Hans Lundmark
Department of Mathematics
Linköping University

2023-10-27

ODEs and Dynamical Systems (TATA71)
Course programme, fall 2023

General information

Ordinary Differential Equations and Dynamical Systems (TATA71) is an optional course for MAT2, Y4, M4, EMM4. It is given in the second half of the fall semester (period ht2). All information is publicly available on the course webpage courses.mai.liu.se/GU/TATA71/. Lisam is currently not used in this course.

Literature

D. K. Arrowsmith & C. M. Place, *Dynamical Systems: Differential Equations, Maps, and Chaotic Behaviour*, Chapman and Hall/CRC (1992), ISBN 9780412390807. Available as an e-book via LiU's library.

Prerequisites

The prerequisites are basic courses in **single-variable** and **multi-variable calculus**, plus **linear algebra**. An honours course in **real analysis** may be helpful for understanding certain subtle details, but that's optional.

Here "single-variable calculus" means that you are supposed to already have seen the basic techniques for solving simple ODEs (as taught, for example, in the course TATA42 here at LiU): separation of variables, integrating factors, the characteristic polynomial, the method of undetermined coefficients, and so on. However, we will spend some time brushing up on this at the beginning of the course.

Teaching

This year the course will follow a new format in the "flipped classroom" style, which means that before each of the 11 classroom sessions marked "Lecture" in the **course outline** below, you are supposed to watch the corresponding pre-recorded video lecture. In class, there will then be a short summary instead of a regular lecture, and the rest of the time will be available for looking at additional examples, working on the exercises, and so on. There are also four "Lessons" in the course outline, which provide some extra time for catching up.

Course evaluations and changes

The main change this year is that we are trying out the flipped classroom format, with the idea of making better use of the classroom time. The course content is unchanged, with only some minor updates to the course programme and to the exercises. Previous course evaluations can be found by searching Evaliuate; no particular changes to the course have been prompted by the responses there.

Examination

The examination consists of two parts:

- **UPG1. Homework assignments, worth 2 hp (= 2 ECTS credits).**

  Some of the exercises are assigned as homework problems (marked with yellow in this course programme), to be handed in continually during the course. These problems are only graded pass/fail, and if you fail a problem, you simply hand in a corrected version later. Discussing the problems with the teacher and with your fellow students is allowed, but please write the solution in your own words; it's not allowed to just copy someone else's solution! Handwritten solutions are fine, and you can write them in English or in Swedish.

  The intent is that the homework assignments should be completed before the first written exam, which takes place Jan 11, 2024. It's a good idea to hand in the last problems before Christmas, to make sure that there is enough time for getting feedback before the exam.

  Homework problems which are handed in after the exam will be graded sooner or later, but probably later, since at that time I will have other duties with higher priority.

- **TEN1. A written test (5 hours), worth 4 hp.**

  The test contains 6 problems, each of which is graded as pass (3 or 2 points) or fail (1 or 0 points). The total grade for the course is determined by the grade for the written test, which in turn is determined as follows: for grade 3/4/5 (respectively), you need 3/4/5 passed problems and in addition at least 8/11/14 points in total.

  The written examination takes place on LiU's campus in Linköping, three times per year (January, March, August). No exams at other locations will be arranged.

What's this course about?

As the name of the course suggests, we will study **ODEs** and **dynamical systems**.

- **ODE** is a standard abbreviation for **ordinary differential equation**, where the function that we seek depends on *one* variable. So an ODE is just a good old differential equation like those which you have already seen in your single-variable calculus course. We will also encounter *systems* of ODEs, involving several unknown functions at once, but each function will still only depend on one variable.

  In contrast, a **partial differential equation** (**PDE**) is a differential equation where one seeks a function depending on *several* variables, so that *partial derivatives* come into play. But that's a different subject; see the course TATA27.

- The idea of a **dynamical system** is rather broad, and it is hard to give a precise mathematical definition which would cover every possible use of the phrase. But it refers to a "system" (whatever that is) which changes in a deterministic way as time passes, and which is "memoryless" in the sense that the future of the system, at any given instant, is uniquely determined by the present state alone; the past is irrelevant.

  We will usually assume when talking about dynamical systems that the laws governing the evolution don't change with time. That is, if we start the system in a given state T units of time from now, we will get the same evolution as if we start it in that state right now (except that everything will be delayed by T time units of course). If this needs to be emphasized, one uses the phrase **autonomous dynamical system** – a system which "runs on its own", in contrast to **non-autonomous** dynamical systems where there may be some external time-dependent factors which influence the evolution of the system.

  The state of the system is represented mathematically by an element of a set called the **state space** or the **phase space**, typically $\mathbf{R}^n$, or maybe some subset of $\mathbf{R}^n$ like a cylinder, a sphere, or a torus. So one pictures the evolution of the system as the motion of a point in the state space.

- A **discrete-time dynamical system** is one where things happen at distinct time steps. We can use integers to label the time steps, so that the system is in the state $x_n \in S$ at time $n \in \mathbf{Z}$, where S is the state space. Then the evolution of the system is simply specified by some function $f\colon S \to S$, like this:

  $$x_{n+1} = f(x_n), \qquad n \in \mathbf{Z}.$$

  (The system is autonomous since the function f is the same for all n.)

  Discrete-time dynamical systems are a very important part of the general theory of dynamical systems, but will not be encountered very much in this course.

- In a **continuous-time dynamical system**, time passes smoothly, so we use real numbers to describe time, and talk about the system being in the state $x(t)$ at time $t \in \mathbf{R}$. In this case, if the state space is $\mathbf{R}^n$ for simplicity, the evolution is determined by a system of first-order ODEs for the state $\mathbf{x}(t) = \big(x_1(t), \dots, x_n(t)\big)$:

  $$\begin{aligned} dx_1/dt &= f_1(x_1, \dots, x_n), \\ dx_2/dt &= f_2(x_1, \dots, x_n), \\ &\vdots \\ dx_n/dt &= f_n(x_1, \dots, x_n), \end{aligned}$$

  or simply $d\mathbf{x}/dt = \mathbf{f}(\mathbf{x})$ for short, with $\mathbf{x} \in \mathbf{R}^n$ and $\mathbf{f}\colon \mathbf{R}^n \to \mathbf{R}^n$. (The system is autonomous since the function $\mathbf{f}$ doesn't depend on t.)

  In order for this to really define a dynamical system, we must impose suitable conditions on $\mathbf{f}$ which will guarantee **existence** and **uniqueness** of the solution to the ODEs for a given initial condition, so that the future is uniquely determined by the present state.

Course outline

[Jump to course outline]

Lecture 1. Basics of first-order ODEs

(Arrowsmith & Place, sections 1.1, 1.2. And your old calculus textbook, if needed.)

Some basic ideas are introduced:

- **Existence and uniqueness theorems** for first-order ODEs $\frac{dx}{dt} = X(t,x)$.

  Proposition 1.1.1 is usually called **Peano's existence theorem**.

  Proposition 1.1.2 is a slightly simplified version of the **Picard–Lindelöf theorem**. (The assumption "$\partial X/\partial x$ exists and is continuous" is stronger than necessary; we'll study this more thoroughly in Lecture 9.)

- How to find **explicit solutions** in simple cases.

  (Hopefully, this will mostly be a question of remembering methods that you have learned in previous courses. See the summary below.)

- How to sketch the **solution curves** directly from the ODE.

- How to draw the **phase portrait** for a one-dimensional (continuous-time & autonomous) dynamical system.

A comment about notation

You might be used to ODEs looking something like this:

$$y''(x) + 5y'(x) + 4y(x) = \cos x \qquad \text{(or simply} \quad y'' + 5y' + 4y = \cos x\text{)},$$

where the independent variable is called x, and $y = y(x)$ is the function that we seek. But since this is a course about ODEs with a "dynamical systems perspective", we will instead call the independent variable t, for "time", and use names like $x(t)$ or $y(t)$ for the sought functions. So the same ODE now instead looks as follows:

$$x''(t) + 5x'(t) + 4x(t) = \cos t \qquad \text{or} \qquad \ddot{x}(t) + 5\dot{x}(t) + 4x(t) = \cos t.$$

(It is common to use **dots** instead of primes to denote **derivatives with respect to time**.)

Two very fundamental examples

- **Exponential growth/decay**:

  $$x'(t) = r\,x(t).$$

  This is a linear equation, and we can solve it in several ways: integrating factor, characteristic polynomial, separation of variables. Either way, the solution with initial condition $x(0) = x_0$ is

  $$x(t) = x_0\,e^{rt}.$$

  This will be encountered again and again in this course, and you will be expected to *instantly* recognize this equation and know its solution. Phase portrait (if $r > 0$): "⟵ 0 ⟶".

- The **logistic equation** with growth rate r and carrying capacity K:

  $$x'(t) = r\,x(t)\left(1 - \frac{x(t)}{K}\right).$$

  This nonlinear equation is often solved via separation of variables (followed by integration using partial fractions), but an easier way is to use the substitution $x(t) = 1/y(t)$, since this is a Bernoulli equation (see below). Solution, with $x(0) = x_0$:

  $$x(t) = \frac{K x_0}{x_0 + (K - x_0)\,e^{-rt}} = \frac{K x_0\,e^{rt}}{K + (e^{rt} - 1)x_0}.$$

Here you don't need to memorize the solution formula, but you should be able to derive it, and you should also know roughly what the *graph* of the solution $x(t)$ looks like for different values of x_0. (In particular, $x(t) = 0$ and $x(t) = K$ are constant solutions.) Phase portrait (if $r > 0$ and $K > 0$): "$\longleftarrow 0 \longrightarrow K \longleftarrow$".

Summary of some exact solution methods

- **Linear first order equations** $x' + ax = b$, where the coefficients a and b may be functions of the time variable t:

  $$x'(t) + a(t)\,x(t) = b(t).$$

  How to solve: Find an antiderivative of $a(t)$; call it $A(t)$. Then multiply both sides of the ODE by the **integrating factor** $e^{A(t)}$, and use the product rule for derivatives (backwards). This gives

  $$\left(e^{A(t)}\,x(t)\right)' = e^{A(t)}\,b(t),$$

  which can now be integrated.

- **Separable equations** $f(x)\,x' = g(t)$.

  What this means is that we seek $x(t)$ such that

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

  Integrating this with respect to t, using the chain rule (backwards), we immediately obtain the solution in the implicit form

  $$F(x(t)) = G(t) + C,$$

  where $F(x)$ is some antiderivative of $f(x)$ and $G(t)$ is some antiderivative of $g(t)$. Usually this is remembered via the trick of writing $x' = dx/dt$ and "separating the variables" by "multiplying by dt", and then attaching integral signs:

  $$f(x)\,\frac{dx}{dt} = g(t) \qquad \Longleftrightarrow \qquad \int f(x)\,dx = \int g(t)\,dt.$$

  It's sometimes convenient to use *definite* integrals instead, particularly if we want to find a solution satisfying a given **initial condition** $x(t_0) = x_0$:

  $$\int_{x_0}^{x(t)} f(\xi)\,d\xi = \int_{t_0}^{t} g(\tau)\,d\tau,$$

  or in other words

  $$F(x(t)) - F(x_0) = G(t) - G(t_0).$$

  As an important special case of separable equations we have **one-dimensional dynamical systems** $x' = a(x)$, which can always be solved (in principle) by separation of variables:

  $$x' = a(x) \quad \Longleftrightarrow \quad x(t) = x^* \quad \text{where } a(x^*) = 0$$
  $$\text{or} \qquad \int \frac{dx}{a(x)} = \int dt = t + C.$$

  Warning! This method is full of pitfalls! Don't forget the case $a(x) = 0$; there is a constant solution $x(t) = x^*$ for each zero x^* of the function $a(x)$. And one should also be *very* careful with handling logarithms and absolute values correctly when integrating and when simplifying the solution. And in principle the constant of integration may be different in the different intervals into which the real line is divided by the zeros of $a(x)$. So if the ODE can be solved by some other method, it may be wise to try that method first.

[Jump to course outline]

- **Linear equations of arbitrary order.**

  A linear ODE of order n has the form

  $$x^{(n)}(t) + a_{n-1}(t)\,x^{(n-1)}(t) + \cdots + a_2(t)\,x''(t) + a_1(t)\,x'(t) + a_0(t)\,x(t) = b(t).$$

  The general solution of such an equation has the structure

  $$x(t) = x_{\text{hom}}(t) + x_{\text{part}}(t)$$

  where $x_{\text{part}}(t)$ is a **particular solution** and $x_{\text{hom}}(t)$ (called the **homogeneous solution** or the **complementary solution**) is the general solution of the corresponding homogeneous equation which has 0 instead of $b(t)$ on the right-hand side.

  Solving higher-order equations with **time-dependent coefficients** is usually a rather hopeless task, except that one may try to find solutions in the form of a **power series** in t.

- **Linear equations of arbitrary order with constant coefficients.**

  The problem becomes tractable when the coefficients $a_k(t) = c_k$ are time-independent:

  $$x^{(n)}(t) + c_{n-1}x^{(n-1)}(t) + \cdots + c_2x''(t) + c_1x'(t) + c_0x(t) = b(t).$$

  Then $x_{\text{hom}}(t)$ can be found by looking at the roots of the **characteristic polynomial**

  $$p(r) = r^n + c_{n-1}\,r^{n-1} + \cdots + c_2r^2 + c_1r + c_0.$$

  If the roots are real and simple, it's straightforward to write down $x_{\text{hom}}(t)$. For repeated and/or complex roots, the rules for constructing $x_{\text{hom}}(t)$ are a bit more complicated.

  To find $x_{\text{part}}(t)$ one usually makes a suitable *Ansatz* ("the method of undetermined coefficients").

  See your calculus book for details about all this.

- A **Bernoulli equation** is an ODE of the form

  $$x'(t) + p(t)\,x(t) = q(t)\,x(t)^k$$

  where k is a constant (not necessarily an integer). Note that this is a **nonlinear** ODE (unless $k = 0$ or $k = 1$) because of the expression $x(t)^k$ on the right-hand side. The nonzero solutions to such an equation can be found by dividing both sides by that factor $x(t)^k$, to get

  $$\frac{x'(t) + p(t)\,x(t)}{x(t)^k} = q(t),$$

  or in other words

  $$x'(t)\,x(t)^{-k} + p(t)\,x(t)^{1-k} = q(t).$$

  Indeed, just let

  $$y(t) = x(t)^{1-k},$$

  which according to the chain rule has the derivative

  $$y'(t) = (1-k)x(t)^{-k}x'(t),$$

  and compare this to what we had in our equation:

  $$\underbrace{x'(t)\,x(t)^{-k}}_{=y'(t)/(1-k)} + p(t)\,\underbrace{x(t)^{1-k}}_{=y(t)} = q(t).$$

  That is, in terms of the new unknown function $y(t)$ we get a first-order **linear** ODE, the type that can be solved using an integrating factor:

  $$\frac{1}{1-k}\,y'(t) + p(t)\,y(t) = q(t).$$

[Jump to course outline]

- Later in this course we will thoroughly study **linear first order systems**

$$\mathbf{x}'(t) = A(t)\,\mathbf{x}(t),$$

where $\mathbf{x}(t) = \big(x_1(t), \ldots, x_n(t)\big)^T$ is a column vector of functions and $A(t)$ is matrix of size $n \times n$. But you may already have seen in your linear algebra course how to solve such a system in the case where A is a **constant** matrix which happens to be **diagonalizable**. In this case, the change of variables $\mathbf{x} = M\mathbf{y}$, where the columns of M are n linearly independent eigenvectors of A, leads to a decoupled system $\mathbf{y}'(t) = D\,\mathbf{y}(t)$, where D is a diagonal matrix with the eigenvalues of A along the diagonal (in the same order as the corresponding eigenvectors were listed in the matrix M).

Exercises

The problems labelled A1, A2, etc., can be found in the section **Additional problems** just below. The remaining problems (1.1, 1.2, etc.) are from the **course textbook** (Arrowsmith & Place), where they are located at the end of each chapter. Problems marked with an **asterisk** might be a bit more challenging.

Complete solutions to the exercises are not provided, only answers (and hints, in some cases). There is a pedagogical idea behind this, namely that you should not be tempted to simply imitate or memorize somebody else's solution, but instead think carefully about the problems and really try to solve them, and resolve any possible doubts through discussions with the teacher and with your fellow students.

- Rehearsal of how to solve ODEs using calculus techniques: 1.1, 1.2, 1.4, A1, A2, A3.
- Sketching solution curves directly from the ODE: 1.11.
- Drawing phase portraits: 1.12, 1.13.
- Phase portraits for parameter-dependent ODEs: 1.14, 1.17*.

  (When the phase portrait changes qualitatively at some particular value of the parameter(s), the system is said to undergo a *bifurcation*.)
- Recovering the ODE from the family of solution curves: A4.

Additional problems

A1 Derive the solution formula given above for the logistic equation $\dot{x} = r\,x\big(1 - x/K\big)$ with $x(0) = x_0$, first by using separation of variables, and then again by treating it as a Bernoulli equation.

A2 Find the general solution of the following constant-coefficient second-order ODEs:

(a) $\ddot{x} + 6\dot{x} + 8x = t + 2e^{-2t}$ [Answer: $x(t) = Ae^{-2t} + Be^{-4t} + (4t-3)/32 + t\,e^{-2t}$.]

(b) $\ddot{x} + 6\dot{x} + 9x = 0$ [Answer: $x(t) = (At + B)e^{-3t}$.]

(c) $\ddot{x} + 6\dot{x} + 10x = 2\,e^{-3t}\cos t$ [Answer: $x(t) = e^{-3t}(A\cos t + B\sin t) + te^{-3t}\sin t$.]

A3 The *Airy equation* $\ddot{x}(t) = t\,x(t)$ is a second-order ODE with non-constant coefficients. Find the solution which satisfies $x(0) = 1$ and $\dot{x}(0) = 0$, in the form of a power series $x(t) = \sum_{k=0}^{\infty} a_k\,t^k$.

[Answer: $x(t) = 1 + \frac{t^3}{2\cdot 3} + \frac{t^6}{2\cdot 3\cdot 5\cdot 6} + \frac{t^9}{2\cdot 3\cdot 5\cdot 6\cdot 8\cdot 9} + \ldots$]

A4 A first-order ODE (usually) has a general solution consisting of a one-parameter family of curves. Here's an opposite problem: Find an ODE $\dot{x} = X(t,x)$ such that the one-parameter family of curves $t^2 + x^2 = 2Ct$ are solution curves.

(Hint: Differentiate $t^2 + x(t)^2 = 2Ct$ with respect to t, and eliminate the parameter C from the two equations.)

[Answer: $\dot{x} = (x^2 - t^2)/2tx$.]

[Jump to course outline]

Lesson 1

Lecture 2. Phase portraits for two-dimensional systems

(Arrowsmith & Place, sections 1.3, 1.4, 1.5.)

Now we turn to **two-dimensional** dynamical systems

$$\dot{x}_1 = X_1(x_1, x_2), \qquad \dot{x}_2 = X_2(x_1, x_2),$$

where we assume that the functions X_1 and X_2 are nice enough to guarantee uniqueness and (local) existence of solutions. In vector notation: $\dot{\mathbf{x}}(t) = \mathbf{X}(\mathbf{x}(t))$, or simply $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$. The function $\mathbf{X}$ should be thought of as a **vector field** in phase space: a vector $\mathbf{X}(\mathbf{x})$ is prescribed at each point $\mathbf{x}$. The solution curves for the system are curves $\mathbf{x} = \mathbf{x}(t)$ whose tangent vector $\dot{\mathbf{x}}$ agrees with the prescribed vector $\mathbf{X}(\mathbf{x})$ at each point on the curve. The solution curves are sometimes called **flow lines** of the vector field.

- In very simple cases, we can solve the system explicitly. Sometimes we can obtain partial information by methods of calculus.

- But we can also get at least a rough idea of what the phase portrait looks like by direct inspection of the signs of the functions $X_1(x_1, x_2)$ and $X_2(x_1, x_2)$.

- Syntax for drawing phase portraits in Mathematica or Wolfram Alpha (www.wolframalpha.com):

  StreamPlot[$X_1(x, y)$, $X_2(x, y)$, {**x,** $x_{\min}$, $x_{\max}$}, {**y,** $y_{\min}$, $y_{\max}$ }]

  For example, if the system is $\dot{x} = -1 - x^2 + y$, $\dot{y} = 1 + x - y^2$:

  ```
  StreamPlot[-1-x^2+y, 1+x-y^2, {x, -3, 3}, {y, -3, 3}]
  ```

  (But don't expect an automatic command like this to produce anything nearly as good as the hand-tuned graphics in the textbook.)

- Arrowsmith & Place use the phrase **fixed point** for a point $\mathbf{x}^*$ such that $\mathbf{X}(\mathbf{x}^*) = \mathbf{0}$, but I will usually say **equilibrium point** or just **equilibrium**, and you may come across many other synonyms too: **rest point**, **critical point**, **steady state**, etc.

  The reason for the terminology is of course that $\mathbf{x}(t) = \mathbf{x}^*$ is a constant solution of the system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$ in this situation. That is, a system starting in the state $\mathbf{x}^*$ remains in the state $\mathbf{x}^*$ forever.

- The **evolution operator** or **flow** φ is the function which maps each point $\mathbf{x}$ in phase space to the place where it will be t units of time later if moving as prescribed by the dynamical system. It is denoted by $\varphi_t(\mathbf{x})$ in the book, but it's also common to write $\varphi(t, \mathbf{x})$, since it is simply a function of t and $\mathbf{x}$.

  A fact which isn't mentioned in the textbook is that the flow φ is as "nice" as the vector field $\mathbf{X}$. More precisely, there's a theorem which says that if the system is

  $$\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x}, \boldsymbol{\lambda}),$$

  where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N)$ is some vector of parameters, and if $\mathbf{X}$ is of class C^k $(1 \le k \le \infty)$ as a function of the variables $(\mathbf{x}, \boldsymbol{\lambda})$, then the flow φ is of class C^k as a function of the variables $(t, \mathbf{x}, \boldsymbol{\lambda})$.

