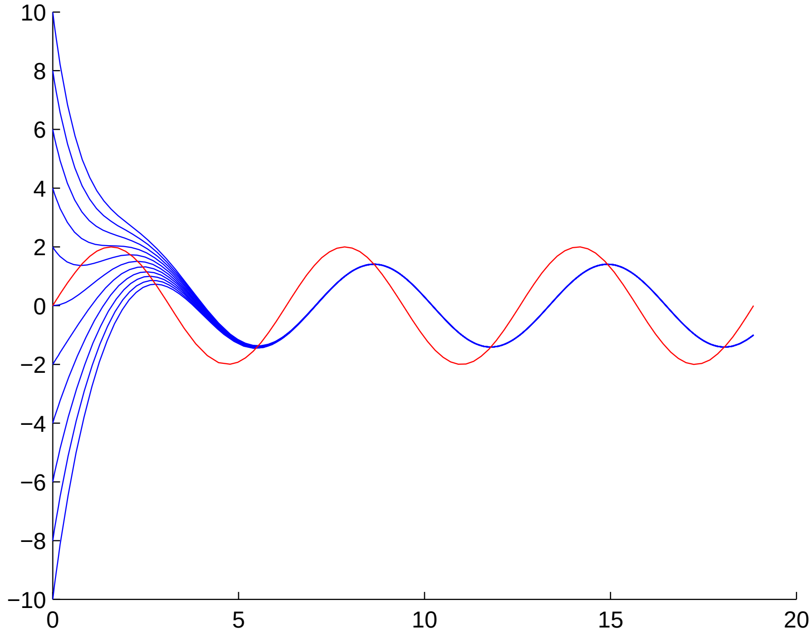

FIGURE 11. The blue curves are numerical solutions of $dx/dt = 2\,\sin(t) - x$ with initial conditions ranging from -10 to 10; the red curve is the plot of $f = 2\sin(t)$. What observations can you make about the solution of the ODE?

6. *RC*-circuit driven by a sine wave

Figure 11 shows several solutions of

$$\frac{dx}{dt} = 2\,\sin(t) - x, \quad x(0) = x_0, \tag{26}$$

with initial conditions ranging from -10 to 10. Regardless of the initial condition, the solution of Equation (26) approaches the same asymptote. Furthermore, that asymptote appears to be a shifted and scaled copy of $2\,\sin(t)$ (the latter is shown in red). This suggests that the solution of (26) is a sum of two terms: a simple harmonic of unit frequency, which is independent of the initial condition, and a rapidly decaying function which depends on the initial condition. Symbolically,

$$x = a\,\cos(t) + b\,\sin(t) + y(t). \tag{27}$$

We will now test the hypothesis (27) numerically.

6.1. **Fitting a harmonic to data.** First, we generate the data using `ode45` as follows:

```
rhs = @(t,x) 2*sin(t) - x;
[t,x] = ode45(rhs,[0 6*pi],10);
```

The choice of the initial condition $x_0 = 10$ was somewhat (but not completely) random: we wanted to ensure that the data has a sizable non-harmonic component and therefore picked a value far from -2.

It is visually evident that x becomes indistinguishable from a simple harmonic when $t > 5$. We therefore chop off the beginning portion of the data:

```
ind = t > 5;
tt = t(ind);
xx = x(ind);
```

The truncated data is shown in Figure 12:

