Math 2552: Ordinary Differential Equations

Dmitrii M. Ostrovskii
ostrov@gatech.edu

Fall 2024

Disclaimer. At a high level, these lectures follow the outline of Brannan and Boyce (2015) [BB15].

Chapter 1

First-order linear equations

Preliminaries

Recap of the introductory lecture: a general **1st-order ordinary differential equation (ODE)** with independent variable t and dependent variable (or unknown function) $u : t \mapsto u(t)$ has the form

$$\Phi(t, u, u') = 0 \tag{1.1}$$

where $\Phi : \mathbb{R}^3 \to \mathbb{R}$ is some function defining the equation.[1] Recall also that any differentiable function $u : I \to \mathbb{R}$ that satisfies (1.1) for all t in some interval I, is called **a (particular) solution** of (1.1) on I, and the set of all such functions is called **the general solution** or **solution set** of (1.1). Sometimes, the algebraic equation $\Phi(t, u, v) = 0$ can be solved for v, in which case (1.2) reduces to

$$u' = F(t, u) \tag{1.2}$$

for some function $F : \mathbb{R}^2 \to \mathbb{R}$. We call (1.2) a 1st-order ODE **in the standard (or normal) form**. Not every ODE can be put to the standard form, but we shall only be concerned with such ODEs.

In fact, there is *no general method* for "solving" an arbitrary first-order ODE.[2] This task can be done *numerically*, i.e. one can *approximate* $u(t)$ with given accuracy; we shall briefly discuss such *numerical methods* at some point later on. In this chapter, we shall focus on some *subclasses* of 1st-order ODEs that *can* be solved exactly: **separable** (Sec. 1.1), **linear** (Sec. 1.2), and **autonomous** (Sec. 1.5) equations. Some other equations that can be reduced to these types; we shall briefly mention those.[3]

1.1 Separable equations

In this section, we shall use x and y instead of t and u; the reason for this shall become clear shortly.

[1] Note that $\Phi(t, u, v) = 0$ is simply an *algebraic* ("usual") equation in variables t, u, v. It becomes an ODE when we impose the dependencies $u = u(t)$ and $v = u'(t)$.

[2] Informally, expressing solutions $u = u(t)$ in terms of "elementary" functions and algebraic operations over them.

[3] The sections in this chapter *roughly* correspond to Secs. 2.1–2.5 in [BB15].

Definition 1. *A 1st-order ODE of the form*

$$y' = \frac{p(x)}{r(y)}, \tag{1.3}$$

where p, r are some functions, is called **separable***.*

In other words, an ODE (1.2) is separable if its rate function F decouples into the product of two terms, each term depending only on its own variable. Of course, both examples from Sections 1.1-1.2 (the Newton law of cooling and the population dynamics equations) are separable. Also, the choice of p and r, for a specific equation, is not unique: multiplying p, r by $c \neq 0$ does not change (1.3).

Separation of variables. This is a general method for solving separable 1st-order ODEs. Let us first present the method itself, before any justification. Let $P(x)$ and $R(y)$ be, respectively, the (arbitrary) antiderivatives for $p(x)$ and $r(y)$; in other words,

$$P(x) = \int p(x)dx, \quad R(y) = \int r(y)dy$$

or, equivalently, $P'(x) = p(x)$ and $R'(y) = r(y)$. Introducing the *differential* $dy := y'(x)dx$ and multiplying by $r(y)$, we rewrite (1.3) in the differential form:

$$r(y)dy = p(x)dx. \tag{1.4}$$

Now let's integrate separately both sides of this equation, treating both x and y as independent variables. This gives $R(y) + C_1 = P(x) + C_2$, where C_1 and C_2 are two arbitrary constants, that is

$$R(y) = P(x) + C \tag{1.5}$$

where C is an arbirary constant. As it turns out,(1.5) *is an equation of the integral curves of* (1.4)*:* in other words, any function $y = \phi(x)$ that satisfies (1.5) $\forall x \in I$, also satisfies (1.1) on I. In this sence, (1.5) already defines the solutions of (1.1) *implicitly.* Furthermore, the solution to the *initial value problem*, i.e. the solution that satisfies $y(x_0) = y_0$—or the integral curve passing through (x_0, y_0)—is (1.5) with

$$C = R(y_0) - P(x_0). \tag{1.6}$$

Proof. We have to prove that a function $y = y(x)$ satisfying (1.5) also satisfies (1.3) (i.e. that (1.5) describes integral curves for (1.3)), and conversely, that *any* integral curve satisfies (1.5) for some C.

1. Let $y = \phi(x)$ satisfy (1.5), that is

$$R(\phi(x)) = P(x) + C.$$

Then, differentiating both sides in x and using the chain rule, we get

$$R'(\phi(x))\,\phi'(x) = P'(x),$$

that is

$$r(\phi(x))\,\phi'(x) = p(x). \tag{1.7}$$

Dividing over $r(\phi(x))$ we get $\phi'(x) = \frac{p(x)}{r(\phi(x))}$ So, $y = \phi(x)$ is indeed a solution of (1.3).

2. Conversely, assume that $y = \phi(x)$ satisfies (1.3). Then it satisfies (1.7). The two sides of (1.7) are functions of x; since these functions are equal (for $x \in I$), their antiderivatives are also equal, up to an additive constant. Meanwhile, the antiderivative of the RHS is $P(x)$, and the antiderivative of the LHS is $R(\phi(x))$. Thus, $R(\phi(x)) = P(x) + C$, as claimed.

3. Moreover, by the fundamental theorem of calculus we also get

$$\int_{x_0}^{x} r(\phi(t))\, \phi'(t)dt = \int_{x_0}^{x} p(t)dt, \quad \text{that is} \quad \int_{\phi(x_0)}^{\phi(x)} r(\phi(t))\, d\phi(t) = \int_{x_0}^{x} p(t)dt,$$

that is

$$R(\phi(x)) - P(x) = R(\phi(x_0)) - P(x_0).$$

In other words, the solution of (1.3) passing through the point (x_0, y_0) is given by (1.5), (1.6). □

Example 1.1.1. *Solve the initial value problem and determine the interval where the solution exists:*

$$\frac{dy}{dx} = -\frac{x}{y}, \quad y(0) = 2.$$

In the differential form, the equation becomes

$$xdx = -ydy,$$

which integrates to

$$x^2 + y^2 = C \quad \forall C \in \mathbb{R}.$$

In fact, since in the LHS we have a sum of squares, $C < 0$ give no integral curves. When $C > 0$, the corresponding integral curve is an origin-centered circle with squared radius $C := r^2 > 0$. This curve consists of two solutions,

$$y(x) = \pm\sqrt{r^2 - x^2},$$

each existing on the interval $I(r) := (-r; r)$. Plugging in the initial condition we get $r = 2$. This also allows to choose the sign: $y(x) = \sqrt{4 - x^2}$ satisfies the condition, while $\sqrt{4 - x^2}$ does not. Hence, the IV problem solution is given by

$$y(x) = \sqrt{4 - x^2}, \quad -2 < x < 2.$$ □

Solve the previous problem replacing the initial condition with (a) $y(0) = 0$; (b) $y(2) = 0$.

In the case (a), we get $C = 0$, and the integral curve reduces to a single point (i.e. $I = \emptyset$). In the case (b), we get $C = 4$, but we cannot choose the sign: $y(x) = \pm\sqrt{4 - x^2}$ both work – in other words, the IVP has *two* solutions. In both cases, the IVP turns out to be ill-posed. Later on, we shall find out assumptions under which the IVP is guaranteed to be well-posed, i.e. admit a *unique* solution existing on a nontrivial interval.

Example 1.1.2 ([BB15, Ex. 2.1.2])**.** *Solve the initial value problem*

$$\frac{dy}{dx} = \frac{3x^2 + 4x + 2}{2(y - 1)}, \quad y(0) = -1$$

and determine the interval I where the solution exists.

Here $p(x) = 3x^2 + 4x + 2$ and $r(y) = 2(y - 1)$; in the differential form the equation is

$$2(y-1)dy = (3x^2 + 4x + 2)dx.$$

Integrating this, we get the equation of integral curves:

$$y^2 - 2y = x^3 + 2x^2 + 2x + C.$$

To identify the curve satisfying the initial condition $y(0) = -1$, we plug in $(x_0; y_0) = (0; -1)$ into this equation, and find $C = y_0^2 - 2y_0 - (x_0^3 + 2x_0^2 + 2x_0) = 3$, that is

$$y^2 - 2y = x^3 + 2x^2 + 2x + 3.$$

Now, let's find the solution explicitly (and find I). Solving the previous equation for y, we get

$$y = 1 \pm \sqrt{x^3 + 2x^2 + 2x + 4}.$$

Note that this describes *two* functions:

$$\phi_+(x) = 1 + \sqrt{x^3 + 2x^2 + 2x + 4},$$
$$\phi_-(x) = 1 - \sqrt{x^3 + 2x^2 + 2x + 4},$$

and we must identify the one satisfying the initial condition. We find that $\phi_+(0) = 3$ and $\phi_-(0) = -1$; hence, $y = \phi_-(x)$ is our solution. Finally, finding the interval I where this solution exists—a.k.a. finding the interval where ϕ_- is well-defined—amounts to solving the inequality $x^3+2x^2+2x+4 \geqslant 0$. While cubic equations can be solved by a general method, the formulas are bulky and no sane person remembers them. Instead, we *guess* one root:

$$x = -2.$$

(This is an "educated" guess: we need "-" because all coefficients in the left-hand side are positive.) Thus, by the fundamental theorem of algebra, we can factor out $x + 2$ from $x^3 + 2x^2 + 2x + 4$:

$$x^3 + 2x^2 + 2x + 4 = (x + 2)(ax^2 + bx + c).$$

Equating the coefficients of the polynomials in the LHS and RHS (this is called "the undetermined coefficients method"), we get $a = 1$, $b = 0$, $c = 2$. That is,

$$x^3 + 2x^2 + 2x + 4 = (x + 2)(x^2 + 2).$$

Since $x^2 + 2 > 0$ for all x, there are no other (real) roots; the sign of $x^3 + 2x^2 + 2x + 4$ is the same as the sign of $x + 2$. Hence, $I = (-2, +\infty)$. □

Example 1.1.3 (*based on [BB15, Ex. 2.1.1])**.** *Solve the initial value problem*

$$\frac{dy}{dx} = \frac{x^2}{1 - y^2}, \quad y(0) = -2.$$

and determine the interval I where the solution exists.

Proceeding as in the previous example, we find the equation of integral curves:

$$y^3 - 3y + x^3 = C.$$

Plugging in the initial conditions we get $C = -2$, i.e. the integral curve passing through $(0; -2)$ is

$$x^3 + y^3 - 3y + 2 = 0.$$

Now, to obtain the IV problem solution explicitly, one could solve this cubic equation for y. But we proceed more elegantly: instead of finding the function $y(x)$, let us find instead $x(y)$. Graphically, passing from $y(x)$ to $x(y)$ corresponds to interchanging the x and y axes, i.e. reflecting the plot w.r.t. the line $y = x$. To that end, solving the equation for x we get:

$$x(y) = -(y^3 - 3y + 2)^{1/3}.$$

To plot the RHS, we first guess one root $y = 1$ of the RHS, then find (by undetermined coefficients or guessing another root $y = -2$) that $y^3 - 3y + 2 = (y + 2)(y - 1)^2$. This allows to sketch the plot of the RHS (see Fig. 1.1): the RHS changes sign only once at $y = -2$, its derivative vanishes at $y = \pm 1$, and $x(-1) = -1$. Redrawing the curve in the xy coordinates, we conclude that $I = (-1, +\infty)$. □

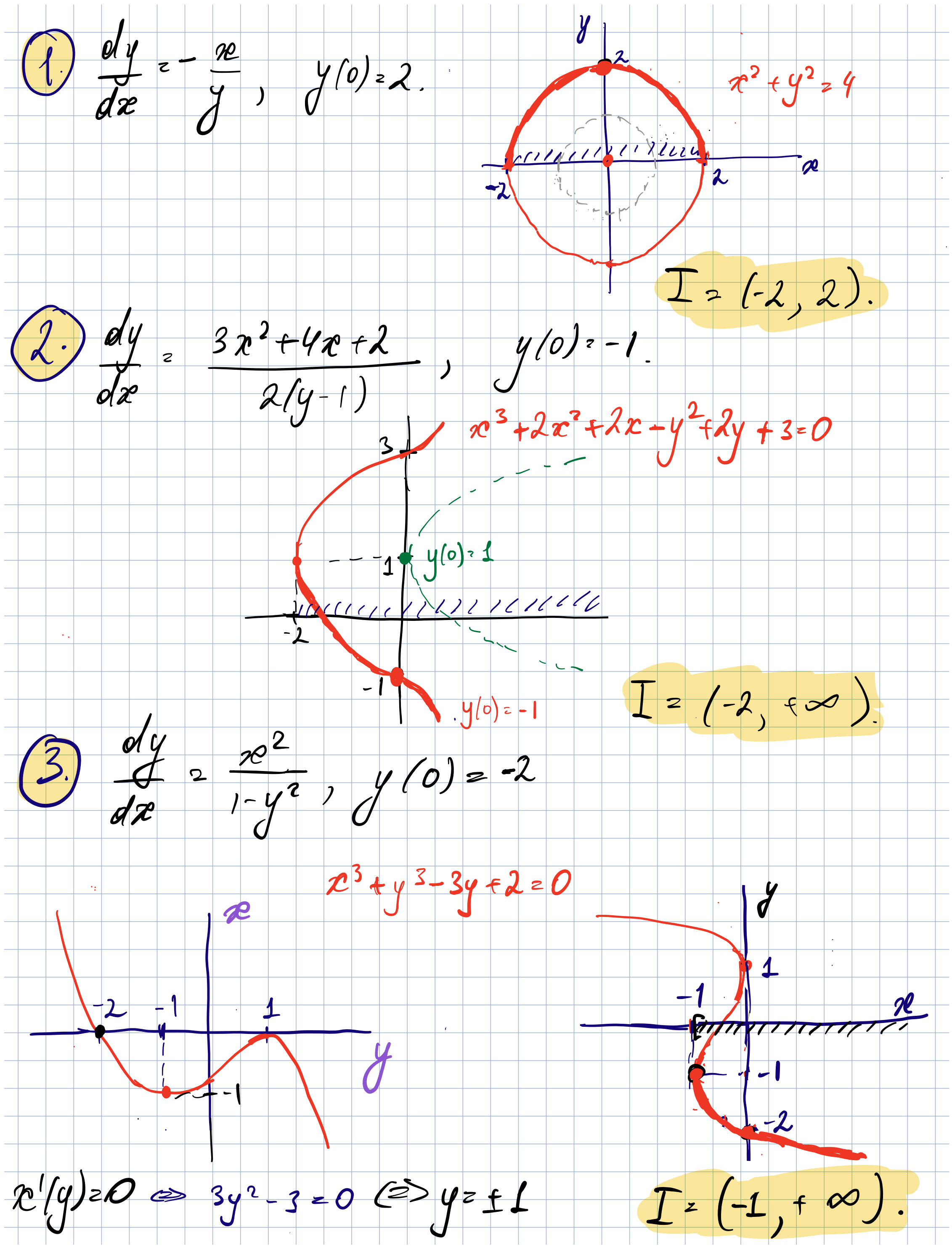

Figure 1.1: Integral curves for separable ODEs in Examples 1.1.1– 1.1.3.

1.2 Linear equations

Recall the form of a **linear** nth-order ODE with independent variable t and function $u = u(t)$:

$$a_0(t)u^{(n)}(t) + a_1(t)u^{(n-1)}(t) + \cdots + a_n(t)u(t) = b(t)$$

for some $a_0(t), ..., a_n(t)$ and $b(t)$. It is called *homogeneous* if $b(t) \equiv 0$, and *nonhomogeneous* otherwise. Such an ODE can always be put to the standard form by dividing over $a_0(t)$. When $n = 1$ this gives

$$u'(t) + p(t)u(t) = g(t). \tag{1.8}$$

Note that in the *homogeneous* case, we have a *separable* equation that can be solved by separating the variables (see Section 1.1). *Nonhomogeneous* linear equations are *not separable*, yet can also be solved. This can be done by a general method explained below.

Method of integrating factors. Let $\mu(t) > 0$ be a differentiable function, yet to be specified. Multiplying (1.8) by $\mu(t)$ we get

$$\mu(t)u'(t) + \mu(t)p(t)u(t) = \mu(t)g(t). \tag{1.9}$$

Now, take $\mu(t)$ of the form

$$\mu(t) = e^{P(t)}$$

where $P(t) = \int p(t)dt$ is an antiderivative of $p(t)$. Note that $\mu(t) > 0$ and, by the chain rule,

$$\mu'(t) = \mu(t)p(t).$$

Thus, (1.9) is reduced to

$$\underbrace{\mu(t)u'(t) + \mu'(t)u(t)}_{v'(t)} = \underbrace{\mu(t)g(t)}_{h(t)} \tag{1.10}$$

where $v(t) := \mu(t)u(t)$ is the new unknown function (even though $\mu(t)$ is known, v is unknown because u is), and $h(t)$ is the new RHS. Now, behold: this is a *separable* equation; we can solve *it* by the separation of variables:

$$v(t) = \int h(t)dt + c = \int \mu(t)g(t)dt + c = \int e^{P(t)}g(t)dt + c,$$

and then recover the solution of (1.8) by dividing over $\mu(t)$:

$$\boxed{u(t) = \frac{1}{\mu(t)}v(t) = e^{-P(t)}\left(\int e^{P(t)}g(t)dt + c\right) \quad \text{where} \quad P(t) = \int p(t)dt.} \tag{1.11}$$

Remark 1.2.1. By the fundamental theorem of calculus, we can rewrite (1.11) with *definite* integrals:

$$u(t) = e^{-P(t)}\left(\int_{t_0}^{t} e^{P(s)}g(s)ds + u_0\right) \quad \text{where} \quad P(t) = \int_{t_0}^{t} p(s)ds, \tag{1.12}$$

for arbitrary $t_0, u_0 \in \mathbb{R}$. At the same time, this is the solution for the IV problem with $u(t_0) = u_0$.

Example 1.2.1 (cf. [BB15, Ex. 2.2.1])**.** *Solve the initial value problem*

$$u' - 2u = 4 - t, \quad u(1) = -\frac{5}{4}.$$

Here $p(t) \equiv -2$ and $g(t) = 4 - t$. Instead of using the ready formula (1.12), let's be incremental and find the general solution first, cf. (1.11). We take $P(t) = -2t$, then $\mu(t) = e^{-2t}$ and

$$v(t) = \int (4-t)e^{-2t}dt = -2e^{-2t} - \int te^{-2t}dt = -2e^{-2t} + \frac{1}{4}(2t+1)e^{-2t} + c = \frac{1}{4}(2t-7)e^{-2t} + c$$

where we took the second integral by parts. This results in

$$u(t) = \frac{1}{4}(2t-7) + ce^{2t}.$$

Plugging in the IV condition we find $c = 0$, so the desired solution is $u(t) = \frac{1}{4}(2t-7)$, on $t \in \mathbb{R}$. □

Example 1.2.2. *Solve the initial value problem*

$$tu' - \frac{2}{t}u = te^{-t}, \quad u(1) = -1.$$

We rewrite the equation in standard form

$$u' - \frac{2}{t^2}u = e^{-t},$$

so that $p(t) = -2t^{-2}$ and $g(t) = e^{-t}$. Then compute $P(t)$ as in (1.12):

$$P(t) = \int_{t_0}^{t} -2s^{-2}ds = 2t^{-1}\Big|_{t_0}^{t} = \frac{2}{t} - \frac{2}{t_0} = \frac{2}{t} - 2.$$

Whence we get

$$\mu(t) = e^{2t^{-1}-2},$$
$$v(t) = \int_{t_0}^{t} \mu(s)g(s)ds + u_0 = e^{-2}\int_{t_0}^{t} e^{2s^{-1}-s}ds + u_0.$$

The last integral cannot be evaluated in terms of elementary functions, so we leave it as is. Thus,

$$u(t) = \frac{v(t)}{\mu(t)} = e^{2-2t^{-1}}\left(e^{-2}\int_{t_0}^{t} e^{-2+2s^{-1}-s}ds + u_0\right) = e^{-2t^{-1}}\left(\int_{t_0}^{t} e^{2s^{-1}-s}ds + e^2 u_0\right)$$
$$= e^{-2t^{-1}}\left(\int_{1}^{t} e^{2s^{-1}-s}ds - e^2\right).$$

If you have ample time, you might want to verify the solution, or at least the initial condition. Finally, we note that the function in the RHS is defined for $t \neq 0$, and diverges to ∞ at $t = 0$. Hence, the solution to the IV problem is defined on $t > 0$ (since this interval includes $t_0 = 1$). □

1.3 Modeling with first-order ODEs

We consider two examples in this section. In both of them, we shall model a physical process/experiment with a first-order linear ODE. In fact, many ODEs arising in mathematical models are lineal, though not first-order. Fortunately, soon we shall learn how to reduce a higher-order linear ODE to a *first-order* linear *system* of ODEs, and also some methods for solving such systems.

Example 1.3.1 (Water tank). *A tank initially contains* 40 *pounds of salt dissolved in* 600 *gallons of water. Starting at time* $t = 0$*, water that contains 1/2 pound of salt per gallon is poured into the tank at the rate of* 4 *gal/min and the mixture is drained from the tank at the same rate (see Fig. 1.2).*

a. *Write down a diff. equation for* $Q(t)$*, the amount of salt (in pounds) in the tank at time* $t > 0$.
b. *Solve the DE to determine an expression for* $Q(t)$.
c. *After a long period of time, what happens to the concentration of salt in the tank?*

To begin with, we observe that the rate of change of the amount of salt in the tank, that is $\frac{dQ}{dt}$, is comprised of the two terms:

$$\frac{dQ}{dt} = \text{“rate in”} - \text{“rate out”}.$$

Naturally, let us measure Q as pounds (pnd), and t in minutes (min); then both sides of the above equation are measured in pounds-per-minute (pnd/min).[4] Now, for the “rate in” we have

$$\text{“rate in”} = \frac{1}{2}\frac{\text{pnd}}{\text{gal}} \cdot 4\frac{\text{gal}}{\text{min}} = 2\frac{\text{pnd}}{\text{min}}.$$

For the “rate out,” we are not told what the concentration term is, so let's figure it out. *What happens here, really?* Well, it is reasonably to assume that salt in the tank gets uniformly dissolved in the water – and it is this “properly mixed” water that flows out of the tank, isn't it? As such, the concentration of salt in the water flowing out, at time t, is given by $\frac{Q(t)\text{pnd}}{600\text{gal}} = \frac{Q(t)}{600}\frac{\text{pnd}}{\text{gal}}$. As the result,

$$\text{“rate out”} = \frac{Q}{600}\frac{\text{pnd}}{\text{gal}} \cdot 4\frac{\text{gal}}{\text{min}} = \frac{Q}{150}\frac{\text{pnd}}{\text{min}}.$$

Omitting the units, we arrive at the following ODE for $Q(t)$

$$Q'(t) + \frac{1}{150}Q(t) = 2.$$

This is, in fact, a first-order linear ODE with constant coefficients. The corresponding IVP has the initial condition that $Q(0) = 40$. We compute the integrating factor $\mu(t) = e^{\int \frac{1}{150}dt} = e^{\frac{t}{150}}$, whence

$$Q(t) = e^{-\frac{t}{150}} \int 2e^{\frac{t}{150}}\,dt = e^{-\frac{t}{150}}(300e^{\frac{t}{150}} + C) = 300 + Ce^{-\frac{t}{150}}$$

and $C = 40 - 300 = -260$ by plugging the initial condition. Thus,

$$Q(t) = 300 - 260e^{-\frac{t}{150}}.$$

[4] Note that units are helpful here, as a sanity check. In particular, it only makes sense to add/subtract two quantities measured in the same units; besides, if the units of f are α (written as $[f] = \alpha$) and those of g are β, then $[fg] = \alpha\beta$.

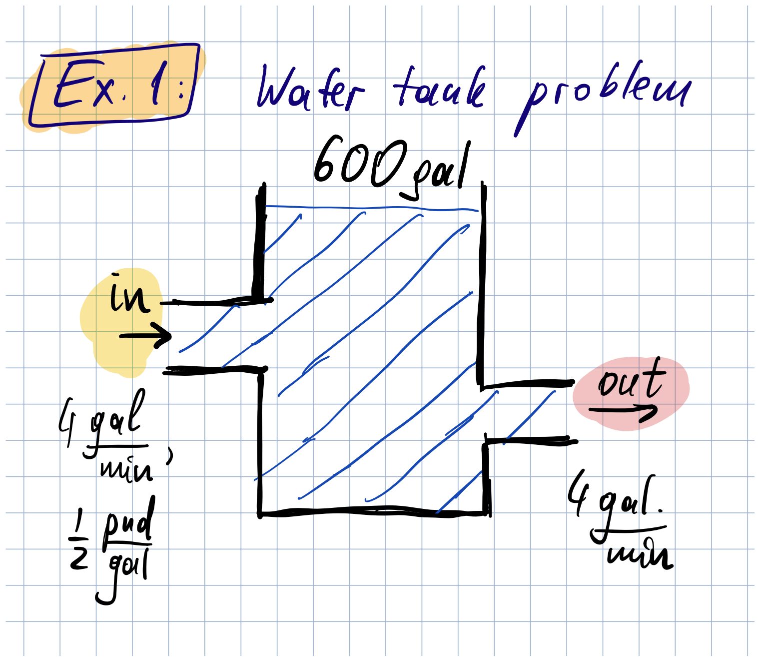

Figure 1.2: Water tank problem (Example 1.3.1).

Clearly, $Q(t) \to 300$ as $t \to \infty$, which answers the last question. Note also that we could arrive at the same conclusion without solving the equation, but via physical intuition. Indeed, since the concentration of salt in the incoming water is kept constant at $\frac{1}{2}\frac{\text{pnd}}{gal}$, and the water in the tank gradually gets replaced with incoming water, after a long time the concentration of salt in the tank will approach $\frac{1}{2}\frac{\text{pnd}}{gal}$, which corresponds to 300 pnd of salt in 600 gal of water. □

The process we used in the previous example roughly followed this outline:

1. Construct a differential equation to model a real-world situation.
2. Solve the differential equation so that we can interpret its solution to characterize a system.
3. Analyze mathematical statements and solutions of differential equations.

