# Math 286
# Introduction to Differential Equations

Thomas Honold

ZJU-UIUC Institute

Fall Semester 2021

# Outline

# Today's Lecture: Higher-Order Linear ODE's with Constant Coefficients

# General (time-dependent) linear ODE's

An $n$-th order linear ODE has the form

$$a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_1(t)y' + a_0(t)y = b(t) \qquad (\star)$$

with coefficient functions $a_0(t), \dots, a_n(t)$, $b(t)$, and $a_n(t) \neq 0$ for at least one $t$ (i.e., $a_n(t)$ is not the all-zero function).

Solutions of $(\star)$ are $n$-times differentiable functions $y\colon I \to \mathbb{R}$ (or $y\colon I \to \mathbb{C}$), where $I$ is an interval on which all coefficient functions are defined, satisfying $a_n(t)y^{(n)}(t) + a_{n-1}(t)y^{(n-1)}(t) + \cdots + a_1(t)y'(t) + a_0(t)y(t) = b(t)$ for all $t \in I$.

As usual, $(\star)$ is said to be *homogeneous* if $b(t) \equiv 0$ (and inhomogeneous if $b(t) \neq 0$ for at least one $t$).

## Notes

- In the homogeneous case $b(t) \equiv 0$, the real solutions of $(\star)$ with fixed domain $J$ form a subspace of $\mathbb{R}^J = \{\phi; \phi\colon J \to \mathbb{R}\}$, and similarly for the complex solutions. This is easily shown using the subspace test: The all-zero function on $J$ is a solution, and linear combinations (with constant coefficients) of solutions are again solutions.

# Notes cont'd

- We can divide ($\star$) by $a_n(t)$ and turn it into an explicit ODE, to which the Existence and Uniqueness Theorem can be applied. If $a_n(t)$ has zeros, this will generally split $I$ into two or more subintervals for which ($\star$) must be solved separately.
- In theory linear ODE's are well understood. There exists a sharpened version of the Existence and Uniqueness Theorem, asserting that an $n$-th order homogeneous linear ODE has an $n$-dimensional solution space (which is a subspace of $\mathbb{R}^I$ resp. $\mathbb{C}^I$) and that there are no obstructions to taking $I$ as large as possible. Further, a particular solution of an inhomogeneous linear ODE can be found using "variation of parameters", and its general solution can be expressed in the usual way in terms of one particular solution and the general solution of the associated homogeneous ODE.
- In practice, however, it is difficult to solve time-dependent linear ODE's of orders $n \geq 2$. There are no known formulas for computing a basis of the associated solution space. For time-independent homogeneous linear ODE's, on the other hand, a basis of the solution space can be computed "algebraically".

# The time-independent (autonomous) case

An $n$-th order linear ODE with constant coefficients (time-independent/autonomous linear ODE) has the form

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = b \tag{DE}$$

with coefficients $a_0, \dots, a_{n-1}, b \in \mathbb{C}$.

It is not necessary to consider the more general form $a_n y^{(n)} + \cdots + a_0 y = b$, $a_n \neq 0$, since we can always divide by $a_n$ to obtain the "monic" form (DE). This doesn't change anything, and neither does rewriting the ODE in explicit form $y^{(n)} = b - a_{n-1}y^{(n-1)} - \cdots - a_0 y$.

Several important physical quantities/systems can be described using ODE's of the form (DE), especially 2nd-order equations. In such applications, the left-hand side of (DE) is usually time-independent, expressing internal characteristics of the system. But $b = b(t)$ may be time-dependent, modeling the influence of an external source that changes over time. Since the basic theory of time-independent linear ODE's applies to this case as well, we will generally allow $b = b(t)$ in (DE).

## The analogy with linear recurrence relations

Recall from Discrete Mathematics (or just take it as a definition) that a linear recurrence relation of order $n$ with constant coefficients has the form

$$y_{i+n} = a_{n-1}y_{i+n-1} + \cdots + a_1 y_{i+1} + a_0 y_i + b_i, \quad i = 0, 1, 2, \dots, \quad \text{(RR)}$$

and that a solution of (RR) is a sequence $(y_0, y_1, y_2, \dots)$ satisfying (RR) for all $i \geq 0$.

In order to make the analogy with linear ODE's more visible, we replace $a_i$ by $-a_i$, write $y(i)$ in place of $y_i$, (after all, a real sequence is just a function $y\colon \mathbb{N} \to \mathbb{R}$, $i \mapsto y_i$), and rename the variable $i$ as $t$. Then (RR) becomes

$$y(t+n) + a_{n-1}y(t+n-1) + \cdots + a_1 y(t+1) + a_0 y(t) = b(t), \quad t \in \mathbb{N}.$$

Thus, compared with (DE), solutions of the recurrence relation (RR) have the "discrete" domain $\mathbb{N}$ (not a "continuous" interval $I$), and the differentiation operator $\mathrm{D}\colon y \mapsto y'$ has been replaced by the shift operator (truncation operator)

$$\mathrm{S}\colon (y_0, y_1, y_2, y_3, \dots) \mapsto (y_1, y_2, y_3, \dots).$$

Because of this striking analogy, it comes as no suprise that the methods for solving linear recurrence relations with constant coefficients and (higher-order) linear ODE's with constant coefficients are very closely related. If you know how to do one of these tasks, you will find it easy to do the other.

## Example (Fibonacci numbers)

The Fibonacci numbers are defined by the order-two recurrence relation

$$f_{i+2} = f_{i+1} + f_i, \quad f_0 = 0,\ f_1 = 1.$$

Of course we can use our brain (or another computer) to compute the Fibonacci numbers successively from this:

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_i$ | 0 | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | 55 | 89 | 144 |

But this leaves several questions open, e.g.

1. How fast do the Fibonacci numbers grow?
2. If we change the initial values $f_0$, $f_1$, how does the Fibonacci sequence change?

These questions can be answered by developing some theory, which yields a closed formula for the Fibonacci numbers as by-product.

## Example (Fibonacci numbers cont'd)

We start with the following

*Question:* How does the collection of all solutions of the recurrence relation $y_{i+2} = y_{i+1} + y_i$ (without specifying initial values) look like?

*Answer:* Since the recurrence relation is homogeneous, sums of solutions and constant multiples of solutions will be again solutions. Moreover, there exist solutions, e.g., the Fibonacci sequence and the all-zero sequence $(0, 0, 0, \dots)$.
$\Longrightarrow$ The solutions form a subspace $S$ of the vector space $\mathbb{R}^{\mathbb{N}}$ of all real sequences (with term-wise addition/scalar multiplication).

*Question:* What is the dimension of $S$?

*Answer:* When specifying a solution, we can choose $y_0 = A$, $y_1 = B$ freely, determining the rest of the sequence. Thus there are two degrees of freedom, which suggests that the dimension is 2. (But this is not a proof, of course.)
Consider the special solutions $f = (f_i)$, $g = (g_i)$ defined as follows:

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_n$ | 0 | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | 55 | 89 | 144 |
| $g_n$ | 1 | 0 | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | 55 | 89 |

## Example (Fibonacci numbers cont'd)

Then for $A, B \in \mathbb{R}$ the sequence $y = Af + Bg$, i.e. $y_i = A\,f_i + B\,g_i$, is also a solution and satisfies

$$y = (Af_0 + Bg_0, Af_1 + Bg_1, \dots) = (A, B, \dots)$$

$\Longrightarrow$ Every solution is uniquely a linear combination of $f$ and $g$.
$\Longrightarrow$ $f, g$ form a basis of $S$; in particular we have $\dim S = 2$.

*Remark:* Since $g_i = f_{i-1}$ for $i \geq 1$, we can express the general solution of $y_{i+2} = y_{i+1} + y_i$ also as $y_i = Af_i + Bf_{i-1}$, using the convention that $f_{-1} = 0$.

The answer obtained so far is not really satisfying—for example we still have no information on the growth of $(f_i)$ and other solutions of $y_{i+2} = y_{i+1} + y_i$, and how these relate to solutions of other linear recurrence relations.

## Key idea

Every homogeneous linear recurrence relation with constant coefficients in $\mathbb{R}$ has solutions of the special form $(1, r, r^2, r^3, \dots)$, i.e., $y_i = r^i$, for some $r \in \mathbb{C}$.

## Example (Fibonacci numbers cont'd)

Using the „Ansatz“ $y_i = r^i$, we get

$$y_{i+2} = y_{i+1} + y_i \iff r^{i+2} = r^{i+1} + r^i \iff r^i(r^2 - r - 1) = 0.$$

$\Longrightarrow (y_i) = (r^i)$ satisfies the recurrence relation iff $r$ is a root of the polynomial $X^2 - X - 1$. This polynomial is called *characteristic polynomial* of the recurrence relation $y_{i+2} = y_{i+1} + y_i$, and $r^2 - r - 1 = 0$ is called *characteristic equation*.

The solutions of $r^2 - r - 1 = 0$ are $r_1 = \frac{1+\sqrt{5}}{2}$, $r_2 = \frac{1-\sqrt{5}}{2}$, giving the two solutions

$$\mathbf{y}^{(1)} = \left(1, \frac{1+\sqrt{5}}{2}, \left(\frac{1+\sqrt{5}}{2}\right)^2, \left(\frac{1+\sqrt{5}}{2}\right)^3, \dots\right),$$
$$\mathbf{y}^{(2)} = \left(1, \frac{1-\sqrt{5}}{2}, \left(\frac{1-\sqrt{5}}{2}\right)^2, \left(\frac{1-\sqrt{5}}{2}\right)^3, \dots\right).$$

Since $r_1 \neq r_2$, the solutions are linearly independent (look at the first two terms of both sequences!) and form a basis of $S$.
$\Longrightarrow$ The general solution is $y_i = c_1 r_1^i + c_2 r_2^i$ with $c_1, c_2 \in \mathbb{R}$.
Plugging in $y_0 = 0$, $y_1 = 1$ gives $c_1 + c_2 = 0$, $c_1 r_1 + c_2 r_2 = 1$ (a linear system of equations for $c_1, c_2$), with solution $c_1 = \frac{1}{\sqrt{5}}$, $c_2 = -\frac{1}{\sqrt{5}}$.

## Example (Fibonacci numbers cont'd)

$\Longrightarrow$The Fibonacci numbers have the closed-form representation

$$f_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right], \qquad n = 0, 1, 2, \dots.$$

Since $\frac{1-\sqrt{5}}{2} \approx -0.62$ has absolute value $< 1$, we have

$$f_n \simeq \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n \approx \frac{1}{\sqrt{5}} \times 1.62^n \quad \text{for large } n,$$

showing that the Fibonacci numbers grow exponentially.

# Example (The "Fibonacci IVP")

By this we mean the IVP

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

Here we don't have an easy method at hand to compute a nontrivial solution, but we can observe at least the following analogy to the Fibonacci recurrence relation: For any interval $I \subseteq \mathbb{R}$ the real solutions $y\colon I \to \mathbb{R}$ of $y'' = y' + y$ form a subspace of $\mathbb{R}^I$.

*Reason:* Rewriting the ODE in the form $y'' - y' - y = 0$ shows that it is homogeneous.

$\Longrightarrow$ The all-zero function on $I$ is a solution, and for solutions $y, z\colon I \to \mathbb{R}$ the sum $y + z\colon I \to \mathbb{R}$, $t \mapsto y(t) + z(t)$, as well as any scalar multiple $cy\colon I \to \mathbb{R}$, $t \mapsto c\,y(t)$ ($c \in \mathbb{R}$) are again solutions:

$$\begin{aligned}
(y+z)'' - (y+z)' - (y+z) &= y'' + z'' - y' - z' - y - z \\
&= y'' - y' - y + (z'' - z' - z) \\
&= 0 + 0 = 0, \\
(cy)'' - (cy)' - cy &= c\,y'' - c\,y' - cy \\
&= c(y'' - y' - y) = c \cdot 0 = 0.
\end{aligned}$$

## Example (The “Fibonacci IVP” cont’d)

*Question:* What is the dimension of the solution space $S$ of $y'' - y' - y = 0$, and how to compute a basis of $S$?

