free parameters such that one is able to satisfy arbitrary initial conditions, one can be sure that this family represents the general solution to this problem.

2.2.2.1 Linear homogeneous 2nd-order ODEs with constant coefficient, $f(x) = 0$

As a special case of (2.8), if $f(x) = 0$ in (2.12) we have the *homogeneous* equation

$$a_2 \frac{d^2y}{dx^2} + a_1 \frac{dy}{dx} + a_0 y = 0\,. \tag{2.13}$$

We will look for solutions in the form of $y = e^{\lambda x}$. Substituting this ansatz into (2.13) and canceling the common factor $e^{\lambda x} \neq 0$ we obtain the **characteristic equation**

$$M_2(\lambda) = a_2\lambda^2 + a_1\lambda + a_0 = 0\,. \tag{2.14}$$

According to the **fundamental theorem of algebra** this *polynomial* equation of degree 2 (i.e., quadratic equation) must have exactly two **roots** including *multiplicities* (for example, $\lambda^2 - 2\lambda + 1 = (\lambda - 1)^2 = 0$ has a single root $\lambda = 1$ of multiplicity two, or equivalently, two equal roots $\lambda_1 = \lambda_2 = 1$). Roots may be either *real* or may come in *complex conjugate pairs*. Accordingly, we have three different cases:

1. We start with the simplest situation where the roots are *real* and, moreover, *distinct*: $\lambda_1 \neq \lambda_2 \in \mathbb{R}$ and $M_2(\lambda_j) = 0,\ j = 1, 2$. In this case we prove that the general solution to (2.13) is given by the family

   $$y_h(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}\,, \tag{2.15}$$

   where c_1, c_2 are real constants.
   First of all, each $y_i(x) = e^{\lambda_i x}$, $i = 1, 2$ is a solution. Hence linearity, cf. (2.10), implies that the linear combination (2.15) is indeed a solution to (2.13). To see that this family provides the general solution we check that using the solution (2.15) one can satisfy *any* initial condition of the form

   $$y(0) = b_1, \quad y'(0) = b_2\,. \tag{2.16}$$

   In other words, one can always find a set of coefficients c_1, c_2 that the solution $y_h(x)$ for such a choice will satisfy (2.16). Plugging (2.16) into (2.15), one sees that the initial conditions generate a system of two linear equations for two unknown coefficients,

   $$\begin{aligned} c_1 + c_2 &= b_1 \\ \lambda_1 c_1 + \lambda_2 c_2 &= b_2\,. \end{aligned} \tag{2.17}$$

   As you know from your course of Linear Algebra, such a system has a (unique) solution provided the matrix of coefficients

   $$\Lambda_2 = \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} \tag{2.18}$$

   is *non-singular* (i.e., invertible), which is equivalent to the condition $D_2 = \det \Lambda_2 \neq 0$. We see that

   $$D_2 = \det \Lambda_2 = \det \begin{pmatrix} 1 & 1 \\ \lambda_1 & \lambda_2 \end{pmatrix} = (\lambda_2 - \lambda_1)\,,$$

   which is obviously nonzero due to our assumption that $\lambda_1 \neq \lambda_2$.

Note:

In all cases considered below it is possible to proceed in a similar way and to demonstrate that arbitrary initial conditions (2.16) can always be satisfied with an appropriate choice of parameters in the proposed general solution. Explicit verification of that fact is however somewhat lengthy and will not be performed in these lecture notes.

2. Suppose now that the roots of the characteristic equation (2.14) come in complex conjugate pairs,
$$\lambda_1 = \alpha + i\beta\,,\ \lambda_2 = \alpha - i\beta \quad \text{with} \quad \beta \neq 0\,. \tag{2.19}$$
The corresponding general real solution of the homogeneous equation (2.13) is given by
$$y_h(x) = (A\cos(\beta x) + B\sin(\beta x))\, e^{\alpha x}\,, \tag{2.20}$$
where A, B are real constants. Alternatively, the solution can be written down in the form of (2.15),
$$y_h(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}\,, \tag{2.21}$$
but here with *complex* coefficients c_i (by imposing the condition of $y_h(x)$ to be real). The equivalence of these two forms can be shown by using **Euler's formula**,
$$e^{i\theta} = \cos\theta + i\sin\theta, \quad \forall\theta \in \mathbb{R}\,. \tag{2.22}$$
Applying it to (2.19) and (2.21) in the form of $e^{\pm i\beta x} = \cos\beta x \pm i\sin\beta x$ yields
$$y_h(x) = [(c_1 + c_2)\cos\beta x + i\,(c_1 - c_2)\sin\beta x]\, e^{\alpha x}\,.$$
Matching this result to (2.20), we see that if we set $A = c_1 + c_2$ and $B = i(c_1 - c_2)$ with A, B real (or, equivalently, $c_1 = (A - iB)/2\,,\ c_2 = (A + iB)/2$), this is equivalent to (2.20).

3. Finally, suppose that the root is real with multiplicity two, $\lambda_1 = \lambda_2 = \lambda$. It is left as an exercise to you to check that the general solution is given by
$$y_h(x) = (c_2 x + c_1)e^{\lambda x}$$
with two real constants c_2, c_1.

MTH5123 Differential Equations

Lecture Notes

Week 5

School of Mathematical Sciences
Queen Mary University of London

Autumn 2021

2.2.2.2 Euler type equation

Note:
Certain second-order linear ODEs with *nonconstant* coefficients can be reduced to corresponding ODEs with constant coefficients by special substitutions. Consider, for example, the *Euler type* equation

$$ax^2y'' + bxy' + cy = 0\ ,\ x > 0,\ a, b, c = \text{const.}\ . \tag{2.28}$$

This equation can be reduced to one with constant coefficients by introducing the new variable $x = e^t$ so that $y(x) = y[x(t)] = z(t)$. By the chain rule we have

$$\dot{z} = \frac{dz}{dt} = \frac{dy}{dx}\frac{dx}{dt} = \frac{dy}{dx}e^t \ \Rightarrow\ y' = e^{-t}\dot{z}\,. \tag{2.29}$$

Differentiating z another time yields

$$\ddot{z} = \frac{d}{dt}\dot{z} = \frac{d}{dt}\left[\frac{dy}{dx}e^t\right] = \left(\frac{d}{dt}\frac{dy}{dx}\right)e^t + \frac{dy}{dx}e^t = \frac{d^2y}{dx^2}\frac{dx}{dt}e^t + \frac{dy}{dx}e^t = \frac{d^2y}{dx^2}e^{2t} + \frac{dy}{dx}e^t\,.$$

Solving this equation for $y'' = \frac{d^2y}{dx^2}$ gives

$$y'' = e^{-2t}\left(\ddot{z} - \dot{z}\right)\ . \tag{2.30}$$

Substituting (2.29)and (2.30) into (2.28) the latter is reduced to the equation with constant coefficients

$$a\ddot{z} + (b-a)\dot{z} + cz = 0\,, \tag{2.31}$$

which can be solved for $z(t)$ by the standard method given above. One then recovers the original solution $y(x)$ by $y(x) = z(t)|_{t=\ln x}$.

2.2.2.3 Linear inhomogeneous 2nd-order ODEs with constant coefficients and $f(x) \neq 0$

The general solution of the inhomogeneous equation (2.17) with $f(x) \neq 0$, viz. (2.14), can be recovered from the general solution of the corresponding homogeneous equation with $f(x) = 0$, viz. (2.13), by extending the *variation of parameter* method that we already applied successfully to solving first-order linear inhomogeneous equations.
As an example, we consider **second-order** ODE's of the form (2.17)

$$a_2\frac{d^2y}{dx^2} + a_1\frac{dy}{dx} + a_0y = f(x)\,. \tag{2.32}$$

First: find $y_h(x)$ to the corresponding homogenous ODE

The characteristic equation corresponding to (2.32) is given by $M_2(\lambda) = a_2\lambda^2 + a_1\lambda + a_0 = 0$. It has the two roots

$$\lambda_1 = \frac{-a_1 + \sqrt{a_1^2 - 4a_2a_0}}{2a_2}\ ,\ \lambda_2 = \frac{-a_1 - \sqrt{a_1^2 - 4a_2a_0}}{2a_2}\ , \tag{2.33}$$

which are real and distinct as long as $a_1^2 - 4a_2a_0 > 0$. Considering for simplicity only this case, we have learned that we can write the general solution to the *homogeneous* equation as

$$y_h(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}, \tag{2.34}$$

where c_1 and c_2 are two real constants.

Second: Find $y_p(x)$ based on The variation of parameter method

According to the variation of paramter method we will look for a solution of the full *inhomogeneous* equation (2.32) in the form of

$$y(x) = c_1(x)e^{\lambda_1 x} + c_2(x)e^{\lambda_2 x}\,. \tag{2.35}$$

We have to show that this ansatz works, and if so, whether it will yield a *particular* or possibly even the *general* solution. Differentiating (2.35) yields

$$\frac{dy}{dx} = c_1(x)\lambda_1 e^{\lambda_1 x} + c_2(x)\lambda_2 e^{\lambda_2 x} + c_1'(x)e^{\lambda_1 x} + c_2'(x)e^{\lambda_2 x} \tag{2.36}$$

with $c_{1,2}' \equiv \frac{dc_{1,2}}{dx}$. To simplify this expression before we proceed further, we impose the additional condition

$$c_1'(x)e^{\lambda_1 x} + c_2'(x)e^{\lambda_2 x} = 0 \tag{2.37}$$

on the two functions $c_{1,2}(x)$. This implies that

$$\frac{dy}{dx} = c_1(x)\lambda_1 e^{\lambda_1 x} + c_2(x)\lambda_2 e^{\lambda_2 x}\,, \tag{2.38}$$

which facilitates the following second differentiation

$$\frac{d^2y}{dx^2} = c_1(x)\lambda_1^2 e^{\lambda_1 x} + c_2(x)\lambda_2^2 e^{\lambda_2 x} + c_1'(x)\lambda_1 e^{\lambda_1 x} + c_2'(x)\lambda_2 e^{\lambda_2 x}\,. \tag{2.39}$$

Now we substitute (2.35), (2.38) and (2.39) into the left-hand side of (2.32) giving

$$a_2\frac{d^2y}{dx^2} + a_1\frac{dy}{dx} + a_0 y = c_1(x)e^{\lambda_1 x}\left(a_2\lambda_1^2 + a_1\lambda_1 + a_0\right) \tag{2.40}$$

$$+c_2(x)e^{\lambda_2 x}\left(a_2\lambda_2^2 + a_1\lambda_2 + a_0\right) + a_2\left(c_1'(x)\lambda_1 e^{\lambda_1 x} + c_2'(x)\lambda_2 e^{\lambda_2 x}\right)\,.$$

Remembering that both λ_1 and λ_2 are roots of the characteristic equation, that is $a_2\lambda_1^2 + a_1\lambda_1 + a_0 = 0$ and $a_2\lambda_2^2 + a_1\lambda_2 + a_0 = 0$ we see that (2.32) and (2.40) together imply the relation

$$c_1'(x)\lambda_1 e^{\lambda_1 x} + c_2'(x)\lambda_2 e^{\lambda_2 x} = f(x)/a_2\,. \tag{2.41}$$

Now we compare (2.37) and (2.41). Multiplying (2.37) with the factor $-\lambda_2$ and adding to (2.41) gives

$$c_1'(x)e^{\lambda_1 x}(\lambda_1 - \lambda_2) = f(x)/a_2\,, \tag{2.42}$$

which allows us to find $c_1(x)$ by straightforward integration,

$$c_1(x) = \frac{1}{(\lambda_1 - \lambda_2)a_2}\left(\int f(x)e^{-\lambda_1 x}dx + C_1\right)\,, \tag{2.43}$$

where C_1 is a real constant. Similarly, multiplying (2.37) with the factor λ_1 and subtracting from (2.41) gives

$$c_2'(x)e^{\lambda_2 x}(\lambda_2 - \lambda_1) = f(x)/a_2 \,. \tag{2.44}$$

Hence

$$c_2(x) = -\frac{1}{(\lambda_1 - \lambda_2)a_2}\left(\int f(x)e^{-\lambda_2 x}dx + C_2\right) \,, \tag{2.45}$$

where C_2 is a real constant. Collecting everything together we find that *a solution* to the *inhomogeneous* equation (2.32) is given by

$$y(x) = \frac{1}{(\lambda_1 - \lambda_2)a_2}\left\{e^{\lambda_1 x}\left(\int f(x)e^{-\lambda_1 x}dx + C_1\right) - e^{\lambda_2 x}\left(\int f(x)e^{-\lambda_2 x}dx + C_2\right)\right\} \,. \tag{2.46}$$

Putting $f(x) = 0$ in (2.46) and introducing the notation

$$\frac{C_1}{\lambda_1 - \lambda_2} = c_1 \,, \; -\frac{C_2}{\lambda_1 - \lambda_2} = c_2$$

we see that the solution (2.46) reduces to

$$y_h(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} \,, \tag{2.47}$$

which is the *general* solution of the corresponding *homogeneous* equation. Accordingly, our corresponding solution $y(x)$ of the *inhomogeneous* equation can be written as $y_g(x) = y_h(x) + y_p(x)$, where

$$y_p(x) = \frac{1}{(\lambda_1 - \lambda_2)a_2}\left\{e^{\lambda_1 x}\int f(x)e^{-\lambda_1 x}dx - e^{\lambda_2 x}\int f(x)e^{-\lambda_2 x}dx\right\} \tag{2.48}$$

is a *particular* solution of the *inhomogeneous* equation. Hence, the variation of parameter method did indeed yield the general solution $y_g(x) = y(x)$ of (2.32).

