

16.A: Math

Vector Basics

Classical mechanics describes the motion of bodies as they move through space. To describe a motion in space it is not sufficient to give a position and a speed: you need a direction as well. Therefore we work with vectors: mathematical objects that have both a magnitude and a direction. If you tell me you're moving, I know something, but not much; I'll know more if you tell me you're moving at walking speed, and have full information of your velocity once you tell me that you're moving at walking speed towards the coffee machine. Although in principle we could make do with specifying a magnitude and direction of every vector in this way, it is often more convenient to express our vectors in a basis. To do so, we choose an (arbitrary) origin, and as many basis vectors as we have spatial dimensions, in such a way that they are not parallel to one another, and usually mutually perpendicular (orthogonal) and of unit length (orthonormal). Then we decompose our vector by giving its components along each of the basis vectors. The most common choice is to use a Cartesian basis, of two or three (depending on spatial dimension) basis vectors of unit length pointing in the standard x, y and z directions, and indicated as \hat{x} , \hat{y} and \hat{z} , or (rather annoyingly) sometimes as \mathbf{i} , \mathbf{j} and \mathbf{k} , the latter especially in American textbooks. Other often encountered systems are polar coordinates (2D) and cylindrical and spherical coordinates (3D), see the mathematical appendix for more background on those. To write our vectors, we now specify the components in each direction, writing for example $\mathbf{v} = 3\hat{x} + 3\hat{y}$ for a vector (in boldface) representing a speed of $3\sqrt{2}$ and a direction making an angle of 45° with the horizontal.

Vectors can be added and subtracted just like scalars - simply add and subtract them by component. Graphically, you add two vectors by putting them head-to-tail: you can find the sum of two vectors \mathbf{v} and \mathbf{w} by putting the start of \mathbf{w} at the end of \mathbf{v} , the sum $\mathbf{v} + \mathbf{w}$ then points from the start of \mathbf{v} to the end of \mathbf{w} . You can also multiply a vector by a scalar, by multiplying every component of the vector with that scalar. Graphically, this means that you extend the length of the vector with the scalar factor you just multiplied with.

You can't take the product of two vectors like you would two scalars. There are however two vector operations that closely resemble the product, known as the inner (or dot) and outer (or cross) product, see Figure 16.A.1. The dot product represents the length of the projection of one vector on another (and thus gives a scalar); it is zero for perpendicular vectors, and the dot product of a vector with itself gives the square of its length. To calculate the dot product of two vectors, sum the products of their components: if $\mathbf{v} = v_x\hat{x} + v_y\hat{y}$ and $\mathbf{w} = w_x\hat{x} + w_y\hat{y}$, then $\mathbf{v} \cdot \mathbf{w} = v_x w_x + v_y w_y$. You can use the dot product to find the angle between two vectors, using standard geometry, which gives

$$\cos \theta = \frac{\mathbf{v} \cdot \mathbf{w}}{|\mathbf{v}||\mathbf{w}|} = \frac{v_x w_x + v_y w_y}{|\mathbf{v}||\mathbf{w}|} \quad (16.A.1)$$

where $|\mathbf{v}|$ and $|\mathbf{w}|$ are the lengths of vectors \mathbf{v} and \mathbf{w} , respectively. The cross product is only defined for three-dimensional vectors, say $\mathbf{v} = v_x\hat{x} + v_y\hat{y} + v_z\hat{z}$ and $\mathbf{w} = w_x\hat{x} + w_y\hat{y} + w_z\hat{z}$. The result is another vector, with a direction perpendicular to the plane spanned by \mathbf{v} and \mathbf{w} , and a magnitude equal to the area of the parallelogram bounded by them. The cross product is most easily expressed in column vector form:

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} \times \begin{pmatrix} w_x \\ w_y \\ w_z \end{pmatrix} = \begin{pmatrix} v_y w_z - v_z w_y \\ v_z w_x - v_x w_z \\ v_x w_y - v_y w_x \end{pmatrix} \quad (16.A.2)$$

The cross product of a vector with itself is zero.