## Key idea

Try functions of the form $y(t) = \mathrm{e}^{rt}$.

Because differentiation “preserves exponentials”, this might work. Indeed, for such functions $y(t)$, defined on $\mathbb{R}$, say, we have

$$\begin{aligned} y''(t) - y'(t) - y(t) &= r^2\,\mathrm{e}^{rt} - r\,\mathrm{e}^{rt} - \mathrm{e}^{rt} \\ &= (r^2 - r - 1)\mathrm{e}^{rt}, \end{aligned}$$

which is zero if $r^2 - r - 1 = 0$.

$\Longrightarrow$ $y'' - y' - y = 0$ has the same *characteristic equation*/*characteristic polynomial* as the Fibonacci recurrence relation (in the sense that this data fully characterizes the solution), and all functions

$$y(t) = c_1\mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t} + c_2\mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t}, \quad c_1, c_2 \in \mathbb{R},$$

are solutions of $y'' - y' - y = 0$.

# Example (The "Fibonacci IVP" cont'd)

Next we fit the initial conditions:

$$\begin{aligned} y(0) &= c_1 + c_2 = 0, \\ y'(0) &= c_1 \tfrac{1+\sqrt{5}}{2} + c_2 \tfrac{1-\sqrt{5}}{2} = 1. \end{aligned}$$

This system is the same as for the Fibonacci recurrence relation and was solved before.

$$\Longrightarrow \ y(t) = \frac{1}{\sqrt{5}} \left( \mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t} - \mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t} \right), \quad t \in \mathbb{R}$$

solves the Fibonacci IVP.

The solution is unique according to the Uniqueness Theorem.

But we can say more: Since for any $A, B \in \mathbb{R}$ the (linear) system

$$\begin{array}{rcrcl} c_1 & + & c_2 & = & A, \\ r_1 c_1 & + & r_2 c_2 & = & B, \end{array}$$

with $r_1 = \frac{1+\sqrt{5}}{2}$, $r_2 = \frac{1-\sqrt{5}}{2}$ can be solved for $c_1, c_2$, we can fit solutions in $S$ to any prescribed initial values $y(0) = A$, $y'(0) = B$.

## Example (The “Fibonacci IVP” cont’d)

$\Longrightarrow$ There are no further solutions (by the Uniqueness Theorem), and hence the solution space $S$ of $y'' - y' - y = 0$ is spanned by $\mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t}$, $\mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t}$ and has dimension 2.
$\Longrightarrow$ $\mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t}$, $\mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t}$ form a basis of $S$, since they generate $S$ and are linearly independent.

We also say that $\mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t}$, $\mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t}$ form a *fundamental system* of solutions of $y'' - y' - y = 0$, according to the following

## Definition

A basis of the solution space of a homogeneous linear ODE (or a homogeneous linear ODE system) is called a *fundamental system* of solutions.

## Example

Determine the general solution of

$$
\begin{aligned}
y_i &= 4y_{i-1} - 4y_{i-2} \quad \text{and} \\
y'' &= 4y' - 4y.
\end{aligned}
$$

First we rewrite the equations in standard form:

$$
\begin{aligned}
y_{i+2} - 4y_{i+1} + 4y_i &= 0, \\
y'' - 4y' + 4y &= 0.
\end{aligned}
$$

The characteristic polynomial is $X^2 - 4X + 4 = (X-2)^2$ and has only one root, viz. $r = 2$. This gives the solutions $y_i = c\,2^i$ in the discrete case and $y(t) = c\,\mathrm{e}^{2t}$ in the continuous case, but these are obviously not enough to fit all possible initial conditions.

*Question:* How to obtain further solutions?

*Answer:* Try the sequence $y_i = i\,2^i$ (for a root $r$ of multiplicity 2 in general $y_i = i\,r^i$, cf. Discrete Mathematics), respectively, the function $y(t) = t\,\mathrm{e}^{2t}$ (in general $y(t) = t\,\mathrm{e}^{rt}$). At least in the continuous case this is reasonable, because $t\,\mathrm{e}^{rt}$ when differentiated also reproduces in a way itself.

## Example (cont'd)

$$\begin{aligned} y(t) &= t\,\mathrm{e}^{2t}, \\ y'(t) &= \mathrm{e}^{2t} + 2t\,\mathrm{e}^{2t} = (1+2t)\mathrm{e}^{2t}, \\ y''(t) &= 2\,\mathrm{e}^{2t} + 2(1+2t)\mathrm{e}^{2t} = (4+4t)\mathrm{e}^{2t}, \\ \implies y'' - 4y' - 4y &= (4+4t)\mathrm{e}^{2t} - 4(1+2t)\mathrm{e}^{2t} + 4t\,\mathrm{e}^{2t} \\ &= (4+4t-4-8t+4t)\mathrm{e}^{2t} = 0. \end{aligned}$$

It works!

One can check that the solutions

$$\begin{aligned} y_i &= c_1 2^i + c_2 i\, 2^i, \\ y(t) &= c_1 \mathrm{e}^{2t} + c_2 t\,\mathrm{e}^{2t} \end{aligned}$$

can be used to fit arbitrary initial conditions ($y_0 = A$, $y_1 = B$ in the discrete case, $y(0) = A$, $y'(0) = B$ in the continuous case). Moreover, the two sequences $(2^i)$, $(i\,2^i)$, respectively, the two functions $\mathrm{e}^{2t}$, $t\,\mathrm{e}^{2t}$ are linearly independent.

$\implies$ They form a basis of the solution space in both cases.

## Example

Solve the inhomogenous recurrence relation

$$g_i = g_{i-1} + g_{i-2} + 1, \quad g_0 = g_1 = 1,$$

and the corresponding IVP $y'' = y' + y + 1$, $y(0) = y'(0) = 1$.

We do the continuous case first.
If we have two solutions $y, z$ of $y'' - y' - y = 1$ then

$$\begin{aligned}(y-z)'' - (y-z)' - (y-z) &= y'' - y' - y - (z'' - z' - z) \\ &= 1 - 1 = 0,\end{aligned}$$

so that $y - z$ solves the associated homogeneous ODE $y'' - y' - y = 0$, which is just the Fibonacci ODE.
This tells us that one particular solution $y_p$ is enough to determine the general solution:

$$y = y - y_p + y_p = \text{sol. of the hom. ODE} + y_p.$$

*Question:* How to find a particular solution?

## Example (cont'd)

*Answer:* The constant function $y(t) \equiv -1$ is clearly a solution.
$\implies$ The general solution of $y'' - y' - y = 1$ is

$$y(t) = -1 + c_1 \, \mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t} + c_2 \, \mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t}, \quad c_1, c_2 \in \mathbb{R}.$$

Fitting the initial conditions gives the linear system

$$\begin{array}{rcrcrcrcl} y(0) & = & -1 & + & c_1 & + & c_2 & = & 1, \\ y'(0) & = & & & r_1 c_1 & + & r_2 c_2 & = & 1, \end{array}$$

which is solved by $c_1 = c_2 = 1$.

$$\implies y(t) = -1 + \mathrm{e}^{\frac{1+\sqrt{5}}{2}\,t} + \mathrm{e}^{\frac{1-\sqrt{5}}{2}\,t}, \quad t \in \mathbb{R}$$

(uniquely) solves the given IVP.

# Example (cont'd)

In the discrete case we can make a table of the numbers $g_n$ and compare with the Fibonacci numbers.

| $n$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $f_n$ | 0 | 1 | 1 | 2 | 3 | 5 | 8 | 13 | 21 | 34 | 55 | 89 | 144 |
| $g_n$ | 1 | 1 | 3 | 5 | 9 | 15 | 25 | 41 | 67 | 109 | 177 | 287 | 465 |

A particular solution of $y_{i+2} - y_{i+1} - y_i = 1$ is $y_i = -1$, i.e. the sequence $y = (-1, -1, -1, \dots)$, and the general solution is therefore

$$y_i = -1 + c_1 \left(\frac{1+\sqrt{5}}{2}\right)^i + c_2 \left(\frac{1-\sqrt{5}}{2}\right)^i, \quad c_1, c_2 \in \mathbb{R}.$$

The initial conditions $y_0 = y_1 = 1$ yield the system

$$\begin{array}{rcrcl} c_1 & + & c_2 & = & 2, \\ r_1 c_1 & + & r_2 c_2 & = & 2, \end{array}$$

which is solved by $c_1 = \frac{1+\sqrt{5}}{\sqrt{5}}$, $c_2 = \frac{1-\sqrt{5}}{-\sqrt{5}}$.

## Example (cont'd)

$$\begin{aligned}\implies g_i &= -1 + \frac{2}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{i+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{i+1}\right],\\ &= 2f_{i+1} - 1, \qquad i = 0, 1, 2, \dots\end{aligned}$$

The relation $g_i = 2f_{i+1} - 1$ is also visible in the table (well, with some effort).
Using the alternative representation $y_i = -1 + Af_i + Bf_{i-1}$ (with $f_{-1} = 0$), we could have found it more quickly: $y_0 = y_1 = 1$ give $A = B = 2$ and hence $y_i = -1 + 2(f_i + 2f_{i-1}) = -1 + 2f_{i+1}$.

## Exercise

In the example we have found that the ODE $y'' - y' - y = 1$ and its discrete "analogue" $y_{i+2} - y_{i+1} - y_i = 1$ both have the constant function $y(t) \equiv -1$ as a solution (of course, with different domains $\mathbb{R}$ resp. $\mathbb{N}$). Is this a pure coincidence or an instance of a more general correspondence between the continuous and discrete case?

*Hint:* It may help to identify the discrete analogue of the exponential function $\mathrm{e}^t$ first.

# Complex Roots

## Example

Determine the general solution of $y'' + y = 0$ and $y_{i+2} + y_i = 0$ from the characteristic equation.

Here the characteristic polynomial is $X^2 + 1 = (X - \mathrm{i})(X + \mathrm{i})$, so that the general complex solutions are

$$\begin{aligned} y(t) &= c_1 \mathrm{e}^{\mathrm{i}t} + c_2 \mathrm{e}^{-\mathrm{i}t}, \quad t \in \mathbb{R}, \\ y_n &= c_1 \mathrm{i}^n + c_2 (-\mathrm{i})^n, \quad n = 0, 1, 2, \dots . \end{aligned}$$

with $c_1, c_2 \in \mathbb{C}$.

*Question:* How can we find the corresponding real solutions?

*Answer:* In the discrete case direct inspection gives that the solution can also be written as

$$y = (A, B, -A, -B, A, B, -A, -B, \dots), \quad A, B \in \mathbb{C},$$

Here we simply need to restrict $A, B$ to real numbers to obtain the general real solution.

## Example (cont'd)

In the continuous case we can argue as follows:

$$\overline{y(t)} = \overline{c}_1 \mathrm{e}^{-\mathrm{i}t} + \overline{c}_2 \mathrm{e}^{\mathrm{i}t} = y(t) \quad \text{iff} \quad \overline{c}_1 = c_2$$

(since $\mathrm{e}^{\mathrm{i}t}$ and $\mathrm{e}^{-\mathrm{i}t}$ are linearly independent).
$\Longrightarrow$ The general real solution is

$$y(t) = c\,\mathrm{e}^{\mathrm{i}t} + \overline{c}\mathrm{e}^{-\mathrm{i}t} = 2\,\mathrm{Re}(c\,\mathrm{e}^{\mathrm{i}t}), \quad c \in \mathbb{C}.$$

Setting $2c = a - b\mathrm{i}$, $a, b \in \mathbb{R}$, we see that the general real solution can also be represented as

$$\begin{aligned} y(t) &= a\,\mathrm{Re}(\mathrm{e}^{\mathrm{i}t}) - b\,\mathrm{Re}(\mathrm{i}\mathrm{e}^{\mathrm{i}t}) = a\,\mathrm{Re}(\mathrm{e}^{\mathrm{i}t}) + b\,\mathrm{Im}(\mathrm{e}^{\mathrm{i}t}) \\ &= a\cos t + b\sin t, \quad a, b \in \mathbb{R}. \end{aligned}$$

A different argument to prove this uses the observation that for a linear ODE with real coefficients the real and imaginary part of any complex solution must be solutions as well.

