MATH 2610

Ordinary Differential Equations

Abbreviations used throughout

DE = differential equations

ODE = ordinary differential equation

PDE = partial differential equation

IVP = initial value problem

IC = initial conditions

iff = if and only if

EX = example

Def = definition

Theo = theorem

Prop = proposition

LHS = left hand side

RHS = right hand side

$\boxtimes$ = indicates the end of a proof.

What is a differential equation?

We are all familiar with algebraic equations, e.g.,

$$x^2 + 2x + 3 = 0$$

In this case the unknown is the variable x and solution to this equation is a number that satisfies it. In this case $x = 1$ and $x = -3$ are solutions because

$$1^2 + 2 \cdot 1 - 3 = 0 \quad \text{and} \quad (-3)^2 + 2(-3) - 3 = 0,$$

whereas $x = 2$ is not a solution since

$$2^2 + 2 \cdot 2 - 3 \neq 0.$$

We can consider similar situations where the unknown is a function:

$$x f(x) - 2 + 3x^2 = 0.$$

Solving for $f(x)$ gives

$$f(x) = \frac{2 - 3x^2}{x} \qquad (x \neq 0)$$

More generally, we can have an equation for an unknown function f where derivatives of f also appear, e.g.,

$$\frac{df}{dx} - 3\cos x = 0.$$

Here, we want to find a function $f(x)$ whose derivative equals $3\cos x$. We know from calculus how to find such a function:

$$\frac{df}{dx} - 3\cos x = 0 \implies \int \frac{df}{dx}\, dx = 3\int \cos x\, dx$$

$$\implies f(x) = 3\sin x + C_1,$$ where C_1 is a constant of integration.

$$\boxed{\text{An equation relating an unknown function and one or more of its derivatives is called a differential equation (DE).}}$$

EX: These are DE:

$$\frac{dy}{dx} + y^2 x = 0 \quad \text{variable: } x, \text{ function } y = y(x)$$

$$\frac{dx}{dt} + e^{-t^2} = 0 \quad \text{variable: } t, \text{ function } x = x(t)$$

These are <u>not</u> DE:

$$x^2 - 4 = 0$$

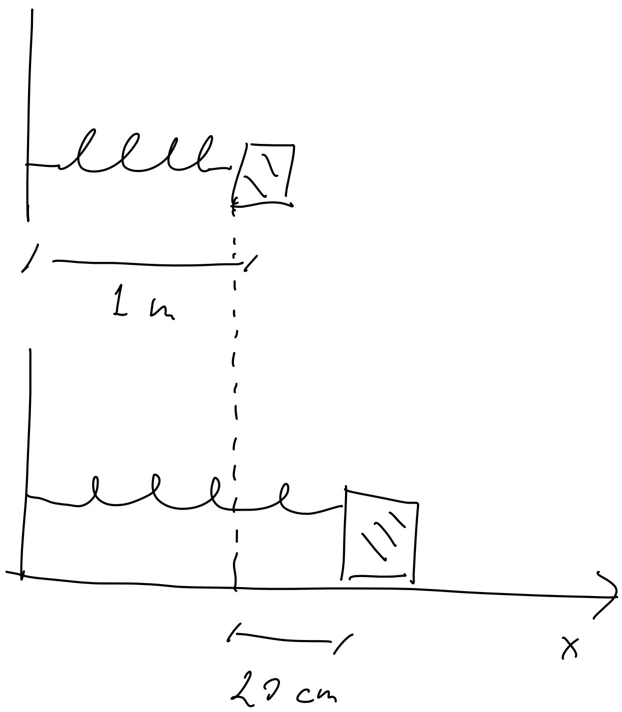
$$\int e^{t^2} y(t)\, dt = \log t - 4$$

$$\int (y(x))^2\, dx = \frac{dy}{dx} + 5x$$

(the second equation is called an integral equation and the third one an integral-differential equation).

Why do we study DE?

Let's investigate the following example.
Consider a spring that has length 1 m when it is not subject to any force. One end of the spring is attached to a wall and the other end to a body of mass 2 kg, as in the figure below.



Suppose you pull the body horizontally, stretching the spring 20 cm, and then release it. The body is going to oscillate back and forth. What is its position after 10 seconds? (disregard friction between the body and the floor. Consider that the spring has constant $k = 50$ N/m)

From Hooke's law we know that the force acting on the body due to the spring is $F = -kx$, where x is the displacement with respect to the equilibrium position, which we identify with $x = 0$.

Since $-kx$ is the only force acting on the body, it equals ma, where m is the block's mass and a its acceleration: $ma = -kx \Rightarrow a = -25x$, since $m = 2\,kg$ and $k = 50\,N/m$.

The position x is a function of time, $x = x(t)$. We want to know $x(10)$ (position at $t = 10\,s$). Since the acceleration is the second time derivative of the position, $a = \frac{d^2x}{dt^2}$, then $\frac{d^2x}{dt^2} + 25x = 0$. This is a DE for the unknown function x. We'll learn later on how to find x. For now, we can verify that $x(t) = 0.2\cos(5t)$ is the desired solution to the above DE since:

$$\frac{d^2}{dt^2}(0.2\cos(5t)) + 25 \cdot 0.2\cos(5t) = -0.2 \cdot 25 \cdot \cos(5t) + 0.2 \cdot 25 \cdot \cos(5t)$$

$$= 0.$$

The factor 0.2 stems from the fact that at time zero the position of the block is $20\,cm = 0.2\,m$, so that $x(0) = 0.2\cos(5 \cdot 0) = 0.2$. We can now calculate

$$x(10) = 0.2\cos(5 \cdot 10) \simeq 0.19\,m.$$

Some terminology and notation

We'll use $\frac{d}{dt}$, $\frac{d}{dx}$, $'$ etc to denote derivative. Hence particular names given to variables and function can change, the same equation might be written in different forms. E.g.,

$x'' - 5x' = e^x$ and $\frac{d^2y}{dt^2} - 5\frac{dy}{dt} = e^y$ both represent the same DE

Def. The order of a DE is the order of the highest derivative that it contains

For example, $y''' + xy^2 = 0$ is a DE of 3^{rd} order.

A solution to a DE is a function that satisfies the equation. E.g., $y = 2x^3$ is a solution of the D.E. $y' - 6x^2 = 0$, but $y = x^2$ is not. Notice that even though it might be difficult to find a solution of a DE it is easy to verify whether or not a given function is a solution: simply plug it into the DE and see if equality is satisfied.

Def. A DE of order n is said to be linear if it has the form:

$$a_n(t)\frac{d^n x(t)}{dt^n} + a_{n-1}(t)\frac{d^{n-1}x(t)}{dt^{n-1}} + \dots + a_1(t)\frac{dx}{dt} + a_0(t)\,x(t) = g(t)$$

where $a_n(t), \dots, a_0(t), g(t)$ are given functions and $a_n(t) \neq 0$. Otherwise, the equation is called non-linear. In the linear case, the functions $a_n(t), \dots, a_0(t)$ are called the coefficients of the equation.

Ex: $\frac{d^2y}{dt^2} + e^t\frac{dy}{dt} - \cos t\, y = 0$ and $x''' - x' = \log t$ are linear, while $(y')^2 = ye^y$ and $e^{y''} + xy = 0$ are non-linear.

Remark Because $a_n(t) \neq 0$ in the definition of a linear DE, we can always divide the equation by $a_n(t)$. Thus, without loss of generality we can say that a linear DE has the form:

$$\frac{d^n x(t)}{dt^n} + a_{n-1}(t)\frac{d^{n-1}x(t)}{dt^{n-1}} + \dots + a_0\,x(t) = g(t).$$

The distinction linear vs. non-linear DE is extremely important. Make sure you fully understand it.

It's important to notice that the unknown function of a DE can depend on more than one variable. For example, if T is a function that describes the temperature inside a room, then T is a function of space and time, so it depends on the three spatial coordinates $x, y,$ and z, and on the time t. Therefore, a DE governing the behavior of T might involve derivatives with respect to $x, y, z,$ and t, and in this case we would need to use partial derivatives, i.e., $\frac{\partial T}{\partial x}, \frac{\partial T}{\partial y}, \frac{\partial T}{\partial z}, \frac{\partial T}{\partial t}$ etc. Such types of DE are called partial differential equations (PDEs), while D.E. involving only one variable are called ordinary differential equations (ODEs).

E.g., $\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} + \frac{\partial^2 T}{\partial z^2} = \frac{\partial T}{\partial t}$ is a PDE for T, while $\frac{d^2 y}{dx^2} + y = 0$ is a ODE for y. In this course we deal only with ODEs, so the term DE will always mean ODE unless stated otherwise.

Initial value problem

Consider the DE $\frac{dy}{dx} = x^3$. We can find a solution by direct integration: $\int \frac{dy}{dx}\,dx = \int x^3\,dx$
$\Rightarrow y = \frac{x^4}{4} + C_1$ where C_1 is a constant of integration. So, instead of a unique solution to the DE, we have a family of solutions, i.e., a different solution for each different choice of C_1. In particular we have infinitely many solutions. Such a family of solutions is called a general solution of the DE.

If we want to determine C_1, we need further information. For example, suppose we want, among all solutions, a solution with the property $y(0) = 5$. Then, plugging $x = 0$ we have $y(0) = \frac{0^4}{4} + C_1 = 5 \Rightarrow C_1 = 5$. So $y = \frac{x^4}{4} + 5$ is the desired solution. In this case we are not solving only the DE $\frac{dy}{dx} = x^3$ but rather the problem

$$\begin{cases} \frac{dy}{dx} = x^3 \\ y(0) = 5 \end{cases}.$$

Such a problem is called an <u>initial value problem (IVP)</u>. The extra conditions given in order to determine the constants appearing in the general solution are called <u>initial conditions (IC)</u> (in the above example, $y(0) = 5$ is the initial condition).

The terminology IVP and IC are used because usually the variable is time. In our first example we investigated not only the DE $x'' + 25x = 0$ but rather the IVP

$$\begin{cases} x'' + 25x = 0 \rightarrow \text{DE} \\ \left.\begin{aligned} x(0) &= 0.2 \\ x'(0) &= 0 \end{aligned}\right\} \text{initial conditions} \end{cases}$$

(the initial condition $x'(0) = 0$ was implicit in the statement of the problem in that we pulled the string and released it, so its velocity $v = \frac{dx}{dt}$ at time zero was zero.)

As we are going to see in detail later on, to solve an IVP we need as many IC as the order of the equation. To have an idea of why this is the case, consider the following simple example: $y'' = e^{2x}$. Since $\int \frac{d^2y}{dx^2}\,dx = \frac{dy}{dx} + \text{constant}$, we have $\int y''\,dx = \int e^{2x}\,dx \Rightarrow y' = \frac{e^{2x}}{2} + C$, where C is a constant. Integrating again yields $y = \frac{e^{2x}}{4} + Cx + D$, where D is another constant. Thus, we have two arbitrary constants. To determine them we need two conditions. For example, we could have $y(0) = 2$ and $y'(0) = 3$. Then $y(0) = \frac{1}{4} + 0 + D = 2 \Rightarrow D = 7/4$. Next, compute $y'(x) = \frac{e^{2x}}{2} + C$, so $y'(0) = \frac{1}{2} + C = 3 \Rightarrow C = 5/2$. Thus $y(x) = \frac{e^{2x}}{4} + \frac{5}{2}x + \frac{7}{4}$ is a solution to the IVP

$$y'' = e^{2x}$$

$$y(0) = 2$$

$$y'(0) = 3$$

Notation. An arbitrary DE of order n for the unknown function $x = x(t)$ will be denoted:

$$F(t, x(t), x'(t), \dots, x^{(n-1)}(t), x^{(n)}(t)) = 0.$$

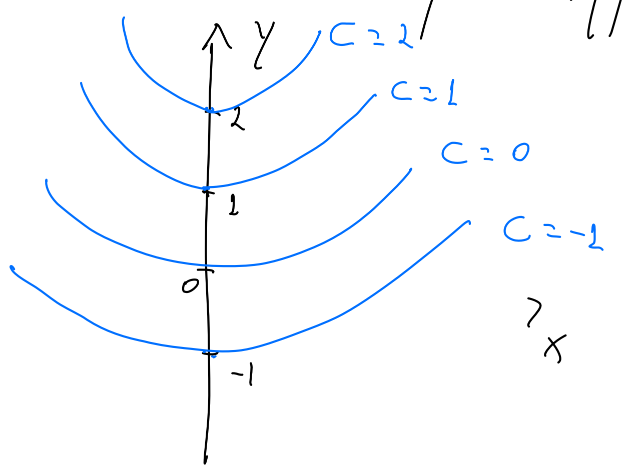
Def. By an initial value problem (IVP) for a DE of order n $F(t, x(t), \dots, x^{(n)}(t)) = 0$, we mean the following problem. Find a solution $x = x(t)$ to the DE defined on an interval (a, b) containing the point t_0 such that $x(t_0) = X_0$, $x'(t_0) = X_1$, ..., $x^{(n-1)}(t_0) = X_{n-1}$ where $X_0, X_1, \dots, X_{n-1}$ are given numbers.

Consider now $y' = \dfrac{2x - e^y}{xe^y + 1}$. We can verify that the function y satisfying the relation $xe^y + y = x^2$ is a solution to the DE. However, we cannot solve this relation explicitly for y. In this case we say we have an implicit solution to the DE.

General and particular solutions

Consider the DE $\frac{dy}{dx} = f(x)$, where f is a known function of x. We can solve this by direct integration: $y(x) = \int f(x)\,dx + C$, where C is an undetermined constant of integration. When a solution to a DE contains such undetermined constants we call it a general solution. When all undetermined constants have been found using IC we call it a particular solution.

A general solution represents a family of solutions. E.g., $\frac{dy}{dx} = 2x \Rightarrow y = x^2 + C$. Below we graph some of these solutions for different values of C:



If we want $y(0) = 0$, then we are selecting one solution in the family of solutions.

Remark. Notice that a general solution might not contain all solutions to a DE. For example, consider $\frac{dy}{dx} = y^2$. If $y \neq 0$, then $\frac{dy}{y^2} = dx \Rightarrow -\frac{1}{y} = x + C$ $\Rightarrow y = -\frac{1}{x+C}$. This is a general solution to the DE. But the function $y = 0$ (i.e., $y(x) = 0$ for all x) is also a solution to the DE, one which is not included in the formula $y = -\frac{1}{x+C}$. When a general solution includes all solutions then we call it the general solution.

Notation. We will use the letter C to denote arbitrary constants in general solutions. Sometimes we use the same letter C to note a different arbitrary constant. E.g., consider the DE $3y' = e^{3x}$, then $3\int \frac{dy}{dx}\,dx = \int e^{3x}\,dx \Rightarrow 3y = \frac{e^{3x}}{3} + C$ $\Rightarrow y = \frac{e^{3x}}{9} + \frac{C}{3}$. Since C is arbitrary so is $\frac{C}{3}$. We can call it another constant $D = C/3$. However, it is cumbersome to keep track of all the relabels of constants, so we denote $C/3$ by C again and write $y = \frac{e^{3x}}{9} + C$.

Existence theorem for first order equations

Theo. Suppose that $f(x,y)$ and $\frac{\partial f}{\partial y}(x,y)$ are continuous on a rectangle $R \subseteq \mathbb{R}^2$ containing the point (a,b). Then, the IVP $\begin{cases} y' = f(x,y) \\ y(a) = b \end{cases}$ has a unique solution defined on some interval I that contains a.

This theorem allows us to say when an IVP admits a unique solution, even though finding a formula for the solution might be very hard.

Ex: $\begin{cases} y' = x^2 e^{\sin[(x-y)^2]} \\ y(0) = 1 \end{cases}$

Here $f(x,y) = x^2 e^{\sin[(x-y)^2]}$. This function is continuous because it is the composition of continuous functions. Compute $\frac{\partial f}{\partial y} = x^2 e^{\sin[(x-y)^2]} \cos[(x-y)^2] \cdot (-2)(x-y)$, which is again a continuous function. Hence, the IVP has a unique solution defined in a neighborhood of $x=0$. Note that it will be very hard to find a formula for such solution.

<u>EX:</u> $\begin{cases} y' = \sqrt{x-y} \\ y(2) = 2 \end{cases}$. In this case $\frac{\partial f}{\partial y} = -\frac{1}{2\sqrt{x-y}}$

which is not continuous (in fact, not even defined) at $(2,2)$. Therefore, the theorem cannot be applied and we cannot guarantee that a unique solution exists.

In the previous example we are <u>not</u> saying that a solution does not exist, only that we cannot use the theorem.

<u>Remark</u>. It is important to verify not only that $\frac{\partial f}{\partial y}$ exist but also that it is continuous. Recall that it is possible for a function to be differentiable but for its derivative not to be continuous. For example, the function $f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & , \ x \neq 0 \\ 0 & , \ x = 0 \end{cases}$ is differentiable

but its derivative at $x = 0$ is not continuous.

Separable equations of first order

A first order DE $\frac{dy}{dx} = F(x,y)$ is called separable if $F(x,y) = g(x)h(y)$, or equivalently, $F(x,y) = \frac{g(x)}{f(y)}$.

In this case, we can find a solution by direct integration:

$$\frac{dy}{dx} = \frac{g(x)}{f(y)} \Rightarrow \int f(y)\,dy = \int g(x)\,dx.$$

Ex: $\frac{dy}{dx} = -6xy \underset{y\neq 0}{\Longrightarrow} \frac{dy}{y} = -6x$. Integrating:

$$\ln|y| = -3x^2 + C \Rightarrow |y| = e^{C}e^{-3x^2} \Rightarrow y = \underbrace{\pm e^{C}}_{=A} e^{-3x^2} = Ae^{-3x^2}.$$

When we divided by y, we had to assume $y \neq 0$. We see $y=0$ is also a solution to this DE. However, the solution $y=0$ is included in the family Ae^{-3x^2} as it corresponds to $A=0$.

Many times when we solve separable equations we have to divide by a function h of y, $h(y)$. This excludes the values where h vanishes. Those must be analyzed separately.

Ex: $\frac{dy}{dx} = y^2$.

If $y \neq 0$, then $\frac{dy}{y^2} = dx \Rightarrow -\frac{1}{y} = x + C_1$

$\Rightarrow y = -\frac{1}{x+C_1}$. This is a general solution to the DE. But the function $y = 0$ (i.e., $y(x) = 0$ for all x) is also a solution to the DE, one which is not included in the formula $y = -\frac{1}{x+C_1}$. Therefore the general solution is $y = -\frac{1}{x+C_1}$, $y = 0$.

Linear first order equations

Consider the DE

$$e^{-x}\frac{dy}{dx} - e^{-x} y = x^3 \quad \text{(linear, first order)}$$

Noting that $e^{-x}\frac{dy}{dx} - e^{-x} y = \frac{d}{dx}(e^{-x}y)$ we have:

$$\frac{d}{dx}(e^{-x}y) = x^3 \Rightarrow \int \frac{d}{dx}(e^{-x}y)\,dx = \int x^3\,dx$$

$$\Rightarrow e^{-x}y = \frac{x^4}{4} + C' \Rightarrow y = \frac{x^4}{4}e^{x} + C'e^{x}.$$

Consider now $\frac{dy}{dx} + y = \cos x$. In this case it is not true that $\frac{dy}{dx} + y = \frac{d}{dx}(\dots)$. But if we multiply the equation by e^x we have:

$$\underbrace{e^{x}\frac{dy}{dx} + e^{x}y}_{\frac{d}{dx}(e^{x}y)} = e^{x}\cos x \Rightarrow \int \frac{d}{dx}(e^{x}y)\,dx = \int e^{x}\cos x\,dx$$

Therefore: $e^{x}y = \frac{1}{2}e^{x}(\cos x + \sin x) + C'$

or $y = \frac{1}{2}(\cos x + \sin x) + C'e^{-x}$.

The idea for solving linear first order DE will be similar to the above example: try to multiply the equation by a suitable function so that the terms in y can be written as the derivative of a product.

A 1st order linear DE can always be written as

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

, where P and Q are known functions.

Multiply by $\mu(x)$, where $\mu(x)$ is a function to be determined.

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

We want the LHS to be the derivative of a product:

$$\underbrace{\mu(x)\frac{dy}{dx}} + \mu(x)\,P(x)\,y = \frac{d}{dx}\big(\mu(x)\,y\big)$$

$$= \frac{d\mu}{dx}y + \underbrace{\mu(x)\frac{dy}{dx}}$$

$$\Rightarrow \mu P \underbrace{y} = \frac{d\mu}{dx}\underbrace{y}$$

Thus $\frac{d\mu}{dx} = \mu P(x)$. This is a separable equation.

$$\frac{d\mu}{\mu} = P(x)\,dx \Rightarrow \int\frac{d\mu}{\mu} = \int P(x)\,dx \Rightarrow \ln|\mu| = \int P(x)\,dx + C'$$

$$\Rightarrow |\mu| = e^{C'}e^{\int P(x)\,dx}$$

, removing the absolute value:

$$\mu(x) = \pm e^{C'}e^{\int P(x)\,dx}.$$

We found a family of functions μ that allow us to write $\mu \frac{dy}{dx} + \mu P y$ as the derivative of a product. But we just need one such function, so we can take $C_1' = 0$ and take the $+$ sign. Thus

$$\frac{d}{dx}\left(\mu(x)\, y\right) = \mu(x)\, Q(x), \quad \text{where } \mu(x) = e^{\int P(x)\,dx}.$$

Integrating: $\int \frac{d}{dx}(\mu(x) y)\, dx = \int \mu(x) Q(x)\, dx$, so

$\mu(x) y(x) = \int \mu(x) Q(x)\, dx + C'$. Dividing by $\mu(x)$ (note that it never vanishes) and using its explicit form:

$$\boxed{y(x) = e^{-\int P(x)\,dx}\left(\int e^{\int P(x)\,dx} Q(x)\, dx + C'\right)}$$

This is an explicit formula for the general solution of $\frac{dy}{dx} + P(x) y = Q(x)$.

Remark. Note that the above formula is for the equation $y' + P(x) y = Q(x)$, i.e., the coefficient of y' must be 1. If we have $a(x) y' + b(x) y = c(x)$, we must first divide by $a(x)$ to use the formula.

Students should not only memorize the above formula for $y(x)$, but also know how to derive it.

Ex: $\dfrac{dy}{dx} - y = \dfrac{11}{8} e^{-\frac{x}{3}}, \quad y(0) = 1$

In this case $P(x) = -1$, $Q(x) = \dfrac{11}{8} e^{-\frac{x}{3}}$. Then $\mu(x) = e^{\int P(x)\,dx} = e^{-x}$, $\int e^{\int P(x)\,dx} Q(x) = \int \dfrac{11}{8} e^{-x} e^{-x/3}\,dx$
$= -\dfrac{33}{32} e^{-\frac{4x}{3}}$. Therefore $y(x) = e^{-(-x)}\left(-\dfrac{33}{32} e^{-\frac{4x}{3}} + C'\right)$
$= e^{x}\left(-\dfrac{33}{32} e^{-\frac{4x}{3}} + C'\right)$. Plugging $y(0) = 1$ we find $C' = \dfrac{65}{32}$,
so $y(x) = \dfrac{65}{32} e^{x} - \dfrac{33}{32} e^{-\frac{x}{3}}$.

A legitimate question is whether our formula for y always works. This is answered by the following theorem:

Theo (existence and uniqueness of solutions for 1st order linear DE).
Assume that $P(x)$ and $Q(x)$ are continuous on an interval (a,b) that contains the point x_0. Then, for any y_0, the IVP
$\begin{cases} y' + P(x)\,y = Q(x) \\ y(x_0) = y_0 \end{cases}$ has a unique solution defined on (a,b).
Moreover, the solution can be written as
$y(x) = e^{-\int P(x)\,dx}\left(\int e^{\int P(x)\,dx} Q(x)\,dx + C'\right)$ for a suitable constant C'.

proof Since $P(x)$ and $Q(x)$ are continuous, the integrals $\int P(x)\,dx$ and $\int e^{\int P(x)dx} Q(x)\,dx$ are well-defined and define differentiable functions on (a,b). Set

$$y(x) = e^{-\int P(x)dx}\left(\int e^{\int P(x)dx} Q(x)\,dx + C'\right),$$

where C' is a constant. Then y is differentiable. Compute:

$$y' = \left(e^{-\int P(x)dx}\right)'\left(\int e^{\int P(x)dx} Q(x)\,dx + C'\right) + e^{-\int P(x)dx}\left(\int e^{\int P(x)dx} Q(x)\,dx + C'\right)'$$

$$= -e^{-\int P(x)dx}\left(\int P(x)\,dx\right)'\left(\int e^{P(x)dx} Q(x)\,dx + C'\right) + e^{-\int P(x)dx}\left(\int e^{\int P(x)dx} Q(x)\,dx\right)'$$

$$= -P(x)\underbrace{e^{-\int P(x)dx}\left(\int e^{\int P(x)dx} Q(x)\,dx + C'\right)}_{=y} + \underbrace{e^{-\int P(x)dx}\, e^{\int P(x)dx}}_{=1} Q(x),$$

where we used the product rule in the first line, the chain rule in the second line, and the fundamental theorem of calculus in the third line.

Thus $y' + P(x)y = Q(x)$ and y satisfies the DE.

Because $e^{-\int P(x)dx}$ never vanishes, we can always solve for C' and determine it so that $y(x_0) = y_0$. $\boxed{\checkmark}$

Exact equations

Let us introduce this topic with the following example.

Consider the DE

$$\left(4y+3x^2-3xy^2\right)\frac{dy}{dx} = y^3-6xy.$$

Write it as

$$(6xy-y^3)\,dx + (4y+3x^2-3xy^2)\,dy = 0$$

Set $M(x,y) = 6xy-y^3$, $N(x,y) = 4y+3x^2-3xy^2$, so that the DE becomes:

$$M(x,y)\,dx + N(x,y)\,dy = 0$$

Now let us ask: is the LHS the differential of a function? In other words, does there exist a $F(x,y)$ such that $dF = M\,dx + N\,dy$? If the answer is yes, then the DE becomes $dF = 0$, which implies that F is constant. In this case the general solution of the DE will be simply $F(x,y) = C$.

Recall from calculus that $dF = M\,dx + N\,dy$ iff $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ (we'll state this more precisely below).

We check!

$$\left.\begin{aligned} \frac{\partial M}{\partial y} &= \frac{\partial}{\partial y}(6xy - y^3) = 6x - 3y^2 \\ \frac{\partial N}{\partial x} &= \frac{\partial}{\partial x}(4y + 3x^2 - 3xy^2) = 6x - 3y^2 \end{aligned}\right\} \Rightarrow \frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$$

Therefore, there exists a function $F = F(x,y)$ such that $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N$. Let's proceed to find F.

$\frac{\partial F}{\partial x} = M = 6xy - y^3$. Integrating with respect to x gives $F(x,y) = \int (6xy - y^3)\,dx = 3x^2y - xy^3 + g(y)$.

After performing the integration we added a function $g(y)$. This is because we must add a constant of integration. But here we are integrating a function of x and y with respect to x so that anything that depends on y only is treated as a constant from the point of view of $\int \dots dx$. Therefore, the "constant" of integration can in principle be a function of y.

To find $g(y)$ we use that

$$\frac{\partial F}{\partial y} = N.$$

Taking $\frac{\partial}{\partial y}$ of the expression we found for F and setting the result equal to N:

$$\frac{\partial}{\partial y} F = \frac{\partial}{\partial y}\left(3x^2y - xy^3 + g(y)\right) = 3x^2 - 3xy^2 + g'(y) = N = 4y + 3x^2 - 3xy^2$$

$$\Longrightarrow \quad x^2 - 3xy^2 + g'(y) = 4y + 3x^2 - 3xy^2 \Longrightarrow g'(y) = 4y$$

This is an equation for $g(y)$ that can be solved by direct integration. Notice how all the x's cancelled and the equation for $g(y)$ involves only y. This must be the case: by construction, g is a function of y only. If we end up with an equation for g involving x, then there is a mistake somewhere.

The equation for g is easily solved, giving $g(y) = 2y^2$. We have not added a constant of integration to g because the solution of the DE already contains an undetermined constant.

Summing up, we have $F(x,y) = 3x^2y - xy^3 + 2y^2$ and the general solution to the DE is $F(x,y) = C$, or:

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

Remark. Above, we found the solution $3x^2y - xy^3 + 2y^2 = C$, but we have not solved explicitly for y. In many cases, it is impossible to find an explicit expression for y. In these cases, i.e., when the solution is given as $F(x,y) = C$, with no explicit expression for y, we say that we have an implicit solution.

We will now streamline the ideas of the previous example.

Def. A first order DE written in the form

$$M(x,y)\,dx + N(x,y)\,dy = 0$$

is called exact if there exists a function $F = F(x,y)$ such that $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N$.

Under appropriate hypotheses, we will show that a DE is exact iff $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Before doing so, we will summarize the method.

Method for solving exact equations

1. Given $y' = f(x,y)$, write it as $M(x,y)\,dx + N(x,y)\,dy = 0$.

2. Test if $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$. If this is not the case, then the method cannot be applied. Otherwise, proceed as follows:

3. If $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$, then define F by

$$F(x,y) = \int M(x,y)\,dx + g(y)$$

where g is a function of y only that needs to be determined.

4. To determine g, take $\dfrac{\partial}{\partial y}$ of F found in step 3, and set it equal to N. This gives an equation for g of the form:

$$g'(y) = \text{expression in } y \text{ containing no } x$$

5. Integrate $g'(y)$ to obtain $g(y)$ and thus $F(x,y)$.

6. The general solution is given by $F(x,y) = C$, where C is an arbitrary constant.

Remark. If the expression for $g'(y)$ found in step 4 involves x, then there is a mistake, and we must re-check the calculations.

Remark. In step 3, we can first integrate in y. I.e., if $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$, then $\dfrac{\partial F}{\partial y} = N$. Integrating with respect to y

produces $F(x,y) = \int N(x,y)\,dy + h(x)$, where h is a function of x only. To find h, we differentiate F with respect to x and set the resulting expression equal to M. This will give an equation for $h'(x)$ involving no y (if it contains y, then there is a mistake). Integrating we find h, and hence F.

In the next example we use this idea of integrating in y first.

<u>Ex</u>: $y' = \tan x \tan y$.

Write the equation as $dy - \tan x \tan y\, dx = 0$. Multiply by $\cos x \cos y$ to obtain $\underbrace{-\sin x \sin y}_{= M(x,y)}\, dx + \underbrace{\cos x \cos y}_{= N(x,y)}\, dy = 0$

Compute $\dfrac{\partial M}{\partial y} = -\sin x \cos y$, $\dfrac{\partial N}{\partial y} = -\sin x \cos y$, so $\dfrac{\partial M}{\partial y} = \dfrac{\partial N}{\partial x}$.

Then $\dfrac{\partial F}{\partial y} = N \Rightarrow F(x,y) = \int N(x,y)\,dy + h(x) = \int \cos x \cos y\, dy + h(x)$
$= \cos x \sin y + h(x)$. Then, $\dfrac{\partial F(x,y)}{\partial x} = -\sin x \sin y + h'(x) = M(x,y) = -\sin x \sin y$.

Therefore $h'(x) = 0$. This means that $h(x)$ is constant. Recalling that we do not include constants of integration at this point, we can take $h(x) = 0$. Thus

$$F(x,y) = \cos x \sin y = C', \quad \text{or} \quad y = \sin^{-1}\left(\frac{C}{\cos x}\right).$$

Remark. In the above example, if we consider the equation written as $dy - \tan x \tan y \, dx = 0$ and take $N(x,y) = 1$, $M(x,y) = -\tan x \tan y$, then we do not obtain $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$. Only after multiplying the equation by $\cos x \cos y$ the condition is satisfied. Thus, how we reorganize the terms can matter.

The next theorem assures that the steps given for solving $M\,dx + N\,dy = 0$ always work if $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$ (and suitable hypotheses are satisfied).

Theo. Suppose the partial derivatives of $M(x,y)$ and $N(x,y)$ exist and are continuous on a rectangle $R \subseteq \mathbb{R}^2$. Then, $M(x,y)\,dx + N(x,y)\,dy = 0$ is exact iff the compatibility condition $\frac{\partial M}{\partial y}(x,y) = \frac{\partial N}{\partial x}(x,y)$ holds for all $(x,y) \in R$.

proof: Assume that the equation is exact, i.e., that there exists a $F = F(x,y)$ such that $dF = M\,dx + N\,dy$. Since $dF = \frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy$, we have $\frac{\partial F}{\partial x} = M$ and $\frac{\partial F}{\partial y} = N$. By assumption, the first derivatives of M and N exist and are continuous, hence the second partial derivatives of F exist and are continuous. Under these circumstances, we have $\frac{\partial^2 F}{\partial y \partial x} = \frac{\partial^2 F}{\partial x \partial y}$. Thus, $\frac{\partial^2 F}{\partial y \partial x} = \frac{\partial}{\partial y}\frac{\partial F}{\partial x} = \frac{\partial M}{\partial y} = \frac{\partial^2 F}{\partial x \partial y}$ $= \frac{\partial}{\partial x}\frac{\partial F}{\partial y} = \frac{\partial N}{\partial x}$, showing that $\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x}$.

Reciprocally, assume the compatibility condition. Let $(x_0, y_0) \in R$. We claim that the expression

$$N(x,y) - \frac{\partial}{\partial y}\int_{x_0}^{x} M(t,y)\,dt$$

is a function of y only. For, compute $\frac{\partial}{\partial x}\left(N(x,y) - \frac{\partial}{\partial y}\int_{x_0}^{x} M(t,y)\,dt\right)$ $= \frac{\partial N(x,y)}{\partial x} - \frac{\partial}{\partial y}\frac{\partial}{\partial x}\int_{x_0}^{x} M(t,y)\,dt = \frac{\partial N(x,y)}{\partial x} - \frac{\partial M(x,y)}{\partial y} = 0$, where we used tha M and N have continuous partial derivatives and the fundamental theorem of calculus. Thus, as the partial derivative with respecto to x of $N(x,y) - \frac{\partial}{\partial y}\int_{x_0}^{x} M(t,y)\,dt$ vanishes, we conclude that it depends on y only.

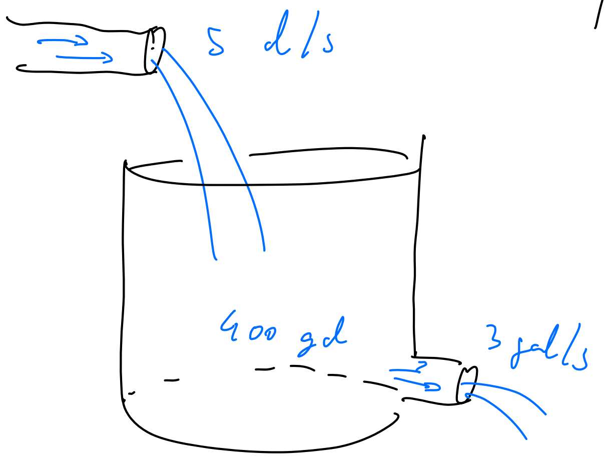
Because of the claim, we can define $g(y)$ as a solution to $g'(y) = N(x,y) - \frac{\partial}{\partial y}\int_{x_0}^{x} M(t,y)\,dt$.

We now define $F(x,y) = \int_{x_0}^{x} M(t,y)\,dt + g(y)$. A direct computation shows that $dF = M\,dx + N\,dy$. $\boxtimes$

Tank problems (compartimental analysis)

We are interested in modeling situations as in the following example.