Note:

1. Although the solution (2.46) (or equivalently the pair (2.47), (2.48)) was formally derived for *real* $\lambda_1 \neq \lambda_2$ it retains its validity for *complex conjugate* roots $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$ with $\beta \neq 0$ as long as one uses complex coefficients c_1, c_2. To bring the solution onto a real form one uses Euler's formula (2.27). One can even use (2.46) in the limit $\lambda_1 \to \lambda_2$ (that is, $\beta \to 0$) by using L'Hopital's rule, as will be demonstrated later on by an example.

2. Although the variation of parameter method is of general validity for arbitrary $f(x)$, its implementation relies on our ability to perform the integrals $\int f(x)e^{-\lambda_2 x}dx$ explicitly. In practical terms, for finding explicit forms of the solution it is sometimes easier to guess a *particular* solution $y_p(x)$ of the inhomogeneous equation and then to combine it with the *general solution* $y_h(x)$ of the corresponding homogeneous equation into the general solution $y_g(x) = y_h(x) + y_p(x)$ of the inhomogeneous one according to our theory.

2.2.2.4 Educated guess method for linear inhomogeneous 2nd-order ODEs with constant coefficients and $f(x) = P(x)e^{ax}$

If the right-hand side has the form $f(x) = P(x)e^{ax}$, where $P(x)$ is a polynomial of degree k, and $a \neq \lambda_1, a \neq \lambda_2$ (which means that e^{ax} is *not* a solution of the *homogeneous* equation), then a particular solution can *always* be found in the form $y_p(x) = Q(x)e^{ax}$ with some polynomial $Q(x) = d_k x^k + \ldots + d_1 x + d_0$ of the *same* degree. We may refer to such a method of finding particular solutions as the **educated guess** method.

Example:
Find a particular solution of the ODE

$$y'' + 2y' - 3y = xe^{2x} .$$

Solution: Here $a = 2$ and $P(x) = x$ is of first degree. First we need to check that e^{2x} is not a solution of $y'' + 2y' - 3y = 0$, which is indeed the case. Then we look for a solution in the form $y_p(x) = (d_1 x + d_0)e^{2x}$. Differentiating gives

$$y_p' = e^{2x}(2d_0 + d_1 + 2d_1 x) \, , \; y_p'' = e^{2x}(4d_0 + 4d_1 + 4d_1 x) \, ,$$

which by substitution into the left-hand side of the inhomogeneous equation and collecting similar terms yields

$$y_p'' + 2y_p' - 3y_p = e^{2x}(5d_0 + 6d_1 + 5d_1 x) \, .$$

Matching the coefficients to the right-hand side xe^{2x} we find $d_1 = 1/5$ and $d_0 = -\frac{6d_1}{5} = -6/25$. Thus a particular solution to the given ODE is

$$y_p(x) = \left(\frac{1}{5}x - \frac{6}{25} \right) e^{2x} \, .$$

Another version of the *educated guess* method exits in the case of two complex conjugate roots $\lambda_1 = \alpha + i\beta, \lambda_2 = \alpha - i\beta$. Here, if the right-hand side has the form $f(x) = P(x)\cos(ax)$ or $f(x) = P(x)\sin(ax)$, where $P(x)$ is a polynomial of degree k, and $ia \neq \lambda_1, ia \neq \lambda_2$ (which means that $e^{iax} = \cos(ax) + i\sin(ax)$ is not a solution of the homogeneous equation), then such a particular solution can always be found in the form

$$y_p(x) = Q(x)(A\cos(ax) + B\sin(ax)) \tag{2.49}$$

with some coefficients A, B and some polynomial $Q(x) = d_k x^k + \ldots + d_1 x + 1$ (note that the last coefficient of the polynomial can be chosen to be equal to one).

MTH5123 Differential Equations

Lecture Notes

Week 6

School of Mathematical Sciences
Queen Mary University of London

Autumn 2021

3 Boundary Value Problems for second-order Linear ODEs

3.1 Definition of B.V.P

So far we have considered the **initial value** problem for second order ODEs

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = f(x) \tag{3.1}$$

by specifying two conditions for the function $y(x)$ and its derivative at **one and the same** value of the independent variable $x = a$ (or, if the independent variable was interpreted as time t, the conditions were specified at $t = a$). We will always assume that all the coefficients $a_0(x), a_1(x), a_2(x)$ and the function $f(x)$ are continuous in some interval $[x_1, x_2]$, and $a_2(x) \neq 0$ in that interval. As we have discussed at the beginning of Section 2.1, according to the *generalised Picard-Lindelöf Theorem* any Initial Value Problem $y(a) = b$, $y'(a) = b_1$ for $a \in [x_1, x_2]$ has one and only one solution in the interval $[x_1, x_2]$.
In this section we are going to consider the different situation when some conditions are specified at the endpoints, or *boundaries*, of an interval of the independent variable, that is, at $x = x_1$ and $x = x_2$ with $x_1 < x_2$. This problem is known as a **Boundary Value Problem** and the conditions are called *boundary conditions*. We are then interested in finding the solution $y(x)$ to a given ODE (which we consider to be linear) inside the interval $x_1 \leq x \leq x_2$.

3.1.1 Linear Boundary Conditions

We will consider only *linear* boundary conditions, where the left-hand sides of the conditions are linear combinations of the function and its derivatives at the same point and the right-hand sides are given constants, for example

$$y(x_1) = b_1\ ,\ y(x_2) = b_2 \quad \text{or} \quad y'(x_1) = b_1\ ,\ y'(x_2) = b_2\ ,$$

or most generally

$$\alpha y'(x_1) + \beta y(x_1) = b_1\ ,\ \gamma y'(x_2) + \delta y(x_2) = b_2\ , \tag{3.2}$$

where $\alpha, \beta, \gamma, \delta$ are given real constants such that $|\alpha| + |\beta| > 0$, $|\gamma| + |\delta| > 0$.

3.1.2 Homogeneous Boundary Value Problem

If the constants b_1, b_2 on the right-hand side are equal to zero, the corresponding boundary condition is called **homogeneous**, otherwise it is **inhomogeneous**. If all boundary conditions are homogeneous and the ODE itself is also homogeneous, the corresponding boundary value problem is called homogeneous as well.

Example:
Consider the B.V.P.

$$y'' + y = f(x)\ ,\ y(0) = 0\ ,\ y'(\pi) = 0\ .$$

Write down the general solution of the above ODE for the special choice $f(x) = e^x$ and use it to solve the corresponding B.V.P.

Solution:
The characteristic equation $\lambda^2 + 1 = 0$ has two complex conjugate roots $\lambda_1 = -i$, $\lambda_2 = i$ so that the general solution of the homogeneous equation can be written as

$$y_h(x) = c_1 \cos x + c_2 \sin x\ .$$

A particular solution for the special choice $f(x) = e^x$ can be found by the *variation of parameter method* where

$$\begin{aligned}
y_p(x) &= \frac{1}{(\lambda_1 - \lambda_2)a_2}\left\{e^{\lambda_1 x}\int f(x)e^{-\lambda_1 x}dx - e^{\lambda_2 x}\int f(x)e^{-\lambda_2 x}dx\right\} \\
&= \frac{1}{(i+i)*1}\left\{e^{ix}\int e^x e^{-ix}dx - e^{-ix}\int e^x e^{ix}dx\right\} \\
&= \frac{1}{(i+i)*1}\left\{e^{ix}\int e^{(1-i)x}dx - e^{-ix}\int e^{(1+i)x}dx\right\} \\
&= \frac{1}{(i+i)*1}\left\{e^{ix}e^{(1-i)x}\frac{1}{1-i} - e^{-ix}e^{(1+i)x}\frac{1}{1+i}\right\} \\
&= \frac{e^x}{2i}\left(\frac{1}{1-i} - \frac{1}{1+i}\right) \\
&= \frac{e^x}{2i}\frac{1+i-(1-i)}{1-i^2} = \frac{e^x}{2}
\end{aligned}$$

Hence, the general solution to the inhomogeneous equation is given by

$$y_g(x) = c_1 \cos x + c_2 \sin x + \frac{1}{2}e^x\ .$$

Differentiating yields

$$y_g'(x) = -c_1 \sin x + c_2 \cos x + \frac{1}{2}e^x\ .$$

Combining these two equations with the boundary conditions leads to

$$y(0) = c_1 + \frac{1}{2} = 0\ ,\ y'(\pi) = -c_2 + \frac{1}{2}e^\pi = 0\ ,$$

which gives $c_1 = -1/2$ and $c_2 = \frac{1}{2}e^\pi$. The solution to the B.V.P. is thus given by

$$y(x) = \frac{1}{2}\left(-\cos x + e^\pi \sin x + e^x\right)\ .$$

3.2 Existence and uniqueness of solutions to B.V.P.

The main difference between a Boundary Value Problem (B.V.P.) and an Initial Value Problem is that the B.V.P. may have *(i) no solution, (ii) a unique solution or (iii) infinitely many solutions.*

Example:

1. B.V.P. $y'' + y = 0\,,\, y(0) = 0\,,\, y(\pi) = 1$ **does not have any** solution, since all solutions satisfying $y(0) = 0$ necessarily have the form $y(x) = c \sin x$ for some c, and they all vanish at $x = \pi$.
2. B.V.P. $y'' + y = 0\,,\ y(0) = 0\,,\ y(\pi) = 0$ has **infinitely many** solutions $y(x) = c \sin x$ for any choice of c.
3. B.V.P. $y'' + y = 0\,,\, y(0) = 1\,,\, y(\pi/2) = 1$ has the **unique** solution $y(x) = \cos x + \sin x$.

This general situation is explained by the following

3.2.1 Theorem of the Alternative

Theorem: Consider the Boundary Value Problem for the second order ODE (3.1), where all functions $a_0(x), a_1(x), a_2(x)$ and $f(x)$ are continuous, $a_2(x) \neq 0$, and all boundary conditions are linear and given by (3.2). Only *two alternative situations* are possible:

1. Either the B.V.P. has a **unique solution** for *any* $f(x)$ and *any* values b_1 and b_2 of the right-hand sides in the boundary conditions (3.2), or
2. the corresponding *homogeneous* problem has **infinitely many solutions**, and the *inhomogenous* problem has **infinitely many solutions** for *some* choices of $f(x)$ and right-hand sides in the boundary conditions, and for *other* choices does **not have solutions at all**.

3.2.1.1 Applications of this theorem

(a) the **homogeneous B.V.P.** has only the trivial **zero** solution
$\Rightarrow$ the **inhomogeneous B.V.P.** has only **one** solution for *any* right-hand side

(b) the **homogeneous B.V.P.** has at least one **non-zero** solution
$\Rightarrow$ the **inhomogeneous B.V.P.** has either **infinitely many** solutions or **none**

Example:

Find the smallest positive value of the parameter $b > 0$ such that the B.V.P.

$$y'' + b^2 y = 0, \quad y(0) = 5\,,\, y(1) = -5$$

does not have any solution.

