26. (Homogeneous) systems of linear differential equations

Passing from individual differential equations to systems is natural, given that ODEs are an important tool for natural scientists and the nature is naturally more dimensional. In this chapter we will look at systems of linear differential equations of order 1 with several variables. A typical linear ODE of order one with two variables may look like as follows.

$$3y' - z' + 13y + 23z = e^x.$$

Here it is understood that $y = y(x)$ and $z = z(x)$ are unknown functions.

However, this is still too complicated for our purposes. We will focus only on differential equations where just one of the unknown function is differentiated. Moreover, we will adopt a different style of writing equations, where we isolate this derivative on one side. An example of a typical linear ODE with three unknown functions in the proper form could be this:

$$y_2' = 13y_1 + 14y_2 - 23y_3 + \cos(x).$$

Equations like these appear naturally in applications, so we are not too restrictive here. Having n unknown functions, it is natural to ask for n linearly independent equations, and we want each of them to have a different derivative on the left. In other words, for each unknown function y_i we expect to see one equation of the form $y_i' = \dots$

Here is a typical example:

$$\begin{aligned} y_1' &= 2y_1 + y_2 - 3 \\ y_2' &= y_1 + 2y_2 + 3x - 4. \end{aligned}$$

Since each equation features the derivative of a different function, these equations are obviously linearly independent.

We expect infinitely many solutions for such a system, and as usual we are interested in a general solution with parameters, namely n parameters in this case.

To determine a specific solution we therefore need n conditions. We will again focus on initial conditions at a given time x_0, and for equations of order one we asked just about the value of a function itself. This fits well with the new setting, since asking for a specific value at time x_0 from each unknown function constitutes n conditions.

We will formalize our understanding.

Definition 26.1.
By a **system of linear ODEs of order 1 with constant coefficients** we mean a system of the form

$$\begin{aligned} y_1' &= a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n + b_1(x) \\ y_2' &= a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n + b_2(x) \\ &\vdots \\ y_n' &= a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n + b_n(x) \end{aligned}$$

where $b_i(x)$ are right hand-sides and $a_{ij} \in \mathbb{R}$ are coefficients.

The system is called **homogeneous** if $b_i(x) = 0$ for all $i = 1, \dots, n$.

An Initial Value Problem (IVP) or Cauchy problem for such a system has initial conditions
$y_1(x_0) = y_{1,0},\ y_2(x_0) = y_{2,0},\ \dots,\ y_n(x_0) = y_{n,0}.$

Note that we allow only for constant coefficients here. It would be possible to work in a more general setting (and in fact structural theorems that follow would still be valid), but we would not know how to solve such systems anyway.

Just like with linear equations, the notation is a bit inconsistent in that we use the variable notation with $b_i(x)$ but not with y_i. The reason for this is the same, we all know that the unknowns are functions so it is not necessary to remind ourselves of that, but we want to emphasize that the right-hand sides may depend on x.

One of the popular ways to solve small systems of algebraic linear equations is by intuitive elimination, where we gradually reduce the number of unknowns while also reducing the number of equations to be solved. The key step is to determine a certain unknown from one of the remaining equations and substitute this formula into the other equations.

This approach also works in our new setting.

Example 26.a: Consider the initial value problem

$$\begin{aligned} y_1' &= 2y_1 + y_2 - 3 \qquad & y_1(0) = 3\\ y_2' &= y_1 + 2y_2 + 3x - 4, \qquad & y_2(0) = 1. \end{aligned}$$

As usual, we start by deriving a general solution of the given system.

1. Looking at the equations, we can either isolate y_2 from the first equation or y_1 from the second. We will go with the former idea, obtaining a key reduction formula

$$y_2 = y_1' - 2y_1 + 3. \qquad (\star)$$

We substitute this into the second equation, obtaining

$$\begin{aligned} [y_1' - 2y_1 + 3]' &= y_1 + 2[y_1' - 2y_1 + 3] + 3x - 4\\ y_1'' - 2y_1' + 0 &= y_1 + 2y_1' - 4y_1 + 6 + 3x - 4\\ y_1'' - 4y_1' + 3y_1 &= 3x + 2 \end{aligned}$$

As expected, the number of equations and the number of unknowns were reduced by one, and we arrived at one linear differential equation of order two. This is perfectly all right, we are trading equations and unknowns for increased degree.

We do know how to solve the resulting equation, so briefly:

a) Solving the homogeneous system through its characteristic equation $\lambda^2 - 4\lambda + 3 = 0$ we obtain a homogeneous solution $y_{1h}(x) = a\,e^x + b\,e^{3x}$.

b) We use the guessing method to find a particular solution. The initial form $y_1(x) = Ax + B$ has special number $\lambda = 0$ and therefore does not require any correction. We substitute this into our equation and obtain $y_{1p}(x) = x + 2$.

We arrive at the general solution $y_1(x) = x + 2 + a\,e^x + b\,e^{3x}$ of the second-order linear ODE.

Now comes the back substitution stage. Having determined the last remaining unknown, we return to reduction formulas and determine the other unknowns. In this case we return to $(\star)$ and obtain

$$\begin{aligned} y_2(x) &= [x + 2 + a\,e^x + b\,e^{3x}]' - 2[x + 2 + a\,e^x + b\,e^{3x}] + 3\\ &= -2x - a\,e^x + b\,e^{3x}. \end{aligned}$$

Now we have a general solution for our system:

$$\begin{aligned} y_1(x) &= x + 2 + a\,e^x + b\,e^{3x},\\ y_2(x) &= -2x - a\,e^x + b\,e^{3x},\ x \in \mathbb{R}. \end{aligned}$$

2. Knowing the functions y_1, y_2, we rewrite the given initial conditions:

$$\begin{aligned} 0 + 2 + a\,e^0 + b\,e^0 &= 3, \\ -2\cdot 0 - ae^0 + be^0 &= 1 \end{aligned} \quad\Longrightarrow\quad \begin{aligned} a + b &= 1,\\ -a + b &= 1. \end{aligned}$$

This yields $a = 0$, $b = 1$. The solution of the given problem is

$$\begin{aligned} y_1(x) &= x + 2 + e^{3x},\\ y_2(x) &= -2x + e^{3x},\ x \in \mathbb{R}. \end{aligned}$$

Remark: If we substitute $y_{1h}(x) = a\,e^x + b\,e^{3x}$ into the reduction formula ($\star$), we obtain the function $y_{2h}(x) = -a\,e^x + b\,e^{3x}$. It is easy to confirm that the pair y_{1h}, y_{2h} actually solves the associated homogeneous system

$$\begin{aligned} y_1' &= 2y_1 + y_2 \\ y_2' &= y_1 + 2y_2. \end{aligned}$$

△

This worked out rather well, and it is a viable approach.

Fact 26.2.
Every system of n linear ODEs of order 1 can be transformed via elimination to one linear ODE of order n.

We restricted our attention to equations with just one derivative present. However, as mathematicians we can easily imagine other situations, for instance this system:

$$\begin{aligned} y_1' + 2y_2' &= 5y_1 - y_2 + 3e^x \\ y_1' - y_2' &= 2y_1 + 5y_2 + 6x. \end{aligned}$$

Here the cure is simple, we apply the Gaussian elimination to the part of the system with derivatives. In fact, we will want to apply the Gauss-Jordan version (so it *is* good for something after all). We start by subtracting the first equation from the second.

$$\begin{aligned} y_1' + 2y_2' &= 5y_1 - y_2 + 3e^x \\ -3y_2' &= -3y_1 + 6y_2 + 6x - 3e^x. \end{aligned}$$

Now we divide the second row by -3, and then subtract it twice from the first row.

$$\begin{aligned} y_1' &= 3y_1 + 3y_2 + 4x + e^x \\ y_2' &= y_1 - 2y_2 - 2x + e^x. \end{aligned}$$

We arrived at a system of the form that we want to see here.

This idea can be generalized. We can simply consider systems of linear equations without any restriction on derivatives, and define the order of such a linear differential equation with more unknown functions as the highest degree of derivative we can see there. Then we have the following statement:

Fact 26.3.
Every system of n linear ODEs of orders n_i with n variables can be transformed via elimination to one linear ODE of order $\sum n_i$.

This sounds good, and one might think that this is could be the end of this chapter: We simply eliminate all systems to single linear equations of higher order and we know how to solve these. Unfortunately, it is not so simple when it comes to actual calculations.

Example 26.b: Consider the system

$$\begin{aligned} y_1' &= y_1 + 2y_2 + y_3 + 1 \\ y_2' &= -y_1 + 2y_2 + 2y_3 \\ y_3' &= 2y_1 + y_2 + y_3 + x. \end{aligned}$$

We can start with the first equation, it offers reduction formulas for y_2 and y_3. The latter looks more appealing (no fractions will appear), so we go this way:

$$y_3 = y_1' - y_1 - 2y_2 - 1. \qquad (\star)$$

We substitute this for y_3 into the second and third equation, work out the derivative in the last one on the left, and the two remaining equations simplify to

$$\begin{aligned} y_2' - 2y_1' &= -3y_1 - 2y_2 - 2 \\ y_1'' - 2y_1' - 2y_2' &= y_1 - y_2 + x - 1. \end{aligned}$$

None of these equations offers a way to express y_1 or y_2. How do we go on?

Now we claimed above that some elimination is possible, but it is not intuitive any more, one needs to play with differential calculus for that, see section 27c. This negates the main advantage of elimination, namely its intuitive ease.

△

The system we played with here is not in any way special, it was a typical complication. The elimination works really swell for two-by-two systems, but it is not very friendly for larger systems. The preferred approach is to solve systems directly. In fact, instead of transforming systems into equations, people actually often go in the opposite way and transform single linear equations into systems.

Example 26.c: Consider the equation $y''' + \sin(x)y'' - xy' + e^x y = \ln(x)$. We start by introducing the nickname y_1 for y and rewrite this equation as

$$y_1''' + \sin(x)y_1'' - xy_1' + e^x y_1 = \ln(x).$$

Note that now we have a 1×1 system of linear ODEs, so we are on the right track, just the order does not fit.

We will therefore reduce it by hiding one derivative into a new unknown:

$$y_1' = y_2 \qquad (\star)$$

Note that in fact, $y' = y_2$, we also have $y_2' = [y_1']' = y_1''$, $y_2'' = [y_1']'' = y_1'''$ and so on, in other words, y_2 really allows us to reduce the order of derivatives for y_1. We substitute into the given equation where possible and obtain

$$y_2'' + \sin(x)y_2' - xy_2 + e^x y_1 = \ln(x).$$

We successfully reduced the order of this equation, but we are not happy with the order yet, so we reduce again. We introduce $y_2' = y_3$, then also $y_3' = y_2''$ and so on. We also note that $y_3 = [y_2]' = [y']' = y''$. In other word, the functions y_i give direct access to derivatives of the original unknown y. Substituting y_3 into the latest equation where possible we obtain

$$y_3' + \sin(x)y_3 - xy_2 + e^x y_1 = \ln(x).$$

We are happy with the order now, so we stop our reduction. As the last step we gather all the reduction equations and the latest version of the reduced equation (rewritten properly) to obtain the system

$$\begin{aligned} y_1' &= \qquad\quad y_2 \\ y_2' &= \qquad\qquad\qquad y_3 \\ y_3' &= -e^x y_1 + xy_2 - \sin(x)y_3 + \ln(x). \end{aligned}$$

This is a standard 3×3 system.

Note that if the order of the original equation were higher, then we would introduce $y_3' = y_4$, and we would have $y_4 = y'''$. You can surely see the pattern in the reduction equations. Consequently, the system created by the reduction formulas has a simple and regular form. Moreover, comparing coefficients of the last transformed equation with the given one we also see a pattern there. In fact, knowing these rules allows one to simply write down the resulting system just by looking at the given equation, without going through all the stages.

We also meet initial value problems. For instance, our equations could have the following conditions attached: $y(0) = 1$, $y'(0) = 13$, $y''(0) = -1$.

However, we observed above that the derivatives of the original unknown function are accessible through the new functions, so we easily rewrite these conditions as

$$\begin{aligned} y_1(0) &= 1 \\ y_2(0) &= 13 \\ y_3(0) &= -1. \end{aligned}$$

This is exactly the right form of initial conditions that we use for systems, so everything fits perfectly.
△

This simple algorithmic procedure has no hidden traps in it, which makes it simple and eminently suitable for computer algebra systems.

Algorithm 26.4.
⟨transforming linear equations into systems⟩
Given: a linear differential equation.

0. Denote the original unknown function as y_1 and rewrite the given equation, obtaining the first transformed equation.
Set $k = 1$.