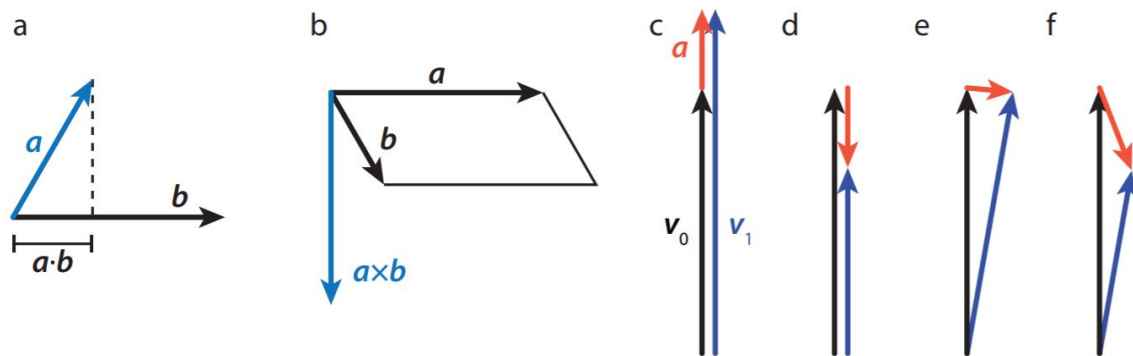


Figure 16.A. 1: Products and derivatives of vectors. (a) The dot product of two vectors \mathbf{a} and \mathbf{b} gives the length of the projection of \mathbf{a} on \mathbf{b} (or vice versa). (b) The cross product of two vectors \mathbf{a} and \mathbf{b} gives a vector perpendicular to the plane spanned by \mathbf{a} and \mathbf{b} and with a magnitude equal to the area of the parallelogram spanned by \mathbf{a} and \mathbf{b} . (c-f) Acceleration as a change in velocity. In every panel, the original velocity is shown in black, the velocity a short time Δt later is shown in blue, and the acceleration is shown in red. (c) Acceleration for an increase in the magnitude of the velocity. (d) Acceleration for a decrease in the magnitude of the velocity. (e) Acceleration for a change in direction of the velocity at constant magnitude. (f) Acceleration for a change in velocity that involves both a change in direction and a decrease in magnitude.

Vectors can be functions, just like scalar quantities: they can depend on one or more parameters, like position or time. Also, again just like scalar functions, you can calculate a rate of change of vector function as you move through parameter values, for instance asking how the velocity of a car changes as a function of time. An instantaneous rate of change is simply a derivative, which is calculated in exactly the same manner as the derivative of a scalar function. For example, the rate of change of the velocity, known as the acceleration \mathbf{a} , is defined as:

$$\mathbf{a} = \lim_{\Delta t \rightarrow 0} \frac{\mathbf{v}(t + \Delta t) - \mathbf{v}}{\Delta t} \quad (16.A.3)$$

Since the velocity itself is the derivative of the position $\mathbf{x}(t)$, the acceleration is also the second derivative of the position. Time derivatives occur so frequently in classical mechanics that we use a special notation for them: a first derivative is indicated by a dot on top of the quantity, and a second derivative by a double dot - so we have $\mathbf{a} = \dot{\mathbf{x}} = \ddot{\mathbf{x}}$.

Vector derivatives are somewhat richer than those of scalar functions, since there are more ways that a vector can change. Like a scalar function, the magnitude of a vector can increase or decrease. Moreover, its direction can also change, which also means that it has a nonzero derivative, and of course, you can have a combination of a change in magnitude and a change in direction, see Figure 16.A.1.

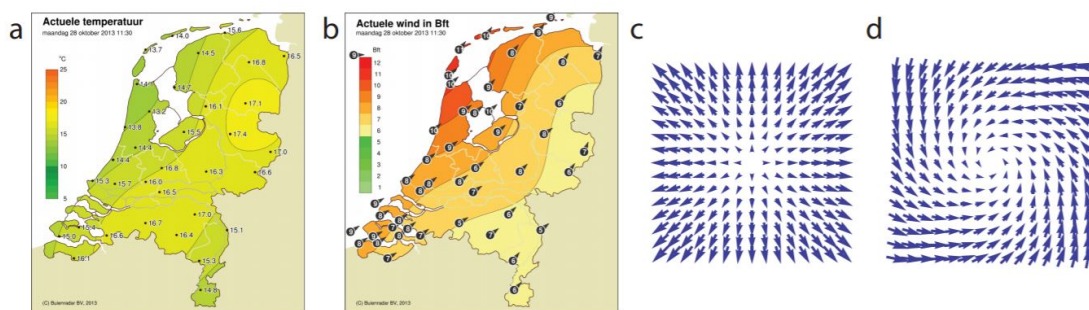


Figure 16.A.2: Examples of scalar and vector fields. (a) Temperature (a scalar) at every point in the Netherlands at 11:30 on October 28, 2013. Lines of equal temperature are drawn in black; the gradient of the temperature field is everywhere perpendicular to these lines. (b) Wind velocity (a vector) at every point in the Netherlands at the same time as the temperature map. Colors correspond to the magnitude of the velocity. (a and b) from [30]. (c) Example of a vector field with a large curl. (d) Example of a vector field with a large divergence and zero curl.

Functions (scalar or vector) that are defined at every point in space are sometimes called fields. Examples are the temperature (scalar) and wind (vector) at every point on the planet, see Figure 16.A.2. Just like you can calculate the rate of change of a function in time, you can also consider how a function changes in space. For a scalar function, this quantity is a vector, known as the gradient, defined as the vector of partial derivatives. For a function $f(x, y, z)$, we have:

$$\nabla f = \frac{\partial f}{\partial x} \hat{x} + \frac{\partial f}{\partial y} \hat{y} + \frac{\partial f}{\partial z} \hat{z} = \begin{pmatrix} \partial f / \partial x \\ \partial f / \partial y \\ \partial f / \partial z \end{pmatrix} \quad (16.A.4)$$

The direction of ∇f is the direction of maximal change, and its magnitude tells you how quickly the function changes in that direction. For a vector field \mathbf{v} , we can't take the gradient, but we can use the 'vector' ∇ of partial derivatives combined with either the dot or cross product. The first option is known as the divergence of \mathbf{v} , and tells you how quickly \mathbf{v} spreads out; the second is the curl of \mathbf{v} and tells you how much \mathbf{v} rotates:

$$\text{div}(\mathbf{v}) = \nabla \cdot \mathbf{v} = \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z} \quad (16.A.5)$$

$$\text{curl}(\mathbf{v}) = \nabla \times \mathbf{v} = \begin{pmatrix} \partial_y v_z - \partial_z v_y \\ \partial_z v_x - \partial_x v_z \\ \partial_x v_y - \partial_y v_x \end{pmatrix} \quad (16.A.6)$$

where $\partial_x = \frac{\partial}{\partial x}$, and so on.