Solution:

According to the *Theorem of the Alternative* the above inhomogeneous problem may not have a solution only if the corresponding *homogeneous problem*

$$y'' + b^2 y = 0\ ,\ y(0) = 0\ ,\ y(1) = 0$$

does have a *non-zero* solution. We know that solutions of $y'' + b^2 y = 0$ must have the form $y(x) = A\sin(bx) + B\cos(bx)$. The first boundary condition $y(0) = 0$ selects $B = 0$, so we must have $y(x) = A\sin(bx)$. The second boundary condition yields $y(1) = A\sin(b) = 0$ and together with $b > 0$ selects the values $b = \pi, 2\pi, 3\pi, \ldots$. For these values the homogeneous problem has a nonzero solution (for example, for $b = \pi$ the solution is $y(x) = A\sin(\pi x)\,\forall A \neq 0$), so we have the second alternative: Either the original inhomogeneous problem has infinitely many solutions, or none at all. Which of these two cases occurs has to be checked case by case by inspecting the corresponding inhomogeneous problem.

Consider first $b = \pi$ so that the general solution of the inhomogeneous problem is $y(x) = A\sin(\pi x) + B\cos(\pi x)$. The condition $y(0) = 5$ yields $B = 5$, and now $y(1) = A\sin\pi + 5\cos\pi = -5$ for *any* choice of A. Thus for $b = \pi$ the B.V.P. has *infinitely many* solutions of the form $y(x) = A\sin(\pi x) + 5\cos(\pi x)$.

Now consider $b = 2\pi$ for which the solution of the inhomogeneous problem must be of the form $y(x) = A\sin(2\pi x) + B\cos(2\pi x)$. Then $y(0) = B = 5$, but in this case we necessarily have $y(1) = A\sin(2\pi) + 5\cos(2\pi) = 5$ in contradiction to the second boundary condition. This implies that the boundary value problem *does not have any solution*. We conclude that $b = 2\pi$ is the required minimal positive value of the parameter b.

3.2.1.2 Proof of this theorem

* This proof is not covered by the lectures and is not examinable (from now until the end of the proof). It is left for students who are interested to work themselves through more mathematical details.

We will provide a proof only in the simplest case of ODEs with *constant coefficients* on the left-hand side,

$$a_2(x) = a_2\ ,\ a_1(x) = a_1\ ,\ a_0(x) = a_0\ \forall x \in [x_1, x_2]\ .$$

We furthermore suppose for simplicity that the associated characteristic equation $a_2\lambda^2 + a_1\lambda + a_0 = 0$ has only distinct real roots $\lambda_1 \neq \lambda_2$. In this case we know that the general solution $y_g(x)$ of the inhomogeneous equation (given by the sum of the general solution of the homogeneous equation $y_h(x)$ and any particular solution of the inhomogeneous equation $y_p(x)$) can be written as

$$y_g(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x} + y_p(x) \tag{3.3}$$

with constants c_1, c_2. Now we should fix these constants by satisfying the linear boundary conditions (3.2). Substituting the solution (3.3) into these boundary conditions and shifting the y_p-dependent terms onto the right-hand side we get a system of two linear algebraic equations for the coefficients c_1, c_2.

For example, the boundary condition $\alpha y'(x_1) + \beta y(x_1) = b_1$ yields

$$\alpha \left(c_1 \lambda_1 e^{\lambda_1 x_1} + c_2 \lambda_2 e^{\lambda_2 x_1} + y_p'(x_1)\right) + \beta \left(c_1 e^{\lambda_1 x_1} + c_2 e^{\lambda_2 x_1} + y_p(x_1)\right) = b_1$$

or equivalently, after rearranging,

$$\left(\beta e^{\lambda_1 x_1} + \alpha \lambda_1 e^{\lambda_1 x_1}\right) c_1 + \left(\beta e^{\lambda_2 x_1} + \alpha \lambda_2 e^{\lambda_2 x_1}\right) c_2 = b_1 - \beta y_p(x_1) - \alpha y_p'(x_1)\,. \tag{3.4}$$

Similarly, the second boundary condition gives at $x = x_2$

$$\left(\delta e^{\lambda_1 x_2} + \gamma \lambda_1 e^{\lambda_1 x_2}\right) c_1 + \left(\delta e^{\lambda_2 x_2} + \gamma \lambda_2 e^{\lambda_2 x_2}\right) c_2 = b_2 - \delta y_p(x_2) - \gamma y_p'(x_2)\,. \tag{3.5}$$

Note that the coefficients of this system on the left-hand side depend on the *left-hand sides* of the boundary conditions but *not* on the right-hand sides b_1, b_2 and *not* on the function $f(x)$.

From the course in Linear Algebra we know that the solution of the system for c_1, c_2 dependends on the value of the *determinant* D of the coefficient matrix associated to this system of linear algebraic equations

$$D = \det \begin{pmatrix} \beta e^{\lambda_1 x_1} + \alpha \lambda_1 e^{\lambda_1 x_1} & \beta e^{\lambda_2 x_1} + \alpha \lambda_2 e^{\lambda_2 x_1} \\ \delta e^{\lambda_1 x_2} + \gamma \lambda_1 e^{\lambda_1 x_2} & \delta e^{\lambda_2 x_2} + \gamma \lambda_2 e^{\lambda_2 x_2} \end{pmatrix}.$$

Namely,

- If $D \neq 0$ the system has a *unique* solution for the coefficients c_1, c_2 for *any* choice of the right-hand sides in equations (3.4),(3.5). Substituting these coefficients into (3.3) we get the **unique** solution of the original B.V.P. Note that for the corresponding *homogeneous* problem (that is, with $b_1 = b_2 = 0$ and $f(x) \equiv 0$) also $y_p(x) = 0$ and the right-hand sides in the corresponding algebraic equations (3.4), (3.5) will be zero. Then the *homogeneous* problem will only have the trivial **zero solution** $y_h(x) = 0$ for $x \in [x_1, x_2]$.
- If $D = 0$ the *homogeneous* linear algebraic system (3.4),(3.5) (i.e., with all right-hand sides zero) will have *infinitely many* non-zero solutions. At the same time the *inhomogeneous* linear systems with the same left-hand side will have either *infinitely many* solutions or *none at all*. If it has infinitely many different solutions for coefficients c_1, c_2, each solution will generate the corresponding different solution to the original B.V.C., which then will have infinitely many solutions as well.

Note:

If the characteristic equation has two complex-conjugate roots $\lambda_{1,2} = a \pm ib$, any particular solution $y_p(x)$ of the inhomogeneous equation can be written as

$$y_g(x) = e^{ax}(c_1 \cos bx + c_2 \sin bx) + y_p(x)\,, \tag{3.6}$$

and the proof can be performed along similar lines.

MTH5123 Differential Equations

Lecture Notes

Week 8

School of Mathematical Sciences
Queen Mary University of London

4 Autonomous systems of two first order ODEs

4.1 General properties of autonomous systems

Definition:
A **system of ODEs** in *normal form*

$$\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}) \ , \ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \ , \ \mathbf{f} = \begin{pmatrix} f_1(t, y_1, y_2) \\ f_2(t, y_1, y_2) \end{pmatrix} \tag{4.1}$$

is called **autonomous** if all functions on the right-hand side of the equation do not depend *explicitly* on the variable t, i.e.,

$$f_1(t, y_1, y_2) = f_1(y_1, y_2) \ , \ f_2(t, y_1, y_2) = f_2(y_1, y_2) \ . \tag{4.2}$$

A part $\mathcal{G}$ of the two-dimensional space $\mathbb{R}^2$ described by the coordinates y_1, y_2, where both functions $f_1(y_1, y_2)$ and $f_2(y_1, y_2)$ are well-defined, is called the **phase space** of this system. We will consider only cases where the phase space is the whole (y_1, y_2) plane $\mathbb{R}^2$.

Furthermore, we will only consider systems where the right-hand sides f_1, f_2 are continuous and where all partial derivatives $\partial f_i / \partial y_j$ are also continuous everywhere in the phase space. The *Picard-Lindelöf Theorem* will then ensure the *uniqueness* of solutions for any initial conditions, that is, globally in the whole phase space.

Dynamical systems.

We will think of t as time, and of the system's dynamics as an evolution in time. However, it is frequently convenient to consider on equal footing not only an evolution from the initial conditions towards the "future" (that is, for $0 \le t < \infty$) but also from the initial conditions towards the "past" (that is, for $-\infty < t \le 0$). Systems like (4.1) are called **dynamical systems**.

Trajectories and equilibra.

Every solution of an autonomous system $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y})$ given by

$$y_1 = y_1(t) \ , \ y_2 = y_2(t) \tag{4.3}$$

describes a **curve** in the phase space, parametrized by the parameter $-\infty < t < \infty$. These curves are called **trajectories** of the dynamical system. In the particular case where the system of ODEs allows a constant solution such that for any time t we have $y_1(t) = a_1 = const$, $y_2(t) = a_2 = const$, the curve degenerates to a single point in the phase space with the coordinate vector $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$. Such a point represented by a constant vector $\mathbf{y}(t) = \mathbf{a}$ can be a solution of (4.1), (4.2) only if simultaneously the right-hand sides of

$$f_1(\mathbf{a}) = 0\ ,\ f_2(\mathbf{a}) = 0$$

vanish for the same vector $\mathbf{a}$. Such special points are called **equilibria**, or synonymously **stationary points**, **singular points** and **fixed points**.

1. *Any two trajectories either completely coincide, or do not have any common points.* This is a consequence of the *uniqueness of solutions* of the initial value problems for (4.1), (4.2) ensured by the Picard-Lindelöf Theorem: If two *different* trajectories had a common point, then using this point as an initial value we would have *two different* solutions to the initial value problem, which is impossible. This **non-intersection property** in turn implies that:

2. *Any solution of (4.1), (4.2) cannot reach an equilibrium point in finite time.* Because if $\mathbf{a}$ is an equilibrium point, the constant solution $\tilde{\mathbf{y}}(t) = \mathbf{a}$ is a (degenerate) trajectory for all times t. But if $\mathbf{y}(t)$ is a solution which does not coincide with $\mathbf{a}$, according to 1. above the trajectory representing $\mathbf{y}(t)$ cannot have a common point with the one representing $\tilde{\mathbf{y}}(t)$, that is, $\forall t\, \mathbf{y}(t) \neq \mathbf{a}$. To be more precise, solutions $\mathbf{y}(t)$ can approach equilibrium points only for $t \to \pm\infty$.

Next our focus will be to understand the typical behaviour of solutions to a general pair of first-order autonomous ODEs (4.1), (4.2). As in this case the **phase space** is the two-dimensional plane, (y_1, y_2), and two autonomous ODEs then takes the form

$$\dot{y}_1 = f_1(y_1, y_2)\ ,\ \dot{y} = f_2(y_1, y_2)\ . \tag{4.4}$$

Our analysis will proceed by establishing typical features of the solutions of (4.4) in the phase space (y_1, y_2). A special role is played by the *equilibria* which, as we know already, in our case are given by the solutions of the pair of equations

$$f_1(y_1, y_2) = 0\ ,\ f_2(y_1, y_2) = 0\ . \tag{4.5}$$

In general, these equations may have several solutions. Our goal will be to investigate typical trajectories of (4.4) in the vicinity of a given solution $y_1 = y_{1_c}$, $y_2 = y_{2_c}$ of (4.5). Here we will assume that such a solution is *isolated*, that is, there exists $R > 0$ such that inside the circle $(y_1 - y_{1_c})^2 + (y_2 - y_{2_c})^2 \leq R^2$ there are no other solutions of (4.5).