<u>Ex</u>: A 400 gal tank initially contains 100 gal of brine containing 50 lbs of salt. Brine containin 1 lb of salt per gallon enters the tank at a rate 5 gal/s and the well-mixed brine flows out at a rate 3 gal/s. How much salt will the tank contain when it is full?



Denote by $x(t)$ the amount of salt in the tank at time t. Note that $x(0) = 50$ lbs. We need to find a DE for $x(t)$, solve it, and compute $x(t_*)$, where t_* is the time when the tank fills up.

To find the DE, let us first think of the process as discrete, i.e, imagine constructing a table with the amount of salt at, say, every second

t	$x(t)$
0	$x(0) = 50$ lb
1	$x(1)$
2	$x(2$
$\vdots$	$\vdots$
t	$x(t)$
$t+\Delta t$	$x(t+\Delta t)$

If we denote by Δt the time interval between two steps, then the amount of salt is the next step is:

$$x(t+\Delta t) = x(t) + \Delta x$$

where Δx = change in the quantity of salt between time t and $t+\Delta t$.
Observe that:

$$\Delta x = \begin{matrix}\text{quantity of salt}\\ \text{coming in during}\\ \text{the interval } \Delta t\end{matrix} - \begin{matrix}\text{quantity of salt}\\ \text{going out during}\\ \text{the interval } \Delta t\end{matrix}$$

If brine flows out at 3 gal/s and the concentration of the solution at time t is $s(t) = \frac{\text{mass}}{\text{volume}} = \frac{x(t)}{V(t)}$, where $V(t)$ = volume at time t, we have that the amount of salt leaving the tank per second.

$$3\ \frac{\text{gal}}{\text{s}}\ s(t)\ \frac{\text{lb}}{\text{gal}} = 3\ \frac{x(t)}{V(t)}\ \frac{\text{lb}}{\text{s}}$$

Because the initial volume is 100 gal, 5 gal come in and 3 gal go out every second, we have:

$$V(t) = 100 + 5t - 3t = 100 + 2t$$

Therefore, the amount of salt leaving the tank per secon is $\frac{3x(t)}{100+2t}\ \frac{lb}{s}$. This is not yet the amount o salt going out during the interval Δt, as the latter is measured in lb and not lb/s. We have:

$$\begin{array}{l}\text{quantity of}\\ \text{salt going out}\\ \text{during the period } \Delta t\end{array} = \frac{3x(t)}{100+2t}\ \frac{lb}{s} \cdot \Delta t\, s = \frac{3x(t)}{100+2t}\,\Delta t\ lb$$

Notice how keeping track of the units (lb/s, s, etc.) is useful to check that we have the right quantities.

Similarly:

$$\begin{array}{l}\text{quantity of}\\ \text{salt coming in}\\ \text{during the interval } \Delta t\end{array} = 1\ \frac{lb}{gal} \cdot 5\ \frac{gal}{s}\ \Delta t\, s = 5\Delta t\ gal$$

Thus $\Delta x = 5\Delta t - \frac{3x(t)}{100+2t}\Delta t$ and

$$x(t+\Delta t) = x(t) + \left(5 - \frac{3x(t)}{100+2t}\right)\Delta t,$$ giving

$$\frac{x(t+\Delta t) - x(t)}{\Delta t} = 5 - \frac{3x(t)}{100+2t}$$

The process is not, in fact, discrete, so we need to take the limit $\Delta t \to 0$. When we do so,

$$\lim_{\Delta t \to 0} \frac{x(t+\Delta t) - x(t)}{\Delta t} = \frac{dx(t)}{dt}$$, and we obtain:

$$\frac{dx}{dt} = 5 - \frac{3x}{100+2t}$$. We thus have that the process is modeled by the IVP

$$\begin{cases} \dfrac{dx}{dt} + \dfrac{3}{100+2t} x = 5 \\ x(0) = 50 \end{cases}$$

The DE is a linear first order equation with $P(t) = \frac{3}{100+2t}$ and $Q(t) = 5$. Compute:

$$\int \frac{3}{100+2t} dt = \frac{3}{2} \ln|100+2t| = \ln|100+2t|^{3/2}$$ so we get

$$e^{\int P(t)\,dt} = (100+2t)^{3/2}$$. Then $\int e^{\int P(t)} Q(t)\,dt = 5\int (100+2t)^{3/2} dt$

$= (100+2t)^{5/2}$. Therefore:

$$x(t) = e^{-\int P(t)\,dt}\left(\int e^{\int P(t)\,dt} Q(t)\,dt + C\right) = (100+2t)^{-3/2}\left((100+2t)^{5/2} + C\right)$$

Using $x(0) = 50 = 100^{-3/2}(100^{5/2} + C) = 10^{-3}(10^5 + C)$, so

$50 \cdot 10^4 - 10^{-5} = C$, $C = 5 \cdot 10^4 - 10 \cdot 10^4 = -5 \cdot 10^4$.

We obtain: $x(t) = (100+2t)^{-3/2}\left((100+2t)^{5/2} - 5.10^4\right)$.

Recall that we want $x(t)$ at the time when the tank is full. This happens when $V(t) = 400$, so $100 + 2t = 400$, $t = 150s$. Finally:

$x(150) = 400^{-3/2}\left(400^{5/2} - 5.10^4\right) \simeq 393.75\ lb$.

We note that there is a more direct way to construct the DE. We know that the change in $x(t)$ is $\frac{dx}{dt} = in - out$. Keeping track of the units, it is eas to figure out the "in" and "out" quantities:

$$\frac{dx}{dt} = 1\,\frac{lb}{gal}\cdot 5\,\frac{gal}{s} - \frac{x(t)}{V(t)}\,\frac{lb}{gal}\cdot 3\,\frac{gal}{s}, \quad V(t) = 100+2t,$$

so that $\frac{dx}{dt} = 5 - \frac{3x}{100+2t}$. $\left(\frac{dx}{dt} \text{ is measured in } \frac{lb}{s}\right)$.

However, students should understand the construction with Δx and Δt. In more complex applications, it is hard to "read off" all quantities directly, and the construction with Δx, Δt, etc. is more appropriate.

The mass-spring oscillator

Suppose a block of mass m is attached to a spring and the other end of the spring is attached to a wall as indicated in the figure:

If we pull the spring and release it, the block will move back and forth.

We want to find a DE modeling the motion of the block.

We assume that the block moves only in the horizontal direction. We choose a coordinate system with the x axis in the direction of the block's motion, with $x=0$ marking the position when the block is at rest.

We denote by $x=x(t)$ the position of the block at time t. The force on the block due to the spring is given by Hooke's law: $F_{spring} = -kx$, where k is a constant depending on the spring. Another force acting on the block is caused by the friction between the block and the floor. The force of friction is usually modeled as proportional to the velocity so we assume $F_{friction} = -p\frac{dx}{dt}$, where p is a non-negative constant. Finally, we assume that the block is also subject to an external force $F_{ext}(t)$ (a known function of t). Newton's law gives:

$$ma = -kx - p\frac{dx}{dt} + F_{ext}(t),$$ where a is the block's acceleration. Since $a = \frac{d^2x}{dt^2}$, we have:

$$m\frac{d^2x}{dt^2} + p\frac{dx}{dt} + kx = F_{ext}(t).$$

This is a second order linear DE for $x(t)$. An IVP for this DE must contain two IC. Physically they correspond to the initial position $x(0)$ and initial velocity $x'(0)$ of the block.

The above example illustrates an important physical situation where 2nd order linear equations appear. There are many other physical scenarios involving 2nd order linear equations. We will next study these equations in detail.

Homogeneous linear second order equations

Consider the DE

$$a x'' + b x' + c x = 0$$

where a, b, c are constants, $a \neq 0$, and $x = x(t)$ is the unknown. This equation is called homogeneous because there is no term without the unknown x. Otherwise we call the equation non-homogeneous (or inhomogeneous).

For example, $2x'' + x = 0$ and $x'' - x' + x = 0$ are homogeneous, whereas $2x'' + x = t^2$ and $x'' - x' + x = 10$ are non-homogeneous. We will study homogeneous equations first.

Ex: Consider $x'' + x' - 2x = 0$.

Let us show that $x(t) = e^{\lambda t}$, λ = constant, is a solution for appropriate values of λ. Plugging in:

$$(e^{\lambda t})'' + (e^{\lambda t})' - 2e^{\lambda t} = 0$$

$\lambda^2 e^{\lambda t} + \lambda e^{\lambda t} - 2e^{\lambda t} = 0$. Since $e^{\lambda t} \neq 0$ for all t, we must have $\lambda^2 + \lambda - 2 = 0$ or

$$(\lambda - 1)(\lambda + 2) = 0 \Rightarrow \lambda = 1 \text{ or } \lambda = -2.$$

Therefore, e^t and e^{-2t} are solutions to the DE. Indeed:

$$(e^t)'' + (e^t)' - 2e^t = e^t + e^t - 2e^t = 0$$

and

$$(e^{-2t})' + (e^{-2t})' - 2e^{-2t} = 4e^{-2t} - 2e^{-2t} - 2e^{-2t} = 0.$$

We will see that this simple idea of plugging $e^{\lambda t}$ is the basis for solving $ax'' + bx' + cx = 0$.

Consider again

$$a \,'' + bx' + cx = 0$$

Let us try to find a solution of the form $x = e^{\lambda t}$. Notice that at this point this is an "educated guess," i.e., we do not really know if $e^{\lambda t}$ in fact solves the DE. Plugging in:

$$a(e^{\lambda t})'' + b(e^{\lambda t})' + c e^{\lambda t} = 0$$

$(a\lambda^2 + b\lambda + c)e^{\lambda t} = 0$. Since $e^{\lambda t} \neq 0$ for all t we conclude that

$$\boxed{a\lambda^2 + b\lambda + c = 0}$$

which is an equation for λ called characteristic equation (also called auxiliary equation).

The roots of the characteristic equation are

$$\lambda_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad \lambda_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}$$

By construction, $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ are solutions to the DE $ax'' + bx' + c = 0$. Are there other solutions? How do we obtain the general solution? To answer these questions we need to develop the theory of second order linear homogeneous equations further. We begin motivating the discussion with the following example:

EX: Consider $x'' - 2x' + x = 0$. The characteristic equation is $\lambda^2 - 2\lambda + 1 = (\lambda - 1)^2 = 0$ giving $\lambda_1 = \lambda_2 = 1$. Thus, $x_1 = e^t$ solves the DE. We can verify that the function $x_2 = te^t$ is also a solution:

$$(te^t)'' - 2(te^t)' + te^t = (e^t + te^t)' - 2(e^t + te^t) + te^t$$

$$= e^t + e^t + te^t - 2e^t - 2te^t + te^t = 0 \text{ as claimed.}$$

The solution te^t did not come solely from the characteristic equation. How do we know if such "extra" solutions exist, and how do we find them? We will now investigate these questions.

Def: Two functions $x_1(t)$ and $x_2(t)$ are said to be linearly independent on an interval I if neither of them is a constant multiple of the other on all of I. Otherwise $x_1(t)$ and $x_2(t)$ are called linearly dependent.

Ex: The functions $\sin t$ and $\cos t$ are linearly independent on $(-\frac{\pi}{2}, \frac{\pi}{2})$. For, suppose that $\sin t = c \cos t$ for some constant c. Then $\tan t = c$. But this would have to hold for all $t \in (-\frac{\pi}{2}, \frac{\pi}{2})$, which would imply that $\tan t$ is constant on $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Ex: The functions $\sin 2t$ and $6 \sin t \cos t$ are linearly dependent on $\mathbb{R}$, because $6 \sin t \cos t = 3 \cdot 2 \sin t \cos t = 3 \sin(2t)$, where we used the trigonometric identity $\sin(\alpha + \beta) = \sin\alpha \cos\beta + \sin\beta \cos\alpha$.

Given two functions $x_1(t)$ and $x_2(t)$, a linear combination of them is the function

$$x(t) = c_1 x_1(t) + c_2 x_2(t)$$

where c_1 and and c_2 are constants. If $x_1(t)$ and $x_2(t)$ are solutions of the DE $ax'' + bx' + cx = 0$, so is any linear combination of x_1 and x_2. To see this, plug in $x(t)$ to find:

$$ax'' + bx' + cx = a(c_1x_1 + c_2x_2)'' + b(c_1x_1 + c_2x_2)' + c(c_1x_1 + c_2x_2)$$

$$= ac_1x_1'' + ac_2x_2'' + bc_1x_1' + bc_2x_2' + cc_1x_1 + cc_2x_2$$

$$= c_1\underbrace{(ax_1'' + bx_1' + cx_1)}_{=0} + c_2\underbrace{(ax_2'' + bx_2' + cx_2)}_{=0} = 0,$$

showing that $x(t)$ is a solution.

In particular, since we can take $c_2 = 0$ above, we also conclude that a multiple of a solution is also a solution.

Def. Let $x_1(t)$ and $x_2(t)$ be two differentiable functions defined on an interval I. The function:

$$W(x_1, x_2)(t) = x_1(t)\, x_2'(t) - x_2(t)\, x_1'(t)$$

is called the Wronskian of x_1 and x_2.

Theo. For any real numbers a, b, c, X_0, X_1, t_0, $a \neq 0$, there exist a unique solution to the IVP

$$\begin{cases} a x'' + b x' + c x = 0 \\ \qquad x(t_0) = X_0 \\ \qquad x'(t_0) = X_1 \end{cases}$$

The solution is valid for all $t \in (-\infty, \infty)$.

Remark. The theorem implies that if x and its derivative both vanish at some point t_0 then $x(t) = 0$ for all t.

Lemma. Let $x_1(t)$ and $x_2(t)$ be two solutions to the DE $ax'' + bx' + cx = 0$ on $(-\infty, \infty)$, $a \neq 0$, a, b, c constants. If $W(x_1, x_2)(\tau) = 0$ holds at some $\tau \in (-\infty, \infty)$, then it vanishes identically and x_1 and x_2 are linearly dependent.

proof: If $x_1(\tau)=0$ and $x_1'(\tau)=0$, then $x_1(t)=0$ for all and $x_1(t)=0\cdot x_2(t)$.

If $x_1(\tau)\neq 0$ then $z(t)=\frac{x_2(\tau)}{x_1(\tau)}x_1(t)$ solves the DE and $z(\tau)=x_2(\tau)$. Moreover,

$$z'(\tau)=\frac{x_2(\tau)}{x_1(\tau)}x_1'(\tau)=x_2'(\tau)\quad \text{since } W(x_1,x_2)(\tau)=0.$$

Hence $z_2(t)=x_2(t)$ by uniqueness and $x_2(t)=\frac{x_2(\tau)}{x_1(\tau)}x_1(t)$.

Finally if $x_1(\tau)=0$ but $x_1'(\tau)\neq 0$, then $W(x_1,x_2)(\tau)=0$ implies $x_2(\tau)=0$. The function $z(t)=\frac{x_2'(\tau)}{x_1'(\tau)}x_1(t)$ satisfies the DE and $z'(\tau)=x_2'(\tau)$. Since $z(\tau)=0$ we conclude by uniqueness that $z(t)=x_2(t)$, finishing the proof.

$\blacksquare$

Theo. If $x_1(t)$ and $x_2(t)$ are two linearly independent solutions to the DE $ax''+bx'+cx=0$ on $(-\infty,\infty)$, a,b,c constants, $a\neq 0$, then unique constants c_1 and c_2 can always be found such that $x(t)=c_1x_1(t)+c_2x_2(t)$ satisfies the IC $x(t_0)=X_0$, $x'(t_0)=X_1$, for any $X_0,X_1\in\mathbb{R}$.

proof: The function $x(t)$ defined in the statement solves the DE. Consider

$$x(t_0) = c_1 x_1(t_0) + c_2 x_2(t_0) = \Xi_0$$
$$x'(t_0) = c_1 x_1'(t_0) + c_2 x_2'(t_0) = \Xi_1$$

We solve this system for c_1 and c_2:

$$c_1 = \frac{\Xi_0 x_2'(t_0) - \Xi_1 x_2(t_0)}{x_1(t_0) x_2'(t_0) - x_1'(t_0) x_2(t_0)}, \quad c_2 = \frac{\Xi_1 x_1(t_0) - \Xi_0 x_1'(t_0)}{x_1(t_0) x_2'(t_0) - x_1'(t_0) x_2(t_0)}$$

provided the the denominator in these expressions is not zero. By the previous lemma and our assumption that $x_1(t)$ and $x_2(t)$ are linearly independent, this is the case.

$\boxed{\text{M}}$

We will now derive some important consequences of the above results.

We first ask the following question: can any solution of $ax'' + bx' + cx = 0$ be written as $c_1 x_1 + c_2 x_2$ for two linearly independent functions x_1 and x_2?

Let x be a solution to $ax'' + bx' + cx = 0$ and x_1 and x_2 be two linearly independent solutions. Pick $t_0 \in \mathbb{R}$. By the previous theorem, we can find c_1 and c_2 such that $c_1 x_1(t_0) + c_2 x_2(t_0) = x(t_0)$ and $c_1 x_1'(t_0) + c_2 x_2'(t_0) = x'(t_0)$. By uniqueness of solutions to the corresponding IVP, we conclude that $x = c_1 x_1 + c_2 x_2$. Thus:

Let x_1 and x_2 be two linearly independent solutions to $ax'' + bx' + cx = 0$, where a, b, c are constants and $a \neq 0$. Then any other solution $x(t)$ can be written as

$$x = c_1 x_1 + c_2 x_2$$

where c_1 and c_2 are constants. In particular, the general solution can be written as $c_1 x_1 + c_2 x_2$.

We saw that we can use the Wronskian to determine that two solutions are linearly dependent if they Wronskian vanishes. It follows that if two solutions are linearly independent, their Wronskian is not zero. We can ask the converse: if the Wronskian is not zero, are the solutions linearly independent?

The answer is yes, and is summarized in the following lemma.

Lemma. Let $x_1(t)$ and $x_2(t)$ be two solutions to the DE $ax''+bx'+cx=0$ on $(-\infty,\infty)$, $a\neq 0$, a,b,c constants. If $W(x_1,x_2)(\tau)\neq 0$ holds at some $\tau\in(-\infty,\infty)$, then it never vanishes and x_1 and x_2 are linearly independent.

Consider now the characteristic equation $a\lambda^2+b\lambda+c=0$ and let λ_1 and λ_2 be its two solutions. If λ_1 and λ_2 are real numbers and $\lambda_1\neq\lambda_2$, then $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ are solutions to the DE, as we have seen.

We now claim that $e^{\lambda_1 t}$ and $e^{\lambda_2 t}$ are linearly independent. For this, we compute the Wronskian:

$$\begin{aligned}W(e^{\lambda_1 t},e^{\lambda_2 t})(t) &= e^{\lambda_1 t}(e^{\lambda_2 t})'-(e^{\lambda_1 t})'e^{\lambda_2 t}\\ &= \lambda_2 e^{\lambda_1 t}e^{\lambda_2 t}-\lambda_1 e^{\lambda_1 t}e^{\lambda_2 t}\\ &= (\lambda_2-\lambda_1)e^{(\lambda_1+\lambda_2)t}\neq 0 \text{ since } \lambda_1\neq\lambda_2\end{aligned}$$

and $e^{(\lambda_1+\lambda_2)t}\neq 0$ for all t.

It follows that the general solution can be written as

$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

where c_1 and c_2 are arbitrary constants.

What if $\lambda_1 = \lambda_2 = \lambda$? In this case, we already know that $e^{\lambda t}$ is a solution. We claim that $te^{\lambda t}$ is also a solution and that $e^{\lambda t}$ and $te^{\lambda t}$ are linearly independent.

To verify the first claim, we plug $te^{\lambda t}$ into the equation:

$$a(te^{\lambda t})'' + b(te^{\lambda t})' + cte^{\lambda t} = a(e^{\lambda t} + \lambda te^{\lambda t})' + b(e^{\lambda t} + te^{\lambda t})$$

$$+ cte^{\lambda t} = (\lambda e^{\lambda t} + \lambda^2 te^{\lambda t} + \lambda e^{\lambda t}) + b(e^{\lambda t} + te^{\lambda t}) + cte^{\lambda t}$$

$$= t(a\lambda^2 + \lambda + c)e^{\lambda t} + (2a\lambda + b)e^{\lambda t} = 0.$$

The last equality holds because $a\lambda^2 + b\lambda + c = 0$ since λ is a root of the characteristic equation, whereas $2a\lambda + b = 0$ because the root is repeated (so $\lambda = -\frac{b}{2a}$).

To verify linear independence, we compute the Wronskian:

$$W(e^{\lambda t}, te^{\lambda t})(t) \quad e^{\lambda t}(te^{\lambda t})' - (e^{\lambda t})' te^{\lambda t}$$

$$= e^{\lambda t}(e^{\lambda t} + t\lambda e^{\lambda t}) - \lambda e^{\lambda t} te^{\lambda t}$$

$$= e^{2\lambda t} \neq 0$$ for all t, hence the two solutions are linearly independent.

We conclude that the general solution can be written as:

$$x(t) = c_1 e^{\lambda t} + c_2 t e^{\lambda t}$$

where c_1 and c_2 are arbitrary constants.

Remark. Students will probably wonder where $te^{\lambda t}$ came from, i.e., how we knew that we had to multiply by t. This comes from developing the theory of DE further, and we will show where it comes from when we study variation of parameters.

It remains to analyze what happens when the roots of the characteristic equation are complex, i.e, when

$$\lambda = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad \text{with} \quad b^2 - 4ac < 0.$$

In this case we can write $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$, where $\alpha = -\frac{b}{2a}$, $\beta = \frac{\sqrt{4ac - b^2}}{2a}$ and i is the imaginary number $i^2 = 1$. Note that $\alpha, \beta \in \mathbb{R}$.

The calculations previously done remain valid here and we have that $e^{\lambda_1 t} = e^{(\alpha + i\beta)t}$ and $e^{\lambda_2 t} = e^{(\alpha - i\beta)t}$ are solutions of the DE $ax'' + bx' + cx = 0$.

These solutions, however, are complex valued, and we would like to have real valued functions as solutions. To do so, we are going to use Euler's formula:

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

We'll prove this formula below. But first let us use it to obtain the desired real solutions.

We have, from Euler's formula:

$$e^{\lambda_1 t} = e^{(\alpha+i\beta)t} = e^{\alpha t + i\beta t} = e^{\alpha t} e^{i\beta t} = e^{\alpha t}(\cos(\beta t) + i\sin(\beta t))$$

$$e^{\lambda_2 t} = e^{(\alpha+i\beta)t} = e^{\alpha t - i\beta t} = e^{\alpha t} e^{-i\beta t} = e^{\alpha t}(\cos(-\beta t) + i\sin(-\beta t))$$

$$= e^{\alpha t}(\cos(\beta t) - i\sin(\beta t)).$$

Lemma. Let $z(t) = u(t) + i v(t)$ be a solution to the DE $a x'' + b x' + c x = 0$, where $a, b, c \in \mathbb{R}$ and $u(t)$ and $v(t)$ are real valued. Then $u(t)$ and $v(t)$ are also solutions.

Proof: We have: $0 = a z'' + b z' + c z = a(u+iv)'' + b(u+iv)' + c(u+iv) = a(u'' + iv'') + b(u' + iv') + c(u+iv)$
$= (a u'' + b u' + c u) + i(a v'' + b v' + c v)$. Since a complex number vanishes iff its real and imaginary parts vanish, we have $a u'' + b u' + c u = 0$ and $a v'' + b v' + c v = 0$. ☐

The lemma implies that $e^{\alpha t}\cos(\beta t)$ and $e^{\alpha t}\sin(\beta t)$ are solutions of the DE. Now let us check that they are linearly independent:

$$W\left(e^{\alpha t}\cos(\beta t), e^{\alpha t}\sin(\beta t)\right)(t) = e^{\alpha t}\cos(\beta t)\left(e^{\alpha t}\sin(\beta t)\right)' - \left(e^{\alpha t}\cos(\beta t)\right)' e^{\alpha t}\sin(\beta t)$$

$$= e^{\alpha t}\cos(\beta t)\left(\alpha e^{\alpha t}\sin(\beta t) + \beta e^{\alpha t}\cos(\beta t)\right) - \left(\alpha e^{\alpha t}\cos(\beta t) - \beta e^{\alpha t}\sin(\beta t)\right)e^{\alpha t}\sin(\beta t)$$

$$= \left(e^{\alpha t}\right)^2\left(\alpha\cos(\beta t)\sin(\beta t) + \beta\cos^2(\beta t) - \alpha\cos(\beta t)\sin(\beta t) + \beta\sin^2(\beta t)\right)$$

$$= \beta\left(e^{\alpha t}\right)^2\left(\cos^2(\beta t) + \sin^2(\beta t)\right) = \beta\left(e^{\alpha t}\right)^2.$$

This expression is never zero because $\beta \neq 0$ (if $\beta = 0$ then λ_1 and λ_2 would not be complex numbers, but we are analyzing the case where they are).

We conclude that the general solution can be written as

$$x(t) = c_1 e^{\alpha t}\cos(\beta t) + c_2 e^{\alpha t}\sin(\beta t)$$

where c_1 and c_2 are arbitrary constants.

Summary of solutions to $ax'' + bx' + cx = 0$.

Consider $ax'' + bx' + cx = 0$, $a, b, c \in \mathbb{R}$, $a \neq 0$. Let λ_1 and λ_2 be the two roots of the characteristic equation

$$a\lambda^2 + b\lambda + c = 0.$$

- If $\lambda_1 \neq \lambda_2$ are real, then the general solution is
$$x(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$
- If $\lambda_1 = \lambda_2 = \lambda$, then the general solution is
$$x(t) = c_1 e^{\lambda t} + c_2 t e^{\lambda t}.$$
- If λ_1 and λ_2 are complex, then we can write $\lambda_1 = \alpha + i\beta$, $\lambda_2 = \alpha - i\beta$, $\alpha, \beta \in \mathbb{R}$, and the general solution is
$$x(t) = c_1 e^{\alpha t} \cos(\beta t) + c_2 e^{\alpha t} \sin(\beta t).$$

Above, c_1 and c_2 are arbitrary constants.

Proof of Euler's formula. Recall from calculus that

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$ Thus $$e^{i\theta} = \sum_{n=0}^{\infty} \frac{(i\theta)^n}{n!}.$$ We separate the sum into even and odd n's:

$$e^{i\theta} = \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} \frac{(i\theta)}{n!} + \sum_{\substack{n=0 \\ n \text{ odd}}}^{\infty} \frac{(i\theta)^n}{n!} = \sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(i\theta)^{2k+1}}{(2k+1)!}.$$

Notice that $i^0 = 1,\ i^1 = i,\ i^2 = -1,\ i^3 = -i,\ i^4 = 1$
$i^5 = i,\ i^6 = -1,\ i^7 = -i,\ i^8 = 1$, so this pattern repeats every four powers. Thus:

$$\sum_{k=0}^{\infty} \frac{(i\theta)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(i\theta)^{2k+1}}{(k+1)!} = \left(\frac{i^0\theta^0}{0!} + \frac{i^2\theta^2}{2!} + \frac{i^4\theta^4}{4!} + \frac{i^6\theta^6}{6!} + \dots \right)$$

$$\left(\frac{i^1\theta^1}{1!} + \frac{i^3\theta^3}{3!} + \frac{i^5\theta^5}{5!} + \frac{i^7\theta^7}{7!} + \dots \right) = \left(\frac{\theta^0}{0!} - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \frac{\theta^6}{6!} + \dots \right)$$

$$+ \left(i\frac{\theta^1}{1!} - i\frac{\theta^3}{3!} + i\frac{\theta^5}{5!} - i\frac{\theta^7}{7!} + \dots \right) = \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{2k+1}}{(2k+1)!}.$$

Recalling from calculus that $$\cos\theta = \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{2k}}{(2k)!}$$ and $$\sin\theta = \sum_{k=0}^{\infty} (-1)^k \frac{\theta^{2k+1}}{(2k+1)!},$$ we have the result.

Linear second order non-homogeneous equations

Consider the equation

$$ax'' + bx' + cx = f(t)$$

where a, b, c are constants, $a \neq 0$, and $f(t)$ is a given function called the non-homogeneous or inhomogeneous term. Let us first proceed by examples.

Ex: Find a solution to $x'' + 3x' + 4x = 3t + 2$.

The given function $f(t) = 3t + 2$ is a polynomial of degree one. We expect that $x(t)$ will be a polynomial as well (we wouldn't get a polynomial by differentiating, say, an exponential). Thus we seek a solution of the form $x(t) = At + B$, where A and B are constants to be determined. Note that we are trying $x(t)$ a polynomial of degree one because $f(t)$ is a polynomial of degree one. Plugging in:

$$(At+B)'' + 3(At+B)' + 4(At+B) = 3t + 2$$

$$0 + 3A + 4At + 4B = 3t + 2$$

$$4At + (3A + 4B) = 3t + 2.$$

Two polynomials are equal iff the corresponding coefficients of the same powers are equal.

So we must have $4A = 3$ and $3A + 4B = 2$, so that
$A = \frac{3}{4}$, $4B = 2 - 3A = 2 - 3 \cdot \frac{3}{4} = -\frac{1}{4} \Rightarrow B = -\frac{1}{16}$.

Therefore, $x(t) = \frac{3}{4}t - \frac{1}{16}$ is a solution.

<u>EX</u>: Find a solution to $x'' - 4x = 2e^{3t}$.

Here the inhomogeneous term $f(t) = 2e^{3t}$ is an exponential. Thus we expect $x(t)$ to be an exponential too (we wouldn't get an exponential differentiating, say, a trigonometric function). Put:
$x(t) = Ae^{t}$, where A is to be found. Plugging in:

$$(Ae^{3t})'' - 4Ae^{3t} = 2e^{3t}$$

$$9Ae^{3t} - 4Ae^{3t} = 2e^{3t} \Rightarrow 5Ae^{3t} = 2e^{3t} \Rightarrow A = \frac{2}{5}$$

Hence $x(t) = \frac{2}{5}e^{3t}$ is a solution.

<u>EX</u>: Find a solution to $3x'' + x' - 2x = 2\cos t$

Here $f(t) = 2\cos t$, so we might try $x(t) = A\cos t$. However, when we plug this in we will obtain some $\sin t$ term, and there is no $\sin t$ on the RHS to compare with. We see, thus, that we should try $x(t) = A\cos t + B\sin t$

Then

$$3(A\cos t + B\sin t)'' + (A\cos t + B\sin t)' - 2(A\cos t + B\sin t) = 2\cos t$$

$$3(-A\cos t - B\sin t) + (-A\sin t + B\cos t) - 2(A\cos t + B\sin t) = 2\cos t$$

$$(-5A + B)\cos + (-A - 5B)\sin t = 2\cos t.$$

Thus, for the equality to hold, we must have

$$-5A + B = 2 \quad \text{and} \quad -A - 5B = 0.$$

This is a system of two equations for the two unknowns A and B. Solving it we find $A = -\frac{5}{13}$, $B = \frac{1}{13}$

Thus $x(t) = -\frac{5}{13}\cos t + \frac{1}{13}\sin t$ is a solution.

Unfortunately, things will not always be this simple, as the next example illustrates.

<u>EX</u>: Find a solution to $x'' - 4x = 2e^{2t}$.

We try $x(t) = Ae^{2t}$. Plugging in:

$$(Ae^{2t})'' - 4Ae^{2t} = 2e^{2t}$$

$$4Ae^{2t} - Ae^{2t} = 2e^{2t}$$

$$0 = 2e^{2t} \quad ???$$

We see that our method did not work in this case. The problem is that e^{2t} is a solution to the equation $x'' - 4x = 0$ (the characteristic equation is $\lambda^2 - 4 = 0$, $\lambda = \pm 2$), and so is any multiple of e^{2t}. Therefore, if the inhomogeneous term happens to be a function that solves the same equation when $f(t) = 0$, then the LHS will always give zero when we plug in, and this idea will not work. We see that to solve $ax'' + bx' + cx = f(t)$ we also need to understand $ax'' + bx' + cx = 0$.

Def. Given $ax'' + bx' + cx = f(t)$, the equation $ax'' + bx' + cx = 0$ is called the associated homogeneous equation. The general solution to the associated homogeneous equation will be denoted x_h.

Observe that if z solves $ax'' + bx' + cx = f$, so does the function $x = x_h + z$ because $a(x_h + z)'' + b(x_h + z)' + c(x_h + z)$

$$= \underbrace{a x_h'' + b x_h' + c x_h}_{=0} + \underbrace{a z'' + b z' + c z}_{=f} = f.$$

It follows that there are two "types" of solutions to $ax'' + bx' + cx = f$: those containing arbitrary constants (because x_h contains arbitrary constants) and those without arbitrary constants (such as the solutions we found in the previous examples).

Def. A solution to $ax'' + bx' + cx = f$ that does not contain arbitrary constants is called a particular solution. Particular solutions will be denoted by x_p.

EX: Let's go back to $x'' - 4x = 2e^{2t}$ and try to find a particular solution. We saw that if we put $x_p(t) = Ae^{2t}$ then it will not work. Let us see that $x(t) = Ate^{2t}$ works:

$$(Ate^{2t})'' - Ate^{2t} = A(2te^{2t} + e^{2t})' - 4Ate^{2t}$$

$$= A(4te^{2t} + 2e^{2t} + 2e^{2t}) - 4Ae^{2t}t = 4Ae^{2t} = 2e^{2t}$$

So we concl. that $A = 1/2$ and $x_p(t) = \frac{1}{2}e^{2t}$.

The idea of multiplying by t can be understood as follows.

We want to satisfying $ax'' + bx' + cx = f$ and we expect x_p to be of similar type as f (since derivatives of polynomials give polynomials, of exponentials give exponentials, etc.). But if f is or contains a term that solves the associated homogeneous equation, this cannot work because it will give a zero on the LHS. How can we find x_p containing f in such a way that after we plug it into the equation, a term with f remains on the LHS? The answer is the product rule, since it produces extra terms containing f. Put $x_p(t) = v(t)\tilde{f}(t)$, where $\tilde{f}$ has the same form of f but contains constants to be determined as in the examples (so $\tilde{f}(t) = Ae^{kt}$ if $f(t)$ is a multiple of e^{kt} and so on), and v is an undetermined function. Then:

$$x_p' = v'\tilde{f} + v\tilde{f}', \quad x_p'' = v''\tilde{f} + 2v'\tilde{f}' + v\tilde{f}''.$$

Then $ax_p'' + bx' + cx_p = a(v''\tilde{f} + 2v'\tilde{f}' + v\tilde{f}'')$
$+ b(v'\tilde{f} + v\tilde{f}') + cv\tilde{f} = \tilde{f}(av'' + bv') + 2av'\tilde{f}'$
$+ v(a\tilde{f}'' + b\tilde{f} + c\tilde{f})$.