[Jump to course outline]

Exercises

As mentioned already on p. 2, problems marked with yellow are **homework problem** to be handed in.

- Finding fixed points: 1.19acd.
- Finding nullclines, and determining the signs of $\dot{x}_1$ and $\dot{x}_2$ in between: A5, A6.
- Drawing families of parametrized curves and finding the corresponding ODEs: 1.20.
- Exact solution of a linear system: 1.23 (draw the phase portrait both in the y_1y_2-plane and in the x_1x_2-plane).
- An easy verification of a claimed solution: 1.24.
- A nonlinear system which can be solved using polar coordinates: 1.25.
- Flows: 1.32, 1.36.

  (Hint for 1.32, if you want to solve it from scratch: It's a Bernoulli equation, so divide by x^3 and then let $y = 1/x^2$. Or just use separation of variables and partial fractions directly. But instead of solving, it might be easier to just *verify* that the given flow is correct, by computing $\varphi_0(x_0)$ and $\frac{d}{dt}\varphi_t(x_0)$.)

Additional problems

A5 For the systems in problem 1.19acd, find the **nullclines** (the curves where $\dot{x}_1 = 0$ or $\dot{x}_2 = 0$) and use this to draw the regions of the phase plane where $\dot{x}_1$ and $\dot{x}_2$ are positive/negative. Clearly mark each of these regions with "DL", "DR", "UL" or "UR" for Down/Up & Left/Right, or, if you prefer, with arrows ↙, ↘, ↖ or ↗. (For example, in 1.19d there will be **11 regions**.)

It might be a good idea to first make a picture for the x_1-nullcline only, marking the regions with "L" and "R", then a separate picture for the x_2-nullcline only, marked with "D" and "U", and finally to combine the two pictures to obtain the answer.

A6 Do the same as in problem A5, but for the system

$$\dot{x} = (x-1)^2 + y + 1, \qquad \dot{y} = (x+1)^2(x+y).$$

Lecture 3. Two-dimensional linear systems

(Arrowsmith & Place, sections 2.1, 2.2, 2.3.)

We will spend some time on understanding **linear** dynamical systems in **two dimensions**:

$$\begin{aligned} \dot{x}_1 &= ax_1 + bx_2 \\ \dot{x}_2 &= cx_1 + dx_2 \end{aligned} \quad \Longleftrightarrow \quad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{=A} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

- Definition of a **simple** linear system: $\det(A) = \left|\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right| = ad - bc$ is **nonzero**. (Equivalently: the eigenvalues of A are nonzero.) This implies that the origin $(x_1, x_2) = (0,0)$ is the **only** equilibrium point. We'll assume simplicity for now, and save non-simple systems for later.
- A linear change of variables $\mathbf{x} = M\mathbf{y}$ turns the system $\dot{\mathbf{x}} = A\mathbf{x}$ into $\dot{\mathbf{y}} = M^{-1}AM\mathbf{y}$.
- "Recipe" for how to choose the columns $\mathbf{m}_1$ and $\mathbf{m}_2$ in the matrix M in order to make $J = M^{-1}AM$ simplify to the **Jordan normal form** of A:
  (a) Suppose A has two distinct real eigenvalues $\lambda_1 > \lambda_2$. Taking $\mathbf{m}_1$ and $\mathbf{m}_2$ to be the corresponding eigenvectors gives $J = \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right)$.

(b) Suppose A has a double real eigenvalue λ_0. Then either $A = \begin{pmatrix} \lambda_0 & 0 \\ 0 & \lambda_0 \end{pmatrix}$ is already in normal form, or else there is just a one-dimensional eigenspace; in this case let $\mathbf{m}_1$ be an eigenvector and take any vector not parallel to $\mathbf{m}_1$ as our preliminary $\mathbf{m}_2$. This will give a preliminary $J = \begin{pmatrix} \lambda_0 & C \\ 0 & \lambda_0 \end{pmatrix}$ with some nonzero constant C. Adjust $\mathbf{m}_2$ by dividing it by this constant C; this will give the Jordan form $J = \begin{pmatrix} \lambda_0 & 1 \\ 0 & \lambda_0 \end{pmatrix}$.

(c) Suppose A has a complex-conjugated pair of eigenvalues $\lambda_{1,2} = \alpha \pm i\beta$ with $\beta > 0$.
This case is done in a very complicated way in the textbook (p. 39). A much simpler way is to let $\mathbf{a} + i\mathbf{b}$ (with $\mathbf{a}$ and $\mathbf{b}$ real) be a complex eigenvector corresponding to the eigenvalue $\lambda_1 = \alpha + i\beta$. Then $\mathbf{m}_1 = \mathbf{b}$ and $\mathbf{m}_2 = \mathbf{a}$ will work, and directly give the Jordan form $J = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$. (These vectors are automatically linearly independent – otherwise they would both be *real* eigenvectors of the real matrix A, with a *non-real* eigenvalue λ_1, which is impossible.)

- How to solve the system $\dot{\mathbf{y}} = J\mathbf{y}$ with a matrix in Jordan form, and how to draw its phase portrait.

  In case (a) the phase portrait is a **node** or a **saddle** depending on the signs of the eigenvalues. In case (b) it is a **star node** or an **improper node**. And in case (c) it is a **spiral** (also called **focus**) if $\alpha \neq 0$, or a **centre** if $\alpha = 0$.

  The phase portrait for $\mathbf{x} = M\mathbf{y}$ will be of the same type, only distorted by the linear transformation M. The columns $\mathbf{m}_1$ and $\mathbf{m}_2$ give the **principal directions** of the phase portrait.

- A solution method which isn't mentioned in the book, but is sometimes convenient, is to **rewrite the system as a single second-order ODE** with constant coefficients, as follows. The derivative of $\dot{x}_1 = ax_1 + bx_2$ is

$$\ddot{x}_1 = a\dot{x}_1 + b\dot{x}_2.$$

  The second term here, $b\dot{x}_2$, can be rewritten using the equations from the system:

$$b\dot{x}_2 = b(cx_1 + dx_2) = bcx_1 + d(bx_2) = bcx_1 + d(\dot{x}_1 - ax_1) = d\dot{x}_1 - (ad - bc)x_1.$$

  So we find that $\ddot{x}_1 = a\dot{x}_1 + \big(d\dot{x}_1 - (ad - bc)x_1\big)$, or in other words

$$\ddot{x}_1 - (a + d)\dot{x}_1 + (ad - bc)x_1 = 0.$$

  This second-order ODE for $x_1(t)$ can now be **solved as in calculus**, with the help of its characteristic polynomial

$$p(\lambda) = \lambda^2 - (a + d)\lambda + (ad - bc),$$

  which (perhaps not surprisingly) coincides with the characteristic polynomial of the system's matrix A, as we can easily check:

$$p(\lambda) = \det(A - \lambda I) = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - bc = \lambda^2 - (a + d)\lambda + (ad - bc).$$

  And once $x_1(t)$ is known, we also get $x_2(t)$ from $x_2 = (\dot{x}_1 - ax_1)/b$, assuming that $b \neq 0$. (If $b = 0$, the system is easy to solve right away: the first equation $\dot{x}_1 = ax_1 + 0x_2$ gives $x_1(t) = x_1(0)\, e^{at}$; plug this into the second equation and solve the resulting ODE for $x_2(t)$ as in calculus, either using an integrating factor or "homogeneous + particular solution".)

Exercises

- A little reminder of how linear transformations work: 2.1.
- Transformation to canonical form: 2.3, 2.4.

  (There's an error in the answer to 2.3c: $\left(\begin{smallmatrix} 3 & 0 \\ 0 & 2 \end{smallmatrix}\right)$ should be $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 3 \end{smallmatrix}\right)$.)
- Solving canonical linear systems (and drawing the phase portrait): 2.8, 2.9.

[Jump to course outline]

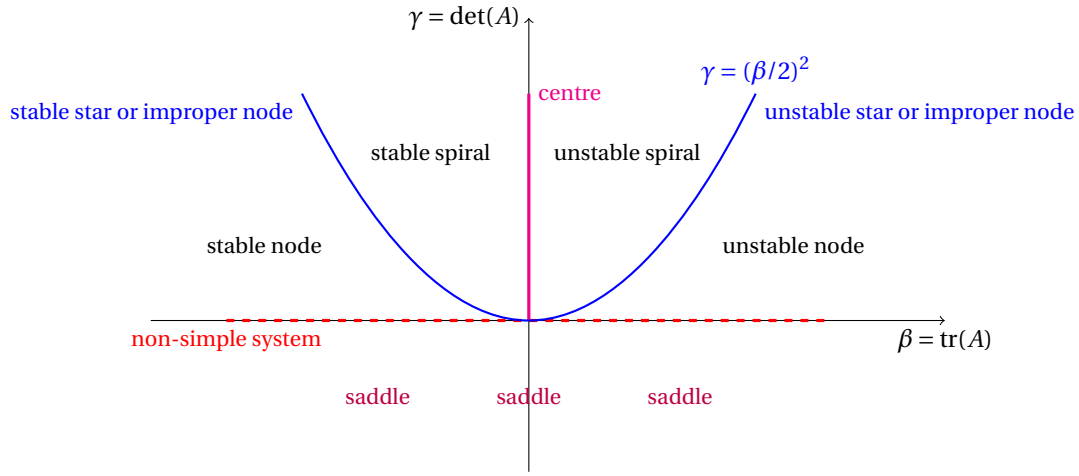
Lecture 4. More about linear systems

(Arrowsmith & Place, sections 2.4, 2.5, 2.6, 2.7.)

- How to tell directly from the coefficients of the matrix $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ what the type of the phase portrait for $\dot{\mathbf{x}} = A\mathbf{x}$ is, without computing the eigenvalues: compute the **trace** and the **determinant**,

$$\beta = \operatorname{tr}(A) = a + d, \qquad \gamma = \det(A) = ad - bc,$$

and locate the point (β, γ) in the following diagram (cf. Figure 2.7 in the book, p. 47):



For systems on the parabola $\gamma = (\beta/2)^2$, the type is a star node if the matrix A equals a constant times the identity matrix, otherwise it's an improper node.

Centres are always stable, and saddles are always unstable (according to the general definitions of "stable equilibrium" and "unstable equilibrium" that will be given later in the course).

- In particular, the zeros of the polynomial $p(z) = z^2 - \beta z + \gamma$ both have *negative real part* if and only if the coefficients satisfy $\beta < 0$ and $\gamma > 0$, i.e., if the point (β, γ) lies in the second quadrant in the diagram above. (This is a special case of the **Routh–Hurwitz criterion**, which determines for a polynomial of arbitrary degree whether all its zeros have negative real part. The name Routh rhymes with "south".)

- The word "type" above refers to a kind of "algebraic type", with two linear systems being of the same type if and only if they are related via a **linear** change of variables. Such changes preserve the eigenvalues, so (for example) an unstable node with $\lambda_1 = 5$ and $\lambda_2 = 3$ cannot be linearly transformed into an unstable node with $\lambda_1 = 43$ and $\lambda_2 = 17$.

  We might want to introduce some looser type of equivalence instead, which will allow (for example) any two unstable nodes to be considered equivalent.

- Classification of **qualitatively equivalent** types of phase portraits for two-dimensional simple linear systems. Under this equivalence relation (**Definition 2.4.1**) there are only four different types:

  (a) Stable node, stable star node, stable improper node, or stable spiral. ("Stable but not centre.")

  (b) Centre.

  (c) Saddle.

  (d) Unstable node, unstable star node, unstable improper node, or unstable spiral. ("Unstable but not saddle.")

[One subtle detail regarding Definition 2.4.1: They say that two systems are qualitatively equivalent if there is a continuous bijection f which maps the phase portrait of one system onto the phase portrait of the other and preserves the orientation of the trajectories. But how do we know that this relation is *symmetric*? (If system A is equivalent to system B, then we want system B to be equivalent to system A as well.) For this to hold, the inverse f^{-1} must also be a *continuous* bijection. This isn't necessarily true for continuous bijections in general; consider for example the function from the interval $(-\pi, \pi]$ in $\mathbf{R}$ to the unit circle in $\mathbf{R}^2$ given by $f(t) = (\cos t, \sin t)$. But there is a famous and rather difficult theorem by Brouwer ("invariance of domain") which implies that for a continuous bijection between two open sets in $\mathbf{R}^n$ the inverse is automatically continuous, so the definition is actually correct as it stands.]

- Qualitative equivalence is also called **topological equivalence**.

  (A stronger condition is C^k**-equivalence**, where one requires the map f and its inverse f^{-1} to be of class C^k instead of just continuous.)

- As an aside, we may also mention the slightly different notion of **conjugacy**, where one also requires the *time parametrization* of the trajectories to be respected. More precisely, two systems with flows φ and ψ are said to be **topologically conjugate** if there's a continuous map f with continuous inverse f^{-1} such that

  $$f \circ \varphi_t = \psi_t \circ f$$

  for all t. That is, following the flow φ of one system for t units of time and then mapping over to the other phase portrait is always the same as first mapping to the other phase portrait and then following *that* flow ψ for t units of time.

  (And the systems are C^k**-conjugate** if f and f^{-1} are of class C^k.)

  For example, a system whose phase portrait consists of concentric circles all traversed with the same period T is *topologically equivalent* to a system whose phase portrait consists of concentric circles traversed with different periods, but the systems are *not topologically conjugate*.

  A conjugacy can also be viewed as a change of variables; if the system with flow φ is described in terms of the variables $\mathbf{x}$, and we make the change of variables $\mathbf{y} = f(\mathbf{x})$, then ψ can be considered as the flow of the same system, just expressed in terms of the new variables $\mathbf{y}$ instead.

- The **exponential function** for square matrices P:

  $$e^P = \exp(P) = I + P + \frac{1}{2!}P^2 + \frac{1}{3!}P^3 + \dots$$

- The solution to $\dot{\mathbf{x}} = A\mathbf{x}$ is

  $$\mathbf{x}(t) = e^{At}\mathbf{x}(0).$$

  (Here it's important that A is a *constant* matrix, not time-dependent.)

- **Affine** systems $\dot{\mathbf{x}} = A\mathbf{x} + \mathbf{h}(t)$, also called **non-homogeneous linear** systems.

  It's easy if $\mathbf{h}$ is time-independent and $A\mathbf{x}_0 = \mathbf{h}$ for some $\mathbf{x}_0$: just set $\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{x}_0$ to get $\dot{\mathbf{y}} = A\mathbf{y}$. But even if this doesn't hold, an affine system can be solved using e^{-At} as an integrating factor.

- Linear systems and Jordan form in n dimensions.

Exercises

- Using the trace–determinant criterion: 2.13.
- Phase portrait via change of variables: 2.14.

  (Warning: It's quite difficult to draw a precise phase portrait in the x_1x_2-plane by hand, since the principal directions for this system are nearly parallel to the nullclines.)

- More phase portraits: A7.

- Computing matrix exponentials: 2.22, 2.23.
- Affine systems: 2.29abcd, 2.30.
- A three-dimensional linear system: 2.33.
- Solution corresponding to a 4×4 Jordan block: 2.35ad.

Additional problems

A7 In each subproblem, draw the phase portrait for the system $\dot{\mathbf{x}} = A\mathbf{x}$ as carefully as you can, but without computing the solution $\mathbf{x}(t)$ explicitly. (Use the trace–determinant criterion or the eigenvalues to determine the type, and compute the the principal directions if there are any. Please indicate the nullclines and the principal directions clearly in your figures.)

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

(b) $A = \begin{pmatrix} 2 & -6 \\ 2 & -1 \end{pmatrix}$.

(c) $A = \begin{pmatrix} 0 & -1 \\ 4 & -4 \end{pmatrix}$.

Lesson 2

Lecture 5. Nonlinear systems, linearization at an equilibrium point

(Arrowsmith & Place, sections 3.1, 3.2, 3.3, 3.4.)

Now back to nonlinear systems $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$, in the plane $\mathbf{R}^2$ for simplicity, but the results are valid in $\mathbf{R}^n$ too. We assume that the vector field $\mathbf{X}(\mathbf{x})$ is of class C^1, so that we have existence and uniqueness of solutions; this assumption is also needed for the linearization theorem below to be valid.

- Suppose that $\mathbf{x}^* = (a_1, a_2)$ is an equilibrium point: $X_1(a_1, a_2) = X_2(a_1, a_2) = 0$. Since the functions X_1 and X_2 are assumed to be of class C^1, they are also differentiable, which by definition means that

$$X_1(a_1 + h_1, a_2 + h_2) = \underbrace{X_1(a_1, a_2)}_{=0} + \frac{\partial X_1}{\partial x_1}(a_1, a_2)\, h_1 + \frac{\partial X_1}{\partial x_2}(a_1, a_2)\, h_2 + \text{remainder},$$

$$X_2(a_1 + h_1, a_2 + h_2) = \underbrace{X_2(a_1, a_2)}_{=0} + \frac{\partial X_2}{\partial x_1}(a_1, a_2)\, h_1 + \frac{\partial X_2}{\partial x_2}(a_1, a_2)\, h_2 + \text{remainder},$$

  where the remainders tend to zero *faster*[*] than $\sqrt{h_1^2 + h_2^2}$ as $(h_1, h_2) \to (0, 0)$.

- If we discard the remainders, we get a **linear** system for $\mathbf{h}(t) = \mathbf{x}(t) - \mathbf{x}^*$:

$$\dot{h}_1 = \frac{\partial X_1}{\partial x_1}(a_1, a_2)\, h_1 + \frac{\partial X_1}{\partial x_2}(a_1, a_2)\, h_2,$$

$$\dot{h}_2 = \frac{\partial X_2}{\partial x_1}(a_1, a_2)\, h_1 + \frac{\partial X_2}{\partial x_2}(a_1, a_2)\, h_2,$$

  or in matrix notation,

$$\begin{pmatrix} \dot{h}_1 \\ \dot{h}_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial X_1}{\partial x_1}(a_1, a_2) & \frac{\partial X_1}{\partial x_2}(a_1, a_2) \\ \frac{\partial X_2}{\partial x_1}(a_1, a_2) & \frac{\partial X_2}{\partial x_2}(a_1, a_2) \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}, \qquad \dot{\mathbf{h}} = \underbrace{\frac{\partial \mathbf{X}}{\partial \mathbf{x}}(\mathbf{x}^*)}_{=A}\, \mathbf{h}.$$

[*] What this means is that the quotient $R(h_1, h_2)/\sqrt{h_1^2 + h_2^2}$ tends to zero as $(h_1, h_2) \to (0, 0)$, where $R(h_1, h_2)$ is the remainder.

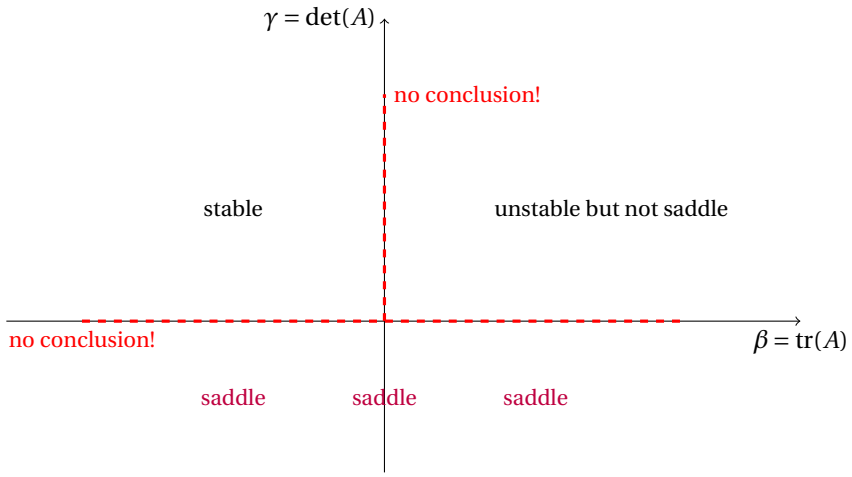
This system is called the **linearization** of the original system at the equilibrium point $\mathbf{x}^*$. Its matrix A is the **Jacobian matrix** of $\mathbf{X}(\mathbf{x})$, *evaluated* at the equilibrium point $\mathbf{x}^*$ (so it's really just a *constant* matrix).

- Our hope is that the linear system $\dot{\mathbf{h}} = A\mathbf{h}$ (which we know how to analyze completely) will tell us something about the behaviour of the original nonlinear system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$ *near* the point $\mathbf{x}^*$.

  And this is indeed the case, provided that $\mathbf{x}^*$ is a **hyperbolic**[†] **equilibrium point**, meaning that the Jacobian matrix $A = \frac{\partial \mathbf{X}}{\partial \mathbf{x}}(\mathbf{x}^*)$ has **no eigenvalues on the imaginary axis** in the complex plane.

  Under that condition, **Theorem 3.3.1** (the **linearization theorem**) says that the nonlinear system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$ is indeed topologically equivalent[‡] to its linearization $\dot{\mathbf{h}} = A\mathbf{h}$ in a neighbourhood of $\mathbf{x}^*$.

  In terms of the trace-determinant diagram of $A = \frac{\partial \mathbf{X}}{\partial \mathbf{x}}(\mathbf{x}^*)$:



The linearization theorem is also called the **Hartman–Grobman theorem**, proved independently by Philip Hartman in the U.S.A. and D. M. Grobman in the Soviet Union around 1960. The proof is rather difficult, and way beyond the scope of this course.[§]

(However, a simpler theorem, not dealing with topological equivalence but only with determining *stability* based on linearization, can be proved using Liapunov's theorems that we will learn about in the next lecture.)

Exercises

- Linearization: 3.5, 3.6, 3.7, A8.
- An isolated, but non-simple, fixed point: 3.8.