# A Stronger Link between the Continuous and Discrete Case

If you were already familiar with the solution methods for linear ODE's in the examples discussed so far, but not with their discrete analogues, you may have wondered where the key idea "try sequences of the form $(1, r, r^2, r^3, \dots)$" in the discrete case comes from. The correct explanation uses the concept of "eigenvectors/eigenvalues" of a linear map (operator).

## Definition

Suppose $V$ is a vector space over a field $K$ and $f\colon V \to V$ a linear map from $V$ into itself (a so-called *endomorphism* of $V$). A nonzero vector $\mathbf{v} \in V$ is said to be an *eigenvector* of $f$ if $f$ maps $\mathbf{v}$ to a scalar multiple of itself, i.e.,

$$f(\mathbf{v}) = \lambda\mathbf{v} \quad \text{for some } \lambda \in K.$$

The scalar $\lambda$ is called the corresponding *eigenvalue*.

These two terms are 50 % loanwords from German: The German adjective „eigen“ translates into "own", "self", or "proper" (for example, „Tor“ means "goal" and „Eigentor“ means "own goal").

Alternative but outdated terms, which nevertheless capture the meaning well, are *characteristic vector* and *characteristic value*.

## Observation

The function $\mathrm{e}^{rt}$, with $r \in \mathbb{C}$ arbitrary, is an eigenvector ("eigenfunction") of the differentiation operator $\mathrm{D}\colon y \mapsto y'$. Likewise, the sequence $(1, r, r^2, r^3, \dots)$ is an eigenvector ("eigensequence") of the shift operator $\mathrm{S}\colon (y_i) \mapsto (y_{i+1})$. In both cases the corresponding eigenvalue is $r$.

Of course in the continuous case you know this already: $\mathrm{D}(\mathrm{e}^{rt}) = r\,\mathrm{e}^{rt}$. In the discrete case we have likewise

$$\mathrm{S}(1, r, r^2, r^3, \dots) = (r, r^2, r^3, \dots) = r(1, r, r^2, \dots).$$

Both $\mathrm{D}$ and $\mathrm{S}$ can be iterated. Writing $\mathrm{D} \circ \mathrm{D} = \mathrm{D}^2$ ("differentiating twice"), $\mathrm{D} \circ \mathrm{D} \circ \mathrm{D} = \mathrm{D}^3$, etc., and similarly for $\mathrm{S}$, we have, e.g.,

$$\begin{aligned}
(y_{i+2} - y_{i+1} - y_i) &= (y_{i+2}) - (y_{i+1}) - (y_i) \\
&= \mathrm{S}^2(y_i) - \mathrm{S}(y_i) - (y_i) = (\mathrm{S}^2 - \mathrm{S} - 1)(y_i), \\
(\mathrm{S}^2 - \mathrm{S} - 1)(r^i) &= \mathrm{S}^2(r^i) - \mathrm{S}(r^i) - (r^i) = r^2(r^i) - r(r^i) - (r^i) \\
&= (r^2 - r - 1)(r^i),
\end{aligned}$$

and similarly

$$\begin{aligned} y'' - y' - y &= \mathrm{D}^2 y - \mathrm{D}y - y = (\mathrm{D}^2 - \mathrm{D} - 1)y, \\ (\mathrm{D}^2 - \mathrm{D} - 1)\mathrm{e}^{rt} &= \mathrm{D}^2\mathrm{e}^{rt} - \mathrm{D}\mathrm{e}^{rt} - \mathrm{e}^{rt} = r^2\,\mathrm{e}^{rt} - r\,\mathrm{e}^{rt} - \mathrm{e}^{rt} \\ &= (r^2 - r - 1)\mathrm{e}^{rt}. \end{aligned}$$

One sees that iterating both operators and taking linear combinations, which corresponds to applying a polynomial in $\mathrm{D}$ or $\mathrm{S}$ (such as $p(\mathrm{D}) = \mathrm{D}^2 - \mathrm{D} - 1$ or $p(\mathrm{S}) = \mathrm{S}^2 - \mathrm{S} - 1$ in the case of $p(X) = X^2 - X - 1$) to the function/sequence and can produce the left-hand side of any higher-order linear ODE/linear recurrence relation, for their eigenfunctions/eigensequences effectively reduces the computation to a scalar multiplication with $p(r)$, where $r$ is the corresponding eigenvalue.

This property will be crucial in the theoretical analysis of general higher-order linear ODE's (or linear recurrence relations) with constant coefficients.

We will discuss eigenvectors/eigenvalues of matrices, which form the central topic of advanced Linear Algebra, later in this course.

# The Homogeneous Case

cf. also [BDM17], Ch. 4.2

We work over $\mathbb{C}$ (because generality does not hurt at this point) and denote by $\mathbb{C}[X]$ the ring of all polynomials in one indeterminate over $\mathbb{C}$. I assume that you know how to add and multiply polynomials, and that equality in $\mathbb{C}[X]$ means coefficient-wise equality.

We will also use the fact that every nonzero polynomial $a(X) \in \mathbb{C}[X]$ splits into linear factors in $\mathbb{C}[X]$, i.e.,

$$a(X) = a_d \prod_{i=1}^{r} (X - \lambda_i)^{m_i},$$

where $d \geq 0$ is the degree of $a(X)$, $a_d \neq 0$ is the leading coefficient of $a(X)$, $\lambda_1, \dots, \lambda_r$ are the distinct roots (zeros) of $a(X)$ in $\mathbb{C}$ and $m_i \geq 1$ the corresponding multiplicities. This "prime factorization" is clearly unique, and its existence follows from the Fundamental Theorem of Algebra; cf. Calculus II.

The Fundamental Theorem of Algebra, asserting that every complex polynomial of degree $d \geq 1$ has a root in $\mathbb{C}$, is a mere existence theorem and doesn't say anything about how to actually compute the roots.

In fact, according to the ABEL-RUFFINI Theorem, the roots of a general polynomial of degree $\geq 5$ (even with integer coefficients) cannot be computed algebraically, and one needs to use numerical approximations instead.

## Definition

For an $n$-th order linear ODE

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

with constant coefficients $a_0, \dots, a_{n-1} \in \mathbb{C}$ (but possibly non-constant right-hand side $b(t)$), the polynomial $a(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in \mathbb{C}[X]$ is called its *characteristic polynomial* (and $r^n + a_{n-1}r^{n-1} + \cdots + a_1 r + a_0 = 0$ the corresponding *characteristic equation*).

*Caution:* In what follows, the letter "$r$" will have a different meaning (number of distinct roots of $a(X)$).

We first consider the homogeneous case $b(t) \equiv 0$. In this case the solutions $y \colon \mathbb{R} \to \mathbb{C}$ form a subspace of $\mathbb{C}^{\mathbb{R}}$ (since sums of solutions and linear combinations of solutions with coefficients in $\mathbb{C}$ are again solutions), and it is reasonable to conjecture that this subspace has dimension $n$ (on the basis of our examples and the Existence and Uniqueness Theorem).

## Theorem

*Suppose the characteristic polynomial $a(X)$ of*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0 \tag{H}$$

*has prime factorization $a(X) = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$. Then the functions*

$$\mathbb{R} \to \mathbb{C},\ t \mapsto t^j \mathrm{e}^{\lambda_i t}, \quad 1 \leq i \leq r,\ 0 \leq j \leq m_i - 1,$$

*form a basis of the complex solution space $S$ of (H) (a so-called fundamental system of solutions); in particular $\dim S = n$.*

## Proof.

First we note that the number of such functions is $\sum_{i=1}^{r} m_i = \deg a(X) = n$, which equals the conjectured dimension of the solution space $S$.

Hence it suffices to prove the following

1. $S$ has dimension at most $n$.
2. The functions $t^j \mathrm{e}^{\lambda_i t}$ actually solve (H), and
3. they are linearly independent.

## Proof cont'd.

(1) Suppose, by contradiction, that $\dim S > n$. Then there exist $n+1$ linearly independent solutions $\phi_1, \dots, \phi_{n+1} \colon \mathbb{R} \to \mathbb{C}$ of (H). Now we try to fit the initial conditions $y(0) = y'(0) = \cdots = y^{(n-1)}(0) = 0$ for a linear combination $y = \sum_{j=1}^{n+1} c_j \phi_j \in S$.

$$\begin{array}{rcrcccrcl}
y(0) & = & c_1\phi_1(0) & + & \dots & + & c_{n+1}\phi_{n+1}(0) & = & 0 \\
y'(0) & = & c_1\phi_1'(0) & + & \dots & + & c_{n+1}\phi_{n+1}'(0) & = & 0 \\
 & \vdots & & & & & & \vdots & \\
y^{(n-1)}(0) & = & c_1\phi_1^{(n-1)}(0) & + & \dots & + & c_{n+1}\phi_{n+1}^{(n-1)}(0) & = & 0
\end{array}$$

This is a linear system of $n$ equations for the $n+1$ unknowns $c_j$. From Linear Algebra we know that such a system must have a solution $\mathbf{c} = (c_1, \dots, c_{n+1}) \neq \mathbf{0}$. The corresponding function $y = c_1\phi_1 + \cdots + c_{n+1}\phi_{n+1}$ is not the all-zero function (since the $\phi_j$ are linearly independent), but satisfies the same initial conditions as the all-zero function. This contradicts the Uniqueness Theorem and proves (1).

# Proof cont'd.

(2) This is the most technical step. As discussed earlier, we can work with polynomial differential operators

$$p(\mathrm{D}) = p_0\mathrm{id} + p_1\mathrm{D} + \cdots + p_d\mathrm{D}^d, \quad \mathrm{D} = \frac{\mathrm{d}}{\mathrm{d}t}, \; p_i \in \mathbb{C}.$$

Such an operator acts on $y(t)$ via

$$p(\mathrm{D})y = p_0 y + p_1\mathrm{D}y + \cdots + p_d\mathrm{D}^d y = p_0 y + p_1 y' + \cdots + p_d y^{(d)},$$

and the ODE can be concisely written as $\boxed{a(\mathrm{D})y = 0.}$

The action is compatible with polynomial addition/multiplication in the following sense:

$$(p_1 + p_2)(\mathrm{D})y = \big(p_1(\mathrm{D}) + p_2(\mathrm{D})\big)y = p_1(\mathrm{D})y + p_2(\mathrm{D})y,$$
$$(p_1 p_2)(\mathrm{D})y = \big(p_1(\mathrm{D})p_2(\mathrm{D})\big)y = p_1(\mathrm{D})\big(p_2(\mathrm{D})y\big).$$

In other words, we can treat $\mathrm{D}$ like an indeterminate (it is also true that $p(\mathrm{D}) = 0$ iff $p_0 = p_1 = \cdots = p_d = 0$; cf. subsequent note), and polynomial addition/multiplication corresponds to addition/composition of the corresponding linear operators $p(\mathrm{D})$.

# Proof cont'd.

In particular we have $p_1(\mathrm{D})p_2(\mathrm{D}) = p_2(\mathrm{D})p_1(\mathrm{D})$ for any two polynomials $p_1(X), p_2(X) \in \mathbb{C}[X]$.