Polar Coordinates

You can specify any point in the plane by specifying its projection on two perpendicular axes - we typically call these the x and y-axes and x and y coordinates. In this Cartesian system (named after Descartes), we identify unit vectors \hat{x} and \hat{y} , pointing along their respective axes, and being of unit length. A position \mathbf{r} can then be decomposed in the two directions: $\mathbf{r} = r_x \hat{x} + r_y \hat{y}$, with $r_x = \mathbf{r} \cdot \hat{x}$ and $r_y = \mathbf{r} \cdot \hat{y}$. Alternatively, we can write $\hat{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\hat{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, which gives for \mathbf{r} :

$$\mathbf{r} = r_x \hat{x} + r_y \hat{y} = \begin{pmatrix} r_x \\ r_y \end{pmatrix} \quad (16.A.7)$$

Instead of specifying the x and y coordinates of our position, we could also uniquely identify it by giving two different numbers: its distance to the origin r , and the angle θ the line to the origin makes with a fixed reference axis (typically the x-axis), see Figure 16.A.3. Invoking the Pythagorean theorem and basic trigonometry,

we readily find $r = \sqrt{r_x^2 + r_y^2}$ and $\tan \theta = \frac{r_y}{r_x}$. We call r the length of the vector \mathbf{r} . We could also invert the relations for r and θ so we can get the Cartesian components if the length and angle are known: $r_x = r \cos \theta$ and $r_y = r \sin \theta$.

Like the Cartesian basis vectors \hat{x} and \hat{y} , which point in the direction of increasing x and y values, we can also define unit vectors pointing in the direction of increasing r and θ . These directions do depend on our position in space, but they do have a clear geometrical interpretation: \hat{r} always points radially outward from the origin, and $\hat{\theta}$ in the direction you'd move if you'd be making a counterclockwise rotation about the origin. Given a position vector \mathbf{r} , finding the vector in the direction of increasing r is very easy: $\hat{r} = \mathbf{r}/r$. The expression for r in our new polar basis ($\hat{r}, \hat{\theta}$) is almost tautological: $\mathbf{r} = r \hat{r}$.

Relating the polar basis vectors to the Cartesian ones is straightforward. We have:

$$\mathbf{r} = r_x \hat{x} + r_y \hat{y} = r \hat{r} \quad (16.A.8)$$

and using $r_x = r \cos \theta$, $r_y = r \sin \theta$ we also have

$$\mathbf{r} = r \cos \theta \hat{x} + r \sin \theta \hat{y} \quad (16.A.9)$$

We thus find that $\hat{r} = \cos \theta \hat{x} + \sin \theta \hat{y}$.

For $\hat{\theta}$ we note that to rotate around the origin, the direction of motion needs to be perpendicular to \hat{r} . There are of course two such directions - we pick the sign by demanding that the counterclockwise rotation is positive. This gives $\hat{\theta} = (r_y/r) \hat{x} - (r_x/r) \hat{y} = \sin \theta \hat{x} - \cos \theta \hat{y}$. Written out as vectors, we have:

$$\hat{r} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}, \quad \hat{\theta} = \begin{pmatrix} \sin \theta \\ -\cos \theta \end{pmatrix} \quad (16.A.10)$$

Note that

$$\hat{r} = \frac{\partial \hat{\theta}}{\partial \theta}, \quad \hat{\theta} = -\frac{\partial \hat{r}}{\partial \theta} \quad (16.A.11)$$

Naturally, we can also express the Cartesian basis in terms of the polar ones:

$$\hat{x} = \cos \theta \hat{r} + \sin \theta \hat{\theta}, \quad \hat{y} = \sin \theta \hat{r} - \cos \theta \hat{\theta} \quad (16.A.12)$$

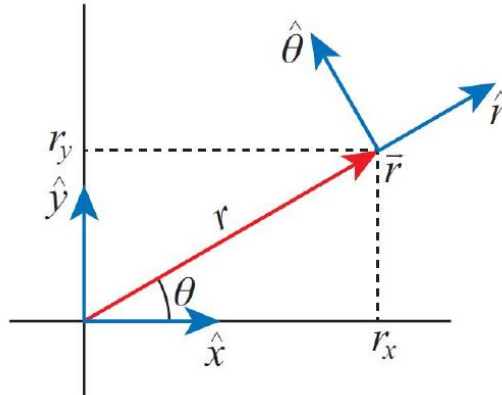


Figure 16.A.3: Polar coordinate system and polar unit vectors. A position \mathbf{r} in the plane can be specified either by giving its projections on two reference axes (along the \hat{x} and \hat{y} direction), or by giving its distance r to the origin, and the angle θ the line to the origin makes with the x axis. The polar basis vectors are defined as pointing in the direction of increasing r (i.e., radially outward), and increasing θ (i.e., rotating counterclockwise around the origin).