Because $\tilde{f}$ contains x_h, the term $a\tilde{f}'' + b\tilde{f}' + c\tilde{f}$ will produce zeros. For simplicity, let us assume we are treating the case when $\tilde{f}$ is proportional to x_h. Then $a\tilde{f}'' + b\tilde{f}' + c\tilde{f} = 0$. Next, recall that $\tilde{f}$ is like f, and we are treating functions that "repeat themselves" after differentiation, like exponentials, polynomials, and sine or cosine (this method will not work for functions that do not repeat themselves in this way). Thus, for the sake of reasoning, we can replace $\tilde{f}'$ by $\tilde{f}$ in the term $2a\sigma'\tilde{f}'$. Thus

$a x_p'' + b x_p' \; c x_p = \tilde{f}(a\sigma'' + b\sigma' + 2a\sigma')$. We want this to be equal to f so: $\tilde{f}(a\sigma'' + b\sigma' + 2a\sigma') = f$. If the term in parenthesis is a constant, then we have $(\text{constant}) \cdot \tilde{f} = f$, and we can solve for the undetermined constants in $\tilde{f}$. The simplest way to accomplish this is to put $\sigma(t) = t$ so $a\sigma'' + b\sigma' + 2a\sigma' = b + 2a$ and $x_p(t) = t\tilde{f}(t)$.

The method outlined above is called the <u>method of undetermined coefficients</u>, summarized as follows: given $ax'' + bx' + cx = f(t)$, a, b, c constants, $a \neq 0$, we seek for a particular solution $x_p(t)$ of the form (below, $m \geq 0$ is an integer, $b_m, \dots, b_0, a_m, \dots, a_0, a, b, r$, and k are constants):

$f(t)$	$x_p(t)$
$b_m t^m + b_{m-1} t^{m-1} + \dots + b_1 t + b_0$	$t^s (B_m t^m + B_{m-1} t^{m-1} + \dots + B_1 t + B_0)$
$a\cos(kt) + b\sin(kt)$	$t^s (A\cos(kt) + B\sin(kt))$
$e^{rt}(a\cos(kt) + b\sin(kt))$	$t^s e^{rt}(A\cos(kt) + B\sin(kt))$
$e^{rt}(b_m t^m + b_{m-1} t^{m-1} + \dots + b_1 t + b_0)$	$t^s e^{rt}(B_m t^m + B_{m-1} t^{m-1} + \dots + B_1 t + B_0)$
$(b_m t^m + b_{m-1} t^{m-1} + \dots + b_0)\cos(kt) + (a_m t^m + a_{m-1} t^{m-1} + \dots + a_0)\sin(kt)$	$t^s (B_m t^m + B_{m-1} t^{m-1} + \dots + B_0)\cos(kt) + t^s (A_m t^m + A_{m-1} t^{m-1} + \dots + A_0)\sin(kt)$
$e^{rt}(b_m t^m + b_{m-1} t^{m-1} + \dots + b_0)\cos(kt) + e^{rt}(a_m t^m + a_{m-1} t^{m-1} + \dots + a_0)\sin(kt)$	$t^s e^{rt}(B_m t^m + B_{m-1} t^{m-1} + \dots + B_0)\cos(kt) + t^s e^{rt}(A_m t^m + A_{m-1} t^{m-1} + \dots + A_0)\sin(kt)$

where s is the smallest non-negative integer such that no term in x_p duplicates a term in x_h.

Ex: Find the form of x_p for

$$x'' + 6x' + 13x = e^{-3t}\cos(2t)$$

The characteristic eq is $\lambda^2 + 6\lambda + 13 = 0 \Rightarrow \lambda = -3 \pm 2i$. Thus $x_h(t) = c_1 e^{-3t}\cos(2t) + c_2 e^{-3t}\sin(2t)$. We see that we cannot try $x_p(t) = Ae^{-3t}\cos(2t) + Be^{-3t}\sin(2t)$ the first term duplicates a term in x_h. We thus multiply by t:

$$x_p(t) = t\left(Ae^{-3t}\cos(2t) + Be^{-3t}\sin(2t)\right).$$

Ex: Find the form of x_p for

$$x'' - 2x' + x = e^t$$

We have $\lambda^2 - 2\lambda + 1 = (\lambda - 1)^2 = 0 \Rightarrow \lambda = 1$ (repeated). Then $x_h(t) = c_1 e^t + c_2 t e^t$. If we put $x_p(t) = Ae^t$, this duplicates the first term in x_h. Multiplying by t gives $x_p(t) = Ate^t$, but this duplicates the second term in x_h, so we multiply by t again:

$$x_p(t) = At^2e^t.$$

The next theorem is known as the superposition principle:

Theo. If x_1 is a solution to $ax'' + bx' + cx = f_1$ and x_2 a solution to $ax'' + bx' + cx = f_2$, then the function $x = c_1 x_1 + c_2 x_2$ is a solution to the DE $ax'' + bx' + cx = c_1 f_1 + c_2 f_2$, where c_1 and c_2 are constants.

proof: Plugging in:

$$ax'' + bx' + cx = a(c_1 x_1 + c_2 x_2)'' + b(c_1 x_1 + c_2 x_2)' + c(c_1 x_1 + c_2 x_2)$$

$$= \underbrace{a x_1'' + b x_1' + c x_1}_{= c_1 f_1} + \underbrace{a x_2'' + b x_2' + c x_2}_{= c_2 f_2} = c_1 f_1 + c_2 f_2. \quad \square$$

It follows that if the inhomogeneous term is of the form $f = f_1 + f_2$, where the method of undetermined coefficients can be applied to f_1 and f_2, then we can find x_p by determining x_{p1} and x_{p2}, the particular solutions for the equation with inhomogeneous terms f_1 and f_2, respectively, and setting $x_p = x_{p1} + x_{p2}$.

Theo. Consider $ax'' + bx' + cx = f$, where a, b, c are constants and $a \neq 0$ Suppose that x_p is a particular solution to the DE in an interval I containing t_0, and let Ξ_0 and Ξ_1 be two given numbers. Then there exists a unique solution in I to the DE satisfying the initial conditions $x(t_0) = \Xi_0$ and $x(t_0) = \Xi_1$.

proof. By the superposition principle, $x = x_h + x_p$ solves the DE. Recall that $x_h = c_1 x_1 + c_2 x_2$, where x_1 and x_2 are linearly independent solutions to the associated homogeneous equation and c_1 and c_2 are constants. Then we need to solve

$$x(t_0) = c_1 x_1(t_0) + c_2 x_2(t_0) + x_p(t_0) = \Xi_0$$
$$x'(t_0) = c_1 x_1'(t_0) + c_2 x_2'(t_0) + x_p'(t_0) = \Xi_1$$

for c_1 and c_2. We find

$$c_1 = \frac{(\Xi_0 - x_p'(t_0))\, x_2'(t_0) - (\Xi_1 - x_p'(t_0))\, x_2(t_0)}{x_1(t_0)\, x_2'(t_0) - x_1'(t_0)\, x_2(t_0)}$$

$$c_2 = \frac{(\Xi_1 - x_p'(t_0))\, x_1(t_0) - (\Xi_0 - x_p'(t_0))\, x_1'(t_0)}{x_1(t_0)\, x_2'(t_0) - x_1'(t_0)\, x_2(t_0)}$$

The denominators in these expressions are non-zero because x_1 and x_2 are linearly independent (so their Wronskian is not zero).

To check uniqueness, suppose z is another solution to the IVP. Put $w = c_1 x_1 + c_2 x_2 + x_p - z$. Then, plugging in, we see that w solves $aw'' + bw' + cw = 0$ with $w(t_0) = 0$, $w'(t_0) = 0$. But we have seen that this IVP, where the IC and the inhomogeneous term are all zero, admits only the zero solution. Thus $w = 0$ and $z = c_1 x_1 + c_2 x_2 + x$.

$\boxed{\text{//}}$

<u>Def.</u> We call a solution $x = x_h + x_p$ the <u>general solution</u> to $ax'' + bx' + cx = f$.

Linear second order non-homogeneous equations: the method of variation of parameters.

The method of undetermined coefficients will not work if the nonhomogeneous term is not of the form listed on the table that summirized the method. This is because the method of undetermined coefficients is based on the property that derivatives of the nonhomogeneous term repeat themselves. The method we will present now, called variation of parameters, deals with more general nonhomogeneous terms. (We will see later that this method applies when a, b, c are not constants, but we will take them constants for now.)

Consider $ax'' + bx' + cx = f$ and let x_1 and x_2 be two linearly independent solutions to the associated homogeneous equation. We will seek a solution of the form:

$$x_p(t) = v_1(t)\,x_1(t) + v_2(t)\,x_2(t)$$

where v_1 and v_2 are functions to be determined.

Compute: $x_p' = v_1' x_1 + v_2' x_2 + v_1 x_1' + v_2 x_2'$.

Next, we reason as follows. Since v_1 and v_2 are two functions to be determined, we expect to have two equations. One equation has to come from $ax'' + bx' + cx = f$, since we want x_p to be a solution. What about the second equation? Because we will plug x_p into $ax'' + bx' + cx = f$, we will obtain another DE involving v_1 and v_2 that is at least as complicated as the equation we are trying to solve, unless we impose some condition that simplifies it. We therefore require:

$$v_1' x_1 + v_2' x_2 = 0$$

which gives our second equation. Thus x_p' becomes:

$$x_p' = v_1 x_1' + v_2 x_2' .$$

Continuing, $x_p'' = v_1' x_1' + v_2' x_2' + v_1 x_1'' + v_2 x_2''$. Thus

$$a x_p'' + b x_p' + c x_p = a(v_1' x_1' + v_2' x_2' + v_1 x_1'' + v_2 x_2'') + b(v_1 x_1' + v_2 x_2') + c(v_1 x_1 + c_2 x_2) = v_1 \underbrace{(a x_1'' + b x_1' + c x_1)}_{=0} + v_2 \underbrace{(a x_2'' + b x_2' + c x_2)}_{=0} + a(v_1' x_1' + v_2' x_2')$$

$$= a(v_1' x_1' + v_2' x_2') = f.$$

Therefore, we have two equations:

$$x_1 v_1' + x_2 v_2' = 0$$

$$x_1' v_1' + x_2' v_2' = \frac{f}{a}$$

This is an algebraic system for v_1' and v_2'. Solving it, we find:

$$v_1' = \frac{-f x_2}{a(x_1 x_2' - x_1' x_2)}, \quad v_2' = \frac{f x_1}{a(x_1 x_2' - x_1' x_2)}.$$

The denominators in these expressions are not zero because x_1 and x_2 are linearly independent. Integrating:

$$v_1(t) = -\frac{1}{a}\int \frac{f(t)\, x_2(t)}{W(x_1, x_2)(t)}\, dt, \quad v_2(t) = \frac{1}{a}\int \frac{f(t)\, x_1(t)}{W(x_1, x_2)(t)}\, dt.$$

We do not add constants to these integrals because x_p does not contain arbitrary constants. Thus, recalling that $x_p = v_1 x_1 + v_2 x$, we find:

$$\boxed{x_p(t) = -\frac{x_1(t)}{a}\int \frac{f(t)\, x_2(t)}{W(x_1, x_2)(t)}\, dt + \frac{x_2(t)}{a}\int \frac{f(t)\, x_1(t)}{W(x_1, x_2)(t)}\, dt}$$

__EX:__ Find x_p for $x'' + 4x = \tan t$.

Note that we cannot apply the method of undetermined coefficients here. To find x_p, we first solve the associated homogeneous equation. The characteristic equation is $\lambda^2 + 4 = 0$, $\lambda = \pm 2i$. Thus $x_1(t) = \cos(2t)$ and $x_2(t) = \sin(2t)$ are two linearly independent solutions. The Wronskian is

$$W(\cos 2t, \sin 2t)(t) = \cos(2t)(\sin(2t))' - (\cos(2t))' \sin(2t)$$
$$= 2\cos^2(2t) + 2\sin^2(2t) = 2.$$

Then

$$x_p(t) = -\cos(2t)\underbrace{\int \frac{\tan t \sin(2t)}{2}\,dt}_{= \frac{t}{2} - \frac{1}{4}\sin(2t)} + \sin(2t)\underbrace{\int \frac{\tan t \cos(2t)}{2}\,dt}_{= -\frac{1}{4}\cos(2t) + \frac{1}{2}\ln|\cos t|}$$

$$x_p(t)) = \frac{1}{2}\left(\frac{1}{2}\sin(2 - t\right)\cos(2t) + \frac{1}{2}\left(\ln|\cos t| - \frac{1}{2}\cos(2t)\right)\sin(2t).$$

__EX:__ Find x_p for $x'' - 2x' + x = \dfrac{e^t}{t}$.

Note that we cannot use the method of undetermined coefficients here.

The characteristic equation is $\lambda^2 - 2\lambda + 1 = 0$, $\lambda = 1$ (repeated). Then $x(t) = e^t$ and $x_2(t) = te^t$ are two linearly independent solutions to the associated homogeneous equation.

$$W(e^t, te^t) = e^t(te^t)' - (e^t)' te^t = e^t(e^t + te^t) - e^t e^t = e^{2t}.$$

$$\int \frac{f(t)\, x_2(t)}{W(x_1, x_2)(t)}\, dt = \int \frac{e^t}{t} \frac{te^t}{e^{2t}}\, dt = t$$

$$\int \frac{f(t)\, x_1(t)}{W(x_1, x_2)(t)}\, dt = \int \frac{e^t}{t} \frac{e^t}{e^{2t}}\, dt = \int \frac{dt}{t} = \ln|t|$$

Since the $f(t)$ is not defined for $t = 0$, we need to work with $t > 0$ or $t < 0$. We consider $t > 0$ so $\ln|t| = \ln t$. Then:

$$x_p(t) = -te^t + te^t \ln t.$$

Remark. We made no restriction on the form of $f(t)$. In particular, t method of variation of parameters can also be applied to equations where $f(t)$ has a form appropriate for the use of undetermined coefficients.

Remark. Inspecting the derivation of the formula for x_p using the method of variation of parameters, we notice that we need not to assume a, b, and c to be constants. If they are not, the only difference is that in the expression for x_p, the term $\frac{1}{a}$ has to be inside the integral.

Second order linear equations with variable coefficients

So far, we studied $ax'' + bx' + cx = f$ under the assumption that a, b, c are constants. Now we will study $a_2(t)\, x''(t) + a_1(t)\, x'(t) + a_0(t)\, x(t) = f(t)$, i.e., the coefficients can be functions of t. We will assume that $a_2(t) \neq 0$ so that, dividing by $a_2(t)$ and relabelling the coefficients and the inhomogeneous term, we can write the equation as $x''(t) + p(t)\, x'(t) + q(t)\, x(t) = g(t)$. To be consistent with our previous notation, we will call the inhomogeneous term $f(t)$ in this case as well. Thus, the equation we will study is

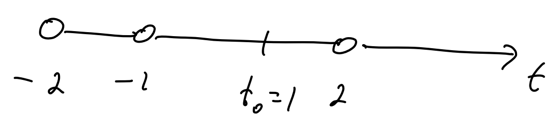
$$x'' + p(t)\, x' + q(t)\, x = f(t).$$

Theo. Let $p(t)$, $q(t)$, and $f(t)$ be continuous functions on the interval (a,b) and $t_0 \in (a,b)$. Given any numbers Ξ_0 and Ξ_1, there exists a unique solution $x(t)$ defined on (a,b) satisfying:

$$\begin{cases} x'' + p(t)\, x' + q(t)\, x = f(t) \\ x(t_0) = \Xi_0 \\ x'(t_0) = \Xi_1 \end{cases}.$$

Ex: Consider $(t^2-4)x'' + x' + x = \frac{1}{t+1}$, $x(1) = 0$, $x'(1) = 1$. What is the maximal interval (a,b) where the previous theorem guarantees the existence of a unique solution?

After dividing by t^2-4, we have $p(t) = q(t) = \frac{1}{t^2-4}$, which are continuous except at $t = \pm 2$, and $f(t) = \frac{1}{(t^2-4)(t+1)}$, which is continuous except for $t = \pm 2$, $t = -1$. Since $t_0 = 1$, the largest interval containing this point is $(a,b) = (-1, 2)$.



As in the constant coefficients case, we will call the equation $x'' + p(t)x + q(t)x = 0$ the <u>associated homogeneous equation</u>. It can be showed that this equation admits two linearly independent solutions x_1 and x_2 (if p and q are continuous). Then, $x_h = c_1 x_1 + c_2 x$, where c_1 and c_2 are arbitrary constants, is also a solution, called the <u>general solution</u> to the DE $x'' + p(t)x' + q(t)x = 0$.

A solution to $x'' + p(t)x' + q(t)x = f(t)$ that does not contain arbitrary constants will be called a particular solution, denoted x_p. As in the constant coefficients case, any solution to the DE can be written as $x = x_h + x_p$ (provided that $p(t)$ and $q(t)$ are continuous).

Many of the theorems for equations with constant coefficients generalize to the case studied here, with the important difference that now statements will in general not hold on $(-\infty, \infty)$ but on an interval (a, b) where $p(t)$ and $q(t)$ are continuous.

Lemma. Let $p(t)$ and $q(t)$ be continuous functions on an interval I. Let $x_1(t)$ and $x_2(t)$ be two solutions of $x'' + p(t)x' + q(t)x = 0$ on I. If the Wronskian $W(x_1, x_2)(t) = x_1(t)x_2'(t) - x_1'(t)x_2(t)$ is zero at some point on I, then it vanishes identically and x_1 and x_2 are linearly dependent. If $W(x_1, x_2)(t)$ is non-zero at some point on I, then it is never zero and the solutions are linearly independent on I.

Theorem. Let $p(t)$, $q(t)$, and $f(t)$ be continuous functions on an interval I and $x_1(t)$ and $x_2(t)$ be two linearly independent solutions to $x'' + p(t)x' + q(t)x = 0$ on I. Let $x_p(t)$ be a particular solution to $x'' + p(t)x' + q(t)x = f(t)$. Then, given $t_0 \in I$ and two real numbers $\mathbb{X}_0, \mathbb{X}_1$, there exist unique constants c_1 and c_2 such that $x = c_1 x_1 + c_2 x_2 + x_p$ satisfies $x'' + p(t)x + q(t)x = f(t)$ with initial conditions $x(t_0) = \mathbb{X}_0$ and $x'(t_0) = \mathbb{X}_1$.

The superposition principle also holds for equations with variable coefficients.

If we go back to the method of variation of parameters and look at how the formula for x_p was derived, we will see that nowhere have we used that the coefficients had to be constants. In other words, variation of parameters applies here as well, i.e., if x_1 and x_2 are two linearly independent solutions of the associated homogeneous equation, then a particular solution is given by

$$x_p(t) = -x_1(t)\int \frac{f(t)\,x_2(t)}{W(x_1,x_2)(t)}\,dt + x_2(t)\int \frac{f(t)\,x_1(t)}{W(x_1,x_2)(t)}\,dt$$

The formula for x_p involves x_1 and x_2. In the constant coefficient case we have a method for finding x_1 and x_2. Here, this might be difficult. However, the next theorem shows that if we know x_1, then we can always determine x_2:

Theo. Let $x_1(t)$ be a solution to $x'' + p(t)x' + q(t)x = 0$ on an interval I, where $p(t)$ and $q(t)$ are continuous functions. Assume that x_1 is not identically zero. Then

$$x_2(t) = x_1(t) \int \frac{e^{-\int p(t)\,dt}}{(x_1(t))^2}\,dt$$

is a second, linearly independent solution.

proof: We look for a solution of the form $x_2(t) = v(t)\,x_1(t)$. Plugging in:

$$x_2'' + p(t)x_2' + q(t)x_2 = (v x_1'' + 2v'x_1' + v''x_1) +$$
$$p(t)(v x_1' + v'x_1) + q(t) v x_1 = v\underbrace{(x_1'' + p(t)x' + q(t)x)}_{=0}$$
$$+ x_1 v'' + (2x_1' + p(t)x_1)v' = 0$$

Set $v' = w$. Then the equation becomes:

$$x_1 w' + (2x_1' + p(t)x_1) w = 0$$

which is a separable equation for w. We find

$$\frac{dw}{w} = -\frac{2x_1'}{x_1} - p(t)$$

Integrating: $\ln|w| = -2\ln|x_1| - \int p(t)\,dt$. Then:

$w = \dfrac{e^{-\int p(t)\,dt}}{x_1^2}$. When we removed the absolute value from w we picked the $+$ sign (this suffices since if w is a solution so is $-w$). Integrating again we find $v(t)$, giving the desired formula.

By construction, $x_2(t)$ is a solution. Let's check that it is linearly independent.

$$W(x_1, x_2)(t) = x_1 x_2' - x_1' x_2 = x_1 (v x_1)' - x_1'(v x_1)$$

$$= x_1(x_1 v' + x_1' v) - x_1' v x_1 = x_1^2 v' = x_1^2 \frac{e^{-\int p(t)\,dt}}{x_1^2}$$

$$= e^{-\int p(t)\,dt} \neq 0.$$

$\boxed{\checkmark}$

EX: Knowing that $\cos t$ is a solution to

$$\sin t\, x'' - 2\cos t\, x' - \sin t\, x = 0, \quad 0 < t < \pi,$$

find a second linearly independent solution.

Here, $p(t) = -2\frac{\cos t}{\sin t} = -2\cot t$.

$$x_2(t) = \cos t \int \frac{1}{\cos^2 t}\, e^{2\underbrace{\int \cot t\, dt}_{= \ln|\sin t| = \ln(\sin t),\ 0<t<\pi}}\, dt$$

$$= \cos t \int \frac{\sin^2 t}{\cos^2 t}\, dt = \cos t\,(\tan t - t).$$

Remark. The formulas we derived above (variation of parameters and second linearly independent solution) assume the equation to be written as $x'' + p(t)x' + q(t)x = f(t)$, i.e., the coefficient of x'' is one. If this is not the case, we have to divide by the coefficient of x'' before applying the formulas, as in the previous examples.

Remark. Recall that in the constant coefficient case, when $\lambda_1 = \lambda_2 = \lambda$, a second linearly independent solution was $t e^{\lambda t}$. We can use the previous theorem to give an alternative justification of this formula.

Cauchy - Euler equation.

The equation

$$a t^2 x'' + b t x' + c x = f(t)$$

where a, b, c are constants and $a \neq 0$, is called Cauchy - Euler equation (aka equidimensional equation).

We will consider the homogeneous Cauchy - Euler equation

$$a t^2 x'' + b t x' + x = 0, \quad t > 0.$$

Because the coefficients involve powers of t, it makes sense to look for a solution $x(t) = t^\lambda$, λ a constant. Thus:

$$a t^2 \lambda(\lambda-1) t^{\lambda-2} + b t \lambda t^{\lambda-1} + c t^\lambda = 0, \text{ or } (t \neq 0)$$

$$\boxed{a \lambda^2 + (b-a)\lambda + c = 0}$$

which is called the characteristic equation for the Cauchy - Euler equation. If λ is a root of the characteristic equation, by construction t^λ is a solution. Denote the roots of the characteristic equation by λ_1 and λ_2.

We need to distinguish the following cases.

Case 1. $\lambda_1 \neq \lambda_2$, λ_1, λ_2 real. Then t^{λ_1} and t^{λ_2} are two linearly independent solutions.

We already know that they are solutions. To verify linear independence:

$$W(t^{\lambda_1}, t^{\lambda_2})(t) = t^{\lambda_1}(t^{\lambda_2})' - (t^{\lambda_1})' t^{\lambda_2} = (\lambda_2 - \lambda_1) t^{\lambda_1 + \lambda_2 - 1} \neq 0$$

for $t \neq 0$.

Case 2. $\lambda_1 = \lambda_2 = \lambda$. Then t^{λ} and $t^{\lambda} \ln t$ are two linearly independent solutions.

We obtain $t^{\lambda} \ln t$ by applying our method to find a second linearly independent solution.

$$x_2(t) = t^{\lambda} \int \frac{e^{-\int p(t)\,dt}}{(t^{\lambda})^2}\,dt = t^{\lambda} \int \frac{e^{-\frac{b}{a}\int \frac{dt}{t}}}{t^{2\lambda}}\,dt$$

$$= t^{\lambda} \int \frac{t^{-\frac{b}{a}}}{t^{2\lambda}}\,t = t^{\lambda} \int t^{-\frac{b}{a}-2\lambda}\,dt.$$

In the case $\lambda_1 = \lambda_2 = \lambda$, the (repeated) roots are given by $\lambda = -\frac{(b-a)}{2a}$, so $-\frac{b}{a} - 2\lambda = -1$.

Thus $x_2(t) = t^\lambda \int t^{-1}\, dt = t^\lambda \ln t$.

Case 3. λ_1, λ complex, so that $\lambda_1 = \alpha + i\beta$ and $\lambda_2 = \alpha - i\beta$, $\alpha, \beta \in \mathbb{R}$. Then $t^\alpha \cos(\beta \ln t)$ and $t^\alpha \sin(\beta \ln t)$ are two linearly independent solutions.

We write $t^{\lambda_1} = t^{\alpha + i\beta} = t^\alpha t^{i\beta} = t^\alpha (e^{\ln t})^{i\beta} = t^\alpha e^{i\beta \ln t}$.
Euler's formula gives $t^{\lambda_1} = t^\alpha \cos(\beta \ln t) + i t^\alpha \sin(\beta \ln t)$. In the constant coefficients case we showed that if $z(t) = u(t) + i v(t)$ is a solution, u, v real, so are $u(t)$ and $v(t)$. The same proof works here, as we conclude that $t^\alpha \cos(\beta \ln t)$ and $t^\alpha \sin(\beta \ln t)$ are solution. We check that they are linearly independent:

$$W(t^\alpha \cos(\beta \ln t), t^\alpha \sin(\beta \ln t)) = t^\alpha \cos(\beta \ln t)(t^\alpha \sin(\beta \ln t))' - (t^\alpha \cos(\beta \ln t))' t^\alpha \sin(\beta \ln t)$$
$$= t^\alpha \cos(\beta \ln t)\left(\alpha t^{\alpha-1} \sin(\beta \ln t) + t^\alpha \frac{\beta}{t} \cos(\beta \ln t)\right) - \left(\alpha t^{\alpha-1} \cos(\beta \ln t) - t^\alpha \frac{\beta}{t} \sin(\beta \ln t)\right).$$
$$t^\alpha \sin(\beta \ln t) = t^{2\alpha-1} \beta \left(\cos^2(\beta \ln t) + \sin^2(\beta \ln t)\right) = \beta t^{2\alpha-1} \neq 0$$

since $t > 0$ and $\beta \neq 0$ (because otherwise the roots would not be complex).

Remark. Above, we solved the Cauchy-Euler equation for $t>0$. If we want to solve it for $t<0$, we proceed as follows. Set $\tau=-t$, so that $\tau>0$. Then

$$x(t)=x(-\tau), \quad x'=\frac{dx}{dt}=\frac{dx}{d\tau}\frac{d\tau}{dt}=-\frac{dx}{d\tau}, \text{ and}$$

$$x''=\frac{d^2x}{dt}=\frac{d}{d\tau}\left(\frac{dx}{dt}\right)\frac{d\tau}{dt}=\frac{d^2x}{d\tau^2}, \text{ and the equation becomes}$$

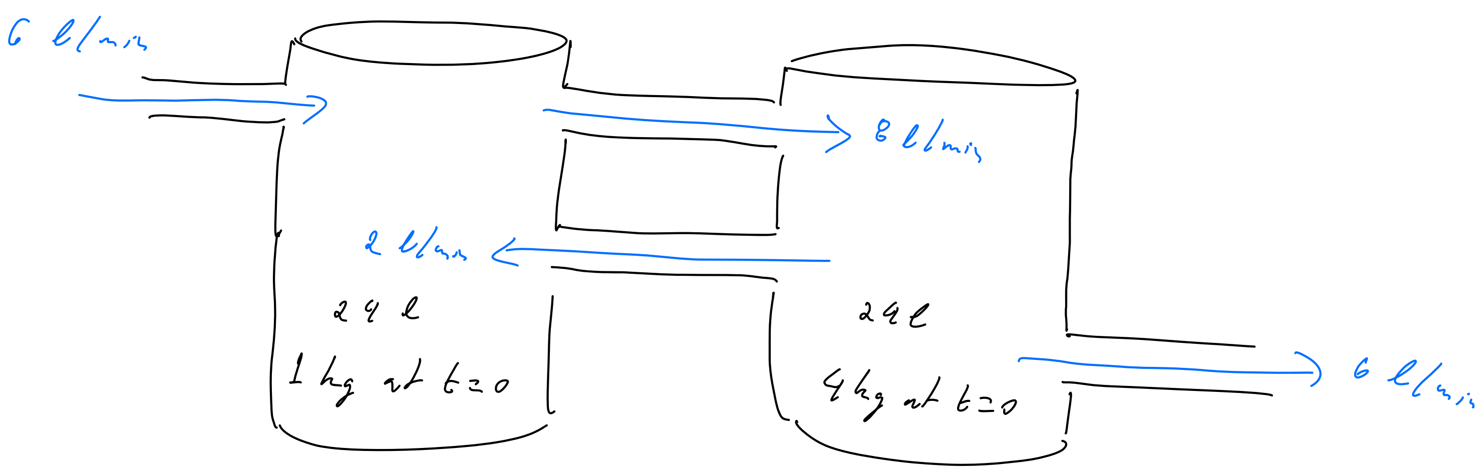
$$at^2x''+btx'+cx=(-\tau)^2\frac{d^2x}{d\tau^2}+b(-\tau)\left(-\frac{dx}{d\tau}\right)+cX=0, \text{ i.e.,}$$

$$a\tau^2\frac{d^2x}{d\tau^2}+b\tau\frac{dx}{d\tau}+cx=0, \quad \tau>0.$$

Now we can apply the above algorithm to find the solutions as functions of τ, and then replace $\tau=-t$ to obtain the result.

Interconnected tanks

Let us study the following situation.
Two tanks containing 24 l of brine are connected by two pipes. Free water flows into tank A at a rate 6 l/min, and fluid is drained out of tank B at the same rate. 8 l/min of fluid are pumped into tank B from tank A through one pipe and 2 l/min from tank B into tank A. Initially tank A contains 1 kg and tank B 4 kg of salt. Find the amount of salt in both tanks as a function of time.



Denote the amount of salt in tanks A and B by $x(t)$ and $y(t)$, respectively. We have

$$\frac{dx}{dt} = \text{in} - \text{out} \qquad \frac{dy}{dt} = \text{in} - \text{out}$$

$$\frac{dx}{dt} = 6\frac{\ell}{\min}\cdot 0\frac{kg}{\ell} + 2\frac{\ell}{\min}\cdot\frac{y}{24}\frac{kg}{\ell} - 8\frac{\ell}{\min}\cdot\frac{x}{24}\frac{kg}{\ell}$$

$$\frac{dy}{dt} = 8\frac{\ell}{\min}\cdot\frac{x}{24}\frac{kg}{\ell} - 2\frac{\ell}{\min}\cdot\frac{y}{24}\frac{kg}{\ell} - 6\frac{\ell}{\min}\cdot\frac{y}{24}\frac{kg}{\ell}$$

Thus

$$x' = -\frac{1}{3}x + \frac{1}{12}y$$

$$y' = \frac{1}{3}x - \frac{1}{3}y$$

This is a system of DE, i.e., we have two DE for two unknown functions. The second equation gives $x = 3y' + y$. Plugging this into the first equation: $(3y'+y)' = -\frac{1}{3}(3y'+y) + \frac{1}{12}y$, $3y'' + y' = -y' - \frac{1}{3}y + \frac{1}{12}y$, or $3y'' + 2y' + \frac{1}{4}y = 0$. This is a second order linear equation with constant coefficients. The characteristic equation is $3\lambda^2 + 2\lambda + \frac{1}{4} = 0$, so $\lambda = \left(-2 \pm \sqrt{4 - 4\cdot 3\cdot\frac{1}{4}}\right)/6$, $\lambda_1 = -\frac{1}{6}$, $\lambda_2 = -\frac{1}{2}$. Then $y = c_1 e^{-\frac{1}{2}t} + c_2 e^{-\frac{1}{6}t}$. We can now plug this into $x = 3y' + y$ to find $x = 3\left(c_1 e^{-\frac{1}{2}t} + c_2 e^{-\frac{1}{6}t}\right)' + c_1 e^{-\frac{1}{2}t} + c_2 e^{-\frac{1}{6}t}$ $= -\frac{3}{2}c_1 e^{-\frac{1}{2}t} - \frac{1}{2}c_2 e^{-\frac{1}{6}t} + c_1 e^{-\frac{1}{2}t} + c_2 e^{-\frac{1}{6}t}$, so

$$x = -\frac{1}{2}c_1 e^{-\frac{1}{2}t} + \frac{1}{2}c_2 e^{-\frac{1}{6}t}.$$

To find c_1 and c_2, we use the IC:

$$x(0) = -\frac{1}{2}c_1 + \frac{1}{2}c_2 = 1, \quad y(0) = c_1 + c_2 = 4.$$ This gives

$c_1 = 1$ and $c_2 = 3$. Hence

$$x(t) = -\frac{1}{2}e^{-\frac{1}{2}t} + \frac{3}{2}e^{-\frac{1}{6}t}$$

$$y(t) = e^{-\frac{1}{2}t} + 3e^{-\frac{1}{6}t}.$$

Many important problems involve systems of DE. We will develop a systematic method for studying systems.

The method of elimination for systems.

We will now study systems of DE, i.e., when we have more than one equation and more than one unknown.

We can think of the derivative $x' = \frac{dx}{dt}$ as the operator $\frac{d}{dt}$ acting on x. Let us denote the operator $\frac{d}{dt}$ by D. Similarly, $\frac{d^2x}{dt^2}$ can be thought as D acting on $\frac{dx}{dt} = Dx$, so $\frac{d^2x}{dt^2} = D(Dx) = D^2x$. We call D, D^2, etc., differential operators to emphasize that they involve derivatives.

We note that we can factor expressions in D in a similar way as we do for numerical expressions.

Ex: Show that $D^2 + D - 2$ is the same as $(D+2)(D-1)$.

For any twice differentiable function x:

$$(D+2)(D-1)x = (D+2)(Dx - x) = D^2x - Dx + 2Dx - 2x$$
$$= D^2x + Dx - 2x = (D^2 + D - 2)x.$$

The same is not true, however, if the coefficients are not constant.

Ex: Show that $(D + 4t)D \neq D(D+4t)x$

$(D+4t)Dx = D^2x + 4tDx$, $D(D+4t)x = D(Dx + 4tx)$

$= D^2x + D(4tx) = D^2x + 4x + 4tDx$ and

$$D^2x + 4tDx \neq D^2x + 4x + 4tDx.$$

Ex: Show that $(D+2)(D-t) \neq D^2 + (2-t)D - 2t$

$(D+2)(D-t)x = (D+2)(Dx - tx) = D^2x - D(tx) + 2Dx - 2tx$

$= D^2x - tDx - x + 2Dx - 2tx = D^2x + (2-t)Dx - (2+t)x$

$= \left(D^2 + (2-t)D - (2+t)\right)x \neq \left(D^2 + (2-t)D - 2t\right)x.$

The method we will present now is for systems with constant coefficients. Consider a 2×2 system of DE with constant coefficients of the form

$$\begin{cases} a_1 x' + a_2 x + a_3 y' + a_4 y = f_1(t) \\ a_5 x' + a_6 x + a_7 y' + a_8 y = f_2(t) \end{cases}$$

We can write it as

$$\begin{cases} (a_1 D + a_2)x + (a_3 D + a_4)y = f_1 \\ (a_5 D + a_6)x + (a_7 D + a_8)y = f_2 \end{cases}$$

Denote $L_1 = a_1 D + a_2$, $L_2 = a_3 D + a_4$, $L_3 = a_5 D + a_6$, $L_4 = a_7 D + a_8$.