[†]The word "hyperbolic" is very over-used in mathematics, and it is perhaps not a very good choice in this context, but unfortunately it has become standard terminology. Arrowsmith & Place avoid this word initially, and express the condition by saying that the linearized system should be **simple** ($\lambda = 0$ is not an eigenvalue) and **not a centre** (we don't have $\lambda = 0 \pm ik$ either). But they introduce it a little later, on p. 80.

[‡]Actually, it's even topologically conjugate, but the book doesn't introduce that concept.

[§]For two-dimensional systems, a stronger result holds (also proved by Hartman): if the vector field $\mathbf{X}(\mathbf{x})$ is of class C^2, then the nonlinear system is actually C^1-conjugate (not just topologically conjugate) to its linearization, and moreover the derivative of the conjugating map at the origin is the identity. This means, for example, that if the linearization is a saddle, the eigenvectors will tell us the correct incoming and outgoing *directions* of the trajectories of the nonlinear system.

It is not true in three or more dimensions that we can always get C^1-conjugacy. However, it *is* true (as was proved as recently as 2003) that the continuous conjugation map is differentiable *at the origin*, with the derivative there equal to the identity. (The proof assumes that $\mathbf{X}(\mathbf{x})$ is of class C^∞, but the claim is conjectured to be true already for class C^2.)

It's also not true (even in two dimensions) that one can get C^k-conjugacy with $k \geq 2$ by assuming the vector field to be nicer; see problem A10 for a counterexample with a vector field of class C^∞ where one doesn't get more than C^1-conjugacy.

- Non-isolated fixed points: 3.11.

  (To avoid confusion: here "a *line* of fixed points" rather means a *curve*.)

- An explicit example of a topological conjugacy: A9*.

- An example regarding smoothness in the Hartman–Grobman theorem: A10**.

Additional problems

A8 Consider the system

$$\dot{x} = 2(x-y)(x+1), \qquad \dot{y} = x - y^2.$$

(a) Linearize the system at the equilibrium point $(0,0)$ and draw the phase portrait of the linearized system in the hk-plane as carefully as you can. (Determine the type of the equilibrium, indicate the principal directions if there are any, take the nullclines into account, etc.)

(b) Do the same for the other equilibrium point $(1,1)$.

(c) Draw the x-nullclines $y = x$ and $x = -1$ for the original nonlinear system, and mark the resulting regions in the xy-plane with "L" and "R" (left/right). Draw the y-nullcline $x = y^2$ in a separate picture, and mark the resulting regions with "D" and "U" (down/up). Overlay the two pictures in a single picture, and mark the regions with "DL", "DR", "UL" or "UR".

(d) Use the information from the previous parts to draw the system's phase portrait as carefully as you can. Include sufficiently many solution curves to give a reasonably complete picture of the system's global behaviour. In particular, your phase portrait should make it clear what happens inside the unit square $[0,1] \times [0,1]$ and in the vicinity thereof. (So don't make your picture too small!) Be careful to make the local phase portraits fit correctly into the global phase portrait – near $(0,0)$ the phase portrait should look like the picture from part (a) and near $(1,1)$ it should look like the picture from part (b).

A9 Find a change of variables $(u,v,w) = h(x,y,z)$ which transforms the nonlinear system

$$\dot{x} = -x, \qquad \dot{y} = -y + xz, \qquad \dot{z} = z$$

into the corresponding linearized system at the origin,

$$\dot{u} = -u, \qquad \dot{v} = -v, \qquad \dot{w} = w.$$

[Answer: For example, $(u,v,w) = (x, y - xz, z)$ works.]

(This shows that the nonlinear system in indeed locally topologically equivalent to its linearization at the hyperbolic equilibrium point $(0,0,0)$, as the linearization theorem promises. In this case, we actually happen to get more than that: the systems are even globally C^∞-conjugate, since the mapping h is bijection of class C^∞ from $\mathbf{R}^3$ to $\mathbf{R}^3$ with inverse of class C^∞.)

A10 (a) Show that the change of variables

$$u = x, \qquad v = y + g(x), \qquad \text{where} \quad g(x) = \begin{cases} x^2 \ln|x|, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

converts the nonlinear system $\dot{x} = -x$, $\dot{y} = -2y + x^2$ to its linearization at the origin, $\dot{u} = -u$, $\dot{v} = -2v$.

(b) Show that the function g belongs to the class $C^1(\mathbf{R})$ but not to $C^2(\mathbf{R})$.
Conclude that the above mapping $(u,v) = (x, y+g(x))$ from the xy-plane to the uv-plane is of class $C^1(\mathbf{R}^2)$, but not of class $C^2(\mathbf{R}^2)$, and likewise for the inverse mapping $(x,y) = (u, v-g(u))$.
(Hence the nonlinear system is C^1-conjugate to its linearization. But not C^2-conjugate; see part (d).)

(c) Write down the explicit solutions for both systems in terms of initial data (x_0, y_0) and (u_0, v_0), respectively, and sketch the phase portraits.

(The solution of the nonlinear system can be obtained either by solving the system directly, or by transforming the solution of the linear system using the change of variables above.)

(d) Finally, prove that there is no C^2-conjugacy between the systems (despite the vector field in the nonlinear system being of class C^∞), by filling in the details in the following outline:

- In order to derive a contradiction, assume that there is such a conjugacy $u = A(x, y)$, $v = B(x, y)$, defined (at least) in some neighbourhood Ω of the origin.
  (That is, assume that the mapping $f = (A, B)$ is of class C^2, with inverse f^{-1} of class C^2, and that it relates the flows of the two systems as in the definition of conjugacy on p. 13).
- Explain why such a mapping must have nonzero Jacobian determinant everywhere. (Hence, in particular, at the origin.)
- Explain why the functions A and B must satisfy the following functional equations for all $(x, y) \in \Omega$ and all t sufficiently close to 0 (so that both sides are defined):

$$A(x, y)\, e^{-t} = A\big(x\, e^{-t}, (y + t x^2)\, e^{-2t}\big),$$
$$B(x, y)\, e^{-2t} = B\big(x\, e^{-t}, (y + t x^2)\, e^{-2t}\big).$$

- Using the assumption that A and B are of class C^2, apply the operator $(\partial/\partial x)^2$ to these identities, and insert $(x, y) = (0, 0)$ afterwards, to get

$$e^{-t} A_{xx}(0,0) = 2t e^{-t} A_y(0,0) + e^{-2t} A_{xx}(0,0),$$
$$e^{-2t} B_{xx}(0,0) = 2t e^{-t} B_y(0,0) + e^{-2t} B_{xx}(0,0).$$

  Conclude, since these identities have to hold for all t in some interval, that $A_y(0,0) = 0$ and $B_y(0,0) = 0$ (and $A_{xx}(0,0) = 0$, but that's not so relevant here).
- This means that the Jacobian determinant is zero at the origin, which is the desired contradiction.

Lecture 6. Stability theorems

(Arrowsmith & Place, sections 3.5, 3.6, 3.7.)

- **Definitions 3.5.1–4**: The definition of what it means for an equilibrium point $\mathbf{x}^*$ to be **(Liapunov) stable**: for every neighbourhood U of $\mathbf{x}^*$ there is a neighbourhood $U' \subseteq U$ of $\mathbf{x}^*$ such that trajectories starting in U' cannot leave U.

  Some stable equilibria are **asymptotically stable**, meaning that (in addition to the above requirement for stability) there is a neighbourhood N of $\mathbf{x}^*$ such that every trajectory starting in N converges to $\mathbf{x}^*$ as $t \to \infty$.

  And those which are stable but not asymptotically stable are called **neutrally stable**. So there are exactly those two types of stable equilibria: asymptotically stable and neutrally stable.

  Unstable equilibrium simply means an equilibrium which is not stable.

- Russian names can be transliterated into the Latin alphabet in many ways. Here I'm writing **Liapunov** like in the textbook, but a very common alternative in English is **Lyapunov**, and one may also come across **Ljapunow**, **Liapounoff**, and so on. Anyway, it's pronounced with the stress on the last syllable: "-OFF".

- **Theorem 3.5.1** is **Liapunov's stability theorem** (1892).

  This theorem is useful for showing stability in situations where linearization is inconclusive. Even more importantly, it also provides a **domain of stability**, which is the textbook's terminology

(although they don't really define it precisely) for a neighbourhood N of an asymptotically stable equilibrium $\mathbf{x}^*$ such that any solution which starts in N stays in N and converges to $\mathbf{x}^*$ as $t \to \infty$. (From linearization we can only say that if all eigenvalues have negative real part, then there is *some* domain of stability, but we don't get any clue about the *size* of that domain.)

However, Arrowsmith & Place don't state exactly how to find a domain of stability, and their proof of the theorem is also rather unclear. I will try to give more precise statements here.

First let us fix some terminology.

Definition. Let I be a proper* interval in $\mathbf{R}$, and let f be a real-valued function whose domain of definition contains I. The function f is said to be **strictly decreasing** on I if

$$f(t_1) > f(t_2)$$

whenever $t_1 \in I$, $t_2 \in I$ and $t_1 < t_2$. It is **weakly decreasing** on I if

$$f(t_1) \geq f(t_2)$$

whenever $t_1 \in I$, $t_2 \in I$ and $t_1 < t_2$.

Remark. Any function which is strictly decreasing is weakly decreasing as well.

Remark. In English it is more common to say **decreasing** and **non-increasing** instead of **strictly decreasing** and **weakly decreasing**, but I have chosen the latter option here to reduce the risk of confusion with the usual terminology in Swedish, which is **strängt/strikt avtagande** and **avtagande**, respectively. (And similarly in German and French.)

From now on we consider some fixed dynamical system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$ in $\mathbf{R}^n$, where the vector field $\mathbf{X}$ is defined in some open set $S \subseteq \mathbf{R}^n$, and we assume that $V\colon \Omega \to \mathbf{R}$ is a differentiable function defined on some open set $\Omega \subseteq S$. (And of course we assume that Ω and S are not the empty set, since that wouldn't be very interesting.)

Definition. The function $\dot{V}\colon \Omega \to \mathbf{R}$ is the dot product of the gradient ∇V and the vector field $\mathbf{X}$:

$$\dot{V}(\mathbf{x}) = \nabla V(\mathbf{x}) \cdot \mathbf{X}(\mathbf{x}), \qquad \mathbf{x} \in \Omega.$$

Theorem. Suppose $\dot{V}(\mathbf{x}) < 0$ for all $\mathbf{x} \in \Omega$. Then, for any solution $\mathbf{x}(t)$ of the system which stays in the set Ω during some nonempty open time interval I, the function

$$f(t) = V(\mathbf{x}(t)), \qquad t \in I$$

is strictly decreasing on I. If instead $\dot{V}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \Omega$, then f is weakly decreasing on I. And if $\dot{V}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \Omega$, then f is constant on I.

Proof. The chain rule gives

$$\begin{aligned} f'(t) &= \nabla V(\mathbf{x}(t)) \cdot \dot{\mathbf{x}}(t) \\ &= \nabla V(\mathbf{x}(t)) \cdot \mathbf{X}(\mathbf{x}(t)) = \dot{V}(\mathbf{x}(t)), \qquad t \in I. \end{aligned}$$

So if we know that $\dot{V} < 0$ everywhere in Ω, and that $\mathbf{x}(t)$ stays in Ω for $t \in I$, then we have $f'(t) < 0$ for $t \in I$, which implies that f is strictly decreasing on I (by a very basic calculus theorem). Similarly if $\dot{V} \leq 0$ or $\dot{V} = 0$. □

Theorem (Liapunov's stability theorem, weak version)**.** Let $\mathbf{x}^*$ be an equilibrium point of the dynamical system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$, $\mathbf{x} \in S \subseteq \mathbf{R}^n$. Suppose that there is a **weak Liapunov function**, i.e., a differentiable function $V\colon \Omega \to \mathbf{R}$ defined on some open set $\Omega \subseteq S$ containing $\mathbf{x}^*$ and satisfying the conditions

*A **proper interval** is an interval (in $\mathbf{R}$) which contains infinitely many points, as opposed to the *degenerate* intervals $[a, a] = \{a\}$ and $\emptyset = \{\}$.

1. $V(\mathbf{x}^*) = 0$, and $V(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega \setminus \{\mathbf{x}^*\}$,
2. $\dot{V}(\mathbf{x}) \leq 0$ for all $\mathbf{x} \in \Omega$.

Then the equilibrium $\mathbf{x}^*$ is **stable**.

Remark. As the theorem is formulated, the function $V\colon \Omega \to \mathbf{R}$ has *exactly* the set Ω as its domain of definition. Usually we have some nice function V (typically a polynomial) which is defined *everywhere* to begin with. What we actually do then is that we compute $\dot{V}$, use that to locate some open set Ω where the assumptions of the theorem are fulfilled, and then apply the theorem with V equal to the *restriction* of the original function V to the set Ω.

Outline of proof. If U is any neighbourhood of $\mathbf{x}^*$, let $B \subset U \cap \Omega$ be a closed ball centered at $\mathbf{x}^*$ and set

$$U' = \{\mathbf{x} \in B : V(\mathbf{x}) < \alpha\},$$

where $\alpha > 0$ is the minimum of V on the boundary sphere ∂B. Then $U' \subset B \subset U$, and U' is a neighbourhood of $\mathbf{x}^*$ such that trajectories starting in U' can't leave U (in fact, they can't even leave U'). □

Detailed proof. Let U be an arbitrary neighbourhood of $\mathbf{x}^*$. To prove stability, we need to find another neighbourhood U' such that solutions starting in U' will never leave U. To find U' we begin by taking a closed ball

$$B = \overline{B(\mathbf{x}^*, \varepsilon)} = \{\mathbf{x} \in \mathbf{R}^n : |\mathbf{x} - \mathbf{x}^*| \leq \varepsilon\}$$

centered at $\mathbf{x}^*$, with radius $\varepsilon > 0$ small enough for B to be contained inside both U and Ω. (This is possible since U and Ω are neighbourhoods of $\mathbf{x}^*$.) The boundary

$$\partial B = \{\mathbf{x} \in \mathbf{R}^n : |\mathbf{x} - \mathbf{x}^*| = \varepsilon\}$$

is a sphere of radius ε centered at $\mathbf{x}^*$. This sphere is a compact set (closed and bounded), and V is continuous by assumption, so according to the extreme value theorem V has a smallest value on ∂B:

$$\alpha = \min_{\mathbf{x} \in \partial B} V(\mathbf{x}).$$

In other words, there is a point $\mathbf{x}_0 \in \partial B$ such that

$$\alpha = V(\mathbf{x}_0) \leq V(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \partial B.$$

Since V is positive definite on Ω (i.e., satisfies condition 1 in the statement of the theorem) and we have chosen B small enough to be a subset of Ω, we have $V(\mathbf{x}) > 0$ for all $\mathbf{x} \in \partial B$, in particular

$$\alpha = V(\mathbf{x}_0) > 0.$$

Now set

$$U' = \{\mathbf{x} \in B : V(\mathbf{x}) < \alpha\}.$$

Then U' contains $\mathbf{x}^*$, since $V(\mathbf{x}^*) = 0 < \alpha$. And U' is an open set, since V is continuous.[†] In other words, U' is an open neighbourhood of $\mathbf{x}^*$. Moreover, a trajectory $\mathbf{x}(t)$ starting in U' (at $t = 0$, say) can't leave U. Here's why: to leave U, the trajectory would have to leave B to begin with (since $U' \subset B \subset U$), and it's a continuous curve so it would have to intersect the boundary sphere ∂B in order to get out. But $f(t) = V(\mathbf{x}(t))$ is a weakly decreasing function of t as long as $\mathbf{x}(t)$ stays in B (since $B \subset \Omega$, and $\dot{V} \leq 0$ in Ω by assumption). Since we start in U', we have $f(0) < \alpha$, and hence

[†]Take any $\mathbf{x} \in U'$, or in other words any $\mathbf{x} \in B$ with $V(\mathbf{x}) < \alpha$. Then actually $\mathbf{x}$ is in the interior of B, since $V \geq \alpha$ on the boundary ∂B. Continuity of V at $\mathbf{x}$ means that there is an open ball $B_2 = B(\mathbf{x}, \delta)$ where $V < \alpha$, and this ball must also be contained in the interior of B, for the same reason. So $B_2 \subseteq U'$. Thus any $\mathbf{x} \in U'$ has an open neighbourhood contained in U', and this is exactly what it means for U' to be open.

$f(t) < \alpha$ for $t \geq 0$. So it's impossible for the trajectory to reach ∂B, since that would mean $f(t) \geq \alpha$ for some $t > 0$. (In fact, the trajectory can't even leave U' – as soon as it did, it would mean that $f(t) \geq \alpha$.)

(One more technical detail: since the trajectory stays inside the compact set B, it must exist for all $t \geq 0$; there can't be any "blowup in finite time". We haven't proved that theorem in this course, so we'll just accept this fact on faith here.) □

Theorem (Liapunov's stability theorem, strong version)**.** Let $\mathbf{x}^*$ be an equilibrium point of the dynamical system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$, $\mathbf{x} \in S \subseteq \mathbf{R}^n$. Suppose that there is a **strong Liapunov function**, i.e., a differentiable function $V : \Omega \to \mathbf{R}$ defined on some open set $\Omega \subseteq S$ containing $\mathbf{x}^*$ and satisfying the conditions

1. $V(\mathbf{x}^*) = 0$ and $V(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega \setminus \{\mathbf{x}^*\}$,
2. $\dot{V}(\mathbf{x}) < 0$ for all $\mathbf{x} \in \Omega \setminus \{\mathbf{x}^*\}$.

Then the equilibrium $\mathbf{x}^*$ is **asymptotically stable**. In fact, for any closed ball $B = \overline{B(\mathbf{x}^*, r)}$ contained in Ω, the set

$$N = \big\{\mathbf{x} \in B : V(\mathbf{x}) < \alpha\big\}, \quad \text{where } \alpha = \min_{\mathbf{x} \in \partial B} V(\mathbf{x}),$$

is a **domain of stability**: solutions starting in N stay in N, and converge to $\mathbf{x}^*$ as $t \to \infty$.

Remark. If $\Omega = \mathbf{R}^n$ and if the additional condition

$$V(\mathbf{x}) \to \infty \quad \text{as} \quad |\mathbf{x}| \to \infty$$

holds, then for *any* given point $\mathbf{x}_0 \in \mathbf{R}^n$ it is true that

$$\min_{\mathbf{x} \in \partial B} V(\mathbf{x}) > V(\mathbf{x}_0)$$

if we take B large enough. This means that $\mathbf{x}_0 \in N$ for that choice of B, causing the trajectory starting at $\mathbf{x}_0$ to converge to $\mathbf{x}^*$. So in this case $\mathbf{x}^*$ is stable and *every* trajectory of the system converges to $\mathbf{x}^*$, which is expressed by saying that $\mathbf{x}^*$ is **globally asymptotically stable**.

Remark. If one studies the proof, it should be clear that the set B in the definition of N doesn't really have to be precisely a closed ball, just something which is topologically equivalent to a ball. For example, if our Liapunov function is $V(x, y) = x^6 + y^4$, then for $k > 0$ the sublevel set $B = \{x^6 + y^4 \leq k\}$ is sufficiently "ball-like" for the proof to work: no continuous curve can pass from the interior to the exterior without crossing the closed level curve $\partial B = \{x^6 + y^4 = k\}$. The same arguments as in the proof then show that trajectories can't leave the set $N = \{\mathbf{x} \in B : V(x, y) < k\} = \{x^6 + y^4 < k\}$, so provided that B is contained in the region Ω, N is a domain of stability for the equilibrium $(0, 0)$.

Outline of proof. Stability follows from the weak version of Liapunov's theorem. Any trajectory starting in N must stay in N, and along such a trajectory the function V decreases strictly towards some limit $L \geq 0$. But $L > 0$ would contradict the continuity of the flow φ_t, so $L = 0$, which in turn implies that the trajectory converges to $\mathbf{x}^*$ (the only point where $V = 0$). □

Detailed proof. Stability follows from the weak version of Liapunov's theorem. As in the proof of that theorem, we see that N (as defined above) is an open neighbourhood of $\mathbf{x}^*$, and that any trajectory $\mathbf{x}(t)$ starting in N stays in N and is defined for all $t \geq 0$. For the constant solution $\mathbf{x}(t) = \mathbf{x}^*$ there is nothing to prove – of course it converges to $\mathbf{x}^*$! So suppose $\mathbf{x}(t)$ is some *other* solution starting in N. Then, by the assumption $\dot{V} < 0$, $V(\mathbf{x}(t))$ is a strictly decreasing function of t on the interval $t \geq 0$, and it's bounded below (since $V \geq 0$), so it has a limit $L \geq 0$ as $t \to \infty$.

We want to show that $L = 0$, so assume $L > 0$ in order to get a contradiction. Take any sequence of positive numbers $t_n \nearrow \infty$; then $\mathbf{x}_n = \mathbf{x}(t_n)$ is a sequence of points in the compact set B. According to the Bolzano–Weierstrass theorem (a standard theorem about compact sets in $\mathbf{R}^n$), this sequence

of points must have a convergent subsequence, i.e., there is a point $\mathbf{y} \in B$ and an integer sequence $n_k \nearrow \infty$ such that $\mathbf{x}_{n_k} \to \mathbf{y}$ as $k \to \infty$. Since V is continuous, we can move the limit outside V and obtain

$$V(\mathbf{y}) = V\left(\lim_{k\to\infty} \mathbf{x}_{n_k}\right) = \lim_{k\to\infty} V(\mathbf{x}_{n_k}) = L.$$

We are assuming $L > 0$, which means that $\mathbf{y} \neq \mathbf{x}^*$ (since V is positive definite), and V will thus continue to decrease *strictly* along the trajectory starting at $\mathbf{y}$. So the flow φ_1, for example (or φ_t for any fixed $t > 0$), will map $\mathbf{y}$ to a point where $V < L$. But φ_1 is a continuous function, so it will also map all sufficiently nearby points $\mathbf{x}_{n_k}$ to points where $V < L$:

$$V(\varphi_1(\mathbf{x}_{n_k})) = V(\mathbf{x}(1 + t_{n_k})) < L, \qquad \text{for all sufficiently large } k.$$

But we know that $V(\mathbf{x}(t)) > L$ for all $t \geq 0$, since $V(\mathbf{x}(t))$ is *decreasing* towards the limit L. This contradiction shows that the assumption $L > 0$ must have been incorrect. Hence $L = 0$.