1. If the latest transformed equation is of order more than one, then

a) Introduce a new unknown function y_{k+1} to hide one derivative of y_k in it by the reduction formula $y_{k+1} = y_k'$.

b) Replace all derivatives of y_k in the transformed equation with derivatives of y_{k+1} in the obvious manner: $y_k^{(i)} = y_{k+1}^{(i-1)}$.
This results in a new transformed equation featuring functions $y_1, \dots, y_k, y_{k+1}$ and with order reduced by one.

c) Increase k by 1 and return to step 1.

2. If the latest transformed equation is of order 1, compose the resulting system of equation by gathering all reduction formulas $y_1' = y_2, \dots, y_{k-1}' = y_k$ and adding the latest transformed equation to this list.
The algorithm stops.

3. If the given problem also features initial conditions of the form $y^{(k)}(x_0) = y_k$, we rewrite them using the fact that $y^{(k)} = y_{k+1}$.
△

If the given equation is

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = b(x),$$

then the resulting system is

$$\begin{aligned} y_1' &= \quad y_2 \\ y_2' &= \qquad y_3 \\ y_3' &= \qquad\quad y_4 \\ &\vdots \\ y_n' &= -a_0(x)y_1 - a_1(x)y_2 - \dots - a_{n-1}(x)y_n + b(x). \end{aligned}$$

This algorithm always succeeds, and this observation actually constitutes the proof of the following statement.

Fact 26.5.
Every linear ODE of order n (and every system of linear ODEs with sum of orders n) can be equivalently transformed into a system of n linear ODEs of order 1 of the type we study here.

Which brings us to an interesting observation. It turns out that high order linear differential equations and systems of special order-one differential equations are connected in both directions, and knowing one gives the knowledge of the other. In chapter 15 we introduced some statements about linear differential equations, in particular an existence and uniqueness theorem. Using the correspondence we have now, we can easily transfer these statements also to systems of linear ODEs. However, it turns out that our special systems of order one are more tractable, and the preferred way is to transfer observation from systems to single equations of higher order. In particular, the existence and uniqueness result for linear differential equations is typically proved as a direct consequence of the existence and uniqueness result for systems of linear ODEs.

Theorem 26.6. (on existence and uniqueness for systems)
Consider a system of linear ODEs of order 1.
If $b_i(x)$ are continuous on an open interval I, then for every $x_0 \in I$ and all $y_{1,0}, y_{2,0}, \dots, y_{n,0} \in \mathbb{R}$ there exists a solution of the corresponding IVP on I and it is unique.

Thus, to put our theory on a firm footing, we should now prove this statement. However, the proof is rather involved and we prefer to leave it to more advanced books.

We will now develop the preferred approach to solving systems of linear differential equations.

26a. Matrix approach

Just like in linear algebra, in order to handle systems well we will capture them using matrix and vector notation.

A system

$$\begin{aligned}
y_1' &= a_{11}y_1 + a_{12}y_2 + \dots + a_{1n}y_n + b_1(x)\\
y_2' &= a_{21}y_1 + a_{22}y_2 + \dots + a_{2n}y_n + b_2(x)\\
&\vdots\\
y_n' &= a_{n1}y_1 + a_{n2}y_2 + \dots + a_{nn}y_n + b_n(x)
\end{aligned}$$

can be written as

$$\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix} + \begin{pmatrix} b_1(x) \\ \vdots \\ b_n(x) \end{pmatrix}.$$

We will naturally denote $\vec{y}(x) = \begin{pmatrix} y_1(x) \\ \vdots \\ y_n(x) \end{pmatrix}$ to have just one unknown (a vector now). Adopting the convention that $\begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}'$, we can write the given system as $\vec{y}\,' = A\vec{y} + b(x)$, where $A = (a_{i,j})_{i,j=1}^n$ is the **matrix of the system** and $\vec{b}(x) = (b_i)_{i=1}^n$ is the **vector of right-hand sides**.

The system is homogeneous if $\vec{b}=\vec{0}$, where $\vec{0}$ is the zero vector in $\mathbb{R}^n$.
Moreover, if initial conditions are given, then they can be expressed as $\vec{y}(x_0)=\vec{y}_0$.
Again, we write $\vec{b}(x)$ to emphasize that this is a vector of functions, although the notation looks a bit inconsistent.

Example 26a.a: Consider the initial value problem

$$\begin{aligned} y_1' &= 2y_1 + y_2 - 3 \qquad & y_1(0) &= 3 \\ y_2' &= y_1 + 2y_2 + 3x - 4, \qquad & y_2(0) &= 1. \end{aligned}$$

It can be written as $\vec{y}\,' = A\vec{y} + \vec{b}(x)$, where

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \qquad A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \vec{b}(x) = \begin{pmatrix} -3 \\ 3x-4 \end{pmatrix}.$$

In example 26.a we found a general solution of the form

$$\vec{y}(x) = \begin{pmatrix} x+2+a\,e^x+b\,e^{3x} \\ -2x-a\,e^x+b\,e^{3x} \end{pmatrix}.$$

Interestingly, it can be written as

$$\vec{y}(x) = \begin{pmatrix} x+2 \\ -2x \end{pmatrix} + \begin{pmatrix} a\,e^x+b\,e^{3x} \\ -a\,e^x+b\,e^{3x} \end{pmatrix}.$$

We can easily check that

$$\vec{y}_p(x) = \begin{pmatrix} x+2 \\ -2x \end{pmatrix}$$

solves the given system, while

$$\vec{y}_h(x) = \begin{pmatrix} a\,e^x+b\,e^{3x} \\ -a\,e^x+b\,e^{3x} \end{pmatrix} = a\begin{pmatrix} e^x \\ -e^x \end{pmatrix} + b\begin{pmatrix} e^{3x} \\ e^{3x} \end{pmatrix}$$

is a solution of the associated homogeneous system. In fact, each of the vectors that we see on the right is a solution of that homogeneous system, so we would guess that is it a basis of the space of homogeneous solutions.
This should not surprise us, as we expect to see the same behavior as we observed in chapter 15.
We also note that the homogeneous solution has two parameters, which suggests that the space of all solutions is two-dimensional.
△

This example shows what we expect to find when we treat a system of (linear) differential equations as a matrix-coded problem. All solutions will be vectors of functions, and we manipulate them in the usual way from linear algebra. We can easily return to the original point of view where we work with individual functions by reading entries in vectors of functions.
As usual, we first focus on the simpler case.

26b. Homogeneous systems

We start our exploration by confirming the expected facts regarding homogeneous systems.

Theorem 26b.1. (on structure of solution set for homogeneous systems)
Consider a homogeneous system of linear ODEs $\vec{y}\,' = A\vec{y}$, where $A\in\mathbb{R}^{n\times n}$.
The set of all solutions of this system on some open interval I is a linear space of dimension n.

The proof is fairly standard in its structure and content, and the convenient matrix notations makes it even simpler (cf. theorem 15.3).

Proof: 1. We will show that the set of all solutions on I (denoted M) is a subset of the space of vector functions on I that is closed under the basic linear operation.

To this end, consider two solutions $\vec{y}_1, \vec{y}_2 \in M$ and a scalar α. Then we have

$$[\alpha\vec{y}_1 + \vec{y}_2]' = \alpha\vec{y}_1{}' + \vec{y}_2{}' = A\vec{y}_1 + A\vec{y}_2 = \vec{0} + \vec{0} = \vec{0}.$$

We see that also $\alpha\vec{y}_1 + \vec{y}_2$ solves the given homogeneous system, therefore it belongs to M.

Consequently, M is a linear space.

2. Choose some $x_0 \in I$.

Take any $i \in \{1, \dots, n\}$. According to the existence theorem, there must be some solution $\vec{y}_i$ of the given system that also satisfies the initial conditions $y_i(x_0) = \vec{e}_i$, where $\vec{e}_i$ are the canonical unit vectors. In this way we obtain vector functions $\{\vec{y}_1, \dots, \vec{y}_n\}$. As solutions they definitely belong to M. We claim that they are linearly independent on I.

Indeed, consider some null linear combination $\sum \alpha_i\vec{y}_i = \vec{0}$ on I. This must be in particular true at x_0, so $\sum \alpha_i\vec{y}_i(x_0) = \vec{0}$, that is, $\sum \alpha_i\vec{e}_i = \vec{0}$. Since the canonical basis $\{\vec{e}_i\}$ consists of linearly independent vectors, we conclude that $\alpha_i = 0$ for all i, in other words, only the trivial linear combination of $\{\vec{y}_i\}$ can produce the zero vector.

We confirmed the desired linear independence, in particular the dimension of M is at least n.

Now take any solution $\vec{y}_0$ of the given system. Then $\vec{\beta} = \vec{y}_0(x_0)$ is some vector from $\mathbb{R}^n$, and this $\vec{y}_0$ in fact solves the initial value problem given by the equation $\vec{y}\,' = A\vec{y}$ and the initial condition $\vec{y}(x_0) = \vec{\beta}$.

Consider also the vector function $\vec{y}_c = \sum \beta_i\vec{y}_i$. This vector function also solves the system $\vec{y}\,' = A\vec{y}$. Moreover, we have

$$\vec{y}_c(x_0) = \sum \beta_i\vec{y}_i(x_0) = \sum \beta_i\vec{e}_i = \vec{\beta}.$$

Consequently, $\vec{y}_c$ solves exactly the same initial value problem as $\vec{y}_0$. By the uniqueness theorem, these two must agree, that is, $\vec{y}_0 = \sum \beta_i\vec{y}_i$.

We just proved that the set $\{\vec{y}_i\}$ generates the space M, so it is in fact a basis and $\dim(M) = n$.

□

A basis of a solution space is naturally very important and there is special terminology related to it.

Definition 26b.2.
Consider a homogeneous system of linear ODEs $\vec{y}\,' = A\vec{y}$, where $A \in \mathbb{R}^{n\times n}$.
By a **fundamental system of solutions** of this system on an open interval I we mean an arbitrary basis of the space of all solutions of this system on I.
If $\{\vec{y}_1, \dots, \vec{y}_n\}$ is a fundamental system of solutions, then we define its **fundamental matrix** on I by $Y(x) = (\,\vec{y}_1(x) \quad \cdots \quad \vec{y}_n(x)\,)$
(an $n \times n$ matrix).

Example 26b.a: Consider the homogeneous system

$$y_1' = 2y_1 + y_2 - 3$$
$$y_2' = y_1 + 2y_2 + 3x - 4.$$

We found its general solution

$$\vec{y}_h(x) = a\begin{pmatrix} e^x \\ -e^x \end{pmatrix} + b\begin{pmatrix} e^{3x} \\ e^{3x} \end{pmatrix}$$

It would seem that there is a basis

$$\left\{\begin{pmatrix} e^x \\ -e^x \end{pmatrix}, \begin{pmatrix} e^{3x} \\ e^{3x} \end{pmatrix}\right\}.$$

This would set up the fundamental matrix
$$Y(x)=\begin{pmatrix} e^x & e^{3x}\\ -e^x & e^{3x}\end{pmatrix}.$$
Notice an interesting thing:
$$\vec{y}_h(x)=\begin{pmatrix} e^x & e^{3x}\\ -e^x & e^{3x}\end{pmatrix}\begin{pmatrix} a\\ b\end{pmatrix}.$$
$\triangle$

This last equality is sometimes very convenient.

Fact 26b.3.
Consider a homogeneous system of linear ODEs $\vec{y}\,'=A\vec{y}$, where $A\in\mathbb{R}^{n\times n}$.
If $Y(x)$ is its fundamental matrix on I, then a general solution of this system on I is $\vec{y}_h(x)=Y(x)\cdot\vec{c}$ for $\vec{c}\in\mathbb{R}^n$.

There also is a convenient analogue of Wronski's result for determining linear independence of solutions.

Theorem 26b.4.
Consider a homogeneous system of linear ODEs $\vec{y}\,'=A\vec{y}$, where $A\in\mathbb{R}^{n\times n}$.
Let $\vec{y}_1,\dots,\vec{y}_n$ be solutions of this system on an open interval I.
$\{\vec{y}_1,\dots,\vec{y}_n\}$ is a fundamental system of solutions of this system on I if and only if $\det(Y(x))\neq 0$ on I, which is true if and only if $\det(Y(x_0))\neq 0$ for some $x_0\in I$.

In short, any fundamental matrix $Y(x)$ is nonsingular for all $x\in I$. We will need this observation in chapter 27.

We have the theory, and now for the practical part: Where do we get that fundamental system?

Fact 26b.5.
Consider a homogeneous system of linear ODEs $\vec{y}\,'=A\vec{y}$ with matrix $A\in\mathbb{R}^{n\times n}$.
If λ_0 is an eigenvalue of A with associated eigenvector $\vec{v}$, then $\vec{y}=\vec{v}\,e^{\lambda_0 x}$ is a solution of the given system on $\mathbb{R}$.
If $\lambda_1,\dots,\lambda_k$ are distinct eigenvalues of the matrix A, then the corresponding solutions form a linearly independent set.