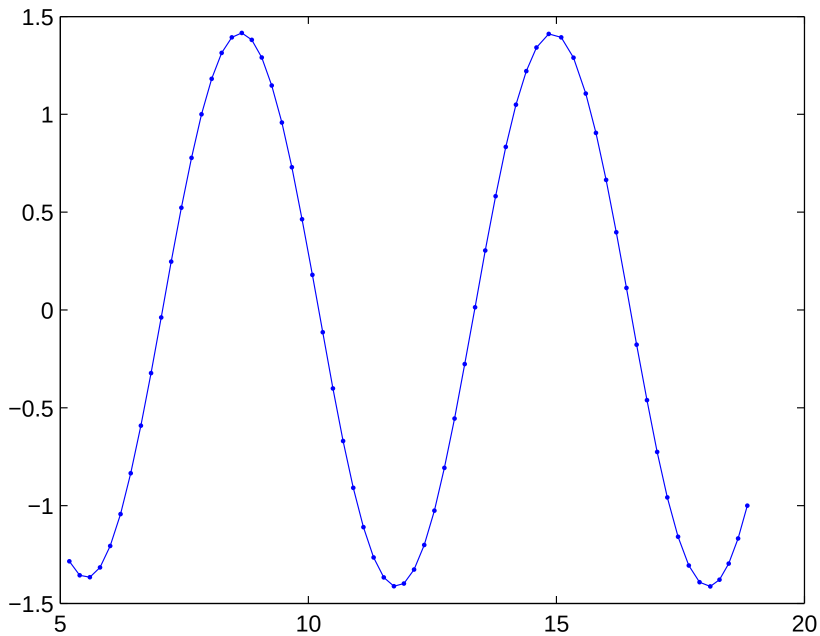

FIGURE 12. The plot of the solution for $t > 5$

.

This is clearly a simple harmonic: $a\ \cos(t)+b\ \sin(t)$. We now proceed to find the coefficients a and b.

Each of the 67 data points shown in Figure 12 corresponds to a linear equation in a and b. Let (t_k, x_k) be the k-th data point. Since data points lie on the harmonic:

$$a\,\cos(t_k) + b\,\sin(t_k) = x_k, \quad k = 1, \ldots, 67.$$

We thus need to solve a system of 67 equations in two unknowns.

In order to solve a linear system of algebraic equations in MATLAB, we need to express it in matrix-vector form. To that end, let us think of t_k and x_k as components of column vectors $\mathbf{t}$ and $\mathbf{x}$, respectively. Incidentally, this is what `ode45` outputs in the first place. Also, let us think of a and b as components of an unknown column vector:

$$\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}.$$

Finally, let A be the matrix with columns $\cos(\mathbf{t})$ and $\sin(\mathbf{t})$: that is, the k-th row of A is $[\cos(t_k)$ and $\sin(t_k)]$. Then

$$A\,\mathbf{v} = \mathbf{x},$$

as follows from the "row-by-column" rule of matrix multiplication. We can now find $\mathbf{v}$ using the linear solver `linsolve`:

```
A = [cos(tt) sin(tt)];
v = linsolve(A,xx)

v =

  -0.9940
   0.9977
```

Note that in the above snippet we work with variables `tt` and `xx` which correspond to truncated data. Evidently, the harmonic we are looking for is very close to: $-\cos(t) + \sin(t)$.

6.2. **Finding the transient.** We now turn our attention to the 'transient' portion of the solution which we denoted $y(t)$. In order to compute it, we simply subtract the harmonic portion of the solution; however, this time we work with the data for $t < 5$:

```
ind = t < 5;
tt = t(ind);
xx = x(ind);
A = [cos(tt) sin(tt)];
yy = xx - A*v;
```

The plot of yy itself is not very revealing: it is just a monotonely decaying function. Since we are solving a model of an RC-circuit, it is sensible to assume that $y(t)$ is an exponential. This is confirmed by the semilogarithmic plot below which was produced by the following code:

```
zz = log(yy);
p = polyfit(tt,zz,1)

p =

 -1.0022    2.4026

figure
plot(tt,zz,'ro')
hold on
plot(tt,polyval(p,tt),'b-')
```