This commuting relation is quite useful. For example, it shows easily that the derivative of a solution is itself a solution:

$$a(\mathrm{D})y = 0 \implies a(\mathrm{D})y' = a(\mathrm{D})\mathrm{D}y = \mathrm{D}a(\mathrm{D})y = \mathrm{D}0 = 0.$$

Keep in mind that $p_1(\mathrm{D})p_2(\mathrm{D})$ is an abbreviation for the composition $p_1(\mathrm{D}) \circ p_2(\mathrm{D})$ ("first apply $p_2(\mathrm{D})$ then $p_1(\mathrm{D})$"), just like $\mathrm{D}^i$ is for $\underbrace{\mathrm{D} \circ \mathrm{D} \circ \cdots \circ \mathrm{D}}_{i \text{ times}}$. The notation $p_1(\mathrm{D})p_2(\mathrm{D})$ makes the analogy with polynomials even more visible.

Also note that composition of differential operators makes only sense after specifying a suitable domain from which $y$ is taken, which in this case is $\mathrm{C}^\infty(\mathbb{R})$, the set of all complex-valued functions $f$ on $\mathbb{R}$ that have derivatives of all orders. The somewhat sloppy notation "id" refers to the identity map with this domain, viz., $\mathrm{C}^\infty(\mathbb{R}) \to \mathrm{C}^\infty(\mathbb{R})$, $y \mapsto y$. For details see a subsequent note.

Next we generalize our observation that the exponentials $\mathrm{e}^{\lambda t}$ form eigenfunctions of $\mathrm{D}$ to arbitrary polynomials $p(\mathrm{D})$.

# Proof cont'd.

$$\mathrm{D}\mathrm{e}^{\lambda t} = \lambda \mathrm{e}^{\lambda t} \implies \mathrm{D}^i \mathrm{e}^{\lambda t} = \lambda^i \mathrm{e}^{\lambda t} \implies p(\mathrm{D})\mathrm{e}^{\lambda t} = p(\lambda)\mathrm{e}^{\lambda t}$$

This is the second useful relation and implies that $y(t) = \mathrm{e}^{\lambda_i t}$ satisfies $a(\mathrm{D})y = a(\lambda_i)y = 0$, i.e., solves the ODE.

Further we have $\mathrm{D}(f(t)\mathrm{e}^{\lambda t}) = f'(t)\mathrm{e}^{\lambda t} + \lambda f(t)\mathrm{e}^{\lambda t}$, giving

$$(\mathrm{D} - \mu\,\mathrm{id})\big(f(t)\mathrm{e}^{\lambda t}\big) = \begin{cases} f'(t)\mathrm{e}^{\lambda t} & \text{if } \mu = \lambda, \\ \big[(\lambda - \mu)f(t) + f'(t)\big]\mathrm{e}^{\lambda t} & \text{if } \mu \neq \lambda. \end{cases}$$

This is the 3rd useful relation, which we will apply to polynomials $f(t)$.

By induction, we get $(\mathrm{D} - \lambda\,\mathrm{id})^m\big(f(t)\mathrm{e}^{\lambda t}\big) = f^{(m)}(t)\mathrm{e}^{\lambda t}$, and hence

$$(\mathrm{D} - \lambda\,\mathrm{id})^m t^j \mathrm{e}^{\lambda t} = \begin{cases} m!\,\mathrm{e}^{\lambda t} & \text{if } m = j, \\ 0 & \text{if } m > j. \end{cases}$$

In particular $(\mathrm{D} - \lambda_i\,\mathrm{id})^{m_i}(t^j \mathrm{e}^{\lambda_i t}) = 0$ for $0 \leq j \leq m_i - 1$.
$\implies a(\mathrm{D})(t^j \mathrm{e}^{\lambda_i t}) = 0$, since $a(\mathrm{D})$ is a multiple of $(\mathrm{D} - \lambda_i\,\mathrm{id})^{m_i}$.
Thus the functions $t \mapsto t^j \mathrm{e}^{\lambda_i t}$, $1 \leq i \leq r$, $0 \leq j \leq m_i - 1$, solve the ODE, completing the proof of (2).

# Note on this part of the proof

The idea behind it is that

$$a(\mathrm{D})y = (\mathrm{D} - \lambda_1 \mathrm{id})^{m_1}(\mathrm{D} - \lambda_2 \mathrm{id})^{m_2} \cdots (\mathrm{D} - \lambda_r \mathrm{id})^{m_r} y$$

can be computed by applying $n$-times an operator of the simple form $\mathrm{D} - \mu\,\mathrm{id}$ with $\mu \in \mathbb{C}$, which acts like this: $(\mathrm{D} - \mu\,\mathrm{id})y = \mathrm{D}y - \mu y = y' - \mu y$. The order in which these operators are applied does not matter, because polynomial multiplication is commutative.

In the following example we write $\mathrm{D} - \mu$ for $\mathrm{D} - \mu\,\mathrm{id}$ (i.e., the identity map is simply denoted by 1). We have used this abbreviation before (when discussing $\mathrm{D}$ and $\mathrm{S}$ together), and it makes the formulas look a little less cluttered.

$y'' + y = (\mathrm{D}^2 + 1)y = (\mathrm{D} + \mathrm{i})(\mathrm{D} - \mathrm{i})y$ can be computed as the composition of $\mathrm{D} + \mathrm{i}$ and $\mathrm{D} - \mathrm{i}$ in either order. Here is one:

$$\begin{aligned}(\mathrm{D} - \mathrm{i})y &= y' - \mathrm{i}y,\\ (\mathrm{D} + \mathrm{i})(y' - \mathrm{i}y) &= (y' - \mathrm{i}y)' + \mathrm{i}(y' - \mathrm{i}y)\\ &= y'' - \mathrm{i}y' + \mathrm{i}y' - \mathrm{i}^2 y = y'' + y, \quad \text{as asserted.}\end{aligned}$$

# Proof cont'd.

(3) A linear dependency relation among the functions $t^j \mathrm{e}^{\lambda_i t}$ amounts to the existence of polynomials $f_i(X) \in \mathbb{C}[X]$ with $\deg f_i(X) \leq m_i - 1$, not all zero, and such that

$$f_1(t)\mathrm{e}^{\lambda_1 t} + f_2(t)\mathrm{e}^{\lambda_2 t} + \cdots + f_r(t)\mathrm{e}^{\lambda_r t} = 0 \quad \text{for all } t \in \mathbb{R}.$$

Write $a(X) = (X - \lambda_1)^{m_1} A_1(X)$, i.e., $A_1(X)$ is the product of all polynomials $(X - \lambda_i)^{m_i}$ with $i \geq 2$.
Since $(\mathrm{D} - \lambda_i \,\mathrm{id})^{m_i}\bigl(f_i(t)\mathrm{e}^{\lambda_i t}\bigr) = 0$, we have $A_1(\mathrm{D})\bigl(f_i(t)\mathrm{e}^{\lambda_i t}\bigr) = 0$ for $i \geq 2$ and hence

$$A_1(\mathrm{D})\bigl(f_1(t)\mathrm{e}^{\lambda_1 t}\bigr) = 0 \quad \text{for all } t \in \mathbb{R}.$$

But each factor $\mathrm{D} - \lambda_i \mathrm{id}$ of $A_1(\mathrm{D})$ preserves the degree of $f_1(X)$ in the product $f_1(t)\mathrm{e}^{\lambda_1 t}$ and hence acts as a bijection on the space consisting of all such functions.

$$\Longrightarrow \; f_1(t)\mathrm{e}^{\lambda_1 t} = 0 \; (t \in \mathbb{R}) \Longrightarrow \; f_1(t) = 0 \; (t \in \mathbb{R}) \; \Longrightarrow \; f_1(X) = 0.$$

The last implication uses the fact that polynomials have only finitely many zeros.
In the same way one shows that $f_2(X) = \cdots = f_r(X) = 0$. $\square$

# Notes

- An analogous theorem holds in the discrete case. The solution space of a homogeneous linear recurrence relation of order $n$ with constant coefficients and characteristic polynomial $a(X) = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$ has a basis consisting of the sequences
$$(k^j \lambda_i^k)_{k\in\mathbb{N}}, \quad 1 \leq i \leq r,\ 0 \leq j \leq m_i - 1.$$
The proof given in the continuous case remains valid—just replace everywhere $\mathrm{D}$ by $\mathrm{S}$ (and functions by sequences, of course).
- The relation $\mathrm{D}f = \lambda f$ is equivalent to $(\mathrm{D} - \lambda\mathrm{id})f = 0$, i.e., to the ODE $y' - \lambda y = 0$ for $f$. The set of all (differentiable) functions $f\colon \mathbb{R} \to \mathbb{C}$ satisfying this relation, the so-called *eigenspace* of $\mathrm{D}$ corresponding to the eigenvalue $\lambda$, is 1-dimensional and spanned by $\mathrm{e}^{\lambda t}$. The set of all functions $f\colon \mathbb{R} \to \mathbb{C}$ satisfying $(\mathrm{D} - \lambda\mathrm{id})^m f = 0$ for some integer $m \geq 1$ is called *generalized eigenspace* of $\mathrm{D}$ corresponding to $\lambda$. In the proof of the theorem we have seen that this space consists precisely of the polynomial multiples $p(t)\mathrm{e}^{\lambda t}$, $p(X) \in \mathbb{C}[X]$.

# Notes cont'd

- In the proof of the theorem we have tacitly used that the Existence and Uniqueness Theorem holds also for complex ODE systems and higher-order ODE's. This can be seen as follows: Apply reduction of order $z_1 = z$, $z_2 = z'$, ..., $z_n = z^{(n-1)}$ to reduce a complex $n$-th order ODE $z^{(n)} = f(t, z, z', \dots, z^{(n-1)})$ to a complex 1st-order system $z_k' = f_k(t, z_1, \dots, z_n)$, $1 \leq k \leq n$. Then, writing $z_k = x_k + \mathrm{i}y_k$ and using $z_k' = x_k' + \mathrm{i}y_k'$, we see that this system is equivalent to

$$\begin{aligned} x_k' &= \operatorname{Re} f_k(t, x_1 + \mathrm{i}y_1, \dots, x_n + \mathrm{i}y_n), && 1 \leq k \leq n, \\ y_k' &= \operatorname{Im} f_k(t, x_1 + \mathrm{i}y_1, \dots, x_n + \mathrm{i}y_n) && 1 \leq k \leq n, \end{aligned}$$

which is a $2n$-dimensional real system. Corresponding IVP's are also equivalent—a vectorial initial condition $\mathbf{z}(t_0) = \mathbf{z}^{(0)} \in \mathbb{C}^n$ translates into $\mathbf{x}(t_0) = \operatorname{Re}\mathbf{z}^{(0)} \wedge \mathbf{y}(t_0) = \operatorname{Im}\mathbf{z}^{(0)}$, which gives $2n$ real initial conditions, matching the dimension of the real system. Finally the real version of the Existence and Uniqueness Theorem can be applied and gives the truth of the corresponding complex version.

# Notes cont'd

- The functions in the span of $\{t^k e^{\lambda t}; k \in \mathbb{N}, \lambda \in \mathbb{C}\}$ are called *exponential polynomials*.
  The theorem implies in particular that any solution of a homogeneous linear ODE with constant coefficients is an exponential polynomial.
  Conversely, every exponential polynomial solves a nontrivial homogeneous linear ODE with constant coefficients. For $t^k e^{\lambda t}$ the corresponding ODE can be taken as $(\mathrm{D} - \lambda\,\mathrm{id})^{k+1} y = 0$, and for a linear combination $\sum_{i=1}^r c_i t^{k_i} e^{\lambda_i t}$ we can then take the ODE as
  $$\left[\prod_{i=1}^r (\mathrm{D} - \lambda_i \mathrm{id})^{k_i+1}\right] y = 0.$$
  It should be noted here that a fixed exponential polynomial $y(t)$ satisfies many different such ODE's, since $a(\mathrm{D})y = 0$ implies $b(\mathrm{D})a(\mathrm{D})y = 0$ for any polynomial $b(X) \in \mathbb{C}[X]$. However, it can be shown that there exists a unique monic polynomial $a(X) \in \mathbb{C}[X]$ of smallest degree such that $a(\mathrm{D})y = 0$ and hence a unique “monic” linear homogeneous ODE of smallest order satisfied by $y(t)$; cf. exercises.