The above "algorithm" is used throughout this course. Oftentimes, the first step is the most difficult.

Example 1.3.2 (World population growth)**.** *The world population in 2018 was about 7.6 billion.*

1. *The world population is increasing at a rate of 1.2% per year. If the growth rate remains fixed at 1.2%, how long will it take for the population of the world to reach 20 billion people?*
2. *Assume Earth cannot support a population beyond 20 billion people. If the population growth rate is* ***also*** *proportional to the difference between this limiting value and the current value. What is the expression that gives the world population as a function of time?*

1. First note that, literally, $\% = 0.01$. Hence, the rate of increase 1.2%/year means $r = 0.012$/year. That is, we have the following ODE for the world population $P(t)$ at time t:

$$P' = rP = 0.012P$$

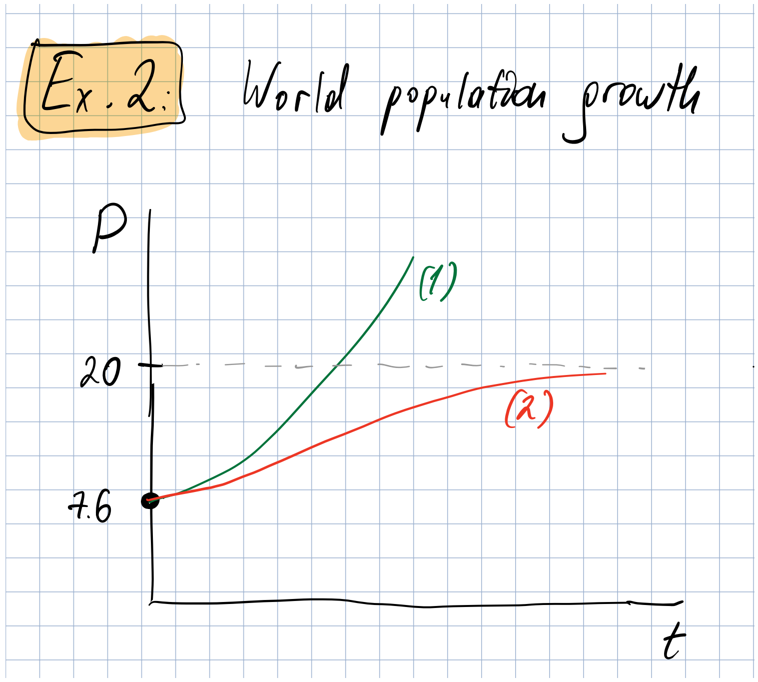

Figure 1.3: World population growth model (Example 1.3.2).

where the units are $[P] = \text{bln}$, $[P'] = \text{bln/year}$, and $[r] = \text{year}^{-1}$; in the ODE the units are hidden. Solving it with the condition that $P \geqslant 0$ we get $P(t) = e^{rt+C}$ for any $C \in \mathbb{R}$, i.e. $P(t) = ce^{0.012t}$ for $c > 0$. The initial condition gives $c = 7.6$. (What are the units of c? And those of the factor $e^{0.012t}$?) Finally, defining the moment of reaching 20 bln amounts to solving the following equation for t:

$$7.6e^{0.012t} = 20.$$

Whence 20 bln will be reached in $\frac{1}{0.012}\ln(20/7.6) \approx 80.632$ years.

2. In the second case, the equation becomes

$$P' = \lambda r P(P_{\max} - P)$$

where $P_{\max} = 20\text{bln}$, and λ the normalization factor. Note that $[\lambda] = \text{bln}^{-1}$. (Why?) Moreover, it would be reasonable to assume that $\lambda = 1/P_{\max}$, since in this case the equation reduces to the previous one for small values of P, i.e. those far away from $P_{\max}$. (Give a motivation for this.) Thus,

$$P' = rP\left(1 - \frac{P}{P_{\max}}\right) = 0.012P\left(1 - \frac{1}{20}P\right).$$

This is a separable ODE, and separating the variables we get

$$\frac{20dP}{P(20-P)} = 0.012\,dt.$$

By partial fractions

$$\frac{20dP}{P(20-P)} = \frac{dP}{P} + \frac{dP}{20-P},$$

and integration gives

$$\ln\left(\frac{P}{20-P}\right) = 0.012t + C, \quad C \in \mathbb{R}.$$

To solve for P we note that $\frac{P}{20-P} = -1 + \frac{20}{20-P}$, whence $\frac{20}{20-P} = 1 + ce^{0.012t}$ for $c > 0$. From the IV condition we find $c = \frac{20}{20-7.6} - 1 = \frac{7.6}{20-7.6} = 0.613$, therefore

$$P(t) = 20\left(1 - \frac{1}{1 + 0.613e^{0.012t}}\right) = 12.26\frac{e^{0.012t}}{1 + 0.613e^{0.012t}}.$$

Note that for small t, the denominator approaches 1, and the population grows nearly exponentially. As t increases, the population saturates at the level $12.26/0.613 = 20$, never reaching it (Fig. 1.3). $\square$

1.4 Existence and uniqueness of solutions in first-order IVPs

In this lecture, we shall explore some differences between linear and nonlinear 1st-order ODEs. Our main focus will be on the conditions for the corresponding IVPs to be "well-posed," i.e. have a unique solution. To this end, we shall formulate two theorems that give sufficient conditions for linear and general first-order ODEs, respectively, and look at some examples that "break" them. Our second goal would be understanding how to determine the intervals of existence for IVP solutions.

But first, consider a motivating example.

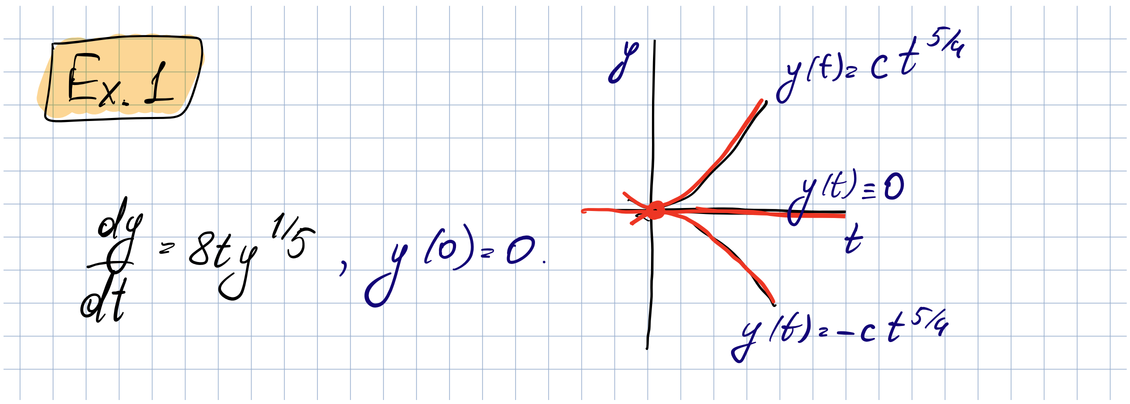

Figure 1.4: Nonlinear IVP with nonunique solutions (Example 1.4.1).

Example 1.4.1. *Consider the IVP*

$$y' = 8ty^{1/5}, \quad y(0) = 0.$$

(a) Is the ODE linear? (b) Solve the IVP explicitly and find the interval where the solution exists.

If course, the ODE is nonlinear. Separation of variables leads to $y^{-1/5}dy = 8tdt$, which results in

$$y^{4/5} = \frac{16}{5}t^2 + C$$

for $C \in \mathbb{R}$. Plugging in the IV condition we find $C = 0$, so the corresponding implicit equation reads

$$y^{4/5} = \frac{16}{5}t^2.$$

Solving for y we get

$$y = \pm\frac{32}{5^{5/4}}t^{5/2}.$$

Now, behold: I cheated when separating the variables! We divided by $y^{1/5}$, but ignored the possibility that $y = 0$. This case gives a stationary solution $y \equiv 0$ that also satisfies the IV condition. Thus, the IVP has three different solutions, each of them existing on the whole real axis (see Fig. 1.4). □

We shall revisit Example 1.4.1, and give other "bad" examples, later. Now we state the first result.

Theorem 1.4.1 (Existence & uniqueness of solutions for linear IVPs)**.** *Let (α, β) be an open interval containing t_0 and such that the functions p, g are defined and continuous on (α, β). Then the IVP*

$$y' + p(t)y = g(t), \quad y(t_0) = y_0 \tag{1.13}$$

has a unique solution for any $y_0 \in \mathbb{R}$.

Note that the premise of Theorem 1.4.1 allows the interval to be infinite, i.e. with $\alpha, \beta \in \{-\infty, +\infty\}$.

Proof. In a nutshell, the key idea is that under the premise of the theorem, each step of the integrating factor method is correct, so the method can be applied and gives the general ODE solution. One can then identify the IVP solution among these solutions, and verify its uniqueness.

1$^{\boldsymbol{o}}$. First note that since p is continuous (and thus integrable) on $I := (\alpha, \beta) \ni t_0$, the function

$$\mu(t) := \exp\left(\int_{t_0}^{t} p(s)ds\right)$$

is defined, positive, and differentiable on I, with $\mu'(t) = p(t)\mu(t)$. In particular, μ is continuous on I; therefore, μg is also continuous, and thus integrable. As such, for any $C \in \mathbb{R}$ the function

$$\phi(t) := \frac{1}{\mu(t)}\left(\int_{t_0}^{t} \mu(s)g(s)ds + C\right)$$

is defined on I, and is differentiable on I with derivative given by

$$\phi'(t) = \frac{1}{\mu(t)}\mu(t)g(t) - \frac{\mu'(t)}{\mu^2(t)}\left(\int_{t_0}^{t} \mu(s)g(s)ds + C\right) = g(t) - \frac{\mu'(t)}{\mu(t)}\phi(t) = g(t) - p(t)\phi(t).$$

Here we first used the product rule, then the chain rule on $1/\mu(t)$. Thus, $\phi(t)$ satisfies the ODE. Moreover, $\phi(t)$ with $C = y_0$ satisfies the IV condition. Thus, we showed that (1.13) has a solution.

2$^{\boldsymbol{o}}$. To show uniqueness, assume that ϕ and $\widetilde{\phi}$ are two solutions of (1.13), and define the "residual"

$$\eta(t) := \phi(t) - \widetilde{\phi}(t).$$

By linearity, η satisfies the following IVP

$$\eta' + p(t)\eta = 0, \quad \eta(t_0) = 0. \tag{1.14}$$

Here, the ODE is homogeneous and separable. It has a stationary solution $\eta \equiv 0$, which clearly solves (1.14). But separation of variables cannot give other solutions of (1.14), as it starts with division over $\eta \neq 0$. (It gives $\eta(t) = ce^{-\int_{t_0}^{t} p(s)ds}$ for $c \neq 0$.) Thus, $\eta(t) \equiv 0$ is the unique solution of (1.14). ☐

Example 1.4.2. *Find the largest interval I over which the solution to the following IVP is defined:*

$$(9 - t^2)y' + 5ty = 3t^2, \quad y(-1) = 1.$$

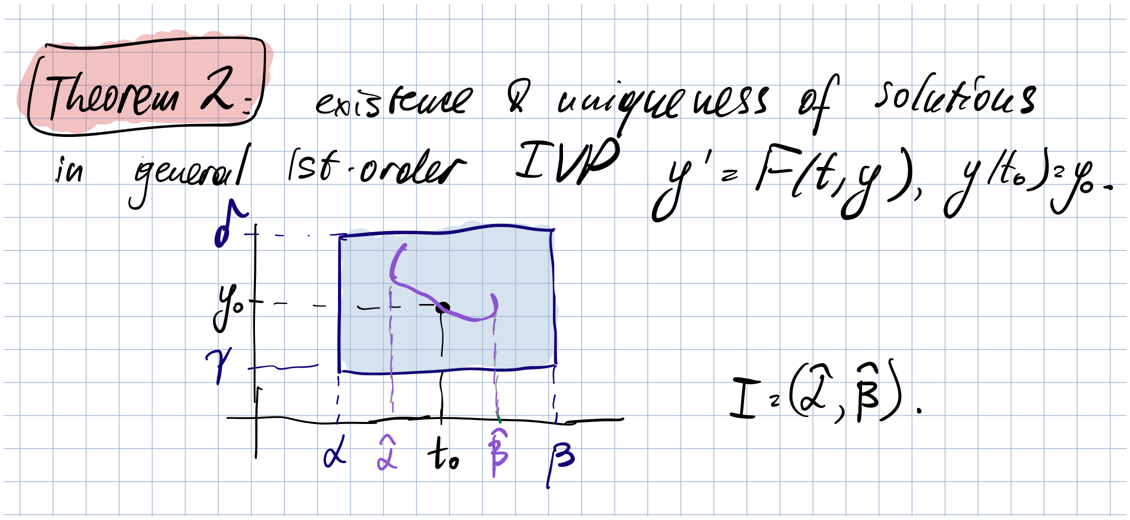

Figure 1.5: Illustration of the conditions in Theorem 1.4.2.

Dividing by $9-t^2$ we put the ODE in standard form $y'+p(t)y = g(t)$ with $p(t) = \frac{5t}{9-t^2}$, $g(t) = \frac{3t^2}{9-t^2}$. Each of the functions p, g has singularities at ± 3 and nowhere else. We can apply Theorem 3.1 with any open interval containing $t_0 = -1$ but neither of the points ± 3. The largest such interval is $(-3, 3)$, which is the answer. Note that we did not have to *solve* the IVP to reach this conclusion. Nor did we use the value $y_0 = 1$ in the IV condition: the answer would not change for any other y_0. These two observations fully agree with Theorem 1.4.1 and generalize for any linear IVP. ☐

Theorem 1.4.2 (Existence & uniqueness of solutions for general first-order IVPs; Fig. 1.5)**.** *Assume that functions F, $\frac{\partial}{\partial y}F$ are continuous in some open rectangle $R = (\alpha, \beta) \times (\gamma, \delta)$ that contains $(t_0; y_0)$. Then the IVP*

$$y' = F(t, y), \quad y(t_0) = y_0$$

has a unique solution on some interval $(\hat{\alpha}, \hat{\beta}) \subseteq (\alpha, \beta)$ that contains t_0. Moreover, in the case where $(\gamma, \delta) = (-\infty, +\infty)$, one can take $\hat{\alpha} = \alpha$ and $\hat{\beta} = \beta$. (I didn't mention the last part in class.)

Theorem (1.4.2) can be proved via the implicit function theorem from calculus; we omit this proof.

- The condition in the premise of the theorem is sufficient, but by no means necessary: there are some nonlinear IVPs where the condition is *not* satisfied, yet solutions exist and are unique.
- Note that in Theorem 1.4.2, existence and uniqueness of solution is not guaranteed on the whole interval (α, β), only on some potentially smaller interval $(\hat{\alpha}, \hat{\beta})$.
- Note that Theorem 1.4.2 includes Theorem 1.4.1 as a special case. Indeed, linear IVP (1.13) is a general one with $F(t, y) = -p(t)y + g(t)$ and $\frac{\partial}{\partial y}F(t, y) = -p(t)$. Assuming that both p and g are continous on the open interval (α, β) that contains t_0, the two functions $\frac{\partial}{\partial y}F$ and F are continuous in the open rectangle $(\alpha, \beta) \times (-\infty, +\infty)$ that contains $(t_0; y_0)$.[5] Thus,

[5] The continuity of $\frac{\partial}{\partial y}F$ is obvious. Meanwhile, in order to conclude that F is continuous, we treated $p(t)$ and y as continuous functions in two variables, and used the result about the continuity of the product of continuous functions.

Theorem 1.4.2 implies that any linear IVP (1.13) with continuous p, g has a unique solution on the interval containing t_0 and in which p, g are continuous. In particular, existence/uniqueness of solutions *does not* depend on y_0 for a linear IVP, while for nonlinear IVP it generally *does* (Example 1.4.1).

Now, let us revisit Example 1.4.1 and see why the sufficient condition of Theorem 1.4.2 is not satisfied. Here $F = 8ty^{1/5}$ is continuous on $\mathbb{R}^2$, yet $\frac{\partial}{\partial y}F = \frac{8}{5}ty^{-4/5}$ is undefined on the straight line $y = 0$ in $\mathbb{R}^2$. As such, $\frac{\partial}{\partial y}F$ cannot be continuous on any rectangle $R = (\alpha, \beta) \times (\gamma, \delta)$ crossing this line, i.e. such that $\gamma < 0 < \delta$, and only such rectangles might contain $(0, 0)$. Thus, Theorem 1.4.2 does not allow to make a conclusion about the IVP in Example 1.4.1. However, if we change the IV condition from $y(0) = 0$ to $y(0) = 1$, then the new IVP will have a unique solution by Theorem 1.4.2.

Example 1.4.3 (cf. Example 1.1.2)**.** *Recall Example 1.1.2. Here F and $\frac{\partial}{\partial y}F$ have a singularity at $y = 1$:*

$$F = \frac{3t^2 + 4t + 2}{2(y-1)}, \qquad \frac{\partial}{\partial y}F = -\frac{3t^2 + 4t + 2}{2(y-1)^2}.$$

Both these functions are continuous on any rectangle that does not cross the horizontal line $y = 1$ (touching the line on a boundary is allowed, since the rectangle is open). Since one can draw such a rectangle around the point $(t_0; y_0) = (0; -1)$, e.g. $(-\infty, +\infty) \times (-\infty, 0)$, it follows that the IVP has a unique solution in some *interval $(\hat{\alpha}, \hat{\beta})$ containing $t_0 = 0$. However, such rectangles cannot be with $(\gamma, \delta) = (-\infty, +\infty)$, since they are not allowed to cross $y = 1$. Hence, we cannot conclude that $(\hat{\alpha}, \hat{\beta}) = (-\infty, +\infty)$, i.e. that the solution exists on $\mathbb{R}$. Instead, we must solve the IVP and examine its solution. And indeed, in Example 1.1.2 we found that the IVP solution exists on $(-2, +\infty)$.*

The last example was not covered. We reiterate that, while for a linear IVP without singularities in p and g the (unique) solution exists on the *whole* real axis, solution of a *nonlinear* IVP might exist only on a *smaller interval* even when there are no singularities in F, F'_y. This is illustrated below.

Example 1.4.4 ([BB15, Example 2.4.4])**.** *Consider the nonlinear IVP*

$$y' = y^2, \quad y(0) = 1.$$

Solving it via integrating factor and fitting the IV condition, we find the solution, existing on $(-\infty, 1)$:

$$y(t) = \frac{1}{1-t}, \quad t < 1.$$

Note that, without solving the IVP, it is impossible to guess from the equation itself that the point $t = 1$ is in any way special: we really have to fit the initial value condition $y(0) = 1$ to discover this. Moreover, if we change $y(0) = 1$ to $y(0) = y_0 > 0$, then the solution $y(t) = \frac{y_0}{1-y_0t}$ exists on $(0, \frac{1}{y_0})$.

1.5 Autonomous equations

Recall that *autonomous equations* are 1st-order ODEs of the form

$$y' = f(y), \tag{1.15}$$

in other words, those where the rate function $F(t, y) = f(y)$. Previously, we have seen how to find and classify *critical points*, or *equilibria* of such equations, i.e. the zeroes of f. In this lecture, we shall learn to extract some additional information on the shape of solution curves near equilibria.

Convexity/concavity. A twice differentiable function y is called (locally) *convex* at t if $y''(t) > 0$, and (locally) *concave* at t if $y''(t) < 0$. If $y''(t) = 0$, then such t is called an *inflection point* of y. Geometrically, the sign of $y''(t)$ corresponds to the direction of the curvature of the graph of y at t:

- If $y''(t) > 0$, the graph is curved upwards (like $\cup$) at t;
- If $y''(t) < 0$, the graph is curved downwards (like $\cap$) at t;
- If $y''(t) = 0$, the direction of curvature changes at t.

This is shown in Fig. 1.5.

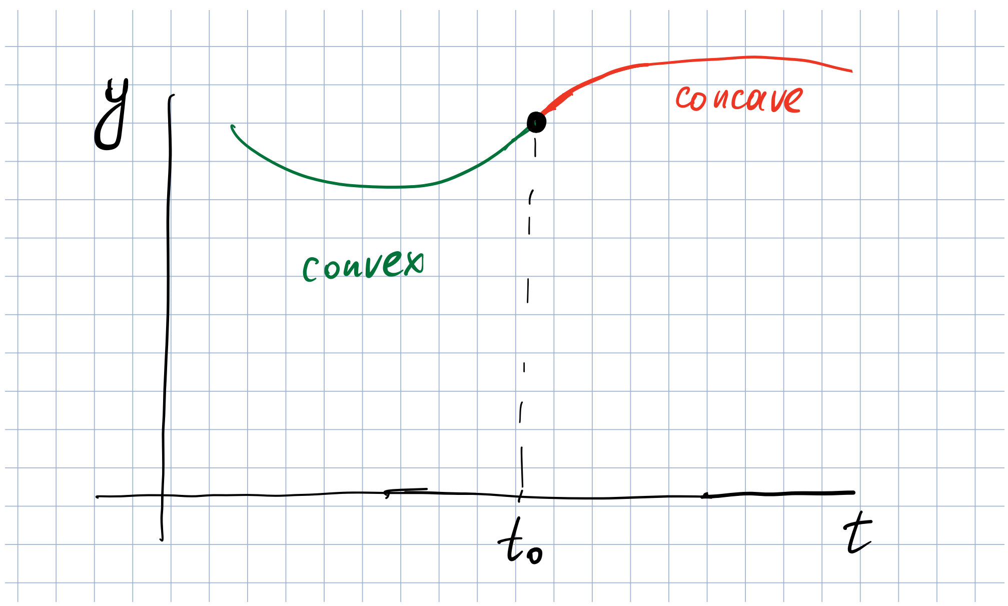

Figure 1.6: Function $y(t)$ is locally convex at $t < t_0$, concave at $t > t_0$, and has inflection point at t_0.

For example:
- e^t and t^2 are convex functions on $\mathbb{R}$;
- function $\ln(t)$ is concave on its domain $t > 0$;
- each of the functions t^3, $\frac{1}{1+e^t}$ is convex at $t > 0$, concave at $t < 0$, has inflection point at $t = 0$.

Curvature of solutions. Going back to ODEs, let $y(t)$ be a particular solution to (1.15). By the chain rule,

$$y'' = f'(y)y' = f'(y)f(y).$$

Thus, we can determine the direction of curvature of a given solution curve at t by only studying f.

> *Any nonstationary solution $y(\cdot)$ of an ODE (1.15) is locally convex at a given t when the signs of $f(y(t))$ and $f'(y(t))$ are the same, and concave when these signs are different.*

Note also that a nonstationary solution can also have an inflection point at some t, since $y''(t)$ might vanish at t because of $f'(y(t)) = 0$, even though $f(y(t)) \neq 0$.

The above observation allows to understand the shape of solution curves close to critical points.

Example 1.5.1 (Logistic equation)**.** *Revisiting the second part of Example 1.3.2, consider the equation*

$$y' = ry(1 - y).$$

Here $f(y) = ry(1-y)$ has zeroes at 0 and 1, which gives two stationary solutions: $y(t) = 0$ and $y = 1$.

1. *Computing $f'(y) = r(1-2y)$, we see that $f'(0) > 0$ and $f'(1) < 0$, hence $y \equiv 1$ is asymptotically stable and $y \equiv 0$ is unstable. (See Fig. 1.5)*

2. *Moreover, $f(y) > 0$ at $y \in (0, 1)$ and $f(y) < 0$ at $y < 0$ and $y > 1$. So, any solution passing through $y_0 \in (0, 1)$ increases, and any solution passing through $y_0 < 0$ or $y_0 > 1$ decreases.*

3. *Finally, for $\varepsilon > 0$ sufficiently small, solutions passing through $y_0 = 0 - \varepsilon$ are concave, those passing through $y_0 = 1 + \varepsilon$ are convex. Moreover, a solution passing through $y_0 = 0 + \varepsilon$ and $y_0 = 1 - \varepsilon$ must have an inflection point at some t_1, be locally convex at $t < t_1$ and locally concave at $t > t_1$. Intuitively, all this should be clear by noting that the first kind of solutions run away from 0, the second kind must approach 1, and the third kind run away from 0 to 1.*

    - *But now we see it directly. For example, at $y_0 = 0-\varepsilon$ we have $f(y_0) < 0$ and $f'(y_0) > 0$.*[6] *Thus, $y'' < 0$ in this case. The reasoning is similar in other cases ($y_0 = 0+\varepsilon$, $y_0 = 1 \pm \varepsilon$).*

[6] The 2nd inequality is since $f'(0) < 0$ and f' is continuous, but we can also get it directly from the formula for f'.

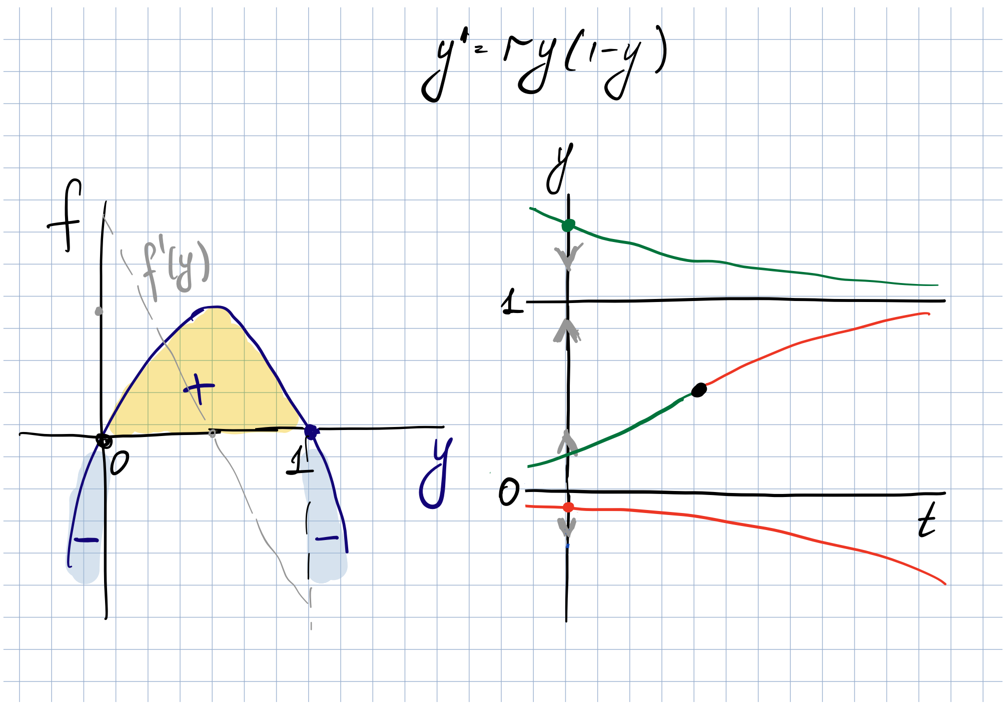

Figure 1.7: Solution curves for the ODE in Example 1.5.1.