Solving Differential Equations

A differential equation is an equation which contains derivatives of the function to be determined. They can be very simple. For example, you may be given the (constant) velocity of a car, which is the derivative of its position, which we'd write mathematically as:

$$v = \frac{dx}{dt} = v_0 \quad (16.A.13)$$

To determine where the car ends up after one hour, we need to solve this differential equation. We also need a second piece of information: where the car was at some reference time (usually $t = 0$), the initial condition. If $x(0) = 0$, you don't need advanced maths skills to figure out that $x(1\text{hour}) = v_0 \cdot (1\text{hour})$. Unfortunately, things aren't usually this easy.

Before we proceed to a few techniques for solving differential equations, we need some terminology. The order of a differential equation is the order of the highest derivative found in the equation; Equation 16.A.13 is thus of first order. A differential equation is called ordinary if it only contains derivatives with respect to one variable, and partial if it contains derivatives to multiple variables. The equation is linear if it does not contain any products of (derivatives of) the unknown function. Finally, a differential equation is homogeneous if it only contains terms that contain the unknown function, and inhomogeneous if it also contains other terms. Equation 16.A.13 is ordinary and inhomogeneous, as the v_0 term on the right does not contain the unknown function $x(t)$. In the sections below, we discuss the various cases you'll encounter in this book; there are many others (many of which can't be solved explicitly) to which a whole subfield of mathematics is dedicated.

A.3.1. FIRST-ORDER LINEAR ORDINARY DIFFERENTIAL EQUATIONS

Suppose we have a general equation of the form

$$a(t) \frac{dx}{dt} + b(t)x(t) = f(t) \quad (16.A.14)$$

where $a(t)$, $b(t)$ and $f(t)$ are known functions of t , and $x(t)$ is our unknown function. Equation 16.A.14 is a first-order, ordinary, linear, inhomogeneous differential equation. In order to solve it we will use two techniques that are tremendously useful: separation of variables and separation into homogeneous and particular solutions.

Suppose we had $f(t) = 0$. Then, if we had two solutions $x_1(t)$ and $x_2(t)$ of Equation 16.A.14, we could construct a third as $x_1(t) + x_2(t)$ (or any linear combination of $x_1(t)$ and $x_2(t)$), since the equation is linear. Now since $f(t)$ is not zero, we can't do

this, but we can do something else. First, we find the most general solution to the equation where $f(t) = 0$, which we call the homogeneous solution $x_h(t)$. Second, we find a solution (any at all) of the full Equation 16.A.14, which we call the particular solution $x_p(t)$. The full solution is then the sum of these two solutions, $x(t) = x_h(t) + x_p(t)$. You may worry that there may be multiple particular solutions: how would we pick the ‘right’ one? Fortunately, we don’t need to worry: the homogeneous solution will contain an unknown variable, which will be set by the initial condition. Changing the particular solution will change the value of the variable, such that the final solution will be the same and satisfy both the differential equation and the initial condition.

To find the solution to the homogeneous equation

$$a(t) \frac{dx_h}{dt} + b(t)x_h(t) = 0 \quad (16.A.15)$$

we’re going to use a technique called separation of variables. There are two variables in this system: the independent parameter t and the dependent parameter x . The trick is to get everything depending on t on one side of the equals sign, and everything depending on x on the other. To do so, we’re going to treat dx/dt as if it were an actual fraction¹. In that case, it’s not hard to see that we can re-arrange Equation 16.A.15 to

$$\frac{1}{x_h} dx_h = -\frac{b(t)}{a(t)} dt \quad (16.A.16)$$

By itself, Equation 16.A.16 means little, but if we integrate both sides, we get something that makes sense:

$$\int \frac{1}{x_h} dx_h = \log(x_h) + C = -\int \frac{b(t)}{a(t)} dt \quad (16.A.17)$$

or

$$x_h(t) = A \exp \left[-\int \frac{b(t)}{a(t)} dt \right] \quad (16.A.18)$$

where $A = \exp(C)$ is an integration constant (the unknown constant that will be set by our initial condition). Of course, in principle it may not be possible to evaluate the integral in Equation 16.A.18, but even then the solution is valid. In practice, you’ll often encounter situations in which $a(t)$ and $b(t)$ are simple functions or even constants, and the evaluation of the integral is straightforward. Now that we have our homogeneous solution, we still need a particular one. Sometimes you’re lucky, and you can easily guess one - for instance one in which $x_p(t)$ doesn’t depend on t at all. In case you’re not lucky, there’s are two other techniques you may try, either using variation of constants or finding an integrating factor. To demonstrate variation of constants, we’ll pick a specific example, to not get lost in a bunch of abstract functions. Let $a(t) = a$ be a constant and $b(t) = bt$ be linear. The homogeneous solution then becomes $x_h(t) = A \exp \left[-\frac{1}{2} \frac{b}{a} t^2 \right]$. The constant we’re going to vary is our integration constant A , so our guess for the particular solution will be

$$x_p(t) = A(t) \exp \left[-\frac{1}{2} \frac{b}{a} t^2 \right] \quad (16.A.19)$$