If you need to refresh you memory regarding eigenvalues, see chapter 30.

Example 26b.b: Consider the homogeneous system
$$\begin{aligned} y_1' &= 2y_1+y_2\\ y_2' &= y_1+2y_2.\end{aligned}$$
We find the eigenvalues of the matrix of the system $A=\begin{pmatrix} 2 & 1\\ 1 & 2\end{pmatrix}$.

We solve the appropriate equation:
$$\begin{aligned} 0=\det(A-\lambda E_n)&=\det\begin{vmatrix} 2-\lambda & 1\\ 1 & 2-\lambda\end{vmatrix}=(2-\lambda)^2-1^2\\ &=\lambda^2-4\lambda+3.\end{aligned}$$
We obtain the eigenvalues $\lambda=1,3$.

Now we find the associated eigenvectors.

$\lambda = 1$: We need to solve the homogeneous system $(A - \lambda E_n)\vec{v} = \vec{0}$, that is,

$$\begin{pmatrix} 2-1 & 1 \\ 1 & 2-1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This reads

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and we see that the equations are linearly dependent, as they should be. It is enough to solve the first equation $v_1 + v_2 = 0$. We can choose one variable, but we have to make sure that the resulting vector will not be trivial, so we have to choose a non-zero number. We set $v_1 = 1$ and obtain $v_2 = -1$.

We have the pair $\lambda = 1$, $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and according to the statement above, we can create one vector of functions for our basis:

$$\vec{y}_a = \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{1\cdot x} = \begin{pmatrix} e^x \\ -e^x \end{pmatrix}.$$

$\lambda = 3$: We need to solve the homogeneous system with the matrix

$$\left(\begin{array}{cc|c} 2-3 & 1 & 0 \\ 1 & 2-3 & 0 \end{array}\right) = \left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right).$$

Again, the rows are linearly dependent, which suggests that perhaps we did not make any mistake yet, and therefore we just look at one equation, for instance the first: $-v_1 + v_2 = 0$. Choosing $v_1 = 1$ we get $v_2 = 1$ and obtain the pair $\lambda = 3$, $\vec{v} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. The second vector of functions is

$$\vec{y}_b = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^{3\cdot x} = \begin{pmatrix} e^{3x} \\ e^{3x} \end{pmatrix}.$$

Now we can form a general solution:

$$\vec{y} = a\vec{y}_a + b\vec{y}_b = a\begin{pmatrix} e^x \\ -e^x \end{pmatrix} + b\begin{pmatrix} e^{3x} \\ e^{3x} \end{pmatrix},$$

that is,

$$\begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} = \begin{pmatrix} a\,e^x + b\,e^{3x} \\ -a\,e^x + b\,e^{3x} \end{pmatrix}, \quad x \in \mathbb{R}.$$

It is nice to write the answer in the same language as the question, so we shed the matrix overcoat and put it like this:

$$\begin{aligned} y_1(x) &= \ a\,e^x + b\,e^{3x}, \\ y_2(x) &= -a\,e^x + b\,e^{3x}, \ x \in \mathbb{R}. \end{aligned}$$

This is the same answer that we obtained in example 26.a.

Remark: Note that the equation $\lambda^2 - 4\lambda - 3 = 0$ that yielded the eigenvalues is exactly the same as the equation that provided us with characteristic numbers when we were solving the corresponding problem using elimination, and thus also the characteristic numbers were the same as the eigenvalues now. This is no surprise.

The polynomial $\lambda^2 - 4\lambda - 3$ and the numbers $\lambda = 1, 3$ capture the substance of the underlying situation (from physics for instance). When we transform a system into one equation or vice versa, we are just changing the language of the mathematical description, but the underlying process is still the same, and therefore the lambdas also stay the same.

△

The procedure that we just tried works for all systems of linear differential equations that have distinct real eigenvalues. To make our knowledge complete we have to find out how to handle other cases. For complex eigenvalues we have the usual answer.

Fact 26b.6.
Consider a homogeneous system of linear ODEs $\vec{y}' = A\vec{y}$ with matrix $A \in \mathbb{R}^{n\times n}$.
Let λ_0 be an eigenvalue of A with associated eigenvector $\vec{v}$.
If λ_0 is a complex number, that is, $\text{Im}(\lambda_0) \neq 0$, then $\text{Re}(\vec{v}e^{\lambda_0 x})$ and $\text{Im}(\vec{v}e^{\lambda_0 x})$ are linearly independent solutions of the given system on $\mathbb{R}$.

We used the same approach with linear equations of higher order, and there we obtained very convenient formulas using this principle. Indeed, from a complex solution $e^{(\alpha+\beta i)x}$ we recovered $e^{\alpha x}\cos(\beta x)$ and $e^{\alpha x}\sin(\beta x)$. For systems of equations the situation is not so simple, because for a complex eigenvalue λ and a real-valued matrix, the associated eigenvector $\vec{v}$ must also have some complex components, and thus we cannot readily identify what the real and imaginary parts of $\vec{v}e^{\lambda x}$ are in general.

Example 26b.c: Consider the oscillation equation $\ddot{x} + \omega^2 x = 0$. We transform it into a system of equation by denoting $\dot{x} = v$. We chose this name because we know that if x is the position of, say, a pendulum (angle or displacement), then $\dot{x}$ is the velocity. We obtain the system

$$\begin{aligned}\dot{x} &= v\\ \dot{v} &= -\omega^2 x.\end{aligned}$$

The matrix of this system is $\begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix}$. We find the eigenvalues:

$$0 = \det\begin{pmatrix} -\lambda & 1 \\ -\omega^2 & -\lambda \end{pmatrix} = \lambda^2 + \omega^2.$$

The roots are $\lambda = \pm\omega i$. Now we find an eigenvector associated with $\lambda = \omega i$:

$$\left(\begin{array}{cc|c} 0-\omega i & 1 & 0 \\ -\omega^2 & 0-\omega i & 0 \end{array}\right) = \left(\begin{array}{cc|c} -\omega i & 1 & 0 \\ -\omega^2 & -\omega i & 0 \end{array}\right)$$

We note that the second row is just the first one multiplied by $-\omega i$, so again it is enough to solve just one equation in this system, for instance the first one: $-\omega i v_1 + v_2 = 0$. We choose $v_1 = 1$ and obtain $v_2 = \omega i$. The eigenvector is $\begin{pmatrix} 1 \\ \omega i \end{pmatrix}$ and the corresponding solution is

$$\begin{aligned}\vec{y} &= \begin{pmatrix} 1 \\ \omega i \end{pmatrix} e^{\omega i t} = \begin{pmatrix} 1 \\ \omega i \end{pmatrix} \big(\cos(\omega t) + i\sin(\omega t)\big)\\ &= \begin{pmatrix} \cos(\omega t) + i\sin(\omega t) \\ \omega i\cos(\omega t) - \omega\sin(\omega t) \end{pmatrix}.\end{aligned}$$

We used t for the free variable, as x is already taken.

Now we take the real and imaginary parts to obtain two vectors for our fundamental system.

$$\vec{y}_a = \begin{pmatrix} \cos(\omega t) \\ -\omega\sin(\omega t) \end{pmatrix}, \qquad \vec{y}_b = \begin{pmatrix} \sin(\omega t) \\ \omega\cos(\omega t) \end{pmatrix}.$$

As usual, there is no need to work out the case $\lambda = -\omega i$, since it would yield the same solutions.

We obtain a general solution of the form

$$\begin{aligned}\begin{pmatrix} x(t) \\ v(t) \end{pmatrix} &= a\begin{pmatrix} \cos(\omega t) \\ -\omega\sin(\omega t) \end{pmatrix} + b\begin{pmatrix} \sin(\omega t) \\ \omega\cos(\omega t) \end{pmatrix}\\ &= \begin{pmatrix} a\cos(\omega t) + b\sin(\omega t) \\ -a\omega\sin(\omega t) + b\omega\cos(\omega t) \end{pmatrix}.\end{aligned}$$

We rewrite it properly:

$$x(t) = a\cos(\omega t) + b\sin(\omega t),$$
$$v(t) = -a\omega\sin(\omega t) + b\omega\cos(\omega t),\ t \in \mathbb{R}.$$

Going back to the original equation, its general solution is

$$x(t) = a\cos(\omega t) + b\sin(\omega t),\ t \in \mathbb{R}.$$

As a bonus for solving it through a system we also obtained information about velocity. Note that our transformation requires that $v = x'$ and we see that our solution satisfies this, which gives us hope that we did not make any mistake when solving it.
△

Eigenvalues of higher multiplicity are a bit more involved.

Fact 26b.7.
Consider a homogeneous system of linear ODEs $\vec{y}\,' = A\vec{y}$ with matrix $A \in \mathbb{R}^{n\times n}$. Let λ_0 be an eigenvalue of A of multiplicity m with associated eigenvector $\vec{v}$. Consider vectors defined as follows:

$\vec{v}_1 = \vec{v}$, so it is a solution of $(A - \lambda_0 E_n)\vec{x} = \vec{0}$,
$\vec{v}_2$ is a solution of $(A - \lambda_0 E_n)\vec{x} = \vec{v}_1$,
$\vec{v}_3$ is a solution of $(A - \lambda_0 E_n)\vec{x} = \vec{v}_2$,
$\vdots$
$\vec{v}_m$ is a solution of $(A - \lambda_0 E_n)\vec{x} = \vec{v}_{m-1}$.

Then the following functions are solutions of the given system on $\mathbb{R}$ and form a linearly independent set:

$$\vec{y} = \vec{v}_1 e^{\lambda_0 x},$$
$$\vec{y} = \Big[\int(\vec{v}_1)\,dx + \vec{v}_2\Big]e^{\lambda_0 x} = (\vec{v}_1 x + \vec{v}_2)e^{\lambda_0 x},$$
$$\vec{y} = \Big[\int(\vec{v}_1 x + \vec{v}_2)\,dx + \vec{v}_3\Big]e^{\lambda_0 x} = \big(\tfrac{1}{2}\vec{v}_1 x^2 + \vec{v}_2 x + \vec{v}_3\big)e^{\lambda_0 x},$$
$$\vdots$$
$$\vec{y} = \big(\tfrac{1}{(m-1)!}\vec{v}_1 x^{m-1} + \tfrac{1}{(m-2)!}\vec{v}_2 x^{m-2} + \cdots + \vec{v}_{m-1}x + \vec{v}_m\big)e^{\lambda_0 x}.$$

This is definitely more complicated than the corresponding procedure for higher order linear differential equations. The vectors $\vec{v}_i$ are called generalized eigenvectors.

Example 26b.d: Consider the initial value problem

$$y_1' = y_1 - y_2 \qquad y_1(0) = 13$$
$$y_2' = y_1 + 3y_2, \qquad y_2(0) = 23.$$

1. First we find its general solution. The matrix of the system is $\begin{pmatrix}1 & -1\\ 1 & 3\end{pmatrix}$ and we will find its eigenvalues:

$$0 = \det\begin{pmatrix}1-\lambda & -1\\ 1 & 3-\lambda\end{pmatrix} = (1-\lambda)(3-\lambda) - 1\cdot(-1)$$
$$= \lambda^2 - 4\lambda + 4 = (\lambda-2)^2.$$

We see that $\lambda = 2$ is an eigenvalue of multiplicity 2.

We will find the necessary chain of generalized eigenvectors, working with the matrix

$$A-2E_n=\begin{pmatrix}-1 & -1\\ 1 & 1\end{pmatrix}.$$

The first one is just the ordinary eigenvector, so we look at the system

$$\left(\begin{array}{cc|c}-1 & -1 & 0\\ 1 & 1 & 0\end{array}\right).$$

The second equation reads $v_1+v_2=0$, choosing $v_1=1$ we get $v_2=-1$ and we have the first vector in the chain $\vec{v}_1=\begin{pmatrix}1\\ -1\end{pmatrix}$.

We now use this vector as the right-hand side for our system:

$$\left(\begin{array}{cc|c}-1 & -1 & 1\\ 1 & 1 & -1\end{array}\right).$$

The rows are linearly dependent again, so it is enough to solve one equation, for instance the second: $v_1+v_2=1$. Note that now we can choose $v_1=0$, since the equation is no longer homogeneous and thus the resulting vector will not be trivial. We get $v_2=1$ and hence $\vec{v}_2=\begin{pmatrix}0\\ 1\end{pmatrix}$.