Chapter 2

First-order linear systems

Preliminaries

Linear ODEs of order n. Recall that a **linear** ODE of order $n > 1$, in the standard form, is

$$u^{(n)} + a_1(t)u^{(n-1)} + ... + a_{n-1}(t)u'(t) + a_n(t)u = g(t) \tag{2.1}$$

where $t \in \mathbb{R}$ is the independent variable; $a_1, ..., a_n, g$ are some functions that specify the equation. In particular, $a_1, ..., a_n$ are called the *coefficient functions*; if they happen to be constants, we say that the corresponding (linear) ODE is with *constant coefficients.* As before, *a particular solution* of (3.2) on some interval $I \subseteq \mathbb{R}$ is any n times differentiable function $u : I \to \mathbb{R}$ that satisfies (3.2) for all t in I, and *the general solution* is the set of all particular solutions. For IVPs, order $n > 1$ requires n conditions to specify a solution. *By definition,* initial conditions for (3.2) are of the form

$$\begin{aligned} u(t_0) &= u_0, \\ u'(t_0) &= u_1, \\ &\vdots \\ u^{(n-1)}(t_0) &= u_{n-1} \end{aligned} \tag{2.2}$$

for some values $u_0, ..., u_{n-1} \in \mathbb{R}$. Note that, compared to the first-order case, we added $n - 1$ additional conditions corresponding to the values of higher derivatives of u, from u' and up to $u^{(n-1)}$.

The theory of linear ODEs of order n can be "handled" by extending the theory of first-order linear ODEs in a rather straightforward fashion. In a nutshell, this is because one may convert such an ODE into a *first-order linear system* of n differential equations through a "vectorization trick," trading the order for dimension. Thus, studying (3.2) reduces to studying first-order systems of DEs. In this chapter, we shall learn some general properties of first-order linear systems of DEs (and nth-order linear ODEs), including the question of solution existence/uniqueness for IVPs. For linear systems of n equations with *constant coefficients*, we obtain general solutions in closed form. Finally, we shall provide a full classification of critical points in the two-dimensional case (i.e. $n = 2$).

2.1 Refresher of linear algebra

Let A be a square $n \times n$ matrix with real entries (which will be denoted as $A \in \mathbb{R}^{n\times n}$ from now on):

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Systems of linear equations. Recall that a system of n linear equations in n variables $x_1, ..., x_n$,

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + \cdots + a_{2n}x_n &= b_2, \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= b_n, \end{aligned} \tag{2.3}$$

can be written as a single vector equation (i.e. in a vector form):

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}, \quad \text{or} \quad A\vec{x} = \vec{b}, \tag{2.4}$$

where $\vec{x} \in \mathbb{R}^n$ is the unknown vector, and $\vec{b} \in \mathbb{R}^n$ is known. (For convenience, all vectors are column-vectors, i.e. $n \times 1$ matrices.) The $n \times (n+1)$ matrix $(A|\vec{b})$ is called the *augmented* matrix of (2.4) Geometrically, kth equation in (2.3) (kth row in (2.4)) specifies a hyperplane with normal vector

$$\begin{pmatrix} a_{k1} & \cdots & a_{kn} \end{pmatrix}^\top = \begin{bmatrix} a_{k1} \\ \vdots \\ a_{kn} \end{bmatrix}$$

and passing trough the origin $\Leftrightarrow b_k = 0$. Solving (2.4) corresponds to finding each point where all these n hyperplanes intersect. A linear system might have a unique solution, no solutions, or an infinite number (continuum) of solutions, with no other possibilities. If $\vec{b} = \vec{0}$, the system is called *homogeneous*, and always has a *trivial solution* $\vec{x} = \vec{0}$. A system with $\vec{b} \neq \vec{0}$ is called *nonhomogeneous*.

For any $\vec{b}$, system $A\vec{x} = \vec{b}$ with a nonsingular $A \in \mathbb{R}^{n\times n}$ has a *unique solution* $A^{-1}\vec{b}$.

Thus, we get a plethora of equivalent conditions for $n \times n$ system $A\vec{x} = \vec{b}$ to have a unique solution:

- $\text{rank}(A) = n$, i.e. A is full-rank;
- $\det(A) \neq 0$;
- $\mathsf{Null}(A) = \{\vec{0}\}$;
- $\mathsf{rref}(A) = I$;
- A has a pivot in each row and a pivot in each column.

To solve the system we can invert A and multiply the result by $\vec{b}$ on the right. Alternatively, we can run Gaussian elimination on the rows of the augmented matrix $(A|\vec{b})$, obtaining the matrix $(I|A^{-1}\vec{b})$.

Case $n = 2$. Clearly, $A = (\vec{u}|\vec{v}) \in \mathbb{R}^{2\times 2}$ is singular $\Leftrightarrow$ $\vec{u}, \vec{v}$ are parallel vectors, i.e. either $\vec{u} = c\vec{v}$ or $\vec{v} = c\vec{u}$ for some $c \in \mathbb{R}$, or both. ($\vec{0}$ is parallel to any vector). Hyperplanes here are straight lines.

Example 2.1.1. *Solve the linear system and determine if the lines intersect, are parallel, or coincide:*

$$\begin{aligned} x_1 - 2x_2 &= -1, \\ -x_1 + 3x_2 &= \ \ 3. \end{aligned}$$

Find normal vectors for the lines.

We run Gaussian elimination on the augmented matrix – first top-to-bottom, then bottom-to-top:

$$(A|\vec{b}) = \left(\begin{array}{cc|c} 1 & -2 & -1 \\ -1 & 3 & 3 \end{array}\right) \underset{R_2 \leftarrow R_2 + R_1}{\sim} \left(\begin{array}{cc|c} 1 & -2 & -1 \\ 0 & 1 & 2 \end{array}\right) \underset{R_1 \leftarrow R_1 + 2R_2}{\sim} \left(\begin{array}{cc|c} 1 & 0 & 3 \\ 0 & 1 & 2 \end{array}\right).$$

The unique solution (and intersection point) is $(x_1; x_2) = (3; 2)$. Normal vectors are, respectively,

$$\begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} -1 \\ 3 \end{bmatrix}.$$

□

Eigenvalues. Recall that *eigenvalues* of $A \in \mathbb{R}^{n\times n}$ are $\lambda \in \mathbb{C}$ such that $A - \lambda I$ is singular, that is

$$\det(A - \lambda I) = 0, \tag{2.5}$$

where I is the identity matrix of appropriate size. Eq. (2.5) is called *the characteristic equation* for A. Using the properties of determinant, one might show that $p_A(\lambda) := \det(A - \lambda I)$ is a polynomial in λ of degree n, called *the characteristic polynomial* of A. Thus, eigenvalues are the roots of the characteristic polynomial. By the fundamental theorem of algebra, any polynomial of degree n has n complex roots counted with multiplicities. Since the entries of A are assumed real, the coefficients of p_A must be real, yet its roots might be complex. Yet, there are some restrictions on this:

- If A is real, and $\lambda = \alpha + i\beta$ is an eigenvalue ($\alpha, \beta \in \mathbb{R}$), then $\bar{\lambda} = \alpha - i\beta$ is also an eigenvalue.
- If A is symmetric (i.e. $A = A^\top$), then its eigenvalues are real.

This is only a sufficient condition: indeed, the eigenvalues of $\begin{pmatrix} \lambda_1 & 0 \\ a & \lambda_2 \end{pmatrix}$ are λ_1, λ_2 – regardless of a.

Eigenvectors and eigenspaces. Let $\lambda \in \mathbb{C}$ be an eigenvalue of A, then the subspace $\mathsf{Null}(A - \lambda I)$ of $\mathbb{C}^n$ must have a positive dimension, i.e. contain a nonzero vector. This subspace is called an *eigenspace of A corresponding to λ*. Any $\vec{v} \in \mathsf{Null}(A - \lambda I)$ such that $\vec{v} \neq \vec{0}$ is called an *eigenvector of A corresponding to λ*. Important results on the dimensionality of eigenspaces are listed below.

- The dimension of eigenspace corresponding to λ is **at most** the algebraic multiplicity of this λ. In particular, if all eigenvalues of A are distinct, then all its eigenspaces are one-dimensional.
- Eigenvectors $\vec{v}_1, \vec{v}_2$ corresponding to different eigenvalues $\lambda_1 \neq \lambda_2$ are linearly independent. More generally, $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ corresponding to $\lambda_1 \neq \lambda_2 \neq ... \neq \lambda_k$ are linearly independent. In particular, if all eigenvalues of A are distinct, then $\vec{v}_1, ..., \vec{v}_n$ form a basis that diagonalizes A:

$$A = V^{-1}\Lambda V \quad \text{where} \quad \Lambda = \mathrm{Diag}(\lambda_1, ..., \lambda_n) \quad \text{and} \quad V = \big(\vec{v}_1 | \cdots | \vec{v}_n\big)$$

The eigenvectors for real eigenvalues are real; for complex eigenvalues they are complex. Moreover:

- If A is symmetric ($A = A^\top$), then its eigenvectors $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^n$ corresponding to $\lambda_1 \neq \lambda_2$ are mutually orthogonal, i.e. $\vec{v}_1^\top \vec{v}_2 = 0$.
- If A is real, and $\vec{u} + i\vec{v}$ is an eigenvector corresponding to $\lambda = \alpha + i\beta$, then its conjugate $\vec{u} - i\vec{v}$ is an eigenvector corresponding to $\bar{\lambda} = \alpha - i\beta$.

Example 2.1.2. *Find the eigenvalues and eigenvectors of* (a) $A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}$ *and* (b) $B = \begin{pmatrix} -5 & -5 \\ 5 & -5 \end{pmatrix}$.

(a). Here $p_A(\lambda) = (5-\lambda)(1-\lambda) - 4 = \lambda^2 - 6\lambda + 1$, so the eigenvalues are $\lambda_{1,2} = \frac{6 \pm \sqrt{32}}{2} = 3 \pm 2\sqrt{2}$. (They *must* be real, since A is symmetric!) They are distinct, so each eigenspace has dimension 1. It suffices to find an eigenvector $\vec{v}_1$ for $\lambda_1 = 3 + 2\sqrt{2}$, then any $\vec{v}_2 \perp \vec{v}_1$ must be an eigenvector for $\vec{v}_2$. Now: to find $\vec{v}_1$, it suffices compute only the first row of $A - \lambda_1 I$, i.e. we do not need the grey stuff:

$$A - \lambda_1 I = \begin{pmatrix} 2 - 2\sqrt{2} & 2 \\ 2 & -2 - 2\sqrt{2} \end{pmatrix} = 2 \begin{pmatrix} 1 - \sqrt{2} & 1 \\ 1 & -1 - \sqrt{2} \end{pmatrix}.$$

Note that the second row is proportional to the first one. (It must be: $A - \lambda_1 I$ is singular.) We can take for $\vec{v}_1$ arbitrary vector orthogonal to the first row, e.g. $\vec{v}_1 = \begin{pmatrix} -1 \\ 1 - \sqrt{2} \end{pmatrix}$; then take $\vec{v}_2 = \begin{pmatrix} 1 - \sqrt{2} \\ 1 \end{pmatrix}$.

(b). Note that

$$B = 5M \quad \text{where} \quad M = \begin{pmatrix} -1 & -1 \\ 1 & -1 \end{pmatrix}.$$

Thus, the eigenvalues of B are $\lambda = 5\mu$ where μ are those of M, i.e. the roots of $p_M(\mu) = (-1-\mu)^2 + 1$. These are $\mu_{1,2} = -1 \pm i$; note that they mutual conjugates, and the corresponding eigenvectors must be mutual conjugates as well. For $\mu_1 = -1 + i$, we find

$$M - \mu_1 I = \begin{pmatrix} -1 - (-1 + i) & -1 \\ * & * \end{pmatrix} = \begin{pmatrix} -i & -1 \\ * & * \end{pmatrix}.$$

We can take $\vec{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$ as it is orthogonal to the first row, and its conjugate $\vec{v}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$ for λ_2. $\square$

Eigenvalue invariants. Recall that for a square matrix A, the determinant $\det(A)$ and the trace,

$$\operatorname{tr}(A) := \sum_{k=1}^{n} a_{kk},$$

can be expressed in terms of the eigenvalues of A, namely $\det(A) = \prod_{k=1}^{n} \lambda_k$ and $\operatorname{tr}(A) = \sum_{k=1}^{n} \lambda_k$. In fact, $\det(A)$ and $\operatorname{tr}(A)$ are, up to a sign, just two coefficients of the characteristic polynomial p_A, and the whole polynomial—i.e. all its n coefficients—can be expressed in terms of eigenvalues only. But even using just $\det(A)$ and $\operatorname{tr}(A)$, one might save a lot of calculations when finding eigenvalues.

Example 2.1.3 (HW3 #3). *Find all eigenvalues of the following matrix*

$$A = \begin{pmatrix} 15 & 4 & -24 \\ 16 & 3 & -24 \\ 12 & 4 & -21 \end{pmatrix}.$$

We will not compute all 3 coefficients of $p_A(\lambda)$. Instead, we first guess one eigenvalue $\lambda_1 = -1$ by noting that the first two rows of A are "almost identical." We guess one more eigenvalue, $\lambda_2 = 3$, by comparing the 1st and 3rd rows. Finally, we find the trace, $\operatorname{tr}(A) = 15 + 3 - 21 = -3$, and the last eigenvalue: $\lambda_3 = \operatorname{tr}(A) - \lambda_1 - \lambda_2 = -3 + 1 - 3 = -5$, and are done. Some variations are possible: for example, we could use $\det(A)$ instead of $\operatorname{tr}(A)$ in the last step (it's longer to compute, though). Or, we could guess just one eigenvalue, then two others from $\det(A)$ and $\operatorname{tr}(A)$, by solving a system of 2 equations in 2 variables. (Even this method is much faster than computing and factorizing p_A.) $\square$

2.2 First-order linear systems of DEs with variable coefficients

2.2.1 Matrix-functions

A matrix-function maps $t \in \mathbb{R}$ into a matrix $M = M(t) \in \mathbb{R}^{n\times m}$, possibly with $m \neq n$. Equivalently, the entries of $M(t)$ are usual functions. You encountered vector-functions ($m = 1$) in vector calculus:

$$\vec{a}(t) = \begin{pmatrix} a_1(t) \\ \vdots \\ a_n(t) \end{pmatrix}.$$

Now we shall also need that of a square matrix:

$$A(t) = \begin{pmatrix} a_{11}(t) & \cdots & a_{1n}(t) \\ \vdots & \ddots & \vdots \\ a_{n1}(t) & \cdots & a_{nn}(t) \end{pmatrix}.$$

A matrix function $A(t)$ is called continuous at t if all its entries $a_{ij}(\cdot)$, $1 \leqslant i, j \leqslant n$ are continuous at t. We can take derivatives and definite integrals of such functions. This is done entrywise: *by definition,*

$$\frac{\partial}{\partial t} A(t) := \begin{pmatrix} a'_{11}(t) & \cdots & a'_{1n}(t) \\ \vdots & \ddots & \vdots \\ a'_{n1}(t) & \cdots & a'_{nn}(t) \end{pmatrix} \quad \text{and} \quad \int_\alpha^\beta A(t)dt := \begin{pmatrix} \int_\alpha^\beta a_{11}(t)dt & \cdots & \int_\alpha^\beta a_{1n}(t)dt \\ \vdots & \ddots & \vdots \\ \int_\alpha^\beta a_{n1}(t)dt & \cdots & \int_\alpha^\beta a_{nn}(t)dt \end{pmatrix}.$$

For example, for $A(t) = \begin{pmatrix} 0 & 1 \\ t & t^2 \end{pmatrix}$ we find

$$\frac{\partial}{\partial t} A(t) = \begin{pmatrix} 0 & 1 \\ 1 & 2t \end{pmatrix} \quad \text{and} \quad \int_0^1 A(t)dt = \begin{pmatrix} \int_0^1 cdt & \int_0^1 tdt \\ \int_0^1 \frac{t^2}{2}dt & \int_0^1 \frac{t^3}{3}dt \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}.$$

Note that both these operations are linear: for any constants c_1, c_2 and functions $A(t), B(t) \in \mathbb{R}^{n\times n}$,

$$\frac{\partial}{\partial t}(c_1 A(t) + c_2 B(t)) = c_1 \frac{\partial}{\partial t} A(t) + c_2 \frac{\partial}{\partial t} B(t),$$
$$\int_\alpha^\beta (c_1 A(t) + c_2 B(t))dt = c_1 \int_\alpha^\beta A(t)dt + c_2 \int_\alpha^\beta B(t)dt.$$

2.2.2 Terminology

A general first-order linear system of DEs of dimension n (with variable coefficients) is of the form

$$\vec{x}' = A(t)\vec{x} + \vec{g}(t) \tag{2.6}$$

where $A(t) \in \mathbb{R}^{n\times n}$ and $g(t) \in \mathbb{R}^n$ are some known matrix-functions, and $\vec{x} \in \mathbb{R}^n$ is the unknown (or dependent) vector. Note that this is an "ordinary" system: there is a single *independent* variable. The system is called *homogeneous* if $g(t) \equiv 0$, and *nohomogeneous* otherwise. Let's define solutions.

Definition 2. *A (particular) solution of* (3.11) *on* $I \subseteq \mathbb{R}$ *is a vector-function* $\vec{x}(t)$ *satisfying* (3.11) *on* I*.*

Definition 3. *The general solution of* (3.11) *is the set of all particular solutions.*

Systems with *constant* coefficients, i.e. with $A(t) = \text{const}$, can be solved by linear algebra. We shall learn how to it in this chapter. For variable coefficients, the methods are more advanced, generalizing the method of integrating factors studied in the previous chapter. We study them later.

Example 2.2.1 (Foxes and rabbits)**.** *Foxes and rabbits live on an island. Their respective numbers* $x_t(1)$*,* $x_2(t)$ *at day* t *is described by the following linear system of DEs with constant coefficients:*

$$\begin{aligned} x_1' &= ax_1 + bx_2 - r, \\ x_2' &= cx_1 + dx_2 \end{aligned}$$

with parameters $c < 0$ *and* $a, b, d, r > 0$*. Interpret this system and write it in a matrix form.*

The matrix form is

$$\vec{x}' = Ax + \vec{g} \quad \text{with} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \vec{g} = \begin{pmatrix} -r \\ 0 \end{pmatrix}.$$

Interpretation: the increase of foxes at day t is the sum of *two* terms: the term proportional to the current number of foxes, plus the one proportional to the number of rabbits. Additionally, r foxes per day are removed from the island. The number of rabbits grows with the number of rabbits, but also decreases proportionally to the number of foxes. Rabbits are neither removed nor imported. □

For autonomous linear systems (i.e., with constant $A, \vec{g}$), we can adapt the notion of critical points.

Definition 4. *A* critical point *for linear system of DEs* $\vec{x}' = A\vec{x} + \vec{g}$ *with constant* A *and* $\vec{g}$ *is a solution to the system of linear equations* $A\vec{x} = -\vec{g}$*.*

For a 1st-order linear system of dimension n, the general solution will typically have n degrees of freedom—arbitrary constants $c_1, ..., c_n$—instead of just one, as it was for first-order ODEs. Equivalently, one can say that solution is defined up to an arbitrary *vector* of constants $\vec{c} \in \mathbb{R}^n$. Intuitively, this is because such a system corresponds to an nth-order ODEs, to solve which one has to "integrate n times." Accordingly, an IVP for a linear system (3.11) has initial condition of the form

$$\vec{x}(t_0) = \vec{x}_0 \tag{2.7}$$

for some initial value vector $\vec{x}_0 \in \mathbb{R}^n$.

Next, we show how to convert nth-order linear ODE (3.2) into a first-order linear system of DEs.[1]

[1] There is a method of conversion in the opposite direction as well. Try to come up with it after reading Section 3.3.3.

2.2.3 Vectorization trick

Example 2.2.2. *Convert a second-order ODE into an equivalent 1st-order linear system of DEs:*

$$u'' - u' \sin t + 7u = e^t \cos t + 1. \tag{2.8}$$

We rewrite the ODE in the standard form: $u'' = u' \sin t - 7u + e^t \cos t + 1$ and define variables

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

In terms of these variables,

$$\begin{aligned} x_1' = u' &= x_2, \\ x_2' = u'' &= u' \sin t - 7u + e^t \cos t + 1 \quad = \quad -7x_1 + x_2 \sin t + e^t \cos t + 1. \end{aligned}$$

where for x_2' we first plugged in the ODE. That is, the equivalent system reads

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -7 & \sin t \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ e^t \cos t + 1 \end{pmatrix}. \tag{2.9}$$

The first row in the matrix is $(0|1)$, and the entries of the last row are the coefficients of the ODE in the standard form, *in the reverse order.* □

Exercise 2.2.1. *Verify that ODE* (3.13) *and system* (3.14) *are indeed equivalent, in the following sense:*

(a) If $u(t)$ satisfies (3.13)*, then* $\vec{x}(t) = \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}$ *satisfies* (3.14)*.*

(b) Conversely, for any solution $\vec{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$ *of* (3.14)*, the first component $x_1(t)$ satisfies* (3.13)*.*

General method. We now explain the general method for nth-order ODE in a standard form:

$$u^{(n)} = a_1(t)u^{(n-1)} + \cdots + a_{n-1}(t)u' + a_n(t)u + b(t).$$

We proceed as follows:

$$\begin{array}{rcl} x_1 = u & & x_1' = u' = x_2 \\ x_2 = u' & & x_2' = u'' = x_3 \\ \vdots & \implies & \vdots \\ x_{n-1} = u^{(n-2)} & & x_{n-1}' = u^{(n-1)} = x_n \\ x_n = u^{(n-1)}. & & x_n' = u^{(n)} = a_n(t)x_1 + a_{n-1}(t)x_2 + \cdots + a_1(t)x_n + b(t). \end{array}$$

As the result, we obtain the system $\vec{x}' = A(t)\vec{x} + \vec{g}(t)$ with

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ a_n(t) & \cdots & \cdots & a_2(t) & a_1(t) \end{pmatrix} \quad \text{and} \quad \vec{g}(t) = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

Here $\vec{g}(t)$ has zeroes in the first $n-1$ rows and $b(t)$ in the last row. In the first $n-1$ rows of $A(t)$, 1 moves from the 2nd to the last position; its last row has the ODE coefficients in the reverse order.

2.2.4 Existence and uniqueness of solutions in linear IVPs

We give a generalization of Theorem 1 (IVP with a $1st$-order linear ODE) from the previous chapter.

Theorem 2.2.1. *Assume that $A(\cdot) \in \mathbb{R}^{n\times n}$ and $\vec{g}(\cdot) \in \mathbb{R}^n$ are continuous in some interval $I = (\alpha, \beta)$ containing t_0. Then the IVP $\vec{x}' = A(t)\vec{x} + \vec{g}(t)$, $\vec{x}(t_0) = \vec{x}_0$ has a unique solution for any $\vec{x}_0 \in \mathbb{R}^n$.*

We omit the proof. Conceptually, it is similar to that of Theorem 1, but uses vector calculus. The theorem might also be used for n-order ODEs, by first converting them to an equivalent system.

Example 2.2.3. *Identify the largest interval on which a solution exists and is unique, for the IVP*

$$(t-2)u'' + 3u = t, \quad u(0) = 0, \quad u'(0) = 1.$$

Dividing by $t-2$, we put equation in the standard form:

$$u'' = -\frac{3}{t-2}u + \frac{t}{t-2}.$$

Then we obtain an equivalent system:

$$\vec{x}' = \begin{pmatrix} 0 & 1 \\ -\frac{3}{t-2} & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ \frac{t}{t-2} \end{pmatrix}.$$

Here both $A(t)$ and $\vec{g}(t)$ are continuous on $(-\infty, 2)$ and $(2, +\infty)$. The point t_0 belongs to the first interval, so the answer is $(-\infty, 2)$ by Theorem 3.3.2. □

In fact, we do not even have to convert an ODE to the system: clearly, the equivalent condition for ODE is that all coefficient functions $a_1(t), ..., a_n(t)$ and $b(t)$ are continuous on $I = (\alpha, \beta) \ni t_0$. Note also that it might be the intervals of continuity for some or all of these functions are distinct, not the same as in the previous example. Then we take all possible intersections of these intervals.