We substitute 16.A.19 back into the full differential Equation 16.A.14, which gives:

$$\left[a \frac{dA}{dt} - aA(t) \frac{bt}{a} + bA(t) \right] \exp \left[-\frac{1}{2} \frac{b}{a} t^2 \right] = a \frac{dA}{dt} \exp \left[-\frac{1}{2} \frac{b}{a} t^2 \right] = f(t) \quad (16.A.20)$$

A big part of the left-hand side thus cancels, and that’s not a coincidence - that’s because it is based on the homogeneous equation. What remains is a differential equation in $A(t)$ that can be trivially solved by direct integration:

$$A(t) = \int \frac{dA}{dt} dt = \frac{1}{a} \int f(t) \exp \left[\frac{1}{2} \frac{b}{a} t^2 \right] dt \quad (16.A.21)$$

Again, it may not be possible to evaluate the integral in Equation 16.A.21, but in principle the solution could be inserted in Equation 16.A.19 to give us our particular solution, and the whole differential equation will be solved.

Alternatively, we may try to find an integration factor for Equation 16.A.14. This means that we try to rewrite the left hand side of the equation as a total derivative, after which we can simply integrate to get the solution. To do so, we first divide the whole equation by $a(t)$, then look for a function $\mu(t)$ that satisfies the condition that

$$\frac{d}{dt}[\mu(t)x(t)] = \mu(t)\frac{dx}{dt} + x(t)\frac{d\mu}{dt} \text{ mod } \mu(t)\frac{dx}{dt} + \mu(t)\frac{b(t)}{a(t)}x(t) \quad (16.A.22)$$

from which we can read off that we need to solve the homogeneous equation

$$\frac{d\mu}{dt} = \frac{b(t)}{a(t)}\mu(t) \quad (16.A.23)$$

We can solve 16.A.23 by separation of constants, which gives us

$$\mu(t) = \exp\left(\int \frac{b(t)}{a(t)} dt\right) \quad (16.A.24)$$

where we set the integration constant to one, as it drops out of the equation for $x(t)$ anyway. With this function $\mu(t)$, we can rewrite Equation 16.A.14 as

$$\frac{d}{dt}[\mu(t)x(t)] = \mu(t)\frac{f(t)}{a(t)} \quad (16.A.25)$$

which we can integrate to find $x(t)$:

$$x(t) = \frac{1}{\mu(t)} \int \mu(t)\frac{f(t)}{a(t)} dt \quad (16.A.26)$$

A.3.2. SECOND-ORDER LINEAR ORDINARY DIFFERENTIAL EQUATIONS WITH CONSTANT COEFFICIENTS

Second order ordinary differential equations are essential for the study of mechanics, as its central equation, Newton's second law of motion (Equation 2.1.4) is of this type. In the case that the equation is also linear, we have some hopes of solving it analytically. There are several examples of this type of equation in the main text, especially in Section 2.6, where we solve the equation of motion resulting from Newton's second law for three special cases, and Section 8.1, where we study a number of variants of the harmonic oscillator. For the case that the equation is homogeneous and has constant coefficients, we can write down the general solution². The equation to be solved is of the form

$$a\frac{d^2x}{dt^2} + b\frac{dx}{dt} + cx(t) = 0 \quad (16.A.27)$$

For the case that $a = 0$, we retrieve a first-order differential equation, whose solution is an exponential (as can be found by separation of variables and integration): $x(t) = C\exp(ct/b)$. In many cases an exponential is also a solution of Equation 16.A.27. To figure out which exponential, let's start with the trial function (or 'Ansatz') $x(t) = \exp(\lambda t)$, where λ is an unknown parameter. Substituting this Ansatz into Equation 16.A.27 yields the characteristic polynomial for this ode:

$$a\lambda^2 + b\lambda + c = 0 \quad (16.A.28)$$

which almost always has two solutions:

$$\lambda_{\pm} = -\frac{b}{2a} \pm \frac{\sqrt{b^2 - 4ac}}{2a} \quad (16.A.29)$$

Note that the solutions can be real or complex. If there are two of them, we can write the general solution³ of Equation 16.A.27 as a linear combination of the Ansatz with the two cases:

$$x(t) = Ae^{\lambda_+ t} + Be^{\lambda_- t} \quad (16.A.30)$$

where A and B are set by either initial or boundary conditions. Since the λ_{\pm} may be complex, so may A and B; it's their combination that should give a real number (as $x(t)$ is real), see problem A.3.1a.