Now we form two solutions according to the theorem above.

$$\vec{y}_a=\begin{pmatrix}1\\ -1\end{pmatrix}e^{2x}=\begin{pmatrix}e^{2x}\\ -e^{2x}\end{pmatrix},$$

$$\vec{y}_b=\left[x\begin{pmatrix}1\\ -1\end{pmatrix}+\begin{pmatrix}0\\ 1\end{pmatrix}\right]e^{2x}=\begin{pmatrix}x\,e^{2x}\\ (-x-1)e^{2x}\end{pmatrix}.$$

This is enough to form a basis, so we write a general solution:

$$\vec{y}(x)=a\begin{pmatrix}e^{2x}\\ -e^{2x}\end{pmatrix}+b\begin{pmatrix}x\,e^{2x}\\ -(x+1)e^{2x}\end{pmatrix}.=\begin{pmatrix}a\,e^{2x}+bx\,e^{2x}\\ -a\,e^{2x}-b(x+1)e^{2x}\end{pmatrix}.$$

As usual, we prefer the form

$$\begin{aligned}y_1(x)&=ae^{2x}+bx\,e^{2x},\\ y_2(x)&=-ae^{2x}-b(x+1)e^{2x},\ x\in\mathbb{R}.\end{aligned}$$

2. Now we address the initial conditions. Knowing the formulas for y_1 and y_2, we can write them as

$$\begin{aligned}a\quad&=13\\ -a-b&=23\end{aligned}\implies a=13,\ b=-36.$$

The solution is

$$\begin{aligned}y_1(x)&=13e^{2x}-36x\,e^{2x},\\ y_2(x)&=-13e^{2x}+36(x+1)e^{2x},\ x\in\mathbb{R}.\end{aligned}$$

△

And that's it, folks. Now we can solve any homogeneous system, that is, as long as we can find the eigenvalues. Indeed, once the matrix is larger than 4×4, then determining eigenvalues in the usual way would require being able to solve polynomials of such high degrees, and we already know that there are no formulas for that.

It would be possible to find roots of characteristic polynomials numerically, but the preferred way is to approximate the eigenvalues directly. The chapter 30 does just that.

27. Nonhomogeneous systems of linear differential equations

We will follow the usual pattern. First, we introduce a structural theorem that shifts the burden of finding a complete solution to the homogeneous case.

Theorem 27.1. (on structure of solution set of systems of linear ODE)
Let $\vec{y}_p$ be some particular solution of a given system of linear ODEs on an open interval I.
A vector function $\vec{y}_0$ is a solution of this equation on I if and only if $\vec{y}_0 = \vec{y}_p + \vec{y}_h$ for some solution $\vec{y}_h$ of the associated homogeneous system in I.
Consequently, if $\vec{y}_h$ is a general solution of the associated homogeneous system on I, then $\vec{y}_p + \vec{y}_h$ is a general solution of the given equation on I.

The proof follows along the same lines as we saw when proving the analogous statement for high order linear equations, see theorem 16.1.

Proof: Assume that $\vec{y}_p$ is a solution of the given system $\vec{y}' = A\vec{y} + \vec{b}(x)$ on I, denote the system (S).

1) Let $\vec{y}_h$ be some homogeneous solution on I, then $\vec{y}_h{}' = A\vec{y}_h$. We claim that $\vec{y} = \vec{y}_p + \vec{y}_h$ is a solution of (S) on I. Indeed, for any $x \in I$ we get

$$\begin{aligned}[\vec{y}_p + \vec{y}_h]'(x) &= \vec{y}_p{}'(x) + \vec{y}_h{}'(x) = (A\vec{y}_p(x) + \vec{b}(x)) + A\vec{y}_h(x)\\ &= A(\vec{y}_p + \vec{y}_h)(x) + \vec{b}(x).\end{aligned}$$

2) Let $\vec{y}_0$ be any solution of (S). Consider the vector function $\vec{y}_h = \vec{y}_0 - \vec{y}_p$. Then obviously $\vec{y}_p + \vec{y}_h = \vec{y}_0$ and $\vec{y}_h$ solves the associated homogeneous system in I. Indeed,

$$\begin{aligned}\vec{y}_h{}'(x) &= [\vec{y}_0 - \vec{y}_p]'(x) = \vec{y}_0{}'(x) - \vec{y}_p{}'(x) = (A\vec{y}_0(x) + \vec{b}(x)) - (A\vec{y}_p(x) + \vec{b}(x))\\ &= A(\vec{y}_0 - \vec{y}_p)(x) = A\vec{y}_h(x).\end{aligned}$$

The claim is proved.

□

Now it remains to figure out how to get one particular solution of the given system.

27a. Variation of parameter

For nonhomogeneous systems of linear differential equations, the variation of parameters is the preferred method, especially where computer algebra systems are concerned.

We start be recalling the variation as applied to first order linear differential equations, see algorithm 9.1. It turns out that it carries over to first-order systems as well. We will show it on an example first and then express it in an algorithm.

Example 27a.a: We return to our favorite example.

$$\begin{aligned}y_1' &= 2y_1 + y_2 - 3\\ y_2' &= y_1 + 2y_2 + 3x - 4.\end{aligned}$$

The variation procedure calls for solving the associated homogeneous equation first, so we should probably interpret this as solving the homogeneous system now. Moreover, the solution should be written as $c \cdot u(x)$.

We actually solved the homogeneous system in example 26b.a and we obtained

$$\begin{aligned}y_{1h}(x) &= a\,e^x + b\,e^{3x},\\ y_{2h}(x) &= a(-e^x) + b\,e^{3x}.\end{aligned}$$

The solutions are of the form $y_i(x) = c_u u_i(x) + c_v v_i(x)$, which is close enough to the original variation. Following the variation way, we should start looking for a solution of the given system in the form

$$y_{1p}(x) = a(x)e^x + b(x)e^{3x},$$
$$y_{2p}(x) = a(x)(-e^x) + b(x)e^{3x}.$$

The original variation then offered two possibilities. We could substitute the new guess into the original equation, experience a miracle of cancelling and obtain an equation for $c'(x)$, or simply remember that this equation is $c'(x)u(x) = b(x)$.

We will see what happens when we substitute our guess into the given system.

$$[a(x)e^x + b(x)e^{3x}]' = 2(a(x)e^x + b(x)e^{3x}) + (-a(x)e^x + b(x)e^{3x}) - 3$$
$$[-a(x)e^x + b(x)e^{3x}]' = (a(x)e^x + b(x)e^{3x}) + 2(-a(x)e^x + b(x)e^{3x}) + 3x - 4.$$

This reads

$$a'(x)e^x + a(x)e^x + b'(x)e^{3x} + b(x)3e^{3x} = a(x)e^x + 3b(x)e^{3x} - 3$$
$$-a'(x)e^x - a(x)e^x + b'(x)e^{3x} + b(x)3e^{3x} = -a(x)e^x + 3b(x)e^{3x} + 3x - 4,$$

that is,

$$a'(x)e^x + b'(x)e^{3x} = -3$$
$$a'(x)(-e^x) + b'(x)e^{3x} = 3x - 4.$$

Our belief in miracles got rewarded again, and all $a(x)$ and $b(x)$ disappeared, leaving us with two equations featuring two unknown functions $a'(x)$, $b'(x)$; this sounds right. By the way, note that if we write the homogeneous solution as $y_i(x) = c_u u_i(x) + c_v v_i(x)$, we can see the equations that we deduced now as $c_u'(x)u_i(x) + c_v'(x)v_i(x) = b_i(x)$, which fits remarkably well with the pattern of variation for one equation.

Now we solve the system. One can use the Cramer rule (which is actually rather convenient when it comes to systems with functions instead of numbers). Elimination is also a good bet, both in the formal shape of Gaussian elimination and in its intuitive forms.

We will go the easy way and observe that adding the equations gets us rid of $a'(x)$. We obtain

$$2b'(x)e^{3x} = 3x - 7 \implies b'(x) = \tfrac{1}{2}(3x - 7)e^{-3x}.$$

Subtracting the second equation from the first we obtain

$$2a'(x)e^x = -3x + 1 \implies a'(x) = \tfrac{1}{2}(-3x + 1)e^{-x}.$$

The last major step is integration, here we have to use integration by parts.

$$a(x) = \int \tfrac{1}{2}(-3x+1)e^{-x}dx = -\tfrac{1}{2}(-3x+1)e^{-x} - \tfrac{1}{2}(-3)e^{-x}$$
$$= \big(\tfrac{3}{2}x + 1\big)e^{-x},$$
$$b(x) = \int \tfrac{1}{2}(3x-7)e^{-3x}dx = -\frac{1}{3}\tfrac{1}{2}(3x-7)e^{-3x} - \frac{1}{9}\tfrac{3}{2}e^{-3x}$$
$$= \big(\tfrac{1}{2}x + 1\big)e^{-3x}.$$

We found functions $a(x), b(x)$ and we can substitute them into our guess for a particular solution.

$$y_{1p}(x) = \big(\tfrac{3}{2}x + 1\big)e^{-x}e^x + \big(\tfrac{1}{2}x + 1\big)e^{-3x}e^{3x}$$
$$= \tfrac{3}{2}x + 1 + \tfrac{1}{2}x + 1 = 2x + 2,$$
$$y_{2p}(x) = -\big(\tfrac{3}{2}x + 1\big)e^{-x}e^x + \big(\tfrac{1}{2}x + 1\big)e^{-3x}e^{3x}$$
$$= -\tfrac{3}{2}x - 1 + \tfrac{1}{2}x + 1 = -x.$$

By a remarkable coincidence, these are exactly the same results that we obtained using elimination in example 26.a. It is just a coincidence, given that this problem has infinitely many particular solutions that we choose from, but it often happens that different methods find the same particular solution, and we know now that this solution is correct.

We will complete this solution by writing a general solution of our problem using the $\vec{y}_p + \vec{y}_h$

structure:

$$y_1(x) = x + 2 + a\,e^x + b\,e^{3x},$$
$$y_2(x) = -x - a\,e^x + b\,e^{3x},\ x \in \mathbb{R}.$$

Remark: We committed the crime of leaving out integration constants in our integrations above. It was a premeditated crime, since we knew that were were looking for just one particular solution, so any antiderivative would do.

There is an alternative approach. Like good boys and girls (or other intelligent galactic creatures), we could include integration constants like this:

$$a(x) = \int \tfrac{1}{2}(-3x+1)e^{-x}dx = \big(\tfrac{3}{2}x+1\big)e^{-x} + A,$$
$$b(x) = \int \tfrac{1}{2}(3x-7)e^{-3x}dx = \big(\tfrac{1}{2}x+1\big)e^{-3x} + B.$$

Now we substitute these into the variational guess:

$$\begin{aligned} y_{1p}(x) &= \big[\big(\tfrac{3}{2}x+1\big)e^{-x} + A\big]e^x + \big[\big(\tfrac{1}{2}x+1\big)e^{-3x} + B\big]e^{3x} \\ &= 2x + 2 + A\,e^x + B\,e^{3x}, \\ y_{2p}(x) &= \big[-\big(\tfrac{3}{2}x+1\big)e^{-x} + A\big]e^x + \big[\big(\tfrac{1}{2}x+1\big)e^{-3x} + B\big]e^{3x} \\ &= -x - A\,e^x + B\,e^{3x}. \end{aligned}$$

In this way we obtain a general solution directly.

We will prefer the first approach here because it ties in with our structural approach and the theory, but this alternative works equally well.

△

The procedure that we outlined above works in general. It is a hands-on version of variation, when we treat the system through individual equations and individual functions. In fact, it is my preferred approach if I am to solve systems by hand.

Recall that a general homogeneous solution can be found in the form $\vec{y}_h = \sum c_j\vec{u}_j$, where $\vec{u}_j$ are vectors of functions that solve the associated homogeneous equation. Each $\vec{u}_j$ is therefore a vector whose ith entry is used when creating the solution y_{ih}. To avoid double indexing, we will write the general solution as

$$\vec{y}_h = c_1\vec{u} + c_2\vec{v} + c_3\vec{w} + \cdots$$

and then we find the functions y_i in rows of the resulting vector: $\vec{y}_{ih} = c_1\vec{u}_i + c_2\vec{v}_i + c_3\vec{w}_i + \cdots$.

Algorithm 27a.1.

⟨variation of parameters method—row version⟩

Given: a system $\vec{y}\,' = A\vec{y} + \vec{b}(x)$, with A a real-valued $n \times n$ matrix.