Example 2.2.4. *Identify the largest interval where the following IVP has a unique solution:*

$$u'' + \frac{1}{t-1}u = \frac{t}{t-3}, \quad u(2) = 0, \quad u'(2) = 1.$$

The coefficient $a_1(t) \equiv 0$ of u' is continuous everywhere; $a_2(t) = -\frac{1}{t-1}$ is continuous on $(-\infty, 1) \cup (1, +\infty)$; finally, $b(t) = \frac{t}{t-3}$ is continuous on $(-\infty, 3) \cup (3, +\infty)$. Hence, we have 3 candidate intervals $(-\infty, 1)$, $(1, 3)$, $(3, +\infty)$. The second interval contains $t_0 = 2$, thus is the one we seek. □

2.2.5 Linear independence of functions, Wronskian, fundamental set of solutions

Linear independence of functions. Given n functions $u_1, ..., u_n$ on some interval $I \subseteq \mathbb{R}$, we call them *linearly dependent* if there exist a tuple of constants $c_1, ..., c_n$, not all zeroes, such that

$$c_1 u_1(t) + \cdots + c_n u_n(t) = 0 \quad \forall t \in I.$$

In other words, $u_1, ..., u_n$ are dependent (on I) if some nontrivial linear combination of them is identically zero on I. Otherwise, $u_1, ..., u_n$ are *linearly independent* on I; equivalently, any nontrivial combination $c_1 u_1(t) + \cdots + c_n(t)$ is nonzero in at least some $t \in I$. To make sense of these definitions, one might view functions $I \to \mathbb{R}$ as *vectors* with continuous index $t \in I$; then the identical zero on I is the analogue of the zero vector.[2]

Example 2.2.5. *Functions t, t^2 are independent on $\mathbb{R}$. Indeed, assume that there are $c_1, c_2 \in \mathbb{R}$ such that $c_1 t + c_2 t^2 \equiv 0$ for all $t \in \mathbb{R}$. Plugging in $t_1 = 1$ and $t_2 = 2$ we get a homogeneous linear system*

$$\begin{aligned} c_1 + c_2 &= 0, \\ c_1 + 4c_2 &= 0 \end{aligned}$$

with a nonsingular matrix, hence with a unique solution $c_1 = c_2 = 0$. Thus, t, t^2 are independent. □

One can show independence for the sequence of monomials $\{1, t, t^2, ...\}$, exponentials $\{e^{\lambda t}, e^{2\lambda t}, ...\}$, trigonometric functions $\{\cos(\omega t), \cos(2\omega t), ...\}$ or $\{\sin(\omega t) \sin(2\omega t), ...\}$. This goes beyond our course.

Example 2.2.6. $\cos^2(t)$ *and* $\sin^2(t) - 1$ *are dependent (on any $I \subseteq \mathbb{R}$) since* $\cos^2(t) + \sin^2(t) - 1 \equiv 0$. □

Wronskian. Let $u_1, ..., u_n$ be $n - 1$ times differentiable on I. We can construct the square matrix

$$M(t) = \begin{pmatrix} u_1(t) & \cdots & u_n(t) \\ u_1'(t) & \cdots & u_n'(t) \\ & \vdots & \\ u_1^{(n-1)}(t) & \cdots & u_n^{(n-1)}(t) \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Its determinant $W(t) = \det(M(t))$, as a function of $t \in I$, is called *the Wronskian* of $u_1, ..., u_n$ (at t).

Example 2.2.7. *The Wronskian of $\{t, t^2\}$ is $W(t) = \begin{vmatrix} t & t^2 \\ 1 & 2t \end{vmatrix} = t^2$. Note that $W(t) = 0$ only at $t = 0$.*

Example 2.2.8. *For $\{\cos^2(t),\ \sin^2(t) - 1\}$, the Wronskian is $W(t) \equiv 0$ on $\mathbb{R}$. Indeed,*

$$\begin{aligned} W(t) = \begin{vmatrix} \cos^2(t) & \sin^2(t) - 1 \\ -2\cos(t)\sin(t) & 2\sin(t)\cos(t) \end{vmatrix} &= 2\sin(t)\cos^3(t) + 2\sin^3(t)\cos(t) - 2\sin(t)\cos(t) \\ &= 2\sin(t)\cos(t)\ [\cos^2(t) + \sin^2(t) - 1] \quad \equiv 0. \end{aligned}$$
□

The Wroskian gives a *necessary* condition of dependence.

[2] This abides the usual definition of linear independence, in the infinite-dimensional vector space of functions on I.

Theorem 2.2.2. *If $u_1, u_2, ..., u_n$ are linearly dependent on I, then $W(t) \equiv 0$ for all $t \in I$*

Proof. The derivative of a linear combination is the same linear combination of the derivatives. So if $c_1u_1(t) + \cdots + c_nu_n(t) \equiv 0$ on I, with $c_1, ..., c_n \in \mathbb{R}$ not all 0, then $c_1u_1^{(k)}(t) + \cdots + c_nu_n^{(k)}(t) \equiv 0$ on I. That is, the columns of $M(t)$ are linearly dependent for all $t \in I$. Thus, $W(t) \equiv 0$ on I. □

Theorem 3.3.3 allows to conclude that $u_1, ..., u_n$ are independent on I by finding some $t_0 \in I$ such that $W(t_0) \neq 0$. E.g. t, t^2 are independent on any open interval $I \in \mathbb{R}$ as follows from Example 3.3.7. The proof also shows why $W(t) \equiv 0$ is necessary for linear dependence, but might be *insufficient*. Indeed, the coefficients of a vanishing linear combination of the columns of $M(t)$ might depend on t, so there might not be any *constants* $c_1, ..., c_n$ that "work" for all $t \in I$ simultaneously. However, the condition actually becomes sufficient if $u_1, ..., u_n$ are *analytic* (infinitely many times differentiable).

Theorem 2.2.3 (Bôcher)**.** *If $u_1, u_2, ..., u_n$ are analytic and $W \equiv 0$ on I, then $u_1, ..., u_n$ are dependent.*

In this class, we only deal with functions analytic on their domains, so $W \equiv 0$ is a *criterion*.

Fundamental set of solutions in ODEs. If the coefficients $a_1, ..., a_n$ of a homogeneous ODE

$$u^{(n)} + a_1(t)u^{(n-1)} + \cdots + a_{n-1}(t)u' + a_n(t)u = 0 \tag{2.10}$$

are continuous on I, then corresponding IVP with $t_0 \in I$ has a unique solution. Yet, to find this solution we first have to obtain the general solution of the ODE. How do we know when we have it?

> *The set of solutions to a homogeneous linear ODE is a vector space: if $u(t)$ and $v(t)$ are solutions, then their linear combination $\alpha u(t) + \beta v(t)$ with $\alpha, \beta \in \mathbb{R}$ is also a solution.*

This vector space is n-dimensional.[3] Its arbitrary basis is called *a fundamental set of solutions* to (3.15). Therefore:

> *The general solution of a homogeneous linear ODE of order n on I is of the form*
>
> $$\sum_{k=1}^{n} c_k u_k(t), \quad t \in I,$$
>
> *where $u_1, ..., u_n$ are linearly independent solutions (a fundamental set of solutions) on I.*

If we found n linearly independent solutions, we are done. And if we have some candidates $u_1, ..., u_n$, their independence can be verified by Theorem 3.3.3: it suffices to find some $t_0 \in I$ such that $W(t_0) \neq 0$.

Fundamental set of solutions for systems of DEs. Consider an equivalent to (3.15) system:

$$\vec{x}' = A(t)\vec{x}. \tag{2.11}$$

[3]We are not proving this.

If $u_1, ..., u_n$ are solutions to (3.15), then the vector-functions $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n$ with

$$\vec{x}_k(t) = \begin{pmatrix} u_1(t) \\ u_1'(t) \\ \vdots \\ u_1^{(n-1)}(t) \end{pmatrix}$$

are solutions to (3.16), see Exercise 3.3.1. Therefore, $M(t)$ is the matrix with columns $\vec{x}_1(t), ..., \vec{x}_n(t)$, and $W(t) \neq 0$ if and only if the vectors $\vec{x}_1(t), ..., \vec{x}_n(t)$ are independent. This warrants a definition:

Definition 5. *Vector-functions $\vec{x}_1, ..., \vec{x}_n$ form a fundamental set of solutions to (3.16) on I if each of them satisfies (3.16) on I, and*

$$\det(\vec{x}_1(t_0) \cdots \vec{x}_n(t_0)) \neq 0 \;\; \textit{for some} \;\; t_0 \in I.$$

As in the case of ODEs, we conclude:

The general solution of a homogeneous 1st-order linear system of dimension n on I is

$$\sum_{k=1}^{n} c_k \vec{x}_k(t), \quad t \in I,$$

where $\vec{x}_1, ..., \vec{x}_k$ are linearly independent solutions (a fundamental set of solutions) on I.

Next, we learn how to find fundamental solutions for ODEs and systems with constant coefficients.

2.3 First-order linear systems with constant coefficients

In this section, we study homogeneous linear systems of dimension n *with constant coefficients*:

$$\vec{x}' = A\vec{x}, \tag{2.12}$$

as well as n-order ODEs that are reduced to such systems via vectorization. As it turns out, solving such systems reduces to linear algebra. We begin with a very simple but crucial observation:

If $\vec{v}$ is an eigenvector of A for eigenvalue $\lambda \in \mathbb{R}$, then $\vec{x}(t) = e^{\lambda t}\vec{v}$ is a solution to (3.17).

To see why this is the case, and explain where $e^{\lambda t}$ comes from, consider the trivial case $n = 1$. The system is then $x' = \lambda x$, with 1×1 matrix (λ). Solving it, we indeed get $x(t) = ce^{\lambda t}$ for $c \in \mathbb{R}$.

Now, in the general case $A\vec{x}(t) = Ae^{\lambda t}\vec{v} = e^{\lambda t}A\vec{v} = \lambda e^{\lambda t}\vec{v}$. But also, if v_k is the kth entry of $\vec{v}$,

$$\frac{d}{dt}\vec{x}(t) = \frac{d}{dt}\begin{pmatrix} e^{\lambda t}v_1 \\ \vdots \\ e^{\lambda t}v_n \end{pmatrix} = \begin{pmatrix} \lambda e^{\lambda t}v_1 \\ \vdots \\ \lambda e^{\lambda t}v_n \end{pmatrix} = \lambda e^{\lambda t}\vec{v}. \tag{2.13}$$

We are done. □

2.3.1 Real and distinct eigenvalues

Consider the case where A has n real eigenvalues $\lambda_1 \neq ... \neq \lambda_n$. In this case, the n corresponding (real) eigenvectors $\vec{v}_1, ..., \vec{v}_n$ are linearly independent. This gives n particular solutions to (3.17):

$$\vec{x}_1(t) = e^{\lambda_1 t}\vec{v}_1, \quad ..., \quad \vec{x}_n(t) = e^{\lambda_n t}\vec{v}_n.$$

It remains to verify that this is a fundamental set of solutions, by examining the determinant:

$$\begin{vmatrix}\vec{x}_1(t) & \cdots & \vec{x}_n(t)\end{vmatrix} = \begin{vmatrix}e^{\lambda_1 t}\vec{v}_1 & \cdots & e^{\lambda_1 t}\vec{v}_1\end{vmatrix} = e^{\lambda_1 t}\cdots e^{\lambda_n t}\begin{vmatrix}\vec{v}_1 & \cdots & \vec{v}_n\end{vmatrix} \neq 0.$$

Here we used that $\det(A)$ is linear in each column of A, then that the exponentials are positive and the vectors $v_1, ..., \vec{v}_n$ are linearly independent. Our result can be summarized as follows

Assume all eigenvalues of A are real and distinct, with respective eigenvectors $\vec{v}_1, ..., \vec{v}_n$. Then the general solution of (3.17) is

$$\vec{x}(t) = \sum_{k=1}^{n} c_k e^{\lambda_k t}\vec{v}_k. \tag{2.14}$$

Solving IVPs

Let us outline the process of solving IVP for (3.17) with initial condition $\vec{x}(t_0) = \vec{x}_0$.

(*i*) Find the general solution (3.19) by finding the eigenvalues and corresponding eigenvectors of A.

(*ii*) Form an $n \times n$ linear system in variables $c_1, ..., c_n$ by plugging the condition $\vec{x}(t_0) = \vec{x}_0$ in (3.19).

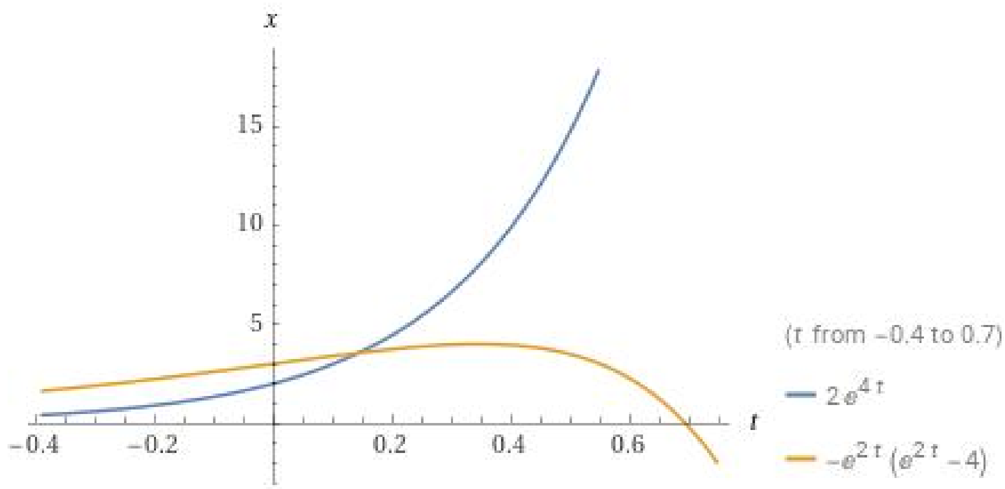

Figure 2.1: Component plot for the IVP solution in Example 3.4.1.

(iii) Identify $c_1, ..., c_n$ by solving this system. Solution is guaranteed to be unique.

Uniqueness is guaranteed since the matrix $(\vec{x}_1(t_0) \cdots \vec{x}_n(t_0))$ of the linear system in (ii) is nonsingular.

Example 2.3.1. *Solve the IVP and sketch the component plots:*

$$\vec{x}' = \begin{pmatrix} 4 & 0 \\ -1 & 2 \end{pmatrix} \vec{x}, \quad \vec{x}(\ln(2)) = \begin{pmatrix} 4 \\ 6 \end{pmatrix}.$$

We find $p_A(\lambda) = (4 - \lambda)(2 - \lambda)$, so the eigenvalues are $\lambda_1 = 4$, $\lambda_2 = 2$. Since $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_1 \neq \lambda_2$, the corresponding eigenvectors $\vec{v}_1, \vec{v}_2$ are real and independent. Now,

$$A - \lambda_1 I = \begin{pmatrix} 0 & 0 \\ -1 & -2 \end{pmatrix} \implies \vec{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \implies \vec{x}_1(t) = e^{4t} \begin{pmatrix} 2 \\ -1 \end{pmatrix};$$

$$A - \lambda_2 I = \begin{pmatrix} 2 & 0 \\ -1 & 0 \end{pmatrix} \implies \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \implies \vec{x}_2(t) = e^{2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We now find the unknown constants c_1, c_2 from the linear system

$$\begin{pmatrix} 2e^{4t_0} & 0 \\ -e^{4t_0} & e^{2t_0} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \vec{x}_0, \quad \text{that is} \quad \begin{pmatrix} 32 & 0 \\ -16 & 4 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 32 \\ 0 \end{pmatrix}.$$

We find $c_1 = 1$, $c_2 = 4$, and the IVP solution is $\vec{x}(t) = \vec{x}_1(t) + 4\vec{x}_2(t) = \begin{pmatrix} 2e^{4t} \\ -e^{4t} + 4e^{2t} \end{pmatrix}$. See Fig. 3.2.

Note that A might be singular, i.e. with 0 as one of the eigenvalues; the theory remains valid.

Example 2.3.2 (Compartment model). *The levels of liquid in two connected tanks (Fig. 3.3) at time t satisfy*

$$\vec{x}' = \begin{pmatrix} -k & k \\ k & -k \end{pmatrix} \vec{x}$$

Figure 2.2: Compartment model.

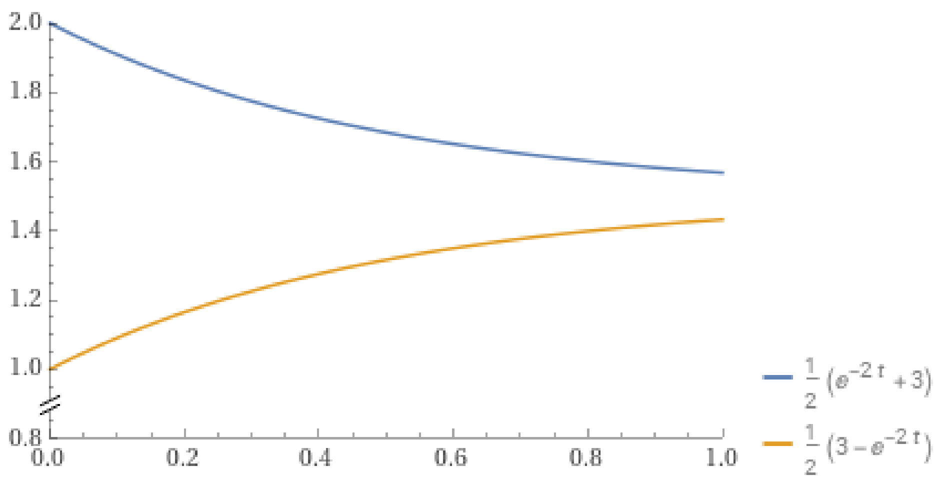

Figure 2.3: Component plot for the IVP solution in Example 3.4.2.

where $k > 0$ is a parameter depending on the liquid Solve the IVP with $\vec{x}(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and plot the component plots for the solution. Can you guess, without computation, the value of $\vec{x}(t)$ for large t?

First, we guess that the two levels should match as $t \to +\infty$, as the equilibria are $(c, c)^\top$ for $c \in \mathbb{R}$. In fact, some intuition from a high-school physics class—the liquid pressure formula $\rho g h$—might hint that the asymptotic level c is the average of the initial levels in the tanks, i.e. $c = \frac{3}{2}$. We now solve the IVP. Note that we can factor out k from the matrix, i.e. $A = kB$ with

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

The eigenvalues of B are found to be $\mu_1 = 0$, $\mu_2 = -2$, so those of A are $\lambda_1 = 0$, $\lambda_2 = -2k$. For $\vec{\lambda}_1 = 0$, we take $\vec{v}_1 = \begin{pmatrix} 1 & 1 \end{pmatrix}^\top$. Since B is symmetric, $\vec{v}_2$ is orthogonal to $\vec{v}_1$, and we take $\vec{v}_2 = \begin{pmatrix} 1 & -1 \end{pmatrix}^\top$. The initial conditions amount to $c_1 + c_2 = 2$ and $c_1 - c_2 = 1$, whence we find $c_1 = \frac{3}{2}$, $c_2 = \frac{1}{2}$, and

$$\vec{x}(t) = \frac{1}{2} \begin{pmatrix} 3 + e^{-2kt} \\ 3 - e^{-2kt} \end{pmatrix}.$$

Our guess is correct, and convergence is exponentially fast. See Fig. 3.4. □

Phase portraits for two real eigenvalues $\lambda_1 \neq \lambda_2 \neq 0$

Note that when A is nonsingular, 0 is not an eigenvalue, and the only critical point is $(0; 0)$. Its type is defined by the signs of λ_1 and λ_2. An equilibrium is asymptotically stable if both eigenvalues are negative, and unstable otherwise. In this section, we assume that $\lambda_1 > \lambda_2$ without loss of generality.

- There is a `Wolfram notebook` that allows to sketch phase portraits in two-dimensional systems!

Nodal source ($\lambda_1 > \lambda_2 > 0$). Unstable equilibrium: trajectories emanate from the origin (see Example 3.4.1). To sketch the phase portrait, we start by plotting the eigenspaces, i.e. straight lines along $\vec{v}_1$ and $\vec{v}_2$, and we put arrows in the direction from the origin. Any other trajectory emanates from the origin to infinity, remaining in its sector. It gets parallel to $\vec{v}_1$ far from the origin and tangent to $\vec{v}_2$ in the origin. In Fig. 3.5, we sketch the phase portrait of the system in Example 3.4.1.

Nodal sink ($\lambda_2 < \lambda_1 < 0$). This is an asymptotically stable equilibrium: trajectories go towards the origin. Note that we can obtain such an equilibrium from a nodal sink by negating A, see Fig. 3.5, which corresponds to time reversal. The arrows on the straight lines point towards the origin. Other trajectories go to the origin, getting parallel to $\vec{v}_2$ far from the origin and tangent to $\vec{v}_1$ in the origin.

Saddle ($\lambda_1 > 0 > \lambda_2$). This is an unstable equilibrium. The arrows point outwards along $\vec{v}_1$ and towards the origin along $\vec{v}_2$. All other trajectories are U-shaped: they approach the origin up to a certain point, then run away from it. Far away from the origin they get parallel to $\vec{v}_1$ or $\vec{v}_2$, and the arrows conform to those on the straight lines. In Fig. 3.5, we sketch the phase portrait for $\vec{x}' = A\vec{x}$ with eigenvalues $\lambda_1 = 8$ and $\lambda_2 = -2$, and respective eigenvectors $\vec{v}_1 = (-6; 1)$ and $\vec{v}_2 = (4; 1)$.

2.3.2 Complex eigenvalues (without repetition)

Complex algebra. Euler's formula extends the exponential to imaginary numbers: *by definition,*

$$e^{i\theta} = \cos(\theta) + i\sin(\theta) \quad \forall \theta \in \mathbb{R}.$$

Then one can extend e^z to $z = x + iy \in \mathbb{C}$: *by definition,* $e^z := e^x e^{iy} = e^x \cos(y) + ie^x \sin(y)$. This is a very natural generalization as it preserves the key algebraic property of exponentials, namely that

$$e^{z_1+z_2} = e^{z_1}e^{z_2} \quad \text{for all} \;\; z_1, z_2 \in \mathbb{C}.$$

(This can be checked by Euler's formula and the formulas for $\cos(\alpha \pm \beta)$ and $\sin(\alpha \pm \beta)$.) Moreover, we can define the derivative of $\varphi : \mathbb{R} \to \mathbb{C}$ by differentiating $\mathsf{Re}\varphi$ and $\mathsf{Im}\varphi$ separately. Then, for example, $(e^{i\omega t})' = (\cos(\omega t) + i\sin(\omega t))' = -\omega\sin(\omega t) + i\omega\cos(\omega t) = i\omega e^{i\omega t}$. More generally,

$$(e^{\lambda t})' = \lambda e^{\lambda t} \quad \text{for all} \quad \lambda \in \mathbb{C}, \tag{2.15}$$

so the key differential property of the exponential function is also preserved. To summarize, "the usual algebra and analysis" still work for $t \mapsto e^{\lambda t}$ with $\lambda \in \mathbb{C}$, plus we can exploit Euler's formula.

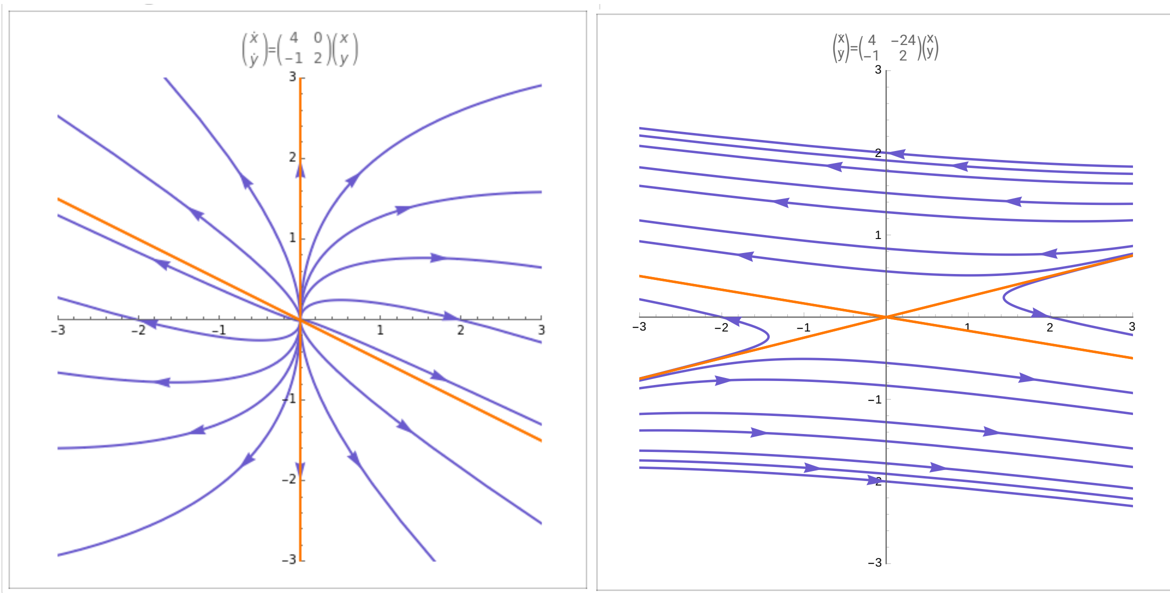

Figure 2.4: **Left:** nodal source (Example 3.4.1). In a nodal sink, the arrows are flipped. **Right:** saddle.

Case $n = 2$

We now consider the case of two complex eigenvalues in dimension $n = 2$. Recall from Section 2.1 that complex eigenvalues (as well as eigenvectors) of a *real* matrix come in mutually conjugate pairs:

$$\begin{aligned} \lambda_1 &= \mu + i\omega, \quad \vec{v}_1 = \vec{a} + i\vec{b}, \\ \lambda_2 &= \mu - i\omega, \quad \vec{v}_2 = \vec{a} - i\vec{b}, \end{aligned} \tag{2.16}$$

for some $\mu, \omega \in \mathbb{R}$; $\vec{a}, \vec{b} \in \mathbb{R}^n$. This holds for all n, but in the case $n = 2$ we have no other eigenvalues, and we should be able to obtain a fundamental set of *real* solutions for (3.17). We prove the following:

Proposition 2.3.1. *System (3.17) with 2 complex eigenvalues (3.21) has a fundamental set of solutions*

$$\boxed{\begin{aligned} \vec{u}(t) &= e^{\mu t}\cos(\omega t)\vec{a} - e^{\mu t}\sin(\omega t)\vec{b}, \\ \vec{v}(t) &= e^{\mu t}\sin(\omega t)\vec{a} + e^{\mu t}\cos(\omega t)\vec{b}. \end{aligned}}$$

As such, the general solution is $\vec{x}(t) = \alpha\vec{u}(t) + \beta\vec{v}(t)$ *for* $\alpha, \beta \in \mathbb{R}$.