In the case that Equation 16.A.29 gives only one solution, the corresponding exponential function is still a solution of Equation 16.A.27, but it is not the most general one, as we only can put a single undetermined constant in front of it. We therefore need a second, independent solution. To guess one, here's a third useful trick⁴: take the derivative of our known solution, $e^{\lambda t}$, with respect to the parameter λ . This gives a second Ansatz: $te^{\lambda t}$, where $\lambda = -b/2a$. Substituting this Ansatz into Equation 16.A.27 for the case that $c = b^2/4a$, we find:

$$\frac{d^2x}{dt^2} + b\frac{dx}{dt} + \frac{b^2}{2a}x(t) = a\left(-\frac{b}{a} + \frac{b^2}{4a^2}t\right)e^{-\frac{bt}{2a}} + b\left(1 - \frac{b}{2a}t\right)e^{-\frac{bt}{2a}} + \frac{b^2}{4a}te^{-\frac{bt}{2a}} = 0 \quad (16.A.31)$$

so our Ansatz is again a solution. For this special case, the general solution is therefore given by

$$x(t) = Ae^{-\frac{bt}{2a}} + Bte^{-\frac{bt}{2a}} \quad (16.A.32)$$

In Section 8.2, where we discuss the damped harmonic oscillator, the special case corresponds to the critically damped oscillator. We get an underdamped oscillator when the roots of the characteristic polynomial are complex, and an overdamped one when they are real.

A.3.3. SECOND-ORDER LINEAR ORDINARY DIFFERENTIAL EQUATIONS OF EULER TYPE

There's a second class of linear ordinary differential equations that we can solve explicitly: those of Euler (or Cauchy-Euler) type, where the coefficient in front of a derivative contains the variable to the power of the derivative, i.e., for a second-order differential equation, we have as the most general form:

$$ax^2\frac{d^2y}{dx^2} + bx\frac{dy}{dx} + cy(x) = 0 \quad (16.A.33)$$

Note that we are now solving for $y(x)$; we do so because this type of equation typically occurs in the context of position- rather than time-dependent functions. An example is the Laplace equation ($\nabla^2 y = 0$) in polar coordinates. Like for the second order ode with constant coefficients, the ode of Euler type can be generalized to higher-order equations.

There are (at least) two ways to solve Equation 16.A.33: through an Ansatz, and through a change of variables. For the Ansatz, note that for any polynomial, the derivative of each term reduces the power by one, and here we're multiplying each such term with the variable to the power the number of derivatives⁵. This suggests we simply try a polynomial, so our Ansatz here will be $y(x) = x^n$. Substituting in Equation 16.A.33 gives:

$$ax^2n(n-1)x^{n-2} + bxn x^{n-1} + cx^n = [an(n-1) + bn + c]x^n = 0 \quad (16.A.34)$$

so we get another second order polynomial to solve, this time in n :

$$an^2 + (b-a)n + c = 0 \Rightarrow n_{\pm} = \frac{1}{2} - \frac{b}{2a} \pm \frac{1}{2}\sqrt{(a-b)^2 - 4ac} \quad (16.A.35)$$

If the roots in Equation 16.A.35 are both real (the most common case in physics problems), we have two independent solutions, and we are done. If the roots are complex, we also have two independent solutions, though they involve complex powers of x ; like for the equation with constant coefficients, we can rewrite these as real functions with Euler's formula (see problem A.3.1b). For the case that we have only one root, we again apply our trick to get a second: we try $\frac{dx^n}{dn} = x^n \ln(x)$, which turns out to be indeed a solution (problem A.3.1c), and the general solution is again a linear combination of the two solutions found.

Alternatively, we could have solved Equation 16.A.33 by a change of variables. Although this method is occasionally useful (and so it's good to be aware of its existence), there is no systematic way of deriving which change of variables will do the trick, so you'll have to go by trial-and-error (without a priori guarantee of success). In this case, this process leads to the following substitution:

$$x = e^t, \quad y(x) = y(e^t) \equiv \phi(t) \quad (16.A.36)$$

where we introduce $\phi(t)$ for convenience. Taking derivatives of $y(x)$ with the chain rule gives

$$\frac{dy}{dx} = \frac{dy}{dt} \frac{dt}{dx} = \frac{1}{x} \frac{d\phi}{dt}, \quad \frac{d^2y}{dx^2} = \frac{1}{x^2} \left(\frac{d^2\phi}{dt^2} - \frac{d\phi}{dt} \right) \quad (16.A.37)$$

which is a second order differential equation with constant coefficients, and thus of the form given in Equation 16.A.27. We therefore know how to find its solutions, and can use Equation 16.A.36 to transform those solutions back to functions $y(x)$.

A.3.4. Reduction of Order

If you find yourself with a non-homogeneous second order differential equation where the homogeneous equation either has constant coefficients or is of Euler type, you can again use the technique of variation of constants to find a particular solution. A similar technique, known as reduction of order, may help you find solutions to a second (or higher) order equation where the

coefficients are not constant. In order to be able to use this technique, you need to know a solution to the homogeneous equation, so it is not as universally applicable as the techniques in the previous two sections, but still frequently very helpful.