1. Find a general solution $\vec{y}_h$ of the associated homogeneous system $\vec{y}\,' = A\vec{y}$, and express it as

$y_{1h}(x) = c_1u_1(x) + c_2v_1(x) + c_3w_1(x) + \cdots,$

$y_{2h}(x) = c_1u_2(x) + c_2v_2(x) + c_3w_2(x) + \cdots,$

$\cdots$

$y_{nh}(x) = c_1u_n(x) + c_2v_n(x) + c_3w_n(x) + \cdots.$

2. Seek a solution of the form

$y_1(x) = c_1(x)u_1(x) + c_2(x)v_1(x) + c_3(x)w_n(x) + \cdots,$

$\vdots$

$y_n(x) = c_1(x)u_n(x) + c_2(x)v_n(x) + c_3(x)w_n(x) + \cdots.$

Unknown functions $c_i(x)$ are found by solving the system of equations

$c_1'(x)u_1(x) + c_2'(x)v_1(x) + c_3'(x)w_n(x) + \cdots = b_1(x),$

$\vdots$

$c_1'(x)u_n(x) + c_2'(x)v_n(x) + c_3'(x)w_n(x) + \cdots = b_n(x).$

Solve this system of equations (using e.g. elimination or Cramer rule) for $c_1'(x), \dots, c_n'(x)$.
Use integration to find some antiderivatives $c_1(x), \dots, c_n(x)$.
Substitute these into modified $y_1, \dots, y_n$ to get $y_{1p}, \dots, y_{np}$.

3. The general solution is $y_i = y_{ip} + y_{ih}$ for $i = 1, \dots, n$.
△

The row-version of variation may be nice for human computers, but it is less convenient to silicon-based computers. They would definitely prefer an abstract, matrix-coded version.

Algorithm 27a.2.
⟨variation of parameters method—matrix version⟩
Given: a system $\vec{y}' = A\vec{y} + \vec{b}(x)$, with A a real-valued $n \times n$ matrix.

1. Find a general solution $\vec{y}_h$ of the associated homogeneous system $\vec{y}' = A\vec{y}$, and express it as $\vec{y} = Y(x) \cdot \vec{c}$.

2. Seek a solution of the form $\vec{y}_p = Y(x) \cdot \vec{c}(x)$.
Substituting it into the original equation one gets the equation $Y(x) \cdot \vec{c}\,'(x) = \vec{b}(x)$.
Then $\vec{c}\,'(x) = Y(x)^{-1}\vec{b}(x)$. Use integration to obtain $\vec{c}(x)$ and substitute it into $\vec{y}_p(x) = Y(x) \cdot \vec{c}(x)$.

3. A general solution is then $\vec{y} = \vec{y}_p + \vec{y}_h$.
△

Note how variation in this matrix point of view strongly resembles variation for one equation in algorithm 9.1. Moreover, all the operations in this procedure can be easily handled by any competent computer algebra system.

In fact, the matrix approach is nothing else but the row approach where we hide the unpleasant details using a convenient notation. To see this we revisit our example.

Example 27a.b: Now we consider the system

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} -3 \\ 3x-4 \end{pmatrix}.$$

We will write A for the system matrix and $\vec{b}$ for the right-hand side vector.

In example 26b.a we found a solution in the matrix form

$$\vec{y}_h(x) = \begin{pmatrix} e^x & e^{3x} \\ -e^x & e^{3x} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = Y(x)\vec{c}.$$

Now we start variation, and focus on vector functions of the form

$$\vec{y}_p(x) = Y(x)\vec{c}(x) = \begin{pmatrix} e^x & e^{3x} \\ -e^x & e^{3x} \end{pmatrix} \begin{pmatrix} a(x) \\ b(x) \end{pmatrix}.$$

According to algorithm, we should now solve the equation

$$Y(x)\vec{c}\,'(x) = \vec{b}(x).$$

When we look at the contents of $Y(x)$ and $\vec{c}(x)$, we find that these are exactly the equations that we obtained for $a'(x)$ and $b'(x)$ when solving this problem via row-variation in the example above. But this time we handle it using matrix tools and derive the following formula.

$$\vec{c}\,'(x) = Y(x)^{-1}\vec{b}(x).$$

We find the inverse matrix in the usual way.

$$\left(\begin{array}{cc|cc} e^x & e^{3x} & 1 & 0\\ -e^x & e^{3x} & 0 & 1\end{array}\right) \sim \left(\begin{array}{cc|cc} e^x & e^{3x} & 1 & 0\\ 0 & 2e^{3x} & 1 & 1\end{array}\right) \sim \left(\begin{array}{cc|cc} e^x & e^{3x} & 1 & 0\\ 0 & e^{3x} & \frac12 & \frac12\end{array}\right)$$
$$\sim \left(\begin{array}{cc|cc} e^x & 0 & \frac12 & -\frac12\\ 0 & e^{3x} & \frac12 & \frac12\end{array}\right) \sim \left(\begin{array}{cc|cc} 1 & 0 & \frac12 e^{-x} & -\frac12 e^{-x}\\ 0 & 1 & \frac12 e^{-3x} & \frac12 e^{-3x}\end{array}\right).$$

We see the inverse matrix on the right, now we can find $\vec{c}\,'(x)$:

$$\begin{pmatrix} a'(x)\\ b'(x)\end{pmatrix} = \begin{pmatrix} \frac12 e^{-x} & -\frac12 e^{-x}\\ \frac12 e^{-3x} & \frac12 e^{-3x}\end{pmatrix} \cdot \begin{pmatrix} -3\\ 3x-4\end{pmatrix}$$
$$= \begin{pmatrix} -3\frac12 e^{-x} - (3x-4)\frac12 e^{-x}\\ -3\frac12 e^{-3x} + (3x-4)\frac12 e^{-3x}\end{pmatrix}$$
$$= \begin{pmatrix} (-\frac32 x + \frac12)e^{-x}\\ (\frac32 x - \frac72)e^{-3x}\end{pmatrix}.$$

Note that the calculations themselves were exactly as with the row variation, as well as the outcome and the integration that we now face. We repeat it to obtain

$$\vec{c}(x) = \begin{pmatrix} a(x)\\ b(x)\end{pmatrix} = \begin{pmatrix} (\frac32 x + 1)e^{-x}\\ (\frac12 x + 1)e^{-3x}\end{pmatrix}.$$

Now we can substitute this into our guess for $\vec{y}_p$ to obtain

$$\vec{y}_p(x) = Y(x)\vec{c}(x) = \begin{pmatrix} e^x & e^{3x}\\ -e^x & e^{3x}\end{pmatrix} \begin{pmatrix} (\frac32 x + 1)e^{-x}\\ (\frac12 x + 1)e^{-3x}\end{pmatrix}$$
$$= \begin{pmatrix} (\frac32 x + 1)e^{-x}e^x + (\frac12 x + 1)e^{-3x}e^{3x}\\ -(\frac32 x + 1)e^{-x}e^x + (\frac12 x + 1)e^{-3x}e^{3x}\end{pmatrix}$$
$$= \begin{pmatrix} 2x+2\\ -x\end{pmatrix}.$$

Again, the calculations were exactly the same as with the row-variation.

△

The procedure seems to work, but as mathematicians we would like to see some justification of the steps above.

When we create our guess $\vec{y}(x) = Y(x)\vec{c}(x)$, we are supposed to substitute it into the given system:

$$[Y(x)\vec{c}(x)]' = AY(x)\vec{c}(x) + \vec{b}(x).$$

What next? We would like to differentiate on the left. Surprisingly enough, the product rule that we know for functions also works for matrices and vectors of functions. Here is a proof for a general matrix $G(x)$ and vector function $\vec{f}(x)$. First,

$$G(x)\vec{f}(x) = \begin{pmatrix} g_{1,1}(x) & \cdots & g_{1,n}(x)\\ \vdots & & \vdots\\ g_{n,1}(x) & \cdots & g_{n,n}(x)\end{pmatrix} \begin{pmatrix} f_1(x)\\ \vdots\\ f_n(x)\end{pmatrix} = \begin{pmatrix} \sum_j g_{1,j}(x)f_j(x)\\ \vdots\\ \sum_j g_{n,j}(x)f_j(x)\end{pmatrix}.$$

Therefore

$$
\begin{aligned}
[G(x)\vec{f}(x)]' &= \left[\begin{pmatrix} \sum_j g_{1,j}(x)f_j(x) \\ \vdots \\ \sum_j g_{n,j}(x)f_j(x) \end{pmatrix}\right]' = \begin{pmatrix} \sum_j g'_{1,j}(x)f_j(x) + \sum_j g_{1,j}(x)f'_j(x) \\ \vdots \\ \sum_j g'_{n,j}(x)f_j(x) + \sum_j g_{n,j}(x)f'_j(x) \end{pmatrix} \\
&= \begin{pmatrix} \sum_j g'_{1,j}(x)f_j(x) \\ \vdots \\ \sum_j g'_{n,j}(x)f_j(x) \end{pmatrix} + \begin{pmatrix} \sum_j g_{1,j}(x)f'_j(x) \\ \vdots \\ \sum_j g_{n,j}(x)f'_j(x) \end{pmatrix} \\
&= \begin{pmatrix} g'_{1,1}(x) & \cdots & g'_{1,n}(x) \\ \vdots & & \vdots \\ g'_{n,1}(x) & \cdots & g'_{n,n}(x) \end{pmatrix} \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix} + \begin{pmatrix} g_{1,1}(x) & \cdots & g_{1,n}(x) \\ \vdots & & \vdots \\ g_{n,1}(x) & \cdots & g_{n,n}(x) \end{pmatrix} \begin{pmatrix} f'_1(x) \\ \vdots \\ f'_n(x) \end{pmatrix} \\
&= G'(x)\vec{f}(x) + G(x)\vec{f}\,'(x).
\end{aligned}
$$

Applying this to our equations we obtain

$$Y'(x)\vec{c}(x) + Y(x)\vec{c}\,'(x) = AY(x)\vec{c}(x) + \vec{b}(x).$$

Now we expect some cancelling to happen. We would like to argue that $Y'(x)\vec{c}(x)$ and $AY(x)\vec{c}(x)$ are the same expression. We could confirm it by rewriting these expressions into rows, in effect going back to the row variation. We prefer to give a matrix-based argument.

The justification starts by recalling a general trick from linear algebra. Given an $n \times n$ matrix $G = (g_{ij})$ and a vector $\vec{h}$ of length n, we can write the product $G\vec{h}$ as

$$G\vec{h} = \sum_i h_i \overrightarrow{(g_{i,j})}_j,$$

where $\overrightarrow{(g_{i,j})}_j$ represents the ith column of the matrix G. Applying it to our situation, we first recall that the ith column of $Y(x)$ is actually $\vec{y}_i$, the ith vector of our fundamental system and, in particular, a solution of the associated homogeneous system. Thus we deduce that

$$Y'(x)\vec{c}(x) = \sum_i c_i(x)\vec{y}_i\,' = \sum_i c_i(x)A\vec{y}_i.$$

The other expression $AY(x)\vec{c}(x)$ represents a linear combination of columns of $AY(x)$, but every column of this matrix is in fact $A\vec{y}_i$, so

$$AY(x)\vec{c}(x) = \sum_i c_i(x)A\vec{y}_i.$$

This confirms that indeed, $Y(x)\vec{c}(x) = AY(x)\vec{c}(x)$ and the cancelling does happen. We obtain the equation

$$Y(x)\vec{c}\,'(x) = \vec{b}(x)$$

as stated by the algorithm.

The next questionable step is the application of $Y(x)^{-1}$. How do we know that this inverse matrix actually exists? For the answer we return to chapter 26, where we recognized linear independence of the proposed basis by checking whether the matrix formed by these vectors is regular. In other words, the matrix $Y(x)$ is invertible.

The last unjustified step is the integration phase. There we can offer two answers. In all the theorems in this chapter we assume that we work on an interval I where all functions involved are continuous, and then also the expressions that we integrate are continuous and hence integrable. That is the theoretical point of view. The practical point of view is that in this step we actually rely on luck, because we are not always able to actually find antiderivatives for continuous functions.

This is the one weak point of variation, and it is not different from variation for one equation of order one as we saw it before. We simply have to get lucky once in a while.

For another example of variation (in its row version) see below.

27b. Method of undetermined coefficients

The method of undetermined coefficients proved to be fairly efficient for solving high order linear differential equations. Could it be applied also to systems?

The answer is in the positive, but with some modifications that make it less pleasant. There are two problems to consider.

First, imagine that there is an expression of suitable type, say, $13e^{5x}$, on the right in one of the equations. We know that we should feed our equation with $A\,e^{5x}$ (if we ignore possible corrections for a moment), but since all unknown functions $y_1,\dots,y_n$ appear (at least potentially) in every equation, we do not know which of the functions y_i should receive the term $A\,e^{5x}$ when we form our guesses. Consequently, we have to include such a term in all function, and each copy with its own constant, because it can easily happen that several functions join forces to create the desired output.

In short, whenever we see some expression on the right in one of the equations, all candidates y_i have to include the corresponding general term. But then it can easily happen that the same type appears in several equations, which would make us add another term of this type into our guesses. However, we do not want unnecessary constants, so we merge terms of the same type and ultimately, the one with the highest degree of polynomial swallows all the others. In effect, when putting a certain type into guesses for y_i, we only consider the one with highest degree of polynomial.