Example 2.3.3. *Solve the IVP*

$$\vec{x}' = \begin{pmatrix} -1 & 2 \\ -1 & -3 \end{pmatrix} \vec{x}, \quad \vec{x}(\pi) = e^{-2\pi} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Here $p_A(\lambda) = \lambda^2 + 4\lambda + 5$, whence $\lambda_{1,2} = -2 \pm i$. The respective eigenvectors can be computed as in (2.1.2), part (b); we will get

$$\vec{v}_{1,2} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \pm i \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In terms of (3.21), $\mu = -2$, $\omega = 1$, and we find the general solution

$$\vec{x}(t) = e^{-2t}\cos t \begin{pmatrix} 2c_1 \\ c_2 - c_1 \end{pmatrix} + e^{-2t}\sin t \begin{pmatrix} 2c_2 \\ -c_1 - c_2 \end{pmatrix} \quad \forall c_1, c_2 \in \mathbb{R}.$$

From the initial condition we get $\begin{pmatrix} -2c_1 \\ c_1 - c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, whence $c_1 = -\frac{1}{2}$, $c_2 = \frac{1}{2}$, and the solution is

$$\vec{x}(t) = e^{-2t}\cos t \begin{pmatrix} -1 \\ 1 \end{pmatrix} + e^{-2t}\sin t \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \qquad \square$$

Proof of Proposition 3.4.1. By (3.20), $y(t) = e^{\lambda t}$ satisfies a homogeneous ODE $y' - \lambda y = 0$ with *complex* coefficient $\lambda \in \mathbb{C}$. Therefore, repeating (3.18), we verify that

$$\vec{x}_1(t) = e^{\lambda_1 t}\vec{v}_1, \quad \vec{x}_2(t) = e^{\lambda_2 t}\vec{v}_2.$$

are solutions to (3.17). However, they are complex, and we need real solutions. Now, we check that

$$\begin{aligned} \vec{x}_1(t) &= \vec{u}(t) + i\vec{v}(t), \\ \vec{x}_2(t) &= \vec{u}(t) - i\vec{v}(t), \end{aligned}$$

that is $\vec{u}(t) = \mathsf{Re}\vec{x}_1(t)$ and $\vec{v}(t) = \mathsf{Im}\vec{x}_1(t)$. Indeed, by Euler's formula

$$\begin{aligned} \vec{x}_1(t) = e^{\mu t}[\cos(\omega t) + i\sin(\omega t)](\vec{a} + i\vec{b}) &= e^{\mu t}[\cos(\omega t)\vec{a} - \sin(\omega t)\vec{b} + i\sin(\omega t)\vec{a} + i\cos(\omega t)\vec{b}] \\ &= \vec{u}(t) + i\vec{v}(t), \end{aligned}$$

similarly for $\vec{x}_2(t)$. This explains where $\vec{u}(t), \vec{v}(t)$ come from, and also shows that they are solutions:

$$\frac{d}{dt}\vec{u}(t) = \frac{d}{dt}\mathsf{Re}(\vec{x}_1(t)) \overset{(i)}{=} \mathsf{Re}\left(\frac{d}{dt}\vec{x}_1(t)\right) \overset{(ii)}{=} \mathsf{Re}(A\vec{x}_1(t)) \overset{(ii)}{=} A\mathsf{Re}(\vec{x}_1(t)) = A\vec{u}(t).$$

Make sure you understand why (i)–(iii) hold. Finally, to verify that $\vec{u}, \vec{v}$ are independent, note that

$$\begin{pmatrix} \vec{u} & \vec{v} \end{pmatrix} = ze^{\mu t}\begin{pmatrix} \vec{a} & \vec{b} \end{pmatrix}\begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}$$

where the rotation matrix in the right-hand side has determinant 1. Hence $\det\begin{pmatrix} \vec{u} & \vec{v} \end{pmatrix} = e^{2\mu t}\det\begin{pmatrix} \vec{a} & \vec{b} \end{pmatrix}$, where $e^{2\mu t} > 0$ since $\mu \in \mathbb{R}$. But if we assume that $\det\begin{pmatrix} \vec{a} & \vec{b} \end{pmatrix} = 0$, then $\vec{v}_1, \vec{v}_2$ must also be dependent due to (3.21), and this contradicts their linear independence (over $\mathbb{C}$) as they correspond to $\lambda_1 \neq \lambda_2$. $\square$

2.3.3 Case $n > 2$.

Here we do not consider the case of more than 2 complex eigenvalues in detail. In a nutshell, any pair of simple complex eigenvalues (i.e. with algebraic multiplicity 1) gives a pair of solutions as in Proposition 3.4.1. If $n > 2$ and a pair has multiplicity $k > 1$, then one can find solutions by combining Proposition 3.4.1 with the Jordan decomposition trick for repeated eigenvalues, to be presented later.

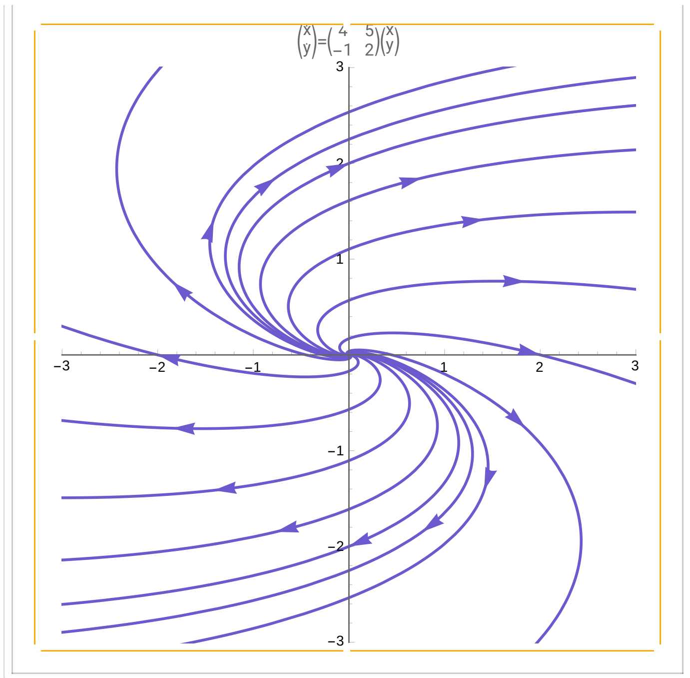

Figure 2.5: Phase portrait for $\vec{x}' = A\vec{x}$ where A has eigenvalues $3 \pm 2i$, so $(0;0)$ is a spiral source.

Phase portraits for two complex eigenvalues

When the eigenvalues are complex, there are three cases, depending on the sign of their real part μ.

Spiral source ($\mu > 0$) – Spiral sink ($\mu < 0$). All trajectories are spirals emanating from (resp. going to) the origin for a source (resp. sink); source is unstable, and sink is stable. (Do you understand why?) In Fig. 3.5, we sketch the phase portrait of a system with a spiral source, and in Example 3.4.3 the equilibrium is a spiral sink. When sketching a spiral source/sink, the only dilemma is to guess the direction of rotation: clockwise or counterclockwise. This can be done by trying a couple of “simple” test values for $\vec{x}$, e.g. $(1;0)$ and $(0;1)$, computing the corresponding values of the flow $\vec{x}'$, and checking if they are consistent with the tentative direction of rotation. For the system

$$\vec{x}' = \begin{pmatrix} 4 & 5 \\ -1 & 2 \end{pmatrix} \vec{x}$$

in Figure 3.6, the eigenvalues are $3 \pm 2i$, so we have a spiral source. From $(1;0)$ we move in the direction $(4;-1)$, and from $(0;1)$ in the direction $(5;2)$; this is consistent with the clockwise rotation.

Circular case ($\mu = 0$). If $\mu = 0$, i.e. both eigenvalues are imaginary, all trajectories are ellipses, and the direction of rotation can be determined by the same method as for spiral source/sink. In fact this is a “neutral” equilibrium: it does not attract nor repell.

2.3.4 Repeated eigenvalues

We now consider the general situation of an n-dimensional system with eigenvalues $\lambda_1, ..., \lambda_n$ that might repeat. The following example demonstrates that this is a practically relevant situation.

Example 2.3.4 (Reversive motion)**.** *The motion of a point in the plane is described by the equation*

$$x' = -x + ky,$$
$$y' = -y,$$

where $k \in \mathbb{R}$ is a parameter, with initial condition $\vec{r}(0) = (1;\, 2)$. Find its position $\vec{r}(t)$ at any $t \in \mathbb{R}$.

Let us deal with this example "ad-hoc," then study the general case. Our equation is $\vec{r}' = Ar$ with

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

The matrix is upper-triangular, and diagonal when $k = 0$. For $k \neq 0$, the variables x, y are coupled.

Diagonal case: $k = 0$. Here $A = -I$, with eigenvalues $\lambda_{1,2} = -1$. Moreover, $A - (-1)I = 0$, so any $\vec{v} \in \mathbb{R}^2$ is an eigenvector; in particular, $\vec{v}_1 = (1; 0)$ and $\vec{v}_2 = (0; 1)$ is a pair of independent ones. On the other hand, in our system with $k = 0$, each of the two equations only concerns its own variable x or y, so these variables do not interact; thus, we can solve the two equations *separately.* The corresponding general solution is given by $x(t) = c_1 e^{-t}$ and $y(t) = c_2 e^{-t}$ for $c_1, c_2 \in \mathbb{R}$, that is

$$\vec{r}(t) = c_1 e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Nondiagonal case: $k \neq 0$. Note that the eigenvalues of A are the same as before: $\lambda_{1,2} = -1$. (This is, in fact, a general result for upper- and lower-triangular matrices.) However, the nullspace of

$$A + I = \begin{pmatrix} 0 & k \\ 0 & 0 \end{pmatrix}$$

has dimension 1; namely, it is the span of $\vec{v} = (1; 0)$. On the other hand, we can solve our system by substitution. Namely, we first solve the second equation that only concerns y; its general solution is $y(t) = ce^{-t}$. Plugging $y(t)$ in the first equation, we get a parameterized first-order ODE in x:

$$x' + x = cke^{-t},$$

that can be solved via the integrating factor method. Doing so, we get $x(t) = (c_1 + ckt)e^{-t}$. Introducing $\vec{w}_1 = (0; \frac{1}{k})$ and $c_2 = ck$, we express the answer in the vector form:

$$\vec{r}(t) = e^{-t} \begin{pmatrix} c_1 + ckt \\ c \end{pmatrix}$$
$$= c_1 e^{-t} \vec{v} + c_2 e^{-t} \left(t\vec{v} + \vec{w} \right).$$

We also note that eigenvalue $\lambda = -1$, respective eigenvector $\vec{v}$, and the vector $\vec{w}$ satisfy the relation

$$(A - \lambda)\vec{w} = \vec{v}.$$

From linear algebra we recall that $\vec{w}$ is called (the first) generalized eigenvector of eigenvalue λ. $\square$

Generalized eigenvectors. Before we proceed further to the general theorem, let's recall a result from linear algebra (see also the beginning of this chapter). Assume A is a square matrix of dimension n, and its characteristic polynomial p_A has some root $\lambda \in \mathbb{C}$ with multiplicity $m \leqslant n$. Then $m(\lambda)$ is called the *algebraic multiplicity* of eigenvalue λ, whereas $s(\lambda) = \dim(\mathsf{Null}(A - \lambda I))$ is the *geometric multiplicity* of eigenvalue λ. For any eigenvalue λ, it holds that $1 \leqslant s(\lambda) \leqslant m(\lambda) \leqslant n$, and the sum of algebraic multiplicities over all (distinct) eigenvalues of A is n. If it happens, for A at hand, that $s(\lambda) = m(\lambda)$, then the subspace $\mathsf{Null}(A - \lambda I)$ has a basic of eigenvectors, and otherwise it's not the case; these two situations corresponded, respectively, to $k = 0$ and $k = 1$ in the above example. We now recall Jordan's theorem from linear algebra – first, for the case of $s(\lambda) = 1$.

Theorem 2.3.1. *Let λ be an eigenvalue of A with algebraic multiplicity $m \geqslant 2$ and single independent eigenvector $\vec{v}$. There exist m independent vectors $\vec{w}_0[= \vec{v}], \vec{w}_1, ..., \vec{w}_{m-1}$ that form the* Jordan chain*:*

$$
\begin{aligned}
(A - \lambda I)\vec{w}_0 &= 0, \\
(A - \lambda I)\vec{w}_1 &= \vec{w}_0, \\
&\vdots \\
(A - \lambda I)\vec{w}_{m-1} &= \vec{w}_{m-2}.
\end{aligned}
$$

Moreover, the chain cannot be continued: the system $(A - \lambda I)\vec{u} = \vec{w}_{m-1}$ has no solutions in $\vec{u} \in \mathbb{C}^n$.

In the previous example with $k \neq 0$, for $\lambda = -1$ we have $m = 2$ and $s = 1$; the Jordan chain is comprised of $\vec{w}_0 = \vec{v} = (1; 0)$ and $\vec{w}_1 = \vec{w} = (0; \frac{1}{k})$, and there exists no $\vec{u}$ such that $(A - \lambda I)\vec{u} = \vec{w}_1$. We can construct a Jordan chain for λ with $m \geqslant 2$ and $s = 1$ as follows.

1. Find some eigenvector $\vec{v}$ by solving $(A - \lambda I)\vec{v} = 0$, and let $\vec{w}_0 = \vec{v}$
2. Repeat the following for $k \in \{1, ..., m-1\}$: given $\vec{w}_{k-1}$, find $\vec{w}_k$ by solving $(A - \lambda I)\vec{u} = \vec{w}_{k-1}$.

In fact, it is guaranteed that the solution in step 2 is unique for each $k \in \{1, ..., m-1\}$; in particular, we can rescale the whole chain by the same scalar, but not its vectors separately from each other.

Remark 2.3.1. *We briefly discuss the case of $1 < s < m$ (which we shall not encounter in this class). In this case, there are s independent eigenvectors $\vec{v}_1, ..., \vec{v}_s$ and s respective Jordan chains, each starting from its own eigenvector. The sum of lengths of these chains is m. To find these chains, we can run the above process for each chain incrementally, cutting it when we cannot find the next link.*

Returning to systems of DEs, we have the following general result.

Theorem 2.3.2. *Let $A \in \mathbb{C}^{n \times n}$ have an eigenvalue λ with algebraic multiplicity m and geometric multiplicity s, with the corresponding Jordan chains*

$$
\left(\vec{w}_0^{(1)}, ..., \vec{w}_{m_1-1}^{(1)}\right); \quad \cdots; \quad \left(\vec{w}_0^{(s)}, ..., \vec{w}_{m_1-1}^{(s)}\right).
$$

Then $\vec{x}\,' = A\vec{x}$ has m solutions of the form

$$
e^{\lambda t}\vec{w}_0^{(j)}, \quad e^{\lambda t}\left(t\vec{w}_0^{(j)} + \vec{w}_1^{(j)}\right), \quad ..., \quad e^{\lambda t}\left(\sum_{k=1}^{m_j} t^{m_j-k}\vec{w}_{k-1}^{(j)}\right) \quad \text{for } j \in \{1, ..., s\}.
$$

The resulting n solutions (for all eigenvalues) are independent, so give a fundamental set of solutions.

Proof. We only consider the simplest case to highlight the mechanism; the general case can be handled by induction. Namely, assume $m = 2$ and $s = 1$, and let $\vec{v}, \vec{w}$ be the Jordan chain of λ. We already know that $\vec{x}_0(t) = e^{\lambda t}\vec{v}$ is a solution of $\vec{x}' = A\vec{x}$, so it remains to verify that $\vec{x}_1(t) = e^{\lambda t}(t\vec{v}+\vec{w})$ is a solution as well, and that the two are independent. For the first claim, we observe that

$$A\vec{x}_1(t) = e^{\lambda t}(tA\vec{v} + A\vec{w}) = e^{\lambda t}(\lambda t\vec{v} + \lambda\vec{w} + \vec{v}) = \vec{x}'(t).$$

As for the second claim, it follows from the independence of $\vec{v}$ and $\vec{w}$. Indeed, the Wronskian reads

$$W(t) = \left|\vec{x}_0(t) \quad \vec{x}_1(t)\right| = \left|e^{\lambda t}\vec{v} \quad e^{\lambda t}(t\vec{v} + \vec{w})\right| = e^{2\lambda t}\left|\vec{v} \quad t\vec{v} + \vec{w}\right| = e^{2\lambda t}\left|\vec{v} \quad \vec{w}\right| \neq 0. \qquad \square$$

Chapter 3

Higher-order linear equations

Preliminaries

In this chapter, we shall mostly focus on linear ODEs of second order

$$y'' + a_1 y' + a_2 y = f(t), \quad t \in I \tag{3.1}$$

with constant or variable coefficients, and, to some extent, on linear ODEs of arbitrary order $n \geqslant 2$,

$$y^{(n)} + a_1 y^{(n-1)} + ... + a_{n-1} y' + a_n y = f(t), \quad t \in I. \tag{3.2}$$

We shall pay special attention to harmonic oscillator, which is an important subclass of second-order systems that appears in several areas of classical physics; in particular, as a spring-mass system presented in the next section. Note that any linear ODE (3.2) with constant coefficients can solve via vectorization, i.e. by reducing it to a second order system and solving that system; this is how we proceeded in the previous chapter. Here, we shall instead propose another method, avoiding vectorization—and any linear-algebraic calculations, seemingly—and solving (3.2) directly. In fact, with some linear algebra (again!), one can show that the two methods are equivalent in a suitable sense; I will not endeavor to do it in these lectures. On the other hand, the method to be presented admits an interpretation in terms of linear operators and their eigenfunctions, that we shall discuss.

3.1 Spring-mass system

In spring-mass system, a mass of m grams[1] is attached to a vertical spring whose other end is attached to a horizontal surface (see Fig. 3.1), and whose stiffness coefficient is k. (The unit of k is Newton-per-meter, or N/m, where Newton is the unit of force. Note that $1\text{N} = 1\frac{\text{kg}\cdot\text{m}}{\text{s}^2} = 10^5\ \frac{\text{g}\cdot\text{cm}}{\text{s}^2}$.) Assuming that air resistance is negligible, and the spring is "ideal" (long and perfectly elastic), what forces are acting on the mass?

- Gravitational force $F_{\text{grv}} = m\text{g}$, where g is the free fall acceleration; directed downwards.
- Stretching force F_{str} that "desires" to put the system back to the equilibrium position.
- Possibly, external force F_{ext} applied to the mass.

[1] We shall use metric units – get used to it.

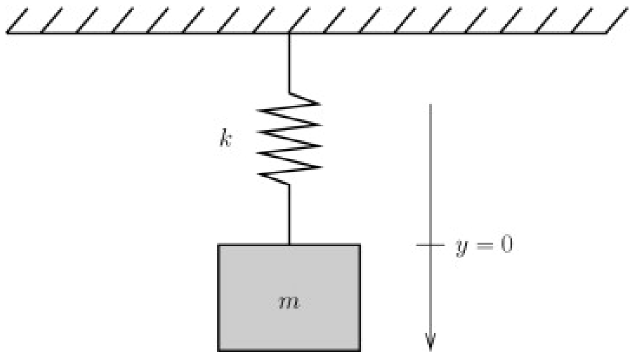

Figure 3.1: Spring-mass system.

Let ℓ_0 be the stretch of the spring by gravity alone, in equilibrium position. By Hook's law, one has

$$k\ell_0 = m\mathrm{g}. \tag{3.3}$$

Since $\mathrm{g} = 9.8$ N/m is known, this allows to find any one of the three quantities k, ℓ_0, m knowing the other two. Now, let y be the vertical position of the mass, measured downwards and in such a way that the equilibrium position corresponds to $y = 0$. (That is, $y + \ell_0$ is the excess length of the spring compared to its length when undeformed (i.e. if $m = 0$), negative if the spring is squeezed.) By Hook's law, the stretching force is $F_{\text{str}} = -k(y + \ell_0)$. On the other hand, by the 2nd Newton's law,

$$my'' = F_{\text{grv}} + F_{\text{str}} + F_{\text{ext}}.$$

Plugging in the equilibrium equation and the value of F_{str}, we get $my'' = k\ell_0 - k(y + \ell_0) + F_{\text{ext}}$, or

$$my'' + ky = F_{\text{ext}}.$$

Dividing over m and defining $\omega_0^2 := \frac{k}{m}$ and $f = \frac{1}{m}F_{\text{ext}}$, we arrive at the standardized equation:

$$y'' + \omega_0^2 y = f. \tag{3.4}$$

This ODE has a special name: *equation of a harmonic oscillator* with *natural frequency* ω_0 and *forcing function* f. Note that in the homogeneous system ($f = 0$), the general real solution writes as

$$y(t) = c_1 \cos(\omega_0 t) + c_2 \sin(\omega_0 t),$$

which explains the name chosen for ω_0. If f is not identically zero, we say that the spring-mass system (or harmonic oscillator) is *forced*; otherwise it is *unforced*. Finally, we can introduce damping, adding a force $F_{\text{dmp}} = -\gamma y'$ where $\gamma \geqslant 0$ is the viscosity coefficient of the medium. Thus, equation

$$y'' + \gamma y' + \omega_0^2 y = f. \tag{3.5}$$

describes an oscillator with damping; if $\gamma = 0$, we say that the oscillator is *undamped*. We shall revisit equations (3.4)–(3.5) later, after discussing general second-order ODEs with constant coefficients.

Example 3.1.1. *A mass weighing* 2 *Newtons is attached to a spring and stretches it by 4 cm. The spring is released from rest at a point that is 3 cm above the equilibrium point. Write down the IVP describing the vertical position of the mass at a given time. The free fall acceleration is* $9.8\frac{m}{s^2} \approx 10\frac{m}{s^2}$.

First, we convert the units. E.g., let's switch to meters and kilograms. Then $\ell_0 = 0.04$ for the initial stretch (in meters), and $mg = 2$ for the weight (in Newtons; recall that $\mathrm{N} = \frac{\mathrm{kg}\cdot\mathrm{m}}{\mathrm{s}^2}$). Whence $m = \frac{2}{9.8} = 0.2\,\mathrm{kg}$, and from the equilibrium condition $k = \frac{mg}{\ell_0} = \frac{2}{0.04} = 50$ N/m. Thus $\omega_0^2 = \frac{k}{m} = \frac{50}{0.2} = 250\ \mathrm{s}^{-2}$. Since there is no damping and no forcing, we write down the IVP:

$$y'' + 250y = 0, \quad y(0) = -0.03,\ y'(0) = 0.$$

Its solution $y(t)$ is the vertical position counted *downwards* relatively to the equilibrium, in *meters*. $\square$

Other harmonic oscillators. In addition to the spring-mass system, there are other physical systems whose behavior is modeled by the harmonic oscillator equation. Two classical examples are:

(a) *pendulum* for small displacement angles θ, such that the approximation $\theta \approx \sin\theta$ can be used;

(b) *LC-circuit* – an electrical circuit that consists of an inductor and a capacitor.

We are not going to consider the physical aspects of any of these systems in this class. Instead, we shall either be reasoning in terms of the abstract equation (3.4), or in terms of a spring-mass system when aiming to highlight some aspect of mathematical modeling and/or use some physical intuition.

3.2 Theory recap: General linear ODEs

In the previous chapter, we have established some general properties of homogeneous linear ODEs,

$$y^{(n)} + a_1(t)y^{(n-1)} + ... + a_{n-1}(t)y' + a_n(t)y = 0, \quad t \in I, \tag{3.6}$$

and IVPs associated to them. We need these results in this chapter, so we briefly summarize them.

Fundamental set of solutions. Assuming that the coefficients $a_1(t), ..., a_n(t)$ are continuous on I, there exist n, and not more, linearly independent solutions $y_1(t), ..., y_n(t)$ on I, and any solution $y(t)$ is their linear combination with constant coefficients: $y(t) = c_1y_1(t) + ... + c_ny_n(t)$. In fact, the set of all solutions of (3.6) is an n-dimensional subspace of the (infinite-dimensional) space $C^n(I)$ of n-times differentiable functions on I, and fundamental sets of solutions are just its bases. E.g., both $\{\cos\omega_0 t,\ \sin\omega_0 t\}$ and $\{\cos\omega_0 t + \sin\omega_0 t,\ \cos\omega_0 t - \sin\omega_0 t\}$ are fundamental sets of solutions of (3.4) with $f = 0$ (provided that $\omega_0 \neq 0$).

Testing independence via Wronskian. Given a set of n functions $f_1, ..., f_n$ that are $n-1$ times continuously differentiable on I, we can define the matrix-function

$$\boldsymbol{M}(t) = \begin{pmatrix} f_1(t) & \cdots & f_n(t) \\ f_1'(t) & \cdots & f_n'(t) \\ \vdots & & \vdots \\ f_1^{(n-1)}(t) & \cdots & f_n^{(n-1)}(t) \end{pmatrix}.$$

Its determinant $W(t)$ is called the Wronskian of $f_1, ..., f_n$. If $f_1, ..., f_n$ are linearly *dependent* on I, then $W \equiv 0$ on I. Rephrasing this, to show that $f_1, ..., f_n$ are linearly *independent* it suffices to find $t_0 \in I$ such that $W(t_0) \neq 0$. This test can be applied to a set of n solutions $y_1, ..., y_n$ of an ODE to check that this set is a fundamental one. For example, two solutions $\{\cos \omega_0 t,\ \sin \omega_0 t\}$ of (3.4) with $f = 0$ and $\omega_0 \neq 0$ indeed form a fundamental set on $\mathbb{R}$, since their Wronskian does not vanish:

$$W(t) = \begin{vmatrix} \cos \omega_0 t & \sin \omega_0 t \\ -\omega_0 \sin \omega_0 t & \omega_0 \cos \omega_0 t \end{vmatrix} = \omega_0 \begin{vmatrix} \cos \omega_0 t & \sin \omega_0 t \\ -\sin \omega_0 t & \cos \omega_0 t \end{vmatrix} = \omega_0 \neq 0.$$

Existence and uniqueness of solutions in IVPs. Finally, recall that an IVP for equation (3.6) with initial conditions $y(t_0) = s_0,\ y'(t_0) = s_1,\ ...,\ y^{(n-1)}(t_0) = s_{n-1}$ has a unique solution, provided that the coefficient functions $a_1, ..., a_n$ are continuous on I. More generally, this holds for nonhomogeneous linear ODEs with some forcing function f, assuming that f is also contunous on I.