Now we know that $V(\mathbf{x}(t)) \searrow L = 0$ as $t \to \infty$. It remains to show that this implies $\mathbf{x}(t) \to \mathbf{x}^*$, i.e., that for any $\varepsilon > 0$ there is a time τ such that $\mathbf{x}(t) \in B_\varepsilon$ for all $t > \tau$, where $B_\varepsilon = B(\mathbf{x}^*, \varepsilon)$ is the open ball of radius ε centered at $\mathbf{x}^*$. We may assume that $0 < \varepsilon < r$, where r is the radius of the closed ball B. Then

$$B \setminus B_\varepsilon = \overline{B(\mathbf{x}^*, r)} \setminus B(\mathbf{x}^*, \varepsilon)$$

is a compact nonempty set, so the continuous function V has a smallest value β on this set (and $\beta > 0$ since V is positive definite). What this means is that if $\mathbf{x} \in B$ and $V(\mathbf{x}) < \beta$, then $\mathbf{x} \in B_\varepsilon$. But we have $\mathbf{x}(t) \in B$ for all $t \geq 0$, and since $V(\mathbf{x}(t)) \searrow 0$ as $t \to \infty$ there is a τ such that $V(\mathbf{x}(t)) < \beta$ for $t > \tau$. Consequently $\mathbf{x}(t) \in B_\varepsilon$ for $t > \tau$, as desired. □

- The very useful **Theorem 3.5.2** is an improvement of Liapunov's theorem which is due to LaSalle (1960). It allows us to conclude *asymptotic* stability using only a *weak* Liapunov function, provided an additional condition is satisfied. The proof is not given in the book, but it is a consequence of something called **LaSalle's invariance principle** (see the next lecture).

  Theorem (LaSalle's stability theorem)**.** Let $\mathbf{x}^*$ be an equilibrium point of the dynamical system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$, $\mathbf{x} \in S \subseteq \mathbf{R}^n$. Suppose that there is a **weak Liapunov function** $V : \Omega \to \mathbf{R}$ on some open set $\Omega \subseteq S$ containing $\mathbf{x}^*$, and in addition suppose that the set

  $$\left\{\mathbf{x} \in \Omega : \dot{V}(\mathbf{x}) = 0\right\}$$

  contains no complete trajectory except the constant solution $\mathbf{x}^*$.

  Then the equilibrium $\mathbf{x}^*$ is **asymptotically stable**, and N (defined as in the strong version of Liapunov's theorem) is a **domain of stability**.

  (And if $\Omega = \mathbf{R}^n$ and $V(\mathbf{x}) \to \infty$ as $|\mathbf{x}| \to \infty$, then $\mathbf{x}^*$ is **globally asymptotically stable**.)

  Proof. See the next lecture. □

- Here is a somewhat subtle point concerning the above theorems and weak Liapunov functions. If we have a function V which is everywhere defined and nice (continuously differentiable), then the set $\Omega_1 = \left\{\mathbf{x} \in \mathbf{R}^n : \dot{V}(\mathbf{x}) \leq 0\right\}$ will be *closed*. But the theorems, as they are formulated above, require us to restrict V to an *open* set Ω. So we have to shrink Ω_1 "by hand" to get an *open* set Ω where $\dot{V} \leq 0$ holds. The purpose of this, as well as all the business with closed balls contained inside Ω, is to avoid accidentally making plausible-sounding claims which are actually false. We know that V is weakly decreasing along trajectories, but only as long as they stay in Ω, so we need to take some precautions to prevent the trajectories from sneaking out of Ω!

  Example. If the system is $\dot{x} = y$, $\dot{y} = -x - y(1 - x^2)$, and if $V(x, y) = x^2 + y^2$, then $\dot{V} = -2y^2(1 - x^2)$. Thus, the set $\Omega_1 = \{\dot{V} \leq 0\}$ is the union of the closed strip $-1 \leq x \leq 1$ and the line $y = 0$. So any trajectory will be moving closer to the origin (or at least not further away from it) as long as it is

inside the strip, but trajectories for $y > 3$ (or so) will enter the strip from the left, *leave it again on the right* (a little further down), and then go off steeply upwards towards infinity instead of converging towards the equilibrium $(0,0)$. So for example, a set like $\{(x,y) \in \Omega_1 : V(x,y) \le 100\}$ is not forward invariant despite V being weakly decreasing on trajectories in Ω_1! But we can take Ω to be the open strip $-1 < x < 1$, let B be any closed ball $x^2 + y^2 \le k$ with $0 < k < 1$ so that it fits inside Ω, and then the set N given by $x^2 + y^2 < k$ will be a domain of attraction by LaSalle's theorem, since there are no trajectories contained in the line $y = 0$ except the equilibrium solution $(x(t), y(t)) = (0,0)$. And since this is true for any $0 < k < 1$, in fact the open unit disk $x^2 + y^2 < 1$ is a domain of attraction.[‡]

- **Theorem 3.5.3** can be formulated as follows:

  Theorem (Liapunov's instability theorem)**.** Let $\mathbf{x}^*$ be an equilibrium point of the dynamical system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$, $\mathbf{x} \in S \subseteq \mathbf{R}^n$. Suppose there is a differentiable function $V : \Omega \to \mathbf{R}$ defined on some open set $\Omega \subseteq S$ containing $\mathbf{x}^*$ and satisfying the conditions

  1. $V(\mathbf{x}^*)$ is not a local maximum,
  2. $\dot{V}(\mathbf{x}) > 0$ for all $\mathbf{x} \in \Omega \setminus \{\mathbf{x}^*\}$.

  Then the equilibrium $\mathbf{x}^*$ is **unstable**.

  Idea of proof. For any closed ball $N = \overline{B(\mathbf{x}^*, r)} \subset \Omega$, there is a point $\mathbf{x}_0$ in the interior of N such that $V(\mathbf{x}_0) > V(\mathbf{x}^*)$, and it is shown that the trajectory starting at such a point $\mathbf{x}_0$ must leave N. □

- We have relied upon the rather deep Hartman–Grobman theorem to show that an equilibrium is asymptotically stable if the linearization there is asymptotically stable. This fact can be proved more directly using Liapunov's stability theorem. The simplest case is when the Jacobian matrix $A = \frac{\partial \mathbf{X}}{\partial \mathbf{x}}(\mathbf{x}^*)$ has distinct real eigenvalues (assumed negative, in order for $d\mathbf{h}/dt = A\mathbf{h}$ to be asymptotically stable):

  $$0 > \lambda_1 > \cdots > \lambda_n.$$

  Make the usual linear change of coordinates $\mathbf{x} = M\mathbf{y}$, where the columns of M form a basis of eigenvectors of A. In terms of these coordinates, we have a system $\dot{\mathbf{y}} = \mathbf{Y}(\mathbf{y})$ with an equilibrium $\mathbf{y}^*$ where the Jacobian is diagonal:

  $$J = \frac{\partial \mathbf{Y}}{\partial \mathbf{y}}(\mathbf{y}^*) = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

  So $\mathbf{k} = \mathbf{y} - \mathbf{y}^*$ satisfies

  $$\dot{\mathbf{k}} = J\mathbf{k} + \text{remainder},$$

  where the remainder tends to zero faster than $|\mathbf{k}|$. It's not too difficult to check that $V(\mathbf{k}) = \sum k_i^2$ is a strict Liapunov function for this system in a neighbourhood of $\mathbf{k} = \mathbf{0}$, which proves asymptotic stability. With repeated and/or non-real eigenvalues things are a bit more complicated, but if all eigenvalues have negative real part, one can find a strong Liapunov function in the form of a sum of squares in those cases too.

  Similarly, one can use Liapunov's instability theorem to prove that if some eigenvalue has positive real part, then the equilibrium is unstable.

- The **flow box theorem**. Global phase portraits.

[‡] We can actually do yet a little better: the closed unit disk $x^2 + y^2 \le 1$ is forward invariant since the open unit disk is; this follows from the continuity of the flow, since if φ_t (for some $t > 0$) would map some point on the unit circle to a point outside the circle, then by continuity it would also have to map some nearby point inside the circle out of the circle, which we know it doesn't. So the closed unit disk is compact and forward invariant, and then we can try applying LaSalle's invariance principle (see next lecture) to it. This turns out to be successful (exercise), so actually the *closed* unit disk is a domain of attraction.

[Jump to course outline]

- **First integrals**, also known as **constants of motion**, **integrals of motion**, **conserved quantities**, **invariants**, etc.

  (Can sometimes be found by writing $dy/dx = \dot{y}/\dot{x} = Y(x,y)/X(x,y)$ and solving the resulting ODE for $y = y(x)$.)

- If $H(q,p)$ is any C^1 function, then the **Hamiltonian system**

$$\begin{pmatrix}\dot{q}\\ \dot{p}\end{pmatrix} = \begin{pmatrix}\partial H/\partial p\\ -\partial H/\partial q\end{pmatrix} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\partial H/\partial q\\ \partial H/\partial p\end{pmatrix} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\nabla H(q,p)$$

  automatically has H as a first integral. A fundamental fact in mechanics is that the Hamiltonian system generated by $H(q,p) = \frac{1}{2}p^2 + V(q)$ is equivalent to the Newton-type equation $\ddot{q} = -V'(q)$. (This works also in higher dimensions, with vectors **p** and **q** instead.)

Exercises

- Using strong Liapunov functions: 3.13abe, 3.14abe.
- Using weak Liapunov functions: 3.15, 3.17bc, A11.
- Finding a domain of stability using a less obvious Liapunov function: 3.19b.

  Warning! Please think *very* carefully before handing in your solution! You will need to use LaSalle's theorem, not Liapunov's. And regarding the domain of stability N, you are required to motivate your answer by giving a detailed description following the "recipe" for N described in the statements of LaSalle's theorem and the strong version of Liapunov's theorem above. (For the best results, let B be a "topological ball" $V(x,y) \le k$ rather than an actual ball $x^2 + y^2 \le k$; cf. the second remark below the strong Liapunov theorem above.)

- And another one: 3.18*.

  Note that there is a **sign error** in the ODE for x_2; the system is supposed to be

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

  [Remark: The answer given in the book is not the only possible one. Since

$$-\tfrac{1}{2}\dot{V} = x_1^2 + x_2^2 + (3-a)x_1x_2 + (1-x_2^2)(bx_1 + cx_2)(x_1 + 2x_2),$$

  it's quite natural to pick $b = 1$ and $c = 2$ to get a term $(1-x_2^2)(x_1+2x_2)^2$ whose sign we have control over, but then we can take any $a \in [1,5]$ to make $\dot{V}$ negative definite in the strip $|x_2| < 1$ (and note also that all these values of a satisfy the conditions $a > 0$ and $2a = ac > b^2 = 1$ for making V positive definite). You might want to draw the phase portrait and your domain of stability on the computer! And why not try this for different values of a? If you take the *union* of the different domains of stability for $1 \le a \le 5$ you get a bigger (=better) domain of stability!]

- And another: A12*.
- An example that $V(\mathbf{x}) \to \infty$ is important in order to get global asymptotic stability: A13.
- An example[§] illustrating why the requirement about stability in the definition of asymptotic stability is necessary: A14.

[§]The system (3.33) in the book is another such example, but that one is much more difficult to analyze rigorously (Exercise 3.12); for details about this, see Section 40 of the book *Stability of Motion* by W. Hahn (Springer, 1967).

- Showing instability: 3.22.

  (First do it as the book suggests, using Liapunov's instability theorem. Then find a much simpler way of showing that the origin is unstable!)

- And one more about Liapunov's instability theorem: A15*.

- Illustration of an explicit transformation which straightens out a nonlinear vector field: 3.24.

- First integrals (constants of motion): 3.28ab, 3.29*.

  In 3.28b, you can do better than the answer in the book, and find a constant of motion which is actually defined for *all* x_1 and x_2!

  In 3.29, you should do much better than the answer in the book, which is actually wrong. A correct constant of motion is $F(x_1, x_2) = x_1(x_1^2 - x_2)/(x_1^2 + x_2)^2$. The level curves of this function are not easy to plot by hand, but you can of course do it on the computer if you want to see what the phase portrait looks like.

Additional problems

A11 Draw the phase portrait (first by hand, and then on the computer for verification) for the system

$$\dot{x} = y, \qquad \dot{y} = -x - y(1 - x^2)$$

from the example above (on p. 21). In the same picture, draw some level sets of the weak Liapunov function $V(x, y) = x^2 + y^2$ and indicate the sets where $\dot{V} < 0$ and $\dot{V} = 0$. Make sure that you understand why the strip $|x| < 1$ is *not* a domain of stability!

A12 Here is an algorithm for finding Liapunov functions for *linear* $n \times n$ systems $d\mathbf{x}/dt = A\mathbf{x}$ such that all eigenvalues of A have negative real part (so that the origin is asymptotically stable):

> Take an arbitrary positive definite $n \times n$ matrix Q (symmetric), and solve for the symmetric $n \times n$ matrix P in the **Liapunov equation**
>
> $$A^T P + PA = -Q.$$
>
> Then the matrix P will be positive definite,[¶] and
>
> $$V(\mathbf{x}) = \mathbf{x}^T P \mathbf{x}$$
>
> will be a strong Liapunov function with $\dot{V}(\mathbf{x}) = -\mathbf{x}^T Q \mathbf{x}$.

Try this out on the system

$$\dot{x} = x + y, \qquad \dot{y} = -5x - 2y,$$

with the positive definite matrix $Q = \mathrm{diag}(2, 4)$. (If you're lazy, you can get some help from Wolfram Alpha.) Verify that the function V that you obtain really is a strong Liapunov function for this system!

[Answer: $V(x, y) = 19x^2 + 8xy + 3y^2$.]

A13 Show that

$$V(x, y) = \frac{x^2}{1 + x^2} + y^2$$

is a global strong Liapunov function for the system

$$\dot{x} = -x(1 - 3x^2 y^2), \qquad \dot{y} = -y(1 + x^2 y^2),$$

[¶] If you do this for a system where the origin is not asymptotically stable, then the matrix P that you get will not be positive definite.

but that the origin is **not** globally asymptotically stable (although it is locally asymptotically stable).

Note that $V(x,y)$ does *not* satisfy the condition that $V(x,y) \to \infty$ as $\sqrt{x^2+y^2} \to \infty$. What do the level sets of V look like, in particular the level set $V = 1$?

[Hints: One can write $\dot{V}$ as

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

and $\big(x(t), y(t)\big) = (e^{2t}, e^{-2t})$ is a particular solution of the system.]

A14 Consider the system

$$\dot{x} = x - y - xr^2 + \frac{xy}{r}, \qquad \dot{y} = x + y - yr^2 - \frac{x^2}{r},$$

where $r = \sqrt{x^2+y^2}$, and where the right-hand sides are interpreted as zero when $(x,y) = (0,0)$.[‖] Rewrite this system in polar coordinates, and use this to draw the phase portrait.

Deduce that, with the exception of the equilibrium solution $(x,y) = (0,0)$, all solutions $(x(t), y(t))$ approach the point $(x,y) = (1,0)$ as $t \to \infty$, but $(1,0)$ is still an **unstable** equilibrium.

[Answer: $\dot{r} = r - r^3$, $\dot{\theta} = 1 - \cos\theta$.]

A15 Let $V(x,y) = -(y-x^3)(y-x^5)$, and consider the system

$$\dot{x} = \partial V/\partial x = -8x^7 + 3x^2y + 5x^4y, \qquad \dot{y} = \partial V/\partial y = x^3 + x^5 - 2y.$$

(a) Have a look at the phase portrait on the computer (Wolfram Alpha link). What do you think – does the origin look stable or unstable?

(b) Prove that the origin is in fact unstable, using Liapunov's instability theorem with the given function V.

Lecture 7. Limit sets

(Arrowsmith & Place, sections 3.8, 3.9.)

- **Definition 3.8.1** and the paragraph just below it:

  The point **y** is an **α-limit point** of a point **x** if there is a sequence $t_n \to -\infty$ such that $\varphi_{t_n}(\mathbf{x}) \to \mathbf{y}$.

  The **α-limit set** $L_\alpha(\mathbf{x})$ is the set of α-limit points of **x**.

  With $t_n \to +\infty$ instead, we obtain **ω-limit points**, and the **ω-limit set** $L_\omega(\mathbf{x})$.

  (As you might know, α and ω are the first and last letters of the Greek alphabet; cf. the following well-known passage from the Bible (Rev. 22:13): "I am the Alpha and the Omega, the first and the last, the beginning and the end." The terminology for limit sets is meant to convey the idea that $L_\alpha(\mathbf{x})$ and $L_\omega(\mathbf{x})$ give information about how the orbit of **x** behaves at the "beginning" of time ($t \to -\infty$) and at the "end" of time ($t \to +\infty$).)

- In three or more dimensions, limit sets can be extremely complicated, since trajectories have room to wind around in space in very strange ways. But in the plane, the possibilities are much more restricted, as shown by **Theorem 3.9.1**, the **Poincaré–Bendixson theorem** (given without proof in the textbook[*]):

[‖] This system is a bit nasty at the origin, since the right-hand sides are not differentiable there, only continuous. If you prefer, you can get a system of class C^1 by multiplying both right-hand sides by r^2; this rescaling of the vector field doesn't change what the phase portrait looks like.

[*] For proofs, see for example H. Amann, *Ordinary Differential Equations*, de Gruyter (1990), p. 333, or C. Chicone, *Ordinary Differential Equations with Applications*, Second Edition, Springer (2006), p. 101.

> If a compact nonempty limit set in the plane contains no equilibrium points, then it must be a periodic orbit.

(There is also a more general version of the theorem, which says what can happen if the limit set contains finitely many equilibrium points; see Wikipedia: Poincaré–Bendixson theorem, for example.)

- Some properties of ω-limit sets:
  - $L_\omega(\mathbf{x})$ is always a **closed** and **invariant** set. (See **Definition 3.9.2**.)
  - Limit sets may be empty, unbounded, disconnected (see problem A16). But if the forward orbit of $\mathbf{x}$ is **bounded**, then $L_\omega(\mathbf{x})$ is **connected**, **compact** and **non-empty**, and[†]

$$\varphi_t(\mathbf{x}) \to L_\omega(\mathbf{x}) \qquad \text{as } t \to \infty.$$

The corresponding properties hold of course for α-limit sets as $t \to -\infty$.

Proof of the invariance property. Suppose $\mathbf{y} \in L_\omega(\mathbf{x})$. By definition, this means that $\varphi_{t_n}(\mathbf{x}) \to \mathbf{y}$ for some sequence $t_n \to \infty$. Fix an arbitrary $t \in \mathbf{R}$. Applying the continuous function φ_t to both sides gives $\varphi_{t+t_n}(\mathbf{x}) \to \varphi_t(\mathbf{y})$, so $\varphi_t(\mathbf{y}) \in L_\omega(\mathbf{x})$. □

(I omit the proofs of the other properties, although they are not very difficult.)

- A **limit cycle** (**Definition 3.8.2**) is a periodic orbit which lies in the α- or ω-limit set of some point not on the orbit.

  To show the existence of a limit cycle for a planar system using the Poincaré–Bendixson theorem, one tries to find a **trapping region containing no equilibria**. A **trapping region** for a system with flow φ_t is a compact, connected set $D \subset \mathbf{R}^2$ such that $\varphi_t(D) \subset D$ for $t > 0$. When reading this definition, it's important to note that Arrowsmith & Place use the symbol "$\subset$" for *strict* set inclusion. The point is that if we only require D to be forward invariant (that is, $\varphi_t(D) \subseteq D$ for $t > 0$, with non-strict set inclusion "$\subseteq$"), then it may be the case that D is a union of periodic orbits[‡], in which case there are no limit cycles in D. But with strict set inclusion, the points in the nonempty set $D \setminus \varphi_t(D)$ (for any fixed $t > 0$) can't lie on a periodic orbit[§], and the ω-limit set of such a point $\mathbf{x}_0$ is a nonempty compact subset of D. If we have chosen the trapping region D such that it contains no equilibria, the Poincaré–Bendixson theorem says that $L_\omega(\mathbf{x}_0)$ must be a periodic orbit. Since $\mathbf{x}_0$ was not on any periodic orbit, $L_\omega(\mathbf{x}_0)$ doesn't contain $\mathbf{x}_0$, so it is a periodic orbit which is the ω-limit set of a point not on the orbit – in other words, it's an ω-limit cycle. Conclusion: the trapping region D contains *at least one* limit cycle.

  (But there may be many limit cycles in D! To prove that there is *at most one* limit cycle is usually much more difficult.)

- Now that we know what an ω-limit set $L_\omega(\mathbf{x})$ is, we can state **LaSalle's invariance principle** that was mentioned in the previous lecture. We consider a system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$ with flow φ_t.

  Theorem. Suppose $V : \Omega \to \mathbf{R}$ is differentiable on the open set $\Omega \subseteq \mathbf{R}^n$, and satisfies $\dot{V}(\mathbf{y}) \le 0$ for each $\mathbf{y}$ in some closed set $M \subseteq \Omega$.

[†] The notation $\varphi_t(\mathbf{x}) \to L_\omega(\mathbf{x})$ means that for any open set $U \supset L_\omega(\mathbf{x})$ there is a time T such that $\varphi_t(\mathbf{x}) \in U$ for all $t > T$.