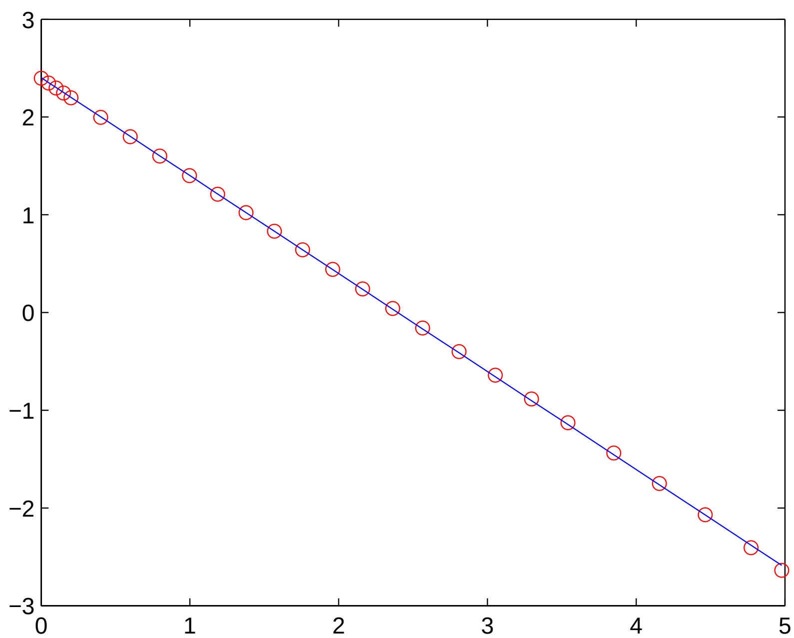

Figure 13. Semi-logarithmic plot of the transient for $t < 5$

.

The linear fit shown in Figure 13 leaves no doubt that $y(t)$ is an exponential: $y = C\,e^{-k\,t}$, with $k \approx 1$ and $C = e^{2.4026} \approx 11$.

6.3. **Conclusions.** We can now reasonably suspect that the solution of (26) with $x_0 = 10$ is given by:

$$x(t) = -\cos(t) + \sin(t) + 11\,e^{-t}.$$

As a simple exercise, verify that this claim is correct. More generally, we can say that the solution of (26) has the general form:

$$x(t) = -\cos(t) + \sin(t) + C\,e^{-t},$$

where C depends on the initial value. In fact, convince yourself that $C = x_0 + 1$.

Exercises.

(1) Use the handout as a template for investigating the following IVP:

$$\frac{dx}{dt} + 2\,x = \cos(3\,t), \quad x(0) = 10.$$

That is:

(a) Plot several solutions of the IVP with different initial conditions as in Figure 11. Use the resulting figure to formulate a hypothesis about the solution.

(b) Investigate the “steady” part of the solution—the asymptote to which all solutions tend as $t \to \infty$. As the handout explains, you will need to truncate the data appropriately, so that it looks like Figure 12; then use `linsolve` to find unknown coefficients.

(c) Investigate the “transient”. Again, truncate the data appropriately and use a semi-logarithmic plot like Figure 13 confirm your guess.

(d) Validate your numerical results symbolically.

(2) Repeat the previous exercise with the IVP:

$$\frac{dx}{dt} + 3\,x = t, \quad x(0) = 10.$$

7. The joy of linearity

Scientific disciplines are often organized around some classification of their objects of study. For instance, chemistry is broadly divided into organic and inorganic; biology is organized according to the tree of life; physics is split into mechanics, thermodynamics, electromagnetism, astronomy, and so on. Most scientific classifications are artificial and, to some extent, crude. Yet they are useful. If nothing else, it is pedagogically helpful to divide objects into classes because that makes learning more focused and efficient. For instance, it is often the case that understanding one member of a particular class leads to broad understanding of the class as a whole.

The study of ordinary differential equations is organized around a classification which, at the top level, has only two classes: linear ODE and nonlinear ODE; these classes are, obviously, disjoint. We hasten to add that linearity is not specific to ordinary differential equations. Partial differential equations, integral equations, and algebraic equations are also divided into linear and nonlinear. What follows is a general discussion of linearity with a strong emphasis on universality of the concept. This will also serve as an introduction to the fascinating subject of Linear Algebra (Math 145) which is at the core of Math 57.