Note that $L_1, \dots, L_4$ are differential operators and that they commute, i.e, $L_i L_j = L_j L_i$, $i,j = 1,2,3,4$ (this would not be true if the coefficients were not constant). Thus

$$\begin{cases} L_1 x + L_2 y = f_1 \\ L_3 x + L_4 y = f_2 \end{cases}$$

Applying L_4 to the first equation and L_2 to the second one, and using that the operators commute:

$$\begin{cases} L_1 L_4 x + L_2 L_4 y = L_4 f_1 \\ L_2 L_3 x + L_2 L_4 y = L_2 f_2 \end{cases}$$

Subtracting, gives

$$L_1 L_4 x - L_2 L_3 x = L_4 f_1 - L_2 f_2$$

Similarly, applying L_3 to the first equation, L_1 to the second, and subtracting, $L_1 L_4 y - L_2 L_3 y = L_1 f_2 - L_3 f_1$. Let $g_1 = L_4 f_1 - L_2 f_2$ and $g_2 = L_1 f_2 - L_3 f_1$ (g_1 and g_2 are known functions). Then

$$\begin{cases} (L_1 L_4 - L_2 L_3)\, x = g_1 \\ (L_1 L_4 - L_2 L_3)\, y = g_2 \end{cases}$$

We obtain two separate equations for x and y only. These are differential equations with constant coefficients that can be solved with methods previously learned. Moreover, the associated homogeneous equation is the same for x and y

<u>EX</u>: Solve
$$x' - 3x + 4y = 1$$
$$y' - 4x + 7y = 10t$$

Write
$$\begin{cases} (D-3)\,x + 4y = 1 \\ -4x + (D+7)\,y = 10t \end{cases}$$

$L_1 = D - 3 \quad L_2 = 4$

$L_3 = -4 \quad L_4 = D + 7$

Then $L_1 L_4 - L_2 L_3 = (D-3)(D+7) - 4 \cdot (-4) = D^2 + 4D - 5$.

Characteristic equation: $\lambda^2 + 4\lambda - 5 = (\lambda - 1)(\lambda + 5) = 0 \quad \lambda_1 = -5, \ \lambda_2 = 1$

There are two possible ways we can proceed now.

<u>Method 1</u> Solve x and y separately.

First we solve $(L_1 L_4 - L_2 L_3) y = g_2$

$y_1 = e^{-5t}$, $y_2 = e^{t}$, $g_2 = L_1 f_2 - L_3 f_1 = (D-3)(10t) - 4 \cdot 1$
$= 10 - 30t - 4 = 6 - 30t$. We seek $y_p = At + B$. Applying the method of undetermined coefficients gives $y_p = 6t + 2$. Thus

$$y = c_1 e^{-5t} + c_2 e^{t} + 6t + 2$$

Next, we find x solving $(L_1 L_4 - L_2 L_3) x = g_1$.

We already know that $x_1 = e^{-5t}$, $x_2 = e^{t}$ (recall that the associated homogeneous equation is the same). $g_1 = L_4 f_1 - L_2 f_2 = (D+7) 1 - 4 \cdot 10t = 7 - 40t$.

We seek $x_p = At + B$. Applying the method of undetermined coefficients gives $x_p = 8t + 5$. Thus

$$x = k_1 e^{-5t} + k_2 e^{t} + 8t + 5$$

We are not done yet. We obtain four constants, c_1, c_2, k_1, and k_2. But we should have only two arbitrary constants because the system we are trying to solve involves two equations of first order (giving one arbitrary constant for each equation). Indeed, as initial condition for the system will contain only two values, $x(0) = X_0$ and $y(0) = Y_0$, hence we only determine two arbitrary constants. This means that there is a relation between k_1, k_2 and c_1, c_2.

To find the relation, we plug our solutions into the first equation of the system:

$$x' - 3x + 4y = 1$$

$$(k_1 e^{-5t} + k_2 e^{t} + 8t + 5)' - 3(k_1 e^{-5t} + k_2 e^{t} + 8t + 5) + 4(c_1 e^{-5t} + c_2 e^{t} + 6t + 2) = 1$$

$$(-8k_1 + 4c_1) e^{-5t} + (-2k_2 + 4c_2) e^{t} + \underbrace{(8t+5)' - 3(8t+5) + 4(6t+2)}_{= 1} = 1$$

Thus $(-8k_1 + 4c_1) e^{-5t} + (-2k_2 + 4c_2) e^{t} = 0$

Since e^{-5t} and e^{t} are linearly independent, we must have $-8k_1 + 4c_1 = 0$ and $-2k_2 + 4c_2 = 0$, so $k_1 = \frac{1}{2} c_1$, $k_2 = 2c_2$

The general solution of the system is

$$x(t) = \frac{1}{2} c_1 e^{-5t} + 2c_2 e^{t} + 8t + 5$$

$$y(t) = c_1 e^{-5t} + c_2 e^{t} + 6t + 2.$$

<u>Method 2</u>. Plug in one solution into one of the equations.

In this approach, we first find one of the unknowns as in the previous method. We have $y = c_1 e^{-5t} + c_2 e^{t} + 6t + 2$. We now plug this into the equation $y' - 4x + 7y = 10t$. We find

$$x = -\frac{10}{4}t + \frac{y'}{4} + \frac{7y}{4} = -\frac{10}{4}t + \frac{1}{4}(c_1 e^{-5t} + c_2 e^{t} + 6t + 2)'$$

$$+ \frac{7}{4}(c_1 e^{-5t} + c_2 e^{t} + 6t + 2) = -\frac{5}{2}t - \frac{5}{4}c_1 e^{-5t} + \frac{c_2}{4}e^{t} + \frac{6}{4} + \frac{7}{4}c_1 e^{-5t} + \frac{7c_2}{4}e^{t}$$

$$+ \frac{21}{2}t + \frac{7}{2} = \frac{1}{2}c_1 e^{-5t} + 2c_2 e^{t} + 8t + 5$$

Remark. It may seem that the second method is simpler than the first one. This was the case in the previous example because we could solve directly for x in $y' - 4x + 7y = 10t$. But if both equations involved x', as it is the case in the general situation, then the resulting equation for x (after plugging in y) will still be a differential equation.

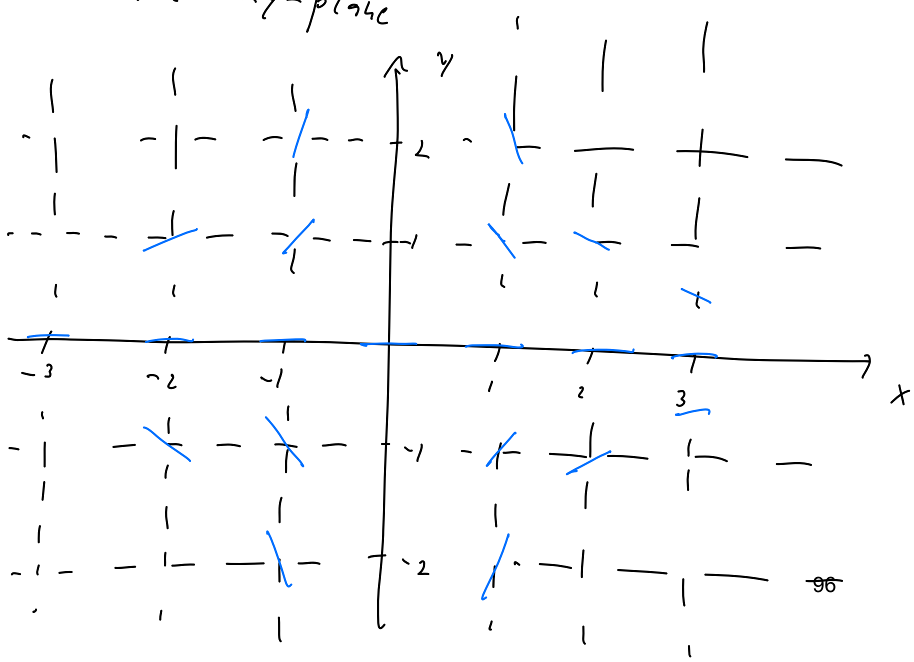
Remark. We can use similar ideas to solve systems with more unknowns and also with higher order equations.

Direction fields

Consider the DE $y' = f(x,y)$. If $f(x,y)$ is very complicated, it might be hard to find the function y. We will develop a method for studying this equation that will allow us to get a good grasp of how y looks like, even when we cannot write it explicitly.

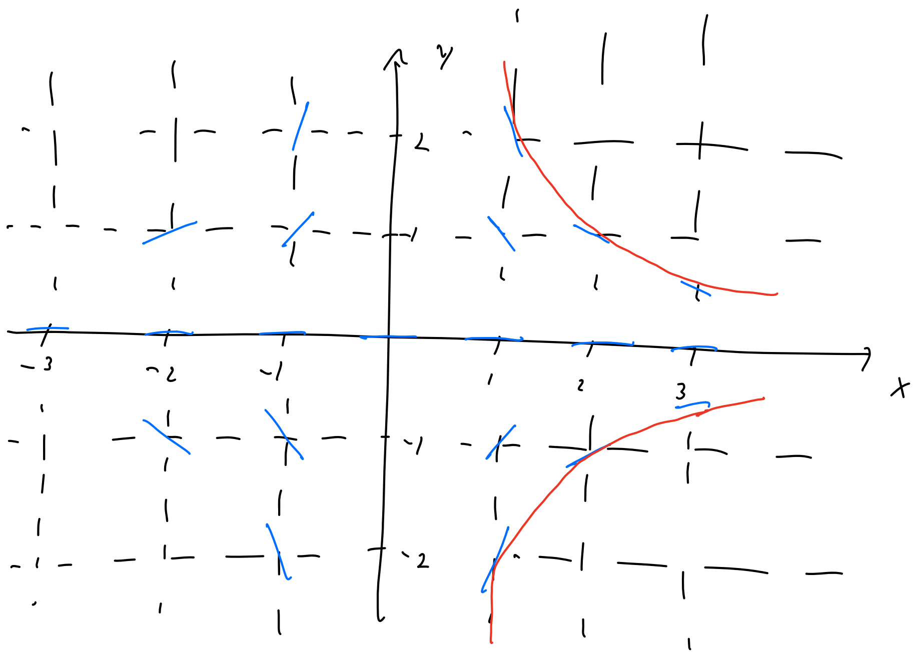
<u>EX</u>: Consider the equation $y' = -\frac{y}{x}$. We can solve this equation, but for the sake of illustrating the new method, let us imagine that we do not know the solution. What the equation tells us is the value of the slope of the tangent to the graph of y (i.e., y') at each point x, y. We construct a table of values. With enough values, we can plot the slopes on the xy-plane

x	y	$y' = -\frac{y}{x}$
1	0	0
1	1	-1
2	1	$-1/2$
1	2	-2
⋮	⋮	⋮



We call such a picture a direction field for the equation $y' = f(x,y)$.

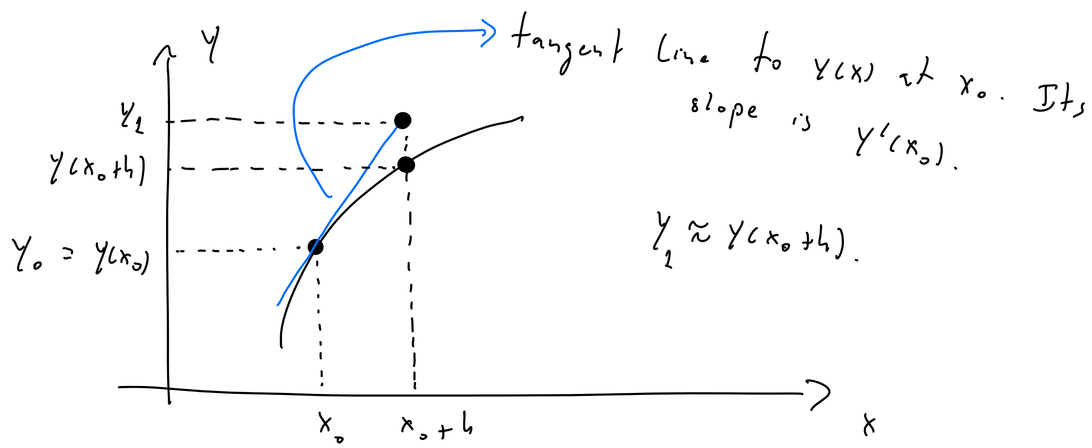
Using enough points, we can sketch solutions. The important thing to remember is that the solutions have their graphs tangent to the line segments we plotted, and that they vary continuously. For example, below we draw the solutions satisfying $y(1) = 2$ and $y(1) = -2$



Euler's method

Consider a DE $y' = f(x,y)$. Depending on what f is, we may not be able to find a formula for the general solution. In this case, we can use direction fields to obtain some qualitative information on the behavior of solutions. Euler's method is a way of finding approximate solutions that provide further, quantitative information.

The idea of Euler's method is that if we know the value of $y = y(x)$ at x_0, then $y(x_0+h)$ can be approximated with the help of the derivative of y at x.



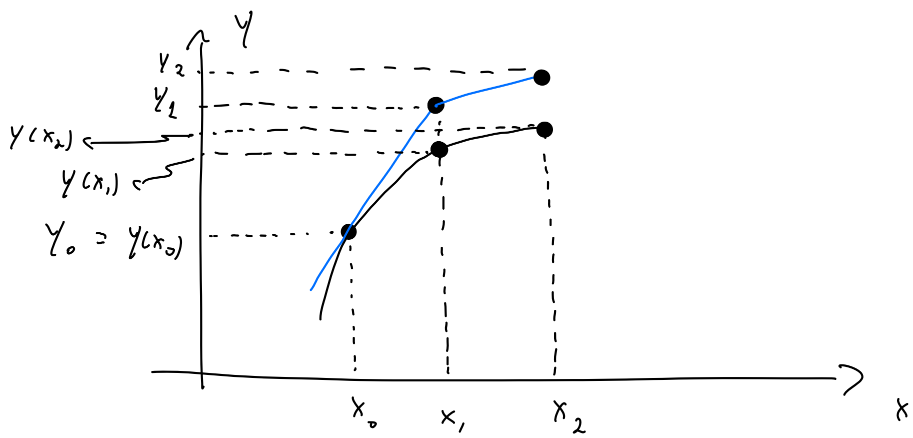
This idea requires knowing y', which in our case we know because we have the DE $y' = f(x,y)$. Thus,

$$y'(x_0) = f(x_0, y_0) \approx \frac{y(x_0+h) - y(x_0)}{h} \Rightarrow y(x_0+h) \approx y(x_0) + h f(x_0, y_0)$$

We can now repeat the process. Starting from the point $x_1 = x_0 + h$, $y_1 \approx y(x_0+h) = y(x_1)$, we find $y_2 \approx y(x_1+h) = y(x_0+2h)$

$$y'(x_1) = f(x_1, y(x_1)) \approx \frac{y(x_1+h) - y(x_1)}{h} \Rightarrow y(x_1+h) \approx y(x_1) + h f(x_1, y(x_1)).$$

This formula is not good because we do not know $y(x_1)$. By we can use $y_1 \approx y(x_1)$ so $y(x_1+h) \approx y_1 + h f(x_1, y_1)$.



We can continue the process and find y_3, y_4, etc., which will be approximate $y(x_3)$, $y(x_4)$, etc., where $x_3 = x_0 + 3h$, $x_4 = x_0 + 4h$, etc.

Summary of Euler's method. Consider $\begin{cases} y' = f(x,y) \\ y(x_0) = y_0 \end{cases}$.

Fix a small number h, called the step size, and set, inductively:

$$x_{m+1} = x_m + h$$

$$y_{m+1} = y_m + h f(x_m, y_m)$$

The points y_m will be approximations for $y(x_m)$.

Remark. Because we have to know the initial point (x_0, y_0), Euler's method is better suited to study IVP. However, we can use it to investigate the general solution upon varying y_0.

Remark. Typically, the smaller the step size h, the better the approximation.

EX: Consider $y' = x\sqrt{y}$, $y(1) = 4$. We can solve this equation exactly. Let us compare the values of the exact solution with those of Euler's method with $h = 0.1$.

n	x_n	y_n (Euler's method)	$y(x_n)$ (exact value)
0	1	4	4
1	1.1	4.2	4.21276
2	1.2	4.42543	4.45210
3	1.3	4.67787	4.71976
4	1.4	4.95904	5.01760
5	1.5	5.27081	5.34766

Numerical solutions of systems

A system of first order DE with k equations for k unknown functions $x_1(t), \dots, x_k(t)$ can be written as

$$x_1'(t) = f_1(t, x_1, \dots, x_k)$$
$$x_2'(t) = f_2(t, x_1, \dots, x_k)$$
$$\vdots$$
$$x_k'(t) = f_m(t, x_1, \dots, x_k)$$

When the system is written in this form, i.e., with the coefficients of all x_i', $i = 1, \dots, k$, equal to one, we say that the system is written in normal form.

EX: The system $\begin{aligned} x' &= 2x + y \\ y' &= xy \end{aligned}$ is in normal form, while $\begin{aligned} yx' &= \cos x \\ y' &= x + y \end{aligned}$ is not.

The IVP problem for a system as above has k IC:

$$x_1(0) = X_{0,1}, \quad x_2(0) = X_{0,2}, \quad \dots, \quad x_n(0) = X_{k,0}.$$

The Euler method for systems is done in the same way as for a single equation. We set:

$$t_{m+1} = t_m + h$$

$$x_{1,m+1} = x_{1,m} + h f_1(t_m, x_{1,m}, x_{2,m}, \dots, x_{k,m})$$
$$x_{2,m+1} = x_{2,m} + h f_2(t_m, x_{1,m}, x_{2,m}, \dots, x_{k,m})$$
$$\vdots$$
$$x_{k,m+1} = x_{k,m} + h f_k(t_m, x_{1,m}, x_{2,m}, \dots, x_{k,m})$$

where h is the step size. Notice that these formulas assume that the system is in normal form.

We can write the above formulas in a compact way upon introducing the vectors:

$$x(t) = (x_1(t), x_2(t), \dots, x_k(t))$$

$$f(t,x) = (f_1(t, x_1, \dots, x_k), f_2(t, x_1, \dots, x_k), \dots, f_k(t, x_1, \dots, x_k))$$

so that we have

$$t_{m+1} = t_m + h$$

$$x_{m+1} = x_m + h f(t_m, x_m).$$

Higher order equations as systems

Besides the fact that systems of DE are common in applications, there is another reason to study them: a DE of order k can always be written as a system of k DE of first order. The procedure is as follows. Given

$$y^{(k)}(t) = f(t, y, y', y'', \ldots, y^{(k-1)})$$

set

$$x_1(t) = y(t), \quad x_2(t) = y'(t), \quad x_3(t) = y''(t), \quad \ldots, \quad x_k(t) = y^{(k-1)}(t)$$

Then

$$x_1'(t) = y'(t) = x_2(t)$$

$$x_2'(t) = y''(t) = x_3(t)$$

$$\vdots$$

$$x_{k-1}'(t) = y^{(k-1)}(t) = x_k(t)$$

$$x_k'(t) = f(t, y, y', \ldots, y^{(k-1)}) = f(t, x_1, x_2, \ldots, x_{k-1}).$$

So we have the system

$$x_1' = x_2$$

$$x_2' = x_3$$

$$\vdots$$

$$x_k' = x_k$$

$$x_k' = f(t, x_1, \ldots, x_{k-1}).$$

Solving this system, we find $x_1, \ldots, x_k$, so in particular we find y because $y = x_1$. If the DE for y comes with IC:

$$y(t_0) = Y_0, \quad y'(t_0) = Y_1, \quad \ldots, \quad y^{(k-1)}(t_0) = Y_{k-1},$$

then we have IC for the system:

$$x_1(t_0) = Y_0, \quad x_2(t_0) = Y_1, \quad \ldots, \quad x_k(t_0) = Y_{k-1}.$$

Since Euler's method can be applied to systems, as a consequence it can be applied to equations of order k as well.

The matrix form of linear systems

We are going to develop methods for studying linear systems of first order DE, i.e., systems of DE where each equation in the system is linear. A linear system of n first order DE for m unknowns can be written as:

$$
\begin{aligned}
x_1' &= a_{11}(t)\, x_1 + a_{12}(t)\, x_2 + \cdots + a_{1m}(t)\, x_m + f_1(t)\\
x_2' &= a_{21}(t)\, x_1 + a_{22}(t)\, x_2 + \cdots + a_{2m}(t)\, x_m + f_2(t)\\
&\vdots\\
x_n' &= a_{n1}(t)\, x_1 + a_{n2}(t)\, x_2 + \cdots + a_{nm}(t)\, x_m + f_n(t).
\end{aligned}
$$

where $a_{ij}(t)$, $i=1,\dots,n$, $j=1,\dots,m$, and $f_1(t),\dots,f_n(t)$ are given functions. The system is called homogeneous if $f_1 = \cdots = f_n = 0$ and inhomogeneous otherwise.

To deal with systems, it is convenient to introduce the following concept.

Def. A $n\times m$ (n by m) rectangular array of numbers is called a (n by n) matrix. If $m=n$ we say that the matrix is square. A m by 1 matrix is called a column vector. If a matrix A has entries a_{ij}, $i=1,\dots,n$, $j=1,\dots,m$, we write $A=[a_{ij}]$.

Ex: $\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ is a 2 by 3 matrix, $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is a 2×2 and a square matrix, and $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ is a 3×1 matrix and a column vector.

Def. An order m-tuple of numbers $(u_1, u_2, \ldots, u_m)$ is called a vector with m components.

Def. The dot product of two vectors $u = (u_1, \ldots, u_m)$ and $v = (v_1, \ldots, v_m)$, denoted $u \cdot v$, is defined as

$$u \cdot v = \sum_{j=1}^{m} u_j v_j = u_1 v_1 + u_2 v_2 + \cdots + u_m v_m .$$

Remark. All the properties of vectors and of the dot product learned in calculus for 2 and 3 component vectors hold for m-component vectors.

Remark. The dot product is only defined between two vectors with the same number of components. Note that the dot product of two vectors is a number, not a vector.

Remark. Given a vector $u = (u_1, \ldots, u_m)$, we can construct out of it the column vector $\begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$. Reciprocally, given the column vector $\begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix}$ we can construct out of it the vector $(u_1, \ldots, u_m)$. Thus, we will not distinguish between column vectors and vectors, referring to column vectors simply as vectors.

Def. Let A be a $n \times m$ matrix, which we can write as

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

We can think of each row of A as a m component vector, so we write

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

where $a_i = (a_{i1}, a_{i2}, \dots, a_{im})$, $i = 1, \dots, n$. Let x be a m component vector, $x = (x_1, \dots, x_m)$. We define the product of A by x, written

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & & & \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$$

as the n component vector given by

$$Ax = \begin{bmatrix} a_1 \cdot x \\ a_2 \cdot x \\ \vdots \\ a_n \cdot x \end{bmatrix}.$$

Remark. Note that Ax is only defined if the number of columns of A equals the number of components of x.

Ex: Find Ax if $A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 2 & 0 \end{bmatrix}$ and $x = (2, 1, -2)$.

In this case $a_1 = (1, -1, 2)$, $a_2 = (1, 2, 0)$ and

$$a_1 \cdot x = 1 \cdot 2 + (-1) \cdot 1 + 2 \cdot (-2) = -3$$

$$a_2 \cdot x = 1 \cdot 2 + 2 \cdot 1 + 0 \cdot (-2) = 4$$

So $Ax = \begin{bmatrix} -3 \\ 4 \end{bmatrix}$.

Consider the system

$$\begin{aligned} x_1' &= a_{11}(t)\, x_1 + a_{12}(t)\, x_2 + \cdots + a_{1m}(t)\, x_m + f_1(t) \\ x_2' &= a_{21}(t)\, x_1 + a_{22}(t)\, x_2 + \cdots + a_{2m}(t)\, x_m + f_2(t) \\ &\vdots \\ x_n' &= a_{n1}(t)\, x_1 + a_{n2}(t)\, x_2 + \cdots + a_{nm}(t)\, x_m + f_n(t). \end{aligned}$$

We can write it as $[x]_n' = A x + f$, where

$$A = A(t) = \begin{bmatrix} a_{11}(t) & a_{12}(t) & \cdots & a_{1m}(t) \\ a_{21}(t) & a_{22}(t) & \cdots & a_{2m}(t) \\ \vdots & & & \\ a_{n1}(t) & a_{n2}(t) & \cdots & a_{nm}(t) \end{bmatrix}$$

is called the coefficient matrix, $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}$, $f = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$

and $[x]_n' = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{bmatrix}$. The system is said then to be written in matrix form.

Remark. Above, we write $[x]_n'$ to emphasize this is the vector of the derivatives of the first n components of x. This is generally different than x', as the latter is $x' = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix}' = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_m' \end{bmatrix}$. In most cases, we will deal with systems where $n = m$, in which case $[x]_n' = x'$ and we can then write

$$x' = Ax + f.$$

Ex: Write $\begin{aligned} x' &= 2x + ty + 6 \\ y' &= 2tx + y + e^t \end{aligned}$ in matrix form.

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 2 & t \\ 2t & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 6 \\ e^t \end{bmatrix}.$$

Linear algebra and algebraic equations

A set of equations

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1$$
$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2$$
$$\vdots$$
$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m$$

is a linear system of m algebraic equations for n unknowns $x_1, \ldots, x_n$. These systems are studied in linear algebra. Here, we briefly review how to solve such systems by Gauss-Jordan elimination.

Ex: Solve

$$\begin{aligned} 2x_1 + 6x_2 + 8x_3 &= 16 \\ 4x_1 + 15x_2 + 19x_3 &= 38 \\ 2x_1 \qquad\quad + 3x_3 &= 6 \end{aligned}$$

Denoting by L_i the i^{th} line of the system, we write $AL_i + BL_j \to L_j$ to indicate the operation where the j^{th} line is replaced by $AL_i + BL_j$.

$$\begin{aligned} -2L_1 + L_2 &\to L_2 \\ -L_1 + L_3 &\to L_3 \end{aligned} \implies \begin{aligned} 2x_1 + 6x_2 + 8x_3 &= 16 \\ 3x_2 + 3x_3 &= 6 \\ -6x_2 - 5x_3 &= -10 \end{aligned}$$

$2L_2 + L_3 \to L_3$ $\qquad$ $2x_1 + 2x_3 = 4$

$-2L_2 + 2L_1 \to L_1$ $\qquad$ $3x_2 + 3x_3 = 6$

$\Longrightarrow$ $\qquad$ $x_3 = 2$

$-2L_3 + L_1 \to L_1$ $\qquad$ $2x_1 = 0$

$-3L_3 + L_2 \to L_2$ $\qquad$ $3x_2 = 0$, $\quad x_1 = 0,\ x_2 = 0,\ x_3 = 2.$

$\Longrightarrow$ $\qquad$ $x_3 = 2$

<u>Ex:</u> Solve

$$\begin{aligned} 2x_1 + 4x_2 + x_3 &= 8 \\ 2x_1 + 4x_2 &= 6 \\ -4x_1 - 8x_2 + x_3 &= -10 \end{aligned}$$

Arguing as above we find

$$\begin{aligned} 2x_1 + 4x_2 &= 6 \\ -x_3 &= -2 \\ 0 &= 0 \end{aligned}$$

We see that x_2 is not determined; it is "free" so the system has infinitely many solutions, given by

$x_1 = -2x_2 + 3$

$x_3 = 2$

$x_2 \in \mathbb{R}$.

Matrices and vectors

The addition of matrices and multiplication by scalars is done entry-wise, i.e., if we denote $A = [a_{ij}]$, $B = [b_{ij}]$ then $A + B = [a_{ij} + b_{ij}]$ and $rA = [ra_{ij}]$ (assuming that A and B have the same size).

Ex: $\begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 1 & 1 \end{bmatrix}$

$3 \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 3 & 0 \\ 3 & 0 & 3 \end{bmatrix}$.

With the above operations, and defining the zero matrix to be the matrix with all entries equal to zero, we have that the set of $m \times n$ matrices forms a vector space. In particular, the defining vector space properties (associativity, etc.) are satisfied.

If A is a $m \times n$ matrix, and B is a $n \times l$ matrix, the product AB is defined as the $m \times l$ matrix whose jth column is given by Ab_j, where b_j is the jth column of B.

So, if $A = [a_{ij}]$, $B = [b_1 \; b_2 \; \cdots \; b_l] = [b_{ij}]$, then

$AB = [Ab_1 \; Ab_2 \; \cdots \; Ab_l]$, or $AB = C$ with $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$.

If A is a $m \times n$ matrix, its transpose, denoted A^T, is the $n \times m$ matrix defined as $[a_{ij}]^T_{n \times m} = [a_{ji}]_{m \times n}$.

The inverse of a square matrix A, denoted A^{-1}, is a matrix such that $AA^{-1} = A^{-1}A = I$, where I is the identity matrix, defined as the matrix with 1 in the diagonal entries and zero everywhere else. If A^{-1} exists, we say that A is invertible.

A linear system:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}
$$

can be written matrix form as $Ax = b$, where

$$
A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}
$$

If $n = m$, and A is invertible, x is then given by $x = A^{-1}b$.

By a row operation on a matrix, we mean any one of the following:

(a) Interchanging two rows of the matrix.

(b) Multiplying a row of the matrix by a non-zero scalar.

(c) Adding a scalar multiple of one row of the matrix to another row (and replacing one of the rows by the result).

If the $n \times n$ matrix A has an inverse A^{-1}, the latter can be determined as follows: we write the $n \times 2n$ matrix $[A \,\vdots\, I]$, and perform row operations until we obtain $[I \,\vdots\, B]$. Then $B = A^{-1}$.

If $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, its determinant is defined as

$$\det A = a_{11}a_{22} - a_{12}a_{21}.$$

If $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$, then we define

$$\det A = a_{11} \det \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{12} \det \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix} + a_{13} \det \begin{bmatrix} a_{21} & a_{2} \\ a_{31} & a_{32} \end{bmatrix}$$

The determinant of a $n \times n$ matrix can be defined inductively, i.e., the determinant of 4×4 matrices is written in terms of determinants of 3×3 submatrices and so on. These are called cofactor expansions (see linear algebra).

Theo. Let A be a $n \times n$ matrix. The following statements are equivalent:

(a) A is singular (does not have an inverse)

(b) $\det A = 0$

(c) $Ax = 0$ has non-trivial solutions ($x \neq 0$).

(d) The columns (rows) of A form are linearly dependent set.

We recall that we say that the vectors $a_1, \ldots, a_n$ are linearly dependent if it is possible to find numbers $c_1, \ldots, c_n$, not all zero, such that $c_1 a_1 + \cdots + c_n a_n = 0$.

Let A be a $n \times n$ matrix. If A is not singular, then the system $Ax = b$ always has a unique solution (given by $x = A^{-1} b$). If A is singular, either $Ax = b$ has no solution, or it has infinitely many solutions. In the latter case, the solutions are given by $x = x_h + x_p$, where x_p is a particular solution satisfying $Ax_p = b$ and x_h are solutions of the homogeneous equation $Ax_h = 0$ (note that there are infinitely many x_h's in this case).

Calculus of matrices.

If the entries $a_{ij}(t)$ of the matrix A are functions of t, then we say that $A = A(t)$ is a matrix function of t. We say that $A(t)$ is continuous (differentiable) at t_0 if each $a_{ij}(t)$ is continuous (differentiable) at t_0. The derivative and integral of $A(t)$ are defined as

$$\frac{d}{dt}A(t) = A'(t) = [a_{ij}'(t)], \quad \int_a^b A(t)\,dt = \left[\int_a^b a_{ij}(t)\,dt\right]$$

It follows that:

$$\frac{d}{dt}(CA) = CA', \quad C \text{ a constant matrix},$$

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

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

In the last formula, note that the order in which the matrices are written matters.

Linear systems in normal form.

We say that a system of n first order linear DE for n unknown functions $x_1, \dots, x_n$ is in normal form if it can be expressed as

$$x'(t) = A(t)\, x(t) + f(t)$$

where $x(t) = (x_1(t), \dots, x_n(t))$, $A(t) = [a_{ij}(t)]$ is the coefficient matrix, and $f(t) = (f_1(t), \dots, f_n(t))$ is the inhomogeneous term. The system is said to have constant coefficients if the matrix $A(t)$ is a constant matrix.

The initial value problem for the system $x' = Ax + f$, with initial condition $x(t_0) = x_0$, consists in finding a solution $x(t)$ defined in a neighborhood of t_0 such that $x(t_0) = x_0$.

Theo. Let $A(t)$ and $f(t)$ be continuous on the interval I that contains t_0, where $A(t)$ is $n \times n$. Then, for any x_0 there exists a unique solution $x(t)$, defined on the whole interval I, to the IVP $x' = Ax + f$, $x(t_0) = x_0$.

If we define $L(x) = Lx = x' - Ax$, then the system can be written as $Lx = f$. We remark that for any differentiable functions $x(t)$ and $y(t)$ and constants a and b, we have

$$L(ax + by) = aLx + bLy.$$

This means that L defines a linear operator. In this case, L maps differentiable functions to continuous functions and this mapping is linear. As a consequence, if $x_1, \dots, x_k$ are solutions of the homogeneous system $Lx = 0$, then $c_1 x_1 + \dots + c_k x_k$ is also a solution, where $c_1, \dots, c_k$ are arbitrary constants.

Def. The vector functions $x_1, \dots, x_k$ are said to be linearly dependent on an interval I if there exist constants $c_1, \dots, c_k$, not all zero, such that

$$c_1 x_1(t) + \dots + c_k x_k(t) = 0 \quad \text{for all } t \in I.$$

Otherwise, they are said to be linearly independent.

Ex: $x_1(t) = (\sin(2t), \cos t)$ and $x_2(t) = (\sin t \cos t, \frac{1}{2} \cos t)$ are linearly dependent on $(-\infty, \infty)$ since $(\sin 2t, \cos t) = (2 \sin t \cos t, \cos t)$, so $x_1(t) - 2 x_2(t) = 0$.

Ex: $x_1(t) = (e^t, 0, e^t)$, $x_2(t) = (e^t, e^t, -e^t)$, $x_3(t) = (e^t, 2e^t, e^t)$ are linearly independent on $(-\infty, \infty)$.

To see this, let c_1, c_2, c_3 be constants such that

$$c_1 x_1 + c_2 x_2 + c_3 x_3 = 0 \quad \text{for all } t \in (-\infty, \infty)$$

Then this holds in particular for $t=0$, so

$$c_1\begin{bmatrix}1\\0\\1\end{bmatrix}+c_2\begin{bmatrix}1\\1\\-1\end{bmatrix}+c_3\begin{bmatrix}1\\2\\1\end{bmatrix}=0$$

Solving for c_1, c_2, and c_3 we find $c_1=c_2=c_3=0$, and we conclude that x_1, x_2, and x_3 are linearly independent on $(-\infty,\infty)$. In fact, this shows that they are linearly independent on any interval containing zero.

<u>Ex:</u> $x_1(t)=\begin{bmatrix}t\\|t|\end{bmatrix}$ and $x_2(t)=\begin{bmatrix}|t|\\t\end{bmatrix}$ are linearly independent on $(-\infty,\infty)$.

To see this, note that $x_1(t)=x_2(t)$ for $t>0$ and $x_1(t)=-x_2(t)$ for $t<0$. If $c_1x_1(t)+c_2x_2(t)=0$ then, for $t>0$ we have $c_1=-c_2$, and for $t<0$ we have $c_1=c_2$, hence $c_1=c_2=0$, giving linear independence.