4.2 Linearization of autonomous systems of two first order ODEs

4.2.1 Linearize a nonlinear ODE system around its equilibrium

When investigating trajectories in the close proximity of an isolated fixed point we can always assume that $y_{1_c} = y_{2_c} = 0$, which is equivalent to placing the origin of the coordinate system in the (y_1, y_2) plane on the chosen equilibrium. This can always be achieved by transforming into new coordinates $(y_1, y_2) \longrightarrow (\tilde{y_1}, \tilde{y_2})$ defined by $\tilde{y_1} \equiv y_1 - y_{1c}$, $\tilde{y_2} \equiv y_2 - y_{2c}$. Hence, by assuming for sake of simplicity that the coordinates (y_1, y_2) are such that there is a fixed point at $(0, 0)$, we will further assume that the functions $f_1(y_1, y_2)$ and $f_2(y_1, y_2)$ can be expanded in a Taylor series around the origin. Taking into account $f_1(0, 0) = f_2(0, 0) = 0$ and denoting

$$a_{11} = \frac{\partial f_1(y_1, y_2)}{\partial x}|_{0,0} \ , \ a_{12} = \frac{\partial f_1(y_1, y_2)}{\partial y}|_{0,0} \ , \tag{4.6}$$

$$a_{21} = \frac{\partial f_2(y_1, y_2)}{\partial x}|_{0,0} \ , \ a_{22} = \frac{\partial f_2(y_1, y_2)}{\partial y}|_{0,0} \tag{4.7}$$

we arrive at the system of two ODEs

$$\dot{y_1} = a_{11}y_1 + a_{12}y_2 + O(y_1^2, y_2y_1, y_2^2) \ , \ \dot{y} = a_{21}y_1 + a_{22}y_2 + O(y_1^2, y_2y_1, y_2^2) \ , \tag{4.8}$$

where $O(...)$ stands for all terms of order (...) or higher. We see that by neglecting these higher-order terms the **local** behaviour of the trajectories close to the chosen isolated fixed point is governed by the system of two **linear** ODEs

$$\dot{y_1} = a_{11}y_1 + a_{12}y_2 \ , \ \dot{y} = a_{21}y_1 + a_{22}y_2 \ . \tag{4.9}$$

They can be rewritten in **matrix form** as

$$\begin{pmatrix} \dot{y_1} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{4.10}$$

and even more concisely as

$$\dot{\mathbf{y}} = A\,\mathbf{y} \ , \ \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \ . \tag{4.11}$$

The above procedure is called the **linearization** of the system of ODEs around a given fixed point.

Further progress with the analysis of such systems heavily relies on understanding the properties of 2×2 matrices. We thus proceed with a very brief review in week 9 on this subject tailored to our goals.

MTH5123 Differential Equations

Lecture Notes

Week 9

School of Mathematical Sciences
Queen Mary University of London

4.2.2 General properties of 2×2 matrices with real entries

First we recall the notion of the **trace** and the **determinant**, which for 2×2 matrices are given by

$$\mathrm{Tr}A = a_{11} + a_{22} \quad \text{and} \quad \det A = a_{11}a_{22} - a_{12}a_{21} \,,$$

respectively. The importance of the determinant is reflected in the fact that the condition $\det A \neq 0$ ensures the possibility to find the *inverse* matrix A^{-1}, which satisfies $A^{-1}A = AA^{-1} = I_d$. The so-called **identity matrix** $I_d = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ plays the same role in matrix algebra as the unity number 1 for usual numbers. Explicitly, the inverse of any 2×2 matrix with $\det A \neq 0$ is given by

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \,. \tag{4.7}$$

The operation of matrix inversion is of fundamental importance, as it allows one to find the unique solution $\mathbf{y}$ of a (non-singular) system of linear equations $A\mathbf{y} = \mathbf{b}$ as $\mathbf{y} = A^{-1}\mathbf{b}$.

A vector $\mathbf{u} = \begin{pmatrix} p \\ q \end{pmatrix} \neq \mathbf{0}$ is called an **eigenvector** of the matrix A if the equality $A\mathbf{u} = \lambda\mathbf{u}$ holds for some value of the parameter λ known as the corresponding **eigenvalue**.

Theorem:

1. There are two eigenvalues of any 2×2 matrix A, which are the roots of the quadratic equation
$$\det(A - \lambda I_d) = \lambda^2 - \mathrm{Tr}A \cdot \lambda + \det A = 0 \,. \tag{4.8}$$
These eigenvalues are either both real or complex conjugate to each other, $\overline{\lambda_1} = \lambda_2$.
2. If $\lambda_1 \neq \lambda_2$ then the two eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ are **linearly independent**.[1]
3. If $a_{12} = a_{21}$ then either $\lambda_1 = \lambda_2$ or the two eigenvectors are orthogonal, $\mathbf{u}_1 \cdot \mathbf{u}_2 \equiv p_1p_2 + q_1q_2 = 0$.

Proof:

This proof is not covered by the lectures and is not examinable. It is left for students who are interested to work themselves through more mathematical details.

To verify the first statement we use $\mathbf{u} = I_d \cdot \mathbf{u} \forall \mathbf{u}$, hence the condition of being an eigenvalue can be rewritten equivalently as

$$A\mathbf{u} = \lambda\, I_d \cdot \mathbf{u} \quad \Leftrightarrow (A - \lambda I_d) \cdot \mathbf{u} = \mathbf{0}$$

for some $\mathbf{u} \neq \mathbf{0}$. Now assuming $\det(A - \lambda I_d) \neq 0$ we immediately see that

$$\mathbf{u} = (A - \lambda I_d)^{-1}\mathbf{0} = \mathbf{0} \,,$$

which is a contradiction, hence necessarily $\det(A - \lambda I_d) = 0$. Furthermore, writing

$$A - \lambda I_d \equiv \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix}$$

[1] Two vectors $\mathbf{u}_1, \mathbf{u}_2$ are *linearly independent* if their linear combination $c_1\mathbf{u}_1 + c_2\mathbf{u}_2$ can be zero only if both c_1 and c_2 are simultaneously zero. Alternatively, two vectors are linearly dependent if there exists a constant $k \neq 0$ such that $\mathbf{u}_2 = k\mathbf{u}_1$.

we see that indeed

$$\det(A-\lambda I_d) = (a_{11}-\lambda)(a_{22}-\lambda) - a_{12}a_{21}$$
$$= \lambda^2 - \lambda(a_{11}+a_{22}) + (a_{11}a_{22}-a_{12}a_{21}) \equiv \lambda^2 - \mathrm{Tr}A\cdot\lambda + \det A$$

as required.
To verify the second statement it is enough to demonstrate that if there exists a $k \neq 0$ such that $\mathbf{u}_2 = k\mathbf{u}_1$ then necessarily $\lambda_1 = \lambda_2$. For this we consider the eigenequation $A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$ and multiply it with $\mathbf{u}_2$ from the left getting: $\mathbf{u_2}A\mathbf{u}_1 = \lambda_1\mathbf{u}_2\cdot\mathbf{u}_1$. In the same way we take the second eigenequation $A\mathbf{u}_2 = \lambda_2\mathbf{u}_2$ and multiply it with $\mathbf{u}_1$ from the left yielding $\mathbf{u_1}A\mathbf{u}_2 = \lambda_2\mathbf{u}_1\cdot\mathbf{u}_2$. Subtracting both equations and using the symmetry of the scalar product $\mathbf{u}_2\cdot\mathbf{u}_1 = \mathbf{u}_1\cdot\mathbf{u}_2$ yields the relation

$$\mathbf{u_2}A\mathbf{u}_1 - \mathbf{u_1}A\mathbf{u}_2 = (\lambda_1-\lambda_2)\mathbf{u}_1\cdot\mathbf{u}_2\,. \tag{4.9}$$

Now, substituting here $\mathbf{u}_2 = k\mathbf{u}_1$ yields on the left-hand side zero, whereas the right-hand side becomes equal to $k(\lambda_1-\lambda_2)\mathbf{u}_1\cdot\mathbf{u}_1$. As $\mathbf{u}_1 \neq 0$ and $k \neq 0$ we conclude that necessarily $\lambda_1 = \lambda_2$ as required. The final part of the theorem follows from the fact that for $a_{12} = a_{21}$ the left-hand side of (4.9) is identically zero for any choice of the vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ (please check!). To make the right-hand side vanishing we therefore have to require either $\lambda_1 = \lambda_2$ or $\mathbf{u}_1\cdot\mathbf{u}_2 = 0$ implying the orthogonality of the eigenvectors.

Examples:
(i) $A = \begin{pmatrix} -4 & 6 \\ -3 & 5 \end{pmatrix}$ with $\lambda_1 = 2\,,\ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\lambda_2 = -1\,,\ \mathbf{u}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.
(ii) $A = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}$ with $\lambda_1 = 1+i\,,\ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1+i \end{pmatrix}$ and $\lambda_2 = 1-i\,,\ \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1-i \end{pmatrix}$.

Note: Eigenvectors are determined up to a nonzero factor.

The last useful property needed to be mentioned is as follows: Suppose a matrix A has two distinct eigenvalues $\lambda_1 \neq \lambda_2$ and an associated pair of two *linearly independent* eigenvectors $\mathbf{u}_1 = \begin{pmatrix} p_1 \\ q_1 \end{pmatrix} \neq \mathbf{0}$ and $\mathbf{u}_2 = \begin{pmatrix} p_2 \\ q_2 \end{pmatrix} \neq \mathbf{0}$. Then the matrix $U = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}$ is nonsingular, that is $\det U \neq 0$, and therefore can be inverted giving U^{-1}. Moreover, the matrix $U^{-1}AU$ turns out to be always *diagonal* and equal to $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$. The last fact follows from

$$AU = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}\begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 p_1 & \lambda_2 p_2 \\ \lambda_1 q_1 & \lambda_2 q_2 \end{pmatrix} = U\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix},$$

where we have used that the condition $A\mathbf{u} = \lambda\mathbf{u}$ for $\mathbf{u} = \begin{pmatrix} p \\ q \end{pmatrix}$ is equivalent to $a_{11}p_i + a_{12}q_i = \lambda p_i$ and $a_{21}p_i + a_{22}q_i = \lambda q_i\,,\ i = 1,2$. Now we will use these properties for solving the system of linear ODE's.

4.2.3 General analysis of the linearised ODE system for $D = (\mathbf{Tr}A)^2 - 4\det A \neq 0$

Solving the characteristic equation $\det(A-\lambda I_d) = \lambda^2 - \mathrm{Tr}A\cdot\lambda + \det A = 0$. for λ yields $\lambda = \frac{\mathrm{Tr}A}{2} \pm \frac{1}{2}\sqrt{(\mathrm{Tr}A)^2 - 4\det A}$. This shows that in this case we must have two distinct roots. In the case $D = (\mathrm{Tr}A)^2 - 4\det A > 0$ these roots are real,

$$\lambda_1 = \frac{1}{2}\left(\mathrm{Tr}A + \sqrt{D}\right) > \lambda_2 = \frac{1}{2}\left(\mathrm{Tr}A - \sqrt{D}\right) ,$$

hence the corresponding eigenvectors $\mathbf{u}_1$ and $\mathbf{u}_2$ must also be real and linearly independent. In the opposite case $D = (\mathrm{Tr}A)^2 - 4\det A < 0$ the two roots are complex conjugate,

$$\lambda_1 = \frac{1}{2}\left(\mathrm{Tr}A + i\sqrt{|D|}\right) , \ \lambda_2 = \frac{1}{2}\left(\mathrm{Tr}A - i\sqrt{|D|}\right) ,$$

and the eigenvectors may be complex but are still linearly independent, as $\lambda_1 \neq \lambda_2$.
Given a pair of linearly independent vectors we know that we can write an arbitrary vector of the same dimension as a linear combination of these two vectors.
We can thus look for a solution $\mathbf{y} \equiv \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ of the ODE system $\begin{pmatrix} \dot{y_1} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $\quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ in the form of $\mathbf{y} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2$, where the coefficients c_1 and c_2 are assumed to depend on t. We then have $\dot{\mathbf{y}} = \dot{c}_1\mathbf{u}_1 + \dot{c}_2\mathbf{u}_2$, which by substituting into the above ODE system gives the chain of identities

$$\dot{c}_1\mathbf{u}_1 + \dot{c}_2\mathbf{u}_2 = A(c_1\mathbf{u}_1 + c_2\mathbf{u}_2) = c_1\lambda_1\mathbf{u}_1 + c_2\lambda_2\mathbf{u}_2 \tag{4.10}$$

or, rearranging,

$$(\dot{c}_1 - \lambda_1 c_1)\mathbf{u}_1 = -(\dot{c}_2 - c_2\lambda_2)\mathbf{u}_2 , \tag{4.11}$$

which must hold at any moment of time t. Using linear independence of the two eigenvectors we conclude that simultaneously we must have

$$\dot{c}_1 = \lambda_1 c_1 \quad \text{and} \quad \dot{c}_2 = \lambda_2 c_2 .$$

These separable equations are immediately solved to produce

$$c_1(t) = c_1 e^{\lambda_1 t} , \ c_2(t) = c_2 e^{\lambda_2 t} , \tag{4.12}$$

where $c_{1,2}$ are constants. We conclude that the **general solution** to this ODE system in such a case is given by

$$\mathbf{y}(t) = c_1 e^{\lambda_1 t}\mathbf{u}_1 + c_2 e^{\lambda_2 t}\mathbf{u}_2 . \tag{4.13}$$

For any initial value problem with the ODE system, $\begin{pmatrix} \dot{y_1} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $\quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, the values of c_1 and c_2 will be determined by the initial conditions $y_1(0) = a$ and $y_2(0) = b$ and the eigenvectors $\mathbf{u}_1 = \begin{pmatrix} p_1 \\ q_1 \end{pmatrix}$ and $\mathbf{u}_2 = \begin{pmatrix} p_2 \\ q_2 \end{pmatrix}$ as, according to (4.13), $a = c_1p_1 + c_2p_2$, $b = c_1q_1 + c_2q_2$.