7.1. **Operational form of an ODE.** Most equations in general and all ordinary differential equations in particular can be written in the form $L(x) = f$, where L is an operation applied to the unknown x with result f. For instance, consider the general RC-circuit equation:

$$\frac{dx}{dt} + k\,x = f. \tag{28}$$

To cast it in the form $L(x) = f$, simply set $L(x) = \frac{dx}{dt} + k\,x$. We will often use the *mapping*, or *operator* notation $L : x \mapsto \frac{dx}{dt} + k\,x$, pronounced "$x$ maps to $\frac{dx}{dt} + k\,x$", to denote operations on functions. Also, we will use the terms "operation", "operator", and "mapping" interchangeably.

Thus at the heart of an ordinary differential equation is a calculus operation L. Consequently, the classification of ordinary differential equations is really the classification of *operations* which define them.

> An (ordinary differential) equation $L(x) = f$ is classified as linear or nonlinear based on whether the defining operation L is linear or nonlinear.

In order to explain what makes an operation linear, we first need to spell out what we allow as inputs and outputs. Linearity makes sense only when operations are applied to *vectors* and produce *vectors.* We therefore need to define vectors in a way that makes sense in the context of ODE.

7.2. **Vectors.** A common misconception is that vectors are directed line segments or, more abstractlt, "things with direction and magnitude." It is certainly true that vectors can be sometimes *depicted* as directed line segments. It is also true that vectors can be endowed with *attributes* such as length and direction. However, vectors are defined neither by their visual representation nor by their attributes. The correct way to think about vectors is in terms of the following (meta) definition.

Definition 1 (Vectors)**.** *Vectors are objects which can be added and scaled.*

We hasten to add that there are certain restrictions on what can be construed as vector addition and scalar multiplication. For instance, vector addition must be commutative: $x+y = y+x$; the associative law must hold: $x+(y+z) = (x+y)+z$; there must exist a zero vector, and so on. We did not spell out these axioms in Definition 1 for two reasons. Firstly, it is better done in linear algebra (Math 145). Secondly, after Calculus 3 you already take all of these axioms for granted anyway. Henceforth, simply think of vectors as objects which can be added and scaled *in the manner of familiar vectors from calculus and physics.*

Observe that Definition 1 is purely algebraic in nature. It says nothing about length, direction, components, and so on; it certainly does not reference line segments. Vectors are purely algebraic objects. As all algebraic objects they are defined *operationally* and the only default vector operations are addition and scalar multiplication. The dot product and the cross product familiar from Calculus 3 are special operations which need to be introduced separately—they do not exist by default.

After months or perhaps even years of thinking about vectors in calculus and physics terms, it may be difficult to switch to the abstract algebraic definition. In fact, this may require some Zen-like practice on your part. Remember not to fixate on how vectors (or other mathematical objects) look or what attributes they may or may not have. As the first Zen exercise, convince yourself that *numbers are vectors.* Next notice that functions are also vectors in the sense of our definition since they can be added and scaled the same way as "ordinary" vectors.

Another general observation which should be made here is that it is impossible to define a single vector; only collections of vectors make sense. Indeed, one needs at least two vectors to be able to speak of vector addition. Moreover, it turns out that not every collection of vectors is mathematically useful: one needs certain closure properties as we will now explain.

7.3. **Vector spaces.** We need some additional definitions before we can define a key concept of a vector space. If we have a collection of vectors, Definition 1 allows us to scale each vector by some number (scalar) and add the resulting scaled vectors. This general operation is so ubiquitous that it deserves a special definition.