3.3 Homogeneous linear ODEs with constant coefficients

In this section, we derive the general solution for *homogeneous* equations with *constant* coefficients:

$$y^{(n)} + a_1 y^{(n-1)} + ... + a_{n-1} y' + a_n y = 0. \tag{3.7}$$

Definition 6. *$p_a(\lambda) = \lambda^{(n)} + a_1 \lambda^{n-1} + ... + a_{n-1}\lambda + a_n$ is the* characteristic polynomial *of equation (3.7).*

At this point, one might be confused with this terminology: we already have $p_A(\lambda) = \det(A - \lambda I)$, the characteristic polynomial of both $A \in \mathbb{C}^{n \times n}$ and the associated first-order linear system of DEs:

$$\vec{x}\,' = Ax. \tag{3.8}$$

This perceived misnomer is "a feature rather than a bug," as the next simple observation shows.

Proposition 3.3.1. *If (3.8) is the equivalent first-order system for ODE (3.7), then $p_a(\lambda) = (-1)^n p_A(\lambda)$.*

Proof. First note that for $n = 1$, the claim holds trivially. In the case $n = 2$,

$$p_A(\lambda) = \begin{vmatrix} -\lambda & 1 \\ -a_2 & -a_1 - \lambda \end{vmatrix} = \lambda^2 + a_1 \lambda + a_2 = p_a(\lambda).$$

The general case can be handled by induction. Recall that A is called the *companion matrix* of p_a. The lemma claims that for the companion matrix of p_a, the characteristic polynomial is p_a itself. □

Another immediate observation is that the roots $\lambda_1, ..., \lambda_n \in \mathbb{C}$ of $p_a(\lambda)$ generate solutions of (3.7).

Proposition 3.3.2. *If λ is a root of p_a, then $y(t) = e^{\lambda t}$ is a solution of (3.7). Moreover, solutions $\{y_j(t) = e^{\lambda_j t}, j \in J\}$ generated by a subset $\{\lambda_j, j \in J\}$ of distinct roots are linearly independent.*

For the proof, we define a special matrix that shall also be used later (and is useful in general).

Definition 7. *The Vandermonde matrix corresponding to $\lambda_1, ..., \lambda_d \in \mathbb{C}$ is a $d \times d$ matrix given by*

$$V_d(\lambda_1, ..., \lambda_d) = \begin{pmatrix} 1 & \cdots & 1 \\ \lambda_1 & \cdots & \lambda_d \\ \vdots & & \vdots \\ \lambda_1^{d-1} & \cdots & \lambda_d^{d-1} \end{pmatrix}.$$

Note that if there is a repetition among $\lambda_1, ..., \lambda_d$, then the corresponding Vandermonde matrix is singular. Moreover, it turns out that this is "if and only if," as the following result implies.

Lemma 3.3.1. *For any $\lambda_1, ..., \lambda_d \in \mathbb{C}$, the determinant of the corresponding Vandermonde matrix is*

$$\det V_d(\lambda_1, ..., \lambda_d) = \prod_{1 \leqslant j < k \leqslant d} (\lambda_k - \lambda_j).$$

Proof. This can be shown by induction: we have $\det V_2(\lambda_1, \lambda_2) = \lambda_2 - \lambda_1$ for the base, and the induction step can be carried out via the cofactor expansion. (Try it, this is a good exercise.) □

We can now prove Proposition 3.3.2. Plugging $y^{(n)}(t) = \lambda^n e^{\lambda t}$ in (3.7), we see that

$$y^{(n)}(t) + a_1 y^{(n-1)}(t) + ... + a_n y(t) = e^{\lambda t} p_a(\lambda) = 0$$

and verify the first claim. For the second claim, assume w.l.o.g. that $J = \{1, 2, ..., d\}$ and note that

$$\begin{vmatrix} y_1(t) & \cdots & y_d(t) \\ y_1'(t) & \cdots & y_d'(t) \\ \vdots & & \vdots \\ y_1^{(d-1)}(t) & \cdots & y_d^{(d-1)}(t) \end{vmatrix} = \begin{vmatrix} e^{\lambda_1 t} & \cdots & e^{\lambda_1 t} \\ \lambda_1 e^{\lambda_1 t} & \cdots & \lambda_d e^{\lambda_d t} \\ \vdots & & \vdots \\ \lambda_1^{d-1} e^{\lambda_1 t} & \cdots & \lambda_d^{d-1} e^{\lambda_d t} \end{vmatrix} = e^{\lambda_1 t} \cdots e^{\lambda_d t} \det V_d(\lambda_1, ..., \lambda_d) \neq 0. \qquad \square$$

General solution in the case of distinct roots. If all characteristic roots $\lambda_1, ..., \lambda_n$ in (3.7) are distinct, invoking Proposition 3.3.2 with $d = n$ implies that the general (complex) solution of (3.7) is

$$c_1 e^{\lambda_1 t} + ... + c_n e^{\lambda_n t} \tag{3.9}$$

with arbitrary (complex) constants $c_1, ..., c_n$. Moreover, if $\lambda_1, ..., \lambda_n \in \mathbb{R}$ then (3.9) with $c_1, ..., c_n \in \mathbb{R}$ is the general real solution. More generally, if $\mu + i\omega$ is a root, its conjugate $\mu - i\omega$ is also a root, and we replace the corresponding pair of terms in (3.9) with $e^{\mu t}(A \cos \omega t + B \sin \omega t)$ to get real solutions.

3.3.1 Repeated roots

As in first-order systems, the case of repeated roots involves functions of the form $t^k e^{\lambda t}$. To handle it and prove the result to be stated next, it is it is convenient to use the formalism of linear operators.

Definition 8. *A linear operator on a vector space of functions $\mathcal{F}$ is a linear transformation of that space, i.e. $\Phi : \mathcal{F} \to \mathcal{F}$ such that $\Phi[\alpha f + \beta g] = \alpha \Phi[f] + \beta \Phi[g]$ for all $f, g \in \mathcal{F}$ and any constants[2] α, β.*

We denote the evaluation of a linear operator with brackets $[\cdot]$, rather than parentheses $(\cdot)$, since $\Phi[f]$ is itself a function that we might want to evaluate at some t; in such cases, we write $\Phi[f](t)$.

[2] For mathematicians: $\mathcal{F}$ is a vector space over $\mathbb{R}$ (resp., over $\mathbb{C}$), then the scalars α, β are real (resp., complex). That is, α, β come from the same field over which $\mathcal{F}$ is a vector space.

Linear ODEs with constant coefficients in terms of linear operators. An important linear operator, that we denote Δ, corresponds to taking the derivative of f. Its powers $\Delta^2, \Delta^3, ...$ corresponds to taking higher derivatives, and $\Delta^0 = \mathsf{Id}$ is the identity operator (mapping f to itself). Now, we can rewrite (3.7) as $\Phi[y] = 0$, where the linear operator Φ in the left-hand side is given by

$$\Phi = \Delta^n + \sum_{k=1}^{n} a_k \Delta^{n-k} = p_a(\Delta).$$

As such, we obtain a geometric characterization of the solution set of (3.7) as the nullspace of Φ.

Theorem 3.3.1. *Each root $\lambda \in \mathbb{C}$ with multiplicity $m \geqslant 2$ generates m solutions of (3.7) as follows:*

$$e^{\lambda t}, te^{\lambda t}, \; ..., \; t^{m-1}e^{\lambda t}. \tag{3.10}$$

The resulting n solutions, for all roots (repeated or not), are independent, so form a fundamental set.

Proof. **1^o**. For the first claim, we have to check that if λ is a root of p_a with multiplicity $m \geqslant 2$, then $te^{\lambda t}$, ..., $t^{m-1}e^{\lambda t}$ are solutions of (3.7). To this end, observe that

$$\Phi = p_a(\Delta) = (\Delta - \lambda_1 \mathsf{Id})\,(\Delta - \lambda_2 \mathsf{Id})\, \cdots \,(\Delta - \lambda_n \mathsf{Id}).$$

In this product, all factors commute, and $\Delta - \lambda\mathsf{Id}$ is repeated m times. Thus, it suffices to verify that for all $k \in \{1, ..., m-1\}$, one has

$$(\Delta - \lambda \mathsf{Id})^m [t^k e^{\lambda t}] = 0.$$

It remains to observe that, since $\Delta[t^k e^{\lambda t}] = kt^{k-1}e^{\lambda t} + \lambda t^k e^{\lambda t}$, it holds that $(\Delta - \lambda \mathsf{Id})[t^k e^{\lambda t}] = kt^{k-1}e^{\lambda t}$, i.e. each application of $(\Delta - \lambda\mathsf{Id})$ lowers the power of t by one. The first claim is proved.
2^o. For the second claim, we only show the independence within the subset of solutions (3.10) corresponding to a single one repeated root λ. Consider first the case $m = 2$ to get some intuition:

$$W(t) = \begin{vmatrix} e^{\lambda t} & te^{\lambda t} \\ \lambda e^{\lambda t} & e^{\lambda t} \end{vmatrix} = e^{2\lambda t} \begin{vmatrix} 1 & t \\ \lambda & 1 \end{vmatrix}$$

The corresponding Wronskian, assuming $m \geqslant 3$ is

$$W(t) = \begin{vmatrix} e^{\lambda t} & te^{\lambda t} & t^2 e^{\lambda t} & \cdots & t^{m-1}e^{\lambda t} \\ \lambda e^{\lambda t} & e^{\lambda t} & 2te^{\lambda t} & \cdots & (m-1)t^{m-2}e^{\lambda t} \\ \lambda^2 e^{\lambda t} & \lambda e^{\lambda t} & 2e^{\lambda t} & \cdots & (m-1)(m-2)t^{m-3}e^{\lambda t} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda^{m-1}e^{\lambda t} & \lambda^{m-2}e^{\lambda t} & 2\lambda^{m-3}e^{\lambda t} & \cdots & (m-1)!e^{\lambda t} \end{vmatrix}.$$

While this seems complicated (and it is), it suffices to find *some* $t \in \mathbb{R}$ for which $W(t) \neq 0$. Now, observe that for $t = 0$ the matrix becomes is lower-triangular: indeed, each entry terms above the main diagonal contain a monomial factor, and so vanishes at $t = 0$. Meanwhile, each diagonal entry is $e^{\lambda t}$ times a positive integer. As such, $W(0) > 0$. □

3.3.2 Terminology

A general first-order linear system of DEs of dimension n (with variable coefficients) is of the form

$$\vec{x}' = A(t)\vec{x} + \vec{g}(t) \tag{3.11}$$

where $A(t) \in \mathbb{R}^{n\times n}$ and $g(t) \in \mathbb{R}^n$ are some known matrix-functions, and $\vec{x} \in \mathbb{R}^n$ is the unknown (or dependent) vector. Note that this is an "ordinary" system: there is a single *independent* variable. The system is called *homogeneous* if $g(t) \equiv 0$, and *nohomogeneous* otherwise. Let's define solutions.

Definition 9. *A (particular) solution of* (3.11) *on $I \subseteq \mathbb{R}$ is a vector-function $\vec{x}(t)$ satisfying* (3.11) *on I.*

Definition 10. *The general solution of* (3.11) *is the set of all particular solutions.*

Systems with *constant* coefficients, i.e. with $A(t) = \text{const}$, can be solved by linear algebra. We shall learn how to it in this chapter. For variable coefficients, the methods are more advanced, generalizing the method of integrating factors studied in the previous chapter. We study them later.

Example 3.3.1 (Foxes and rabbits)**.** *Foxes and rabbits live on an island. Their respective numbers $x_t(1)$, $x_2(t)$ at day t is described by the following linear system of DEs with constant coefficients:*

$$\begin{aligned} x_1' &= ax_1 + bx_2 - r, \\ x_2' &= cx_1 + dx_2 \end{aligned}$$

with parameters $c < 0$ and $a, b, d, r > 0$. Interpret this system and write it in a matrix form.

The matrix form is

$$\vec{x}' = Ax + \vec{g} \quad \text{with} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \vec{g} = \begin{pmatrix} -r \\ 0 \end{pmatrix}.$$

Interpretation: the increase of foxes at day t is the sum of *two* terms: the term proportional to the current number of foxes, plus the one proportional to the number of rabbits. Additionally, r foxes per day are removed from the island. The number of rabbits grows with the number of rabbits, but also decreases proportionally to the number of foxes. Rabbits are neither removed nor imported. □

For autonomous linear systems (i.e., with constant $A, \vec{g}$), we can adapt the notion of critical points.

Definition 11. *A* critical point *for linear system of DEs $\vec{x}' = A\vec{x} + \vec{g}$ with constant A and $\vec{g}$ is a solution to the system of linear equations $A\vec{x} = -\vec{g}$.*

For a 1st-order linear system of dimension n, the general solution will typically have n degrees of freedom—arbitrary constants $c_1, ..., c_n$—instead of just one, as it was for first-order ODEs. Equivalently, one can say that solution is defined up to an arbitrary *vector* of constants $\vec{c} \in \mathbb{R}^n$. Intuitively, this is because such a system corresponds to an nth-order ODEs, to solve which one has to "integrate n times." Accordingly, an IVP for a linear system (3.11) has initial condition of the form

$$\vec{x}(t_0) = \vec{x}_0 \tag{3.12}$$

for some initial value vector $\vec{x}_0 \in \mathbb{R}^n$.

Next, we show how to convert nth-order linear ODE (3.2) into a first-order linear system of DEs.[3]

[3]There is a method of conversion in the opposite direction as well. Try to come up with it after reading Section 3.3.3.

3.3.3 Vectorization trick

Example 3.3.2. *Convert a second-order ODE into an equivalent 1st-order linear system of DEs:*

$$u'' - u' \sin t + 7u = e^t \cos t + 1. \tag{3.13}$$

We rewrite the ODE in the standard form: $u'' = u' \sin t - 7u + e^t \cos t + 1$ and define variables

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

In terms of these variables,

$$\begin{aligned} x_1' = u' &\quad = \quad x_2, \\ x_2' = u'' &\quad = \quad u' \sin t - 7u + e^t \cos t + 1 \quad = \quad -7x_1 + x_2 \sin t + e^t \cos t + 1. \end{aligned}$$

where for x_2' we first plugged in the ODE. That is, the equivalent system reads

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -7 & \sin t \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ e^t \cos t + 1 \end{pmatrix}. \tag{3.14}$$

The first row in the matrix is $(0|1)$, and the entries of the last row are the coefficients of the ODE in the standard form, *in the reverse order.* □

Exercise 3.3.1. *Verify that ODE (3.13) and system (3.14) are indeed equivalent, in the following sense:*

(a) If $u(t)$ satisfies (3.13), then $\vec{x}(t) = \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}$ satisfies (3.14).

(b) Conversely, for any solution $\vec{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}$ of (3.14), the first component $x_1(t)$ satisfies (3.13).

General method. We now explain the general method for nth-order ODE in a standard form:

$$u^{(n)} = a_1(t)u^{(n-1)} + \cdots + a_{n-1}(t)u' + a_n(t)u + b(t).$$

We proceed as follows:

$$\begin{aligned} x_1 &= u & & & x_1' &= u' = x_2 \\ x_2 &= u' & & & x_2' &= u'' = x_3 \\ &\vdots & &\Longrightarrow & &\vdots \\ x_{n-1} &= u^{(n-2)} & & & x_{n-1}' &= u^{(n-1)} = x_n \\ x_n &= u^{(n-1)}. & & & x_n' &= u^{(n)} = a_n(t)x_1 + a_{n-1}(t)x_2 + \cdots + a_1(t)x_n + b(t). \end{aligned}$$

As the result, we obtain the system $\vec{x}' = A(t)\vec{x} + \vec{g}(t)$ with

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ a_n(t) & \cdots & \cdots & a_2(t) & a_1(t) \end{pmatrix} \quad \text{and} \quad \vec{g}(t) = \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

Here $\vec{g}(t)$ has zeroes in the first $n-1$ rows and $b(t)$ in the last row. In the first $n-1$ rows of $A(t)$, 1 moves from the 2nd to the last position; its last row has the ODE coefficients in the reverse order.

3.3.4 Existence and uniqueness of solutions in linear IVPs

We give a generalization of Theorem 1 (IVP with a 1*st*-order linear ODE) from the previous chapter.

Theorem 3.3.2. *Assume that $A(\cdot) \in \mathbb{R}^{n\times n}$ and $\vec{g}(\cdot) \in \mathbb{R}^n$ are continuous in some interval $I = (\alpha, \beta)$ containing t_0. Then the IVP $\vec{x}' = A(t)\vec{x} + \vec{g}(t)$, $\vec{x}(t_0) = \vec{x}_0$ has a unique solution for any $\vec{x}_0 \in \mathbb{R}^n$.*

We omit the proof. Conceptually, it is similar to that of Theorem 1, but uses vector calculus. The theorem might also be used for n-order ODEs, by first converting them to an equivalent system.

Example 3.3.3. *Identify the largest interval on which a solution exists and is unique, for the IVP*

$$(t-2)u'' + 3u = t, \quad u(0) = 0, \quad u'(0) = 1.$$

Dividing by $t-2$, we put equation in the standard form:

$$u'' = -\frac{3}{t-2}u + \frac{t}{t-2}.$$

Then we obtain an equivalent system:

$$\vec{x}' = \begin{pmatrix} 0 & 1 \\ -\frac{3}{t-2} & 0 \end{pmatrix} \vec{x} + \begin{pmatrix} 0 \\ \frac{t}{t-2} \end{pmatrix}.$$

Here both $A(t)$ and $\vec{g}(t)$ are continuous on $(-\infty, 2)$ and $(2, +\infty)$. The point t_0 belongs to the first interval, so the answer is $(-\infty, 2)$ by Theorem 3.3.2. □

In fact, we do not even have to convert an ODE to the system: clearly, the equivalent condition for ODE is that all coefficient functions $a_1(t), ..., a_n(t)$ and $b(t)$ are continuous on $I = (\alpha, \beta) \ni t_0$. Note also that it might be the intervals of continuity for some or all of these functions are distinct, not the same as in the previous example. Then we take all possible intersections of these intervals.

Example 3.3.4. *Identify the largest interval where the following IVP has a unique solution:*

$$u'' + \frac{1}{t-1}u = \frac{t}{t-3}, \quad u(2) = 0, \quad u'(2) = 1.$$

The coefficient $a_1(t) \equiv 0$ of u' is continuous everywhere; $a_2(t) = -\frac{1}{t-1}$ is continuous on $(-\infty, 1) \cup (1, +\infty)$; finally, $b(t) = \frac{t}{t-3}$ is continuous on $(-\infty, 3) \cup (3, +\infty)$. Hence, we have 3 candidate intervals $(-\infty, 1)$, $(1, 3)$, $(3, +\infty)$. The second interval contains $t_0 = 2$, thus is the one we seek. □

3.3.5 Linear independence of functions, Wronskian, fundamental set of solutions

Linear independence of functions. Given n functions $u_1, ..., u_n$ on some interval $I \subseteq \mathbb{R}$, we call them *linearly dependent* if there exist a tuple of constants $c_1, ..., c_n$, not all zeroes, such that

$$c_1 u_1(t) + \cdots + c_n u_n(t) = 0 \quad \forall t \in I.$$

In other words, $u_1, ..., u_n$ are dependent (on I) if some nontrivial linear combination of them is identically zero on I. Otherwise, $u_1, ..., u_n$ are *linearly independent* on I; equivalently, any nontrivial combination $c_1 u_1(t) + \cdots + c_n(t)$ is nonzero in at least some $t \in I$. To make sense of these definitions, one might view functions $I \to \mathbb{R}$ as *vectors* with continuous index $t \in I$; then the identical zero on I is the analogue of the zero vector.[4]

Example 3.3.5. *Functions t, t^2 are independent on $\mathbb{R}$. Indeed, assume that there are $c_1, c_2 \in \mathbb{R}$ such that $c_1 t + c_2 t^2 \equiv 0$ for all $t \in \mathbb{R}$. Plugging in $t_1 = 1$ and $t_2 = 2$ we get a homogeneous linear system*

$$\begin{aligned} c_1 + c_2 &= 0, \\ c_1 + 4c_2 &= 0 \end{aligned}$$

with a nonsingular matrix, hence with a unique solution $c_1 = c_2 = 0$. Thus, t, t^2 are independent. □

One can show independence for the sequence of monomials $\{1, t, t^2, ...\}$, exponentials $\{e^{\lambda t}, e^{2\lambda t}, ...\}$, trigonometric functions $\{\cos(\omega t), \cos(2\omega t), ...\}$ or $\{\sin(\omega t) \sin(2\omega t), ...\}$. This goes beyond our course.

Example 3.3.6. $\cos^2(t)$ *and* $\sin^2(t) - 1$ *are dependent (on any $I \subseteq \mathbb{R}$) since* $\cos^2(t) + \sin^2(t) - 1 \equiv 0$. □

Wronskian. Let $u_1, ..., u_n$ be $n - 1$ times differentiable on I. We can construct the square matrix

$$M(t) = \begin{pmatrix} u_1(t) & \cdots & u_n(t) \\ u_1'(t) & \cdots & u_n'(t) \\ & \vdots & \\ u_1^{(n-1)}(t) & \cdots & u_n^{(n-1)}(t) \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

Its determinant $W(t) = \det(M(t))$, as a function of $t \in I$, is called *the Wronskian* of $u_1, ..., u_n$ (at t).

Example 3.3.7. *The Wronskian of $\{t, t^2\}$ is $W(t) = \begin{vmatrix} t & t^2 \\ 1 & 2t \end{vmatrix} = t^2$. Note that $W(t) = 0$ only at $t = 0$.*

Example 3.3.8. *For $\{\cos^2(t),\ \sin^2(t) - 1\}$, the Wronskian is $W(t) \equiv 0$ on $\mathbb{R}$. Indeed,*

$$\begin{aligned} W(t) = \begin{vmatrix} \cos^2(t) & \sin^2(t) - 1 \\ -2\cos(t)\sin(t) & 2\sin(t)\cos(t) \end{vmatrix} &= 2\sin(t)\cos^3(t) + 2\sin^3(t)\cos(t) - 2\sin(t)\cos(t) \\ &= 2\sin(t)\cos(t)\ [\cos^2(t) + \sin^2(t) - 1] \quad \equiv 0. \end{aligned}$$ □

The Wroskian gives a *necessary* condition of dependence.

[4]This abides the usual definition of linear independence, in the infinite-dimensional vector space of functions on I.

Theorem 3.3.3. *If $u_1, u_2, ..., u_n$ are linearly dependent on I, then $W(t) \equiv 0$ for all $t \in I$*

Proof. The derivative of a linear combination is the same linear combination of the derivatives. So if $c_1u_1(t) + \cdots + c_nu_n(t) \equiv 0$ on I, with $c_1, ..., c_n \in \mathbb{R}$ not all 0, then $c_1u_1^{(k)}(t) + \cdots + c_nu_n^{(k)}(t) \equiv 0$ on I. That is, the columns of $M(t)$ are linearly dependent for all $t \in I$. Thus, $W(t) \equiv 0$ on I. $\square$

Theorem 3.3.3 allows to conclude that $u_1, ..., u_n$ are independent on I by finding some $t_0 \in I$ such that $W(t_0) \neq 0$. E.g. t, t^2 are independent on any open interval $I \in \mathbb{R}$ as follows from Example 3.3.7. The proof also shows why $W(t) \equiv 0$ is necessary for linear dependence, but might be *insufficient*. Indeed, the coefficients of a vanishing linear combination of the columns of $M(t)$ might depend on t, so there might not be any *constants* $c_1, ..., c_n$ that "work" for all $t \in I$ simultaneously. However, the condition actually becomes sufficient if $u_1, ..., u_n$ are *analytic* (infinitely many times differentiable).

Theorem 3.3.4 (Bôcher)**.** *If $u_1, u_2, ..., u_n$ are analytic and $W \equiv 0$ on I, then $u_1, ..., u_n$ are dependent.*

In this class, we only deal with functions analytic on their domains, so $W \equiv 0$ is a *criterion*.

Fundamental set of solutions in ODEs. If the coefficients $a_1, ..., a_n$ of a homogeneous ODE

$$u^{(n)} + a_1(t)u^{(n-1)} + \cdots + a_{n-1}(t)u' + a_n(t)u = 0 \tag{3.15}$$

are continuous on I, then corresponding IVP with $t_0 \in I$ has a unique solution. Yet, to find this solution we first have to obtain the general solution of the ODE. How do we know when we have it?

> *The set of solutions to a homogeneous linear ODE is a vector space: if $u(t)$ and $v(t)$ are solutions, then their linear combination $\alpha u(t) + \beta v(t)$ with $\alpha, \beta \in \mathbb{R}$ is also a solution.*

This vector space is n-dimensional.[5] Its arbitrary basis is called *a fundamental set of solutions* to (3.15). Therefore:

> *The general solution of a homogeneous linear ODE of order n on I is of the form*
>
> $$\sum_{k=1}^{n} c_k u_k(t), \quad t \in I,$$
>
> *where $u_1, ..., u_n$ are linearly independent solutions (a fundamental set of solutions) on I.*

If we found n linearly independent solutions, we are done. And if we have some candidates $u_1, ..., u_n$, their independence can be verified by Theorem 3.3.3: it suffices to find some $t_0 \in I$ such that $W(t_0) \neq 0$.

Fundamental set of solutions for systems of DEs. Consider an equivalent to (3.15) system:

$$\vec{x}' = A(t)\vec{x}. \tag{3.16}$$

[5] We are not proving this.

If $u_1, ..., u_n$ are solutions to (3.15), then the vector-functions $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n$ with

$$\vec{x}_k(t) = \begin{pmatrix} u_1(t) \\ u_1'(t) \\ \vdots \\ u_1^{(n-1)}(t) \end{pmatrix}$$

are solutions to (3.16), see Exercise 3.3.1. Therefore, $M(t)$ is the matrix with columns $\vec{x}_1(t), ..., \vec{x}_n(t)$, and $W(t) \neq 0$ if and only if the vectors $\vec{x}_1(t), ..., \vec{x}_n(t)$ are independent. This warrants a definition:

Definition 12. *Vector-functions $\vec{x}_1, ..., \vec{x}_n$ form a fundamental set of solutions to (3.16) on I if each of them satisfies (3.16) on I, and*

$$\det(\vec{x}_1(t_0) \cdots \vec{x}_n(t_0)) \neq 0 \text{ for some } t_0 \in I.$$

As in the case of ODEs, we conclude:

The general solution of a homogeneous 1st-order linear system of dimension n on I is

$$\sum_{k=1}^{n} c_k \vec{x}_k(t), \quad t \in I,$$

where $\vec{x}_1, ..., \vec{x}_k$ are linearly independent solutions (a fundamental set of solutions) on I.