<u>Remark.</u> Note that linear dependence/independence depends on the interval. E.g., $\begin{bmatrix}t\\|t|\end{bmatrix}$ and $\begin{bmatrix}|t|\\t\end{bmatrix}$ are linearly independent on $(-\infty,\infty)$ but they are linearly dependent on $(0,\infty)$.

Def. The Wronskian of n vector functions $x_1(t) = (x_{11}(t), \dots, x_{n1}(t)), \dots, x_n(t) = (x_{1n}(t), \dots, x_{nn}(t))$ is defined as the function:

$$W(x_1, \dots, x_n)(t) = \det \begin{bmatrix} x_{11}(t) & x_{12}(t) & & x_{1n}(t) \\ x_{21}(t) & x_{22}(t) & \cdots & x_{2n}(t) \\ \vdots & \vdots & & \vdots \\ x_{n1}(t) & x_{n2}(t) & & x_{nn}(t) \end{bmatrix}$$

Theo. If $W(x_1, \dots, x_n)(t_0) \neq 0$, then $x_1, \dots, x_n$ linearly independent on any interval (a,b) containing t_0.

proof: Consider $c_1 x_1(t) + \dots + c_n x_n(t) = 0$. If this holds for any $t \in (a,b)$ then in particular it holds for $t = t_0$, so $c_1 x_1(t_0) + \dots + c_n x(t_0) = 0$. If not all c_i's are zero, this means that the system $Ac = 0$ has a non-trivial solution $c = (c_1, \dots, c_n) \neq 0$, where

$$A = \begin{bmatrix} x_{11}(t_0) & x_{12}(t_0) & & x_{1n}(t_0) \\ x_{21}(t_0) & x_{22}(t_0) & \cdots & x_{2n}(t_0) \\ \vdots & \vdots & & \vdots \\ x_{n1}(t_0) & x_{n2}(t_0) & & x_{nn}(t_0) \end{bmatrix}$$

But then $\det A = 0$ (by a previous theorem), contradicting the assumption. $\boxed{\checkmark}$

Theo. Let A be a $n \times n$ continuous matrix function. If $x_1, \dots, x_n$ are linearly independent solutions of $x' = Ax$ on an interval I, then their Wronskian never vanishes on I.

proof: Suppose that $W(x_1, \dots, x_n)(t_0) = 0$ for some $t_0 \in I$. Then the vectors $x_1(t_0), \dots, x_n(t_0)$ are linearly dependent, so we can find $c_1, \dots, c_n$ such that $c_1 x_1(t_0) + \dots + c_n x_n(t_0) = 0$. The functions $c_1 x_1(t) + \dots + c_n x_n(t)$ and $z(t) \equiv 0$ are both solutions to the IVP $x' = Ax$, $x(t_0) = 0$, so by uniqueness we have $z(t) = c_1 x_1(t) + \dots + c_n x_n(t) = 0$ for all $t \in I$, contrary to the assumption. $\square$

Theo. Let $x_1, \dots, x_n$ be solutions to $x' = Ax$ defined on an interval I. Then either their Wronskian vanishes identically on I or it is never zero on I.

Def. An expression of the form $c_1 x_1(t) + \dots + c_n x_n(t)$, where $c_1, \dots, c_n$ are constants, is called a linear combination of $x_1, \dots, x_n$.

Theo. Let $x_1, \dots, x_n$ be n linearly independent solutions to $x' = Ax$ on an interval I, where A is a $n \times n$ continuous matrix function. Then, any solution to $x' = Ax$ on I can be written as a linear combination of $x_1, \dots, x_n$.

Def. A set of $\{x_1, \ldots, x_n\}$ of n linearly independent solutions to $x' = Ax$, (A $n \times n$) is called a fundamental solution set to $x' = Ax$. The linear combination

$$x(t) = c_1 x_1(t) + \cdots + c_n x_n(t)$$

where $c_1, \ldots, c_n$ are constants, is called the general solution to $x' = Ax$. The matrix

$$\Phi(t) = \begin{bmatrix} x_1(t) & \cdots & x_n(t) \end{bmatrix} = \begin{bmatrix} x_{11}(t) & x_{12}(t) & & x_{1n}(t) \\ x_{21}(t) & x_{22}(t) & \cdots & x_{2n}(t) \\ \vdots & \vdots & & \vdots \\ x_{n1}(t) & x_{n2}(t) & & x_{nn}(t) \end{bmatrix}$$

is called the fundamental matrix of $x' = Ax$.

Note that the general solution x can be written as $x(t) = \Phi(t) c$, where $c = (c_1, \ldots, c_n)$ is a constant vector, and that $W(x_1, \ldots, x_n)(t) = \det \Phi(t)$.

The superposition principle for linear systems says that if x_1 and x_2 are solutions to $x_1' = Ax_1 + f_1$ and $x_2' = Ax_2 + f_2$, then $c_1 x_1 + c_2 x_2$ is a solution to $x' = Ax + c_1 f_1 + c_2 f_2$, where c_1 and c_2 are constants.

Theo. The fundamental matrix $\Phi(t)$ satisfies:

$$\Phi'(t) = A(t)\,\Phi(t).$$

proof: We have

$$\Phi'(t) = [\, x_1'(t) \; \cdots \; x_n'(t) \,].$$ Each $x_i(t)$, $i=1,\dots,n$, satisfies $x_i'(t) = A(t)\,x_i(t)$, so

$$\Phi'(t) = [\, A(t)\,x_1(t) \; \cdots \; A(t)\,x_n(t) \,].$$

From the formula for multiplication of matrices we see that

$$[\, A(t)\,x_1(t) \; \cdots \; A(t)\,x_n(t) \,] = A(t)\,[\, x_1(t) \cdots x_n(t) \,] = A(t)\,\Phi(t). \;\blacksquare$$

Def. Given $x' = Ax + f$, we call $x' = Ax$ the associated homogeneous system. The general solution of the associated homogeneous system is denoted x_h.

Theo. If $x_p(t)$ is a particular solution to $x' = Ax + f$ on the interval I, where A is a $n \times n$ continuous matrix function, and $\{x_1(t), \dots, x_n(t)\}$ is a fundamental solution set to the system $x' = Ax$ on I, then every solution to $x' = Ax + f$ on I can be written as $x(t) = x_h(t) + x_p(t)$.

Homogeneous linear systems with constant coefficients

Here we will study the system

$$x' = Ax$$

where A is a $n \times n$ real (constant) matrix. We first recall some definitions from linear algebra.

Def. Let A be a $n \times n$ matrix. The eigenvalues of A are those (real or complex) numbers λ for which the equation $(A - \lambda I)u = 0$ has at least one non-trivial (i.e., non-zero) solution u, where I is the $n \times n$ identity matrix. Note that u is possibly a complex vector. Any non-trivial u satisfying $(A - \lambda I)u = 0$ is called an eigenvector (associated to the eigenvalue λ) of A.

For λ to be an eigenvalue of A, the equation $(A - \lambda I)u = 0$ needs to admit non-trivial solutions, so its determinant must vanish. The equation

$$\det(A - \lambda I) = 0$$

is called the characteristic equation of A. It is a polynomial of degree n and its roots are the eigenvalues, so we find the eigenvalues by finding the roots of the characteristic determinant.

Returning to $x' = Ax$, we try a solution of the form $x(t) = e^{\lambda t} u$, where λ and u have to be determined. Plugging in:

$$(e^{\lambda t} u)' = \lambda e^{\lambda t} u = A e^{\lambda t} u \Rightarrow (A - \lambda I) u = 0.$$

Thus, λ is an eigenvalue of A and u an eigenvector. Said differently, if λ is an eigenvalue of A and u is a corresponding eigenvector, then $x = e^{\lambda t} u$ solves $x' = Ax$.

Theorem. Suppose the constant $n \times n$ matrix A has n linearly independent eigenvectors $u_1, \dots, u_n$. Let λ_i be the eigenvalue corresponding to u_i. Then $\{e^{\lambda_1 t} u_1, e^{\lambda_2 t} u_2, \dots, e^{\lambda_n t} u_n\}$ is a fundamental solution set for $x' = Ax$ on $(-\infty, \infty)$. Thus, the general solution of $x' = Ax$ is

$$x = c_1 e^{\lambda_1 t} u_1 + \dots + c_n e^{\lambda_n t} u_n,$$ where $c_1, \dots, c_n$ are arbitrary constants.

Proof: Set $\mathbb{X}(t) = [\, e^{\lambda_1 t} u_1 \ \cdots \ e^{\lambda_n t} u_n \,]$. Then $\det \mathbb{X}(t) = e^{(\lambda_1 + \dots + \lambda_n)t} \det [u_1 \cdots u_n]$ which is never zero since $u_1, \dots, u_n$ are linearly independent, hence the result by one of our previous theorems. $\square$

We now recall some useful results from linear algra. In what follows, A is a constant $n \times n$ matrix.

- If $\lambda_1, \dots, \lambda_n$ are distinct eigenvalues of A, then the vectors $u_1, \dots, u_n$ are linearly independent, where u_i is an eigenvector associated with λ_i.

- Any non-zero multiple of an eigenvector is also an eigenvector.

- If A is real and symmetric, i.e., all entries of A are real and $A^T = A$, where A^T is the transpose of A, then A admits n linearly independent eigenvectors.

The case of complex eigenvalues

Consider $x' = Ax$, where A is a $n \times n$ real (constant) matrix. We saw that if λ is an eigenvalue and u an associated eigenvalue, then $z = e^{\lambda t} u$ is a solution. We can write

$$z = e^{(\alpha + i\beta)t} (a + ib)$$

where $\alpha, \beta \in \mathbb{R}$ and a and b are real vectors. The complex conjugate of λ, $\bar{\lambda}$, is also an eigenvalue and $\bar{z}$ is a corresponding eigenvector, so we write

$$\bar{z} = e^{(\alpha - i\beta)t} (a - ib)$$

Using Euler's formula, we can write

$$z = e^{\alpha t} (\cos(\beta t) + i \sin(\beta t)) (a + ib)$$

$$= e^{\alpha t} (\cos(\beta t)\, a - \sin(\beta t)\, b) + i\, e^{\alpha t} (\sin(\beta t)\, a + \cos(\beta t)\, b)$$

giving two linearly independent real solutions (note that the same conclusion holds if we use $\bar{z}$).

Summarizing: if a real (constant) matrix A has complex conjugate eigenvalues $\alpha \pm i\beta$ with corresponding eigenvectors $a \pm ib$, a, b real vectors, then two linearly independent real solutions are given by

$$x_1(t) = e^{\alpha t} \cos(\beta t)\, a - e^{\alpha t} \sin(\beta t)\, b, \quad x_2(t) = e^{\alpha t} \sin(\beta t)\, a + e^{\alpha t} \cos(\beta t)\, b.$$

The method of undetermined coefficients for systems

We will now discuss methods for solving non-homogeneous systems of DE, starting with the method of undetermined coefficients.

The method of undetermined coefficients for systems of DE is very similar to the case of a single equation. We will illustrate the method with examples.

<u>EX:</u> Find the general solution of

$$x' = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} x + \begin{bmatrix} -4\cos t \\ -\sin t \end{bmatrix}$$

First, we solve the associated homogeneous equation $x' = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} x$. The matrix $A = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}$ has eigenvalues $\lambda_1 = 0$ and $\lambda_2 = 4$. Corresponding eigenvectors are $u_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $u_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. We conclude that $x_1 = e^{0t}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $x_2 = e^{4t}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ are two linearly independent solutions of $x' = Ax$.

We now seek for a particular solution x_p. Mimicking what we have done for single equations, we look for a solution of the form $x_p = \cos t\, a + \sin t\, b$, except that now a and b

are vectors to be determined rather than real valued constants, i.e., $a, b \in \mathbb{R}^2$. Compute

$$x_p' = -\sin t\, a + \cos t\, b.$$

We want this to equal $Ax_p + \begin{bmatrix} -4\cos t \\ -\sin t \end{bmatrix}$, so

$$-\sin t\, a + \cos t\, b = A(\cos t\, a + \sin t\, b) + \underbrace{\begin{bmatrix} -4\cos t \\ -\sin t \end{bmatrix}}$$

$$= \cos t \begin{bmatrix} -4 \\ 0 \end{bmatrix} + \sin t \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

which we write as

$$(Aa - b)\cos t + (Ab + a)\sin t = \cos t \begin{bmatrix} 4 \\ 0 \end{bmatrix} + \sin t \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Setting the coefficients of $\cos t$ (and $\sin t$) on both sides equal to each other gives:

$$Aa - b = \begin{bmatrix} 4 \\ 0 \end{bmatrix} \quad \text{and} \quad Ab + a = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Recalling the definition of A, we can write this explicitly as:

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} -4 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{l} 2a_1 + 2a_2 - b_1 = 4 \\ 2a_1 + 2a_2 - b_2 = 0 \end{array}$$

$$\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \Rightarrow \begin{array}{l} 2b_1 + 2b_2 + a_1 = 0 \\ 2b_1 + 2b_2 + a_2 = 1, \end{array}$$

where $a = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$, $b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$. This is a system for the unknowns a_1, a_2, b_1, and b_2:

$$\begin{aligned} 2a_1 + 2a_2 - b_1 \qquad\quad &= 4 \\ 2a_1 + 2a_2 \qquad\quad - b_2 &= 0 \\ a_1 \qquad + 2b_1 + 2b_2 &= 0 \\ a_2 + 2b_1 + 2b_2 &= 1 \end{aligned}$$

Using Gauss-Jordan elimination, we find: $a_1 = 0$, $a_2 = 1$, $b_1 = -2$, $b_2 = 2$,

i.e., $a = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $b = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$. Thus

$x_p = \cos t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \sin t \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ and the general solution is

$$x = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 e^{4t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \cos t \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \sin t \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} c_1 + c_2 e^{4t} - 2\sin t \\ -c_1 + c_2 e^{4t} + \cos t + 2\sin t \end{bmatrix}.$$

EX: Give the form of the particular solution to

$$x' = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} x + \begin{bmatrix} e^{4t} \\ 3e^{4t} \end{bmatrix}$$

(You do not need to find x_p.)

First we solve the associated homogeneous equation:

$$x' = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} x \Rightarrow \lambda_1 = 0, \Rightarrow u_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \Rightarrow x_1 = e^{0t}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\lambda_2 = -5 \qquad u_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \qquad x_2 = e^{-5t}\begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

The non-homogeneous term does not repeat any term in the homogeneous solution, thus

$$x_p = e^{4t} a = e^{4t} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}.$$

Ex: Find x_p in the previous example.

We plug x_p in:

$$\left(e^{4t} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \right)' = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{4t} + \begin{bmatrix} e^{4t} \\ 3e^{4t} \end{bmatrix}$$

Canceling e^{4t} we can rewrite this as

$$\begin{bmatrix} 8 \\ -2 \end{bmatrix} a_1 + \begin{bmatrix} -2 \\ 5 \end{bmatrix} a_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \Rightarrow \begin{aligned} 8a_1 - 2a_2 &= 1 \\ -2a_1 + 5a_2 &= 3. \end{aligned}$$

Solving the system we find $a_1 = \frac{11}{36}$ and $a_2 = 13/18$. Then:

$$x_p = e^{4t} \begin{bmatrix} 11/36 \\ 13/18 \end{bmatrix}.$$

Remark. A multiple of a particular solution is not, in general, a particular solution. Thus, in the previous example we cannot multiply x_p by, say, 36, to get rid of the fractions.

Ex: Find x_p for $x' = Ax + f$, with

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 0 & 3 & 2 \\ 0 & -1 & 2 \end{bmatrix}, \quad f(t) = \begin{bmatrix} e^{-t} \\ 2 \\ 1 \end{bmatrix}$$

After some algebra, we find that the solution of the associated homogeneous system is

$$x_h = c_1 e^{2t}\begin{bmatrix}1\\0\\0\end{bmatrix} + c_2\left(e^{\frac{5t}{2}}\cos\left(\frac{\sqrt{7}t}{2}\right)\begin{bmatrix}11\\-2\\4\end{bmatrix} - e^{\frac{5t}{2}}\sin\left(\frac{\sqrt{7}t}{2}\right)\begin{bmatrix}-3\sqrt{7}\\-2\sqrt{7}\\0\end{bmatrix}\right)$$

$$+ c_3\left(e^{\frac{5t}{2}}\sin\left(\frac{\sqrt{7}t}{2}\right)\begin{bmatrix}11\\-2\\4\end{bmatrix} + e^{\frac{5t}{2}}\cos\left(\frac{\sqrt{7}t}{2}\right)\begin{bmatrix}-3\sqrt{7}\\-2\sqrt{7}\\0\end{bmatrix}\right).$$

Now write

$$f(t) = \begin{bmatrix}e^{-t}\\3\\1\end{bmatrix} = e^{-t}\begin{bmatrix}1\\0\\0\end{bmatrix} + \begin{bmatrix}0\\3\\1\end{bmatrix} = f_1(t) + f_2(t).$$

By the superposition principle, we seek x_p in the form $x_p = x_{p1} + x_{p2}$, where x_{p1} is a particular solution associated with f_1 and x_{p2} a particular solution associated with f_2. f_1 is an exponential that does not repeat any term in x_h, so we put $x_{p1} = e^t a = e^t\begin{bmatrix}a_1\\a_2\\a_3\end{bmatrix}$. f_2 is a polynomial of degree zero (i.e., a constant vector) that does not repeat any term in x_h, so we put $x_{p2} = b = \begin{bmatrix}b_1\\b_2\\b_3\end{bmatrix}$.

Plugging x_p into the equation and proceeding as in the above examples, we find:

$$x_p = \begin{bmatrix}-\frac{1}{3}e^t + 11/16\\ -\frac{1}{4}\\ -\frac{5}{8}\end{bmatrix}.$$

Remark. By the superposition principle, we can first find x_{p1} upon plugging it into $x' = Ax + f_1$. In this case we will find $x_{p1} = e^t \begin{bmatrix} -1/3 \\ 0 \\ 0 \end{bmatrix}$. Similarly, we plug x_{p2} into $x' = Ax + f_2$, finding $x_{p2} = \begin{bmatrix} 11/16 \\ -1/4 \\ -5/8 \end{bmatrix}$.

Ex: Find x_p for $x' = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-5t}$.

The associated homogeneous equation for this system was solved as an example above. We found:

$$x_h = c_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 e^{-5t} \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

Because $e^{-5t} \begin{bmatrix} 2 \\ -1 \end{bmatrix}$ solves the associated homogeneous system, we suspect, based on our experience with single equations, that $x_p = a e^{-5t}$ will not work. Let's verify that this is indeed the case. Plugging in:

$$\left(\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{-5t} \right)' = \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{-5t} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-5t}$$

which gives, after differentiating and canceling the exponential:

$$\begin{aligned} a_1 + 2a_2 &= 1 \\ 2a_1 + 4a_2 &= 1 \end{aligned}$$

. Multiplying the first equation by -2 and adding to the second yields:

$$\begin{aligned} a_1 + 2a_2 &= 1 \\ 0 &= -1, \end{aligned}$$

which is of course inconsistent.

Based on our experience with single equations, we are tempted to try $x_p = te^{-5t}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$. However, this will not work either. Indeed, plugging $te^{-5t}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ into the equation gives:

$$x_p' = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{-5t} - 5t \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} e^{-5t} = t \begin{bmatrix} -4a_1 + 2a_2 \\ 2a_1 - a_2 \end{bmatrix} e^{-5t} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} e^{-5t}.$$

Setting the terms with and without t on each side equal to each other gives two systems:

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -4a_1 + 2a_2 \\ 2a_1 - a_2 \end{bmatrix} = \begin{bmatrix} -5a_1 \\ -5a_2 \end{bmatrix}.$$

It is impossible to satisfy both systems at the same time: the first system gives $a_1 = a_2 = 1$, which is not a solution of the second system (although each system, separately, is consistent).

The difference from the single equation case is that for systems the terms in t do not necessarily cancel out. This is because we know we have two constants a_1 and a_2 (or more if the vectors had more components) leading to more conditions necessary for cancelation.

Let us show that

$$x_p = te^{-5t}a + e^{-5t}b = te^{-5t}\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + e^{-5t}\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

works.

Plugging in gives:

$$-5t\begin{bmatrix}a_1\\a_2\end{bmatrix}e^{-5t} + \begin{bmatrix}a_1\\a_2\end{bmatrix}e^{-5t} - 5\begin{bmatrix}b_1\\b_2\end{bmatrix}e^{-5t} = t\begin{bmatrix}-4a_1+2a_2\\2a_1-a_1\end{bmatrix}e^{-5t} + \begin{bmatrix}-4b_1+2b_2\\2b_1-b_2\end{bmatrix}e^{-5t}.$$

Setting the terms with and without t on each side equal to each other:

$$\begin{aligned}
a_1 + 2a_2 \quad &= 0\\
2a_1 + 4a_2 \quad &= 0\\
a_1 \quad - b_1 - 2b_2 &= 1\\
a_2 - 2b_1 - 4b_2 &= 1
\end{aligned}$$

Solving, we find $a_1 = \frac{2}{5}$, $a_2 = -\frac{1}{5}$, $b_1 = -2b_2 - \frac{3}{5}$ and b_2 is a free variable. Thus:

$$X_p = \begin{bmatrix}2/5\\-1/5\end{bmatrix}te^{-5t} + \begin{bmatrix}-2b_2 - 3/5\\b_2\end{bmatrix}e^{-5t}.$$ Because we want the particular solution to contain no free variables, we set $b_2 = 0$. Alternatively, we can write

$$\begin{bmatrix}-2b_2 - 3/5\\b_2\end{bmatrix}e^{-5t} = -b_2\begin{bmatrix}2\\-1\end{bmatrix}e^{-5t} + \begin{bmatrix}-3/5\\0\end{bmatrix}e^{-5t}$$ and combine

$-b_2\begin{bmatrix}2\\-1\end{bmatrix}e^{-5t}$ with X_h. Thus:

$$X_p = \begin{bmatrix}2/5\\-1/5\end{bmatrix}te^{-5t} + \begin{bmatrix}-3/5\\0\end{bmatrix}e^{-5t}.$$

Summary of the method of undetermined coefficients for systems

Consider the system

$$x'(t) = Ax(t) + f(t)$$

where A is a $n \times n$ constant matrix and

$$f(t) = e^{\alpha t}\cos(\beta t)\, P_m(t) + e^{\alpha t}\sin(\beta t)\, Q_m(t)$$

where α and β are real numbers and $P_m(t)$ and $Q_m(t)$ are vector polynomials of degree m, i.e, $P_m(t) = a_0 + a_1 t + \dots + a_m t^m$, $Q_m(t) = b_0 + b_1 t + \dots + b_m t^m$, where $a_0, \dots, a_n, b_0, \dots, b_n$ are n component vectors.

Under the above conditions, the form of the particular solution $x_p(t)$ is given as follows:

- If $\beta = 0$, then $x_p(t) = e^{\alpha t} R_{m+k}(t)$, where $R_{m+k}(t)$ is a vector polynomial of degree $m+k$, with $k=0$ if α is not an eigenvalue of A, and k = multiplicity of α if α is an eigenvalue of A.
- If $\beta \neq 0$, then $x_p(t) = e^{\alpha t}\cos(\beta t)\, R_{m+k}(t) + e^{\alpha t}\sin(\beta t)\, S_{m+k}(t)$, where $R_{m+k}(t)$ and $S_{m+k}(t)$ are vector polynomials of degree $m+k$, with $k=0$ if $\alpha + i\beta$ is not an eigenvalue of A, and k = multiplicity of $\alpha + i\beta$ if $\alpha + i\beta$ is an eigenvalue of A.

Variation of parameters for systems

Now we show how to generalize the method of variation of parameters to systems of DE.

Consider

$$x'(t) = A(t)\,x(t) + f(t).$$

Suppose that $\Phi(t)$ is a fundamental matrix for $x'(t) = A(t)\,x(t)$. Following what we did for single equations, we look for a particular solution in the form $x_p(t) = \Phi(t)\,v(t)$, where the vector valued function $v(t)$ is to be determined. Plugging in:

$$x_p' = (\Phi v)' = \underbrace{\Phi'}_{= A\Phi} v + \Phi v' = A x_p + f = A \Phi v + f$$

Canceling $A\Phi v$ on both sides: $\Phi v' = f$. Thus $v' = \Phi^{-1} f$. Integrating we find $v(t) = \int (\Phi(t))^{-1} f(t)\,dt$, where we do not add a constant since we only need to find a particular solution. Therefore, we have:

$$\boxed{x_p(t) = \Phi(t) \int (\Phi(t))^{-1} f(t)\,dt}$$

Next, consider the IVP $\begin{cases} x'(t) = A(t)\,x(t) + f(t) \\ x(t_0) = x_0 \end{cases}$.

From the previous formula and the fact that the general solution to the associated homogeneous equation can be written as $x_h = \Phi c$, where $c = (c_1, \dots, c_n)$, we have that the solution of the IVP can be written as:

$$x(t) = \Phi(t)\,c + \Phi(t) \int_{t_0}^{t} (\Phi(s))^{-1} f(s)\,ds$$

where c is to be determined. Plugging $t = t_0$ we find

$$x(t_0) = \Phi(t_0)\,c + \underbrace{\Phi(t) \int_{t_0}^{t_0} (\Phi(s))^{-1} f(s)\,ds}_{=0} = x_0 \Rightarrow c = (\Phi(t_0))^{-1} x_0$$

Thus

$$\boxed{x(t) = \Phi(t)\,(\Phi(t_0))^{-1} x_0 + \Phi(t) \int_{t_0}^{t} (\Phi(s))^{-1} f(s)\,ds}$$

<u>Ex</u>! Use variation of parameters to find x_p for

$$x' = \frac{1}{3}\begin{bmatrix} 7 & 2 \\ 4 & 5 \end{bmatrix} x - \begin{bmatrix} 5 \\ 8 \end{bmatrix} e^{t}$$

Using the techniques we learned for homogeneous systems, we find:

$$\mathbf{X}(t) = \begin{bmatrix} -e^{t} & e^{3t} \\ 2e^{t} & e^{3t} \end{bmatrix}, \text{ so } (\mathbf{X}(t))^{-1} = \frac{1}{3}\begin{bmatrix} -e^{-t} & e^{-t} \\ 2e^{-3t} & e^{-3t} \end{bmatrix}.$$

Thus

$$x_p = \mathbf{X}\int \mathbf{X}^{-1} f = \begin{bmatrix} -e^{t} & e^{3t} \\ 2e^{t} & e^{3t} \end{bmatrix} \int \frac{1}{3}\begin{bmatrix} -e^{-t} & e^{-t} \\ 2e^{-3t} & e^{-3t} \end{bmatrix}\left(-\begin{bmatrix} 5 \\ 8 \end{bmatrix}\right)e^{t}\,dt$$

$$= \begin{bmatrix} -e^{t} & e^{3t} \\ 2e^{t} & e^{3t} \end{bmatrix} \int \begin{bmatrix} -1 \\ -6e^{-2t} \end{bmatrix} dt$$

$$= \begin{bmatrix} -e^{t} & e^{3t} \\ 2e^{t} & e^{3t} \end{bmatrix} \begin{bmatrix} -t \\ 3e^{-2t} \end{bmatrix} = \begin{bmatrix} (t+3)e^{t} \\ (3-2t)e^{t} \end{bmatrix}$$

Remark. Above we employed the following useful formula for the inverse of a 2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The matrix exponential function

If $a \in \mathbb{R}$, the solution to $x' = ax$ is $x(t) = C_1 e^{at}$. For a system $x' = Ax$, where A is a constant matrix, we would like to make a similar statement. The first step is to define e^A when A is a matrix.

Def. If M is a $n \times n$ matrix, we define e^M by

$$e^M = \sum_{k=0}^{\infty} \frac{M^k}{k!} = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \dots$$

where $M^0 = I = n \times n$ identity matrix.

To show that this definition makes sense, we need to show that the series converges. For this, need some way of talking about the "length" or "size" of a matrix. We do this by defining a norm on the space of $n \times n$ matrices (recall that the space of $n \times n$ matrices is a vector space, so it makes sense to talk about a norm).

Def. For a $n \times n$ matrix M we define its norm, denoted $\|M\|$, by

$$\|M\| = \sup_{\|x\|=1} \|Mx\|.$$

Prop. $\|M\|$ is indeed a norm, i.e., it satisfies:

a) $\|M\| \in \mathbb{R}$

b) $\|aM\| = |a|\,\|M\|$, $a \in \mathbb{R}$

c) $\|M\| \geq 0$ and $\|M\| = 0$ iff $M = 0$

d) $\|M + N\| \leq \|M\| + \|N\|$.

Moreover, $\|MN\| \leq \|M\|\|N\|$.

We can now show that e^M is well defined:

$$\|e^M\| = \Big\|\sum_{k=0}^{\infty} \frac{M^k}{k!}\Big\| \leq \sum_{k=0}^{\infty} \frac{\|M^k\|}{k!} \leq \sum_{k=0}^{\infty} \frac{\|M\|^k}{k!} = e^{\|M\|} < \infty,$$

where we used the above proposition (in particular in the step $\|M^k\| \leq \|M\|^k$).

Some properties of the exponential matrix. If M and N are $n \times n$ matrices and $t, s \in \mathbb{R}$, then

a) $e^{M0} = e^0 = I$

b) $e^{M(t+s)} = e^{Mt} e^{Ms}$

c) $(e^M)^{-1} = e^{-M}$

d) $e^{(M+N)t} = e^{Mt} e^{Nt}$ if $MN = NM$

e) $e^{It} = e^t I$

For diagonal matrices it is easy to compute the exponential. For example, let $M = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$. Then $M^2 = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$, $M^3 = \begin{bmatrix} 8 & 0 \\ 0 & 27 \end{bmatrix}$, etc. Then

$$e^M = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}^k = \begin{bmatrix} \sum_{k=0}^{\infty} \frac{2^k}{k!} & 0 \\ 0 & \sum_{k=0}^{\infty} \frac{3^k}{k!} \end{bmatrix} = \begin{bmatrix} e^2 & 0 \\ 0 & e^3 \end{bmatrix}$$

Now we compute:

$$\frac{d}{dt} e^{At} = \frac{d}{dt} \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} = \sum_{k=0}^{\infty} \frac{k A^k t^{k-1}}{k!} = A \sum_{k=1}^{\infty} \frac{A^{k-1} t^{k-1}}{(k-1)!}$$

$$= A \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} = A e^{At}$$

Consequently, e^{At} is a solution of the matrix DE $\Phi' = A \Phi$. Because e^{At} is invertible, its columns are linearly independent, thus e^{At} is a fundamental matrix for the system $x' = Ax$, where A is a $n \times n$ constant matrix. If Φ and Ψ are two fundamental matrices for $x' = Ax$, there always exist a constant matrix M such that $\Psi = \Phi M$. In particular $e^{At} = \Phi(t)(\Phi(0))^{-1}$. If A has n linearly independent eigenvectors $u_1, \dots, u_n$, then

$$e^{At} = [\, e^{\lambda_1 t} u_1 \; \dots \; e^{\lambda_n t} u_n \,] \, [\, u_1 \dots u_n \,]^{-1}.$$

Solving $x' = Ax$ when A does not have n linearly independent eigenvectors

In what follow, A is a (constant) $n \times n$ matrix.

We saw that $e^{\lambda_1 t} u_1, \ldots, e^{\lambda_n t} u_n$ give n linearly independent solutions to $x' = Ax$ if $u_1, \ldots, u_n$ are n linearly independent eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$. Now we will see how to use e^{At} to solve $x' = Ax$ when A does not have n linearly independent eigenvectors.

Def. A non-zero vector u satisfying

$$(A - \lambda I)^m u = 0$$

for some λ and some integer m is called a generalized eigenvector of the matrix A.

Remark. The number λ in the above definition must be an eigenvalue of A since $(A - \lambda I)^{m-1} u$ is an eigenvector associated to λ. Every eigenvector is a generalized eigenvector.

A matrix that does not have n linearly independent eigenvectors is called defective. A defective matrix always has n linearly independent generalized eigenvectors. In fact, if λ is an eigenvalue of multiplicity k, then there always exist k linearly independent generalized eigenvectors associated with λ.

If u is a generalized eigenvector associated to λ, then

$$e^{At}u = e^{\lambda I t} e^{(A-\lambda I)t} u$$

$$= e^{\lambda t}\left[I u + t(A-\lambda I)u + \dots + \frac{t^{m-1}}{(m-1)!}(A-\lambda I)^{m-1}u + \underbrace{\frac{t^m}{m!}(A-\lambda I)^m u}_{=0} + \underbrace{\dots}_{=0} \right]$$

$$= e^{\lambda t}\left[u + t(A-\lambda I)u + \dots + \frac{t^{m-1}}{(m-1)!}(A-\lambda I)^{m-1}u \right]$$

(Note that we computed $e^{At}u$ without knowing e^{At}.)

On the other hand, $e^{At}u$ is a solution to $x' = Ax$. This is because e^{At} is a fundamental matrix, thus the general solution is $e^{At}c$, where $c = (c_1, \dots, c_n)$ is an arbitrary non-zero vector. In particular we have a solution upon choosing $c = u$.

From the above we conclude the following: let $u_1, \dots, u_n$ be n linearly independent generalized eigenvectors (which always exist) corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$ (not necessarily distinct). Then $x_1(t) = e^{At}u_1, \dots, x_n(t) = e^{At}u_n$ are n linearly independent solutions to $x' = Ax$, where each $e^{At}u_i$ is computed as above (without the need to know e^{At}).

Summary for solving $x' = Ax$

1) Compute the characteristic polynomial $p(\lambda) = \det(A - \lambda I)$.

2) Find the roots of $p(\lambda) = 0$. Let the distinct roots be $\lambda_1, \dots, \lambda_k$, and let $m_1, \dots, m_k$ be their multiplicities.

3) For each λ_i, $i = 1, \dots, k$, find m_i linearly independent generalized eigenvectors by solving $(A - \lambda I)^{m_i} u = 0$.

4) Form $n = m_1 + \dots + m_k$ linearly independent solutions to $x' = Ax$ by computing

$$x(t) = e^{At} u = e^{\lambda t}\left[u + t(A - \lambda I)u + \frac{t^2}{2!}(A - \lambda I)^3 u + \dots\right]$$

for each generalized eigenvector u, corresponding to each eigenvalue λ, found in part 3. If λ has multiplicity m, the series terminates after m terms.

Ex: Find a general solution to $x' = Ax$, where $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$.

The eigenvalues of A are $\lambda_1 = 0$ and $\lambda_2 = 1$ with multiplicity two.

$u_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is an eigenvector associated with $\lambda_1 = 0$, thus

$$x_1 = e^{0t}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$ is a solution.