Example:

Consider the ODE system, $\begin{pmatrix} \dot{y_1} \\ \dot{y} \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, $\quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, with the particular choice of A

$$\begin{pmatrix} \dot{y_1} \\ \dot{y_2} \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} , \ A = \begin{pmatrix} -4 & 6 \\ -3 & 5 \end{pmatrix} . \tag{4.14}$$

In exercise (i) on p.5 we have found that $\lambda_1 = 2$, $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\lambda_2 = -1$, $\mathbf{u}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$. According to (4.13), the solution is given by

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = c_1 e^{2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} .$$

The initial conditions $x(0) = a, y(0) = b$ can now be written as

$$a = c_1 + 2c_2 \ , \ b = c_1 + c_2 \ .$$

Solving these equations we get $c_2 = a - b$, $c_1 = 2b - a$ so that the explicit expressions for the time dependence of the coordinates in the (y_1, y_2) plane are given by

$$y_1 = (2b - a)e^{2t} + 2(a - b)e^{-t}, \quad y_2 = (2b - a)e^{2t} + (a - b)e^{-t} \ .$$

4.3 Phase portraits for linearised systems

The eigenvalues of the linearised system are $\lambda = \frac{\mathrm{Tr}A}{2} \pm \frac{1}{2}\sqrt{D}$, where $D = (\mathrm{Tr}A)^2 - 4 \det A$. We can classify the phase portraits of our linear systems based the different situations of the eigenvalues.

4.3.1 Transformation of phase portraits between coordinates through invariant manifolds

In our example in last section, we have two distinct real eigenvalues, where $\lambda_1 = 2$ and $\lambda_2 = -1$. How do the trajectories look like for different initial conditions, which means when a, b have different values?
Consider when $b = a/2$, which implies $y_1 = ae^{-t}$, $y_2 = \frac{a}{2}e^{-t}$. We conclude that $y_2 = y_1/2 \ \forall t$, i.e., this trajectory corresponds to motion along a straight line with slope 1/2 towards the origin (since both y_1 and y_2 tend to zero for $t \to \infty$). Similarly, for $b = a$ we have $y_2 = y_1 = a\,e^{2t} \ \forall t$, which describes motion along a straight line away from the origin. These two special straight lines are known as **invariant manifolds**, which correspond to the two lines on top of the two eigenvectors. They intersect at the origin and partition the phase space, which is (y_1, y_2) plane, into four sectors, see Fig. 4.1 (left). Asymptotically the trajectories which start from initial conditions such that $b > a/2$ tend to approach to the line $y_2 = y_1 \to +\infty$ for $t \to \infty$, whereas the trajectories with initial conditions such that $b < a/2$ tend to approach to the line $y_2 = y_1 \to -\infty$ for $t \to \infty$.
In the above example we have fully understood the structure of the trajectories in the (y_1, y_2) phase space, which is called a **phase portrait**. This picture is looking rather complicated, but it simplifies if we transform into specific coordinates. For this purpose we introduce the vector $\tilde{\mathbf{y}} \equiv \begin{pmatrix} \tilde{y}_1 \\ \tilde{y}_2 \end{pmatrix}$ of new coordinates $\tilde{y}_1, \tilde{y}_2$ related to the vector of old coordinates $\mathbf{y} = \begin{pmatrix} x \\ y \end{pmatrix}$ via $\tilde{\mathbf{y}} = U^{-1}\mathbf{y}$, or equivalently $\mathbf{y} = U\tilde{\mathbf{y}}$. The columns of the 2×2 matrix U are chosen to be the eigenvectors $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{u}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, which implies

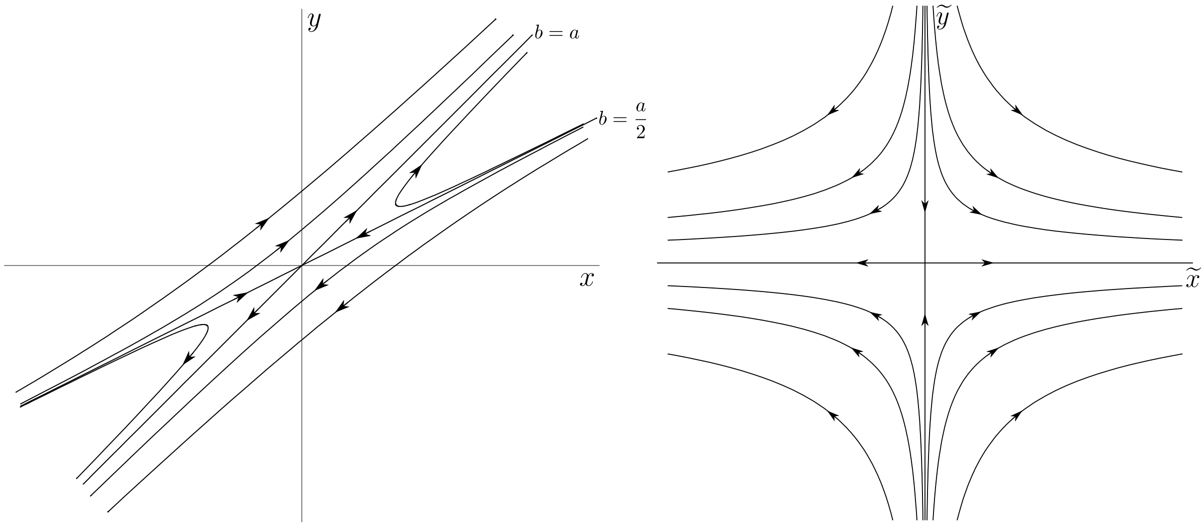

Figure 4.1: Phase portraits of the ODE (4.14) in old (left) and in new (right) coordinates.

$U^{-1} = -\begin{pmatrix} 1 & -2 \\ -1 & 1 \end{pmatrix}$. This yields the chain of identities

$$\frac{d}{dt}\tilde{\mathbf{y}} = U^{-1}\frac{d}{dt}\mathbf{y} = U^{-1}A\mathbf{y} = U^{-1}AU\tilde{\mathbf{y}} \tag{4.15}$$

and accordingly the system of ODEs in new coordinates

$$\frac{d}{dt}\tilde{\mathbf{y}} = \tilde{A}\tilde{\mathbf{y}}\ ,\ \tilde{A} = U^{-1}AU = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}\ . \tag{4.16}$$

Hence, this system is equivalent to $\dot{\tilde{y_1}} = 2\tilde{y_1}$, $\dot{\tilde{y_2}} = -\tilde{y_2}$. Specifying initial conditions in new coordinates as $\tilde{y_1}(0) = \tilde{a}, \tilde{y_2}(0) = \tilde{b}$ we solve these equations to $\tilde{y_1}(t) = \tilde{a}e^{2t}, \tilde{y_2}(t) = \tilde{b}e^{-t}$. Eliminating the time variable t we find the trajectories to be given by $\tilde{y_2} = \tilde{b}\left(\frac{\tilde{y_1}}{\tilde{a}}\right)^{-1/2}$ if $\tilde{a} \neq 0$ and $\tilde{y_1} = 0$ if $\tilde{a} = 0$. Moreover, for $t \to \infty$ we have $\tilde{y_2} \to 0$ whereas $\tilde{y_1} \to \infty$ for $\tilde{a} > 0$ and $\tilde{y_1} \to -\infty$ for $\tilde{a} < 0$. This allows us to sketch the phase portrait given in Fig. 4.1 (right).

Comparing these two phase portraits in old and new coordinates we see that *qualitatively* they look similar. However, the one in old coordinates (y_1, y_2) is a kind of rotated, twisted and stretched version of the one in new coordinates $\tilde{y_1}, \tilde{y_2}$. In particular, in new coordinates the important straight lines which partition the plane into four sectors coincide with the two coordinate axes simplifying the picture. These two pictures are called **topologically equivalent** meaning that one can be transformed into the other by a continuous set of transformations without the need to cut or tear the plane apart (imagine the plane to be made of plasticine, which can be easily distorted). Phase portraits for other types of systems may not be deformed that way and are then of a different *topological type*, as we shall see in the next example.

MTH5123 Differential Equations

Lecture Notes

Week 10

School of Mathematical Sciences
Queen Mary University of London

Autumn 2021

4.3.2 Phase portrait shape when $D = (\mathbf{Tr}A)^2 - 4\det A > 0\,\&\det A < 0$: saddle-type (unstable)

But first we generalize our analysis: The last example fulfills $D = (\mathrm{Tr}A)^2 - 4\det A > 0$ for which the matrix A has two distinct real eigenvalues. For systems with $D > 0$ specific coordinates can always be defined as in this example: Having two real linearly independent eigenvectors $\mathbf{u}_1 = \begin{pmatrix} p_1 \\ q_1 \end{pmatrix} \neq \mathbf{0}$ and $\mathbf{u}_2 = \begin{pmatrix} p_2 \\ q_2 \end{pmatrix} \neq \mathbf{0}$ of the matrix A allows us to build a non-singular matrix $U = \begin{pmatrix} p_1 & p_2 \\ q_1 & q_2 \end{pmatrix}$, $\det U \neq 0$, which can be inverted giving U^{-1}. We then change the coordinates (y_1, y_2) into new coordinates $(\tilde{y_1}, \tilde{y_2})$ via the transformation $\tilde{\mathbf{y}} = U^{-1}\mathbf{y}$ such that the system of two ODEs in the new coordinates takes the form $\frac{d}{dt}\tilde{\mathbf{y}} = \tilde{A}\tilde{\mathbf{y}}$ with $\tilde{A} = U^{-1}AU$ generalizing (4.16). As $\tilde{A}$ was shown to be diagonal, $\tilde{A} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$, the two equations of the system defined by $\tilde{A}$ are now uncoupled and can be solved straightforwardly. Assuming the initial conditions $\tilde{y_1}(0) = \tilde{a}, \tilde{y_2}(0) = \tilde{b}$ we arrive at $\tilde{y_1}(t) = \tilde{a}e^{\lambda_1 t}$, $\tilde{y_2}(t) = \tilde{b}e^{\lambda_2 t}$. We see that the asymptotic behaviour and the phase portrait is determined by the *signs of the (real) eigenvalues* λ_1 and λ_2: If they are of different signs, $\lambda_2 < 0 < \lambda_1$, we have $\tilde{y_2} \to 0$ $(t \to \infty)$ whereas $\tilde{y_1} \to \infty$ for $\tilde{a} > 0$ and $\tilde{y_1} \to -\infty$ for $\tilde{a} < 0$ for $(t \to \infty)$. Nearby trajectories are given by *hyperbolic curves*

$$\tilde{y_2} = \tilde{b}\left(\frac{\tilde{y_1}}{\tilde{a}}\right)^{\lambda_2/\lambda_1} \tag{4.21}$$

if $\tilde{a} \neq 0$ and by the straight line $\tilde{y_1} = 0$ if $\tilde{a} = 0$. After this transformation the phase portrait qualitatively always looks like the one in Fig.4.1 (right), which is called a **saddle**. In the original coordinates the saddle will retain its topology, but generally it may be rotated and distorted; see Fig. 4.2 for two examples.