[‡] For example (in polar coordinates), if the system is $\dot{r} = 0$, $\dot{\theta} = 1$, and D is the annulus $1 \le r \le 2$.

[§] Suppose that $\mathbf{x}_0 \in D \setminus \varphi_t(D)$ for some $t > 0$ and that $\mathbf{x}_0$ lies on a periodic orbit with period $T > 0$. Let n be a positive integer such that $nT > t$. Then

$$\mathbf{y} = \varphi_{nT-t}(\mathbf{x}_0) \in \varphi_{nT-t}(D) \subset D$$

so that

$$\mathbf{x}_0 = \varphi_{nT}(\mathbf{x}_0) = \varphi_t(\mathbf{y}) \in \varphi_t(D),$$

which is a contradiction.

1. If $\mathbf{x}$ is a point in M whose forward orbit $\mathcal{O}^+(\mathbf{x})$ never leaves M, then there is an $\alpha \in \mathbf{R}$ such that

$$L_\omega(\mathbf{x}) \subseteq \{\mathbf{y} \in M : V(\mathbf{y}) = \alpha\}.$$

This implies that $L_\omega(\mathbf{x})$ is an invariant set contained in the set

$$C = \{\mathbf{y} \in M : \dot{V}(\mathbf{y}) = 0\},$$

and hence $L_\omega(\mathbf{x}) \subseteq E$, where E is the *largest invariant subset* of C (i.e., E is the union of all trajectories which stay in C for all $t \in \mathbf{R}$).

Proof. This is trivially true if $L_\omega(\mathbf{x}) = \emptyset$, since the empty set is a subset of *every* set. So assume $L_\omega(\mathbf{x}) \neq \emptyset$. (In particular, this assumption entails that the solution $\varphi_t(\mathbf{x})$ exists for all $t \geq 0$.) To begin with,

$$\begin{aligned} L_\omega(\mathbf{x}) &\subseteq \overline{\mathcal{O}^+(\mathbf{x})} && \text{(follows from def. of } L_\omega(\mathbf{x})) \\ &\subseteq \overline{M} && \text{(since } \mathcal{O}^+(\mathbf{x}) \subseteq M \text{ by assumption)} \\ &= M && \text{(since } M \text{ is closed).} \end{aligned}$$

Next, the assumptions that $\dot{V} \leq 0$ on M and that $\varphi_t(\mathbf{x})$ stays in M for $t \geq 0$ imply that $V(\varphi_t(\mathbf{x}))$ is a weakly decreasing function of t for $t \geq 0$, so the limit

$$\alpha = \lim_{t \to \infty} V(\varphi_t(\mathbf{x}))$$

exists, either as a real number $\alpha \in \mathbf{R}$ or in the improper sense $\alpha = -\infty$. But if $\mathbf{y}$ is any element in the nonempty set $L_\omega(\mathbf{x})$, meaning that $\varphi_{t_n}(\mathbf{x}) \to \mathbf{y}$ for some sequence $t_n \nearrow \infty$, then

$$V(\mathbf{y}) = V\Big(\lim_{n \to \infty} \varphi_{t_n}(\mathbf{x})\Big) = \lim_{n \to \infty} V\big(\varphi_{t_n}(\mathbf{x})\big) = \alpha,$$

since V is continuous. This shows that α equals the real number $V(\mathbf{y})$, not $-\infty$.

The above calculation holds for an arbitrary $\mathbf{y} \in L_\omega(\mathbf{x})$, so $V = \alpha$ on all of $L_\omega(\mathbf{x})$. And $L_\omega(\mathbf{x})$ is an invariant set (general property of limit sets), so $\varphi_t(\mathbf{y}) \in L_\omega(\mathbf{x})$ for all t if $\mathbf{y} \in L_\omega(\mathbf{x})$. Thus $V(\varphi_t(\mathbf{y})) = \alpha$ for all t, and hence

$$\dot{V}(\mathbf{y}) = 0 \quad \text{if } \mathbf{y} \in L_\omega(\mathbf{x}).$$

What we have shown now is that $L_\omega(\mathbf{x})$ is an invariant set which is contained in the set $C \subseteq M$ where $\dot{V} = 0$. Therefore, it must trivially be contained in E, the *largest* invariant set contained in C. □

2. If moreover the forward orbit $\mathcal{O}^+(\mathbf{x})$ is **bounded**, then $L_\omega(\mathbf{x})$ is nonempty and $\varphi_t(\mathbf{x}) \to L_\omega(\mathbf{x})$ as $t \to \infty$. So $\varphi_t(\mathbf{x}) \to E$ as $t \to \infty$.

Proof. The first sentence was one of the general properties of limit sets stated at the beginning of the lecture. The second sentence follows at once from the property $L_\omega(\mathbf{x}) \subseteq E$ that we proved in item 1. (The conclusion that $\varphi_t(\mathbf{x}) \to E$ is of course a bit weaker than $\varphi_t(\mathbf{x}) \to L_\omega(\mathbf{x})$, but the point is that the set E does not depend on $\mathbf{x}$.) □

3. If M is **compact** and **forward invariant**, then items 1 and 2 apply to *every* point $\mathbf{x} \in M$. So in this case, $\varphi_t(\mathbf{x}) \to E$ as $t \to \infty$, for every $\mathbf{x} \in M$.

Proof. Trivial. □

The point of this theorem is that the set E is often quite easy to determine. We find the set C simply by computing $\dot{V}$ and checking where it's zero. This is typically some curve, if we are in $\mathbf{R}^2$. Then we study what the vector field $\mathbf{X}$ is doing at each point of the set C – if the vector field is pointing out from C at some point, then that point can't be part of a trajectory completely contained in C, so it can't belong to E.

A typical application is the situation described in **Theorem 3.5.2** (see the previous lecture), where we have only managed to find a *weak* Liapunov function V, but the "bad" set C where we have $\dot{V} = 0$ instead of $\dot{V} < 0$ doesn't contain any trajectories except the equilibrium point $\mathbf{x}^*$.

[Jump to course outline]

Proof of Theorem 3.5.2. As usual, let $B = \overline{B(\mathbf{x}^*, r)}$ be a closed ball (or some other neighbourhood of $\mathbf{x}^*$ topologically equivalent to a closed ball) contained in Ω, and define

$$N = \{\mathbf{x} \in B : V(\mathbf{x}) < \alpha\}, \quad \text{where } \alpha = \min_{\mathbf{x} \in \partial B} V(\mathbf{x}) > 0.$$

Stability of $\mathbf{x}^*$ follows from the weak version of Liapunov's theorem, so we just need to show that N is a domain of stability. To apply LaSalle's invariance principle, we need a compact and forward invariant set M, so N (which is open) won't do. Instead, take β with $0 \le \beta < \alpha$, and let

$$M = \{\mathbf{x} \in B : V(\mathbf{x}) \le \beta\}.$$

Then M is closed (and hence compact) since V is continuous.[¶] And it's forward invariant; indeed, a trajectory starting in M can't leave M, since then it would enter the part of B where $V > \beta$, so V wouldn't be weakly decreasing along that trajectory, and that would contradict the assumption that $\dot{V} \le 0$ in Ω. Now the invariance principle says that every trajectory starting in M converges to E, the largest invariant subset of $C = \{\mathbf{x} \in M : \dot{V}(\mathbf{x}) = 0\}$. But by assumption, there are no trajectories even in the larger set $C_2 = \{\mathbf{x} \in \Omega : \dot{V}(\mathbf{x}) = 0\}$ except for the equilibrium $\mathbf{x}^*$. Hence $E = \{\mathbf{x}^*\}$, and every trajectory starting in M converges to $\mathbf{x}^*$. So every such set M is a domain of stability.

To show that N is a domain of stability, just note that any $\mathbf{x}_0 \in N$ belongs to the set $M \subseteq N$ defined using $\beta = V(\mathbf{x}_0) < \alpha$. Therefore the trajectory starting at $\mathbf{x}_0$ stays in N and converges to $\mathbf{x}^*$. □

- The **Poincaré map** (or **first-return map**) associated with a periodic orbit is defined on p. 104.
- **Theorem 3.9.2** is called the **Bendixson criterion**.

  A simple generalization (with virtually the same proof) is the **Bendixson–Dulac criterion**, which gives the same conclusion provided that that there is a function $f(x_1, x_2)$ of class C^1 such that the divergence of the rescaled vector field $f\mathbf{X}$,

  $$\nabla \cdot (f\mathbf{X}) = \frac{\partial}{\partial x_1}(fX_1) + \frac{\partial}{\partial x_2}(fX_2),$$

  is of constant sign (positive or negative) in D.

  (One can in fact weaken the hypotheses, by allowing $\nabla \cdot (f\mathbf{X})$ to be zero on a set of measure zero; this does not alter the fact that the double integral in the proof must be nonzero.)

Exercises

- α- and ω-limit sets: 3.35, A16*.

  (There's an error in the answer to 3.35b; it should say $L_\alpha(\mathbf{x}) = \{\mathbf{0}\}$ for $0 < r < 1$.)
- Poincaré map: 3.36.

  (In this problem, it's understood that $a > 0$.)
- Invariant sets, trapping regions: 3.42, 3.43, A17.
  - Don't miss the follow-up question that's formulated at the end of problem 3.42, after part (e).
  - **Problem 3.42a is incorrect**; the upper half-plane $x_2 \ge 0$ is actually **not** a positively invariant set for that system! The reason is rather subtle; can you see what it is that goes wrong compared to Definition 3.9.2?

[¶] If $\mathbf{x}_n$ is a sequence of points in M converging to $\mathbf{x}$, then $\mathbf{x} \in \overline{B} = B$, and by continuity

$$\underbrace{V(\mathbf{x}_n)}_{\le \beta} \to V(\mathbf{x}),$$

so $V(\mathbf{x}) \le \beta$. Hence $\mathbf{x} \in M$, which means that M is closed.

  - In problem 3.42d, it's fine to "cheat" a little by using the computer to draw the level curves of $3(x_1^2+x_2^2)-2x_1^3$.
  - In the last part of problem 3.43 ("Show that the system has a limit cycle when $F=0$") it is assumed that $w \neq 0$.

- The Bendixson criterion: A18.

Additional problems

A16 (a) Show that for the equation $\dot{x}=1$, the ω-limit set of each point is the **empty** set.

(b) Sketch the phase portrait for the system

$$\dot{x} = -y + \frac{x}{1+x^2}, \qquad \dot{y} = x(1-y^2).$$

Show that the ω-limit set of any point $(x,y) \neq (0,0)$ in the strip $|y|<1$ is the union of the lines $y=\pm 1$, and hence is **unbounded** and **disconnected**. (Note also that for these points it's *not* true that $\varphi_t(x,y) \to L_\omega(x,y)$ as $t \to \infty$.)

[Hint: To show that the solution curves spiral outwards, consider in what direction they deviate from the closed solution curves of the conservative system $\dot{x}=-y$, $\dot{y}=x(1-y^2)$.]

A17 Show that the parabola $y=x^2$ is an invariant set for the system

$$\dot{x} = x^2 - x - y, \qquad \dot{y} = x^2 - 3y.$$

(Don't forget to show that the solutions starting on the parabola *exist* for all $t \in \mathbf{R}$, since this is part of the definition of "invariant set".) Sketch the phase portrait.

A18 Show that the following systems have no closed orbits:

(a) $\dot{x}=y+x^3$, $\dot{y}=x+y+y^3$. [Hint: Bendixson.]

(b) $\dot{x}=y$, $\dot{y}=-x-y+x^2+y^2$. [Hint: Bendixson–Dulac with $f(x,y)=e^{-2x}$.]

Lesson 3

Lecture 8. Some applications

(Arrowsmith & Place, sections 5.1, 5.2, 5.3, 5.4.)

- In class we will look at a selection of the applications from Chapter 5, but there will not be time to cover everything, so you'll have to read the rest for yourself.

- The analysis in Section 5.3.3 of the **Holling–Tanner predator–prey model**

$$\frac{dx}{dt} = rx\left(1-\frac{x}{K}\right) - \frac{wxy}{D+x}, \qquad \frac{dy}{dt} = sy\left(1-\frac{y}{x/J}\right)$$

can be made simpler by writing the system in **dimensionless variables** (τ,u,v) instead of (t,x,y). This **nondimensionalization** is a very useful technique for reducing the number of parameters in a system, and since it's not described in the textbook, I'll explain it here instead. Let

$$t = c_0\tau, \qquad x = c_1 u, \qquad y = c_2 v,$$

where c_0, c_1 and c_2 are constants that we will specify soon. Inserting this into the differential equations, we get

$$\frac{c_1}{c_0}\frac{du}{d\tau} = rc_1u\left(1-\frac{c_1u}{K}\right) - \frac{wc_1uc_2v}{D+c_1u}, \qquad \frac{c_2}{c_0}\frac{dv}{d\tau} = sc_2v\left(1-\frac{c_2v}{c_1u/J}\right),$$

which we can simplify to

$$\frac{du}{d\tau} = (rc_0)u\left(1 - \frac{c_1}{K}u\right) - \frac{\frac{wc_0c_2}{c_1}uv}{\frac{D}{c_1} + u}, \qquad \frac{dv}{d\tau} = (sc_0)v\left(1 - \frac{c_2J}{c_1}\frac{v}{u}\right).$$

At this stage, we have some freedom of choice, but for example we can get rid of the coefficients indicated by the arrows, if we choose c_0, c_1 and c_2 such that

$$rc_0 = 1, \qquad \frac{c_1}{K} = 1, \qquad \frac{c_2J}{c_1} = 1.$$

In other words, we take

$$c_0 = \frac{1}{r}, \qquad c_1 = K, \qquad c_2 = \frac{K}{J}. \tag{1}$$

Then the equations become

$$\frac{du}{d\tau} = u(1-u) - \frac{\frac{w}{rJ}uv}{\frac{D}{K} + u}, \qquad \frac{dv}{d\tau} = \frac{s}{r}v\left(1 - \frac{v}{u}\right),$$

and if we now give names to the remaining coefficients appearing in the formulas, for example

$$\alpha = \frac{w}{rJ}, \qquad \beta = \frac{s}{r}, \qquad \delta = \frac{D}{K}, \tag{2}$$

then the system in its final form is

$$\frac{du}{d\tau} = u(1-u) - \frac{\alpha uv}{\delta + u}, \qquad \frac{dv}{d\tau} = \beta v\left(1 - \frac{v}{u}\right). \tag{3}$$

Note that this system contains only three parameters (α, β, δ), instead of the original six parameters (r, K, w, D, s, J). By using the possibility of rescaling the three variables, we have reduced the number of parameters by three.

Our choice (1) of the constants c_k means that the new variables the we have introduced are actually

$$\tau = \frac{t}{c_0} = rt, \qquad u = \frac{x}{c_1} = \frac{x}{K}, \qquad v = \frac{y}{c_2} = \frac{Jy}{K}.$$

Considering that the parameter r in the original ODEs must have the dimension $[\text{time}]^{-1}$ (it's a per capita growth rate), and that t of course has dimension [time], we see that the rescaled time variable $\tau = rt$ is actually dimensionless. Similarly, the carrying capacity K for the prey has the same dimension as the prey population size x (whatever unit we happen to use for this, like number of millions of individuals, or biomass in kilograms, or something else), so the variable $u = x/K$ is dimensionless. And so is v, as you can check.

Moreover, the new parameters given by (2) are also dimensionless. For example, r and s are both per capita growth rates and have the same units, so $\beta = s/r$ is a dimensionless quantity which measures the *ratio* between the intrinsic growth rates of the two species. It is rather meaningless to say something like "r is small", since this depends on what time unit we are using – if we switched from measuring time in nanoseconds to measuring it in centuries, we would get a very different numerical value for r. But the statement "β is small" (say much less than 1) expresses a fact which is meaningful regardless of scale, namely that the predators reproduce much slower than the prey.

With the system in the simpler form (3), we can now carry out the same analysis as in Section 5.3.3, but it will be cleaner, since there is less to write, and we also don't get all the original parameters scattered among our formulas, but we always keep them gathered in the meaningful combinations α, β and δ. In the textbook, they do a bit of rescaling at the end, namely taking x/x^* and y/y^* as new variables, where (x^*, y^*) is the nontrivial equilibrium point, but they don't use the possibility of rescaling time to get rid of one more parameter.

[Jump to course outline]

- The very last sentence in Section 5.3 (on p. 188) is (with my emphasis) "Thus the phase portrait corresponding to Fig. 5.22(a) **has no limit cycle**; (y_1^*, y_2^*) is simply a stable focus." However, the claim that there cannot exist any limit cycles in this case is **false**; there are actually parameter values such that the stable focus is surrounded by two limit cycles, the inner one unstable and the outer one stable.*

Exercises

- Damped harmonic oscillator: 5.2, 5.3, 5.10.
- Population models: 5.15, 5.16, 5.17.

  The system in 5.15 is the same as in 1.19d that you have done before. Since we view it as population model here, we can restrict ourselves to investigating the region where $x_1 \geq 0$ and $x_2 \geq 0$.
- Epidemics: 5.21, 5.23, A19.

  In problem 5.21, contemplate very carefully what the constant of motion says about what the phase portrait looks like, and especially where the orbits go as $t \to \infty$ (this should be clearly indicated in your phase portrait). Also read carefully how the constant c_0 is actually defined. Just because you happen to give some constant the name c_0, that doesn't mean that it's the same as what's called c_0 in the problem – that's something that you will have to *show*.
- Chemostat: A20. (Please be advised that this is a large exercise with many parts, so it will require quite a lot more time than the other homework problems!)

Additional problems

A19 As in problem 5.23, let S, I and R denote the fractions of the total population (so that $S+I+R=1$) that are **susceptible** to infection, **infected**, and **recovered** (immune). If we assume that immunity is only temporary, so that recovered individuals may go back to the susceptible class after a while, we get an **SIRS model**:

$$dS/dt = -\alpha IS + \gamma R, \qquad dI/dt = \alpha IS - \beta I, \qquad dR/dt = \beta I - \gamma R,$$

where the parameters α, β and γ are all positive.

(a) Show that the equations are consistent with $S+I+R$ staying equal to 1 for all t.

(b) Use $R = 1-I-S$ to eliminate R and get a two-dimensional system for S and I only.

The rest of this problem will concern this two-dimensional system only. Note that since $S \geq 0$, $I \geq 0$ and $R = 1-S-I \geq 0$, we may restrict our attention to the triangular region

$$D = \{(S,I) \in \mathbf{R}^2 : S \geq 0,\ I \geq 0,\ S+I \leq 1\}.$$

(c) Show that the S-nullcline for the two-dimensional system can be written as $I = (1-S)/(1+\frac{\alpha}{\gamma}S)$, with I as a function of S. Draw this curve in the whole SI-plane, **carefully motivating** your drawing with the help of your calculus skills (or otherwise). Then indicate the part of the curve which lies in the triangle D.

(d) Consider the case $0 < \alpha < \beta$. Draw the nullclines (for both S and I together) in the SI-plane. Are there any equilibrium points in the triangle D? If so, analyze them using linearization. Sketch the phase portrait in the region D, and give a biological interpretation of the results.

(e) Do the same for the case $0 < \beta < \alpha$.

*A. Gasull, R. E. Kooij & J. Torregrosa, Limit cycles in the Holling–Tanner model, Publicacions Matemàtiques, Vol. 41 (1997), 149–167.

A20 A **chemostat** is a bioreactor for growing microorganisms in the laboratory. The liquid in the reactor tank is kept well stirred at all times, and contains the microbial culture as well as nutrients needed for the microorganisms to grow. All nutrients are supplied in excess, except for one, called the *limiting nutrient* (since it's the available amount of this nutrient that will limit how fast the microorganism population can grow). Let $C = C(t)$ denote the concentration of the limiting nutrient in the tank, as a function of time t, and let $X = X(t)$ be the concentration of microorganisms; both are measured in units of mass per volume. The volume V of liquid in the tank is kept constant by continuously harvesting microorganism–nutrient solution at a constant rate F (units: volume per time), and resupplying fresh nutrient solution at the same rate F; the ratio $D = F/V$ is called the *dilution rate* (unit: 1/time). By adjusting the concentration C_0 of the limiting nutrient in the solution being pumped into the reactor, the growth rate of the microorganisms can be controlled. See for example the Wikipedia article for schematic diagrams and more information.

In this problem, we will investigate the following mathematical model of a chemostat:

$$\underbrace{\frac{dX}{dt}}_{\substack{\text{organism}\\ \text{rate of}\\ \text{change}}} = \underbrace{f(C)X}_{\substack{\text{organism}\\ \text{growth}\\ \text{rate}}} - \underbrace{DX}_{\substack{\text{organism}\\ \text{outflux}\\ \text{rate}}}, \qquad \underbrace{\frac{dC}{dt}}_{\substack{\text{nutrient}\\ \text{rate of}\\ \text{change}}} = \underbrace{DC_0}_{\substack{\text{nutrient}\\ \text{influx}\\ \text{rate}}} - \underbrace{DC}_{\substack{\text{nutrient}\\ \text{outflux}\\ \text{rate}}} - \underbrace{f(C)X/\gamma}_{\substack{\text{nutrient}\\ \text{consumption}\\ \text{rate}}},$$

where the per-capita growth rate of the microorganisms is given by

$$f(C) = \frac{K_{\max} C}{K_m + C}.$$

The dimensionless constant γ is called the *yield*, since it determines how many mass units of microorganisms that are obtained per mass unit of nutrient consumed. The formula for the function $f(C)$ is an empirical expression suggested by the famous French biochemist Jacques Monod. The parameters in this expression are $K_{\max} = \lim_{C\to\infty} f(C)$, which is the greatest possible growth rate of the microorganisms, obtained when there is an infinite supply of the limiting nutrient, and K_m, which is the value of C for which the growth rate is half the maximal rate: $f(K_m) = \frac{1}{2}K_{\max}$.