Let us compute the eigenvectors for $\lambda_2 = 1$:

$$A - \lambda_2 I = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus $\begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ gives $a = c$, b free, $c = 0 \Rightarrow a = 0$

Hence there is only one linearly independent eigenvector, $u_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.
Since λ_2 has multiplicity 2, a generalized eigenvector can be found by solving

$(A - \lambda_2 I)^2 u = 0$. Compute

$$(A - \lambda_2 I)^2 = \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Solve $\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow a = c$, b, c free, thus

$u = \begin{bmatrix} c \\ b \\ c \end{bmatrix} = c\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. This gives us two linearly independent generalized eigenvectors, $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. We already knew that the latter is a generalized eigenvector since it is an eigenvector. We can thus take $u_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

Now we compute

$$e^{At}u_3 = e^{\lambda_2 t}\Big[u_3 + \underbrace{t(A-\lambda_2 I)u_3 + \frac{t^2}{2!}(A-\lambda_2 I)^2 u_3}_{=0} + \underbrace{\cdots}_{=0}\Big]$$

$$= e^t\left[\begin{bmatrix}1\\0\\1\end{bmatrix} + t\begin{bmatrix}-1 & 0 & 1\\ 0 & 0 & 2\\ 0 & 0 & 0\end{bmatrix}\begin{bmatrix}1\\0\\1\end{bmatrix}\right]$$

$$= e^t\left[\begin{bmatrix}1\\0\\1\end{bmatrix} + \begin{bmatrix}0\\2t\\0\end{bmatrix}\right] = e^t\begin{bmatrix}1\\2t\\1\end{bmatrix}.$$

The general solution is

$$x(t) = c_1\begin{bmatrix}1\\0\\0\end{bmatrix} + c_2 e^t\begin{bmatrix}0\\1\\0\end{bmatrix} + c_3 e^t\begin{bmatrix}1\\2t\\1\end{bmatrix}$$

where c_1, c_2, and c_3 are arbitrary constants.

Remark. A common mistake is to forget to compute $e^{At}u$, and write the "solution" corresponding to a generalized eigenvector u as $e^{\lambda t}u$. In the above example, it would be $e^t\begin{bmatrix}1\\0\\1\end{bmatrix}$, which is not a solution.

The phase plane

Consider the system $\begin{cases} \dfrac{dx}{dt} = f(x,y,t) \\ \dfrac{dy}{dt} = f(x,y,t) \end{cases}$. When f and g do not depend explicitly on t, the system is called autonomous. We will focus on autonomous systems.

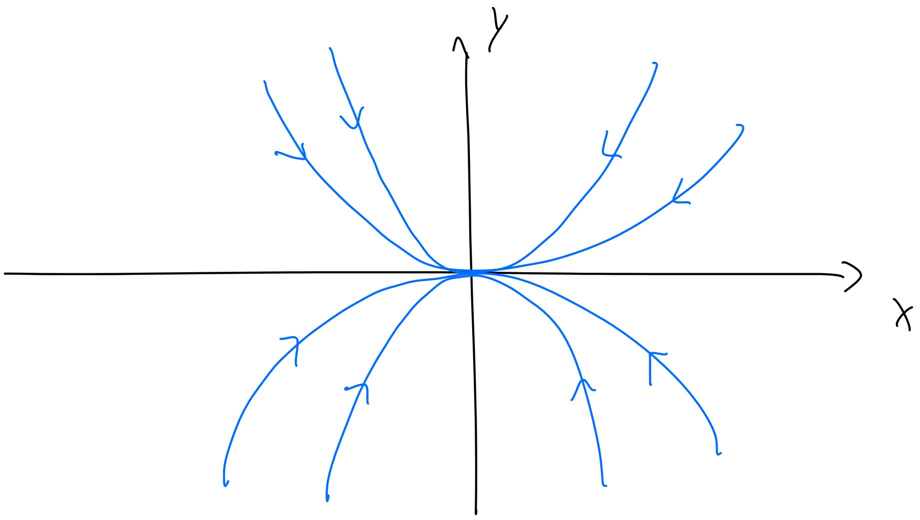
Notation. We will often denote time derivatives by a dot, i.e., $\dfrac{dx}{dt} = \dot{x}$.

Def. Consider the autonomous system $\begin{cases} \dot{x} = f(x,y) \\ \dot{y} = g(x,y) \end{cases}$. Let $(x(t), y(t))$ be a solution defined on some interval I. A plot in the xy-plane of the parametrized curve $x = x(t)$, $y = y(t)$, along with arrows indicating the direction of increasing t, is called a trajectory of the system. In this context we call the xy-plane the phase plane and a representative set of trajectories in this plane is called phase portrait of the system.

Ex: The system $\dot{x} = -2x$, $\dot{y} = -8y$ has solutions $x(t) = c_1 e^{-2t}$, $y(t) = c_2 e^{-8t}$, c_1, c_2 arbitrary constants. To draw the trajectories, we write

$$\frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{dy}{dx} = \frac{8y}{2x} \Rightarrow \frac{dy}{y} = 4\frac{dx}{x} \Rightarrow \ln|y| = \ln x^4 + C \Rightarrow y = Cx^4$$

The phase portrait is illustrated in the following figure:



The arrows indicating the direction of time (increasing t) all point toward the origin because $x(t) = c_1 e^{-2t} \to 0$ and $y(t) = c_2 e^{-8t} \to 0$ as $t \to \infty$.

To draw the phase portrait, it is useful to note that we can rewrite $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$, as $\frac{\dot{y}}{\dot{x}} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{dy}{dx} = \frac{g(x,y)}{f(x,y)}$, as in the previous example.

If $(x(t), y(t))$ is a solution to an autonomous system, so is the time-shifted pair $(x(t+c), y(t+c))$ for any constant c.

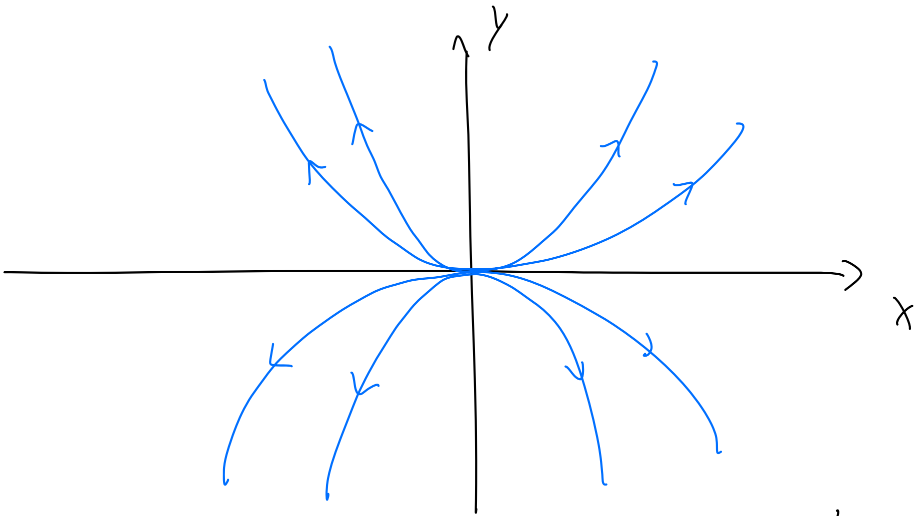
To see this, set $X(t) = x(t+c)$, $Y(t) = y(t+c)$ and note that $\dot{X}(t) = \dot{x}(t+c) = f(x(t+c), y(t+c)) = f(X(t), Y(t))$ and $\dot{Y}(t) = \dot{y}(t+c) = g(x(t+c), y(t+c)) = g(X(t), Y(t))$.

If (x_0, y_0) is a point such that $f(x_0, y_0) = 0 = g(x_0, y_0)$, then the constant function $x(t) = x_0$, $y(t) = y_0$ are solutions. This a solution that does not change over time, motivativating the following definition:

Def. Consider the system $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$. A point (x_0, y_0) such that $f(x_0, y_0) = 0 = g(x_0, y_0)$ is called a critical point of the system and the solution $x(t) = x_0$, $y(t) = y_0$ is called an equilibrium solution (or simply an equilibrium).

We are interested not only in determining the equilibria/critical points of autonomous systems, but also in studying their stability properties. For example, in the example above, $(0,0)$ is a critical point with the property that all trajectories converge to it as $t \to \infty$. Such a critical point is called asymptotically stable.

Consider now $\dot{x} = 2x$, $\dot{y} = 8y$. Then $x(t) = c_1 e^{2t}$ and $y(t) = c_2 e^{8t}$. The trajectories can be found by solving $\frac{dy}{dx} = \frac{4y}{x}$, which again gives $y = Cx^4$. The point $(0,0)$ is a critical point for this system, but now the trajectories move away from $(0,0)$ as $t \to \infty$, so the arrows are reversed as compared to the previous example:



In fact, no matter how close to the origin we start, the trajectories will move away from $(0,0)$ as t increases. Such a critical point is called unstable.

EX: (ex 3, sec. 5.4). Find the critical points of the system

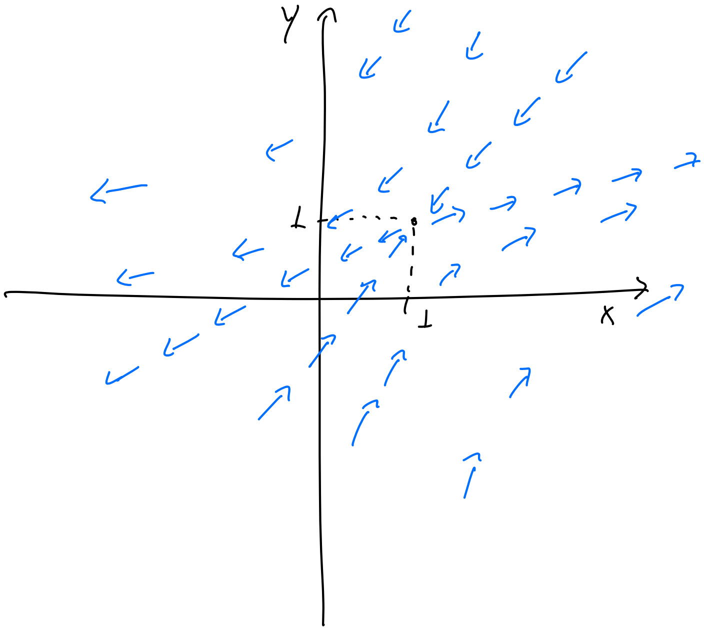
$$\dot{x} = 4x - 3y - 1$$
$$\dot{y} = 5x - 3y - 2,$$

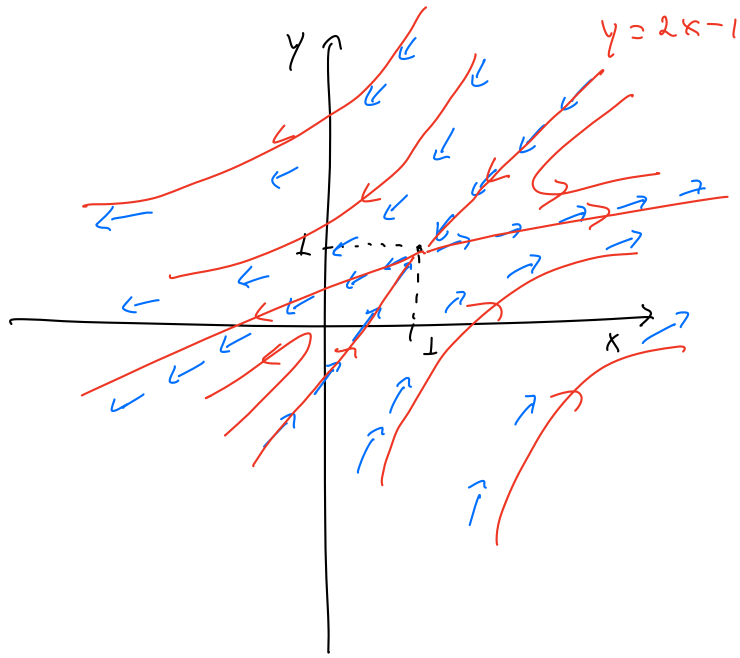
and sketch the phase portrait.

To find the critical points (x_0, y_0), we solve $f(x_0, y_0) = 0 = g(x_0, y_0)$, i.e.,

$$\left.\begin{aligned} 4x_0 - 3y_0 - 1 &= 0 \\ 5x_0 - 3y_0 - 2 &= 0 \end{aligned}\right\} \Rightarrow x_0 = 1,\ y_0 = 1.$$

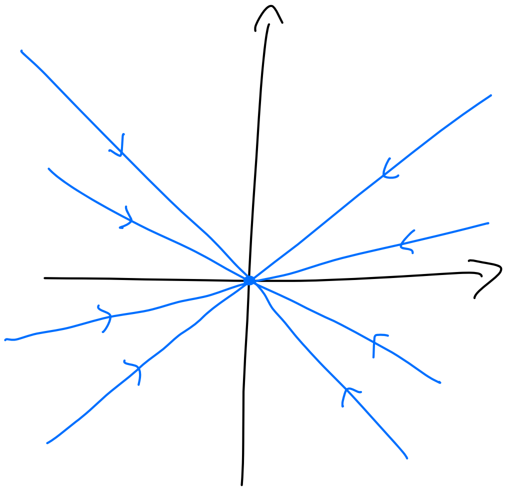
To draw the phase portrait, we sketch the direction field of the system:



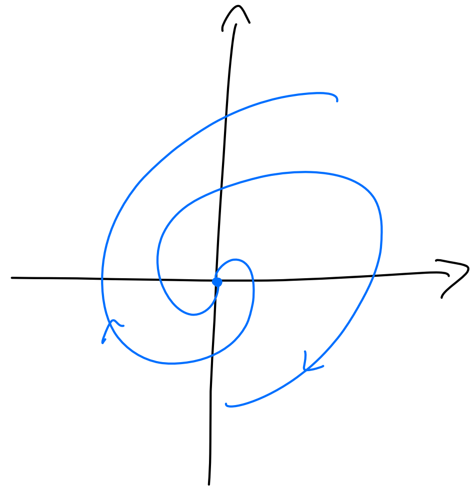


From the sketch, we see that if we start very near the critical point, trajectories will move away from it as the time passes (i.e., as t increases), with one exception: trajectories converge to the critical point along the line $y = 2x - 1$. Such a critical point is unstable (because most trajectories flow away from the critical point) and is called a <u>saddle point</u>.

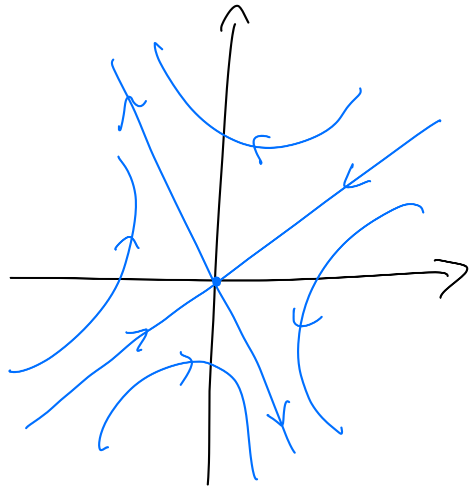
We will give the precise definitions of the critical points illustrated above, and others, later on. For now, let us simply illustrate the types of critical points we will encounter with the following pictures:



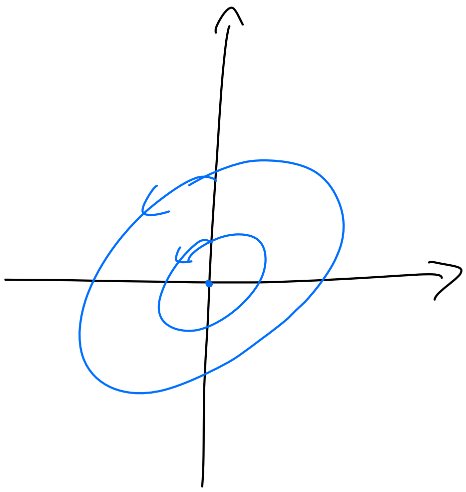
Node
(asymptotically stable)



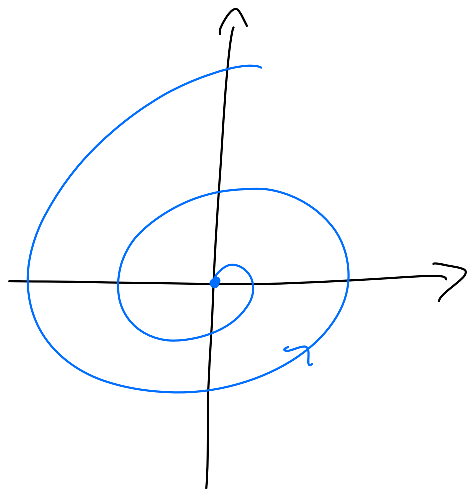
Spiral
(unstable)



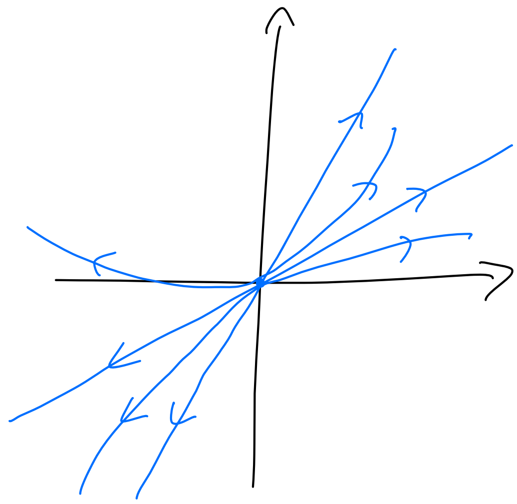
Saddle
(unstable)



Center
(stable)



Spiral
(asymptotically stable)



Node
(unstable)

When we draw phase portraits, it is useful to remember that $\frac{dy}{dx} = \frac{g(x,y)}{f(x,y)}$ admits unique solutions with initial conditions within a region R of the plane if g/f is continuously differentiable there.

Therefore, under these conditions, trajectories of the system cannot intersect.

The following theorem is useful to determine critical points:

Theo. Let $x(t), y(t)$ solve $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$, where f and g are continuous. If the limits $x_0 = \lim_{t\to\infty} x(t)$ and $y_0 = \lim_{t\to\infty} y(t)$ exist (and are finite) then (x_0, y_0) is a critical point.

EX: (sec. 12.1) Consider the system

$$\dot{x} = x(a_1 - b_1 x - c_1 y)$$

$$\dot{y} = y(a_2 - b_2 y - c_2 x)$$

where $a_1, a_2, b_1, b_2, c_1, c_2$ are positive constants. This system models the dynamics of two competing species with populations x and y.

Let us find the critical points and analyze the system. The critical points are solutions to

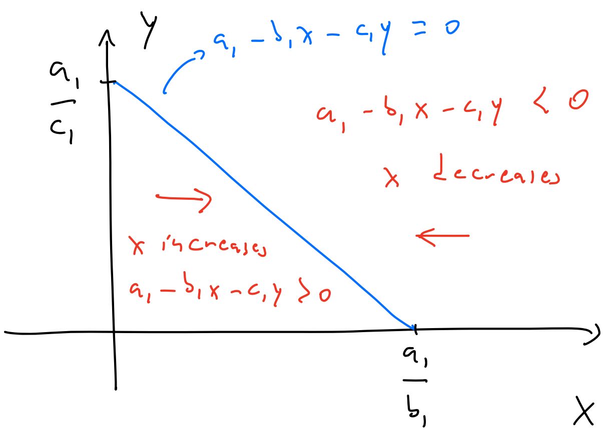
$$x(a_1 - b_1 x - c_1 y) = 0$$

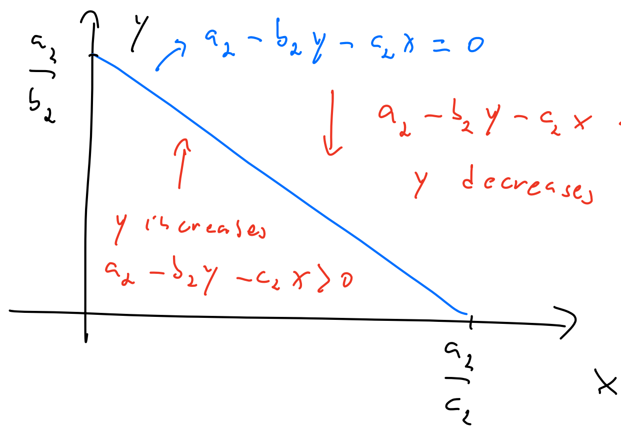
$$y(a_2 - b_2 y - c_2 x) = 0$$

There are four possibilities: $x=0$, $y=0$, or $x=0$, $a_2 - b_2 y - c_2 x = 0$, or $a_1 - b_1 x - c_1 y = 0$, $y=0$ or $a_1 - b_1 x - c_1 y = 0$, $a_2 - b_2 y - c_2 x = 0$, giving

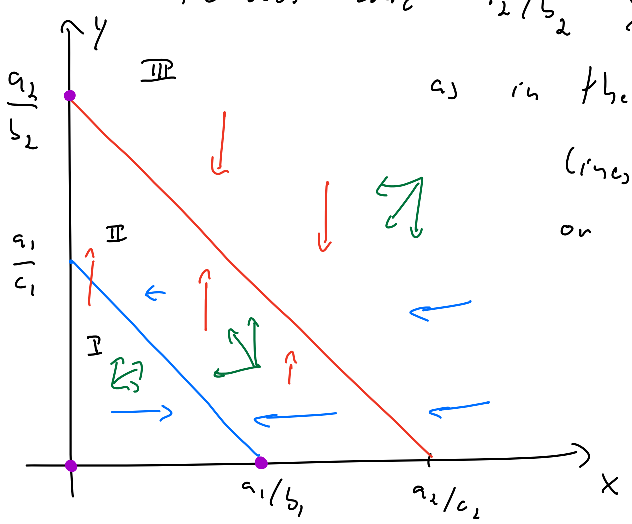
$(0,0)$, $\left(0, \frac{a_2}{b_2}\right)$, $\left(\frac{a_1}{b_1}, 0\right)$, or $\left(\frac{a_1 b_2 - a_2 c_1}{b_1 b_2 - c_1 c_2}, \frac{a_2 b_1 - a_1 c_2}{b_1 b_2 - c_1 c_2}\right)$. This last critical point is well defined for $b_1 b_2 - c_1 c_2 \neq 0$. It corresponds to the intersection of $a_1 - b_1 x - c_1 y = 0$ and $a_2 - b_2 y - c_2 x = 0$. If $b_1 b_2 - c_1 c_2 = 0$, these lines do not intersect.

To analyze the dynamics, let us indicate the regions in the phase plane (i.e., the xy-plane) where x and y increase/decrease, i.e., the regions where $\dot{x}, \dot{y}$ are positive/negative. Because $x, y \geq 0$ (as they represent populations), we consider only the first quadrant.





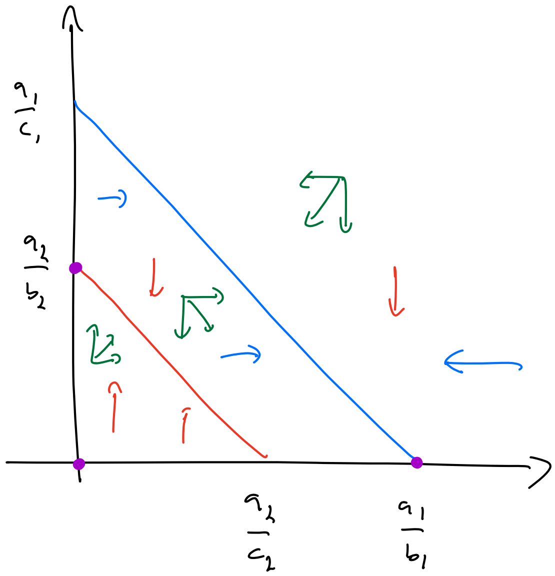
Consider first the case when $a_1 - b_1 x - c_1 y = 0$ and $a_2 - b_2 y - c_2 x = 0$ do not intersect and $a_2/b_2 > a_1/c_1$. This gives regions I, II, and III as in the picture. The blue (red) horizontal (vertical) lines indicate the directions that x (y) increases or decreases. The combined result of x, y increasing / decreasing is indicated by the green triforks. The critical points are colored in pink.



For any initial condition in region III, trajectories will move toward the line $a_2 - b_2 y - c_2 x = 0$ (red line) as indicated by the trifork. These trajectories cannot cross the x-axis because this requires $\dot{y} \neq 0$ on the x-axis, but $\dot{y} = 0$ when $y = 0$. Since y is decreasing in region III, we see that all trajectories in region III will eventually cross into region II. For initial conditions in region I, trajectories will move away from it (trifork) eventually entering into region II. For initial conditions in region II, trajectories will move up and to the left (trifork).

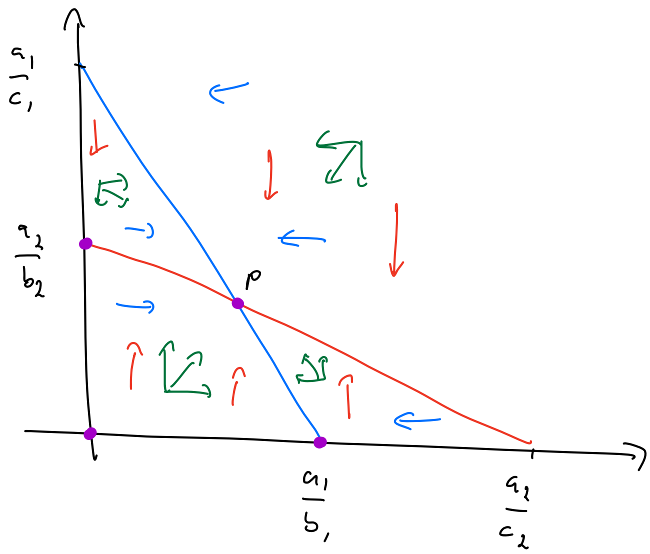
These trajectories cannot cross into region I as this would contradict the analysis we did for region I. They cannot cross into region III either, as this would require x to be increasing in region II (which it is not) or y to increase across the line $a_2 - b_2 x - c_2 y = 0$ (red line), which it cannot because y decreases in region III. We conclude that for any trajectory starting at (x_0, y_0) with $x_0 > 0$, $y_0 > 0$, it will converge to the critical point a_2/b_2. Thus, the population x will die off and the population y will approach the value a_2/b_2.

The analysis for the case when the lines $a_1 - b_1 x - c_1 y = 0$ and $a_2 - b_2 y - c_2 x = 0$ do not intersect and $a_2/b_2 < a_1/c_1$ is similar. The picture below illustrates the situation.



The conclusion in this case is that solutions starting with x_0, y_0 will converge to $\left(\frac{a_1}{b_1}, 0\right)$: the population y will die off and the population x will approach $\frac{a_1}{b_1}$.

Consider now the situation where the lines do intersect (so now there are four critical points), and let us take $a_1/c_1 > a_2/b_2$

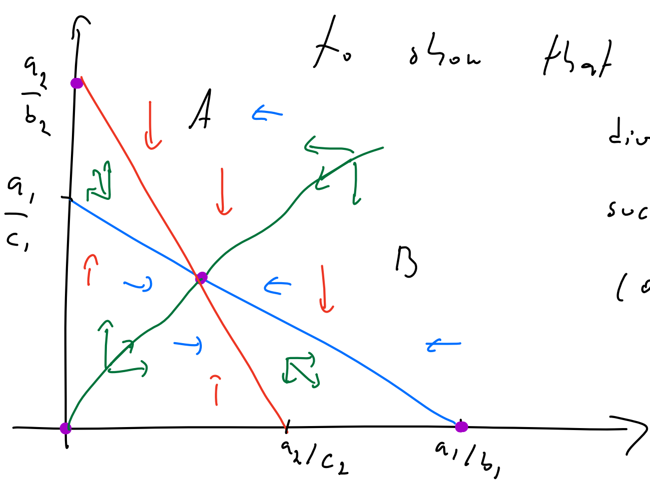


The triforks in the picture show that for any initial condition (x_0, y_0) with $x_0, y_0 > 0$, trajectories will converge to

$$p = \left(\frac{a_1 b_2 - a_2 c_1}{b_1 b_2 - c_1 c_2}, \frac{a_2 b_1 - a_1 c_2}{b_1 b_2 - c_1 c_2} \right).$$

This is an equilibrium where both species survive.

Finally, for $\frac{a_2}{b_2} > \frac{a_1}{c_1}$ with the lines intersecting, it is possible to show that there exist a line (called separatrix) dividing the plane into two regions A and B, such that trajectories starting in A approach $\left(0, \frac{a_2}{b_2}\right)$ and trajectories starting in B approach $\left(\frac{a_1}{b_1}, 0\right)$.



Linear systems in the plane

Consider the autonomous system

$$\dot{x} = a_{11} x + a_{12} y + b_1$$
$$\dot{y} = a_{12} x + a_{22} y + b_2$$

where $a_{11}, a_{12}, a_{21}, a_{22}$ and b_1, b_2 are constants. Suppose that (x_0, y_0) is a critical point for the system above. Setting $\tilde{x} = x - x_0$, $\tilde{y} = y - y_0$, we find

$$\dot{\tilde{x}} = \dot{x} = a_{11} x + a_{12} y + b_1 = a_{11} \tilde{x} + a_{12} \tilde{y} + \underbrace{a_{11} x_0 + a_{12} y_0 + b_1}_{\substack{= 0 \text{ because } (x_0, y_0) \\ \text{is a critical point}}}$$

$$\dot{\tilde{y}} = \dot{y} = a_{21} x + a_{22} y + b_2 = a_{12} \tilde{x} + a_{12} \tilde{y} + \underbrace{a_{12} x_0 + a_{22} y_0 + b_2}_{\substack{= 0 \text{ because } (x_0, y_0) \\ \text{is a critical point}}}$$

Thus, without loss of generality, we can assume that the system is written as

$$\dot{x} = a x + b y$$
$$\dot{y} = c x + d y \quad , \quad a, b, c, d \text{ constants}$$

in which case $(0,0)$ is a critical point. We will henceforth assume that $ad - bc \neq 0$, which implies that $(0,0)$ is the only critical point.

The methods previously developed give that solutions x and y are of the form $x(t) = Ae^{\lambda t}$, $y(t) = Be^{\lambda t}$, where u, v, λ are constants. Plugging in:

$$\begin{aligned} (Ae^{\lambda t})' &= aAe^{\lambda t} + bBe^{\lambda t} \\ (Be^{\lambda t})' &= cAe^{\lambda t} + dBe^{\lambda t} \end{aligned} \implies \begin{aligned} (\lambda - a)A - bB &= 0 \\ -cA + (\lambda - d)B &= 0 \end{aligned}$$

which is a system determining the eigenvalues λ and corresponding eigenvectors.

We are interested in questions of stability of the critical point $(0,0)$. E.g. Do solutions that start near $(0,0)$ remain close to $(0,0)$? If they do, do they converge to $(0,0)$ as $t \to \infty$? And if they don't, what happens when $t \to \infty$? As we will see, answers to these questions depend on the nature of the eigenvalues. We will consider separate cases.

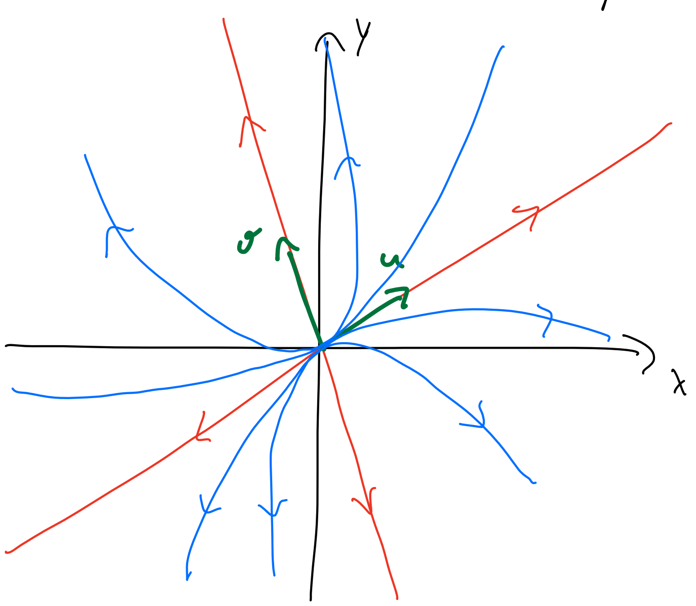
Case 1: $0 < \lambda_1 < \lambda_2$ (i.e., λ_1, λ_2 real, distinct, and positive).

In this case solutions are given by

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} e^{\lambda_1 t} + c_2 \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} e^{\lambda_2 t} = c_1 u e^{\lambda_1 t} + c_2 v e^{\lambda_2 t}$$

where c_1, c_2 are constants and $u = (u_1, u_2)$ and $v = (v_1, v_2)$ are linearly independent eigenvectors corresponding to λ_1, λ_2. (Note that such eigenvectors exist because $\lambda_1 \neq \lambda_2$). Each choice of c_1, c_2 corresponds to a different initial condition.

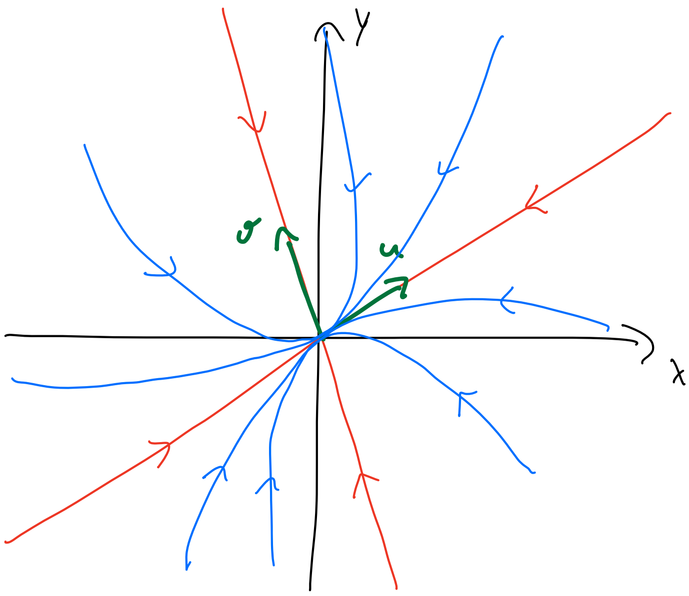
Because $\lambda_1, \lambda_2 > 0$, we see that any trajectory not starting at $(0,0)$ will move away from the origin, indicating that the critical point is unstable. We also see that for initial conditions such that $c_2 = 0$, trajectories remain on the line spanned by u, and for initial conditions such that $c_1 = 0$, trajectories remain on the line spanned by v. Furthermore, if c_1 and c_2 are both non-zero, then trajectories tend to become parallel to v when t becomes large (because $\lambda_2 > \lambda_1$). Finally, to understand what happens near $(0,0)$, we look at the limit $t \to -\infty$, because in this limit trajectories will converge to $(0,0)$ (since $\lambda_1, \lambda_2 > 0$). Because $\lambda_2 > \lambda_1$, the component of trajectories in the direction of v vanishes faster than the component in the direction of u (except when $c_1 = 0$). Therefore, unless $c_1 = 0$, trajectories become tangent to u at $(0,0)$. The phase portrait is illustrated below.



The critical point $(0,0)$ is called an unstable improper node in this case. The lines spanned by u and v are sometimes called the transformed axes.

Case 2: $\lambda_2 < \lambda_1 < 0$ (i.e., λ_1, λ_2 real, distinct, and negative).

The analysis in this case is like in case 1, but now trajectories converge to $(0,0)$ as $t \to \infty$. The origin is an asymptotically stable improper node. The phase portrait is illustrated below.