# Notes cont'd

Here we supply the yet missing precise definition of polynomial differential operators $p(\mathrm{D}) = p_0 + p_1\mathrm{D} + p_2\mathrm{D}^2 + \cdots + p_d\mathrm{D}^d$ corresponding to polynomials $p(X) \in \mathbb{C}[X]$.

We have defined $p(\mathrm{D})$ as the map $y \mapsto p_0 y + p_1 \mathrm{D}y + \cdots + p_d\mathrm{D}^d)y$, but this is incomplete without specifying the domain and codomain of $p(\mathrm{D})$. Since we want to compose differential operators as maps, domain and codomain should be equal. Now care must be taken to avoid the following problem: If $f\colon \mathbb{R} \to \mathbb{C}$ is differentiable but $f'$ is not, $\mathrm{D}(\mathrm{D}f) = \mathrm{D}f'$ is undefined. (In other words, $\mathrm{D}$ doesn't map the space of differentiable functions into itself.)
The problem can be cured by taking as domain of $\mathrm{D}$ the set of functions $f \in \mathbb{C}^{\mathbb{R}}$ that have derivatives of all orders. This set is commonly denoted by $\mathbb{C}^\infty(\mathbb{R})$ and forms a subspace of $\mathbb{C}^{\mathbb{R}}$. For $f \in \mathbb{C}^\infty(\mathbb{R})$ we have $f' \in \mathbb{C}^\infty(\mathbb{R})$ as well (check it!), and hence $\mathrm{D}\colon \mathbb{C}^\infty(\mathbb{R}) \to \mathbb{C}^\infty(\mathbb{R})$, $f \mapsto f'$ is well-defined.
Finally, we check that solutions of $a(\mathrm{D})y = 0$ are in fact in $\mathbb{C}^\infty(\mathbb{R})$. Writing the ODE in the form $y^{(n)} = a_0 y + a_1 y' + \cdots + a_{n-1}y^{(n-1)}$ and differentiating gives $y^{(n+1)} = a_0 y' + a_1 y'' + \cdots + a_{n-1}y^{(n)}$, showing that $y^{(n+1)}$ exists. Iterating this argument gives $y \in \mathbb{C}^\infty(\mathbb{R})$.

# The Real Case

## Corollary

*If $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$ then the complex solution space of*

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = 0 \qquad \text{(H)}$$

*has a basis consisting of n real solutions, and these form a basis of the real solution space as well. In particular the real solution space has dimension n as well.*

The subsequent proof shows how to actually obtain a basis of real solutions. One simply takes the real and imaginary parts of the (possibly complex) solutions $t^j \mathrm{e}^{\lambda_i t}$, discarding repetitions.

## Proof.

Writing a complex solution as $z(t) = x(t) + \mathrm{i}y(t)$, or $z = x + \mathrm{i}y$ for short, we have

$$\begin{aligned} 0 &= z^{(n)} + a_{n-1}z^{(n-1)} + \cdots + a_1 z' + a_0 z \\ &= x^{(n)} + \mathrm{i}y^{(n)} + \cdots + a_1(x' + \mathrm{i}y') + a_0(x + \mathrm{i}y) \\ &= x^{(n)} + \cdots + a_1 x' + a_0 x + \mathrm{i}\left(y^{(n)} + \cdots + a_1 y' + a_0 y\right). \end{aligned}$$

## Proof cont'd.

By assumption, $x^{(n)} + \cdots + a_1 x' + a_0 x$ and $y^{(n)} + \cdots + a_1 y' + a_0 y$ are real for each $t \in \mathbb{R}$, and hence both must be zero.
$\implies$ The real and imaginary parts of a complex solution are itself solutions.

Applying this to a basis of the complex solution space $S$, we obtain $2n$ real solutions, which generate $S$ and from which we can then select $n$ linearly independent real solutions $\phi_1, \dots, \phi_n$ forming a basis of $S$. Now suppose $c_1, \dots, c_n \in \mathbb{C}$ are such that

$$\begin{aligned} & y(t) = c_1\phi_1(t) + \cdots + c_n\phi_n(t) \in \mathbb{R} \quad \text{for all } t \in \mathbb{R}. \\ \implies \; & \overline{y(t)} = \overline{c}_1\phi_1(t) + \cdots + \overline{c}_n\phi_n(t) = y(t) \end{aligned}$$

for all $t$, and hence $c_i = \overline{c}_i \in \mathbb{R}$ for $1 \leq i \leq n$ by the linear independency of $\phi_i$. This shows that the real solution space is generated by $\phi_1, \dots, \phi_n$.
Moreover, since these functions are linearly independent over $\mathbb{C}$ and $\mathbb{R} \subset \mathbb{C}$, they must also be linearly independent over $\mathbb{R}$. Hence they form a basis of the real solution space. $\square$

## Notes

- Without the condition $a_0, a_1, \dots, a_{n-1} \in \mathbb{R}$ the conclusion in the corollary is false, as the example $y' - \mathrm{i}y = 0$ shows. The complex solutions are $c\,\mathrm{e}^{\mathrm{i}t}$, $c \in \mathbb{C}$, and the only real solution among these is the all-zero function.

- The relation between the real and complex solution space of $y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1y' + a_0y = 0$ just described actually holds in more generality and can be formulated in a pure Linear Algebra setting (and for arbitrary fields $E, F$ with $E \supset F$ in place of $\mathbb{C}$, $\mathbb{R}$). If $S$ is a subspace of some function space $\mathbb{C}^I$ (considered as a vector space over $\mathbb{C}$), we can consider the subset $S_{\mathbb{R}} = S \cap \mathbb{R}^I$ consisting of all real-valued functions in $S$ (*field reduction*). The set $S_{\mathbb{R}}$ forms a vector space over $\mathbb{R}$ (a subspace of $\mathbb{R}^I$). Conversely, starting with a subspace $T$ of $\mathbb{R}^I$, we can consider the span $T_{\mathbb{C}}$ of $T$ over $\mathbb{C}$, which is a subspace of $\mathbb{C}^I$ (*field extension*).
  Then $\dim(T_{\mathbb{C}}) = \dim(T)$ holds in general, and any basis of $T$ over $\mathbb{R}$ forms a basis of $T_{\mathbb{C}}$ over $\mathbb{C}$; moreover, $T = (T_{\mathbb{C}})_{\mathbb{R}}$. But $\dim(S_{\mathbb{R}}) = \dim(S)$ (equivalently, $S = (S_{\mathbb{R}})_{\mathbb{C}}$ iff $S$ has a basis consisting of real-valued functions, and $\dim(S_{\mathbb{R}}) < \dim(S)$ otherwise; the latter case actually occurs if $|I| > 1$.

## Exercise

a) Prove the assertions about field extension $T \mapsto T_{\mathbb{C}}$ in the previous note.
*Hint:* The key fact to be established is that functions $f_1, \dots, f_r \colon I \to \mathbb{R}$ that are linearly independent over $\mathbb{R}$ remain linearly independent over the larger field $\mathbb{C}$.

b) Prove the assertions about field reduction $S \mapsto S_{\mathbb{R}}$ in the previous note, including for $|I| > 1$ an example of a subspace $S$ of $\mathbb{C}^I$ for which $\dim(S_{\mathbb{R}}) < \dim(S)$.
*Hint:* For the example it suffices to consider the case $|I| = 2$, i.e., $\mathbb{C}^I \cong \mathbb{C}^2$.

c) Show that $T_{\mathbb{C}}$ forms a vector space of dimension $2\dim(T)$ over $\mathbb{R}$.

## Exercise

Show that for any matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ the following are equivalent.

1. The real solution space and the complex solution space of $\mathbf{Ax} = \mathbf{0}$ have the same dimension.
2. The row space of $\mathbf{A}$ has a basis consisting of vectors in $\mathbb{R}^n$.

How about the column space of $\mathbf{A}$ in this regard?

## Example

We solve the 3rd-order ODE

$$y''' - y'' - 2y' = 0.$$

The characteristic polynomial is

$$a(X) = X^3 - X^2 - 2X = X(X+1)(X-2)$$

with roots $\lambda_1 = 0$, $\lambda_2 = -1$, $\lambda_3 = 2$ and all multiplicities equal to 1.

$$\implies \mathrm{e}^{0\,t} = 1,\ \mathrm{e}^{-t},\ \mathrm{e}^{2t}$$

form a fundamental system of solutions.

$\implies$ The general solution is

$$y(t) = c_1 + c_2\mathrm{e}^{-t} + c_3\mathrm{e}^{2t}, \quad c_1, c_2, c_3 \in \mathbb{C}$$

(or "$c_1, c_2, c_3 \in \mathbb{R}$" if only real solutions are considered).

## Example

We solve the homogeneous 4th-order ODE

$$y^{(4)} + 8y'' + 16y = 0.$$

The characteristic polynomial is

$$X^4 + 8X^2 + 16 = (X^2 + 4)^2 = (X - 2\mathrm{i})^2 (X + 2\mathrm{i})^2.$$

$\Longrightarrow \lambda_1 = 2\mathrm{i}, \lambda_2 = -2\mathrm{i}$ with multiplicities $m_1 = m_2 = 2$.

$$\Longrightarrow \quad \mathrm{e}^{2\mathrm{i}t},\ t\,\mathrm{e}^{2\mathrm{i}t},\ \mathrm{e}^{-2\mathrm{i}t},\ t\,\mathrm{e}^{-2\mathrm{i}t}$$

form a complex fundamental system.

A real fundamental system is then obtained by taking the real and imaginary parts of one function from each complex conjugate pair, i.e.,

$$\cos(2t),\ t\cos(2t),\ \sin(2t),\ t\sin(2t).$$

The general real solution of $y^{(4)} + 8y'' + 16y = 0$ is therefore $y(t) = c_1\cos(2t) + c_2 t\cos(2t) + c_3\sin(2t) + c_4 t\sin(2t)$ with $c_1, c_2, c_3, c_4 \in \mathbb{R}$ (and the general complex solution is of the same form with $c_1, c_2, c_3, c_4 \in \mathbb{C}$).

# Note

If you are wondering why the real and imaginary parts of any complex fundamental system form a real fundamental system—in particular why the number of functions in both systems is the same, here is the argument in more detail:

If $a(X)$ is real, its non-real roots (if any) come in complex-conjugate pairs $\mu, \overline{\mu}$, which must have the same multiplicity, say $m$. The corresponding $2m$ functions in the complex fundamental system are $\{t^k e^{\mu t}, t^k e^{\overline{\mu} t}; 0 \leq k \leq m-1\}$. Writing $\mu = \alpha + \beta\mathrm{i}$, $\overline{\mu} = \alpha - \beta\mathrm{i}$, we have

$$\begin{aligned} t^k e^{\mu t} &= t^k e^{\alpha t} \cos(\beta t) + \mathrm{i}\, t^k e^{\alpha t} \sin(\beta t), \\ t^k e^{\overline{\mu} t} &= t^k e^{\alpha t} \cos(\beta t) - \mathrm{i}\, t^k e^{\alpha t} \sin(\beta t). \end{aligned}$$

$\Longrightarrow$ The $2m$ real and imaginary parts of both kinds of functions are the same (except for a sign change in the imaginary parts). Discarding these “repetitions”, we obtain the correct number $2m$ of real fundamental solutions, viz.
$\{t^k e^{\alpha t} \cos(\beta t), t^k e^{\alpha t} \sin(\beta t); 0 \leq k \leq m-1\}$.

## Example (Harmonic oszillator)

The corresponding ODE is

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \omega^2 y = 0, \quad \omega > 0.$$

Here the characteristic polynomial is $X^2 + \omega^2 = (X - \mathrm{i}\omega)(X + \mathrm{i}\omega)$ and a fundamental system is $\{\mathrm{e}^{\mathrm{i}\omega t}, \mathrm{e}^{-\mathrm{i}\omega t}\}$.