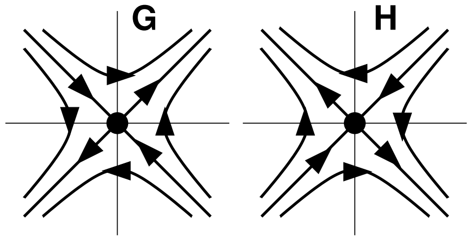

Figure 4.2: Two examples of phase portraits of saddle-type.

Example:
Find the general solution of the system of ODEs

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = A\begin{pmatrix} x \\ y \end{pmatrix}\ ,\ A = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{3}{2} \end{pmatrix} \tag{4.22}$$

and sketch the trajectories in phase space.

Solution:

We find (exercise for you) $\lambda_1 = 1$, $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\lambda_2 = 2$, $\mathbf{u}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. According to (4.19) the solution is given by

$$\begin{pmatrix} x \\ y \end{pmatrix} = c_1 e^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{2t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} .$$

The initial conditions $y_1(0) = a, y_2(0) = b$ yield $a = c_1 + c_2, b = c_1 - c_2$. Solving for c_1, c_2 we get $c_1 = (a+b)/2, c_2 = (a-b)/2$. The explicit expressions for the time dependence of the coordinates in the (y_1, y_2) plane are thus given by

$$x = \frac{a+b}{2}\, e^t + \frac{a-b}{2}\, e^{2t} \,,\; y = \frac{a+b}{2}\, e^t - \frac{a-b}{2}\, e^{2t} \,.$$

4.3.3 Phase portrait shapes when $D = (\mathbf{Tr}A)^2 - 4\det A > 0 \& \det A > 0$: unstable node ($\mathbf{Tr}A > 0$) and stable ($\mathbf{Tr}A < 0$) node

How do the trajectories look like for different initial conditions? We start with $b = a$, which implies $y_1 = ae^t$, $y_2 = ae^t$. We conclude that $y_2 = y_1\ \forall t$, i.e., this trajectory corresponds to motion along a straight line with slope one away from the origin, since both x and y tend to infinity when $t \to \infty$. Similarly, for $b = -a$ we have $y_2 = -y_1 = -a\, e^{2t}\ \forall t$, which describes motion along a straight line away from the origin. These two special straight lines, which intersect at the origin and partition the (y_1, y_2) plane into four sectors, define the invariant manifolds for the present system; see Fig. 4.3. Trajectories that start from initial conditions away from the invariant manifolds tend to be asymptotically parallel to one of the manifolds $y_2 = -y_1$ for $t \to \infty$, whereas close to the origin they are tangent to the other manifold $y_2 = y_1$. Such a phase portrait is topologically non-equivalent to the previous case.

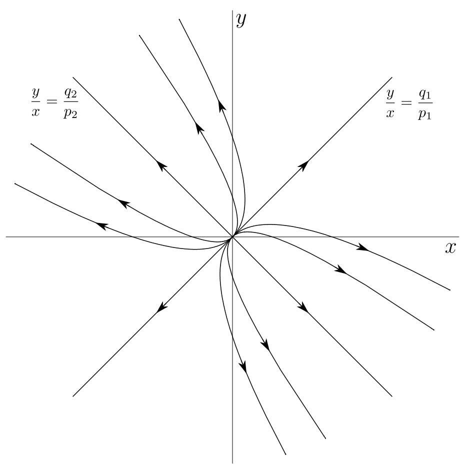

Figure 4.3: Phase portrait of the system (4.22) for which $q_1/p_1 = 1$, $q_2/p_2 = -1$.

As before, we now generalize our analysis: The last example still fulfills $D > 0$ like the previous one, but while in the first example the two eigenvalues had *different* signs, in the second one they have the *same* sign. In this case one can again perform a transformation into new coordinates $\tilde{y_1}, \tilde{y_2}$ exactly as before, and the trajectories are again given by (4.21) $\tilde{y_2} = \tilde{b}\left(\frac{\tilde{y_1}}{\tilde{a}}\right)^{\lambda_2/\lambda_1}$. However, due to both eigenvalues having the same sign in this case the curves are no longer hyperbolic. Instead, for $0 > \lambda_1 > \lambda_2$ they look like a set of parabolas tangent to the horizontal axis $\tilde{y_2} = 0$ at the origin of the $(\tilde{y_1}, \tilde{y_2})$ plane while for $\lambda_1 > \lambda_2 > 0$ they are tangent to the vertical axis $\tilde{y_1} = 0$; see Fig. 4.4 for examples.

The direction of motion along these curves is towards the origin if the eigenvalues are both negative, in which case the phase portrait is called a **stable node**. If the eigenvalues are both positive we have an **unstable node** with motion away from the origin; see again Fig. 4.4. Finally, if the initial conditions are chosen on the coordinate axis $\tilde{y_1} = 0$ or $\tilde{y_2} = 0$, the trajectory will coincide with the corresponding axis, which in turn defines an invariant manifold.

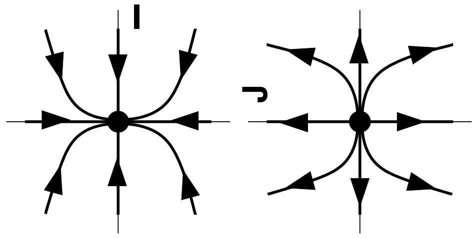

Figure 4.4: Phase portraits for a stable node with $0 > \lambda_1 > \lambda_2$ (left) and an unstable node with $\lambda_1 > \lambda_2 > 0$ (right).

In the original coordinates (y_1, y_2) the corresponding phase portraits retain these main features. Here the role of the invariant manifolds will be played by two straight lines intersecting at the origin defined by the corresponding eigenvectors, $y_2 = \frac{q_1}{p_1} y_1$ and $y_2 = \frac{q_2}{p_2} y_1$; see Fig. 4.3.

4.3.4 Phase portrait shape when $D = (\mathbf{Tr}A)^2 - 4\det A < 0$: Centre (stable, $\mathbf{Tr}A = 0$), Spiral in (stable, $\mathbf{Tr}A < 0$) and Spiral out (unstable, $\mathbf{Tr}A > 0$)

Now we consider the second general case $D = (\mathrm{Tr}A)^2 - 4\det A < 0$ starting with another example.

Example:
Find the general solution of the system of ODEs

$$\dot{y_1} = y_2 \,, \; \dot{y_2} = -2y_1 + 2y_2$$

and visualize the trajectory which corresponds to the initial conditions $y_1(0) = 0\,, y_2(0) = 1$.

Solution:

Rewriting the system in matrix form we find $A = \begin{pmatrix} 0 & 1 \\ -2 & 2 \end{pmatrix}$. Its eigenvalues and eigenvectors were already given in example (ii) on p.5 of the lecture notes of week 9 as $\lambda_1 = 1+i$, $\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1+i \end{pmatrix}$ and $\lambda_2 = 1-i$, $\mathbf{u}_2 = \begin{pmatrix} 1 \\ 1-i \end{pmatrix}$. Since the eigenvectors are linearly independent, according to (4.19) the general solution is

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = c_1 e^{(1+i)t} \begin{pmatrix} 1 \\ 1+i \end{pmatrix} + c_2 e^{(1-i)t} \begin{pmatrix} 1 \\ 1-i \end{pmatrix} \tag{4.23}$$

or equivalently

$$y_1 = e^t \left(c_1 e^{it} + c_2 e^{-it} \right) \ , \ y_2 = e^t \left(c_1 (1+i) e^{it} + c_2 (1-i) e^{-it} \right) \ .$$

The coefficients c_1, c_2 are determined by the above initial values,

$$0 = c_1 + c_2, \ 1 = c_1(1+i) + c_2(1-i) \ \Rightarrow \ c_1 = \frac{1}{2i}, \ c_1 = -\frac{1}{2i} \ .$$

This leads to the trajectory

$$y_1 = e^t \sin t \ , \ y_2 = e^t (\sin t + \cos t) \ ,$$

which describes a spiral in the form of a rotation in the (y_1, y_2) plane around the origin with period π, with the distance to the origin increasing exponentially in time. For example, the trajectory crosses the vertical axis $y_1 = 0$ periodically at times $t_n = 0, \pi, 2\pi, \ldots, \pi n, \ldots$, and the coordinates of the points of intersections are given by $(-1)^n e^{t_n}(0,1)$. Similarly, the trajectory intersects the diagonal $y_1 = y_2$ periodically at times $t_n^* = \frac{\pi}{2}, \frac{3}{2}\pi, \ldots, \left(n + \frac{1}{2}\right)\pi, \ldots$, and the coordinates of points of intersections are given by $(-1)^n e^{t_n^*}(1,1)$. We can find the direction of the tangent vector to the trajectory at $t = 0$ from the system of ODEs combined with the initial conditions yielding $\dot{y_1}(0) = y_2(0) = 1$, $\dot{y_2}(0) = -2y_1(0) + 2y_2(0) = 2$. Hence, the trajectory starts pointing towards the direction $(1, 2)$. The resulting trajectory is sketched in Fig. 4.5.

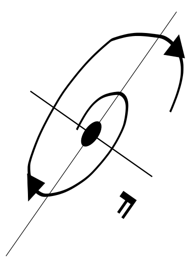

Figure 4.5: Sketch of the trajectory solving the initial value problem in the above example by spiraling away from the origin (please ignore the tilted axes).

Again we generalize our analysis for systems with $D < 0$ where the characteristic equation has two complex conjugate roots,

$$\lambda_{1,2} = r \pm i\omega \ , \ r = \frac{1}{2}Tr(A) \ , \ \omega = \sqrt{\det A - r^2} \ . \tag{4.24}$$

As we will see in a moment, the problem simplifies if we change variables $(x, y) \rightarrow (\tilde{y_1}, \tilde{y_2})$ by using the linear transformation

$$\tilde{\mathbf{y_1}} = W\mathbf{y_1} \ , \ W = \begin{pmatrix} a_{21} & r - a_{11} \\ 0 & \omega \end{pmatrix} \ . \tag{4.25}$$

In the new coordinates $(\tilde{y_1}, \tilde{y_2})$ the system of ODEs takes the form

$$\frac{d}{dt}\tilde{\mathbf{y_1}} = \tilde{A}\tilde{\mathbf{y_1}} \ , \ \tilde{A} = WAW^{-1} = \begin{pmatrix} r & -\omega \\ \omega & r \end{pmatrix} \ , \tag{4.26}$$

After this new transformation the matrix $\tilde{A}$ is not diagonal. Its eigenvalues and eigenvectors can be found to $\lambda_1 = r + i\omega$ with eigenvector $\mathbf{u}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ and $\lambda_2 = r - i\omega$ with $\mathbf{u}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$. According to (4.19) we can write the general solution to (4.26) as

$$\begin{pmatrix} \tilde{y_1} \\ \tilde{y_2} \end{pmatrix} = c_1 e^{(r+i\omega)t} \begin{pmatrix} 1 \\ -i \end{pmatrix} + c_2 e^{(r-i\omega)t} \begin{pmatrix} 1 \\ i \end{pmatrix} \ . \tag{4.27}$$

The values of c_1 and c_2 can be determined by the two initial values $\tilde{y_1}(0) = \tilde{a}$ and $\tilde{y_2}(0) = \tilde{b}$, which gives $\tilde{a} = c_1 + c_2, \tilde{b} = -i(c_1 - c_2)$ leading to $c_1 = \frac{\tilde{a}+i\tilde{b}}{2}, c_2 = \frac{\tilde{a}-i\tilde{b}}{2}$. Using Euler's formula $e^{\pm i\omega t} = \cos\omega t \pm i \sin\omega t$ we obtain the solution for given initial conditions

$$\tilde{y_1} = e^{rt}\left(\tilde{a}\cos(\omega t) - \tilde{b}\sin(\omega t)\right) \ , \ \tilde{y_2} = e^{rt}\left(\tilde{b}\cos(\omega t) + \tilde{a}\sin(\omega t)\right) \ . \tag{4.28}$$

These two equations can be combined to yield the identity

$$\tilde{y_1}^2 + \tilde{y_2}^2 = e^{2rt}(\tilde{a}^2 + \tilde{b}^2) \ , \tag{4.29}$$

which gives us the key for the phase portrait of the system.