Second, imagine that we have a term of a certain type on the right of some equation, and its special number λ also happens to be an eigenvalue of the system. Then we have to correct, and if this eigenvalue is of higher multiplicity, then we use stronger correction. However, due to mutual interaction of equations, it may turn out that the original guess actually survives passage through the system. In other words, we never know whether the correction is actually needed. This means that we should add to our guess (in fact to all guesses for y_i per the previous observation) not just the term with the proper correction, but also corresponding terms with smaller corrections all the way to a term without any correction at all.

In effect, this creates a full polynomial of a raised degree. To see this, imagine that the basic form of guess is $(Ax+B)e^{ax}$, and that a happens to be a double eigenvalue, that is $m=2$. Then we should add into our guesses the terms

$$x^2(Ax+B)e^{ax}+x(Cx+D)e^{ax}+(Ex+F)e^{ax}=(Ax^3+(B+C)x^2+(D+E)x+F)e^{ax}.$$

Since we do not want unnecessary constants, we simply treat this as a full polynomial of degree $1+2$.

All this can make the guessing method very labor-intensive.

Algorithm 27b.1.
⟨method of undetermined coefficients for systems of linear ODEs⟩
Given: a system $\vec{y}\,'=A\vec{y}+\vec{b}(x)$, with A a real-valued $n\times n$ matrix.
Assumptions: Right-hand sides are linear combinations of quasipolynomials, that is, of expressions of the type $p(x)e^{\alpha x}\cos(\beta x)+q(x)e^{\alpha x}\sin(\beta x)$

1. Find a general solution $y_{1h},\dots,y_{nh}$ of the associated homogeneous system $\vec{y}\,'=A\vec{y}$, in particular determine the eigenvalues λ of the matrix A.

2. Scan the right-hand sides for all quasipolynomials and make a list.
Collate all terms of the same type as determined by the special number $\lambda=\alpha+\beta i$, and in each group, leave only the term with the highest power of polynomial. If there is a tie, pick just any.

3. Create candidates $y_1,\dots,y_n$ for a particular solution as follows:

a) For each quasipolynomial on the list, guess the basic form of a solution. Assume that the degree of the polynomial in it is d.
b) For the special number of this term, determine its multiplicity m as an eigenvalue, including the case $m = 0$ if it is not an eigenvalue at all. Modify the basic form to feature a full polynomial of degree $d + m$.
c) Add the modified form of the guess to all candidates for y_i.

4. Substitute the created guesses $y_1, \dots, y_n$ into the given system and by comparing terms, set up a system of linear equations that determines unknown constants.
Solve this system, substitute the known constants into the guessed forms $y_1, \dots, y_n$ and obtain particular solutions $y_{1p}, \dots, y_{np}$.

5. The general solution is $y_i = y_{ip} + y_{ih}$ for $i = 1, \dots, n$.
△

Example 27b.a: Consider the system

$$\begin{aligned} y_1' &= \dots + \cos(2x) + 2x + 1 - 13e^{2x} \\ y_2' &= \dots + x^2 + x\,e^{2x} + 23\sin(x) \\ y_3' &= \dots + e^x + 14\sin(2x) - 13. \end{aligned}$$

We actually do not care much about the actual matrix of the system here, but we need to know that its eigenvalues are $\lambda = 0, 2, 2$.

What types do we see among the right-hand sides?

• There are two terms of the trigonometric type with frequency $\beta = 2$ (special number $\lambda = 2i$), namely $\cos(2x)$ and $14\sin(2x)$. Both have a polynomial of degree zero attached, so the basic form of the guess should be $A\cos(2x) + B\sin(2x)$. Since the special number does not match any eigenvalue, we do not have to increase the degree of the polynomial.

• There are three polynomial terms (special number $\lambda = 0$), namely $2x + 1$, x^2 and -13. The highest degree is two, so the basic form of guess would be $Ax^2 + Bx + C$. However, the special number matches one of the eigenvalues (multiplicity 1), so we have to increase the degree by one and include a full cubic polynomial in all guesses.

• There are two terms of exponential type with factor $\alpha = 2$ (special number $\lambda = 2$), namely $-13e^{2x}$ and $x\,e^{2x}$. The higher degree wins, so the basic form of our guess is $(Ax+B)e^{2x}$. However, the special number matches a double eigenvalue, so we have to increase the degree of the polynomial by two, and a cubic polynomial should be attached to this exponential.

• There is the term $13\sin(x)$ of the trigonometric type with frequency $\beta = 1$ (special number $\lambda = i$), so the basic form of the guess should be $A\cos(x) + B\sin(x)$. Since the special number does not match any eigenvalue, we do not have to increase the degree of the polynomial.

• Finally, there is the term e^x of exponential type with factor $\alpha = 1$ (special number $\lambda = 1$), so the basic form of our guess is $A\,e^x$. Since the special number does not match any eigenvalue, we do not have to increase the degree of the polynomial.

We therefore seek a solution of our system of the form

$$\begin{aligned} y_1 &= A\cos(2x) + B\sin(2x) + Cx^3 + Dx^2 + Ex + F + (Gx^3 + Hx^2 + Ix + J)e^{2x} \\ &\quad + K\cos(x) + L\sin(x) + M\,e^x, \\ y_2 &= O\cos(2x) + P\sin(2x) + Qx^3 + Rx^2 + Sx + T + (Ux^3 + Vx^2 + Wx + X)e^{2x} \\ &\quad + Y\cos(x) + Z\sin(x) + a\,e^x, \\ y_3 &= b\cos(2x) + c\sin(2x) + dx^3 + fx^2 + gx + h + (jx^3 + kx^2 + lx + m)e^{2x} \\ &\quad + n\cos(x) + o\sin(x) + p\,e^x. \end{aligned}$$

Substituting these into the given system, we would hope to obtain 39 equations for 39 unknown coefficients. This explains why we did not really care about the actual system, we are definitely

not going through this.
△

Example 27b.b: For the last time we return to the example 26.a,

$$y_1' = 2y_1 + y_2 - 3$$
$$y_2' = y_1 + 2y_2 + 3x - 4.$$

Its homogeneous solution is

$$y_{1h}(x) = a\,e^x + b\,e^{3x},$$
$$y_{2h}(x) = -a\,e^x + b\,e^{3x},\ x \in \mathbb{R}.$$

Now we find candidates for a particular solution. Surveying the right-hand sides we see that they feature just one type, namely a polynomial (special number $\lambda = 0$), in two incarnations, one with degree zero and the other with degree one. According to our observations above, the higher degree wins. The special number does not match any eigenvalue, we therefore do not need to increase the degree of polynomial, and thus we should include polynomials of order 1 in our guesses:

$$y_1(x) = Ax + B,$$
$$y_2(x) = Cx + D.$$

To find the constants, we substitute our guesses into the given system:

$$\begin{matrix} [Ax+B]' = 2(Ax+B) + (Cx+D) - 3 \\ [Cx+D]' = (Ax+B) + 2(Cx+D) + 3x - 4 \end{matrix} \implies \begin{matrix} (-2A-C)x + (A-2B-D) = -3 \\ (-A-2C)x + (-B+C-2D) = 3x-4. \end{matrix}$$

Comparing both sides we see four equations that we easily solve:

$$\begin{aligned} -2A - C &= 0 \\ A - 2B - D &= -3 \\ -A - 2C &= 3 \\ -B + C - 2D &= -4 \end{aligned} \implies \begin{aligned} A &= 1 \\ C &= -2 \\ -2B - D &= -4 \\ -B - 2D &= -2 \end{aligned} \implies \begin{aligned} A &= 1 \\ C &= -2 \\ B &= 2 \\ D &= 0. \end{aligned}$$

We found the following particular solution.

$$y_{1p}(x) = x + 2,$$
$$y_{2p}(x) = -2x.$$

It agrees with the one obtained by elimination and variation, so we are quite confident that it is correct.
△

Example 27b.c: Consider the following initial value problem:

$$y_1' = y_1 + y_2 + 1 \qquad y_1(0) = 2$$
$$y_2' = y_1 + y_2 + 1 - e^x, \qquad y_2(0) = 0.$$

1. First we find its general solution.

a) Homogeneous version: the matrix is $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. In the usual way we find $\lambda = 0, 2$. We determine associated eigenvectors and obtain the homogeneous solution

$$y_{1h} = a + b\,e^{2x},$$
$$y_{2h} = -a + b\,e^{2x}.$$

b) Now we find a particular solution.

We start with the variation method (row version), and look for a solution of the form

$$y_{1p} = a(x) + b(x)e^{2x},$$
$$y_{2p} = -a(x) + b(x)e^{2x}.$$

Substituting our candidates into the given system we obtain equations

$$\begin{aligned} a'(x) + b'(x)e^{2x} &= 1\\ -a'(x) + b'(x)e^{2x} &= 1 - e^x. \end{aligned}$$

Normally we would solve this by adding and subtracting these equations, but to show an alternative, we will use the Cramer rule:

$$\begin{aligned} D &= \det\begin{pmatrix} 1 & e^{2x}\\ -1 & e^{2x} \end{pmatrix} = e^{2x} + e^{2x} = 2e^{2x},\\ D_a &= \det\begin{pmatrix} 1 & e^{2x}\\ 1-e^x & e^{2x} \end{pmatrix} = e^{2x} - e^{2x}(1-e^x) = e^{3x},\\ D_b &= \det\begin{pmatrix} 1 & 1\\ -1 & 1-e^x \end{pmatrix} = 1 - e^x + 1 = 2 - e^x. \end{aligned}$$

Note that D is in fact the determinant of the fundamental matrix $Y(x)$, so it will never be zero. We can therefore find

$$\begin{aligned} a'(x) &= \frac{D_a}{D} = \tfrac{1}{2}e^x\\ &\implies a(x) = \tfrac{1}{2}e^x,\\ b'(x) &= \frac{D_b}{D} = e^{-2x} - \tfrac{1}{2}e^{-x}\\ &\implies b(x) = -\tfrac{1}{2}e^{-2x} + \tfrac{1}{2}e^{-x}. \end{aligned}$$

We can substitute into our candidates and obtain

$$\begin{aligned} y_{1p} &= \big(\tfrac{1}{2}e^x\big) + \big(-\tfrac{1}{2}e^{-2x} + \tfrac{1}{2}e^{-x}\big)e^{2x} = e^x - \tfrac{1}{2},\\ y_{2p} &= -\big(\tfrac{1}{2}e^x\big) + \big(-\tfrac{1}{2}e^{-2x} + \tfrac{1}{2}e^{-x}\big)e^{2x} = -\tfrac{1}{2}. \end{aligned}$$

Using the formula $\vec{y}_p + \vec{y}_h$ we obtain a general solution

$$\begin{aligned} y_1 &= e^x - \tfrac{1}{2} + a + b\,e^{2x},\\ y_2 &= -\tfrac{1}{2} - a + b\,e^{2x},\ x \in \mathbb{R}. \end{aligned}$$

Next, we will try the guessing method. Scanning the right-hand sides we see a polynomial term appearing twice, each time as just the constant 1, so the basic form of our guess is just a constant polynomial A. However, the corresponding special number $\lambda = 0$ matches one of the eigenvalues, so we have to make a correction by one and use a general polynomial of order one in our guess.

We also see an exponential term $-e^x$, leading to the basic guess $A\,e^x$. The corresponding special number $\lambda = 1$ does of match eigenvalues, so there will be no correction.

We therefore form the following guess for solutions:

$$\begin{aligned} y_{1p} &= A\,e^x + Bx + C,\\ y_{2p} &= D\,e^x + Ex + F. \end{aligned}$$

We substitute these into the given system:

$$\begin{aligned} [A\,e^x + Bx + C]' &= (A\,e^x + Bx + C) + (D\,e^x + Ex + F) + 1\\ [D\,e^x + Ex + F]' &= (A\,e^x + Bx + C) + (D\,e^x + Ex + F) + 1 - e^x \end{aligned}$$

$$\implies \begin{aligned} (-B-E)x + (B-C-F) - D\,e^x &= 1\\ (-B-E)x + (-C+E-F) - A\,e^x &= 1 - e^x. \end{aligned}$$

Comparing both sides we obtain the following system:

$$\begin{aligned} -B - E &= 0 & -B - E &= 0\\ B - C - F &= 1 & -C + E - F &= 1\\ -D &= 0 & -A &= -1. \end{aligned}$$

We immediately see that one equation is twice there, which means that this system us undetermined and we are free to choose one variable as we wish. This sometimes happens when solving systems by guessing, in fact it can be expected when we have an eigenvalue match and a resulting correction happening.

However, it is not clear which unknown can be freely chosen, and if we pick the wrong one, we may create an unsolvable system that way. We have to simplify the system before we see how things stand.