The general real solution may be written in either of the two forms

1. $y(t) = c_1 \cos(\omega t) + c_2 \sin(\omega t), \quad c_1, c_2 \in \mathbb{R}$;
2. $y(t) = A \cos(\omega t + \alpha), \quad A \geq 0, \alpha \in [0, 2\pi)$.

The second form arises from the general complex solution $y(t) = c_1 \mathrm{e}^{\mathrm{i}\omega t} + c_2 \mathrm{e}^{-\mathrm{i}\omega t}$ by observing that $y(t)$ is real iff $c_2 = \overline{c}_1$, and setting $2c_1 = A\mathrm{e}^{\mathrm{i}\alpha}$.

## Example (Harmonic oszillator with damping)

The corresponding ODE is

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 2\mu\,\frac{\mathrm{d}y}{\mathrm{d}t} + \omega_0^2 y = 0.$$

The quantity $2\mu > 0$ is the damping factor, and $\omega_0 > 0$ is the (suitably normalized) characteristic frequency of the undamped system, whose solutions are generated by $\cos(\omega_0 t)$, $\sin(\omega_0 t)$.

This is a time-independent 2nd-order linear ODE with characteristic polynomial $X^2 + 2\mu X + \omega_0^2$, whose roots are

$$\lambda_{1,2} = -\mu \pm \sqrt{\mu^2 - \omega_0^2}.$$

<u>Case 1: $\mu < \omega_0$.</u>

In this case we have $\lambda_{1,2} = -\mu \pm \mathrm{i}\sqrt{\omega_0^2 - \mu^2}$, and a real fundamental system of solutions is

$$\mathrm{e}^{-\mu t}\cos(\omega t),\ \mathrm{e}^{-\mu t}\sin(\omega t), \quad \omega = \sqrt{\omega_0^2 - \mu^2} < \omega_0.$$

## Example (cont'd)

Solutions form periodic oszillations with lower frequency and exponentially decreasing amplitude.

Case 2: $\mu = \omega_0$.

In this case $\lambda_1 = \lambda_2 = -\mu$, and a (real) fundamental system of solutions is

$$e^{-\mu t}, \; te^{-\mu t}.$$

Solutions ultimately approach zero exponentially, but may have one maximum or minimum.

Case 3: $\mu > \omega_0$.

In this case $\lambda_1$ and $\lambda_2$ are distinct negative real numbers, and a fundamental system of solutions is

$$e^{-\mu_1 t}, \; e^{-\mu_2 t}, \quad \mu_{1,2} = \mu \pm \sqrt{\mu^2 - \omega_0^2} > 0.$$

All solutions approach zero exponentially.

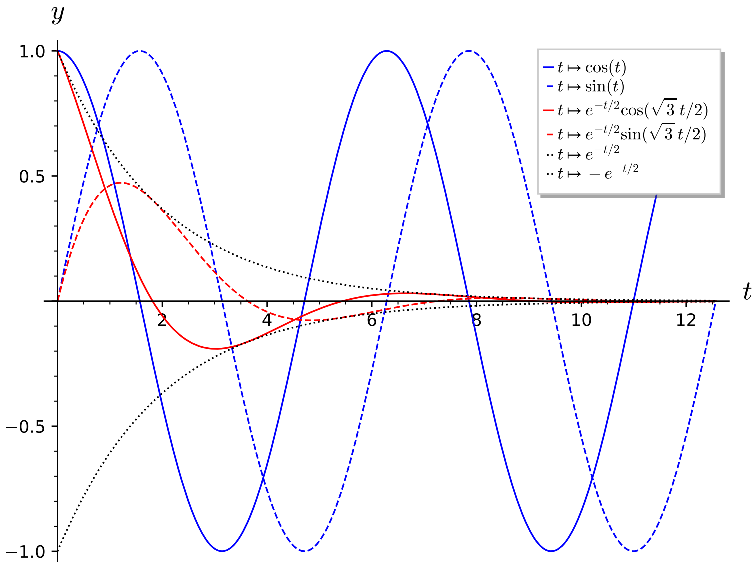


Figure: Fundamental system (in red) for $\omega_0 = 1$, $\mu = 1/2$

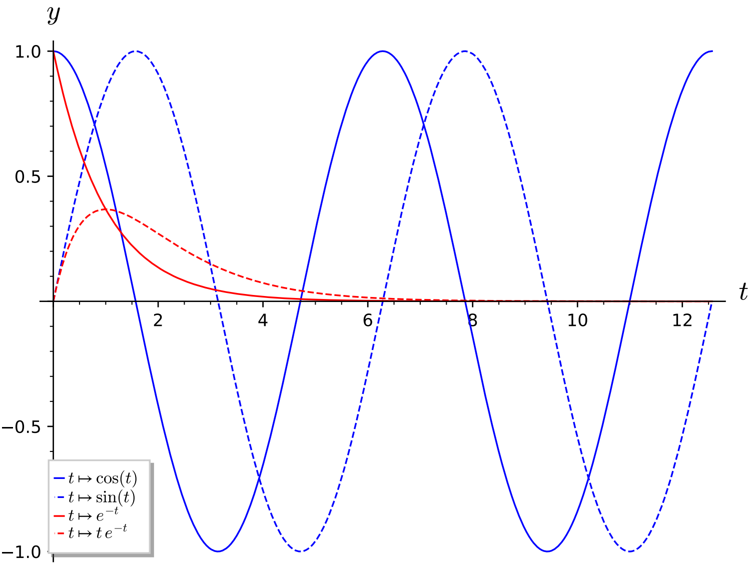


Figure: Fundamental system (in red) for $\omega_0 = 1$, $\mu = 1$

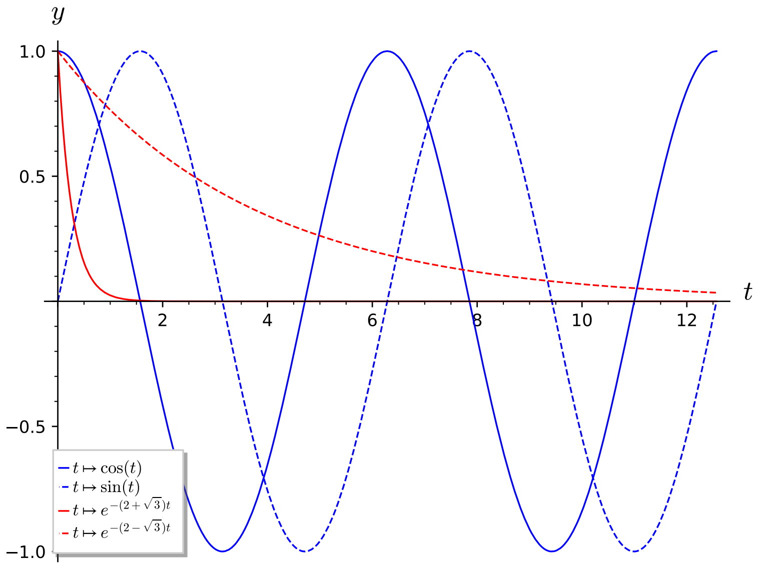


Figure: Fundamental system (in red) for $\omega_0 = 1$, $\mu = 2$

# Notes on the preceding figures

- In the first figure you can see that the "period" (obtained by neglecting the decay factor $\mathrm{e}^{-t/2}$) of the two fundamental solutions is larger than for the corresponding undamped system, whose solutions are $\cos t$, $\sin t$ ($4\pi/\sqrt{3} \approx 7.255$ vs. $2\pi$). Every nonzero solution $y(t)$ inherits the "period" and the decay factor, as can be seen by writing it in the form $y(t) = \mathrm{e}^{-t/2}\big(c_1 \cos(\omega t) + c_2 \sin(\omega t)\big)$, $\omega = \sqrt{3}/2$.
- The figures show that, contrary to the case of 1st-order ODE's, solution graphs of 2nd-order ODE's may intersect but can't touch. In fact, the Existence and Uniqueness Theorem tells us that in the cases under consideration for any point $(t_0, y_0) \in \mathbb{R}^2$ and any $m_0 \in \mathbb{R}$ there is exactly one solution passing through this point and having slope $m_0$ there. For the case $\mu = \omega_0 = 1$ this is illustrated on the next slide.

## Exercise

It appears that in the first figure the first fundamental solution $y_1(t)$ (that involving $\cos$) doesn't have a maximum at $t = 0$. Verify this property, and describe the extrema of $y_1(t)$ in terms of those of $\cos$.

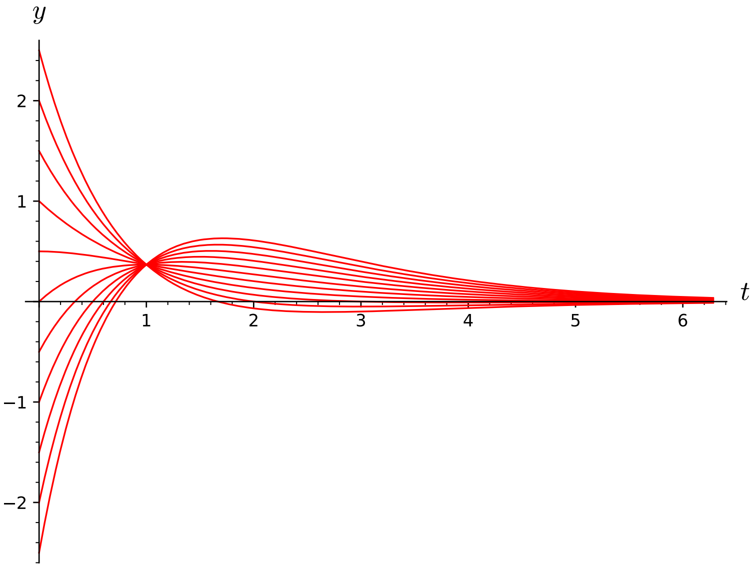


Figure: The solutions of $y'' + 2y' + y = 0$ satisfying $y(1) = 1/e \approx 0.368$ form a 1-parameter family, viz. $y(t) = c\,\mathrm{e}^{-t} + (1-c)t\,\mathrm{e}^{-t}$, $c \in \mathbb{R}$.

# The Inhomogeneous Case

cf. also [BDM17], Ch. 4.3

The general solution of

$$y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_1 y' + a_0 y = b(t) \tag{I}$$

($a_0, \dots, a_{n-1} \in \mathbb{C}$, $b\colon I \to \mathbb{C}$) has the form $y(t) = y_h(t) + y_p(t)$, where $y_p$ denotes one particular solution and $y_h$ the general solution of the associated homogeneous ODE (H). This is proved in the same way as in a previously considered example case.

For a general continuous "source" $b(t)$ order reduction and variation of parameters in the general solution of the resulting 1st-order system provide a method for finding a particular solution. This will be discussed later. Here we consider only the case

$$b(t) = f(t)\mathrm{e}^{\mu t} \quad \text{with } \mu \in \mathbb{C},\ f(X) \in \mathbb{C}[X].$$

The solution of this special case allows us to solve $a(\mathrm{D})y = b(t)$ for any exponential polynomial $b(t) = \sum_{i=1}^{s} f_i(t)\mathrm{e}^{\mu_i t}$ according to the

## Superposition principle

If $y_1(t)$ solves $a(\mathrm{D})y = b_1(t)$ and $y_2(t)$ solves $a(\mathrm{D})y = b_2(t)$ then $y(t) = c_1 y_1(t) + c_2 y_2(t)$ solves $a(\mathrm{D})y = c_1 b_1(t) + c_2 b_2(t)$ ($c_1, c_2 \in \mathbb{C}$).