Definition 2 (Linear Combinations)**.** *Let $x_1, \ldots, x_N$ be a collection of N vectors and let $c_1, \ldots, c_N$ be a collection of N scalars (numbers). The vector*

$$c_1\, x_1 + \ldots + c_N\, x_N = \sum_{n=1}^{N} c_n\, x_n$$

is called a linear combination of vectors $x_1, \ldots, x_N$ with coefficients $c_1, \ldots, c_N$.

As a quick example, recall from Calculus 3 that every vector in $\mathbb{R}^3$ is a linear combination of unit vectors $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$: $\mathbf{x} = a\,\mathbf{i} + b\,\mathbf{j} + c\,\mathbf{k}$; here we typeset vectors in boldface only to aid recognition.

As another example, consider polynomials of degree two. These are linear combinations of monomials. Indeed, every quadratic can be written in the form: $q = a\,1 + b\,x + c\,x^2$. Note how similar the monomials $1, x, x^2$ seem to be to vectors $\mathbf{i}, \mathbf{j}, \mathbf{k}$. As we will explain later, from the point of view of linear algebra there is, in fact, no distinction between quadratics and three dimensional vectors in $\mathbb{R}^3$!

Now let us get back to the question of which collections of vectors are mathematically useful. According to Definition 2, any vectors can be linearly combined into new vectors. However, if we just have a collection of only two vectors, say, their linear combination may not be present in the collection. And this may be a problem. This suggests that we should require our collections of vectors to be *closed* under linear combinations. And that simple and natural requirement brings us to the extremely important idea of vector spaces.

Definition 3 (Vector Space)**.** *A vector space is any collection of vectors that is closed under linear combinations. That is, a linear combination of any vectors in a vector space must necessarily belong to the vector space.*

As an example, the collection of *all* vectors in $\mathbb{R}^3$ is a vector space called, unsurprisingly, $\mathbb{R}^3$. As a counterexample, consider all *nonzero* vectors in $\mathbb{R}^3$. This is not a vector space because the *trivial* linear combination $0\,\mathbf{i} + 0\,\mathbf{j} + 0\,\mathbf{k} = \mathbf{0}$ is not contained in the collection. Incidentally, this shows that every vector space necessarily contains the zero vector.

7.4. **Linear operations.** As we have already remarked, linearity makes sense only when operations are applied to vectors and produce vectors. More precisely, both the domain and the range of a linear operation must be some vector spaces (see Definition 3).

Suppose that L is an operation which maps vectors in one space into vectors in another space; henceforth, 'space' is short for *vector* space. The basic operation in a vector space is that of taking a linear combination (see Definition 2) and linearity has to do with the special interaction between L and linear combinations.

Definition 4. *Let $L : X \mapsto Y$ be an operation taking vectors from vector space X into vectors in vector space Y. L is linear if*

$$L(c_1\,x_1 + \ldots + c_N\,x_N) = c_1\,L(x_1) + \ldots + c_N\,L(x_N).$$

In words, L can be distributed over linear combinations.

Another way to remember Definition 4 is as the following sentence: "linear operations take linear combinations of inputs into linear combinations of outputs." Nonlinearity is, obviously, the breakdown of linearity.

You can now see why we could not explain linearity right away. We needed the idea of a linear combination which only applies to vectors. Furthermore, for linear combinations to be always well-defined, the vectors must form a vector space. This is not an apology, by the way. We are going through a highly abstract and very difficult stretch of the course and there is no way around it: the power of mathematics lies in abstraction. Make a note to reread this handout and then re-reread it several times, if necessary. In the meantime, let us consider a few examples of linear operations related to ODE.

Let us get back to the forced RC-circuit modeled by Equation (28). We write the latter in operational form as $L(x) = f$ with $L : x \mapsto \frac{dx}{dt} + k\,x$. Let us now show that this equation is linear. To this end, we need to show that the *operation* L is linear in the sense of Definition 4.

First, note that L acts on functions and produces functions. Now functions are vectors according to Definition 1. If we consider, say, all infinitely differentiable functions, the resulting collection is a vector space in the sense of Definition 3. We thus have an operation whose domain and range is that vector space (called C^∞), so the question about linearity is sensible.