Let us write the general non-homogeneous second-order linear differential equation as

$$\frac{d^2 y}{dx^2} + p(x) \frac{dy}{dx} + q(x)y(x) = r(x) \quad (16.A.38)$$

Note that this is the most general form: if there is a coefficient (constant or otherwise) in front of the second derivative, we simply divide the whole equation by that coefficient and redefine the coefficients to match Equation 16.A.38. Now suppose we have a solution $y_1(x)$ of the homogeneous equation (so for the case that $r(x) = 0$). As the equation is homogeneous, for any constant v the function $vy_1(x)$ will also be a solution. As an Ansatz for the second solution, we'll try a variant of variation of constants, and take

$$y_2(x) = v(x)y_1(x) \quad (16.A.39)$$

where $v(x)$ is an arbitrary function. Substituting 16.A.39 back into 16.A.38, we find

$$y_1(x) \frac{d^2 v}{dx^2} + \left[2 \frac{dy_1}{dx} + p(x)y_1(x) \right] \frac{dv}{dx} + \left[\frac{d^2 y_1}{dx^2} + p(x) \frac{dy_1}{dx} + q(x)y_1(x) \right] v(x) = r(x) \quad (16.A.40)$$

We recognize the prefactor of $v(x)$ as exactly the homogeneous equation, which $y_1(x)$ satisfies, so this term vanishes. Now defining $w(x) = \frac{dv}{dx}$, we are left with a first-order equation for $w(x)$:

$$y_1(x) \frac{dw}{dx} + \left[2 \frac{dy_1}{dx} + p(x)y_1(x) \right] w(x) = r(x) \quad (16.A.41)$$

Equation 16.A.41 is a first-order linear differential equation, and can be solved by the techniques from Section A.3.1. Integrating the equation $w(x) = \frac{dv}{dx}$ then gives us $v(x)$, and hence the second solution 16.A.39 of the (inhomogeneous) second order differential equation.

A.3.5. POWER SERIES Solutions

If none of the techniques in the sections above apply to your differential equation, there's one last Ansatz you can try: a power series expansion of your solution. To illustrate, we'll again pick a concrete example: Legendre's differential equation, given by

$$\frac{d}{dx} \left[(1-x^2) \frac{dy}{dx} \right] + n(n+1)y(x) = (1-x^2) \frac{d^2 y}{dx^2} - 2x \frac{dy}{dx} + n(n+1)y(x) = 0 \quad (16.A.42)$$

where n is an integer. As an Ansatz for the solution, we'll try a power series expansion of $y(x)$:

$$y(x) = \sum_{k=0}^{\infty} a_k x^k \quad (16.A.43)$$

Our task is now to find numbers a_k (many of which may be zero) such that 16.A.43 is a solution of 16.A.42. Fortunately, we can simply substitute our trial solution and re-arrange to get

$$0 = (1-x^2) \frac{d^2}{dx^2} \left(\sum_{k=0}^{\infty} a_k x^k \right) - 2x \frac{d}{dx} \left(\sum_{k=0}^{\infty} a_k x^k \right) + n(n+1) \sum_{k=1}^{\infty} a_k x^k \quad (16.A.44)$$

$$= (1-x^2) \left(\sum_{k=0}^{\infty} k(k-1) a_k x^{k-2} \right) - 2x \left(\sum_{k=0}^{\infty} k a_k x^{k-1} \right) + n(n+1) \sum_{k=1}^{\infty} a_k x^k \quad (16.A.45)$$

$$= \sum_{k=0}^{\infty} [(-k(k-1) - 2k + n(n+1)) a_k x^k + k(k-1) a_k x^{k-2}] \quad (16.A.46)$$

$$= \sum_{k=0}^{\infty} [(-k(k+1) + n(n+1)) a_k + (k+2)(k+1) a_{k+2}] x^k \quad (16.A.47)$$

where in the last line, we 'shifted' the index of the last term⁶. We do so in order to get at an expression for the coefficient of x^i for any value of k . As the functions x^k are linearly independent⁷ (i.e., you can't write x^k as a linear combination of other functions x^m

where $m \neq k$), the coefficient of each of the powers in the sum in Equation 16.A.47 has to vanish for the sum to be identically zero. This gives us a recurrence relation between the coefficients a_k :

$$a_{k+2} = \frac{k(k+1) - n(n+1)}{(k+2)(k+1)} a_k \quad (16.A.48)$$

Given the values of a_0 and a_1 (the two degrees of freedom that our second-order differential equation allows us), we can repeatedly apply Equation 16.A.48 to get all coefficients. Note that for $k = n$ the coefficient equals zero. Therefore, if for an even value of n , we set $a_1 = 0$, and for an odd value of n , we set $a_0 = 0$, we get a finite number of nonzero coefficients. The resulting solutions are polynomials, characterized by the number n ; in this case, they're known as the Legendre polynomials, typically denoted $P_n(x)$, and normalized (by setting the value of the remaining free coefficient) such that $P_n(1) = 1$. Table A.1 lists the first five, which are also plotted in Figure 16.A.4a.

Legendre polynomials have many other interesting properties (many of which can be found in either math textbooks or on their Wikipedia page). They occur frequently in physics, for example in solving problems involving Newtonian gravity or Laplace's equation from electrostatics.