Next, we learn how to find fundamental solutions for ODEs and systems with constant coefficients.

3.4 First-order linear systems with constant coefficients

In this section, we study homogeneous linear systems of dimension n *with constant coefficients*:

$$\vec{x}' = A\vec{x}, \tag{3.17}$$

as well as n-order ODEs that are reduced to such systems via vectorization. As it turns out, solving such systems reduces to linear algebra. We begin with a very simple but crucial observation:

If $\vec{v}$ is an eigenvector of A for eigenvalue $\lambda \in \mathbb{R}$, then $\vec{x}(t) = e^{\lambda t}\vec{v}$ is a solution to (3.17).

To see why this is the case, and explain where $e^{\lambda t}$ comes from, consider the trivial case $n = 1$. The system is then $x' = \lambda x$, with 1×1 matrix (λ). Solving it, we indeed get $x(t) = ce^{\lambda t}$ for $c \in \mathbb{R}$.

Now, in the general case $A\vec{x}(t) = Ae^{\lambda t}\vec{v} = e^{\lambda t}A\vec{v} = \lambda e^{\lambda t}\vec{v}$. But also, if v_k is the kth entry of $\vec{v}$,

$$\frac{d}{dt}\vec{x}(t) = \frac{d}{dt}\begin{pmatrix} e^{\lambda t}v_1 \\ \vdots \\ e^{\lambda t}v_n \end{pmatrix} = \begin{pmatrix} \lambda e^{\lambda t}v_1 \\ \vdots \\ \lambda e^{\lambda t}v_n \end{pmatrix} = \lambda e^{\lambda t}\vec{v}. \tag{3.18}$$

We are done. □

3.4.1 Real and distinct eigenvalues

Consider the case where A has n real eigenvalues $\lambda_1 \neq ... \neq \lambda_n$. In this case, the n corresponding (real) eigenvectors $\vec{v}_1, ..., \vec{v}_n$ are linearly independent. This gives n particular solutions to (3.17):

$$\vec{x}_1(t) = e^{\lambda_1 t}\vec{v}_1, \quad ..., \quad \vec{x}_n(t) = e^{\lambda_n t}\vec{v}_n.$$

It remains to verify that this is a fundamental set of solutions, by examining the determinant:

$$\begin{vmatrix}\vec{x}_1(t) & \cdots & \vec{x}_n(t)\end{vmatrix} = \begin{vmatrix}e^{\lambda_1 t}\vec{v}_1 & \cdots & e^{\lambda_1 t}\vec{v}_1\end{vmatrix} = e^{\lambda_1 t}\cdots e^{\lambda_n t}\begin{vmatrix}\vec{v}_1 & \cdots & \vec{v}_n\end{vmatrix} \neq 0.$$

Here we used that $\det(A)$ is linear in each column of A, then that the exponentials are positive and the vectors $v_1, ..., \vec{v}_n$ are linearly independent. Our result can be summarized as follows

Assume all eigenvalues of A are real and distinct, with respective eigenvectors $\vec{v}_1, ..., \vec{v}_n$. Then the general solution of (3.17) is

$$\vec{x}(t) = \sum_{k=1}^{n} c_k e^{\lambda_k t}\vec{v}_k. \tag{3.19}$$

Solving IVPs

Let us outline the process of solving IVP for (3.17) with initial condition $\vec{x}(t_0) = \vec{x}_0$.

(*i*) Find the general solution (3.19) by finding the eigenvalues and corresponding eigenvectors of A.

(*ii*) Form an $n \times n$ linear system in variables $c_1, ..., c_n$ by plugging the condition $\vec{x}(t_0) = \vec{x}_0$ in (3.19).

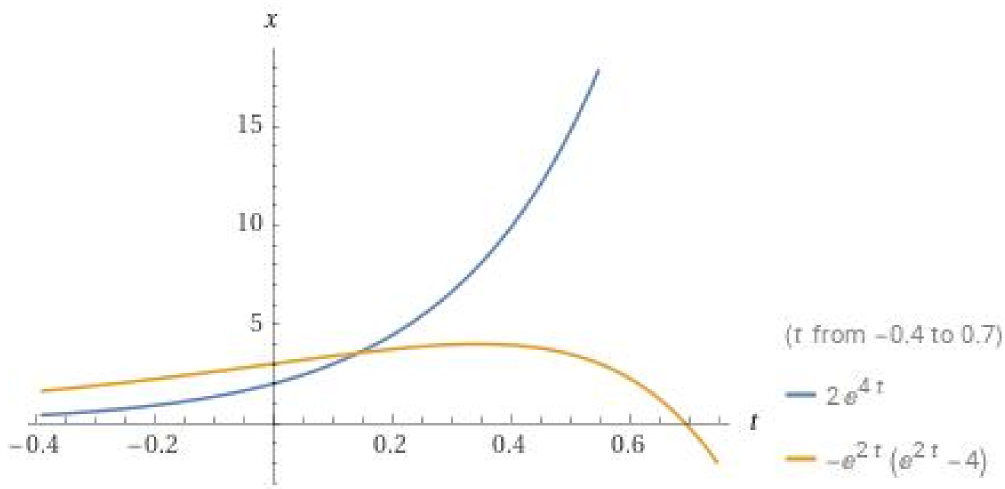

Figure 3.2: Component plot for the IVP solution in Example 3.4.1.

(iii) Identify $c_1, ..., c_n$ by solving this system. Solution is guaranteed to be unique.

Uniqueness is guaranteed since the matrix $(\vec{x}_1(t_0)\cdots\vec{x}_n(t_0))$ of the linear system in (ii) is nonsingular.

Example 3.4.1. *Solve the IVP and sketch the component plots:*

$$\vec{x}' = \begin{pmatrix} 4 & 0 \\ -1 & 2 \end{pmatrix}\vec{x}, \quad \vec{x}(\ln(2)) = \begin{pmatrix} 4 \\ 6 \end{pmatrix}.$$

We find $p_A(\lambda) = (4-\lambda)(2-\lambda)$, so the eigenvalues are $\lambda_1 = 4$, $\lambda_2 = 2$. Since $\lambda_1, \lambda_2 \in \mathbb{R}$ with $\lambda_1 \neq \lambda_2$, the corresponding eigenvectors $\vec{v}_1, \vec{v}_2$ are real and independent. Now,

$$A - \lambda_1 I = \begin{pmatrix} 0 & 0 \\ -1 & -2 \end{pmatrix} \implies \vec{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \implies \vec{x}_1(t) = e^{4t}\begin{pmatrix} 2 \\ -1 \end{pmatrix};$$
$$A - \lambda_2 I = \begin{pmatrix} 2 & 0 \\ -1 & 0 \end{pmatrix} \implies \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \implies \vec{x}_2(t) = e^{2t}\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

We now find the unknown constants c_1, c_2 from the linear system

$$\begin{pmatrix} 2e^{4t_0} & 0 \\ -e^{4t_0} & e^{2t_0} \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \vec{x}_0, \quad \text{that is} \quad \begin{pmatrix} 32 & 0 \\ -16 & 4 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 32 \\ 0 \end{pmatrix}.$$

We find $c_1 = 1$, $c_2 = 4$, and the IVP solution is $\vec{x}(t) = \vec{x}_1(t) + 4\vec{x}_2(t) = \begin{pmatrix} 2e^{4t} \\ -e^{4t} + 4e^{2t} \end{pmatrix}$. See Fig. 3.2.

Note that A might be singular, i.e. with 0 as one of the eigenvalues; the theory remains valid.

Example 3.4.2 (Compartment model)**.** *The levels of liquid in two connected tanks (Fig. 3.3) at time t satisfy*

$$\vec{x}' = \begin{pmatrix} -k & k \\ k & -k \end{pmatrix}\vec{x}$$

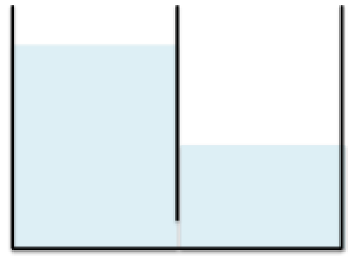

Figure 3.3: Compartment model.

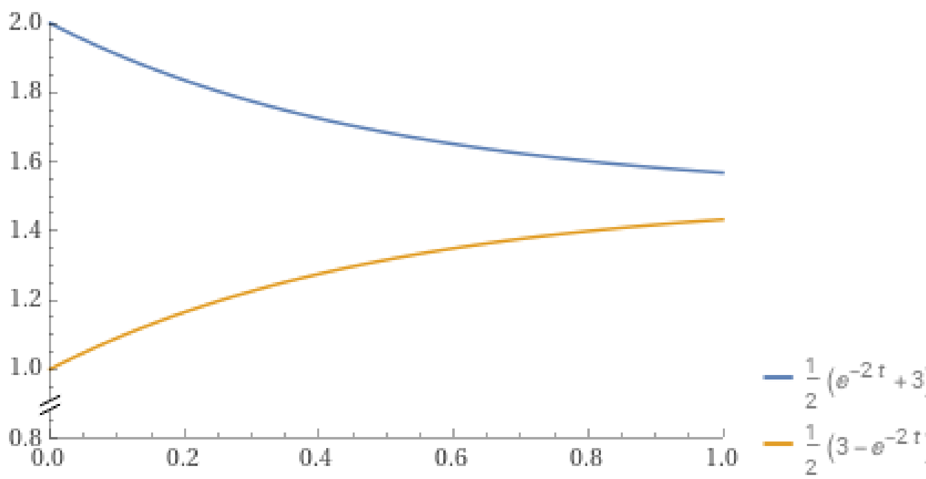

Figure 3.4: Component plot for the IVP solution in Example 3.4.2.

where $k > 0$ is a parameter depending on the liquid Solve the IVP with $\vec{x}(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and plot the component plots for the solution. Can you guess, without computation, the value of $\vec{x}(t)$ for large t?

First, we guess that the two levels should match as $t \to +\infty$, as the equilibria are $(c, c)^\top$ for $c \in \mathbb{R}$. In fact, some intuition from a high-school physics class—the liquid pressure formula $\rho g h$—might hint that the asymptotic level c is the average of the initial levels in the tanks, i.e. $c = \frac{3}{2}$. We now solve the IVP. Note that we can factor out k from the matrix, i.e. $A = kB$ with

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

The eigenvalues of B are found to be $\mu_1 = 0$, $\mu_2 = -2$, so those of A are $\lambda_1 = 0$, $\lambda_2 = -2k$. For $\vec{\lambda}_1 = 0$, we take $\vec{v}_1 = \begin{pmatrix} 1 & 1 \end{pmatrix}^\top$. Since B is symmetric, $\vec{v}_2$ is orthogonal to $\vec{v}_1$, and we take $\vec{v}_2 = \begin{pmatrix} 1 & -1 \end{pmatrix}^\top$. The initial conditions amount to $c_1 + c_2 = 2$ and $c_1 - c_2 = 1$, whence we find $c_1 = \frac{3}{2}$, $c_2 = \frac{1}{2}$, and

$$\vec{x}(t) = \frac{1}{2} \begin{pmatrix} 3 + e^{-2kt} \\ 3 - e^{-2kt} \end{pmatrix}.$$

Our guess is correct, and convergence is exponentially fast. See Fig. 3.4. □

Phase portraits for two real eigenvalues $\lambda_1 \neq \lambda_2 \neq 0$

Note that when A is nonsignular, 0 is not an eigenvalue, and the only critical point is $(0; 0)$. Its type is defined by the signs of λ_1 and λ_2. An equilibrium is asymptotically stable if both eigenvalues are negative, and unstable otherwise. In this section, we assume that $\lambda_1 > \lambda_2$ without loss of generality.

- There is a `Wolfram notebook` that allows to sketch phase portraits in two-dimensional systems!

Nodal source ($\lambda_1 > \lambda_2 > 0$). Unstable equilibrium: trajectories emanate from the origin (see Example 3.4.1). To sketch the phase portrait, we start by plotting the eigenspaces, i.e. straight lines along $\vec{v}_1$ and $\vec{v}_2$, and we put arrows in the direction from the origin. Any other trajectory emanates from the origin to infinity, remaining in its sector. It gets parallel to $\vec{v}_1$ far from the origin and tangent to $\vec{v}_2$ in the origin. In Fig. 3.5, we sketch the phase portrait of the system in Example 3.4.1.

Nodal sink ($\lambda_2 < \lambda_1 < 0$). This is an asymptotically stable equilibrium: trajectories go towards the origin. Note that we can obtain such an equilibrium from a nodal sink by negating A, see Fig. 3.5, which corresponds to time reversal. The arrows on the straight lines point towards the origin. Other trajectories go to the origin, getting parallel to $\vec{v}_2$ far from the origin and tangent to $\vec{v}_1$ in the origin.

Saddle ($\lambda_1 > 0 > \lambda_2$). This is an unstable equilibrium. The arrows point outwards along $\vec{v}_1$ and towards the origin along $\vec{v}_2$. All other trajectories are U-shaped: they approach the origin up to a certain point, then run away from it. Far away from the origin they get parallel to $\vec{v}_1$ or $\vec{v}_2$, and the arrows conform to those on the straight lines. In Fig. 3.5, we sketch the phase portrait for $\vec{x}' = A\vec{x}$ with eigenvalues $\lambda_1 = 8$ and $\lambda_2 = -2$, and respective eigenvectors $\vec{v}_1 = (-6; 1)$ and $\vec{v}_2 = (4; 1)$.

3.4.2 Complex eigenvalues (without repetition)

Complex algebra. Euler's formula extends the exponential to imaginary numbers: *by definition,*

$$e^{i\theta} = \cos(\theta) + i\sin(\theta) \quad \forall \theta \in \mathbb{R}.$$

Then one can extend e^z to $z = x + iy \in \mathbb{C}$: *by definition,* $e^z := e^x e^{iy} = e^x\cos(y) + ie^x\sin(y)$. This is a very natural generalization as it preserves the key algebraic property of exponentials, namely that

$$e^{z_1+z_2} = e^{z_1}e^{z_2} \quad \text{for all} \;\; z_1, z_2 \in \mathbb{C}.$$

(This can be checked by Euler's formula and the formulas for $\cos(\alpha \pm \beta)$ and $\sin(\alpha \pm \beta)$.) Moreover, we can define the derivative of $\varphi : \mathbb{R} \to \mathbb{C}$ by differentiating $\mathsf{Re}\varphi$ and $\mathsf{Im}\varphi$ separately. Then, for example, $(e^{i\omega t})' = (\cos(\omega t) + i\sin(\omega t))' = -\omega\sin(\omega t) + i\omega\cos(\omega t) = i\omega e^{i\omega t}$. More generally,

$$(e^{\lambda t})' = \lambda e^{\lambda t} \quad \text{for all} \quad \lambda \in \mathbb{C}, \tag{3.20}$$

so the key differential property of the exponential function is also preserved. To summarize, "the usual algebra and analysis" still work for $t \mapsto e^{\lambda t}$ with $\lambda \in \mathbb{C}$, plus we can exploit Euler's formula.

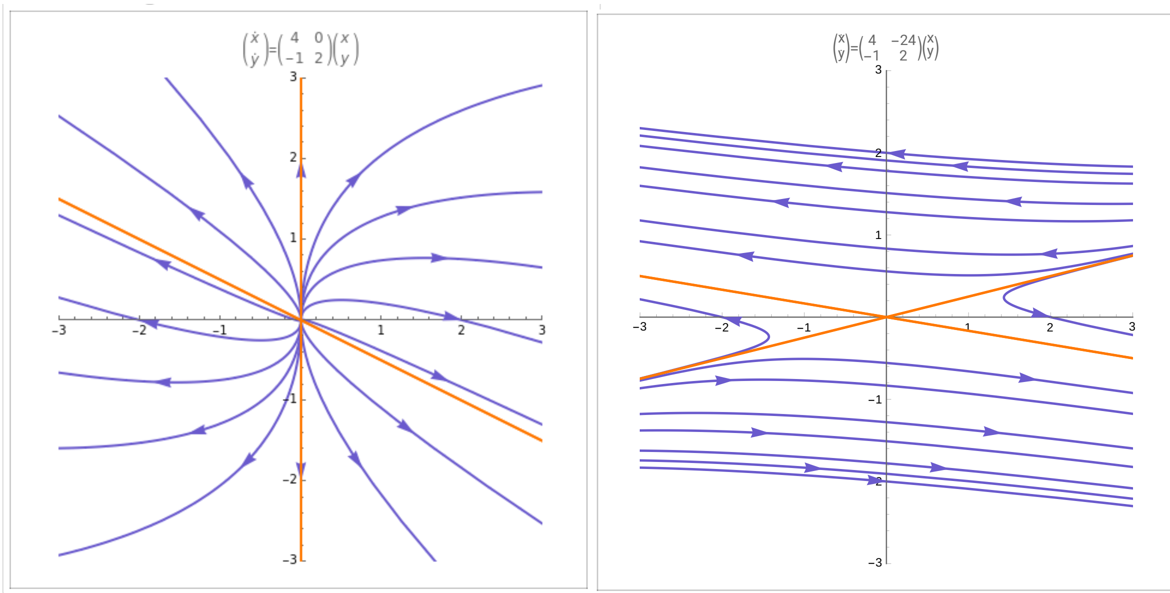

Figure 3.5: **Left:** nodal source (Example 3.4.1). In a nodal sink, the arrows are flipped. **Right:** saddle.

Case $n = 2$

We now consider the case of two complex eigenvalues in dimension $n = 2$. Recall from Section 2.1 that complex eigenvalues (as well as eigenvectors) of a *real* matrix come in mutually conjugate pairs:

$$\begin{aligned} \lambda_1 &= \mu + i\omega, \quad \vec{v}_1 = \vec{a} + i\vec{b}, \\ \lambda_2 &= \mu - i\omega, \quad \vec{v}_2 = \vec{a} - i\vec{b}, \end{aligned} \tag{3.21}$$

for some $\mu, \omega \in \mathbb{R}$; $\vec{a}, \vec{b} \in \mathbb{R}^n$. This holds for all n, but in the case $n = 2$ we have no other eigenvalues, and we should be able to obtain a fundamental set of *real* solutions for (3.17). We prove the following:

Proposition 3.4.1. *System (3.17) with 2 complex eigenvalues (3.21) has a fundamental set of solutions*

$$\boxed{\begin{aligned} \vec{u}(t) &= e^{\mu t}\cos(\omega t)\vec{a} - e^{\mu t}\sin(\omega t)\vec{b}, \\ \vec{v}(t) &= e^{\mu t}\sin(\omega t)\vec{a} + e^{\mu t}\cos(\omega t)\vec{b}. \end{aligned}}$$

As such, the general solution is $\vec{x}(t) = \alpha\vec{u}(t) + \beta\vec{v}(t)$ *for* $\alpha, \beta \in \mathbb{R}$.

Example 3.4.3. *Solve the IVP*

$$\vec{x}' = \begin{pmatrix} -1 & 2 \\ -1 & -3 \end{pmatrix}\vec{x}, \quad \vec{x}(\pi) = e^{-2\pi}\begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Here $p_A(\lambda) = \lambda^2 + 4\lambda + 5$, whence $\lambda_{1,2} = -2 \pm i$. The respective eigenvectors can be computed as in (2.1.2), part (b); we will get

$$\vec{v}_{1,2} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \pm i\begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In terms of (3.21), $\mu = -2$, $\omega = 1$, and we find the general solution

$$\vec{x}(t) = e^{-2t}\cos t \begin{pmatrix} 2c_1 \\ c_2 - c_1 \end{pmatrix} + e^{-2t}\sin t \begin{pmatrix} 2c_2 \\ -c_1 - c_2 \end{pmatrix} \quad \forall c_1, c_2 \in \mathbb{R}.$$

From the initial condition we get $\begin{pmatrix} -2c_1 \\ c_1 - c_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, whence $c_1 = -\frac{1}{2}$, $c_2 = \frac{1}{2}$, and the solution is

$$\vec{x}(t) = e^{-2t}\cos t \begin{pmatrix} -1 \\ 1 \end{pmatrix} + e^{-2t}\sin t \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \qquad \square$$

Proof of Proposition 3.4.1. By (3.20), $y(t) = e^{\lambda t}$ satisfies a homogeneous ODE $y' - \lambda y = 0$ with *complex* coefficient $\lambda \in \mathbb{C}$. Therefore, repeating (3.18), we verify that

$$\vec{x}_1(t) = e^{\lambda_1 t}\vec{v}_1, \quad \vec{x}_2(t) = e^{\lambda_2 t}\vec{v}_2.$$

are solutions to (3.17). However, they are complex, and we need real solutions. Now, we check that

$$\begin{aligned} \vec{x}_1(t) &= \vec{u}(t) + i\vec{v}(t), \\ \vec{x}_2(t) &= \vec{u}(t) - i\vec{v}(t), \end{aligned}$$

that is $\vec{u}(t) = \mathsf{Re}\vec{x}_1(t)$ and $\vec{v}(t) = \mathsf{Im}\vec{x}_1(t)$. Indeed, by Euler's formula

$$\begin{aligned} \vec{x}_1(t) = e^{\mu t}[\cos(\omega t) + i\sin(\omega t)](\vec{a} + i\vec{b}) &= e^{\mu t}[\cos(\omega t)\vec{a} - \sin(\omega t)\vec{b} + i\sin(\omega t)\vec{a} + i\cos(\omega t)\vec{b}] \\ &= \vec{u}(t) + i\vec{v}(t), \end{aligned}$$

similarly for $\vec{x}_2(t)$. This explains where $\vec{u}(t), \vec{v}(t)$ come from, and also shows that they are solutions:

$$\frac{d}{dt}\vec{u}(t) = \frac{d}{dt}\mathsf{Re}(\vec{x}_1(t)) \overset{(i)}{=} \mathsf{Re}\left(\frac{d}{dt}\vec{x}_1(t)\right) \overset{(ii)}{=} \mathsf{Re}(A\vec{x}_1(t)) \overset{(ii)}{=} A\mathsf{Re}(\vec{x}_1(t)) = A\vec{u}(t).$$

Make sure you understand why (i)–(iii) hold. Finally, to verify that $\vec{u}, \vec{v}$ are independent, note that

$$\begin{pmatrix} \vec{u} & \vec{v} \end{pmatrix} = ze^{\mu t}\begin{pmatrix} \vec{a} & \vec{b} \end{pmatrix}\begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix}$$

where the rotation matrix in the right-hand side has determinant 1. Hence $\det\begin{pmatrix} \vec{u} & \vec{v} \end{pmatrix} = e^{2\mu t}\det\begin{pmatrix} \vec{a} & \vec{b} \end{pmatrix}$, where $e^{2\mu t} > 0$ since $\mu \in \mathbb{R}$. But if we assume that $\det\begin{pmatrix} \vec{a} & \vec{b} \end{pmatrix} = 0$, then $\vec{v}_1, \vec{v}_2$ must also be dependent due to (3.21), and this contradicts their linear independence (over $\mathbb{C}$) as they correspond to $\lambda_1 \neq \lambda_2$. $\square$

3.4.3 Case $n > 2$.

Here we do not consider the case of more than 2 complex eigenvalues in detail. In a nutshell, any pair of simple complex eigenvalues (i.e. with algebraic multiplicity 1) gives a pair of solutions as in Proposition 3.4.1. If $n > 2$ and a pair has multiplicity $k > 1$, then one can find solutions by combining Proposition 3.4.1 with the Jordan decomposition trick for repeated eigenvalues, to be presented later.

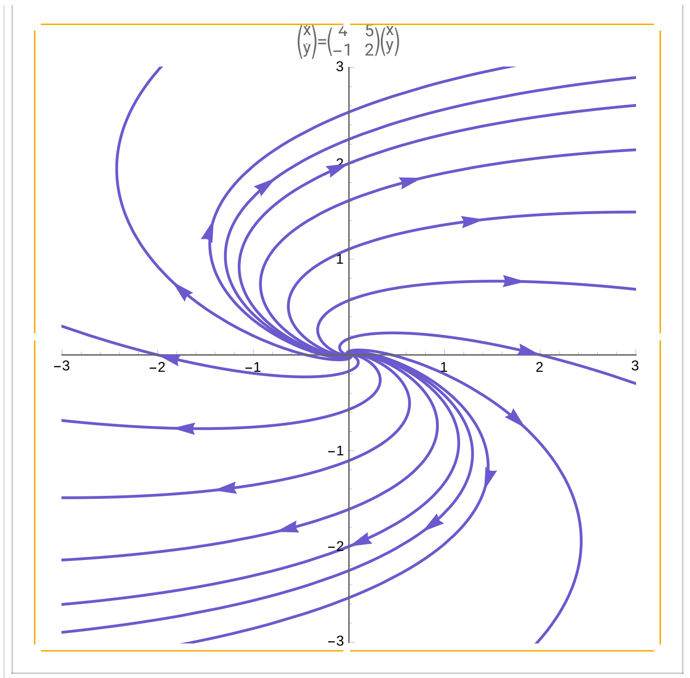

Figure 3.6: Phase portrait for $\vec{x}' = A\vec{x}$ where A has eigenvalues $3 \pm 2i$, so $(0; 0)$ is a spiral source.

Phase portraits for two complex eigenvalues

When the eigenvalues are complex, there are three cases, depending on the sign of their real part μ.

Spiral source ($\mu > 0$) – Spiral sink ($\mu < 0$). All trajectories are spirals emanating from (resp. going to) the origin for a source (resp. sink); source is unstable, and sink is stable. (Do you understand why?) In Fig. 3.5, we sketch the phase portrait of a system with a spiral source, and in Example 3.4.3 the equilibrium is a spiral sink. When sketching a spiral source/sink, the only dilemma is to guess the direction of rotation: clockwise or counterclockwise. This can be done by trying a couple of "simple" test values for $\vec{x}$, e.g. $(1; 0)$ and $(0; 1)$, computing the corresponding values of the flow $\vec{x}'$, and checking if they are consistent with the tentative direction of rotation. For the system

$$\vec{x}' = \begin{pmatrix} 4 & 5 \\ -1 & 2 \end{pmatrix} \vec{x}$$

in Figure 3.6, the eigenvalues are $3 \pm 2i$, so we have a spiral source. From $(1; 0)$ we move in the direction $(4; -1)$, and from $(0; 1)$ in the direction $(5; 2)$; this is consistent with the clockwise rotation.