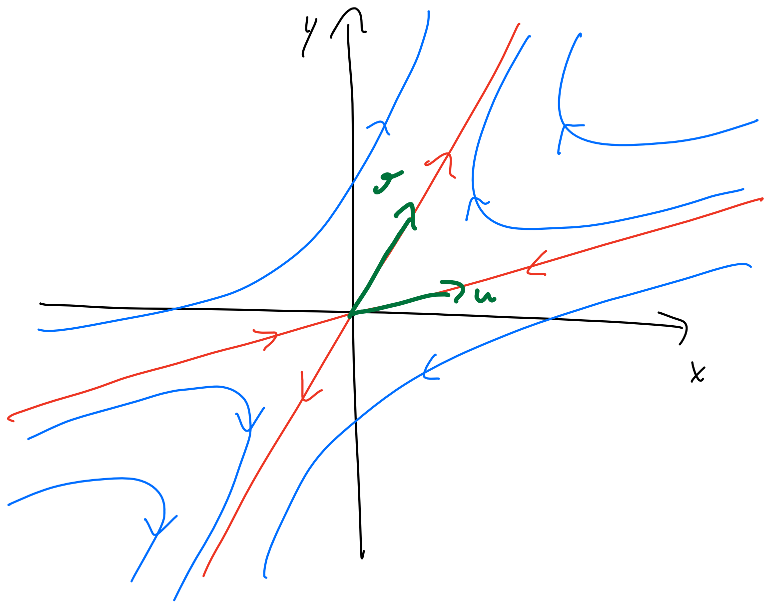
$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 u e^{\lambda_1 t} + c_2 v e^{\lambda_2 t}.$$

As before, each choice of c_1, c_2 corresponds to a different initial condition.

Case 3: $\lambda_1 < 0 < \lambda_2$ (i.e, λ_1, λ_2 real, distinct, opposite signs).

Solutions are given by $\begin{bmatrix} x \\ y \end{bmatrix} = c_1 u e^{\lambda_1 t} + c_2 v e^{\lambda_2 t}$, where u, v are linearly independent eigenvectors associated to λ_1 and λ_2 (which we know to exist because $\lambda_1 \neq \lambda_2$) and c_1, c_2 are constants, each choice of c_1, c_2 corresponding to a different initial condition. Trajectories stay on the line spanned by v if $c_1 = 0$ and on the line spanned by u if $c_2 = 0$, and they move away from $(0,0)$ along the line spanned by v (because $\lambda_2 > 0$) and toward $(0,0)$ along the line spanned by u (because $\lambda_1 < 0$).

Since $e^{\lambda_1 t} \to 0$ and $e^{\lambda_2 t} \to \infty$ as $t \to \infty$ (since $\lambda_1 < 0$ and $\lambda_2 > 0$), trajectories tend to become parallel to v for large times. Moreover, for any initial condition with $c_2 \neq 0$, trajectories will move away from the origin. The critical point $(0,0)$ is an unstable critical point. The phase portrait is illustrated below.

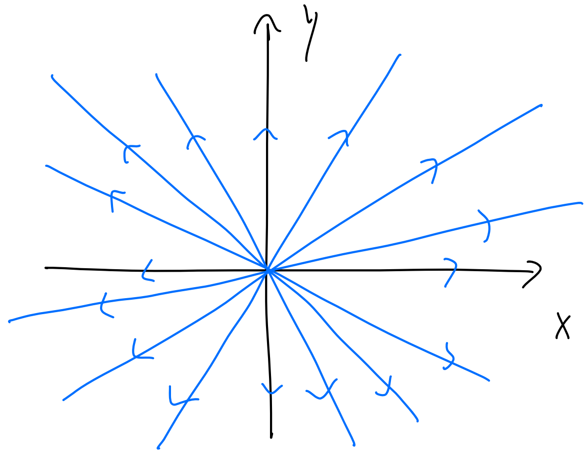


Case 4: $\lambda_1 = \lambda_2$ (i.e. λ_1, λ_2 real and equal)

Let us first consider $\lambda_1 = \lambda_2 = \lambda > 0$. If there exist two linearly independent eigenvectors u and v, then we can write

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 u e^{\lambda t} + c_2 v e^{\lambda t} = (c_1 u + c_2 v) e^{\lambda t}.$$

We see that for each c_1, c_2, not both zero, trajectories move away from $(0,0)$ along the line $c_1 u + c_2 v$. The critical point is called an unstable proper node and the phase portrait is illustrated below.



If a second linearly independent eigenvector does not exist, then solutions are written as

$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 u e^{\lambda t} + c_2 \left(v + t (A - \lambda I) v \right) e^{\lambda t}$$

$$= \Big(c_1 u + c_2 v + c_2 t \underbrace{(A - \lambda I) v}_{= w} \Big) e^{\lambda t}$$

$$= \left((c_1 u + c_2 v) + c_2 t w \right) e^{\lambda t},$$

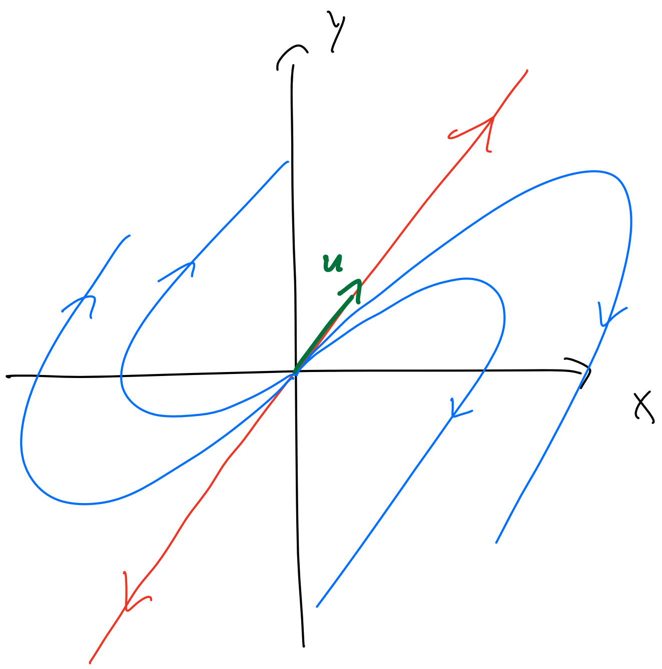
where v is a generalized eigenvector.

As $t \to \infty$, all trajectories move away from $(0,0)$. They do so along the line spanned by u for initial condition such that $c_2 = 0$. To understand the behavior of trajectories near $(0,0)$ we look at the limit $t \to -\infty$. In this limit, the term $c_2 t w e^{\lambda t}$ dominates the term $(c_1 u + c_2 v) e^{\lambda t}$ (for $c_2 \neq 0$), so trajectories tend to become parallel to w for very negative t. But we know that w is an eigenvector of A (since $(A - \lambda I) w = (A - \lambda I)^2 v = 0$)

So it must be parallel to u. But at the same time, in the limit $t \to -\infty$, trajectories converge to $(0,0)$. We conclude that trajectories must be tangent to the line spanned by u at the origin. Finally, considering the trajectories in the form $y = y(x)$, we see that along each trajectory there exists one, and only one, point x_0 such that $y_0 = y'(x_0) = 0$ and $y''(x_0) \neq 0$, thus trajectories always turn around at (x_0, y_0). Indeed, $\frac{dy}{dx} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}}$ so $y'(x) = 0$ iff $\frac{dy}{dt} = 0$. But $y(t) = (c_1 u_1 + c_2 v_2 + c_2 t w_2) e^{\lambda t}$, thus

$\dot{y}(t) = \lambda (c_1 u_1 + c_2 u_2 + c_2 t w_2 + c_2 w_2) e^{\lambda t}$ and we find one, and only one, t_0 such that $\dot{y}(t_0) = 0$. We also see that $y'(t)$ changes sign at t_0, so it increases (decreases) before (after) t_0, preventing $y''(x_0) = 0$.

The phase portrait is illustrated below. The critical point is an unstable improper node.



The case $\lambda_1 = \lambda_2 = \lambda < 0$ is analyzed in the same fashion and gives a asymptotically stable proper/improper node (it corresponds essentially to inverting the arrows in the case $\lambda > 0$).

Remark. Above, we used $w_2 \neq 0$ when we solved for t_0 to find $\dot{y}(t_0) = 0$. If $w_2 = 0$, then we consider $x = x(y)$ and find $x'(y_0) = 0$ instead, thus computing $\dot{x}(t_0) = 0$ ($w_1 \neq 0$ in this case since $w \neq 0$).

Case 5: $\lambda = \alpha \pm i\beta$, $\alpha \neq 0$, $\beta \neq 0$

In this case

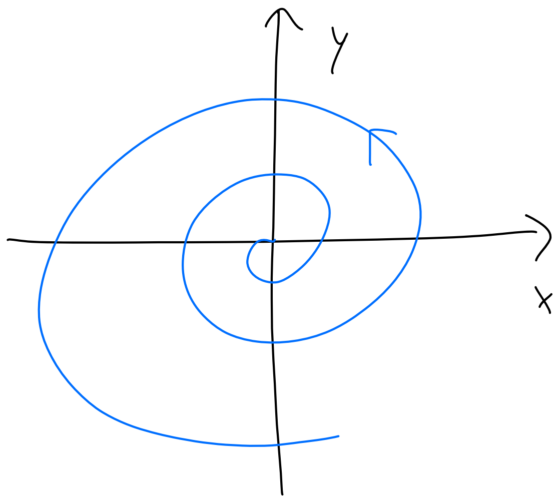
$$\begin{bmatrix} x \\ y \end{bmatrix} = c_1 e^{\alpha t}\left(\cos(\beta t)\, a - \sin(\beta t)\, b\right) + c_2 e^{\alpha t}\left(\sin(\beta t)\, a + \cos(\beta t)\, b\right)$$

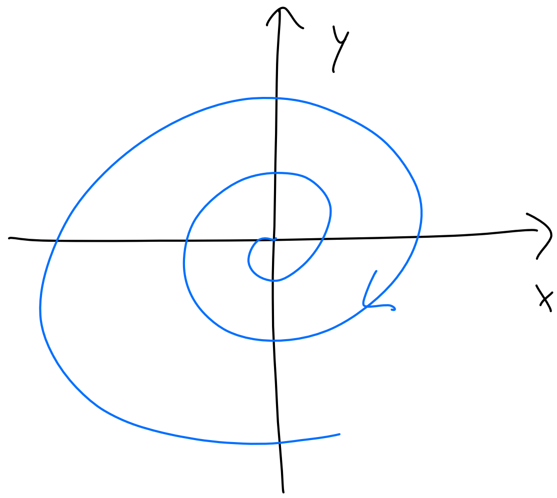
where $a \pm ib$ are eigenvectors corresponding to $\alpha \pm i\beta$. Let us write the system in polar coordinates:

$$r^2 = x^2 + y^2 = \left[c_1 e^{\alpha t}\left(\cos(\beta t)\, a_1 - \sin(\beta t)\, b_1\right) + c_2 e^{\alpha t}\left(\sin(\beta t)\, a_1 + \cos(\beta t)\, b_1\right)\right]^2$$

$$+ \left[c_1 e^{\alpha t}\left(\cos(\beta t)\, a_2 - \sin(\beta t)\, b_2\right) + c_2 e^{\alpha t}\left(\sin(\beta t)\, a_2 + \cos(\beta t)\, b_2\right)\right]^2$$

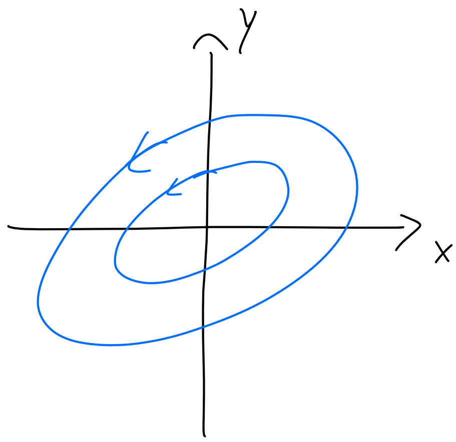
$$= e^{2\alpha t}\left[\cdots\right],$$ where $a = (a_1, a_2)$, $b = (b_1, b_2)$ and the term $[\cdots]$ is a positive function of t (unless $c_1 = c_2 = 0$). We see that $r \to \infty$ or 0 depending on whether $\alpha > 0$ or $\alpha < 0$. The periodic character of the solution also tells us that the term $[\cdots]$ causes x and y to oscillate between a positive and negative value. The critical point $(0,0)$ is an unstable spiral for $\alpha > 0$ and an asymptotically stable spiral for $\alpha < 0$.





Case 6: $\lambda = \pm i\beta$, $\beta \neq 0$

The analysis in this case is similar to case 5, but now r remains bounded. The trajectories in this case are closed and we have a stable center.



Remark. Note that we do not have any case with $\lambda = 0$ because $ad - bc \neq 0$.

Summary of stability analysis for linear systems. Consider the system $\dot{x} = Ax$, where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a real matrix, and let λ_1, λ_2 be its eigenvalues. The stability of the critical point $(0,0)$ is as follows:

Eigenvalues	Type of critical point	Stability
$0 < \lambda_1 < \lambda_2$	improper node	unstable
$\lambda_2 < \lambda_1 < 0$	improper node	asymptotically stable
$\lambda_1 < 0 < \lambda_2$	saddle point	unstable
$\lambda_1 = \lambda_2 = \lambda > 0$	proper or improper node	unstable
$\lambda_1 = \lambda_2 = \lambda < 0$	proper or improper node	asymptotically stable
$\lambda = \alpha \pm i\beta,\ \beta \neq 0$, $\alpha > 0$	spiral	unstable
$\lambda = \alpha \pm i\beta,\ \beta \neq 0$, $\alpha < 0$	spiral	asymptotically stable
$\lambda = \pm i\beta,\ \beta \neq 0$	center	stable

Above, the terms on the second and third columns are defined by the given conditions on the eigenvalues listed in the first column.

Almost linear systems

Def. Consider the autonomous system $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$. Let (x_0, y_0) be a critical point. The system is called <u>stable</u> if, given any $\varepsilon > 0$, there exists a $\delta > 0$ such that every solution $x(t), y(t)$ of the system that satisfies

$$\sqrt{(x(0) - x_0)^2 + (y(0) - y_0)^2} < \delta$$

also satisfies

$$\sqrt{(x(t) - x_0)^2 + (x(t) - y_0)^2} < \varepsilon$$

for all $t \geq 0$.

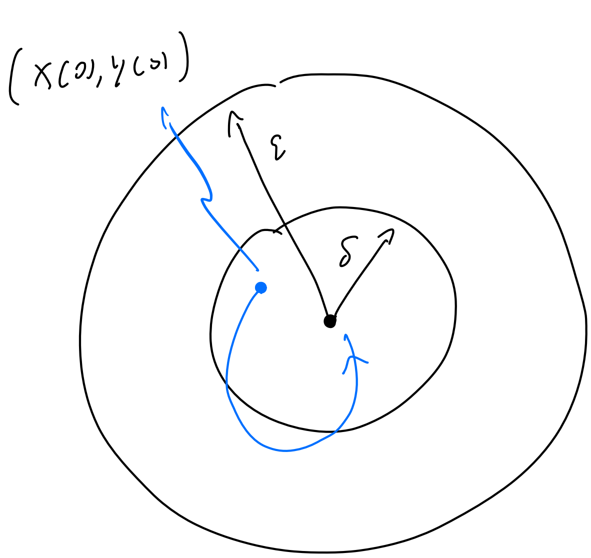
If (x_0, y_0) is stable and there exists a $\eta > 0$ such that any solution $x(t), y(t)$ that satisfies

$$\sqrt{(x(0) - x_0)^2 + (y(0) - y_0)^2} < \eta$$

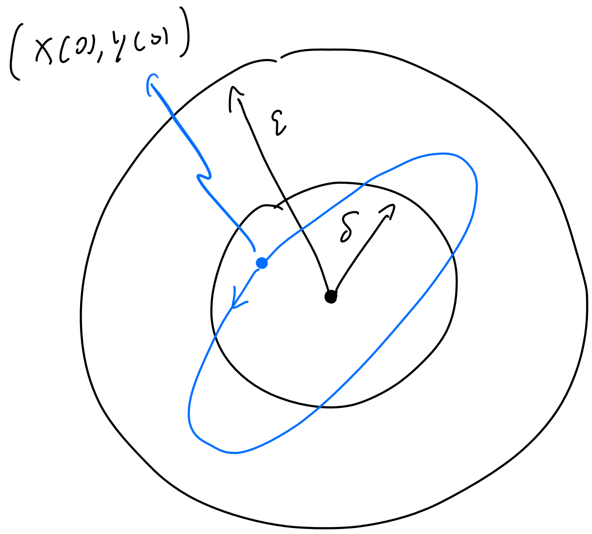
also satisfies $\lim_{t \to \infty} (x(t), y(t)) = (x_0, y_0)$, then the critical point is <u>asymptotically stable.</u> A critical point that is not stable is called <u>unstable</u>.

(Sometimes a critical point (x_0, y_0) is implicitly understood, e.g., (x_0, y_0) is the only critical point, and then we simply talk about the system being stable (unstable.)

The interpretation of this definition is as follows. A critical point is stable if any trajectory that begins near (within δ of) (x_0, y_0) stays near (within ε of) (x_0, y_0). If trajectories not only stay near but converge to (x_0, y_0) as $t \to \infty$, then the critical point is asymptotically stable.



asymptotically stable



stable, but not asymptotically stable

For linear systems, the definition of (asymptotically) stable/unstable critical points based on the eigenvalues of the system agrees with the previous definition.

<u>Ex</u>: Suppose that $(0,0)$ is a center of the linear system $\dot{x} = ax + by$, $\dot{y} = cx + dy$. Show that $(0,0)$ is stable in the sense of the above definition.

Let $x(t)$ and $y(t)$ be a solution. Then

$$(x(t)-0)^2 + (y(t)-0)^2 = (x(t))^2 + (y(t))^2$$

$$= \left[c_1\left(\cos(\beta t)a_1 - \sin(\beta t)b_1\right) + c_2\left(\sin(\beta t)a_1 + \cos(\beta t)b_1\right)\right]^2$$

$$+ \left[c_1\left(\cos(\beta t)a_2 - \sin(\beta t)b_2\right) + c_2\left(\sin(\beta t)a_2 + \cos(\beta t)b_2\right)\right]^2,$$

where $a+ib = (a_1,a_2) + i(b_1,b_2)$ is an eigenvector associated to the eigenvalue $i\beta$, $\beta \neq 0$, and c_1 and c_2 are constants. Using $AB \leq \frac{A^2}{2} + \frac{B^2}{2}$, $(A+B)^2 \leq A^2 + B^2 + 2|A||B| \leq 2(A^2+B^2)$, we have

$$\left[c_1\left(\cos(\beta t)a_1 - \sin(\beta t)b_1\right) + c_2\left(\sin(\beta t)a_1 + \cos(\beta t)b_1\right)\right]^2$$

$$\leq 2c_1^2\left(\cos(\beta t)a_1 - \sin(\beta t)b_1\right)^2 + 2c_2^2\left(\sin(\beta t)a_1 + \cos(\beta t)b_1\right)^2$$

$$\leq 4c_1^2\left(\cos^2(\beta t)a_1^2 + \sin^2(\beta t)b_1^2\right) + 4c_2^2\left(\sin^2(\beta t)a_1^2 + \cos^2(\beta t)b_1^2\right)$$

$\leq 4(c_1^2+c_2^2)(a_1^2+b_1^2)$, where we used $\cos^2(\beta t) \leq 1$, $\sin^2(\beta t) \leq 1$.

Similarly,

$$\left[c_1\left(\cos(\beta t)a_2 - \sin(\beta t)b_2\right) + c_2\left(\sin(\beta t)a_2 + \cos(\beta t)b_2\right)\right]^2,$$

$\leq 4(c_1^2+c_2^2)(a_2^2+b_2^2)$ so that

$$(x(t))^2 + (y(t))^2 \leq 4(c_1^2+c_2^2)(a_1^2+a_2^2+b_1^2+b_2^2)$$

We also have:

$$x(0) = c_1 a_1 + c_2 b_1, \quad y(0) = c_1 a_2 + c_2 b_2.$$

Solving for c_1, c_2 in terms of $x(0)$, $y(0)$

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} x(0) \\ y(0) \end{bmatrix} \Rightarrow \begin{array}{l} c_1 = \dfrac{b_2 x(0) - b_1 y(0)}{a_1 b_2 - b_1 a_2} \\[2ex] c_2 = \dfrac{-a_2 x(0) + a_1 y(0)}{a_1 b_2 - b_1 a_2} \end{array}$$

Thus:

$$(x(t))^2 + (y(t))^2 \leq 4(c_1^2 + c_2^2)(a_1^2 + a_2^2 + b_1^2 + b_2^2)$$

$$\leq 8\left((a_2^2 + b_2^2)(x(0))^2 + (a_1^2 + b_1^2)(y(0))^2\right) \frac{(a_1^2 + a_2^2 + b_1^2 + b_2^2)}{(a_1 b_2 - b_1 a_2)^2}$$

$$\leq 8 \frac{(a_1^2 + b_1^2 + a_2^2 + b_2^2)^2}{(a_1 b_2 - a_2 b_1)^2} \left((x(0))^2 + (y(0))^2\right),$$ which gives:

$$\sqrt{(x(t))^2 + (y(t))^2} \leq \sqrt{8} \frac{|a_1^2 + b_1^2 + a_2^2 + b_2^2|}{|a_1 b_2 - a_2 b_1|} \sqrt{(x(0))^2 + (y(0))^2}$$

Let $\varepsilon > 0$ be given. We want to find $\delta > 0$ such that $\sqrt{(x(t))^2 + (y(t))^2} < \varepsilon$ if $\sqrt{(x(0))^2 + (y(0))^2} < \delta$. From the above we see that this is the case if we choose $\delta < \dfrac{|a_1 b_2 - a_2 b_1|}{\sqrt{8}\, |a_1^2 + b_1^2 + a_2^2 + b_2^2|} \varepsilon$.

In practice, determining the stability/instability of non-linear systems can be very difficult. For almost linear systems, defined below, however, the stability/instability can in general be determined.

Def. Let $(0,0)$ be a critical point of the system

$$\dot{x} = ax + by + F(x,y),$$

$$\dot{y} = cx + dy + G(x,y),$$

where $a, b, c,$ and d are constants and F and G are continuous in a neighborhood of the origin. Assume that $ad - bc \neq 0$. The system is called almost linear near the origin if

$$\frac{F(x,y)}{\sqrt{x^2+y^2}} \to 0 \quad \text{and} \quad \frac{G(x,y)}{\sqrt{x^2+y^2}} \to 0 \quad \text{as } \sqrt{x^2+y^2} \to 0.$$

The assumption $ad - bc \neq 0$ implies that the corresponding linear system (obtained by setting $F = G = 0$) has only $(0,0)$ as critical point. The definition implies that $F(0,0) = 0 = G(0,0)$. Moreover, if F and G are differentiable and we write the system as $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$, then the partial derivatives of F and G vanish at $(0,0)$ and from Taylor's expansion we have that $f_x(0,0) = a$, $f_y(0,0) = b$, $g_x(0,0) = c$, $g_y(0,0) = d$. Given $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$, the linear system

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} f_x(0,0) & f_y(0,0) \\ g_x(0,0) & g_y(0,0) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

is called the linearization of the system.

Typical examples of almost linear systems involve powers of x and y.

<u>EX</u>: Show that $\dot{x} = 2x + y + x^2 + y^2$, $\dot{y} = x - y + y^3$ is an almost linear system.

We have $F(x,y) = x^2+y^2$, $G(x,y) = y^3$. We have $\frac{F(x,y)}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2} \to 0$ as $\sqrt{x^2+y^2} \to 0$. For $G(x,y)$, we have, for $|y|<1$, $|y^3| \leq y^2$, thus

$$0 \leq \frac{|G(x,y)|}{\sqrt{x^2+y^2}} \leq \frac{y^2}{\sqrt{x^2+y^2}} \leq \frac{x^2+y^2}{\sqrt{x^2+y^2}} \to 0 \text{ as } \sqrt{x^2+y^2} \to 0.$$

<u>EX</u>: Is $\dot{x} = 2x+y+x^2+y^2$, $\dot{y} = x - y + \frac{\sin\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}}$ almost linear?

No because $\frac{\sin\theta}{\theta} \to 1$ as $\theta \to 0$.

The idea of almost linear systems is that they are a perturbation of the corresponding linear system (or, if we write $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$, that the full system is a pertubation of its linearization). It is reasonable to expect that in this case the stability of the system should be the same of very similar to that of the corresponding linear system. This is the case (with one exception).

<u>Theo.</u> Consider an almost linear system and let λ_1, λ_2 be the eigenvalues of the corresponding linear system. Then the stability properties of the critical point $(0,0)$ for the almost linear system are the same as those of the corresponding linear system, with one exception: if λ_1 and λ_2 are purely imaginary then the stability of the almost linear system cannot be deduced from the corresponding linear system.

EX: Show that the system

$$\dot{x} = -2x + 2xy$$
$$\dot{y} = x - y + x^2$$

is almost linear near the origin and determine its stability.

We have $F(x,y) = 2xy$, $G(x,y) = x^2$. We find

$$0 \le \frac{|F(x,y)|}{\sqrt{x^2+y^2}} = \frac{2|x||y|}{\sqrt{x^2+y^2}} \le \frac{x^2+y^2}{\sqrt{x^2+y^2}} \to 0 \text{ as } \sqrt{x^2+y^2} \to 0$$

$$0 \le \frac{G(x,y)}{\sqrt{x^2+y^2}} = \frac{x^2}{\sqrt{x^2+y^2}} \le \frac{x^2+y^2}{\sqrt{x^2+y^2}} \to 0 \text{ as } \sqrt{x^2+y^2} \to 0$$

The corresponding linear system is $\dot{x} = -2x$, $\dot{y} = x - y$. Its eigenvalues are -2 and -1, giving an asymptotically stable improper node. By the above theorem, $(0,0)$ is an asymptotically stable critical point for the original (almost linear) system.

EX: Show that the system $\dot{x} = \sin(y-3x)$, $\dot{y} = \cos x - e^y$ is almost linear near the origin and determine its stability.

To write the system as $\dot{x} = ax + by + F(x,y)$, $\dot{y} = cx + dy + G(x,y)$, we consider the linearization. Put $f(x,y) = \sin(y-3x)$, $g(x,y) = \cos x - e^y$ and compute $f_x(0,0) = -3$, $f_y(0,0) = 1$, $g_x(0,0) = 0$, $g_y(0,0) = -1$. Then

$$\dot{x} = -3x + y + (3x - y + \sin(y-3x)) = -3x + y + F(x,y),$$

$$\dot{y} = -y + (y + \cos x - e^y) = -y + G(x,y).$$

To study the limits $\frac{F(x,y)}{\sqrt{x^2+y^2}}$ and $\frac{G(x,y)}{\sqrt{x^2+y^2}}$ as $\sqrt{x^2+y^2}$, we use Taylor's expansion:

$$\sin\theta = \theta - \frac{\theta^3}{3!} + O(\theta^5) \qquad e^y = 1 + y + \frac{y^2}{2!} + O(y^3)$$

$$\cos\theta = 1 - \frac{\theta^2}{2!} + O(\theta^4)$$

where $O(z^m)$ means terms in powers of at least z^m.

Then

$$F(x,y) = 3x - y + \sin(y-3x) = 3x - y + y - 3x - \frac{(y-3x)^3}{3!} + O((y-3x)^5)$$

We can assume that $|y-3x| < 1$ so that $O((y-3x)^5) \leq O((y-3x)^3)$ thus

$$0 \leq \frac{|F(x,y)|}{\sqrt{x^2+y^2}} \leq \frac{O(|y-3x|^3)}{\sqrt{x^2+y^2}} \leq \frac{O(|y|^3 + |x|^3)}{\sqrt{x^2+y^2}} \leq \frac{O(x^2+y^2)}{\sqrt{x^2+y^2}}$$

where we also used that we can assume $|y| < 1$, $|x| < 1$. Thus $\frac{F(x,y)}{\sqrt{x^2+y^2}} \to 0$ as $\sqrt{x^2+y^2} \to 0$. Similarly

$$G(x,y) = y + \cos x - e^y = y + 1 - \frac{x^2}{2!} + O(x^4) - \left(1 + y + \frac{y^2}{2!} + O(y^3)\right)$$

$$= -\frac{x^2}{2!} - \frac{y^2}{2!} + O(x^4) + O(y^3)$$ and arguing as above we find $\frac{G(x,y)}{\sqrt{x^2+y^2}} \to 0$ as $\sqrt{x^2+y^2} \to 0$.

The eigenvalues of the corresponding linear system are -3 and -1, giving an asymptotically stable improper node. Thus, $(0,0)$ is an asymptotically stable critical point for the original system.

Remark. Recalling that a second order DE can be written as a 2×2 first order system, we can also analyze the stability of second order DE.

Summary of stability for almost linear systems. Consider an almost linear system near the origin and let λ_1, λ_2 be the eigenvalues of the corresponding linear system. The table below summarizes the stability properties of $(0,0)$. We underlined the cases that are different than linear systems.

Eigenvalues	Type of critical point	Stability
$0 < \lambda_1 < \lambda_2$	improper node	unstable
$\lambda_2 < \lambda_1 < 0$	improper node	asymptotically stable
$\lambda_1 < 0 < \lambda_2$	saddle point	unstable
$\lambda_1 = \lambda_2 = \lambda > 0$	proper or improper node <u>or spiral</u>	unstable
$\lambda_1 = \lambda_2 = \lambda < 0$	proper or improper node <u>or spiral</u>	asymptotically stable
$\lambda = \alpha \pm i\beta,\ \beta \neq 0$, $\alpha > 0$	spiral	unstable
$\lambda = \alpha \pm i\beta,\ \beta \neq 0$, $\alpha < 0$	spiral	asymptotically stable
$\lambda = \pm i\beta,\ \beta \neq 0$	center <u>or spiral</u>	<u>indetermine</u>

Energy methods

Consider Newton's law (force = mass × acceleration):

$$F = ma = m\ddot{x}.$$

When the force $F = F(t, x, \dot{x})$ depends only on x, $F = F(x)$, the system is called conservative. In this case we define the potential energy $U = U(x)$ by $\frac{dU(x)}{dx} = -F(x)$ or

$$U(x) = -\int F(x)\,dx + K$$

where K is a constant. K is chosen according to a pre-determined convention of where one sets the value of the potential energy to be zero. (Only differences of potential energy are physically meaningful thus one is free to choose a x_{zero} such that $U(x_{zero}) = 0$).

We can now rewrite Newton's law as

$$m\ddot{x} + \frac{dU}{dx} = 0.$$

We now compute:

$$\frac{d}{dt}\left(\frac{1}{2}m(\dot{x})^2 + U\right) = m\dot{x}\ddot{x} + \frac{dU}{dx}\dot{x} = \overbrace{\left(m\ddot{x} + \frac{dU}{dx}\right)}^{=0}\dot{x} = 0,$$

showing that the quantity $E = \frac{1}{2}m(\dot{x})^2 + U(x)$ is constant during the motion. The quantity $\frac{1}{2}m(\dot{x})^2$ is called the kinetic energy of the system and E is called the total energy.

In other words, to say that E is constant means that it is conserved (hence the name conservative). Defining $g(x) = \frac{U'(x)}{m}$, we obtain

$$\ddot{x} + g(x) = 0$$

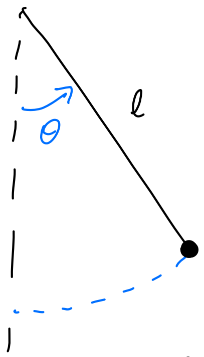
which is called the standard form of the DE for a conservative system. We can rewrite this equation as a system in the phase plane for x and $v = \dot{x}$:

$$\begin{cases} \dot{x} = v, \\ \dot{v} = -g(x). \end{cases}$$

We now introduce the potential function $G(x) = \int g(x)\,dx + C$, where C is a constant, and the energy function $E(x,v) = \frac{1}{2}v^2 + G(x)$. The constant C is chosen according to a pre-determined convention for the values where E equals to zero.

The fact that the total energy is conserved means that the level curves of $E(x,v)$, i.e., the curves in the xv-plane satisfying $E(x,v) = k$, where k is a constant, contain the phase plane trajectories of the system. (Note that different trajectories can have different energies. E.g., if (x_1, v_1) and (x_2, v_2) are two different solutions, then $E(x_1, v_1) = k_1$, $E(x_2, v_2) = k_2$, where k_1 and k_2 are constants, but in general $k_1 \neq k_2$).

__Ex:__ Consider the motion of a frictionless pendulum of length ℓ as in the figure. The evolution of the angle θ with the vertical is described by



$$\ddot{\theta} + \frac{g}{\ell} \sin\theta = 0$$

where g is the gravitational acceleration (see page 208 of the textbook for a derivation of this equation). Assume that $g/\ell = 1$. Find $E(\theta, v)$ and choose it so that $E(0,0) = 0$.

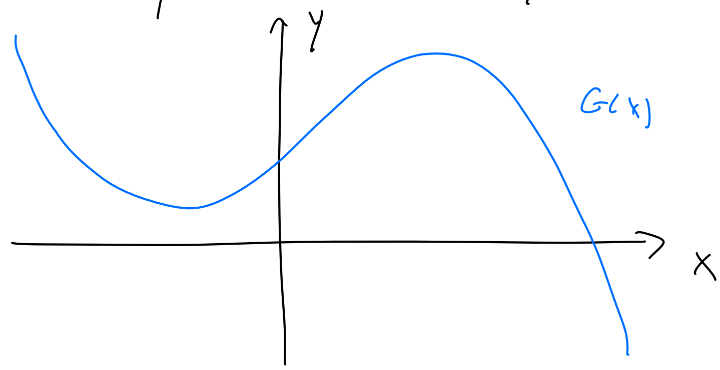
We have $g(\theta) = \sin\theta$, so $G(\theta) = -\cos\theta + C$. Thus

$$E(\theta, v) = \frac{1}{2} v^2 + C - \cos\theta.$$

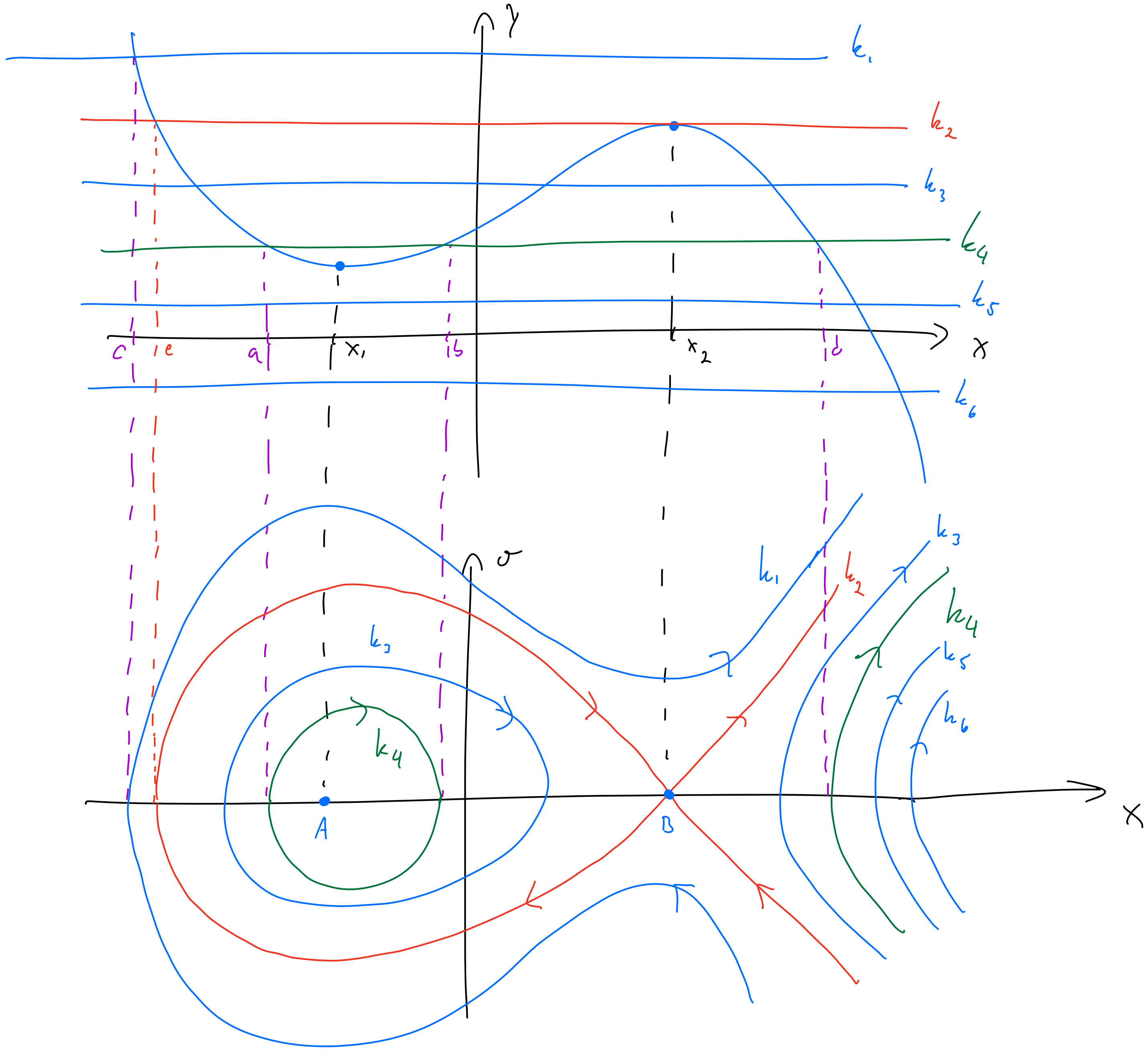
Plugging $\theta = 0 = v$ we find $E(0,0) = C - 1 = 0 \Rightarrow C = 1$, so $E(\theta, v) = \frac{1}{2} v^2 + 1 - \cos\theta$.