We immediately obtain $D = 0$ and $A = 1$, and thus our system reduces to

$$\begin{aligned} B + E &= 0 \\ B - C - F &= 1 \\ -C + E - F &= 1 \end{aligned} \Longrightarrow \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 1 \\ 0 & -1 & 1 & -1 & 1 \end{array}\right)$$

We use elimination to reduce the matrix:

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 1 & -1 & 0 & -1 & 1 \\ 0 & -1 & 1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & -1 & 1 & -1 & 1 \end{array}\right) \sim \left(\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & -1 \\ 0 & 0 & 2 & 0 & 0 \end{array}\right)$$

The third row shows that $E = 0$ is given, hence also $B = 0$ by the first equation. Surprisingly, the corrected terms Bx and Ex are not needed, although the special number $\lambda = 0$ did match an eigenvalue. This sometimes happens with systems, but you never know until you try it, so we always have to include general corrected terms when a match occurs.

The shape of the matrix suggests that we can choose F. We opt for the traditional $F = 0$, then $C = -1$ and we get the particular solution

$$\begin{aligned} y_{1p} &= e^x - 1, \\ y_{2p} &= 0. \end{aligned}$$

We see that y_{2p} has nothing which is a bit unfair, perhaps we should have chosen something else for F, but it is too late now, life is tough.

What is more interesting, we obtained a different particular solution compared to our previous attempt using variation. This can happen, there are infinitely many possible solutions after all. This time we get a general solution

$$\begin{aligned} y_1 &= e^x - 1 + a + b\,e^{2x}, \\ y_2 &= -a + b\,e^{2x},\ x \in \mathbb{R}. \end{aligned}$$

This is actually incorrect. The way we wrote it, it would seem that the homogeneous parts of this and the previous attempt match while particular solutions differ, so they would describe different sets of functions. In fact, the parameters a, b here are different from those in the previous solution, so properly we should write

$$\begin{aligned} y_1 &= e^x - 1 + \tilde{a} + \tilde{b}\,e^{2x}, \\ y_2 &= -\tilde{a} + \tilde{b}\,e^{2x},\ x \in \mathbb{R}. \end{aligned}$$

Does this general formula describe the same set of solutions as the previous one? In other words, when we create a specific pair of functions using one description, can we also obtain the same pair from the other description (perhaps using different values of parameters)?

We start by observing that if we use $a = -\frac{1}{2}$ and $b = 0$ in the first general solution, we obtain $y_1 = e^x - 1$ and $y_2 = 0$, which is exactly the particular solution we have now. Conversely, the particular solution obtained by variation can be obtained by substituting $\tilde{a} = \frac{1}{2}$ and $\tilde{b} = 0$ into the new solution.

This is an inspiration for a general transformation: We claim that if the first formula generates some solution using specific values a, b, then the new solution creates the same functions when using parameters $\tilde{a} = a - \frac{1}{2}$, $\tilde{b} = b$. Indeed, when we substitute these values into the new formula, we obtain exactly the previous version of general solutions. In this way we confirm that these two formulas describe the same set of solutions.

2. To handle the initial conditions we use the second version of a general solution, because it is fraction-free. We will omit those tildas for simplicity.

$$\begin{matrix} y_1(0)=2 \\ y_2(0)=0 \end{matrix} \implies \begin{matrix} e^0-1+a+be^0=2 \\ -a+be^0=0 \end{matrix} \implies \begin{matrix} a+b=2 \\ -a+b=0. \end{matrix}$$

We conclude that $a=b=1$ and the answer to the given question is

$$y_1(x)=e^{2x}+e^x,$$
$$y_2(x)=e^{2x}-1,\ x\in\mathbb{R}.$$

If we used the first form of a general solution, we would have to use parameters $a=\frac{1}{2}$, $b=1$ to obtain this solution.

△

27c. Bonus: Differential elimination

In chapter 20 we talked of differentiation as if it were a mapping. We will push this idea further here.

We start by denoting the differential operator $f\mapsto f'$ with the letter D. So $D[y]=y'$. It is actually customary to skip the brackets, people would often write just Dy. We can rewrite differential equations using this notation, for instance

$$y''-3y'+2y=0 \implies D[D[y]]-3D[y]+2y=0.$$

This is where it gets interesting. The second derivative means that we apply derivative to an outcome of the previous derivative, $y''=[y']'$. In the mapping language, we use the output of the mapping D as an argument of another edition of D, $D[D[y]]$. This is a familiar concept, namely the composition of mappings. NOrmally we would write $D\circ D$, but when composing a mapping with itself, we call it a power of this mapping and denote it using the square notation. Therefore y'' can be captured as $D^2[y]$ or simply D^2y for short.

Similarly, $D^3y=y'''$ and so on. We can therefore write

$$y''-3y'+2y=0 \implies D^2y-3Dy+2y=0.$$

Time for another review When we see $2\ln(x)+\sin(x)$, we can interpret it in two ways. We can see two outcomes of functions, namely $\ln(x)$ and $\sin(x)$, combined into a linear combination. Or, we can see it as the function $2\ln+\sin$ applied to x. Indeed, when we define the addition and multiplication of functions as objects, we do it by the formula

$$(f+g)(x)=f(x)+g(x), \qquad (f\cdot g)(x)=f(x)\cdot g(x).$$

Therefore we can see $D^2y-3Dy+2y$ as the outcome of the mapping D^2-3D+2 being applied to a function y.

$$y''-3y'+2y=0 \implies (D^2-3D+2)[y]=0.$$

Now it gets interesting. If D was just a number we could write $D^2-3D+2=(D-2)(D-1)$. The interesting thing is that this factorization is also valid for the differential operator D if we interpret the multiplication of mappings in this formula as a composition. In other words, we will understand it as follows:

$$(D-2)(D-1)[y]=(D-2)[(D-1)[y]].$$

Does it work? Let's see:

$$\begin{aligned}(D-2)[(D-1)[y]]&=(D-2)[Dy-y]=D[y'-y]-2(y'-y)=[y'-y]'-2y'+2y\\ &=y''-y'-2y'+2y=y''-3y'+2y=(D^2-3D+2)[y].\end{aligned}$$

Moreover, we even have a commutative law in place. For instance, we easily check that

$$(D-2)(D-1)[y]=(D-1)(D-2)[y].$$

We can express the key property as follows:

• If we interpret compositions of D as multiplications of D, then linear combinations of the operator D behave just like polynomials.

This can be used to treat systems of linear differential equations in a completely novel way. For the start, every such system can be rewritten so that the unknown functions appear on the left.

Example 27c.a: We return to our favorite example 26.a:

$$\begin{aligned} y_1' &= 2y_1 + y_2 - 3 \\ y_2' &= y_1 + 2y_2 + 3x - 4. \end{aligned}$$

We rewrite it as

$$\begin{aligned} y_1' - 2y_1 - y_2 &= -3 \\ y_2' - y_1 - 2y_2 &= 3x - 4. \end{aligned}$$

Using our new notation we can write it also as

$$\begin{aligned} (D-2)y_1 - \quad y_2 &= -3 \\ -y_1 + (D-2)y_2 &= 3x - 4. \end{aligned}$$

Now for the punchline:

$$\begin{pmatrix} (D-2) & -1 \\ -1 & D-2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -3 \\ 3x-4 \end{pmatrix}.$$

△

At this point we may really look at all linear equations, not just those where only one unknown function is allowed to be differentiated. For instance, the system

$$\begin{aligned} y_1'' + y_2' - 2y_1 + 3y_2 &= e^x \\ y_1' + y_2' + 4y_1 - y_2 &= 13x \end{aligned}$$

can be written as

$$\begin{pmatrix} (D^2-2) & D+3 \\ D+4 & D-1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} e^x \\ 13 \end{pmatrix}.$$

The beauty of this notation is that we can now work with this using normal tools from linear algebra, treating individual entries that feature D as if they were polynomials in D. For instance, systems can be solved using elimination. However, certain special strategies should be followed, and they correspond to strategies related to polynomial matrices.

When working with a certain column, we compare the candidates for the pivot. Treating the entries as polynomials, we choose for the pivot the one with the lowest degree. Using row operations, it should be possible to make sure that in the new matrix, the entries in the pivot column feature only lower powers of D than the degree of the pivot. When removing higher powers of D in a certain entry, we usually start from the highest and work our way down, as far as possible. Sometimes, with smart multiplication of the pivot row, we can remove several powers in one step.

If all the entries below the pivot become zero, then this stage is done, and we also reduce the terms above the pivot as much as possible. Otherwise, we choose for our new pivot the row with the lowest degree in the pivot column and repeat the above process. In this way it should be possible to achieve zeros below the pivot, and have the pivot with the smallest possible degree, which is very desirable.

To illustrate the main points of this type of elimination we will focus on just the pivot column now. This means that we are addressing a certain unknown function in our system. So consider a certain system where y_1 appears as follows:

$$\begin{matrix} y_1'' - y_1 \\ y_1'' + 2y_1' + y_1 \\ y_1''' + y_1'' \end{matrix} \implies \begin{pmatrix} D^2 - 1 \\ D^2 + 2D + 1 \\ D^3 + D^2 \end{pmatrix}.$$

For the pivot we should choose the term with lowest degree, and we have two there. Then it is better to use the simpler one, so in fact the term in the first row is the best candidate for our pivot already. Now we should be able to remove D^2 and also all higher powers from the remaining terms in the column. We do so easily in the second row simply by subtracting the first row from it.

With the third row we first focus on D^3. We remove it when we we subtract the first row "multiplied" by D. In effect this means that we add another order of derivative to the operator represented by the term D^2-1. Obviously, if we were solving a system, all these operations would be acting in parallel on all other columns. We obtain

$$\begin{pmatrix} D^2-1 \\ D^2+2D+1 \\ D^3+D^2 \end{pmatrix} \sim \begin{pmatrix} D^2-1 \\ 2D+2 \\ D^2+D \end{pmatrix}.$$

We got rid of D^3 in the third row, and now we focus on D^2 there. This is easy to handle, we just subtract the first row.

$$\begin{pmatrix} D^2-1 \\ 2D+2 \\ D^2+D \end{pmatrix} \sim \begin{pmatrix} D^2-1 \\ 2D+2 \\ D+1 \end{pmatrix}.$$

This is as far as our pivot can get us, and we start the whole process again. First we choose a new pivot, the term with the lowest degree, and again there are two candidates. We will choose the one from the third row and then use it to reduce the middle row by subtracting the new pivot row twice.

$$\begin{pmatrix} D^2-1 \\ 2D+2 \\ D^2+D \end{pmatrix} \sim \begin{pmatrix} D+1 \\ 2D+2 \\ D^2-1 \end{pmatrix} \sim \begin{pmatrix} D+1 \\ 0 \\ D^2-1 \end{pmatrix}.$$

Now we address the third row. Normally we would first reduce the order by subtracting the first row multiplied by D, and then handle the first power by a suitable subtraction again. However, we may notice that D^2-1 is in fact a multiple (in the world of polynomials in D) of the pivot, so we can use the usual approach from elimination. That is, we multiply the pivot row by $D-1$ and subtract it from the third row. We obtain

$$\begin{pmatrix} D+1 \\ 0 \\ D^2-1 \end{pmatrix} \sim \begin{pmatrix} D+1 \\ 0 \\ 0 \end{pmatrix}.$$

Obviously, all these operations would be acting in parallel on all other columns in an actual matrix of a system. In the end we obtain a new system where y_1 appear only in the first equation, namely in the form $y_1'+y_1$. We can do nothing more with it, and it is enough.

Example 27c.b: We return to our favorite example.

$$\begin{aligned} y_1' &= 2y_1+y_2-3 \\ y_2' &= y_1+2y_2+3x-4 \end{aligned} \implies \begin{aligned} (D-2)y_1 - y_2 &= -3 \\ -y_1+(D-2)y_2 &= 3x-4 \end{aligned} \implies \left(\begin{array}{cc|c} D-2 & -1 & -3 \\ -1 & D-2 & 3x-4 \end{array}\right).$$

In the first column, the second entry has a lower degree, so we take it for the pivot.

$$\left(\begin{array}{cc|c} D-2 & -1 & -3 \\ -1 & D-2 & 3x-4 \end{array}\right) \sim \left(\begin{array}{cc|c} -1 & D-2 & 3x-4 \\ D-2 & -1 & -3 \end{array}\right).$$

We multiply the first row by $D-2$, so it becomes $(-D+2\ (D-2)^2 \mid (D-2)[3x-4])$. The term in the third column is actually

$$(D-2)[3x-4] = [3x-4]' - 2(3x-4) = 3-6x+8 = 11-6x.$$

We add this to the second row and obtain

$$\left(\begin{array}{cc|c} -1 & D-2 & 3x-4 \\ D-2 & -1 & -3 \end{array}\right) \sim \left(\begin{array}{cc|c} -1 & D-2 & 3x-4 \\ 0 & (D-2)^2-1 & 8-6x \end{array}\right) \sim \left(\begin{array}{cc|c} -1 & D-2 & 3x-4 \\ 0 & D^2-4D+3 & 8-6x \end{array}\right).$$

This concludes the first stage, reduction of the first column, and we move to the second column. Since the pivot there does not have a lower degree than the term above it, we cannot do any simplification there, so the elimination stops. We obtained the system

$$\begin{aligned} -y_1 + (D-2)[y_2] &= 3x-4 \\ (D^2-4D+3)[y_2] &= -6x+8 \end{aligned} \implies \begin{aligned} -y_1 + y_2' - 2y_2 &= 3x-4 \\ (D^2-4D+3)y_2'' = 4y_2' + 3y_2 &= -6x+8. \end{aligned}$$

This is already solvable. The second equation is just a single linear equation of order two that we can handle. The first equation than represents a formula that allows us to derive y_1 from y_2.