Centre (stable) phase portrait

First suppose that $r = 0$. In this case every trajectory starting at $t = 0$ from the point $(\tilde{a}, \tilde{b})$ in the plane is represented by a circle of radius $\sqrt{\tilde{a}^2 + \tilde{b}^2}$ centered at the origin. The fixed point in such a phase portrait is called a **centre**; see Fig. 4.6 (left). In original coordinates circular trajectories are generally deformed into ellipses in the plane, see Fig. 4.6 (right). The arrows show the direction of rotation along the ellipses with increasing time. They can be inferred from the initial tangent vector $\dot{\mathbf{y}} = (\dot{y_1}(0), \dot{y_2}(0))^T$.

Stable focus (spiral in and stable) phase portait

If $r < 0$ we see that asymptotically both $\tilde{y_1} \rightarrow 0$ $(t \rightarrow \infty)$ and $\tilde{y_2} \rightarrow 0$ $(t \rightarrow \infty)$. Hence, every trajectory approaches the origin along a spiral of shrinking radius.

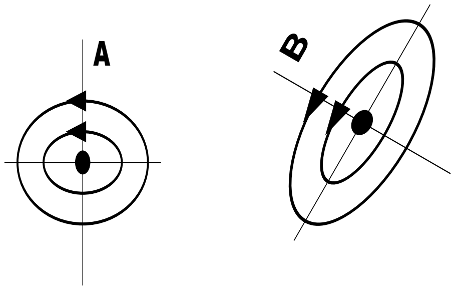

Figure 4.6: Phase portrait for a centre-type fixed point in transformed coordinates (left) and in original coordinates (right).

Unstable focus (spiral out and unstable) phase portrait
Similarly, if $r > 0$ the trajectory spirals away from the origin with ever increasing radius.

These two types of phase portraits are known as a **stable focus**, and respectively an **unstable focus**; see Fig. 4.7 for examples. Again, in original coordinates (y_1, y_2) the phase portraits retain their topological features, but the trajectories may be distorted, i.e., circular spirals may deform into an elliptic type, as we have already seen in the last example.

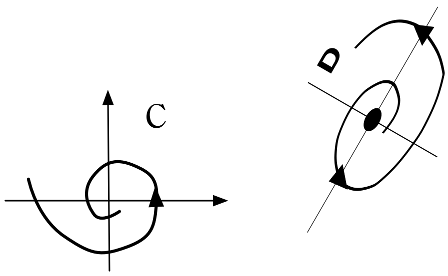

Figure 4.7: Phase portraits of fixed points of focus type: a stable one (left) and a distorted unstable one (right).

The only remaining case is $D = 0$, i.e., $(\mathrm{Tr}A)^2 = 4\det A$ where both eigenvalues are equal and real, $\lambda_1 = \lambda_2 = \lambda = \frac{a_{11}+a_{22}}{2}$. This is a specific case yielding a special type of nodes, and in this module we will not dwell upon its analysis. This completes our classification of the different types of phase portraits in autonomous systems of two linear ODEs. With more advanced mathematical techniques it is possible to prove that the phase portrait of a nonlinear system (4.5) in the *vicinity* of an isolated fixed point is *topologically equivalent* to the phase portrait of the corresponding linear approximation **provided the real parts of all eigenvalues are nonzero**. Some information about this fact will be given in the next chapter.

MTH5123 Differential Equations

Lecture Notes

Week 11

School of Mathematical Sciences
Queen Mary University of London

5 Stability of Solutions of ODEs

The subject of stability studies is to understand how a *change of initial conditions* or a *change in parameter values* of equations defining a dynamical system (e.g., coefficients in front of derivatives) affects the behaviour of the solutions, especially when the independent variable (usually interpreted as time t) tends to infinity, $t \to \infty$. The main goal is to establish criteria ensuring that the solution will change only slightly if a small (in an appropriate sense) change in the initial conditions or parameters is implemented. This type of question is of great importance for practical applications, as parameters of differential equations which govern real-life processes as, e.g., the functioning of mechanical aggregates or electronic devices, are known only approximately due to unpredictable changes in temperature, humidity or other properties of the environment. We will only discuss the stability of solutions of systems of two coupled first-order ODEs, but the basic ideas can be extended to any number of equations.
As usual we will use vector notation by writing down a system of two general non-autonomous ODEs in normal form as

$$\dot{\mathbf{y}} = \mathbf{f}(t, \mathbf{y}) \, , \; \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \, , \; \mathbf{f} = \begin{pmatrix} f_1(t, y_1, y_2) \\ f_2(t, y_1, y_2) \end{pmatrix} \, . \tag{5.1}$$

Stability theory is mainly based on the following two definitions:

Definition: Lyapunov stability
A solution $\mathbf{y}_*(t)$ of (5.1) corresponding to the initial condition $\mathbf{y}_*(0) = \mathbf{a_1} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}$ is called **Lyapunov stable** (or simply **stable**) if for any (arbitrarily small) $\epsilon > 0$ we can find a $\delta > 0$ such that if another initial condition $\mathbf{y}(0) = \mathbf{a_2} = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix}$ is chosen inside a circle of radius δ around the initial point $\mathbf{y}_*(0)$ then for any time $t > 0$ the solution $\mathbf{y}(t)$ corresponding to the initial condition $\mathbf{y}(0)$

1. exists, and
2. will stay inside a "tube" of radius ϵ around the solution $\mathbf{y}_*(t)$, see Fig. (5.1).

In mathematical shortcut notation this definition reads

$$\forall \epsilon > 0 \, \exists \delta > 0 \, \text{s.t.} \, \forall t > 0 \quad |\mathbf{y}(0) - \mathbf{y}_*(0)| < \delta \; \Rightarrow \; |\mathbf{y}(t) - \mathbf{y}_*(t)| < \epsilon \, .$$

Definition: asymptotic stability
The solution $\mathbf{y}_*(t)$ of (5.1) corresponding to the initial condition $\mathbf{y}_*(0) = \mathbf{a}$ is called **asymptotically stable** if it is

1. Lyapunov stable, and

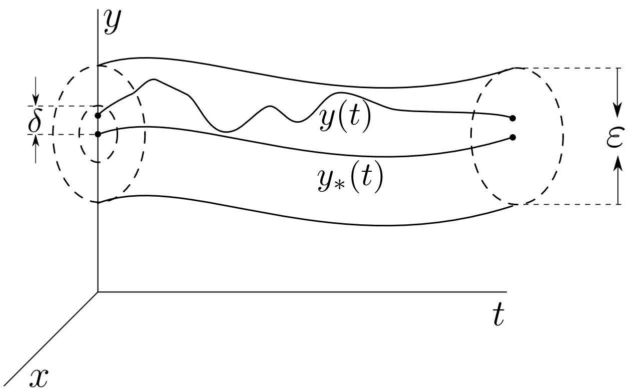

Figure 5.1: A sketch of the *stability tube* according to the definition of Lyapunov stability.
y

2. there exists a $\delta > 0$ such that the condition $|\mathbf{y}(0) - \mathbf{y}_*(0)| < \delta$ implies $|\mathbf{y}(t) - \mathbf{y}_*(t)| \to 0$ for $t \to \infty$.

y

Note:
1. Conditions 1. and 2. are *independent*, i.e., neither does 1. imply 2., nor does 2. imply 1. For the first direction see the counterexample below; the second direction is less obvious, but there are also counterexamples.
2. Making in (5.1) the change of variables $\mathbf{z}(t) \equiv \mathbf{y}(t) - \mathbf{y}_*(t)$ one can show that investigating the stability of any solution $\mathbf{y}_*(t)$ of the system (5.1) can always be reduced to investigating the stability of the *zero solution* $\mathbf{z}(t) = 0$ of the transformed system $\dot{\mathbf{z}} = \mathbf{f}(t, \mathbf{z})$; see our previous discussion at the beginning of Section 4.1.1.

Reformulating the definition of (Lyapunov) stability in terms of $\mathbf{z}(t)$ we arrive at the following expression: Stability of the zero solution $\mathbf{z} = \mathbf{0}$, i.e., of the fixed point at $\mathbf{z} = \mathbf{0}$, means that for any $\epsilon > 0$ we can find a $\delta > 0$ such that $\forall t > 0$ $|\mathbf{z}(0)| < \delta$ implies $|\mathbf{z}(t)| < \epsilon$. The definition of asymptotic stability can be reformulated accordingly. From now on we will concentrate on stability of the zero solution only.

Example:
Is the zero solution $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ of the system $\dot{y_1} = -4y_2$, $\dot{y_2} = y_1$ (Lyapunov) stable? Is it asymptotically stable?

Solution:
The system can be written as $\dot{\mathbf{y}} = A\mathbf{y}$ with $A = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix}$. The eigenvalues are purely imaginary, $\lambda_1 = 2i$, $\lambda_2 = -2i$, and the associated eigenvectors are $\mathbf{u}_{1,2} = \begin{pmatrix} 2 \\ \mp i \end{pmatrix}$. According to (4.19) the general solution of this system is given by

$$\mathbf{y}(t) = c_1 e^{2it} \begin{pmatrix} 2 \\ -i \end{pmatrix} + c_2 e^{-2it} \begin{pmatrix} 2 \\ i \end{pmatrix} .$$

Determining the two constants c_1, c_2 by imposing the initial conditions $y_1(0) = a, y_2(0) = b$ and expressing the solution in terms of real functions yields

$$y_1(t) = a\cos(2t) - 2b\sin(2t)\ ,\ y_2(t) = \frac{a}{2}\sin(2t) + b\cos(2t)\ .$$

As we have seen before, trajectories of this type are ellipses, $y_1^2 + 4y_2^2 = a^2 + 4b^2$. But this implies: Given any $\epsilon > 0$ let us choose $\delta = \epsilon/2$. Then by choosing the initial conditions (a, b) to be inside a circle of radius δ, that is, $a^2 + b^2 < \delta^2 = \epsilon^2/4$, we find that $\mathbf{y}^2(t) = y_1^2 + y_2^2 < y_1^2 + 4y_2^2 = a^2 + 4b^2 < 4(a^2 + b^2) < \epsilon^2$, hence $|\mathbf{y}| = \sqrt{y_1^2 + y_2^2} < \epsilon$ for any time t. We have thus shown that the zero solution $y_1 = y_2 = 0$ is Lyapunov stable. However, it is not asymptotically stable, as each solution rotates around its ellipse without approaching the origin for $t \to \infty$, see Fig. 5.

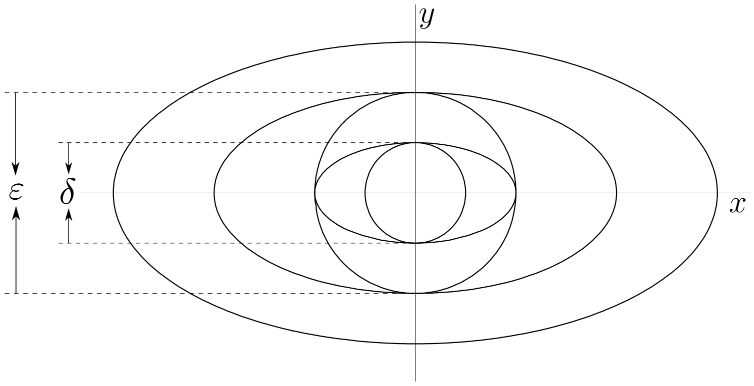

Figure 5.2: Sketch of the stability of the zero solution for elliptic trajectories providing an example that Lyapunov stability does not imply asymptotic stability.