## Theorem

*Suppose $\mu$ is a root of $a(X)$ of multiplicity $m$ (“$m = 0$” means “not a root of $a(X)$”) and $f(X)$ has degree $k$. Then $a(\mathrm{D})y = f(t)\mathrm{e}^{\mu t}$ has a (unique) solution $y(t)$ of the form*

$$y(t) = t^m(c_0 + c_1 t + \cdots + c_k t^k)\mathrm{e}^{\mu t}.$$

## Proof.

We may assume $a(X) = \prod_{i=1}^{r}(X - \lambda_i)^{m_i}$ with $\mu = \lambda_1$, $m_1 = m$ (provided $m \geq 1$). We have seen that $\mathrm{D} - \lambda_i \,\mathrm{id}$ acts on the spaces

$$E_k = \left\{t \mapsto g(t)\mathrm{e}^{\mu t}; g(X) \in \mathbb{C}[X], \deg g(X) \leq k\right\}, \quad t = 0, 1, 2, \dots,$$

bijectively if $i \geq 2$ and maps $E_k$ onto $E_{k-1}$ (with the convention $E_{-1} = \{0\}$) if $i = 1$.

Since $a(\mathrm{D})$ is the composition of such operators, with exactly $m$ of them equal to $\mathrm{D} - \mu\,\mathrm{id}$, it is clear that $a(\mathrm{D})$ maps $E_{m+k}$ surjectively onto $E_k$.

In other words, there exists $g(X) \in \mathbb{C}[X]$ of degree $\leq m + k$ such that $a(\mathrm{D})\big(g(t)\mathrm{e}^{\mu t}\big) = f(t)\mathrm{e}^{\mu t}$.

Moreover, since $a(\mathrm{D})$ annihilates $t^j\mathrm{e}^{\mu t}$ for $0 \leq j \leq m - 1$, we can choose $g(X)$ of the form $g(X) = X^m(c_0 + c_1 X + \cdots + c_k X^k)$. $\square$

# Notes

- The proof remains valid for $m = 0$, if we omit the normalization $\mu = \lambda_1$.
- In the case $m = k = 0$, i.e. $b(t) = \mathrm{e}^{\mu t}$ with $\mu$ not a root of $a(X)$, we have $a(\mu) \neq 0$ and we can solve $a(\mathrm{D})y = \mathrm{e}^{\mu t}$ directly as follows:

$$a(\mathrm{D})\mathrm{e}^{\mu t} = a(\mu)\mathrm{e}^{\mu t} \implies a(\mathrm{D})\left(\frac{1}{a(\mu)}\,\mathrm{e}^{\mu t}\right) = \mathrm{e}^{\mu t}.$$

- If you have difficulties to understand the argument using the differential operators $\mathrm{D} - \mu\,\mathrm{id}$, consider first the special case $\mu = 0$, in which the operator is just $\mathrm{D}\colon y \to y'$ and $E_k$ is the space of polynomials of degree $\leq k$. From Calculus I we know that differentiation decreases the degree of a polynomial by 1 (except for the constant case) and hence maps $E_k$ onto $E_{k-1}$. On the other hand, if $\lambda \neq 0$ then

$$\mathrm{D}(t^k\mathrm{e}^{\lambda t}) = kt^{k-1}\mathrm{e}^{\lambda t} + t^k\lambda\mathrm{e}^{\lambda t} = (\lambda t^k + kt^{k-1})\mathrm{e}^{\lambda t},$$

showing that in this case the degree of any non-constant polynomial factor is preserved and $\{g(t)\mathrm{e}^{\lambda t}; \deg g(X) \leq k\}$ is mapped bijectively onto itself.

# Example

We determine the general solution of the 3rd-order ODE

$$y''' + 2y'' + y' = t + 2\mathrm{e}^{-t}.$$

The characteristic polynomial is $a(X) = X^3 + 2X^2 + X = X(X+1)^2$.
$\Longrightarrow$ $1$, $\mathrm{e}^{-t}$, $t\mathrm{e}^{-t}$ form a fundamental system of solutions of $y''' + 2y'' + y' = 0$.

A particular solution of the inhomogeneous ODE can be obtained by solving

1. $y''' + 2y'' + y' = t$,
2. $y''' + 2y'' + y' = \mathrm{e}^{-t}$,

and applying superposition.

(1) Here $\mu = 0$, which is a root of $a(X)$ of multiplicity 1.
$\Longrightarrow$ The „Ansatz“ $y_1(t) = c_1 t + c_2 t^2$ yields a solution.
Substituting this in the ODE (1) gives

$$2(2c_2) + (c_1 + 2c_2 t) = c_1 + 4c_2 + 2c_2 t = t.$$

## Example (cont'd)

The solution is $c_2 = \frac{1}{2}$, $c_1 = -2$, i.e., $y_1(t) = -2t + \frac{1}{2}t^2$.

(2) Here $\mu = -1$, which is a root of $a(X)$ of multiplicity 2.

$\Longrightarrow$ The „Ansatz“ $y_2(t) = c\,t^2\mathrm{e}^{-t}$ yields a solution.

$$a(\mathrm{D})y_2(t) = c\,\mathrm{D}(\mathrm{D}+\mathrm{id})^2(t^2\mathrm{e}^{-t}) = 2c\,\mathrm{D}(\mathrm{D}+\mathrm{id})(t\mathrm{e}^{-t}) = 2c\,\mathrm{D}\mathrm{e}^{-t} = -2c\,\mathrm{e}^{-t}$$

$\Longrightarrow$ $y_2(t) = -\frac{1}{2}t^2\mathrm{e}^{-t}$ solves the ODE (2).

Finally, superposition gives that

$$y(t) = y_1(t) + 2y_2(t) = -2t + \frac{1}{2}t^2 - t^2\mathrm{e}^{-t}$$

solves the original ODE $a(\mathrm{D})y = t + 2\mathrm{e}^{-t}$.

The general complex (real) solution of the original ODE is therefore

$$y(t) = c_1 + c_2\mathrm{e}^{-t} + c_3t\mathrm{e}^{-t} - 2t + \frac{1}{2}t^2 - t^2\mathrm{e}^{-t}$$

with constants $c_1, c_2, c_3 \in \mathbb{C}$ (respectively, $c_1, c_2, c_3 \in \mathbb{R}$).

# Further Notes

- Pay attention to the fact that in the monomial case $b(t) = t^k \mathrm{e}^{\mu t}$ the correct „Ansatz" is $y(t) = t^m(c_0 + c_1 t + \cdots + c_k t^k)\mathrm{e}^{\mu t}$ (i.e., a full exponential polynomial).
- Before applying superposition, collect monomials with the same factors $\mathrm{e}^{\mu t}$. For example, when solving $a(\mathrm{D})y = 1 + t + 2\,\mathrm{e}^{-t}$, use $b_1(t) = (1 + t)\mathrm{e}^{0\,t}$, $b_2(t) = \mathrm{e}^{-t}$ (and not a superposition of three solutions corresponding to $1$, $t$, $\mathrm{e}^{-t}$). This saves computation time.

## Example (Harmonic oszillator with periodic source)

The corresponding ODE is

$$\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + \omega_0^2 y = A\cos(\omega t), \quad \omega_0, \omega, A > 0.$$

$\omega_0$ denotes the characteristic frequency of the oszillator and $\omega$ the frequency of the external source.

In order to apply the machinery developed, we consider the "complexified" ODE

$$y'' + \omega_0^2 y = A\,\mathrm{e}^{\mathrm{i}\omega t}.$$

The real part of any particular solution of the complex ODE will then solve the real ODE.

The characteristic polynomial $a(X) = X^2 + \omega_0^2 = (X - \mathrm{i}\omega_0)(X + \mathrm{i}\omega_0)$ has roots $\lambda_{1,2} = \pm\mathrm{i}\omega_0$ (the same as in the homogeneous case $b(t) = 0$).

Hence we need to distinguish the cases $\omega = \omega_0$ (the so-called *resonance* case) and $\omega \neq \omega_0$.

# Example (cont'd)

Case 1: $\omega \neq \omega_0$.

In this case the „Ansatz“ $y(t) = c\,\mathrm{e}^{\mathrm{i}\omega t}$ yields a solution.

$$a(\mathrm{D})y(t) = c\,a(\mathrm{i}\omega)\mathrm{e}^{\mathrm{i}\omega t} = c(\omega_0^2 - \omega^2)\mathrm{e}^{\mathrm{i}\omega t} = A\,\mathrm{e}^{\mathrm{i}\omega t}$$

$\Longrightarrow c = \frac{A}{a(\mathrm{i}\omega)} = \frac{A}{\omega_0^2 - \omega^2}$, and a real particular solution is

$$y(t) = \frac{A}{\omega_0^2 - \omega^2}\cos(\omega t).$$

Case 2: $\omega = \omega_0$.

In this case the „Ansatz“ $y(t) = c\,t\mathrm{e}^{\mathrm{i}\omega_0 t}$ yields a solution.

$$\begin{aligned} a(\mathrm{D})y(t) &= c(\mathrm{D} + \mathrm{i}\omega_0\,\mathrm{id})(\mathrm{D} - \mathrm{i}\omega_0\,\mathrm{id})(t\,\mathrm{e}^{\mathrm{i}\omega_0 t}) = c(\mathrm{D} + \mathrm{i}\omega_0\,\mathrm{id})(\mathrm{e}^{\mathrm{i}\omega_0 t}) \\ &= c(2\mathrm{i}\omega_0)\mathrm{e}^{\mathrm{i}\omega_0 t} = A\,\mathrm{e}^{\mathrm{i}\omega_0 t} \end{aligned}$$

$\Longrightarrow c = \frac{A}{2\mathrm{i}\omega_0}\left(= \frac{A}{a'(\mathrm{i}\omega_0)}\right)$, and a real particular solution is

$$y(t) = \frac{A}{2\omega_0}\,t\sin(\omega_0 t).$$

$\Longrightarrow$ The general real solution of $y'' + \omega_0^2 y = A\cos(\omega t)$ is

$$y(t) = c_1\cos(\omega_0 t) + c_2\sin(\omega_0 t) + \begin{cases} \frac{A}{\omega_0^2-\omega^2}\cos(\omega t) & \text{if } \omega \neq \omega_0, \\ \frac{A}{2\omega_0}\, t\sin(\omega_0 t) & \text{if } \omega = \omega_0 \end{cases}$$

with constants $c_1, c_2 \in \mathbb{R}$.

## Notes

- The resonance phenomenon must, e.g., be taken into account when constructing bridges, which are subject to vertical vibrations caused by the airflow around the bridge. If the frequency of the external force (which is periodic) matches the natural frequency of the bridge's material (steel), vibrations are amplified—leading ultimately to disaster ($\longrightarrow$ *Tacoma Narrows Bridge*).
- In the preceding examples we have sometimes used the factorization of $a(D)$, which we knew from solving the associated homogeneous ODE $a(D)y = 0$, to speed up the computation of $a(D)y$ for certain functions $y$. The standard method for obtaining $a(D)y$ is of course to compute all derivatives $y', y'', \dots, y^{(n)}$ and then form the linear combination $y^{(n)} + \dots + a_2y'' + a_1y' + a_0y$.

# Notes cont'd

- Higher-order linear ODE's with constant coefficients are discussed in [BDM17], Ch. 4.2–4.4. Our theorem for the inhomogeneous case can be viewed as a generalization of the "method of undetermined coefficients" in Ch. 4.3; cf. also Ex. 14 in this chapter. The "method of annihilators", discussed in exercises for the same chapter, provides an alternative but equivalent approach to our result. The "method of variation of parameters" in Ch. 4.4, which is more powerful, will be discussed within the framework of linear ODE systems later in the course.
- The explicit formula $y(t) = \frac{1}{a(\mu)}\,\mathrm{e}^{\mu t}$ for the solution of $a(\mathrm{D})y = \mathrm{e}^{\mu t}$, which requires $\mu$ to be a nonzero of the characteristic polynomial $a(X)$, admits the generalization $y(t) = \frac{1}{a^{(m)}(\mu)}t^m\mathrm{e}^{\mu t}$ in the case where $\mu$ is a root of multiplicity $m$ of $a(X)$; cp. the solution in Case 2 of the example. For the proof write $a(X) = (X-\mu)^m A(X)$ and compute $a(\mathrm{D})[t^m\mathrm{e}^{\mu t}] = A(\mathrm{D})(\mathrm{D}-\mu)^m[t^m\mathrm{e}^{\mu t}] = A(\mathrm{D})[m!\mathrm{e}^{\mu t}] = m!A(\mu)\mathrm{e}^{\mu t}$. This yields the solution $y(t) = \frac{1}{m!A(\mu)}t^m\mathrm{e}^{\mu t}$, and the Leibniz formula for the $m$th derivative $\mathrm{D}^m(fg)$ of a product of two functions can then be used to show that $a^{(m)}(\mu) = m!A(\mu)$.