The critical points (x_0, v_0) of $\dot{x} = v$, $\dot{v} = -g(x)$ are given by $v_0 = 0$, $g(x_0) = 0$. So the critical points of the system are always along the x-axis, i.e., $(x_0, 0)$. Recalling that $g(x) = \frac{1}{m} \frac{dU(x)}{dx} = G'(x)$, we have $g(x_0) = 0 = \frac{1}{m} U'(x_0) = G'(x_0)$. Thus, x_0 must be a critical point (in the sense learned in calculus) of the potential function $G(x)$ and the potential energy $U(x)$. We will now see how to use this information to sketch the trajectories of the system.

Ex: Sketch the phase portrait of a conservative system with potential function whose graph is



We draw the phase plane below the graph of G and recall that the critical points of the system are those $(x_0, 0)$ with $G'(x_0) = 0$.



We see that the system has two critical points, A and B.

Let us look at the level curves of the energy function:

$$\frac{1}{2}v^2 + G(x) = k.$$

Since $v = \pm\sqrt{2(k - G(x))}$, v exists (is real-valued) only for $k - G(x) \geq 0$.

Consider a strict local minimum of G at x_1, and take a level curve $E(x,v) = k_4$, where k_4 is slightly greater than $G(x_1)$. There is an interval (a,b) containing x_1 such that such that $G(a) = k_4 = G(b)$ and $G(x) < k_4$ for $a < x < b$. Note that $v = 0$ for $x = a$ and $x = b$, and that $v = \pm\sqrt{2(k_4 - G(x))}$ is well defined and nonzero for $x \in (a,b)$ (and undefined for $x \notin [a,b]$). Thus the two curves $v = +\sqrt{2(k_4 - G(x))}$ and $v = -\sqrt{2(k_4 - G(x))}$ join at $x = a$ and $x = b$ to produce a closed curve about A. This is the case for any k such that $G(x_1) < k < k_5$, such as k_3 in the picture, where the value k_5 is indicated in the picture (red line). Hence A is a center.

For the level curves with $E(x,v) = k_4$, there is a region that corresponds to no curve because $G(x) > k_4$ there (between $x = b$ and $x = d$). But for $x \geq d$, v is well-defined, with $|v|$ increasing without bound as x increases and $v = 0$ at $x = d$. Similarly for k_3.

Consider next the strict local maximum x_2 and the level curve $E(x,v) = h_1$, with $h_1 > G(x_2)$. We see that v is not defined for $x < c$, thus the trajectory lies to the right of $x = c$. $v = 0$ only for $x = c$, i.e., the trajectory only touches the x-axis for $x = c$. For the part of the trajectory with v positive, $v = +\sqrt{2(h_1 - G(x))}$, as x varies from $x = c$ to x_1 and then from x_1 to x_2, $G(x)$ decreases and then increases, hence v increases and then decreases. From x_2 on, $G(x)$ decreases without bound. Similarly, for v negative, $v = -\sqrt{2(h_1 - G(x))}$, v decreases from $x = c$ to x_1 and increases from x_1 to x_2, decreasing again for $x > x_2$. The corresponding trajectory is illustrated above.

Consider now values h such that $h < G(x_1)$, such as h_5 and h_6 in picture. We see that trajectories exist (v is real-valued) only for x to the right of the intersection of $y = h$ with $G(x)$.

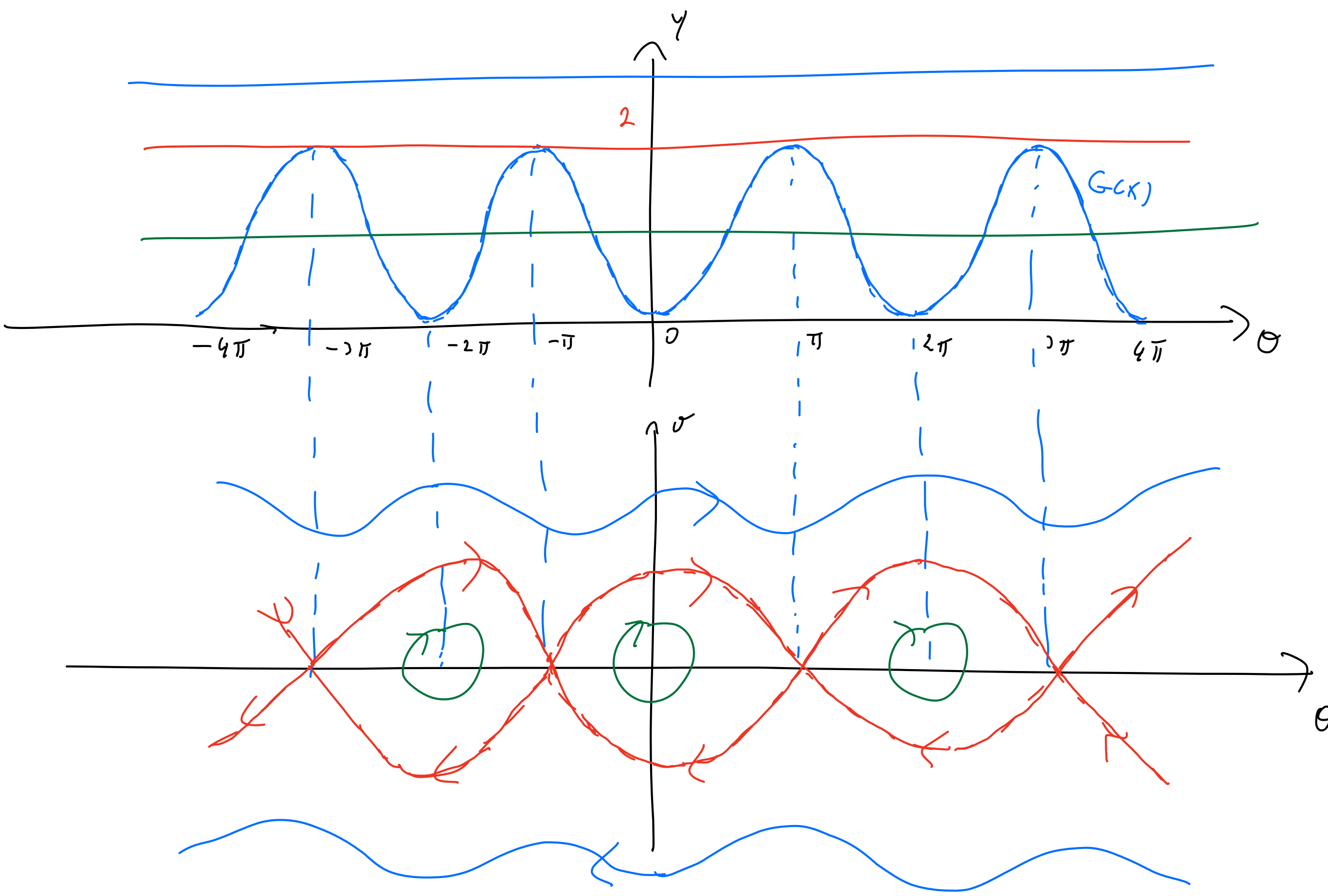
Finally, we see that h_2 corresponding to a "limiting case" separating the closed curves from the unbounded ones. We see that B is a saddle point.

Remark. To draw the arrows indicating the direction of increasing time, it is useful to recall $v = \dot{x}$, thus $v > 0$ gives that x increases along the trajectory and $v < 0$ that x decreases along the trajectory.

EX: Sketch the phase portrait of the pendulum $\ddot{\theta} + \sin\theta = 0$.

We take $E(0,0) = 0$, so $G(\theta) = 1 - \cos\theta$.

G has strict local minima at $\theta = \pm 2n\pi$, $n = 0, 1, 2, \ldots$ and strict local maxima at $\theta = \pm(2n+1)\pi$, $n = 0, 1, 2, \ldots$ Arguing as in the previous example we conclude that the critical points $(\pm 2n\pi, 0)$ are centers and $(\pm(2n+1)\pi, 0)$ are saddle points. Level curves $E(\theta, v) = k$ with $k > 2$ do not cross the θ axis and correspond to trajectories that are not closed curves.



Lyapunov's method

For conservative systems, we saw that a great deal of information can be obtained by considering the energy function $E(x,v)$. The Lyapunov method generalizes the energy method to autonomous systems $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$. In this case we no longer have an energy. Instead, we look for an appropriate function that generalizes E.

Def. Let $W = W(x,y)$ be a function that is continuous on a disk D containing $(0,0)$ and assume that $W(0,0) = 0$. We call W:

- positive definite on D if $W(x,y) > 0$ for all $(x,y) \in D$, $(x,y) \neq (0,0)$.
- positive semi-definite on D if $W(x,y) \geq 0$ for all $(x,y) \in D$.
- negative definite on D if $W(x,y) < 0$ for all $(x,y) \in D$, $(x,y) \neq (0,0)$.
- negative semi-definite on D if $W(x,y) \leq 0$ for all $(x,y) \in D$.

The theorems below apply to systems $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$ where the origin is an isolated critical point. I.e., $f(0,0) = 0 = g(0,0)$ and there exists a disk D about $(0,0)$ such that no other critical point other than $(0,0)$ exists within D.

Theo (Lyapunov's stability theorem). Let V be a positive definite function on an open disk D containing the origin. Suppose that $(0,0)$ is an isolated critical point for the system $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$. Set

$$W(x,y) = V_x(x,y) f(x,y) + V_y(x,y) g(x,y).$$

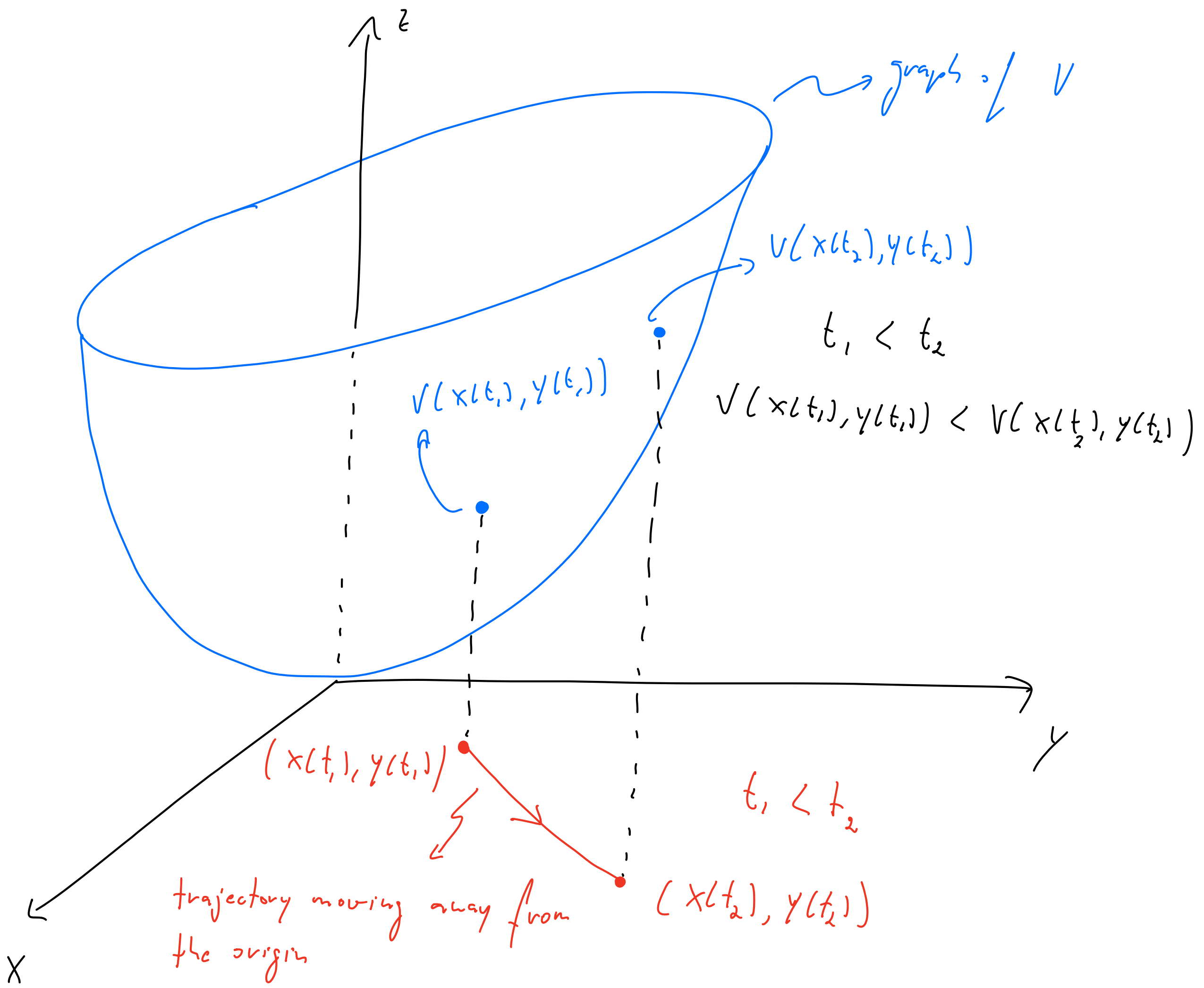
(1) If W is negative semi-definite on D, then $(0,0)$ is stable.

(2) If W is negative definite, then $(0,0)$ is asymptotically stable.

The idea behind this theorem is very simple. Let $(x(t), y(t))$ be a solution starting near the origin. Compute:

$$\begin{aligned}
\frac{d}{dt} V(x(t), y(t)) &= V_x(x(t), y(t))\, \dot{x}(t) + V_y(x(t), y(t))\, \dot{y}(t) \\
&= V_x(x(t), y(t)) f(x(t), y(t)) + V_y(x(t), y(t)) g(x(t), y(t)) \\
&= W(x(t), y(t)).
\end{aligned}$$

If W is negative semi-definite on D, this means that $\frac{d}{dt} V(x(t), y(t)) \le 0$. Hence, the function $F = F(t) = V(x(t), y(t))$ is a non-increasing function of t. On the other hand, V is positive definite, V increases when x and y move away from $(0,0)$. Now, if a trajectory $(x(t), y(t))$ were to escape the vicinity of $(0,0)$, then the function $F(t)$ would have to increase with t, contradicting the fact that $F'(t) \le 0$. Moreover, if W is negative definite, then $F(t)$ has to be strictly decreasing, and this cases trajectories cannot be closed about $(0,0)$ and the only possibility is that they converge to $(0,0)$. These ideas are illustrated in the picture below.



The function V in the theorem is called a Lyapunov function.

The main drawback of the theorem is that it gives no idea of how to find the V. However, experience shows that in many cases expressions of the form $V(x,y) = a x^l + b y^m$ with l,m positive even numbers and a,b constants appropriately chosen, produce Lyapunov functions. Note that such V's are positive definite for $a,b > 0$.

Ex: Consider $\dot{x} = -2y^3$, $\dot{y} = x - 3y^3$. We see that the origin is a critical point and in fact the only critical point since $-2y^3 = 0$, $x - 3y^3 = 0 \Rightarrow (x,y) = (0,0)$, so the critical point is isolated.

We seek a Lyapunov function in the form $V(x,y) = ax^2 + by^2$. Then

$$
\begin{aligned}
W(x,y) &= V_x(x,y) f(x,y) + V_y(x,y) g(x,y) \\
&= 2ax(-2y^3) + 2by(x-3y^3) \\
&= -4axy^3 + 2bxy - 6by^4
\end{aligned}
$$

This is not negative semi-definite because along the line $y = x$ we have:

$$W(x,-x) = -4ax^4 + 2bx^2 - 6bx^4$$

For x very small, x^4 is much smaller than x^2 and the term $+2bx^2$ dominates the remaining ones, giving $W(x,x) > 0$ (recall that $a,b > 0$).

We now try $V(x,y) = ax^2 + by^4$. Then

$$
\begin{aligned}
W(x,y) &= 2ax(-2y^3) + 4by^3(x - 3y^3) \\
&= -4axy^3 + 4by^3x - 12by^6.
\end{aligned}
$$

If we put $a = b = 1$, then $W(x,y) = -12y^6$ which is negative semi-definite. By the above theorem, $(0,0)$ is a stable critical point.

Remark. We could of course have chosen any positive $a = b$. This shows that Lyapunov functions are not unique.

Remark. The last example cannot be treated with the method of almost linear systems. This is because the corresponding linear system is $\dot{x} = 0$, $\dot{y} = x$, which does not satisfy $ad-bc \neq 0$. In other words, although $(0,0)$ is an isolated critical point of the system, it is not an isolated critical point of the corresponding linear system.

The next theorem is a criterion for instability.

Theo (Lyapunov's instability theorem). Suppose that the origin is an isolated critical point for the system $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$. Let $V = V(x,y)$ be a continuous function defined on an open disk D containing $(0,0)$ and assume that $V(0,0) = 0$. Suppose that

$$W(x,y) = V_x(x,y) f(x,y) + V_y(x,y) g(x,y)$$

is positive definite on D. Finally, assume that for every disk D' centered at the origin, there exist a $(x_0, y_0) \in D'$ such that $V(x_0, y_0) > 0$. Then $(0,0)$ is unstable.

As in the previous theorem, the main difficulty to apply this theorem consists in finding the function V.

EX! Show that $\dot{x} = -y^3$, $\dot{y} = -x^3$ is unstable using $V(x,y) = -xy$.

First, note that $(0,0)$ is an isolated critical point.

The function $V(x,y)$ is continuous, $V(0,0)$, and every disk about the origin contains a point where V is positive (any point where x and y have opposite signs). Compute:

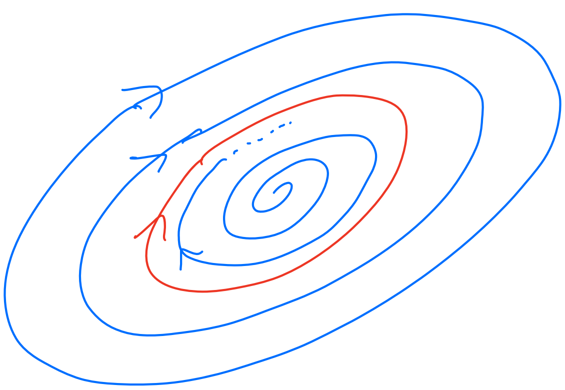
$$\begin{aligned} W(x,y) &= V_x(x,y) f(x,y) + V_y(x,y) g(x,y) \\ &= -y(-y^3) + (-x)(-x^3) \\ &= y^4 + x^4 \text{, which is positive definite.} \end{aligned}$$

Hence, $(0,0)$ is an unstable critical point.

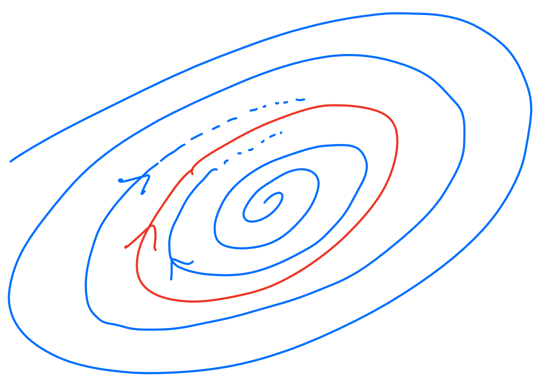
Limit cycles and periodic solutions

Def. A non-trivial closed trajectory with at least one other trajectory spiraling into it (as time approaches plus or minus infinity) is called a limit cycle. When nearby trajectories approach a limit cycle, we call it stable, and unstable when they recede. If trajectories approach a limit cycle from one side and recede from the other, it is called semi-stable.

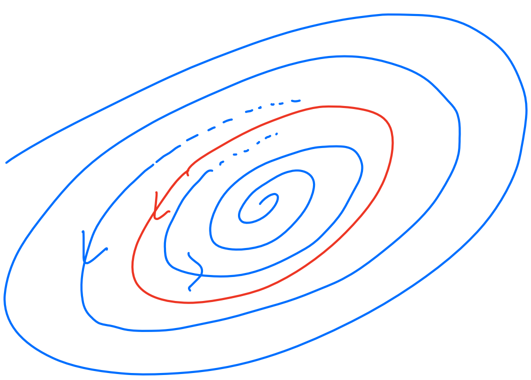
Remark. In the above definition non-trivial means not a single point (since critical points are closed trajectories).



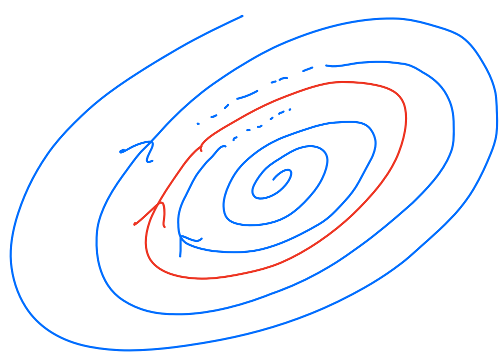
limit cycle



stable limit cycle



unstable limit cycle



semi-stable limit cycle

An important fact about limit cycles is the following:

A limit cycle must enclose at least one critical point. Moreover, any critical point enclosed by a limit cycle cannot be a saddle point.

For the next theorem, we recall that a simply connected domain in the plane is an open connected set that has no "holes"

simply connected — not simply connected — not connected (thus not simply connected)

Theo (Bendixson negative criterion). Let $f(x,y)$ and $g(x,y)$ have continuous first partial derivatives in a simply connected domain D and assume that $f_x(x,y) + g_y(x,y)$ does not change sign in D. Then there are no non-constant periodic solution to $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$ that lie entirely in D. In particular, if $D = \mathbb{R}^2$, then the system has no non-constant periodic solutions.

EX: Show that $\dot{x} = -2x - y - xy^2$, $\dot{y} = x - 3y - x^2y$ has no closed trajectories (other possibly than critical points).

Since a closed trajectory corresponds to a periodic solution, we will apply Bendixson's criterion with $D = \mathbb{R}^2$. Compute

$$f_x(x,y) + g_y(x,y) = -2 - y^2 - 3 - x^2$$

which is always negative so by Bendixson's criterion the system cannot have (non-constant) periodic solutions.

The next theorem gives a sufficient condition for the existence of periodic solutions (i.e., closed trajectories) that are not constants.

Theo (Poincaré-Bendixson theorem). Consider the system $\dot{x} = f(x,y)$, $\dot{y} = g(x,y)$, and assume that f and g have continuous partial derivatives on a closed bounded region R. Suppose that there are no critical points within R. Then any solution that stays within R for all $t \geq t_0$ for some t_0 is either a periodic solution or it approaches a limit cycle. Consequently, the system has a non-constant periodic solution.

To apply this theorem, we need to find a region R that "traps" trajectories as illustrated in the next example.

Ex: Show that the equation

$$\ddot{x} + \left(4x^2 + (\dot{x})^2 - 4\right)\dot{x} + x^3 = 0$$

has a non-constant periodic solution.

We set $y = \dot{x}$ and write the equation as the system:

$$\dot{x} = y$$

$$\dot{y} = -x^3 - (4x^2 + y^2 - 4)y$$

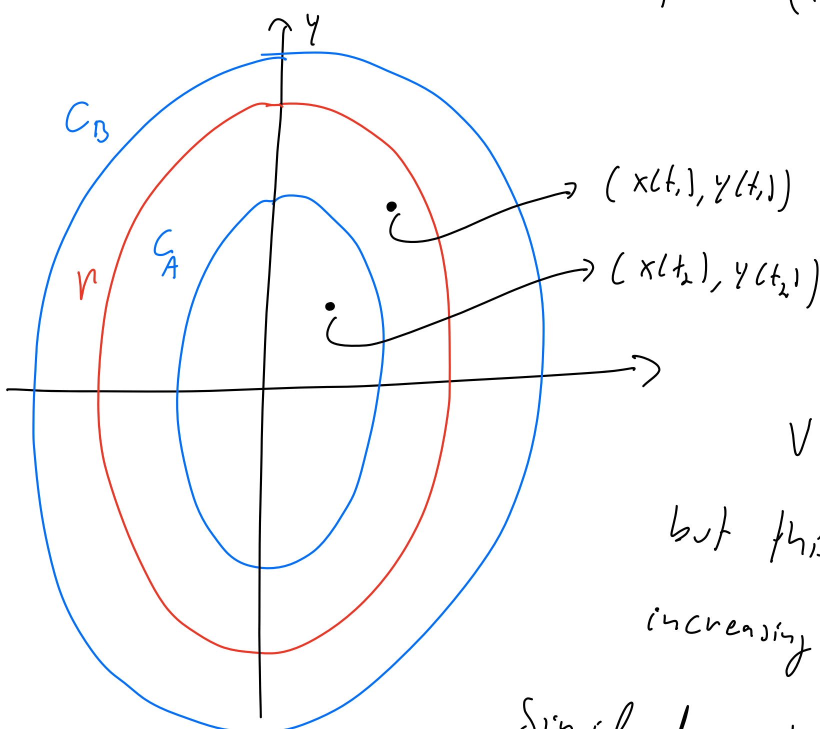
The origin is the only critical point of this system. To find the region R, we will construct a function $V(x,y)$ that increases in x and y, and such that $\frac{d}{dt} V(x(t), y(t))$ is ≥ 0 inside a curve Γ enclosing the origin and ≤ 0 outside Γ. This implies that trajectories outside Γ move toward it from the outside and trajectories inside Γ move toward it from the inside. Thus, trajectories have to remain in an annulus containing Γ (see picture).

Put $V(x,y) = \frac{1}{4}x^4 + \frac{1}{2}y^2$. Then

$$\frac{d}{dt} V(x(t), y(t)) = x^3\dot{x} + y\dot{y} = x^3 y - y\left(x^3 + (4x^2 + y^2 - 4)y\right)$$

$$= -(4x^2 + y^2 - 4)y^2$$

Thus, the function $V(x(t), y(t))$ is increasing in the variable t inside the ellipse Γ given by $4x^2+y^2=4$ (since $4x^2+y^2-4<0$ inside Γ) and decreasing outside Γ (since $4x^2+y^2-4>0$ outside Γ). Pick a number A such that the ellipse C_A given by $4x^2+y^2=A$ lies inside Γ. A trajectory starting outside C_A but inside Γ cannot cross C_A. For, suppose that $(x(t), y(t))$ lies outside C_A and inside Γ at time t_1, and inside C_A at a later time t_2. Then, since V is increasing in x, y, we would have



$$V(x(t_1), y(t_1)) > V(x(t_2), y(t_2)),$$

but this contradicts $V(x(t), y(t))$ being increasing in t inside Γ.

Similarly, choosing B such that the ellipse $4x^2+y^2=B$ is outside Γ, we conclude that a trajectory starting inside C_B but outside Γ cannot cross C_B. Thus, a trajectory that is outside C_A and inside C_B has to stay between these two curves for all future time. Taking R to be the annular region between C_A and C_B, the Poincaré-Bendixson theorem now gives the result.

Remark. We are not saying that such a closed trajectory is given by the ellipses Γ, C_A or C_B.

Stability of higher dimensional systems

We will now generalize some of the stability results discussed for 2×2 systems to $n\times n$ systems.

For $x \in \mathbb{R}^n$, instead of working with the usual norm given by $\sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$, it is convenient to define

$$\|x\| = \max_{i=1,\dots,n} |x_i|,$$

where $\max\limits_{i=1,\dots,n}$ is the maximum when i varies from 1 to n.

For a $n\times n$ matrix A we define

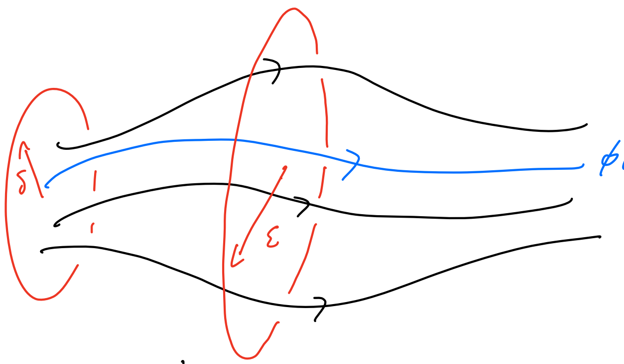
$$\|A\| = \max_{i,j=1,\dots,n} |a_{ij}|$$

where a_{ij} are the entries of A. It follows that

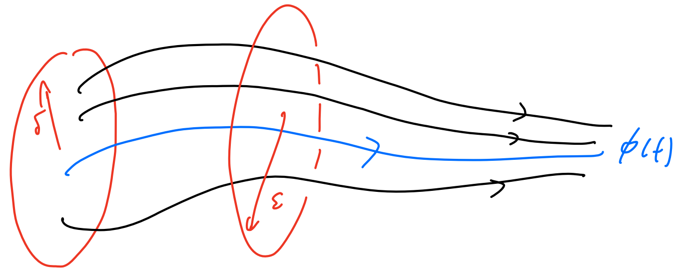
$$\|Ax\| = \max_{i=1,\dots,n} \left|\sum_{l=1}^{n} a_{il} x_l\right| \leq \max_{i=1,\dots,n} \sum_{l=1}^{n} |a_{il}|\,|x_l|$$

$$\leq \sum_{l=1}^{n} \|A\|\,|x_l| \leq \sum_{l=1}^{n} \|A\|\,\|x\| = n\|A\|\,\|x\|.$$

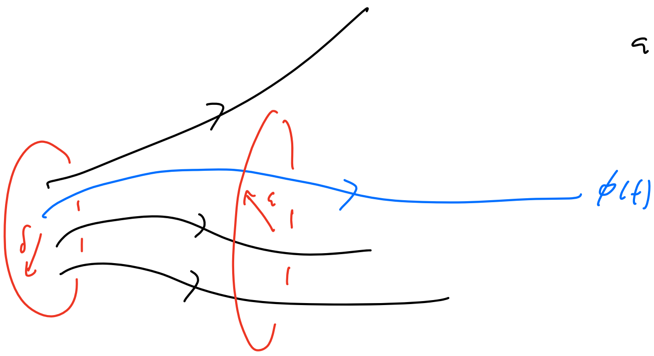
Def. Consider the system $\dot{x} = f(t,x)$, where $x = x(t)$ is a n-component vector function, $f(t,x) = (f_1(t,x), \dots, f_n(t,x))$ with each $f_i(t,x)$ a real valued function with continuous partial derivatives. We say that a solution $\phi(t)$ to this system is stable (also called Lyapunov stable) for $t \ge t_0$, if for any $\varepsilon > 0$ there exists a $\delta = \delta(t_0, \varepsilon) > 0$ such that if $\| x(t_0) - \phi(t_0) \| < \delta$, where $x(t)$ is any solution to $\dot{x} = f(t,x)$, then $\| x(t) - \phi(t) \| < \varepsilon$ for all $t \ge t_0$. If, in addition, for any such $x(t)$ we have that $\lim_{t\to\infty} \| x(t) - \phi(t) \| = 0$, then $\phi(t)$ is called asymptotically stable. If ϕ is not stable, then we call it unstable.



stable



asymptotically stable



unstable

This definition generalizes the stability of critical points previously introduced. Indeed, when the solution ϕ is a critical point and $t_0 = 0$, the above definition reduces to that of a critical point, except that the norm $\|\cdot\|$ employed is different. This is not an issue because both norms in question are equivalent, i.e., there exist constants A, B, C, and D such that

$$A\sqrt{x_1^2 + \dots + x_n^2} \le \max_{i=1,\dots,n} |x_i| \le B\sqrt{x_1^2+\dots+x_n^2} \quad \text{and} \quad C \max_{i=1,\dots,n} |x_i| \le \sqrt{x_1^2+\dots+x_n^2} \le D \max_{i=1,\dots,n} |x_i|$$

Theo. Consider $\dot{x}(t) = A(t)\,x(t) + f(t)$. A solution $\phi(t)$ is stable (asymptotically stable) if and only if the zero solution is a stable (asymptotically stable) solution to $\dot{x}(t) = A(t)\,x(t)$.

Theo. Let $A = A(t)$ be a $n \times n$ continuous matrix function. Let Φ be a fundamental matrix for the system $\dot{x} = Ax$, $t \geq t_0$. If there exists a constant $k > 0$ such that $\|\Phi(t)\| \leq k$ for all $t \geq t_0$, then the zero solution is stable. Moreover, if $\lim_{t\to\infty} \|\Phi(t)\| = 0$, the zero solution is asymptotically stable.

We conclude this brief description of higher dimensional systems with almost (higher dimensional) linear systems.

Def. Let A be a $n \times n$ matrix with non-zero determinant. Let $f = f(t,x)$ be continuously differentiable for $t \geq 0$ and $\|x\| < k$ for some $k > 0$. Suppose that $f(t,0) = 0$ for all $t \geq 0$. Assume that for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $0 < \|z\| < \delta$ implies $\|f(t,z)\|/\|z\| < \varepsilon$ for all $t \geq 0$ (i.e., $\lim_{\|z\|\to 0} \frac{\|f(t,z)\|}{\|z\|} = 0$ uniformly in t). Under these conditions, we call the system $\dot{x} = Ax + f$ almost linear.

Theo. Let $\dot{x} = Ax + f$ be an almost linear system. If all eigenvalues of A are negative, then the zero solution is asymptotically stable. If at least one eigenvalue of A has positive real part, then the zero solution is unstable.