If we replace the $n(n+1)$ factor in the Legendre differential equation with an arbitrary number λ , the series solution remains a solution, but it no longer terminates⁸. There are many other differential equations that lead to both infinite series and polynomial solutions. A well-known example is the Bessel differential equation:

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2) y(x) = 0 \quad (16.A.49)$$

The solutions to this equation are known as the Bessel functions of the first and second kind (see Problem A.3.3, where you'll prove that for these functions the series never terminates). These functions generalize the sine and cosine function and occur in the vibrations of two-dimensional surfaces. Other examples include the Hermite and Laguerre polynomials, which feature in quantum mechanics, and Airy functions, which you can encounter when studying optics.

Table A.1: : The first five (and zeroth, for good measure) Legendre polynomials, the solutions of Equation 16.A.42 for the given value of n and the appropriate choice of a_0 and a_1 .

n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x - 1)$
3	$\frac{1}{2}(5x^3 - 3x)$
4	$\frac{1}{8}(35x^4 - 30x^2 + 3)$
5	$\frac{1}{8}(63x^5 - 70x^3 + 15x)$

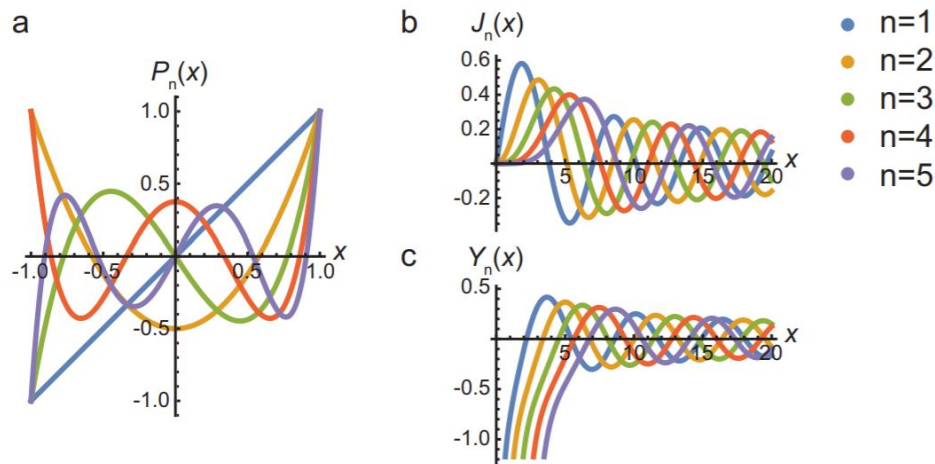


Figure 16.A. 4: Solutions to the Legendre and Bessel differential equations. (a) The first five Legendre polynomials (Table A.1). Note that the polynomials with even n are all even, and those with odd n are all odd. (b-c) The first five Bessel functions of the first (b) and second (c) kind.

A.3.6. Problems

A.3.1

- Suppose we have a solution of Equation 16.A.27 where the roots λ_{\pm} of the characteristic polynomial (Equation 16.A.29) are complex, so $\lambda_{\pm} = \alpha \pm i\beta$. Rewrite the general solution 16.A.30 in real functions with real coefficients C and D , and express C and D in terms of A and B . Hint: use Euler's formula $e^{ix} = \cos(x) + i \sin(x)$.
- Suppose we have a solution of Equation 16.A.35 where the roots n_{\pm} are complex, so $n_{\pm} = \alpha \pm i\beta$. To get a solution of Equation 16.A.33 without complex numbers, we make the substitution $x = e^t$, so

$$x^{n_{\pm}} = x^{\alpha \pm i\beta} = e^{\alpha t} e^{\pm i\beta t} \quad (16.A.50)$$

Use Euler's formula again to rewrite the complex exponential in terms of sines and cosines, and make the back-substitution to x to show that the general solution of Equation 16.A.33 in this case is given by

$$y(x) = x^{\alpha} [A \cos(\beta \ln(x)) + \sin(\beta \ln(x))] \quad (16.A.51)$$

- Suppose we have a solution of Equation 16.A.35 for which there is only a single root n . Show that the derivative of x^n with respect to n is in this case also a solution of Equation 16.A.33 and that the general solution is given by

$$y(x) = x^n [A + B \ln(x)] \quad (16.A.52)$$

A.3.2 Use the method of reduction of order to obtain a second solution of Equation 16.A.27 for the case that the characteristic polynomial (Equation 16.A.29) has only a single root.

A.3.3

- Use the power series technique to find a solution to Bessel's differential Equation 16.A.49. Why doesn't the series terminate in this case? Why do you only get one family of solutions? We'll call these solutions 'Bessel functions of the first kind' and label them as $J_n(x)$ (see Figure 16.A.4b).
- Use the method of reduction of order to find a second family of solutions to the Bessel differential equation, known as 'Bessel functions of the second kind' ($Y_n(x)$, see Figure 16.A.4c).

This page titled 16.A: Math is shared under a CC BY-NC-SA 4.0 license and was authored, remixed, and/or curated by Timon Idema (TU Delft Open) via source content that was edited to the style and standards of the LibreTexts platform.