(a) Nondimensionalize the model by a change of variables of the form

$$t = a_0\tau, \qquad X = a_1 x, \qquad C = a_2 c,$$

to obtain

$$\frac{dx}{d\tau} = \left(\frac{\beta_1 c}{\beta_2 + c} - 1\right)x, \qquad \frac{dc}{d\tau} = 1 - c - \frac{\beta_1 cx}{\beta_2 + c}.$$

Write down the expressions for the scaling factors a_0, a_1, a_2 and the new parameters β_1, β_2 in terms of the old parameters. Check that β_1 and β_2 are really dimensionless, and that a_0, a_1 and a_2 have the dimensions that they should have.

(Also note that β_1 and β_2 are obviously positive.)

(b) Now let's investigate the nondimensionalized model!

To begin with, show that $(x, c) = (0, 1)$ is always an equilibrium point. Explain why it's called a "washout" equilibrium, and why it's undesirable to end up there!

(c) Assuming that $\beta_1 \neq 1$, there is another equilibrium point $(x, c) = (x^*, c^*)$. Show that it is given by the formulas

$$(x^*, c^*) = \left(1 - \frac{\beta_2}{\beta_1 - 1}, \frac{\beta_2}{\beta_1 - 1}\right).$$

(d) Show that (x^*, c^*) lies in the positive quadrant (i.e., $x^* > 0$ and $c^* > 0$) if and only if β_1 lies in the interval $(1 + \beta_2, \infty)$. We will call the "good" case.

(Below we will also look at two "bad" cases, where β_1 lies in the interval $(0, 1)$ or $(1, 1 + \beta_2)$. We will ignore the borderline cases $\beta_1 = 1$ and $\beta_1 = 1 + \beta_2$, since "in reality they only happen with probability zero".)

(e) Use linearization to investigate the stability of (x^*, c^*) in the good case $\beta_1 \in (1+\beta_2, \infty)$.

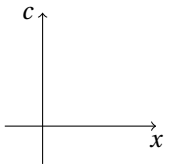
(f) Investigate the stability of the washout equilibrium $(x, c) = (0, 1)$ using linearization, both in the good case $\beta_1 \in (1+\beta_2, \infty)$ and in the two bad cases $\beta_1 \in (1, 1+\beta_2)$ and $\beta_1 \in (0, 1)$.
(Hint: The Jacobian is lower triangular, so it's very easy to find its eigenvalues and use them to determine stability.)

(g) Show that the point $(x(\tau), c(\tau))$ moves in such a way that its perpendicular distance to the line $x + c = 1$ decreases towards zero (or is identically zero).
(Hint: Write a formula for the distance $y(\tau)$ in terms of $x(\tau)$ and $c(\tau)$, and compute $dy/d\tau$.)
Explain why this implies that no equilibrium can be a focus!

(h) The c-nullcline can clearly be rewritten as

$$x = \frac{(1-c)(\beta_2 + c)}{\beta_1 c},$$

with x as a function of c. Using your calculus skills, draw this curve in the xc-plane. (In the end we will only be interested in nonnegative x and c, but draw the curve in the whole xc-plane just for practice.) Please **motivate clearly** why you draw the curve the way you do!
Also draw the line $x + c = 1$ in the same figure.
Here, and in the other cases below where you are asked to draw something in the xc-plane, **please draw the x-axis horizontally and the c-axis vertically**, in order to facilitate the marking of the homework. Like this:



(i) Sketch the phase portrait in the first quadrant of the xc-plane, in the good case $\beta_1 \in (1+\beta_2, \infty)$.
(As usual, consider the nullclines and the signs of $dx/d\tau$ and $dc/d\tau$. But try to also use all the other information that you have gathered above, in particular regarding the line $x + c = 1$.)

(j) The same, but for the bad case $\beta_1 \in (0, 1)$.

(k) And the same again, but for the other bad case $\beta_1 \in (1, 1+\beta_2)$.

(l) By now it should hopefully be clear what is "good" and "bad" about these cases. Express the condition $\beta_1 > 1$ in terms of the original parameters and explain why this condition should obviously be necessary for avoiding washout.

(m) But $\beta_1 > 1$ is not sufficient to avoid washout; we need the stronger condition $\beta_1 > 1 + \beta_2$. In order to understand this condition, express it in terms of the original parameters, and show that it can be rewritten as $C_0 > C^*$, where $C^* = K_m D/(K_{\max} - D)$.

(n) Prove that C^* is the value of C that corresponds to $c = c^*$, and use this to explain why C_0 ought to be greater than this if we want to avoid washout.

(o) Also show that $f(C^*) = D$, and use this to explain the (perhaps surprising) fact that the equilibrium nutrient concentration C^* (and consequently also the equilibrium microorganism concentration X^*) is independent of the nutrient supply concentration C_0 (as long as $C_0 > C^*$, of course).

Lecture 9. More about existence and uniqueness

(Not covered in Arrowsmith & Place; see notes below instead.)

Our goal this time is to use **Picard iteration**, also known as the **method of successive approximations**, to prove the **Picard–Lindelöf theorem**, the fundamental existence and uniqueness theorem for a system of (non-autonomous) first order ODEs $\dot{\mathbf{x}} = \mathbf{X}(t, \mathbf{x})$ with a given initial condition $\mathbf{x}(t_0) = \mathbf{c}$.

[Jump to course outline]

Preliminaries from analysis: uniform convergence

We will need some theorems about convergence of sequences and series of *functions* (not just *numbers*).

Definition (Pointwise and uniform convergence)**.** Suppose f and $f_0, f_1, f_2, \ldots$ are real-valued functions all defined on the same set I (for example an interval).

- The sequence $(f_n)_{n=0}^{\infty}$ converges to the function f **pointwise** on I if

  $$\lim_{n\to\infty} f_n(x) = f(x)$$

  for each $x \in I$.

  [Equivalently: for each $x \in I$ and for each $\varepsilon > 0$ there is an N (which may depend on x and ε) such that $\left|f_n(x) - f(x)\right| < \varepsilon$ for all $n \geq N$.]

- The sequence $(f_n)_{n=0}^{\infty}$ converges to the function f **uniformly** on I if

  $$\lim_{n\to\infty} \sup_{x\in I} \left|f_n(x) - f(x)\right| = 0.$$

  [Equivalently: for each $\varepsilon > 0$ there is an N (which may depend on ε) such that $\left|f_n(x) - f(x)\right| < \varepsilon$ for all $x \in I$ and all $n \geq N$.]

- The function *series* $\sum_{n=0}^{\infty} f_n(x)$ converges pointwise/uniformly to the function $s(x)$ if the *sequence of partial sums* $s_n(x) = \sum_{k=0}^{n} f_k(x)$ converges pointwise/uniformly to $s(x)$.

Remark. The same definitions also apply to complex-valued or vector-valued functions, etc., with the suitable interpretation of what $\left|f_n(x) - f(x)\right|$ means.

Theorem. Uniform convergence implies pointwise convergence, but not the other way around.

Proof. If $f_n \to f$ uniformly, then for a given $\varepsilon > 0$ one can find an N which works for all x, so the same number N will work for each particular x in the definition of pointwise convergence.

An example showing that the converse fails is the sequence $f_n(x) = x^n$ on the interval $[0,1]$, which converges pointwise, but *not uniformly*, to the discontinuous function

$$f(x) = \begin{cases} 0, & 0 \leq x < 1, \\ 1, & x = 1. \end{cases}$$

□

Theorem (The uniform limit theorem)**.** If each f_n is continuous on I, and $f_n \to f$ *uniformly*, then f is continuous on I.

Proof. Suppose $a \in I$. Let $\varepsilon > 0$. Since $f_n \to f$ uniformly, there is an N such that $\left|f_N(x) - f(x)\right| < \varepsilon/3$ for all $x \in I$. Since f_N is continuous, there is a $\delta > 0$ such that $\left|f_N(x) - f_N(a)\right| < \varepsilon/3$ for all $x \in I$ such that $|x - a| < \delta$. The triangle inequality gives

$$\begin{aligned} \left|f(x) - f(a)\right| &= \left|f(x) - f_N(x) + f_N(x) - f_N(a) + f_N(a) - f(a)\right| \\ &\leq \left|f(x) - f_N(x)\right| + \left|f_N(x) - f_N(a)\right| + \left|f_N(a) - f(a)\right| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

for all $x \in I$ such that $|x - a| < \delta$. Thus f is continuous at a. □

Theorem (The Weierstrass M-test)**.** If the numerical series $\sum_{n=0}^{\infty} M_n$ converges, and if $\left|f_n(x)\right| \leq M_n$ for all $x \in I$, then the function series $\sum_{n=0}^{\infty} f_n(x)$ converges uniformly (and absolutely) on I to some function $S(x)$.

Proof. (Omitted.) □

Remark. In the M-test, if each f_n is *continuous*, then the sum S is also a continuous function. This follows by applying the uniform limit theorem to the sequence of partial sums.

[Jump to course outline]

Equivalent integral equation

Lemma. Let $I \subseteq \mathbf{R}$ be an open interval (bounded or unbounded), and assume that $\mathbf{X}\colon I \times \mathbf{R}^n \to \mathbf{R}^n$ is continuous. Let $t_0 \in I$. Then the function $\mathbf{x}(t)$ $(t \in I)$ is a continuously differentiable solution of the initial value problem

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{X}\big(t, \mathbf{x}(t)\big) \quad \text{for } t \in I, \\ \mathbf{x}(t_0) &= \mathbf{c}, \end{aligned} \tag{A}$$

if and only if it is a continuous solution of the integral equation

$$\mathbf{x}(t) = \mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s, \mathbf{x}(s)\big)\, ds \quad \text{for } t \in I. \tag{B}$$

The same statement holds also for closed intervals I, provided that the derivative $\dot{\mathbf{x}}(t)$ is interpreted as a *one-sided* derivative when t is an *endpoint* of I.

Proof. This is an immediate consequence of the fundamental theorem of calculus. □

Picard iteration

Picard's idea for proving the existence of a solution to problem (A) is to recursively define an infinite sequence of functions

$$\mathbf{x}_0(t), \quad \mathbf{x}_1(t), \quad \mathbf{x}_2(t), \quad \ldots \qquad (t \in I)$$

by the formulas

$$\begin{aligned} \mathbf{x}_0(t) &= \mathbf{c}, \\ \mathbf{x}_n(t) &= \mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s, \mathbf{x}_{n-1}(s)\big)\, ds \qquad \text{for } n \geq 1, \end{aligned}$$

and to show that this sequence converges (under some conditions) to a continuous function $\mathbf{x}(t)$ which satisfies the integral equation (B), and hence also the initial value problem (A). The uniqueness of this solution is proved by separate argument (but under the same conditions).

We may note right away that each function $\mathbf{x}_n(t)$ in the sequence is differentiable on I. This is obvious for $n = 0$ since $\mathbf{x}_0$ is just a constant function, and for $n \geq 1$ it follows from the fundamental theorem of calculus:

$$\frac{d\mathbf{x}_n}{dt}(t) = \mathbf{X}\big(t, \mathbf{x}_{n-1}(t)\big).$$

And since the functions $\mathbf{x}_n(t)$ are differentiable, they are automatically continuous as well.

The Lipschitz condition

How to prove that the sequence defined by Picard iteration converges? Answer: We write

$$\begin{aligned} \mathbf{x}_n &= \big(\mathbf{x}_n - \mathbf{x}_{n-1}\big) + \cdots + \big(\mathbf{x}_2 - \mathbf{x}_1\big) + \big(\mathbf{x}_1 - \mathbf{x}_0\big) + \mathbf{x}_0 \\ &= \mathbf{c} + \sum_{k=1}^{n} \big(\mathbf{x}_k - \mathbf{x}_{k-1}\big), \end{aligned}$$

and apply the Weierstrass M-test to show that the function series

$$\mathbf{c} + \sum_{k=1}^{\infty} \big(\mathbf{x}_k - \mathbf{x}_{k-1}\big)$$

converges (uniformly, on some interval).

For this, we will need to estimate the differences $\mathbf{x}_k - \mathbf{x}_{k-1}$. For $k \geq 2$ we have

$$\begin{aligned}\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t) &= \left(\mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s, \mathbf{x}_{k-1}(s)\big)\, ds\right) \\ &\quad - \left(\mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s, \mathbf{x}_{k-2}(s)\big)\, ds\right) \\ &= \int_{t_0}^{t} \Big(\mathbf{X}\big(s, \mathbf{x}_{k-1}(s)\big) - \mathbf{X}\big(s, \mathbf{x}_{k-2}(s)\big)\Big)\, ds.\end{aligned}$$

To get anything interesting out of this expression, it's necessary to assume something about the function $\mathbf{X}$. The natural assumption in this context is that $\mathbf{X}$ satisfies the following so-called **Lipschitz condition** with respect to $\mathbf{x}$: there is some set $\Omega \subseteq \mathbf{R}^n$ and some constant $L > 0$ such that*

$$|\mathbf{X}(t, \mathbf{a}) - \mathbf{X}(t, \mathbf{b})| \leq L\,|\mathbf{a} - \mathbf{b}| \qquad \text{for all } t \in I \text{ and for all } \mathbf{a}, \mathbf{b} \in \Omega. \tag{Lip}$$

This assumption allows us to make an estimate where we get rid of the terms containing $\mathbf{X}$, as follows: if

$$\mathbf{x}_{k-1}(t) \in \Omega \quad \text{and} \quad \mathbf{x}_{k-2}(t) \in \Omega \quad \text{for all } t \in I,$$

then for $t_0 \leq t \in I$ we have

$$\begin{aligned}|\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)| &= \left|\int_{t_0}^{t} \Big(\mathbf{X}\big(s, \mathbf{x}_{k-1}(s)\big) - \mathbf{X}\big(s, \mathbf{x}_{k-2}(s)\big)\Big)\, ds\right| && \text{(put } |\cdots| \text{ around the equality above)} \\ &\leq \int_{t_0}^{t} \Big|\mathbf{X}\big(s, \mathbf{x}_{k-1}(s)\big) - \mathbf{X}\big(s, \mathbf{x}_{k-2}(s)\big)\Big|\, ds && \text{(triangle inequality for integrals)} \\ &\leq L \int_{t_0}^{t} |\mathbf{x}_{k-1}(s) - \mathbf{x}_{k-2}(s)|\, ds && \text{(because of the Lipschitz condition),}\end{aligned}$$

and similarly for $t_0 \geq t \in I$ with the bounds of integration in the opposite order:

$$|\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)| \leq L \int_{t}^{t_0} |\mathbf{x}_{k-1}(s) - \mathbf{x}_{k-2}(s)|\, ds.$$

In the proofs below, these inequalities will allow us to use knowledge about one difference $\mathbf{x}_{k-1} - \mathbf{x}_{k-2}$ to say something about the next difference $\mathbf{x}_k - \mathbf{x}_{k-1}$.

The Picard–Lindelöf theorem

Theorem (Picard–Lindelöf theorem, **global** version)**.** Let $I \subseteq \mathbf{R}$ be an open interval, and assume that $\mathbf{X}\colon I \times \mathbf{R}^n \to \mathbf{R}^n$ is continuous and satisfies the Lipschitz condition (Lip) on the whole space $\mathbf{R}^n$:

$$|\mathbf{X}(t, \mathbf{a}) - \mathbf{X}(t, \mathbf{b})| \leq L\,|\mathbf{a} - \mathbf{b}| \qquad \text{for all } t \in I \text{ and for all } \mathbf{a}, \mathbf{b} \in \mathbf{R}^n.$$

Then for any $t_0 \in I$ and any $\mathbf{c} \in \mathbf{R}^n$, the initial value problem (A),

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{X}\big(t, \mathbf{x}(t)\big) \quad \text{for } t \in I, \\ \mathbf{x}(t_0) &= \mathbf{c},\end{aligned}$$

has exactly one solution $\mathbf{x}(t)$. (Note that the interval I may be bounded or unbounded, and that the solution is defined on the whole interval $t \in I$.)

*For example, in the one-dimensional case, if the partial derivative $\frac{\partial X}{\partial x}(t, x)$ exists and satisfies the boundedness condition

$$\left|\frac{\partial X}{\partial x}(t, x)\right| \leq L \qquad \text{for all } t \in I \text{ and for all } a, b \in \Omega,$$

then the mean value theorem for derivatives implies that

$$|X(t, a) - X(t, b)| = \left|(b - a)\,\frac{\partial X}{\partial x}(t, \xi)\right| \leq L\,|b - a| \qquad \text{for all } t \in I \text{ and for all } a, b \in \Omega,$$

so that the Lipschitz condition holds. Similarly in higher dimensions. For simplicity, one often uses the stronger assumption that $\mathbf{X}(t, \mathbf{x})$ is of class C^1; this gives the Lipschitz condition automatically.

[Jump to course outline]

Theorem (Picard–Lindelöf theorem, **local** version)**.** Let $I \subseteq \mathbf{R}$ be an open interval, and assume that $\mathbf{X}\colon I \times \Omega \to \mathbf{R}^n$ is continuous and satisfies the Lipschitz condition (Lip) on some open set $\Omega \subseteq \mathbf{R}^n$:

$$|\mathbf{X}(t,\mathbf{a}) - \mathbf{X}(t,\mathbf{b})| \le L|\mathbf{a}-\mathbf{b}| \qquad \text{for all } t \in I \text{ and for all } \mathbf{a},\mathbf{b} \in \Omega.$$

Given any $t_0 \in I$ and any $\mathbf{c} \in \Omega$, take $h > 0$ and $r > 0$ small enough that the interval $J = [t_0 - h, t_0 + h]$ is contained in I and the closed ball $B = \overline{B(\mathbf{c}, r)}$ is contained in Ω, and let

$$C = \max_{(t,\mathbf{x}) \in J \times B} \left|\mathbf{X}\big(t,\mathbf{x}\big)\right|.$$

(This maximum exists by the extreme value theorem.) Then the initial value problem (A),

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \mathbf{X}\big(t,\mathbf{x}(t)\big) \quad \text{for } t \in [t_0 - \varepsilon, t_0 + \varepsilon] \text{ where } \varepsilon = \min(h, r/C), \\ \mathbf{x}(t_0) &= \mathbf{c}, \end{aligned}$$

has exactly one solution $\mathbf{x}(t)$. (Note that we cannot in general guarantee that the solution is defined on the whole interval I, only on a subinterval.)

*Proof of the **global** version.* Define the sequence $\big(\mathbf{x}_n(t)\big)_{n=0}^{\infty}$ for $t \in I$ by Picard iteration as above.

Take any $S \in I$ and $T \in I$ with $S < t_0 < T$. We will show that there is a unique solution defined on the interval $[S,T]$, and since S and T are arbitrary, this implies that there is a unique solution on the whole interval I.

Since the function $\mathbf{X}$ is continuous and the interval $[S,T]$ is closed and bounded, the maximum

$$M = \max_{t \in [S,T]} |\mathbf{X}(t,\mathbf{c})|$$

exists, by the extreme value theorem.

Let $t \in [t_0, T]$ to begin with. Then we have

$$|\mathbf{x}_1(t) - \mathbf{x}_0(t)| = \left|\left(\mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{c}\big)\,ds\right) - \mathbf{c}\right| \le \int_{t_0}^{t} \left|\mathbf{X}\big(s,\mathbf{c}\big)\right| ds \le M(t - t_0),$$

Now that we have an estimate for the first difference $\mathbf{x}_1 - \mathbf{x}_0$, we can start estimating the other differences $\mathbf{x}_k - \mathbf{x}_{k-1}$ successively, using the inequality that we derived in the section about the Lipschitz condition above. This gives (still for $t \in [t_0, T]$):

$$\begin{aligned} |\mathbf{x}_2(t) - \mathbf{x}_1(t)| &\le L\int_{t_0}^{t} |\mathbf{x}_1(s) - \mathbf{x}_0(s)|\,ds \le L\int_{t_0}^{t} M(s - t_0)\,ds = \frac{LM}{2}(t-t_0)^2, \\ |\mathbf{x}_3(t) - \mathbf{x}_2(t)| &\le L\int_{t_0}^{t} |\mathbf{x}_2(s) - \mathbf{x}_1(s)|\,ds \le L\int_{t_0}^{t} \frac{L^2 M}{2}(s-t_0)^2\,ds = \frac{L^2 M}{2\cdot 3}(t-t_0)^3, \\ |\mathbf{x}_4(t) - \mathbf{x}_3(t)| &\le L\int_{t_0}^{t} |\mathbf{x}_3(s) - \mathbf{x}_2(s)|\,ds \le L\int_{t_0}^{t} \frac{L^2 M}{2\cdot 3}(s-t_0)^3\,ds = \frac{L^3 M}{2\cdot 3\cdot 4}(t-t_0)^4, \end{aligned}$$

and so on, with an obvious pattern emerging. To get a uniform estimate, let $t = T$:

$$|\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)| \le \frac{L^{k-1}M}{k!}(T - t_0)^k \qquad \text{for all } t \in [t_0, T] \text{ and } k \ge 1.$$

If we instead consider $t \in [S, t_0]$, we find in the same way that

$$|\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)| \le \frac{L^{k-1}M}{k!}(t_0 - S)^k \qquad \text{for all } t \in [S, t_0] \text{ and } k \ge 1.$$

We can combine these two estimates into a single uniform estimate over the whole interval $[S,T]$, less sharp but still good enough for our purposes:

$$|\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)| \le \frac{L^{k-1}M}{k!}(T - S)^k \qquad \text{for all } t \in [S, T] \text{ and } k \ge 1.$$

The numerical series

$$\sum_{k=1}^{\infty} \frac{L^{k-1}M}{k!}(T-S)^k = \frac{M}{L}\sum_{k=1}^{\infty} \frac{\big(L(T-S)\big)^k}{k!} = \frac{M}{L}\big(e^{L(T-S)}-1\big)$$

converges, so the Weierstrass M-test shows the uniform convergence on $[S,T]$ of the function series that we have majorized,

$$\mathbf{c} + \sum_{k=1}^{\infty} \Big(\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)\Big).$$

The partial sums of this series are just the functions

$$\mathbf{x}_n(t) = \mathbf{c} + \sum_{k=1}^{n} \Big(\mathbf{x}_k(t) - \mathbf{x}_{k-1}(t)\Big),$$

so what we have shown is that the function sequence $(\mathbf{x}_n)_{n=0}^{\infty}$ converges uniformly on $[S,T]$ to some function $\mathbf{x}$. And since each function $\mathbf{x}_n$ is continuous, the uniform limit theorem shows that this function $\mathbf{x}$ is continuous.