5.1 Stability criteria for systems of two first-order linear ODEs with constant coefficents

Our goal is to formulate the stability conditions for the fixed point at $y_1 = y_2 = 0$ of any system of the form

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\ , \tag{5.2}$$

where the matrix A is time independent. We will furthermore assume that A is characterized by distinct eigenvalues $\lambda_1 \neq \lambda_2$. One can then prove the following statement:

Theorem:
Define $s \equiv \max\{Re\lambda_1, Re\lambda_2\}$. Then the zero solution $\mathbf{y} = 0$ of (5.2) is

1. unstable for $s > 0$,
2. stable for $s = 0$, and
3. asymptotically stable for $s < 0$.

Instead of providing a proof we just outline the basic idea of it: If $s > 0$ then $e^{st} \to \infty$ for $t \to +\infty$. Hence at least one of the factors $|e^{\lambda_1 t}|, |e^{\lambda_2 t}|$ (or both) grows without bound in time, and the modulus of the solution must increase as well implying instability. Similarly, if $s < 0$ then $e^{st} \to 0$ for $t \to +\infty$ implying that *both* $|e^{\lambda_1 t}|, |e^{\lambda_2 t}| \to 0$ as $t \to \infty$. This means that the modulus of the solution must vanish asymptotically for $t \to \infty$ so that system is asymptotically stable. Finally, for $s = 0$ the eigenvalues are purely imaginary and complex conjugate. We know that in this case the trajectories are ellipses that neither approach zero nor go to infinity. Instead, they remain at a finite distance from the origin, hence this case is stable but not asymptotically stable. For a proof these ideas need to be formalized in terms of equations by starting from the general solution (4.19) for the general initial value problem $y_1(0) = a, y_2(0) = b$.

Note:
Although we considered only the case of distinct eigenvalues $\lambda_1 \neq \lambda_2$ the theorem can be generalized to $\lambda_1 = \lambda_2 = \lambda$ showing that also in this case the zero solution is unstable for $\lambda > 0$, stable for $\lambda = 0$ and asymptotically stable for $\lambda < 0$.

5.2 Lyapunov function method for investigating stability

Consider again the general system of two first order ODEs written in normal form (5.1). Suppose that $\mathbf{y}(t)$ is a solution of (5.1). Then for any continuously differentiable function $V(\mathbf{y})$ defined on the same domain as $\mathbf{y}(t)$ one can define its values at any moment of time t *on the solution* $\mathbf{y}(t)$ as $v(t) \equiv V(y_1(t), y_2(t))$. We will need the expression for the time derivative $\dot{v} = \frac{dv}{dt}$ of such a function, which by using the chain rule of differentiation can be obtained to

$$\dot{v} = \frac{\partial V}{\partial y_1}\dot{y_1} + \frac{\partial V}{\partial y_2}\dot{y_2} \,. \tag{5.3}$$

Using (5.1) in the form of $\dot{y_1} = f_1(t, y_1, y_2), \dot{y_2} = f_2(t, y_1, y_2)$ we obtain

$$\dot{v} = \frac{\partial V}{\partial y_1} f_1(t, y_1, y_2) + \frac{\partial V}{\partial y_2} f_2(t, y_1, y_2) \equiv \mathcal{D}_f(V) \,, \tag{5.4}$$

where we introduced the notation $\mathcal{D}_f(V)$. In Calculus II you have learned that this equation defines the *directional derivative of* V *along* $\mathbf{f}$. Within our specific context, the above expression is called the **orbital derivative**. Note that $\mathcal{D}_f(V)$ is determined for any value of $\mathbf{y}$ solely by the functional form of $V(\mathbf{y})$ and the form of the right-hand side of the system (5.1) without the need to know the explicit solution of the latter system.
$\mathcal{D}_f(V)$ can be used to formulate the following statement, which we give without proof:

Theorem: Lyapunov Stability Theorem
Let $\mathbf{y}(t) = 0$ be a solution of (5.1) and assume that inside the circle $0 < |\mathbf{y}| < R$ there exists a continuously differentiable function $V(\mathbf{y})$ satisfying

1. $V(\mathbf{y} = 0) = 0$
2. $V(\mathbf{y} \neq 0) > 0$
3. The derivative of V along $\mathbf{f}$ is non-positive, $\mathcal{D}_f(V) \leq 0$ for $(y_1, y_2) \neq (0, 0)$.

Then the zero solution $\mathbf{y}(t) = 0$ is **stable**.

The function $V(\mathbf{y})$ featuring in this theorem is called the **Lyapunov function** of the system (5.1). While such a function can be found for certain classes of differential equations, see the following example, unfortunately there does not exist a systematic way of how to construct it for a given system of differential equations.

Note:
If the third condition is replaced by $\mathcal{D}_f(V) < 0$ being *strictly negative* one can prove that the zero solution $\mathbf{y}(t) = 0$ is **asymptotically stable**, which is called the **Lyapunov Asymptotic Stability Theorem**.

Example:
Verify that the function $V(y_1, y_2) = y_1^2 + y_2^2$ is a valid Lyapunov function for the system

$$\dot{y}_1 = -y_2 - y_1^3\,,\; \dot{y}_2 = y_1 - y_2^3\,.$$

Is the zero solution asymptotically stable?

Solution:
$V(y_1, y_2)$ satisfies the first and the second condition in the Lyapunov Stability Theorem. For any $(y_1, y_2) \neq (0,0)$ we have

$$\mathcal{D}_f(V) = \frac{\partial V}{\partial x}\dot{y}_1 + \frac{\partial V}{\partial y}\dot{y}_2 = 2y_1\left(-y_2 - y_1^3\right) + 2y_2(y_1 - y_2^3) = -2(y_1^4 + y_2^4) < 0$$

so that the zero solution is not only stable but even asymptotically stable.

Theorem: Consider a continuously differentiable function $V(\mathbf{y})$ satisfying

1. $V(\mathbf{y} = 0) = 0$
2. $V(\mathbf{y} \neq 0) > 0$

If the system (5.1) is autonomous and can be written as

$$\dot{y}_1 = -\frac{\partial V}{\partial y_1} \tag{5.5}$$

$$\dot{y}_2 = -\frac{\partial V}{\partial y_2} \tag{5.6}$$

then the dynamical system is called a **gradient flow**, $\mathbf{y} = \mathbf{0}$ is an equilibrium solution of the dynamical system which is Lyapunov stable. The function $V(\mathbf{y})$ is a Lyapunov function of the dynamical system called **potential**.

Proof:
In order to show that $\mathbf{y} = \mathbf{0}$ is an equilibrium solution we note that if condition 1. and 2. are satisfied then $\mathbf{y} = \mathbf{0}$ is a minimum of $V(\mathbf{y})$, thus

$$\left.\frac{\partial V}{\partial y_1}\right|_{\mathbf{y}=\mathbf{0}} = 0, \quad \left.\frac{\partial V}{\partial y_2}\right|_{\mathbf{y}=\mathbf{0}} = 0. \tag{5.7}$$

Hence $\mathbf{y} = \mathbf{0}$ is an equilibrium solution of the gradient flow.

The function $V(\mathbf{y})$ is a Lyapunov function for the gradient flow.Indeed it satisfies conditions 1. and 2. and its the orbital derivative is non-negative, as

$$D_f(V) = \frac{\partial V}{\partial y_1}\dot{y}_1 + \frac{\partial V}{\partial y_2}\dot{y}_2 = -\left[\left(\frac{\partial V}{\partial y_1}\right)^2 + \left(\frac{\partial V}{\partial y_2}\right)^2\right] \leq 0. \tag{5.8}$$

From the Lyapunov stability theorem it follows that the equilibrium solution $\mathbf{y} = \mathbf{0}$ is Lyapunov stable.

Example: Verify that the following dynamical system is a gradient flow and determine its Lyapunov function (potential):

$$\dot{y}_1 = -y_1 - y_1 y_2^2, \quad \dot{y}_2 = -2y_2 - y_1^2 y_2 \tag{5.9}$$

We want to express the dynamical system as a gradient flow of $V(y_1, y_2)$, i.e.

$$f_1(y_1, y_2) = -y_1 - y_1 y_2^2 = -\frac{\partial V}{\partial y_1} \tag{5.10}$$

$$f_2(y_1, y_2) = -2y_2 - y_1^2 y_2 = -\frac{\partial V}{\partial y_2}. \tag{5.11}$$

Integrating the first equation we get

$$V(y_1, y_2) = -\int f_1(y_1, y_2) dy_1 + g(y_2) = \frac{1}{2}y_1^2 + \frac{1}{2}y_1^2 y_2^2 + g(y_2). \tag{5.12}$$

Imposing $f_2(y_1, y_2) = -2y_2 - y_1^2 y_2 = -\frac{\partial V}{\partial y_2}$ we obtain $g'(y_2) = -2y_2$ leading to $g(y_2) = y_2^2 + C$. Imposing $V(\mathbf{0}) = 0$ we get

$$V(y_1, y_2) = \frac{1}{2}y_1^2 + y_2^2 + \frac{1}{2}y_1^2 y_2^2 \tag{5.13}$$

It is easy to check that $V(\mathbf{y}) > 0$ for $\mathbf{y} \neq \mathbf{0}$. Therefore $V(y_1, y_2)$ conditions 1. and 2. and is the Lyapunov function (potential function) of the considered dynamical system. Moreover, the considered dynamical system is the gradient flow of $V(\mathbf{y})$ and its equilibrium solution $\mathbf{y} = \mathbf{0}$ is Lyapunov stable.

The Lyapunov function method enables to investigate the stability of whole classes of systems of ODEs. For example, along these lines one can prove the following important generalization of the theorem on p.4, which we state without proof:

Theorem:
Let us consider a nonlinear system of two ODEs of the form

$$\begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \end{pmatrix} = A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \text{higher order nonlinear terms}\,, \tag{5.14}$$

where the matrix A is time independent and characterized by the two eigenvalues λ_1, λ_2. Then

1. if both $Re\lambda_1 < 0$ and $Re\lambda_2 < 0$ then the zero solution of (5.14) is asymptotically stable.

2. If at least one of $Re\lambda_1, Re\lambda_2$ is positive then the zero solution of (5.14) is unstable.

3. If $max\{Re\lambda_1, Re\lambda_2\} = 0$ then the stability of the zero solution is determined not only by A but also by the properties of the nonlinear terms, i.e., the zero solution may be stable for some nonlinear terms but unstable for others.

Note:
The third case implies that **linear stability analysis does not work** if $max\{Re\lambda_1, Re\lambda_2\} = 0$.

Example:
Determine the maximal range of the values of the parameter a for which the zero solution of the system

$$\dot{y_1} = y_1 + (2-a)y_2, \quad \dot{y_2} = ay_1 - 3y_2 + (a^2 - 2a - 3)y_1^2$$

is (**i**) unstable, (**ii**) stable.

Solution:
The linear part of the system is obtained by simply discarding the nonlinear terms in the second equation, as can be verified by Taylor expansion. Hence it is described by the matrix $A = \begin{pmatrix} 1 & 2-a \\ a & -3 \end{pmatrix}$ whose characteristic equation is $\lambda^2 + 2\lambda + (a^2 - 2a - 3) = 0$. The two roots are given by

$$\lambda_1 = -1 + \sqrt{-a^2 + 2a + 4}\,, \ \lambda_2 = -1 - \sqrt{-a^2 + 2a + 4}\,.$$

If the two roots are complex conjugate we have $Re\{\lambda_{1,2}\} = -1 < 0$, hence the zero solution is asymptotically stable. We conclude that an instability may occur only for values of a where both roots are real and $\lambda_1 > 0$. This implies $\sqrt{-a^2 + 2a + 4} > 1$ so that $-a^2 + 2a + 4 > 1$ or equivalently

$$a^2 - 2a - 3 = (a+1)(a-3) < 0\,,$$

hence $-1 < a < 3$. Thus for $a \in (-1, 3)$ the zero solution is unstable. Correspondingly, for $a < -1$ or $a > 3$ the zero solution must be asymptotically stable. For $a = -1$ or $a = 3$ we have $\lambda_1 = 0$ while $\lambda_2 = -2 < 0$. Therefore in this case the situation depends on the nonlinear terms. But precisely for these parameter values of a the nonlinear term in our original system vanishes, due to $a^2 - 2a - 3 = 0$. The system thus becomes linear with eigenvalues $\lambda_1 = 0, \lambda_2 = -2$, hence by our theorem on page 4 of this document for $a = -1, 3$ the zero solution is stable but not asymptotically stable.