Next let x_1 and x_2 be two arbitrary functions of time and let c_1 and c_2 be some numbers. Apply L to the linear combination $c_1\,x_1 + c_2\,x_2$:

$$L\,(c_1\,x_1 + c_2\,x_2) = \frac{d}{dt}\,(c_1\,x_1 + c_2\,x_2) + k\,(c_1\,x_1 + c_2\,x_2)\,.$$

Since differentiation and multiplication by a constant can be distributed over linear combinations, we get

$$c_1\,\frac{dx_1}{dt} + c_2\,\frac{dx_2}{dt} + k\,c_1\,x_1 + k\,c_2\,x_2.$$

Now collect the terms containing c_1 and c_2. This leads to:

$$c_1\left(\frac{dx_1}{dt} + k\,x_1\right) + c_2\left(\frac{dx_2}{dt} + k\,x_2\right) = c_1\,L(x_1) + c_2\,L(x_2).$$

Thus $L(c_1\,x_1 + c_2\,x_2) = c_1\,L(x_1) + c_2\,L(x_2)$ and it should be evident that

$$L\left(\sum_{n=1}^{N} c_n\,x_n\right) = \sum_{n=1}^{N} c_n\,L(x_n),$$

the proof is exactly the same. Therefore L satisfies Definition 4 and is linear.

Note that in order to prove linearity it suffices to test an operation on a linear combination of just *two* arbitrary inputs. Note also that the above L is linear largely as a consequence of linearity of differentiation. Can you explain why $\frac{d}{dt}$ is a linear operation? If you cannot, it is time to reread and rethink you favorite Calculus I text.

As another example, consider the following second order ODE:

$$m\,\frac{d^2x}{dt^2} = -k\,(x - l) - r\,\frac{dx}{dt}. \tag{29}$$

In the next section we will show that Equation (29) is a model of a simple mass-spring system based on Newton's second law. In the meantime, let us demonstrate linearity. First, rewrite (29) as $L(x) = k\,l$ with

$$L : x \mapsto m\,\frac{d^2x}{dt^2} + r\,\frac{dx}{dt} + k\,x.$$

Next apply L to a linear combination $c_1\,x_1 + c_2\,x_2$. Following the same steps as in the previous example, convince yourself that $L(c_1\,x_1 + c_2\,x_2) = c_1\,L(x_1) + c_2\,L(x_2)$ and, more generally,

$$L\left(\sum_{n=1}^{N} c_n\,x_n\right) = \sum_{n=1}^{N} c_n\,L(x_n),$$

The ODE (29) is therefore linear.

We conclude this section by pointing out that some care needs to be exercised when you cast an ODE into operational form $L(x) = f$. For instance, we could have rewritten Equation (29) as $L(x) = -k$ with

$$L : x \mapsto \frac{m\,\frac{d^2x}{dt^2} + r\,\frac{dx}{dt}}{(x - l)}$$

This L is actually nonlinear! Yet the ODE is linear because it *can be* written as $L(x) = f$ with a *linear* L. The moral is that a linear ODE can be algebraically rearranged so as to appear nonlinear. Therefore when casting ODE into operational form, avoid nonlinear operations, such as division, if possible. The converse, by the way, is not true: a nonlinear ODE cannot be algebraically rearranged so that it becomes linear. To wit, linearity can be destroyed but not created.

7.5. **Fundamental Theorem of Math 57.** Believe it or not, we have reached the joyous part! For we will now show that linearity can be actively used to construct solutions of linear ODE.

First, we need to introduce some additional terminology. Let $L(x) = f$ be a linear equation. If $f \neq 0$ we will say that the equation is non-*homogeneous* (note the emphasis in pronunciation: homogéneous). In the very special case of $L(x) = 0$ we will say that the equation is homogeneous. Homogeneity only makes sense when the equation is linear and has to do with the right-hand side. We note that there is a much better way to define homogeneity and we will return to it later. Yet "homogeneity means zero right-hand side" will be sufficient for now. On to the joy of linearity!