Moreover, for $t \in [S,T]$,

$$\begin{aligned}
&\mathbf{x}(t) - \mathbf{c} - \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}(s)\big)\,ds \\
&= \mathbf{x}(t) - \mathbf{x}_n(t) + \mathbf{x}_n(t) - \mathbf{c} - \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}(s)\big)\,ds && \text{(add and subtract } \mathbf{x}_n\text{)} \\
&= \mathbf{x}(t) - \mathbf{x}_n(t) + \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}_{n-1}(s)\big)\,ds - \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}(s)\big)\,ds && \text{(use definition of } \mathbf{x}_n\text{)} \\
&= \mathbf{x}(t) - \mathbf{x}_n(t) + \int_{t_0}^{t} \Big(\mathbf{X}\big(s,\mathbf{x}_{n-1}(s)\big)\,ds - \mathbf{X}\big(s,\mathbf{x}(s)\big)\Big)\,ds,
\end{aligned}$$

so

$$\begin{aligned}
\left|\mathbf{x}(t) - \mathbf{c} - \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}(s)\big)\,ds\right| &\le |\mathbf{x}(t) - \mathbf{x}_n(t)| + \left|\int_{t_0}^{t} \Big|\mathbf{X}\big(s,\mathbf{x}_{n-1}(s)\big)\,ds - \mathbf{X}\big(s,\mathbf{x}(s)\big)\Big|\,ds\right| \\
&\le |\mathbf{x}(t) - \mathbf{x}_n(t)| + L\left|\int_{t_0}^{t} |\mathbf{x}_{n-1}(s) - \mathbf{x}(s)|\,ds\right|,
\end{aligned}$$

where the right-hand side can be made arbitrarily small by taking n large enough, due to the *uniform* convergence

$$\max_{t\in[S,T]} |\mathbf{x}_n(t) - \mathbf{x}(t)| \to 0 \qquad \text{as } n \to \infty.$$

Therefore, since the left-hand side is nonnegative and independent of n, it must be zero:

$$\mathbf{x}(t) - \mathbf{c} - \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}(s)\big)\,ds = 0.$$

In other words, the continuous function $\mathbf{x}(t)$ satisfies the integral equation (B) on the interval $[S,T]$, and thus it also satisfies the equivalent initial value problem (A) on $[S,T]$.

To show uniqueness, assume that the function $\mathbf{y}(t)$ is also a continuous solution of (B) on $[S,T]$. Then the maximum

$$A = \max_{t\in[S,T]} \big|\mathbf{y}(t) - \mathbf{c}\big|$$

exists, by the extreme value theorem. For $t \in [t_0,T]$ we first estimate

$$\begin{aligned}
\big|\mathbf{y}(t) - \mathbf{x}_1(t)\big| &= \left|\left(\mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{y}(s)\big)\,ds\right) - \left(\mathbf{c} + \int_{t_0}^{t} \mathbf{X}\big(s,\mathbf{x}_0(s)\big)\,ds\right)\right| && \text{(def. of } \mathbf{y} \text{ and } \mathbf{x}_1\text{)} \\
&\le \int_{t_0}^{t} \Big|\mathbf{X}\big(s,\mathbf{y}(s)\big) - \mathbf{X}\big(s,\mathbf{c}\big)\Big|\,ds && \text{(triangle inequality for integrals)} \\
&\le L\int_{t_0}^{t} \big|\mathbf{y}(s) - \mathbf{c}\big|\,ds && \text{(Lipschitz condition)} \\
&\le LA(t-t_0) && \text{(definition of } A\text{)},
\end{aligned}$$

and then successively (in a manner very similar to what we did earlier)

$$\begin{aligned}\bigl|\mathbf{y}(t)-\mathbf{x}_2(t)\bigr| &\le \frac{L^2A(t-t_0)^2}{2},\\ \bigl|\mathbf{y}(t)-\mathbf{x}_3(t)\bigr| &\le \frac{L^3A(t-t_0)^3}{2\cdot 3},\\ \bigl|\mathbf{y}(t)-\mathbf{x}_4(t)\bigr| &\le \frac{L^4A(t-t_0)^4}{2\cdot 3\cdot 4},\end{aligned}$$

and so on. Together with the similar estimates for $t \in [S, t_0]$ we get

$$\bigl|\mathbf{y}(t)-\mathbf{x}_n(t)\bigr| \le \frac{L^nA(T-S)^n}{n!} \qquad \text{for all } t\in[S,T] \text{ and } n \ge 0.$$

The right-hand side tends to zero as $n \to \infty$, so the limit of the left-hand side, $\bigl|\mathbf{y}(t)-\mathbf{x}(t)\bigr|$, must also be zero. Thus $\mathbf{y}(t) = \mathbf{x}(t)$ for all $t \in [S, T]$, and uniqueness is proved. □

*Idea of proof of the **local** version.* Do more or less the same thing, but also use the restrictions to make sure that the Picard iteration doesn't take us outside of the region where the Lipschitz condition holds. □

Exercises

- Integral equations and Picard iteration: A21.
- Non-uniqueness and non-existence: A22, A23.
- Grönwall's lemma: A24, A25.

Additional problems

A21 (a) Solve the integral equation

$$x(t) = 1 + \int_0^t x(s)\,ds$$

exactly. For comparison, also compute the sequence of Picard approximations

$$x_n(t) = 1 + \int_0^t x_{n-1}(s)\,ds$$

starting with the constant function $x_0(t) = 1$.

[Answer: Exact solution $x(t) = e^t$. Picard iterates $x_n(t) = \sum_{k=0}^{n} \frac{t^k}{k!}$.]

(b) Do the same for

$$x(t) = 3 + \int_0^t 4s\,x(s)\,ds.$$

A22 Consider the ODE $t\,\dot{x} = 2x$. (Notice that the coefficient of $\dot{x}$ equals zero at $t = 0$, so we might expect some "trouble" there; it's a *singular point* of the equation.)

(a) Verify that

$$x(t) = \begin{cases} t^2, & t \ge 0,\\ Ct^2, & t < 0\end{cases}$$

satisfies the ODE for any constant C. Thus there are infinitely many solutions satisfying the condition $x(1) = 1$.

(b) Show that there are also infinitely many solutions satisfying the condition $x(0) = 0$, but no solutions satisfying $x(0) = b$ with $b \neq 0$.

[Jump to course outline]

A23 Find *all* functions $x(t)$, $t \in \mathbf{R}$ which satisfy

$$\dot{x} = 2\sqrt{|x|}, \qquad x(0) = 0.$$

A24 There are several variants of **Grönwall's lemma** (or **Grönwall's inequality**), which all boil down to the fact that if a function satisfies a differential or integral *inequality* of a certain form, then it can be no bigger than the solution of the corresponding differential or integral *equation*.

(The inequality limits how fast the function can grow, and to push those limits to the maximum and *grow as fast as possible*, the function should satisfy the inequality *with equality*.)

Here your task is to prove (with guidance) the following version of Grönwall's lemma:

Theorem. Let $a(t)$ and $b(t)$ be continuous functions with $b(t) \geq 0$, and suppose that $u(t)$ satisfies the integral inequality

$$u(t) \leq a(t) + \int_0^t b(s)\, u(s)\, ds \qquad \text{for } t \geq 0.$$

Then

$$u(t) \leq y(t) \qquad \text{for } t \geq 0,$$

where $y(t)$ is the solution of the corresponding integral equation

$$y(t) = a(t) + \int_0^t b(s)\, y(s)\, ds,$$

namely

$$y(t) = a(t) + \int_0^t a(s)\, b(s)\, e^{B(t)-B(s)}\, ds, \qquad \text{where} \quad B(t) = \int_0^t b(\tau)\, d\tau.$$

Outline of proof. Follow these steps:

- First rewrite the integral equation for $y(t)$ as an ODE for the integral $I(t) = \int_0^t b(s)\, y(s)\, ds = y(t) - a(t)$ appearing on the right-hand side:

$$I'(t) - b(t)\, I(t) = a(t)\, b(t), \qquad I(0) = 0.$$

- Multiply both sides by the integrating factor $e^{-B(t)}$ and integrate from 0 to t.
- After having solved for $I(t)$ in this way, you also know what $y(t)$ is, namely $y(t) = a(t) + I(t)$. Verify that your expression for $y(t)$ agrees with the formula for $y(t)$ given in the theorem above.
- Next consider the inequality for $u(t)$. Let $J(t) = \int_0^t b(s)\, u(s)\, ds$ be the integral appearing on the right-hand side. It satisfies $J'(t) = b(t)\, u(t)$. By assumption we have $u \leq a + J$ and $b \geq 0$, and therefore $J' = bu \leq b(a + J)$. So we know that J satisfies

$$J'(t) - b(t)\, J(t) \leq a(t)\, b(t), \qquad J(0) = 0.$$

- Convince yourself that you can now perform exactly the same steps as when you solved the equation for $I(t)$ (multiply by the integrating factor, integrate, etc.), but with J instead of I and "$\leq$" instead of "$=$". (Here it's important that $t \geq 0$.)
- And since you do the same steps, you will get the same result in the end, except with "$u(t) \leq$" instead of "$y(t) =$".
- This proves that $u(t) \leq y(t)$ for $t \geq 0$. Done! □

As a bonus question, see if you can prove this variant of Grönwall's lemma in a similar way:

Theorem. Let g be a continuous function (which need not be positive). If $u(t)$ is continuous for $t \geq 0$ and differentiable for $t > 0$, and satisfies the differential inequality

$$\begin{aligned} u'(t) &\leq g(t)\, u(t) \qquad \text{for } t > 0, \\ u(0) &= c, \end{aligned}$$

then

$$u(t) \leq y(t) \qquad \text{for } t \geq 0,$$

where $y(t)$ is the solution of the corresponding differential equation

$$\begin{aligned} y'(t) &= g(t)\, y(t) \qquad \text{for } t > 0, \\ y(0) &= c, \end{aligned}$$

namely

$$y(t) = c\, e^{G(t)}, \qquad G(t) = \int_0^t g(s)\, ds.$$

A25 (a) Compute the solution $x(t)$ of the initial value problem $\dot{x} = x^2$, $x(0) = c > 0$, and note that it blows up after finite time (as $t \nearrow 1/c$).

(b) In part (a), the right-hand side of the ODE was obviously quadratic in x.
In contrast, show that if the right-hand side of the system $\dot{\mathbf{x}} = \mathbf{X}(\mathbf{x})$ is **linearly bounded**, meaning that there are constants $a \geq 0$ and $b \geq 0$ such that

$$|\mathbf{X}(\mathbf{x})| \leq a + b|\mathbf{x}| \quad \text{for all } \mathbf{x} \in \mathbf{R}^n,$$

then the solution $\mathbf{x}(t)$ with initial value $\mathbf{x}(0) = \mathbf{x}_0$ cannot blow up in finite time, so it is defined for all $t \geq 0$ (regardless of $\mathbf{x}_0$).
(We are assuming that the vector field $\mathbf{X}(\mathbf{x})$ is nice enough for the flow to exist at least locally, say of class C^1. It can be proved that if $\mathbf{X}(\mathbf{x})$ is defined for all $\mathbf{x} \in \mathbf{R}^n$, then the only way for solutions to cease existing after finite time is that $|\mathbf{x}| \to \infty$. Let us take this fact for granted here.)
[Hint: Assume, for a contradiction, that some solution $\mathbf{x}(t)$ only exists for $0 \leq t < t_0$. From

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_0^t \mathbf{X}(\mathbf{x}(s))\, ds,$$

deduce that

$$|\mathbf{x}(t)| \leq |\mathbf{x}_0| + at + \int_0^t b|\mathbf{x}(s)|\, ds \qquad \text{for } 0 \leq t < t_0,$$

and apply Grönwall's lemma from problem A24 to show that $|\mathbf{x}(t)|$ cannot tend to ∞ as $t \nearrow t_0$.]

Lecture 10. Linear equations with non-constant coefficients

(Not covered in Arrowsmith & Place; see notes below instead.)

Second-order linear ODEs

Many ODE books with a more "classical" flavour allocate plenty of space to the topic of second order (inhomogeneous) linear ODES,

$$\ddot{x} + p_1(t)\, \dot{x} + p_0(t)\, x = f(t),$$

which appear in many applications, and are also of historical importance. Some particular such ODEs have been studied so much that one could easily spend several courses on them alone, like the **Bessel equation** (with parameter $\alpha \in \mathbf{C}$),

$$\ddot{x} + \tfrac{1}{t}\, \dot{x} + \tfrac{t^2 - \alpha^2}{t^2}\, x = 0.$$

[Jump to course outline]

Any second-order ODE can be rewritten as a system of two first-order ODEs, for example by letting $x_1 = x$ and $x_2 = \dot{x}$:

$$\ddot{x} + p_1(t)\,\dot{x} + p_0(t)\,x = f(t) \qquad \Longleftrightarrow \qquad \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ f(t) - p_1(t)\,x_2 - p_0(t)\,x_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -p_0(t) & -p_1(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ f(t) \end{pmatrix}.$$

Here we will focus more on systems of first-order (inhomogeneous) linear ODEs in general:

$$\dot{\mathbf{x}} = A(t)\,\mathbf{x} + \mathbf{f}(t),$$

where $\mathbf{x}(t) \in \mathbf{R}^n$ and $A(t)$ is some $n \times n$ matrix. From the general results about systems we can then obtain some of the basic facts about second-order linear ODEs simply by considering the special case when $n = 2$ and $\dot{x}_1 = x_2$.

First-order systems of linear ODEs

Theorem. Assume that $A(t)$ and $\mathbf{f}(t)$ are *continuous* on some time interval I (bounded or not) containing $t = 0$. Then the initial value problem

$$\dot{\mathbf{x}} = A(t)\,\mathbf{x} + \mathbf{f}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0$$

has a unique solution defined on all of I.

Proof. Let $J = [t_1, t_2]$ be an arbitrary compact subinterval of I containing $t = 0$. By the extreme value theorem, each matrix entry $A_{ij}(t)$ is bounded on J, and since there are only finitely many matrix entries, there must be a common constant which bounds all of them:

$$\left|A_{ij}(t)\right| \le C \qquad \text{for all } i \text{ and } j, \text{ and all } t \in J.$$

Then the ith entry in the matrix product $A(t)\,\mathbf{y}$ can be estimated using the triangle inequality and the Cauchy–Schwarz inequality:

$$\left|\big(A(t)\,\mathbf{y}\big)_i\right| = \left|A_{i1}(t)\,y_1 + \cdots + A_{in}(t)\,y_n\right| \le C\left|y_1\right| + \cdots + C\left|y_n\right| \le C\sqrt{n}\,|\mathbf{y}|.$$

Squaring, summing over i, and taking the square root, we get

$$\left|A(t)\,\mathbf{y}\right| \le Cn\,|\mathbf{y}|,$$

for any vector $\mathbf{y}$.

Using this property we can show that the right-hand side $\mathbf{X}(t,\mathbf{x}) = A(t)\,\mathbf{x} + \mathbf{f}(t)$ satisfies the global Lipschitz condition with Lipschitz constant $L = Cn$. Indeed,

$$\mathbf{X}(t,\mathbf{a}) - \mathbf{X}(t,\mathbf{b}) = \Big(A(t)\,\mathbf{a} + \mathbf{f}(t)\Big) - \Big(A(t)\,\mathbf{b} + \mathbf{f}(t)\Big) = A(t)\,\mathbf{a} - A(t)\,\mathbf{b} = A(t)\,(\mathbf{a} - \mathbf{b})$$

implies that

$$|\mathbf{X}(t,\mathbf{a}) - \mathbf{X}(t,\mathbf{b})| = |A(t)(\mathbf{a} - \mathbf{b})| \le Cn\,|\mathbf{a} - \mathbf{b}| \qquad \text{for all } t \in J \text{ and for all } \mathbf{a}, \mathbf{b} \in \mathbf{R}^n.$$

By the global Picard–Lindelöf theorem there is therefore a unique solution defined on all of J, and since the subinterval J was arbitrary, we can extend this solution as far as we like inside I. □

Theorem. The general solution of the inhomogeneous system $\dot{\mathbf{x}} - A(t)\,\mathbf{x} = \mathbf{f}(t)$ has the form

$$\mathbf{x}(t) = \mathbf{x}_{\text{hom}}(t) + \mathbf{x}_{\text{part}}(t),$$

where $\mathbf{x}_{\text{part}}$ is some particular solution, and $\mathbf{x}_{\text{hom}}$ is the general solution of the corresponding homogeneous system $\dot{\mathbf{x}} - A(t)\,\mathbf{x} = \mathbf{0}$.

Proof. This is just a fact about linearity, and not really about differential equations. If $\mathbf{x} = \mathbf{x}_1$ and $\mathbf{x} = \mathbf{x}_2$ are two particular solutions, then by linearity their difference $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$ satisfies the homogeneous system. □

[Jump to course outline]

The homogeneous case

Let's look at the homogeneous system $\dot{\mathbf{x}} = A(t)\,\mathbf{x}$ first. If $A(t) = A$ is a *constant* matrix, we have seen before that the solution is simply

$$\mathbf{x}(t) = e^{tA}\mathbf{x}(0),$$

but if $A(t)$ is *time-dependent*, we can't in general find an explicit solution formula like that. But we can still say a few things in principle about the structure of the solution. To formulate the theorem, let

$$(\mathbf{e}_1, \ldots, \mathbf{e}_n)$$

denote the standard basis for $\mathbf{R}^n$, i.e., $\mathbf{e}_k = (0, \ldots, 0, 1, 0, \ldots, 0)^T$ with a 1 in the kth position.

Theorem (Solution space)**.** The homogeneous system $\dot{\mathbf{x}} = A(t)\,\mathbf{x}$ has an n-dimensional solution space. For example, a basis is given by the functions $\mathbf{g}_1(t), \ldots, \mathbf{g}_n(t)$, where $\mathbf{x}(t) = \mathbf{g}_k(t)$ is the (unique) solution of the initial value problem starting at $\mathbf{e}_k$:

$$\dot{\mathbf{x}} = A(t)\,\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{e}_k.$$

In terms of the $n \times n$-matrix

$$\Phi(t) = \Big[\mathbf{g}_1(t), \ldots, \mathbf{g}_n(t)\Big],$$

any solution has the form

$$\mathbf{x}(t) = \Phi(t)\,\mathbf{x}(0). \tag{SOL}$$

Proof. Let $\mathbf{x}(t) = \mathbf{z}(t)$ be any solution of $\dot{\mathbf{x}} = A(t)\,\mathbf{x}$, and let $\mathbf{c} = \mathbf{z}(0)$. Then the functions $\mathbf{x}(t) = \mathbf{z}(t)$ and

$$\mathbf{x}(t) = c_1\,\mathbf{g}_1(t) + \cdots + c_n\,\mathbf{g}_n(t)$$

both satisfy the initial value problem

$$\dot{\mathbf{x}} = A(t)\,\mathbf{x}, \qquad \mathbf{x}(0) = \mathbf{c},$$

so by uniqueness they must be the same function:

$$\mathbf{z}(t) = c_1\,\mathbf{g}_1(t) + \cdots + c_n\,\mathbf{g}_n(t) = \Phi(t)\,\mathbf{c}.$$

Thus an arbitrary solution $\mathbf{z}$ can be written as a linear combination of the functions $\mathbf{g}_k$, which shows that they span the solution space, and the formula (SOL) also follows.

To show that the functions $\mathbf{g}_k$ are linearly independent, suppose that the linear combination $\mathbf{x}(t) = \sum c_k\,\mathbf{g}_k(t)$ is the zero function; then in particular it's zero when $t = 0$:

$$\mathbf{0} = \mathbf{x}(0) = \sum c_k\,\mathbf{g}_k(0) = \sum c_k\,\mathbf{e}_k = (c_1, \ldots, c_n)^T.$$

In other words, $c_1 = \cdots = c_n = 0$. □

Proposition. The columns of an $n \times n$ matrix $\Phi(t)$ are solutions of the linear system $\dot{\mathbf{x}} = A(t)\,\mathbf{x}$ if and only if the matrix itself is a solution of the matrix-valued linear ODE

$$\dot{\Phi}(t) = A(t)\,\Phi(t).$$

Proof. This is an immediate consequence of how matrix multiplication works: the kth column in $A\Phi$ equals A times the kth column in Φ. □

Definition. Any time-dependent $n \times n$ matrix whose columns are *linearly independent* solutions of $\dot{\mathbf{x}} = A(t)\,\mathbf{x}$ is called a **fundamental matrix** for the system.

Remark. The matrix $\Phi(t)$ in the theorem above is thus one particular fundamental matrix, distinguished by the property that $\Phi(0) = I$. Any other fundamental matrix $\Psi(t)$ has the form $\Psi(t) = \Phi(t)\,M$ where $M = \Psi(0)$ is a nonsingular constant $n \times n$ matrix; indeed, this is just a change of basis in the solution space. In terms of such a Ψ, the general solution of the initial value problem is $\mathbf{x}(t) = \Phi(t)\,\mathbf{x}(0) = \Psi(t)\,M^{-1}\mathbf{x}(0)$.