## Exercise

Determine a real fundamental system of solutions for the following ODE's:

a) $y'' - 4y' + 4y = 0$;

b) $y''' - 2y'' - 5y' + 6y = 0$;

c) $y''' - 2y'' + 2y' - y = 0$;

d) $y''' - y = 0$;

e) $y^{(4)} + y = 0$;

f) $y^{(8)} + 4y^{(6)} + 6y^{(4)} + 4y'' + y = 0$.

## Exercise

Determine the general real solution of

a) $y'' + 3y' + 2y = 2$;

b) $y'' + y' - 12y = 1 + t^2$;

c) $y'' - 5y' + 6y = 4t\mathrm{e}^t - \sin t$;

d) $y''' - 2y'' + y' = 1 + \mathrm{e}^t \cos(2t)$;

e) $y^{(4)} + 2y'' + y = 25\mathrm{e}^{2t}$;

f) $y^{(n)} = t\mathrm{e}^t$, $n \in \mathbb{N}$.

## Exercise

For $a, b \in \mathbb{C}$ consider the ODE

$$y'' + \frac{a}{t}\,y' + \frac{b}{t^2}\,y = 0 \qquad (t > 0). \tag{1}$$

a) Show that $\phi\colon \mathbb{R}^+ \to \mathbb{C}$ is a solution of (1) iff $\psi\colon \mathbb{R} \to \mathbb{C}$ defined by $\psi(s) = \phi(\mathrm{e}^s)$ is a solution of

$$y'' + (a-1)y' + by = 0. \tag{2}$$

b) Determine the general solution of (1) for $(a, b) = (6, 4)$ and $(a, b) = (3, 1)$.

## Exercise

Solve the initial value problem

$$\mathbf{y}' = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \mathbf{y} + \begin{pmatrix} t \\ \sin t \end{pmatrix}, \quad \mathbf{y}(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

## Definition

Suppose $\mathbf{y} = (y_0, y_1, y_2, \dots)$ is a sequence of (complex) numbers. The (formal) power series

$$f(t) = \sum_{n=0}^{\infty} \frac{y_n}{n!} t^n = \frac{y_0}{0!} + \frac{y_1}{1!} t + \frac{y_2}{2!} t^2 + \frac{y_3}{3!} t^3 + \cdots$$

is called *exponential generating function* of $\mathbf{y}$ and denoted by $\operatorname{egf}(\mathbf{y})$.

## Notes

- $\operatorname{egf}(\mathbf{y}) = f(t)$ contains all information about the sequence $\mathbf{y}$. This is true at least in a formal sense, but in the case where $f(t)$ has radius of convergence $R > 0$ can also be seen by term-wise differentiation: $y_n = f^{(n)}(0)$ is determined by $f$.
- Exponential generating functions (and likewise their ordinary counterparts $g(t) = \sum_{n=0}^{\infty} y_n t^n$) are used with great success in Enumerative Combinatorics. The most famous example is the sequence $\mathbf{d} = (d_0, d_1, d_2, \dots) = (1, 0, 1, 2, 9, \dots)$ of fixed-point free permutations of $n$ letters (so-called *derangements*), whose exponential generating function turns out to be $\frac{\mathrm{e}^{-t}}{1-t}$, showing that $d_n = n! \sum_{k=0}^{n} \frac{(-1)^k}{k!}$.

# Notes cont'd

- If the sequence $\mathbf{y} = (y_0, y_1, y_2, \dots)$ grows at most exponentially, i.e., there exists a constant $C > 1$ such that $|y_n| \leq C^n$ for sufficiently large $n$, then $\mathrm{egf}(\mathbf{y})$ has radius of convergence $R = \infty$. This follows from the elementary estimate $n! \geq (n/\mathrm{e})^n$, which implies $\sqrt[n]{|y_n|/n!} \leq C\mathrm{e}/n \to 0$ for $n \to \infty$. All homogeneous linear recurring sequences with constant coefficients have this property (as we know from Discrete Mathematics). The same is true in the inhomogeneous case, provided that the right-hand side $\mathbf{b} = (b_0, b_1, b_2, \dots)$ grows at most exponentially, as is easily proved.

## Theorem

1. *For all sequences* $\mathbf{y} \in \mathbb{C}^{\mathbb{N}}$ *and polynomials* $p(X) \in \mathbb{C}[X]$ *we have*
$$p(\mathrm{D})\,\mathrm{egf}(\mathbf{y}) = \mathrm{egf}\big(p(\mathrm{S})\mathbf{y}\big).$$
2. *Now suppose* $p(X)$ *is monic of degree* $n$*. The sequence* $\mathbf{y} = (y_0, y_1, y_2, \dots)$ *solves the linear recurrence relation* $p(\mathrm{S})\mathbf{y} = \mathbf{b}$ *iff the function* $y(t) = \mathrm{egf}(\mathbf{y})$ *solves the IVP* $p(\mathrm{D})y = \mathrm{egf}(\mathbf{b})$, $y^{(i)}(0) = y_i$ *for* $0 \leq i \leq n-1$*.*

## Proof.

(1) We have

$$\mathrm{D}[\mathrm{egf}(\mathbf{y})] = \frac{\mathrm{d}}{\mathrm{d}t}\sum_{n=0}^{\infty}\frac{y_n}{n!}\,t^n = \sum_{n=1}^{\infty}\frac{n\,y_n}{n!}\,t^{n-1} = \sum_{n=0}^{\infty}\frac{y_{n+1}}{n!}\,t^n = \mathrm{egf}(\mathrm{S}\mathbf{y})$$

This implies $\mathrm{D}^k\,\mathrm{egf}(\mathbf{y}) = \mathrm{egf}(\mathrm{S}^k\mathbf{y})$ for $k \in \mathbb{N}$ and, since egf is $\mathbb{C}$-linear, further $p(\mathrm{D})\,\mathrm{egf}(\mathbf{y}) = \sum_{k=0}^{d} p_k\,\mathrm{D}^k\,\mathrm{egf}(\mathbf{y}) = \sum_{k=0}^{d} p_k\,\mathrm{egf}(\mathrm{S}^k\mathbf{y}) = \mathrm{egf}\left(\sum_{k=0}^{d} p_k\mathrm{S}^k\mathbf{y}\right) = \mathrm{egf}\big(p(\mathrm{S})\mathbf{y}\big)$.

## Proof cont'd.

(2) If $p(\mathrm{S})\mathbf{y} = \mathbf{b}$ then $\operatorname{egf}\big(p(\mathrm{S})\mathbf{y}\big) = \operatorname{egf}(\mathbf{b})$, which on account of Part (1) means $p(\mathrm{D}) \operatorname{egf}(\mathbf{y}) = \operatorname{egf}(\mathbf{b})$. Thus $y(t) := \operatorname{egf}(\mathbf{y})$ solves $p(\mathrm{D})y = \operatorname{egf}(\mathbf{b})$, and as remarked after the definition of $\operatorname{egf}(\mathbf{y})$ we have $y^{(i)}(0) = y_i$.

Conversely, suppose $y(t) = \sum_{k=0}^{\infty} \frac{y_k}{k!} t^k$ solves $p(\mathrm{D})y = \operatorname{egf}(\mathbf{b})$. Then $p(\mathrm{D}) \operatorname{egf}(\mathbf{y}) = \operatorname{egf}(\mathbf{b})$ and hence $\operatorname{egf}\big(p(\mathrm{S})\mathbf{y}\big) = \operatorname{egf}(\mathbf{b})$ by Part (1). Since egf maps sequences bijectively onto formal power series, this implies $p(\mathrm{S})\mathbf{y} = \mathbf{b}$. $\square$

## Note

The theorem merely expresses the fact that term-wise differentiation of an exponential generating function amounts to shifting and truncating the corresponding sequence and that its derivatives evaluated at zero are just the entries of the sequence. If $\mathbf{b}$ grows at most exponentially then $b(t) = \operatorname{egf}(\mathbf{b})$ represents a function with domain $\mathbb{R}$ and $y(t) = \operatorname{egf}(\mathbf{y})$, which has also domain $\mathbb{R}$, forms a solution of the "real" ODE $p(\mathrm{D})y = b(t)$.

## Example

We consider again $y'' = 4y' - 4y$, this time with particular initial values $y(0) = y'(0) = 1$. The corresponding characteristic polynomial is $X^2 - 4X + 4 = (X-2)^2$, so that the general solution of $y'' = 4y' - 4y$ has the form $y(t) = c_1\mathrm{e}^{2t} + c_2 t\mathrm{e}^{2t}$.
Since $y'(t) = 2c_1\mathrm{e}^{2t} + c_2(1+2t)\mathrm{e}^{2t}$, we obtain the system $c_1 = 2c_1 + c_2 = 1$, $c_2 = -1$, and hence $y(t) = \mathrm{e}^{2t} - t\mathrm{e}^{2t}$.

The corresponding discrete IVP is $y_{k+2} = 4y_{k+1} - 4y_k$, $y_0 = y_1 = 1$. Here we have $y_i = c_1 2^k + c_2 k 2^k$, and the initial conditions give $c_1 = 1$, $c_2 = -1/2$, so that $y_k = 2^k - k2^{k-1}$.

By the theorem the solutions must be related by $y(t) = \mathrm{egf}(\mathbf{y})$. Indeed, we have

$$\begin{aligned} y(t) &= \sum_{k=0}^{\infty} \frac{2^k t^k}{k!} - \sum_{k=0}^{\infty} \frac{2^k t^{k+1}}{k!} = \sum_{k=0}^{\infty} \frac{2^k}{k!}\, t^k - \sum_{k=1}^{\infty} \frac{2^{k-1}}{(k-1)!}\, t^k \\ &= \sum_{k=0}^{\infty} \frac{2^k - k2^{k-1}}{k!}\, t^k = \mathrm{egf}(\mathbf{y}). \end{aligned}$$

# Example (cont'd)

The coefficients $c_1, c_2$ in the discrete and continuous case are not the same. This is due to the fact that the chosen fundamental systems are not mapped onto each other by $\mathbf{y} \mapsto \mathrm{egf}(\mathbf{y})$. Rather, an exponential monomial

$$
\begin{aligned}
t^k e^{\lambda t} &= \sum_{n=0}^{\infty} \frac{\lambda^n}{n!}\, t^{n+k} = \sum_{n=k}^{\infty} \frac{\lambda^{n-k}}{(n-k)!}\, t^n \\
&= \sum_{n=0}^{\infty} \frac{n(n-1)\cdots(n-k+1)\lambda^{n-k}}{n!}\, t^n \\
&= \lambda^{-k} \sum_{n=0}^{\infty} \frac{n(n-1)\cdots(n-k+1)\lambda^{n}}{n!}\, t^n
\end{aligned}
$$

is (up to a constant factor) the egf of the sequence $y_n = n(n-1)\cdots(n-k+1)\lambda^n$, which involves falling factorials instead of the powers $n^k$. In our example the difference is not really visible, but it becomes apparent if there are fundamental solutions with $k = 2$ (i.e., the characteristic polynomial has a zero of multiplicity $\geq 3$).