Theorem 1 (Fundamental Theorem of Math 57)**.** *The general solution of a linear nonhomogeneous equation $L(x) = f$ is the sum*

$$x = x_1 + x_2$$

where x_1 is the general solution of the homogeneous equation $L(x) = 0$ and x_2 is a particular solution of the nonhomogeneous equation $L(x) = f$.

Here is how this works. Consider Equation (28), which we know to be linear, in which we set $f(t) = t$ for definitiveness:

$$\frac{dx}{dt} + k\,x = t.$$

Begin with the homogeneous equation:

$$\frac{dx}{dt} + k\,x = 0.$$

The general solution, $x_1 = C\,e^{-k\,t}$, can be easily obtained using separation of variables.

Next we now need to find *any* solution of the nonhomogeneous equation. Let us "guess" it to be a first degree polynomial: $x_2 = a\,t + b$. The guess, of course, becomes obvious once we look at the plot of the solution (exercise). For the guess to work, we must have

$$L(x_2) = a + k\,a\,t + k\,b = t$$

which implies $a = 1/k$ and $b - 1/k^2$. Therefore

$$x_2 = \frac{1}{k}\,t - \frac{1}{k^2}.$$

Now the final step is, simply, to invoke Theorem 1 and add x_1 with x_2:

$$x = C\,e^{-k\,t} + \frac{1}{k}\,t - \frac{1}{k^2}.$$

As another example, consider

$$\frac{d^2x}{dt^2} = t.$$

Here $L : x \mapsto \frac{d^2x}{dt^2}$ which is conspicuously linear. The general solution of the homogeneous equation

$$\frac{d^2x}{dt^2} = 0,$$

as you may remember from Calculus, is $x_1 = C_1 + C_2\,t$. Meanwhile, a particular solution of the nonhomogeneous equation can be found by integrating the right-hand side twice:

$$x_2 = \frac{t^3}{6}.$$

Therefore the general solution of the nonhomogeneous equation is

$$x = x_1 + x_2 = C_1 + C_2\,t + \frac{t^3}{6}.$$

Nice and simple!

In order to fully appreciate the joy of Theorem 1, let us conclude with a counterexample. The ODE

$$\frac{dx}{dt} + x^2 = 1 + t^2$$

has particular solution $x_2 = t$. The general solution of the "homogeneous" equation

$$\frac{dx}{dt} + x^2 = 0$$

can be easily found using separation of variables:

$$x_1 = \frac{1}{t + C}.$$

Unfortunately, the ODE is nonlinear and its general solution has nothing to do with the sum of x_1 and x_2:

$$x \neq \frac{1}{t + C} + t.$$

We thus are in a position where we have some partial information about the solution but there is no way to combine the parts due to nonlinearity.

Exercises.

(1) Use linearity to construct the solution of the following IVP:

$$\frac{dx}{dt} + x = \sum_{n=1}^{10} \frac{\sin((2\,n-1)\,t)}{2\,n-1}, \quad x(0) = 10.$$

This is a model of an RC-circuit driven by a square wave. Once you find a formula for $x(t)$, validate it on the interval $[0, 20]$ using `ode45`. Include both the code and the plot in your solution.

(2) Find the general solution of

$$m\,\frac{d^2x}{dt^2} + r\,\frac{dx}{dt} = A\,\sin(\omega\,t)$$

Here m, r, A, and ω are positive constants. Validate your solution symbolically.

(3) Repeat the previous exercise with

$$m\,\frac{d^2x}{dt^2} + k\,x = A\,\sin(\omega\,t).$$

Here k is another positive constant. Validate your solution symbolically. Also, discuss the difficulty of numerical validation.