Rephrasing what we said before the theorem: If A is *constant*, then $\Phi(t) = e^{tA}$ is the fundamental matrix satisfying $\Phi(0) = I$, but if A is *time-dependent* we cannot in general compute the fundamental matrix $\Phi(t)$ explicitly; we only know from the Picard–Lindelöf theorem that it exists and is unique. But curiously enough, we can always compute its determinant:

Theorem (Liouville's identity)**.** If $\Phi(t)$ is a solution of the matrix ODE $\dot{\Phi}(t) = A(t)\,\Phi(t)$, then its determinant*

$$W(t) = \det \Phi(t)$$

satisfies the scalar ODE

$$\dot{W}(t) = \operatorname{tr}\big(A(t)\big)\, W(t),$$

and hence

$$W(t) = \exp\left(\int_0^t \operatorname{tr}\big(A(s)\big)\, ds\right) W(0).$$

Proof. A down-to-earth way of proving this is to just compute. Consider the 3×3 case, for ease of notation, so that

$$W(x) = \det\Big(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big) = \begin{vmatrix} g_{11} & g_{21} & g_{31} \\ g_{12} & g_{22} & g_{32} \\ g_{13} & g_{23} & g_{33} \end{vmatrix} = g_{11}g_{22}g_{33} + \cdots.$$

When differentiating this using the product rule for derivatives, each of the $n!$ terms in the determinant will give rise to n terms, and the resulting sum can be rearranged back into a sum of n determinants:

$$\begin{aligned} \dot{W} &= \frac{d}{dt}\det\Big(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big) \\ &= \Big(\frac{dg_{11}}{dt}g_{22}g_{33} + g_{11}\frac{dg_{22}}{dt}g_{33} + g_{11}g_{22}\frac{dg_{33}}{dt}\Big) + \cdots \\ &= \det\Big(\frac{d\mathbf{g}_1}{dt}, \mathbf{g}_2, \mathbf{g}_3\Big) + \det\Big(\mathbf{g}_1, \frac{d\mathbf{g}_2}{dt}, \mathbf{g}_3\Big) + \det\Big(\mathbf{g}_1, \mathbf{g}_2, \frac{d\mathbf{g}_3}{dt}\Big) \\ &= \det\Big(A\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big) + \det\Big(\mathbf{g}_1, A\mathbf{g}_2, \mathbf{g}_3\Big) + \det\Big(\mathbf{g}_1, \mathbf{g}_2, A\mathbf{g}_3\Big). \end{aligned}$$

If we expand all these determinants, what terms would we get that contain A_{11}? Answer: the terms appearing in the expression

$$\begin{vmatrix} A_{11}g_{11} & 0 & 0 \\ 0 & g_{22} & g_{32} \\ 0 & g_{23} & g_{33} \end{vmatrix} + \begin{vmatrix} 0 & A_{11}g_{21} & 0 \\ g_{12} & 0 & g_{32} \\ g_{13} & 0 & g_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & A_{11}g_{31} \\ g_{12} & g_{22} & 0 \\ g_{13} & g_{23} & 0 \end{vmatrix} = A_{11}\det\Big(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big).$$

And the terms that contain A_{12} are

$$\begin{vmatrix} A_{12}g_{12} & 0 & 0 \\ 0 & g_{22} & g_{32} \\ 0 & g_{23} & g_{33} \end{vmatrix} + \begin{vmatrix} 0 & A_{12}g_{22} & 0 \\ g_{12} & 0 & g_{32} \\ g_{13} & 0 & g_{33} \end{vmatrix} + \begin{vmatrix} 0 & 0 & A_{12}g_{32} \\ g_{12} & g_{22} & 0 \\ g_{13} & g_{23} & 0 \end{vmatrix} = A_{12}\begin{vmatrix} g_{12} & g_{22} & g_{32} \\ g_{12} & g_{22} & g_{32} \\ g_{13} & g_{23} & g_{33} \end{vmatrix},$$

but this is zero since two rows are equal, so there will be no terms containing A_{12} in the expansion. Similarly, we get contributions $A_{22}\det\Big(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big)$ and $A_{33}\det\Big(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big)$, but no terms containing an A_{ij} with $i \neq j$. Thus,

$$\dot{W} = (A_{11} + A_{22} + A_{33})\det\Big(\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\Big) = \operatorname{tr}(A)\, W,$$

as desired. The formula for $W(t)$ is obtained by solving this ODE for W using an integrating factor. □

Remark. Liouville's identity shows that if the columns of the matrix $\Phi(t)$ are *solutions of some linear system* $\dot{\mathbf{x}} = A(t)\mathbf{x}$, then $W(t) = \det\Phi(t)$ is either **identically zero** or **never zero**. The case $W(t) = 0$ occurs when the columns of Φ are linearly dependent functions, and $W(t) \neq 0$ when they are linearly independent. But beware that if we just look at the determinant of some arbitrary time-dependent matrix $\Phi(t)$, then the question of linear independence of the columns is not this simple! (See problem A29.)

*Called **the Wronskian determinant**, or simply **the Wronskian**.

[Jump to course outline]

Consider the special case of second-order linear homogeneous ODEs

$$\ddot{x} + p_1(t)\,\dot{x} + p_0(t)\,x = 0,$$

which can be written as $\dot{\mathbf{x}} = A\mathbf{x}$ with

$$\mathbf{x}(t) = \begin{pmatrix} x(t) \\ \dot{x}(t) \end{pmatrix}, \qquad A(t) = \begin{pmatrix} 0 & 1 \\ -p_0(t) & -p_1(t) \end{pmatrix},$$

as in the beginning of this lecture. Since $\operatorname{tr}\big(A(t)\big) = -p_1(t)$ in this case, Liouville's identity takes the following form:

Theorem (Abel's identity)**.** If $x = y(t)$ and $x = z(t)$ are two solutions of $\ddot{x} + p_1\dot{x} + p_0 x = 0$, then their Wronskian

$$W(t) = \begin{vmatrix} y(t) & z(t) \\ \dot{y}(t) & \dot{z}(t) \end{vmatrix} = y(t)\,\dot{z}(t) - \dot{y}(t)\,z(t)$$

satisfies

$$\dot{W}(t) = -p_1(t)\,W(t),$$

and hence

$$W(t) = \exp\left(-\int_0^t p_1(s)\,ds\right) W(0).$$

The inhomogeneous case

Now to the question of finding a **particular solution** $\mathbf{x}_{\text{part}}(t)$ of the *inhomogeneous* system

$$\dot{\mathbf{x}} - A(t)\,\mathbf{x} = \mathbf{f}(t),$$

supposing that we already know the general solution $\mathbf{x}_{\text{hom}}(t)$ of the *homogeneous* equation. (In other words, supposing that we know a fundamental matrix.) This can be done by a method called **variation of constants**, **variation of parameters** or **Lagrange's method**, as follows:

Theorem (Variation of constants)**.** If $\Phi(t)$ is a fundamental matrix for the homogeneous system, then

$$\mathbf{x}_{\text{part}}(t) = \Phi(t) \int_0^t \Phi(s)^{-1}\,\mathbf{f}(s)\,ds$$

is a particular solution of the inhomogeneous system.

Proof. The fundamental matrix satisfies $\dot{\Phi} = A\Phi$ (by definition). Make the change of variables

$$\mathbf{x}(t) = \Phi(t)\,\mathbf{y}(t).$$

When we substitute this into the system, together with $\dot{\mathbf{x}} = \dot{\Phi}\,\mathbf{y} + \Phi\,\dot{\mathbf{y}}$, we obtain

$$\begin{aligned}
& \dot{\mathbf{x}}(t) - A(t)\,\mathbf{x}(t) = \mathbf{f}(t) \\
\iff \quad & \dot{\Phi}(t)\,\mathbf{y}(t) + \Phi(t)\,\dot{\mathbf{y}}(t) - A(t)\,\Phi(t)\,\mathbf{y}(t) = \mathbf{f}(t) \\
\iff \quad & \underbrace{\big(\dot{\Phi}(t) - A(t)\,\Phi(t)\big)}_{=0}\,\mathbf{y}(t) + \Phi(t)\,\dot{\mathbf{y}}(t) = \mathbf{f}(t) \\
\iff \quad & \Phi(t)\,\dot{\mathbf{y}}(t) = \mathbf{f}(t) \\
\iff \quad & \dot{\mathbf{y}}(t) = \Phi(t)^{-1}\,\mathbf{f}(t).
\end{aligned}$$

(We know that $\Phi(t)^{-1}$ exists for every t, since the Wronskian $W(t) = \det\Phi(t)$ is never zero, by Liouville's identity.) Now just change t to s, integrate both sides from (say) 0 to t to find a particular $\mathbf{y}(t)$ which works, and go back to the old variables $\mathbf{x}$. Done! □

Remark. There is no need to memorize the solution formula; just repeat the procedure in the proof every time you need it! The reason for the funny name "variation of constants" is that the general solution of the homogeneous system is

$$\mathbf{x}_{\text{hom}}(t) = \Phi(t)\,\mathbf{c},$$

where $\mathbf{c} = (c_1, \dots, c_n)^T$ is an arbitrary *constant* vector, and the idea here is to try to find a particular solution by "letting the constants c_k vary", i.e., by replacing them with the *time-dependent* quantities $\mathbf{y}(t) = \big(y_1(t), \dots, y_n(t)\big)^T$:

$$\mathbf{x}_{\text{part}}(t) = \Phi(t)\,\mathbf{y}(t).$$

Back to single second-order (or higher-order) ODEs

Now consider again a single linear ODE of order n, with time-dependent coefficients:

$$\frac{d^n x}{dt^n} + p_{n-1}(t)\,\frac{d^{n-1}x}{dt^{n-1}} + \cdots + p_1(t)\,\frac{dx}{dt} + p_0(t)\,x = f(t).$$

The general solution of this inhomogeneous equation equals one particular solution plus the general solution of the homogeneous equation

$$\frac{d^n x}{dt^n} + p_{n-1}(t)\,\frac{d^{n-1}x}{dt^{n-1}} + \cdots + p_1(t)\,\frac{dx}{dt} + p_0(t)\,x = 0.$$

The first question is how we may find a basis for the n-dimensional solution space of the homogeneous equation. Sometimes we might get lucky and find one solution (by inspired guessing or power series methods or something else). This can then be used for finding *other* solutions, by the following simple trick:

Theorem (Reduction of order)**.** Suppose $x_0(t)$ is a known solution of the homogeneous equation. Then the substitution $x(t) = Y(t)\,x_0(t)$ leads to a homogeneous equation of order $n-1$ for $y(t) = \dot{Y}(t)$.

Proof. Substitute

$$\begin{aligned} x &= Y x_0, \\ \dot{x} &= \dot{Y} x_0 + Y \dot{x}_0, \\ \ddot{x} &= \ddot{Y} x_0 + 2\dot{Y}\dot{x}_0 + Y \ddot{x}_0, \\ &\vdots \end{aligned}$$

into the homogeneous ODE for x. Then the coefficient of Y will be $x_0^{(n)} + p_{n-1}\,x_0^{(n-1)} + \cdots + p_1\,\dot{x}_0 + p_0\,x_0$, which equals zero by assumption. Thus only $\dot{Y}, \ddot{Y}, \dots, Y^{(n)}$ appear in the equation, so if we let $y = \dot{Y}$ we get an equation involving only $y, \dot{y}, \dots, y^{(n-1)}$. □

Remark. The reduced equation always has the trivial solution $y(t) = 0$, which gives $Y(t) = C$. But this is quite uninteresting, since $x(t) = Y(t)\,x_0(t) = C x_0(t)$ is then just a constant multiple of the already known solution $x_0(t)$.

Remark. If we manage to find a solution of a *second-order* homogeneous linear ODE, then reduction of order gives a *first-order* equation, which means that we can find a nontrivial solution with the help of an integrating factor. We can then integrate this solution $y(t)$ to find a non-constant $Y(t)$, and hence find a second *linearly independent* solution $x(t) = Y(t)\,x_0(t)$. So in this case $x_0(t)$ and $Y(t)\,x_0(t)$ will be a basis of the solution space.

For finding a particular solution of the inhomogeneous equation, we have the method of **variation of constants**, as a special case of what we did for systems. Rewrite the ODE as a first-order system by letting $x_1 = x$, $x_2 = \dot{x}$, etc.:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ \vdots \\ x_n \\ f(t) - \sum_{k=1}^{n} p_{k-1}(t)\,x_k \end{pmatrix} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & & \\ & & & 0 & 1 \\ -p_0(t) & -p_1(t) & \cdots & -p_{n-2}(t) & -p_{n-1}(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ f(t) \end{pmatrix}$$

[Jump to course outline]

Let's look at the case $n = 3$ to simplify notation:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ f(t) - p_0(t)\,x_1 - p_1(t)\,x_2 - p_2(t)\,x_3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -p_0(t) & -p_1(t) & -p_2(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ f(t) \end{pmatrix}$$

The method requires that we already know the general solution x_{hom} of the homogeneous equation, say

$$x_{\text{hom}}(t) = c_1 g_1(t) + c_2 g_2(t) + c_3 g_3(t)$$

where (g_1, g_2, g_3) is a known basis for the solution space. We are going to seek a particular solution by "letting the constants c_k vary", i.e., replacing c_k by $y_k(t)$:

$$x_{\text{part}}(t) = y_1(t)\,g_1(t) + y_2(t)\,g_2(t) + y_3(t)\,g_3(t).$$

In terms of the first-order system, this means that

$$\Phi(t) = \begin{pmatrix} g_1 & g_2 & g_3 \\ \dot{g}_1 & \dot{g}_2 & \dot{g}_3 \\ \ddot{g}_1 & \ddot{g}_2 & \ddot{g}_3 \end{pmatrix}$$

is a fundamental matrix, and we are making the substitution $\mathbf{x}(t) = \Phi(t)\,\mathbf{y}(t)$. Simply remembering what we did for systems above, we know that this leads to $\Phi(t)\,\dot{\mathbf{y}} = \mathbf{f}(t)$:

$$\begin{pmatrix} g_1 & g_2 & g_3 \\ \dot{g}_1 & \dot{g}_2 & \dot{g}_3 \\ \ddot{g}_1 & \ddot{g}_2 & \ddot{g}_3 \end{pmatrix} \begin{pmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ f(t) \end{pmatrix}.$$

This system of equations determines $\dot{\mathbf{y}}$, which we can then integrate to find $\mathbf{y}$, and hence x_{part}.

Remark. Books which deal only with single higher-order ODEs (rather than systems of first-order ODEs) usually present the method of variation of constants in the following way, which in my opinion is rather obscure. Start with

$$x = y_1 g_1 + y_2 g_2 + y_3 g_3.$$

Take the first derivative:

$$\dot{x} = (\dot{y}_1 g_1 + \dot{y}_2 g_2 + \dot{y}_3 g_3) + (y_1 \dot{g}_1 + y_2 \dot{g}_2 + y_3 \dot{g}_3).$$

For some mysterious reason ("just because it works"), we require the first bracket to be zero: $\dot{y}_1 g_1 + \dot{y}_2 g_2 + \dot{y}_3 g_3 = 0$. So only the second bracket remains in $\dot{x} = y_1 \dot{g}_1 + y_2 \dot{g}_2 + y_3 \dot{g}_3$, and taking the derivative of this gives

$$\ddot{x} = (\dot{y}_1 \dot{g}_1 + \dot{y}_2 \dot{g}_2 + \dot{y}_3 \dot{g}_3) + (y_1 \ddot{g}_1 + y_2 \ddot{g}_2 + y_3 \ddot{g}_3).$$

Again we require the first bracket to be zero: $\dot{y}_1 \dot{g}_1 + \dot{y}_2 \dot{g}_2 + \dot{y}_3 \dot{g}_3 = 0$. This leaves $\ddot{x} = y_1 \ddot{g}_1 + y_2 \ddot{g}_2 + y_3 \ddot{g}_3$, so

$$\dddot{x} = (\dot{y}_1 \ddot{g}_1 + \dot{y}_2 \ddot{g}_2 + \dot{y}_3 \ddot{g}_3) + (y_1 \dddot{g}_1 + y_2 \dddot{g}_2 + y_3 \dddot{g}_3)$$

Now using that g_1, g_2 and g_3 satisfy the homogeneous equation, we find after some computation that if we want x to satisfy the inhomogeneous equation, then we must require the first bracket here to satisfy $\dot{y}_1 \ddot{g}_1 + \dot{y}_2 \ddot{g}_2 + \dot{y}_3 \ddot{g}_3 = f$. These three requirements are exactly the matrix equation $\Phi\,\dot{\mathbf{y}} = \mathbf{f}$ that we found above in a much simpler way, using systems and matrix algebra.

Exercises

- Reduction of order: A26, A27.
- Variation of constants: A28 (A28c).
- Wronskians: A29.

[Jump to course outline]

Additional problems

A26 Determine α^2 so that that the Bessel equation

$$t^2\ddot{x} + t\dot{x} + (t^2 - \alpha^2)\,x = 0 \qquad (t > 0)$$

has a solution $x_0(t) = t^{-1/2}\sin t$, and use reduction of order to find another (linearly independent) solution.

[Answer: $\alpha^2 = 1/4$, $x(t) = t^{-1/2}\cos t$.]

A27 (a) The usual rule based on the characteristic polynomial says that $x(t) = Ae^{2t} + Be^{3t}$ is the general solution of $\ddot{x} - 5\dot{x} + 6x = 0$. Give a direct proof that this *really* is the most general solution, by applying reduction of order using the known solution $x_0(t) = e^{2t}$.

(b) Similarly, show that $x(t) = (At + B)e^{-3t}$ is the general solution of $\ddot{x} + 6\dot{x} + 9x = 0$, by applying reduction of order with $x_0(t) = e^{-3t}$.

(If you look back in your calculus book, you will probably find arguments like these in the section explaining the theory behind the characteristic polynomial.)

A28 Find the general solution. (Variation of constants is useful for finding a particular solution, at least in the more difficult cases. But just for practice, use it in the simpler cases as well.)

(a) $\ddot{x} + 2\dot{x} + x = 2\sin t$ [Answer: $x(t) = (At + B)e^{-t} - \cos t$.]

(b) $\ddot{x} + 9x = \cos 3t$ [Answer: $x(t) = A\cos 3t + B\sin 3t + \frac{1}{6}t\sin 3t$.]

(c) $\ddot{x} + x = \tan t$ [Answer: $x(t) = A\cos t + B\sin t + \frac{1}{2}\cos t \cdot \ln\left|\frac{1-\sin t}{1+\sin t}\right|$.]

(d) $\dot{x}_1 = 2x_1 - 5x_2 + 4t$, $\dot{x}_2 = x_1 - 2x_2 + 1$ [Answer: $x_1(t) = A(\cos t + 2\sin t) + B(-5\sin t) + 8t - 1$, $x_2(t) = A\sin t + B(\cos t - 2\sin t) + 4t - 2$.]

(e) $(t^2 - 1)\ddot{x} - 2t\dot{x} + 2x = t^2 - 1$ (To find $x_{\text{hom}}(t)$, try a power series solution.) [Answer: $x(t) = At + B(1 + t^2) - t^2 + t\ln\left|\frac{1+t}{1-t}\right| + \frac{1}{2}(1 + t^2)\ln\left|1 - t^2\right|$.]

(And if you haven't had enough, you can look back at problem A2.)

A29 (a) Show that there is no second-order ODE $\ddot{x} + p_1(t)\,\dot{x} + p_0(t)\,x = 0$ whose general solution has the form $x(t) = At + B\cos t$, if we require the coefficients p_0 and p_1 to be defined and continuous for all $t \in \mathbf{R}$.

[Hint: Consider the Wronskian $W = y\dot{z} - \dot{y}z$ of $y(t) = t$ and $z(t) = \cos t$.]

(b) Show that there is such an ODE if we remove the requirement that p_0 and p_1 be defined on the whole real line.

[Hint: Just plug $x(t) = t$ and $x(t) = \cos t$ into the ODE and see what $p_0(t)$ and $p_1(t)$ must be in order for the ODE to be satisfied. In your answer, you should be able to see why there's a problem when the Wronskian becomes zero.]

(c) Find a 2×2 matrix $\Phi(t)$ such that its columns are linearly independent (as functions of t), but $\det\Phi(t)$ is zero at some points. [Hint: Same as for (a).]

(d) Show that things can be even worse than in (b): the functions $y(t) = t^3$ and $z(t) = |t|^3$ ($t \in \mathbf{R}$) are linearly independent, but their Wronskian is *identically* zero.

Lecture 11. Outlook: Poincaré maps, attractors, chaotic systems

(Arrowsmith & Place, selected parts of chapters 4 & 6.)

This material is not strictly a part of the course, but is provided as "edutainment", and to give a rough idea of some additional topics which are important in the theory of dynamical systems.

Lesson 4

[Jump to course outline]

All problems on one page

Lecture 1. 1.1, 1.2, 1.4, A1, A2, A3, 1.11, 1.12, 1.13, 1.14, 1.17*, A4

Lecture 2. 1.19acd, A5, A6, 1.20, 1.23, 1.24, 1.25, 1.32, 1.36

Lecture 3. 2.1, 2.3, 2.4, 2.8, 2.9

Lecture 4. 2.13, 2.14, A7, 2.22, 2.23, 2.29abcd, 2.30, 2.33, 2.35ad

Lecture 5. 3.5, 3.6, 3.7, A8, 3.8, 3.11, A9*, A10**

Lecture 6. 3.13abe, 3.14abe, 3.15, 3.17bc, A11, 3.19b, 3.18*, A12*, A13, A14, 3.22, A15*, 3.24, 3.28ab, 3.29*

Lecture 7. 3.35, A16*, 3.36, 3.42, 3.43, A17, A18

Lecture 8. 5.2, 5.3, 5.10, 5.15, 5.16, 5.17, 5.21, 5.23, A19, A20

Lecture 9. A21, A22, A23, A24, A25

Lecture 10. A26, A27, A28, A28c, A29

Lecture 11. —