Solving the second equation we obtain, $y_2(x) = -2x + a\,e^x + b\,e^{3x}$. We substitute it into the first and get $y_1(x) = x + 2 - a\,e^x + b\,e^{3x}$. These are not exactly the same formulas as those that we obtained in example 26.a, but they describe the same set of solutions. Indeed, after substituting $-a$ for a here the formulas match.

△

Remarkably, we can also use the Cramer rule. For systems of linear algebraic equations we use it in the form

$$x_i = \frac{\det(A_i)}{\det(A)}.$$

For systems if linear differential equations, the form $\det(A)y_i = \det(A_i)$ is preferable, as we do not quite know how to divide with differential operators. However, there is an unexpected twist there.

Example 27c.c: We return to our favorite system

$$\left(\begin{array}{cc|c} D-2 & -1 & -3 \\ -1 & D-2 & 3x-4 \end{array}\right).$$

We apply the formulas, first for y_1:

$$\begin{aligned} \det\begin{pmatrix} D-2 & -1 \\ -1 & D-2 \end{pmatrix} y_1 &= \det\begin{pmatrix} -3 & -1 \\ 3x-4 & D-2 \end{pmatrix} \\ &\implies [(D-2)^2-1]y_1 = (D-2)[-3] - (-1)(3x-4) \\ &\implies [D^2-4D+3]y_1 = [-3]' - 2\cdot(-3) + (3x-4). \end{aligned}$$

We obtain the equation $y_1'' - 4y_1' + 3y_1 = 3x + 2$. By a remarkable coincidence, we arrived at this equation when solving this system by elimination in example 26.a, and we obtained the solution $y_1(x) = x + 2 + a\,e^x + b\,e^{3x}$.

Now we address the second unknown function.

$$\begin{aligned} \det\begin{pmatrix} D-2 & -1 \\ -1 & D-2 \end{pmatrix} y_2 &= \det\begin{pmatrix} D-2 & -3 \\ -1 & 3x-4 \end{pmatrix} \\ &\implies [(D-2)^2-1]y_2 = (D-2)[3x-4] - (-1)(-3). \end{aligned}$$

We obtain the equation $y_2'' - 4y_2' + 3y_2 = -6x + 8$. We solved it above, and we know that $y_2(x) = -2x + c\,e^x + d\,e^{3x}$.

Which brings us to the unexpected twist. We found the functions y_1, y_2 independently, and thus there is no reason to assume any connection between the parameters a, b in y_1 and c, d in y_2. However, there is a relationship (after all, the space of solutions is not four-dimensional), and we find this by substituting our solutions into the original system. We obtain

$$\begin{aligned} [x+2+a\,e^x+b\,e^{3x}]' &= 2[x+2+a\,e^x+b\,e^{3x}] + [-2x+c\,e^x+d\,e^{3x}] - 3 \\ [-2x+c\,e^x+d\,e^{3x}]' &= [x+2+a\,e^x+b\,e^{3x}] + 2[-2x+c\,e^x+d\,e^{3x}] + 3x - 4, \end{aligned}$$

that is,

$$\begin{aligned} a\,e^x + 3b\,e^{3x} &= 2a\,e^x + 2b\,e^{3x} + c\,e^x + d\,e^{3x} \\ c\,e^x + 3d\,e^{3x} &= a\,e^x + b\,e^{3x} + 2c\,e^x + 2d\,e^{3x} \end{aligned} \implies \begin{aligned} (a+c)e^x + (-b+d)e^{3x} &= 0 \\ (a+c)e^x + (b-d)e^{3x} &= 0. \end{aligned}$$

Since the functions e^x, e^{3x} are linearly independent, the first equality can be true only if $a+c=0$ and $-b+d=0$. Similarly, the second equality demands that $a+c=0$ and $b-d=0$, but these are actually the same equations as from the first equality. We conclude that $d=b$ and $c=-a$. We substitute these into the formula for y_2 and obtain the solution

$$\begin{aligned} y_1(x) &= x+2+a\,e^x+b\,e^{3x},\\ y_2(x) &= -2x-a\,e^x+b\,e^{3x},\ x\in\mathbb{R}. \end{aligned}$$

This agrees with our previous results.
△

We see the using the Cramer rule, solving an $n\times n$ system would require solving n linear differential equations of higher order. This is very likely more work than elimination, where with a bit of luck it is enough to solve just one such equation and the remaining functions can be obtained by substituting into direct formulas. However, this is not always the case, and sometimes we also have to solve more differential equations when using elimination.

The operator approach for elimination is a tool that can be also applied to the problem of reducing a system to just one equation, see chapter 26.

Example 27c.d: We return to example 26.b. First we rewrite it properly.

$$\begin{aligned} Dy_1 &= y_1+2y_2+y_3+1 & & (D-1)y_1-2y_2-y_3=1\\ Dy_2 &= -y_1+2y_2+2y_3 & \implies\quad & y_1+(D-2)y_2-2y_3=0\\ Dy_3 &= 2y_1+y_2+y_3+x & & -2y_1-y_2+(D-1)y_3=x \end{aligned}$$

$$\begin{aligned} & y_1+(D-2)y_2-2y_3=0\\ \implies\quad & (D-1)y_1-2y_2-y_3=1\\ & -2y_1-y_2+(D-1)y_3=x \end{aligned}$$

If we get lucky, an unknown appears in one of the equations without the differential operator D, then we can use that equation to eliminate this unknown from the other equations. In this particular case we decided to use the second equation with isolated y_1, so we moved it to the top.

Informally, we could express y_1 from the first equation and substitute into others, here we will show a formal elimination approach. We add double of the first row to the third, and apply $D-1$ to the first row before we subtract it from the second. The first row then looks like this:

$$(D-1)y_1+(D-1)(D-2)y_2-2(D-1)y_3=0.$$

We obtain

$$\begin{aligned} y_1+(D-2)y_2-2y_3&=0\\ [-(D-1)(D-2)-2]y_2+[2(D-1)-1]y_3&=1\\ [(2D-4)-1]y_2+[(D-1)-4]y_3&=x\\ y_1+(D-2)y_2-2y_3&=0\\ \implies\quad (-D^2+3D-4)y_2+(2D-3)y_3&=1\\ (2D-5)y_2+(D-5)y_3&=x \end{aligned}$$

Now we have a problem, because none of the equations features an unknown without the operator D, which stopped us when attempting an intuitive evaluation. HOwever, this time we look at it as a general elimination. We choose a pivot, and the third row looks more appealing (just one derivative instead of the second order).

$$\begin{aligned} y_1+(D-2)y_2-2y_3&=0\\ (2D-5)y_2+(D-5)y_3&=x\\ (-D^2+3D-4)y_2+(2D-3)y_3&=1 \end{aligned}$$

If D^2-3D+4 had $2D-5$ as a factor, we would easily get rid of the corresponding term by subtracting a suitable multiple of the second row from the third. Unfortunately, this is not true. We will therefore multiply the third row by $2D-5$, and then add the second row multiplied by D^2-3D+4. This second row then becomes

$$(2D^3-11D^2+23D-20)y_2+(D^3-8D^2+19D-20)y_3=(D^2-3D+4)x=4x-3$$

Here we go,

$$\begin{aligned} y_1+(D-2)y_2-2y_3&=0\\ (2D-5)y_2+(D-5)y_3&=x\\ (2D-5)(-D^2+3D-4)y_2+(2D-5)(2D-3)y_3&=1 \end{aligned}$$

$$\Longrightarrow\quad \begin{aligned} y_1+(D-2)y_2-2y_3&=0\\ (2D-5)y_2+(D-5)y_3&=x\\ (-2D^3+11D^2-23D+20)y_2+(4D^2-16D+15)y_3&=-5 \end{aligned}$$

$$\Longrightarrow\quad \begin{aligned} &y_1+(D-2)y_2-2y_3=0\\ &(2D-5)y_2+(D-5)y_3=x\\ &(D^3-4D^2+3D-5)y_3=4x-8 \end{aligned}$$

The last row is the equation we expect after elimination: It only features one unknown function and has order three:

$$y_3'''-4y_3''+3y_3'-5y_3=4x-8.$$

The first row offers a reduction formula that will determine y_1 once we find the other two:

$$y_1=-y_2'+2y_2+2y_3.$$

In the second row we would expect a reduction formula for y_2, and we do get one:

$$2y_2'-5y_2=x-y_3'+5y_3.$$

Unfortunately, it has the form of a differential equation, but we cannot do any better than this.

Obviously, elimination has its problems even using differential calculus.

27c.1 Differential calculus and homogeneous equations

Assume that we were given some equation of the form $L[y]=0$ with constant coefficient. Since the left-hand side is linear and has constant coefficients, we should be able to write it as $L[y]=p(D)[y]$, where $p(D)$ is some polynomial with the differential operator D.

For instance, given the equation $y''-4y'+3y=0$, we can write it as $(D^2-4D+3)[y]=0$.

We can factor the polynomial into linear factors and obtain

$$(D-3)(D-1)[y]=0.$$

We easily observe that if some function y satisfies $(D-1)[y]=0$, then it also satisfies $(D-3)(D-1)[y]=0$. Since $(D-3)(D-1)=(D-1)(D-3)$ also for differential operators, similar argument tells us that solutions of the equation $(D-3)[y]=0$ solve the given equation. Remarkably, this also works the other way around, every solution of the equation $(D-3)(D-1)[y]=0$ must solve one of the simpler equations.

This is the key to our favorite method of solving homogeneous equations. Given some equation of the form $p(D)[y]=0$, if the polynomial $p(D)$ factors into linear factors, that is, if the polynomial $p(\lambda)$ has only real roots, then we obtain all solutions of this equation simply by solving equations corresponding to the linear factors. And since equations of the form $(D-a)[y]=0$ have the obvious solution e^{ax}, we see why we build our fundamental systems out of exponentials.

27c.2 Differential calculus and the guessing method

The differential calculus offers a way to transform nonhomogeneous equations into homogeneous. To be of any use, we only care about the case when the original equation and the new are both linear.

Assume that we were given some equation of the form $L[y] = b(x)$ with constant coefficient, so it is of the form $p(D)[y] = b(x)$ for some polynomial $p(D)$ with the differential operator D.

For instance, given the equation $y'' - 4y' + 3y = e^{2x}$, we can write it as $(D^2 - 4D + 3)[y] = e^{2x}$.

Assume further that we are able to find some differential operation that can turn the right-hand side $b(x)$ into zero, and to be of any use, assume that this operator is also some polynomial in D, we can call it $q(D)$.

For instance, we know that e^{2x} solves $y' - 2y = 0$, that is, $(D-2)[y] = 0$, which tells us that $(D-2)[e^{2x}] = 0$. In this case therefore $q(D) = D - 2$.

Now we apply this $q(D)$ to the original equation $p(D)[y] = b(x)$ and obtain a new equation $q(D)p(D)[y] = 0$. This new equation is homogeneous, and it is still linear with constant coefficients. Moreover, every solution of the original equation is also a solution of the new equation. This means that we can actually find the solutions of the original nonhomogeneous problems among solutions of the new homogeneous problem.

If p and q have distinct roots, and therefore decompose into distinct linear factors, we can easily separate solutions from the original homogeneous equation (these solve $p(D)[y] = 0$) from the solutions that are new, they must come from the factors of $q(D)$. Thus we look for our particular solutions among the solutions of the equation $q(D)[y] = 0$, which in our particular example means among solutions of $y' - 2y = 0$. This leads to the guess $A\,e^{2x}$.

Things get more problematic when $p(D)$ and $q(D)$ share linear factors, and careful analysis shows the need for a corrective factor.

This approach works for any right-hand side for which we can find a linear differential operator that turns it into zero, and when we make a survey, we end up with the usual suspects: polynomials, exponentials, sines and cosines, and products of the three.