Circular case ($\mu = 0$). If $\mu = 0$, i.e. both eigenvalues are imaginary, all trajectories are ellipses, and the direction of rotation can be determined by the same method as for spiral source/sink. In fact this is a "neutral" equilibrium: it does not attract nor repell.

3.4.4 Repeated eigenvalues

We now consider the general situation of an n-dimensional system with eigenvalues $\lambda_1, ..., \lambda_n$ that might repeat. The following example demonstrates that this is a practically relevant situation.

Example 3.4.4 (Reversive motion)**.** *The motion of a point in the plane is described by the equation*

$$x' = -x + ky,$$
$$y' = -y,$$

where $k \in \mathbb{R}$ is a parameter, with initial condition $\vec{r}(0) = (1;\, 2)$. Find its position $\vec{r}(t)$ at any $t \in \mathbb{R}$.

Let us deal with this example "ad-hoc," then study the general case. Our equation is $\vec{r}' = Ar$ with

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

The matrix is upper-triangular, and diagonal when $k = 0$. For $k \neq 0$, the variables x, y are coupled.

Diagonal case: $k = 0$. Here $A = -I$, with eigenvalues $\lambda_{1,2} = -1$. Moreover, $A - (-1)I = 0$, so any $\vec{v} \in \mathbb{R}^2$ is an eigenvector; in particular, $\vec{v}_1 = (1; 0)$ and $\vec{v}_2 = (0; 1)$ is a pair of independent ones. On the other hand, in our system with $k = 0$, each of the two equations only concerns its own variable x or y, so these variables do not interact; thus, we can solve the two equations *separately.* The corresponding general solution is given by $x(t) = c_1 e^{-t}$ and $y(t) = c_2 e^{-t}$ for $c_1, c_2 \in \mathbb{R}$, that is

$$\vec{r}(t) = c_1 e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Nondiagonal case: $k \neq 0$. Note that the eigenvalues of A are the same as before: $\lambda_{1,2} = -1$. (This is, in fact, a general result for upper- and lower-triangular matrices.) However, the nullspace of

$$A + I = \begin{pmatrix} 0 & k \\ 0 & 0 \end{pmatrix}$$

has dimension 1; namely, it is the span of $\vec{v} = (1; 0)$. On the other hand, we can solve our system by substitution. Namely, we first solve the second equation that only concerns y; its general solution is $y(t) = ce^{-t}$. Plugging $y(t)$ in the first equation, we get a parameterized first-order ODE in x:

$$x' + x = cke^{-t},$$

that can be solved via the integrating factor method. Doing so, we get $x(t) = (c_1 + ckt)e^{-t}$. Introducing $\vec{w}_1 = (0; \frac{1}{k})$ and $c_2 = ck$, we express the answer in the vector form:

$$\begin{aligned} \vec{r}(t) &= e^{-t} \begin{pmatrix} c_1 + ckt \\ c \end{pmatrix} \\ &= c_1 e^{-t} \vec{v} + c_2 e^{-t} \left(t\vec{v} + \vec{w} \right). \end{aligned}$$

We also note that eigenvalue $\lambda = -1$, respective eigenvector $\vec{v}$, and the vector $\vec{w}$ satisfy the relation

$$(A - \lambda)\vec{w} = \vec{v}.$$

From linear algebra we recall that $\vec{w}$ is called (the first) generalized eigenvector of eigenvalue λ. □

Generalized eigenvectors. Before we proceed further to the general theorem, let's recall a result from linear algebra (see also the beginning of this chapter). Assume A is a square matrix of dimension n, and its characteristic polynomial p_A has some root $\lambda \in \mathbb{C}$ with multiplicity $m \leqslant n$. Then $m(\lambda)$ is called the *algebraic multiplicity* of eigenvalue λ, whereas $s(\lambda) = \dim(\mathsf{Null}(A - \lambda I))$ is the *geometric multiplicity* of eigenvalue λ. For any eigenvalue λ, it holds that $1 \leqslant s(\lambda) \leqslant m(\lambda) \leqslant n$, and the sum of algebraic multiplicities over all (distinct) eigenvalues of A is n. If it happens, for A at hand, that $s(\lambda) = m(\lambda)$, then the subspace $\mathsf{Null}(A - \lambda I)$ has a basic of eigenvectors, and otherwise it's not the case; these two situations corresponded, respectively, to $k = 0$ and $k = 1$ in the above example. We now recall Jordan's theorem from linear algebra – first, for the case of $s(\lambda) = 1$.

Theorem 3.4.1. *Let λ be an eigenvalue of A with algebraic multiplicity $m \geqslant 2$ and single independent eigenvector $\vec{v}$. There exist m independent vectors $\vec{w}_0[= \vec{v}], \vec{w}_1, ..., \vec{w}_{m-1}$ that form the* Jordan chain*:*

$$
\begin{aligned}
(A - \lambda I)\vec{w}_0 &= 0,\\
(A - \lambda I)\vec{w}_1 &= \vec{w}_0,\\
&\vdots\\
(A - \lambda I)\vec{w}_{m-1} &= \vec{w}_{m-2}.
\end{aligned}
$$

Moreover, the chain cannot be continued: the system $(A - \lambda I)\vec{u} = \vec{w}_{m-1}$ has no solutions in $\vec{u} \in \mathbb{C}^n$.

In the previous example with $k \neq 0$, for $\lambda = -1$ we have $m = 2$ and $s = 1$; the Jordan chain is comprised of $\vec{w}_0 = \vec{v} = (1; 0)$ and $\vec{w}_1 = \vec{w} = (0; \frac{1}{k})$, and there exists no $\vec{u}$ such that $(A - \lambda I)\vec{u} = \vec{w}_1$. We can construct a Jordan chain for λ with $m \geqslant 2$ and $s = 1$ as follows.

1. Find some eigenvector $\vec{v}$ by solving $(A - \lambda I)\vec{v} = 0$, and let $\vec{w}_0 = \vec{v}$
2. Repeat the following for $k \in \{1, ..., m - 1\}$: given $\vec{w}_{k-1}$, find $\vec{w}_k$ by solving $(A - \lambda I)\vec{u} = \vec{w}_{k-1}$.

In fact, it is guaranteed that the solution in step 2 is unique for each $k \in \{1, ..., m-1\}$; in particular, we can rescale the whole chain by the same scalar, but not its vectors separately from each other.

Remark 3.4.1. *We briefly discuss the case of $1 < s < m$ (which we shall not encounter in this class). In this case, there are s independent eigenvectors $\vec{v}_1, ..., \vec{v}_s$ and s respective Jordan chains, each starting from its own eigenvector. The sum of lengths of these chains is m. To find these chains, we can run the above process for each chain incrementally, cutting it when we cannot find the next link.*

Returning to systems of DEs, we have the following general result.

Theorem 3.4.2. *Let $A \in \mathbb{C}^{n \times n}$ have an eigenvalue λ with algebraic multiplicity m and geometric multiplicity s, with the corresponding Jordan chains*

$$
\left(\vec{w}_0^{(1)}, ..., \vec{w}_{m_1-1}^{(1)}\right); \quad \cdots; \quad \left(\vec{w}_0^{(s)}, ..., \vec{w}_{m_1-1}^{(s)}\right).
$$

Then $\vec{x}\,' = A\vec{x}$ has m solutions of the form

$$
e^{\lambda t}\vec{w}_0^{(j)}, \quad e^{\lambda t}\left(t\vec{w}_0^{(j)} + \vec{w}_1^{(j)}\right), \quad ..., \quad e^{\lambda t}\left(\sum_{k=1}^{m_j} t^{m_j - k}\vec{w}_{k-1}^{(j)}\right) \quad \text{for } j \in \{1, ..., s\}.
$$

The resulting n solutions (for all eigenvalues) are independent, so give a fundamental set of solutions.

Proof. We only consider the simplest case to highlight the mechanism; the general case can be handled by induction. Namely, assume $m = 2$ and $s = 1$, and let $\vec{v}, \vec{w}$ be the Jordan chain of λ. We already know that $\vec{x}_0(t) = e^{\lambda t}\vec{v}$ is a solution of $\vec{x}' = A\vec{x}$, so it remains to verify that $\vec{x}_1(t) = e^{\lambda t}(t\vec{v}+\vec{w})$ is a solution as well, and that the two are independent. For the first claim, we observe that

$$A\vec{x}_1(t) = e^{\lambda t}(tA\vec{v} + A\vec{w}) = e^{\lambda t}(\lambda t\vec{v} + \lambda\vec{w} + \vec{v}) = \vec{x}'(t).$$

As for the second claim, it follows from the independence of $\vec{v}$ and $\vec{w}$. Indeed, the Wronskian reads

$$W(t) = \begin{vmatrix}\vec{x}_0(t) & \vec{x}_1(t)\end{vmatrix} = \begin{vmatrix}e^{\lambda t}\vec{v} & e^{\lambda t}(t\vec{v} + \vec{w})\end{vmatrix} = e^{2\lambda t}\begin{vmatrix}\vec{v} & t\vec{v} + \vec{w}\end{vmatrix} = e^{2\lambda t}\begin{vmatrix}\vec{v} & \vec{w}\end{vmatrix} \neq 0. \qquad \square$$

Bibliography

[BB15] J. R. Brannan and W. E. Boyce. *Differential equations: An introduction to modern methods and applications*. John Wiley & Sons, 2015.

Appendix A

Recap of calculus on $\mathbb{R}$

A.1 Limits

Composition rule. Let f, g, φ be functions on $\mathbb{R}$. The key property of the limit operation is that

$$\varphi\left(\lim_{x\to a} f(x)\right) = \lim_{x\to a} \varphi(f(x)),$$

provided that the limit in the LHS exists and φ is continuous at $f(a)$. This generalizes to the multivariate situation where $f : \mathbb{R}^n \to \mathbb{R}^m$ and $\varphi : \mathbb{R}^m \to \mathbb{R}^k$. For example, this implies the rules

$$\lim_{x\to a} f(x) + g(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x), \qquad \lim_{x\to a} f(x)g(x) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$$

where $x \in \mathbb{R}$, and the LHS exists whenever the RHS does.

A.2 Differentiation

Definitions. Recall that the **derivative** of f at x is defined as

$$f'(x) := \lim_{\delta\to 0} \frac{f(x+\delta) - f(x)}{\delta}$$

whenever the limit in the RHS exists; we also say that f is **differentiable at** x if this is the case. We can also define the right and left derivatives $f'(x+0), f'(x-0)$ by requiring that $\delta \to +0$ or $\delta \to -0$ respectively. Geometrically, $f'(x)$ is the *slope* (of the tangent line) of the graph of f at x. If f is differentiable at x, then the graph is "smooth" at x, so the tangent line is unique (and not vertical). If, say, $f'(x+0)$ and $f'(x-0)$ both exist but are different, then the graph has a "kink" at x, so there tangent line is not unique, and f is *not* differentiable at x (e.g. $f(x) = |x|$ at $x = 0$).

Common derivatives:

- Constant: $c' = 0$ for any $c \in \mathbb{R}$.
- Powers: $(x^p)' = px^{p-1}$ for $p \neq 0$.
- Exponential: $(e^x)' = e^x$.

- Natural logarithm: $(\ln(x))' = 1/x$ for $x > 0$ and $(\ln(-x))' = 1/x$ for $x < 0$.
- Trigonometric functions:[1]

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x),$$
$$\tan'(x) = \frac{1}{\cos^2(x)}, \quad \cot'(x) = -\frac{1}{\sin^2(x)}.$$

Chain rule:

$$(\ \varphi(f(x))\ )' = \varphi'(f(x)) \cdot f'(x).$$

Product rule:

$$(fg)' = f'g + fg'.$$

Quotient rule:

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

Combining the chain and product rules with common derivatives, we can handle complicated functions, or refresh our memory. For example, if we remember that $\sin'(x) = \cos(x)$, we verify that

$$(\cos(x))' = \left(\sin\left(\frac{\pi}{2} - x\right)\right)' = \left(\frac{\pi}{2} - x\right)' \cdot \sin'\left(\frac{\pi}{2} - x\right) = (-1) \cdot \cos\left(\frac{\pi}{2} - x\right) = -\sin(x).$$

Another exercise: recover the quotient rule from the product and chain rules by introducing $h = 1/g$. Other useful formulas that can be obtained in this fashion are

$$(a^x)' = \ln(a) \cdot a^x, \quad (\log_a(x))' = \frac{1}{x \ln(a)}.$$

Inverse function. Recall that function φ is the *inverse* of function f on I (denoted $\varphi = f^{-1}$) if

$$\varphi(f(x)) = x \quad \forall x \in I.$$

That is: if f maps $x \in I$ to some $y \in J$, then $\varphi = f^{-1}$ maps y back into x. Drawing the plot of f^{-1} amounts to "mirror-reflecting" the plot of f across $y = x$ (or swapping x and y). E.g., $\varphi(y) = y^{1/3}$ is the inverse of $f(x) = x^3$ on $\mathbb{R}$; more generally, $f(x) = x^p$ has the inverse $\varphi(y) = y^{1/p}$ on $I = (0, +\infty)$, and on $I = \mathbb{R}$ when $p \in \{1, 3, 5, ...\}$. Recall also that $\ln(\cdot)$ is *defined* as the inverse function of e^x:

$$\ln(e^x) = x,$$

see Fig. A.1. Differentiating both sides, we derive the formula $\ln'(y) = 1/y$ from $(e^x)' = e^x$. More generally, differentiating $\varphi(f(x)) = x$ we obtain an explicit formula for the derivative of the inverse:

$$\varphi'(y) = \frac{1}{f'(\varphi(y))}.$$

[1] Recall that $\tan(x) := \frac{\sin(x)}{\cos(x)}$ and $\cot(x) := \frac{1}{\tan(x)} = \frac{\cos(x)}{\sin(x)}$.

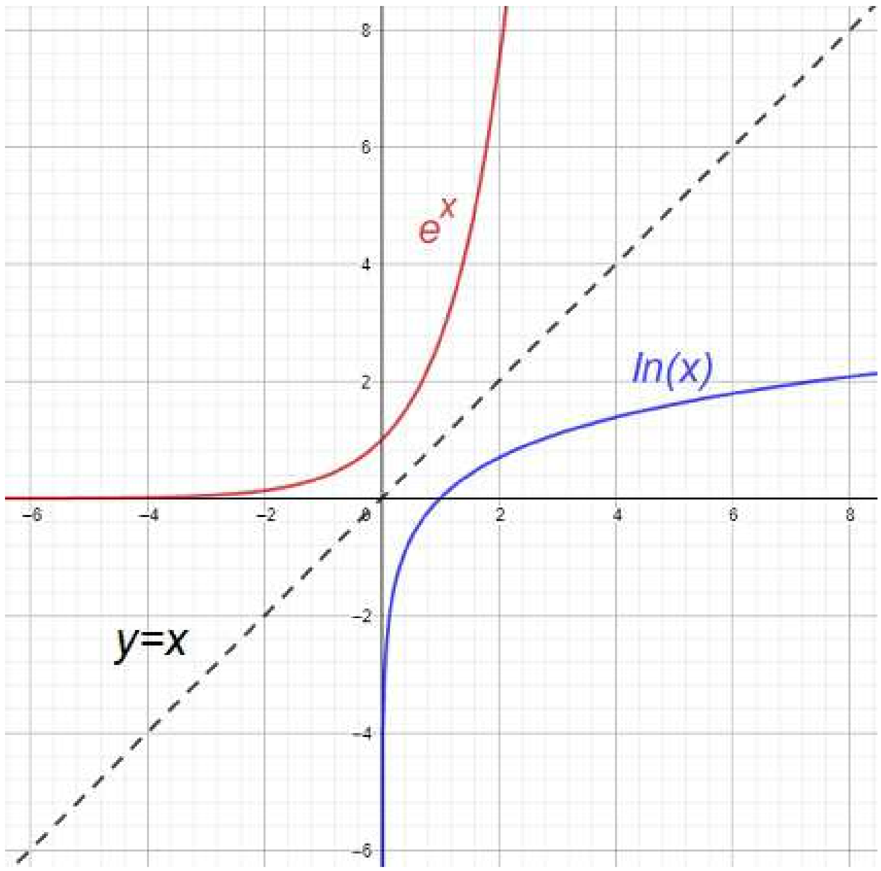

Figure A.1: $\varphi(\cdot) = \ln(\cdot)$ is the inverse function for $f(x) = e^x$ on $\mathbb{R}$.

A.3 Integration

There are two different notions of integral, connected through the fundamental theorem of calculus: *indefinite integrals (a.k.a. antiderivatives)* and *definite integrals.*

Indefinite integrals (antiderivatives). We say that F is an *antiderivative* of f on I when

$$F'(x) = f(x) \quad \forall x \in I.$$

In other words, finding an antiderivative is the inverse operation for differentiation. Note that if F is an antiderivative for f, then so is $f + c$ for any $c \in \mathbb{R}$. Moreover, we do not "lose" anything: the set of *all* antiderivatives of f is given by $\{F(x) + c | c \in \mathbb{R}\}$, where one can choose F as any *specific* antiderivative of f. Sometimes, we cannot afford to "lose" the constant, as we need *all possible* antiderivative, not just a *specific* one. In this case, we write

$$\int f(x)dx = F(x) + C \quad (\forall C \in \mathbb{R})$$

and call the LHS *the (indefinite) integral* of f. Specification $\forall C \in \mathbb{R}$ is often omitted – but assumed.

Common integrals:

- Exponential: $\int (e^x)dx = e^x + C$ and $\int a^x dx = \frac{1}{\ln(a)}a^x + C$.
- Powers: $\int x^r dx = \frac{1}{r+1}x^{r+1} + C$ for $r \neq -1$.
- Linear reciprocal: $\int \frac{dx}{x} = \ln|x| + C$, or more generally

$$\int \frac{dx}{x+a} = \ln|x+a| + C. \tag{A.1}$$

- Positive quadratic reciprocal ($a \neq 0$):

$$\int \frac{dx}{x^2+a^2} = \frac{1}{a}\arctan\left(\frac{x}{a}\right) + C. \tag{A.2}$$

- Rational functions: $\int \frac{P(x)}{Q(x)}dx$, where P, Q are polynomials. There is no explicit formula, but such integrals can be found by the **method of partial fractions**, as explained below.
- Trigonometric functions:

$$\int \sin(x)dx = -\cos(x) + C, \quad \int \cos(x)dx = \sin(x) + C,$$
$$\int \frac{dx}{\cos^2(x)} = \tan(x) + C, \quad \int \frac{dx}{\sin^2(x)} = -\cot(x) + C.$$

  Many trigonometric integrals can be done by combining these with **change of variable** and trigonometric identities (see `en.wikipedia.org/wiki/List_of_trigonometric_identities`).

Change of variable (integration by substitution). This is essentially the reverse chain rule:

$$\int f(g(x))g'(x)dx = F(g(x))$$

where $F(u) = \int f(u)du$. It allows to evaluate the integral in the LHS if we know how to integrate f. Note also that the formula becomes transparent in Leibniz notation, i.e. by writing $dg(x) = g'(x)dx$.

Fundamental theorem of calculus. This theorem connects *definite* integral with antiderivative:

$$\int_a^b f(x)dx = F(b) - F(a).$$

A.3.1 Integration by partial fractions

Reciprocal of a quadratic. Let us first explain how to evaluate an integral of the form

$$\int \frac{dx}{ax^2+bx+c}. \tag{A.3}$$

By the fundamental theorem of algebra,

$$ax^2 + bx + c = a(x - x_1)(x - x_2)$$

where x_1, x_2 are the *roots* of $ax^2 + bx + c$, i.e. solutions of the equation $ax^2 + bx + c = 0$. We assume that the roots are real. Multiplying for convenience by a, we get

$$\int \frac{a}{ax^2 + bx + c} dx = \int \frac{dx}{(x - x_1)(x - x_2)}, \tag{A.4}$$

and our task reduces to evaluating the RHS. Now, if $x_1 = x_2$, then $\int \frac{dx}{(x-x_1)(x-x_2)} = \int \frac{dx}{(x-x_1)^2}$ is reduced to the integral $\int \frac{1}{t^2} dt = -\frac{1}{t} + C$ by the change of variables $t = x - x_1$. Hence, we can assume that $x_1 \neq x_2$ and use the method of partial fractions: find the coefficients A_1, A_2 such that

$$\frac{1}{(x - x_1)(x - x_2)} \equiv \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2}. \tag{A.5}$$

Essentially, this operation is reverse for reducing to a common denominator: A_1 and A_2 must satisfy

$$A_1(x - x_2) + A_2(x - x_1) = 1$$

for all x, in particular $x \in \{x_1, x_2\}$. Plugging these values of x, we find $A_1 = \frac{1}{x_1 - x_2}$ and $A_2 = \frac{1}{x_2 - x_1}$. Thus, by (A.4)–(A.5) we reduce the reciprocal-quadratic integral (A.3) to a sum of integrals of the form (A.1).

Example A.3.1. *Compute* $\int \frac{dx}{x^2+3x+2}$.

Proof. We find the roots $x_1 = -1$ and $x_2 = 2$, by a lucky guess or the discriminant formula. Thus,

$$x^2 + 3x + 2 = (x + 1)(x - 2),$$

and we can find A_1, A_2 such that

$$\frac{1}{x^2 + 3x + 2} = \frac{1}{(x + 1)(x - 2)} = \frac{A_1}{x + 1} + \frac{A_2}{x - 2}$$

by solving

$$A_1(x - 2) + A_2(x + 1) = 1.$$

Plugging in $x = -1$ and $x = 2$, we recover $A_1 = -1/3$ and $A_2 = 1/3$, and then compute the integral:

$$\int \frac{dx}{x^2 + 3x + 2} = \int \left(\frac{A_1}{x + 1} + \frac{A_2}{x - 2} \right) dx = \frac{1}{3} \left(\int \frac{dx}{x - 2} dx - \int \frac{dx}{x + 1} dx \right) = \frac{1}{3} \ln \left(\frac{|x - 2|}{|x + 1|} \right) + C.$$

□

Reciprocal of a polynomial. The method of partial fractions can be generalized for computing

$$\int \frac{dx}{Q(x)} \tag{A.6}$$

where Q is a polynomial of degree n. Assuming that the coefficient of x^n in Q is 1, it holds that

$$Q(x) = (x - x_1)(x - x_2) \cdots (x - x_n)$$

where $x_1, ..., x_n$ are the roots of Q. We can then try to find $A_1, ..., A_n$ that ensure the identity

$$\frac{1}{(x - x_1)(x - x_2) \cdots (x - x_n)} \equiv \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \cdots + \frac{A_n}{x - x_n}$$

or, equivalently,

$$A_1(x - x_2) \cdots (x - x_n) + A_2(x - x_1)(x - x_3) \cdots (x - x_n) + \cdots + A_n(x - x_1) \cdots (x - x_{n-1}) \equiv 1.$$

The LHS is a polynomial of degree $n - 1$, and equating its coefficients to $(1, 0, ..., 0)$ gives a system of n *linear equations* in $A_1, ..., A_n$. In fact, one can show that this system will have a unique solution when the roots are distinct: $x_1 \neq ... \neq x_n$. Solving it, we recover $A_1, ..., A_n$ and so reduce (A.6) to (A.1).

Rational functions. Finally, let us see how to integrate a rational function, i.e. compute

$$\int \frac{P(x)}{Q(x)} dx. \tag{A.7}$$

where P and Q are polynomals, respectively, of degrees m and n. We can assume that $m \geqslant 1$; otherwise we are in the previous case. We also assume, as before, that the highest-degree coefficient in Q is 1 and its roots $x_1, ..., x_n$ are distinct. Proceeding as before, we can find $A_1, ..., A_n$ such that

$$\frac{1}{Q(x)} \equiv \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} + \cdots + \frac{A_n}{x - x_n}.$$

Multiplying by $P(x)$ we get

$$\frac{P(x)}{Q(x)} \equiv A_1 \frac{P(x)}{x - x_1} + A_2 \frac{P(x)}{x - x_2} + \cdots + A_n \frac{P(x)}{x - x_n}.$$

Now recall that any polynomial $P(x)$ of degree $m \geqslant 1$ can be divided over linear polynomial $x - x_0$ with remainder of the form $\frac{r_0}{x - x_0}$. That is, for any $x_0 \in \mathbb{C}$ and polynomial $P(x)$ of degree m, one has

$$\frac{P(x)}{x - x_0} = R(x) + \frac{r_0}{x - x_0}$$

with some $r_0 \in \mathbb{C}$ and polynomial $R(x)$ of degree $m - 1$. Here $R(x)$ and r_0 can be computed by algebra. Dividing $P(x)$ over n linear polynomials $x - x_1$, ..., $x - x_n$ in turn, we reduce (A.7) to (A.1).

Example A.3.2. *Compute* $\int \frac{3x}{x^2+3x+2} dx$.

Proof. From Example A.3.1 we recall that

$$\frac{3x}{x^2+3x+2} = \frac{x}{x-2} - \frac{x}{x+1}.$$

Here $Q(x) = x^2 + 3x + 2$ and $P(x) = 3x$. Since $m = 1$, polynomial division is trivial here:

$$\frac{x}{x-2} = \frac{x-2+2}{x-2} = 1 + \frac{2}{x-2} \quad \text{and} \quad \frac{x}{x+1} = \frac{x+1-1}{x+1} = 1 - \frac{1}{x+1}.$$

Therefore

$$\int \frac{3x dx}{x^2+3x+2} = \int \frac{2dx}{x-2} - \int \frac{dx}{x+1} = 2\ln(|x-2|) - \ln(|x+1|) + C. \qquad \square$$

Example A.3.3. *Divide the polynomial* $3x^2+x-4$ *over the linear polynomial* $x-2$ *with a remainder.*

Proof. We observe that

$$\frac{3x^2+x-4}{x-2} = \frac{3x(x-2)+6x+x-4}{x-2} = 3x + \frac{7x-4}{x-2} = 3x + \frac{7(x-2)+10}{x-2} = 3x+7+\frac{10}{x-2}. \quad \square$$