February 5

1. Solutions to differential equations

1.1. Introduction. A differential equation (DE) is an equation relating an unknown function and some of its derivatives. DEs arise in engineering, chemistry, biology, physics, economics, etc., because many laws of nature describe the instantaneous rate of change of a quantity in terms of current conditions.

Overall goals: to learn how to

- model real-world problems with DEs.
- solve DEs exactly when possible, or else solve numerically (get an approximate solution)
- extract qualitative information from a DE, whether or not it can be solved.

1.2. A secret function.

Example: Can you guess my secret function $y(t)$? Clue: It satisfies the DE

$$\dot{y} = 3y.$$  

(This might model population growth in some biological system.) (Note: $\dot{y}$ is physicist’s notation for derivative with respect to time; $\dot{y}$, $y'$, and $\frac{dy}{dt}$ all mean the same thing here.)
Maybe you guessed $y = e^{3t}$. This is a solution to the differential equation (1), because substituting it into the DE gives $3e^{3t} = 3e^{3t}$. But it’s not the function I was thinking of! Some other solutions are $y = 7e^{3t}$, $y = -5e^{3t}$, $y = 0$, etc. Later we’ll see that the general solution to (1) is

$$y = ce^{3t},$$

where $c$ is a parameter;

saying this means that

- for each number $c$, the function $y = ce^{3t}$ is a solution, and
- there are no other solutions besides these.

So there is a 1-parameter family of solutions to (1).

You still haven’t guessed my secret function.

Clue 2: My function satisfies the initial condition $y(0) = 5$.

Solution: There is a number $c$ such that $y(t) = ce^{3t}$ holds for all $t$; we need to find $c$. Plugging in $t = 0$ shows that $5 = ce^0$, so $c = 5$. Thus, among the infinitely many solutions to the DE, the particular solution satisfying the initial condition is $y(t) = 5e^{3t}$. □

**Important:** Checking a solution to a DE is usually easier than finding the solution in the first place, so it is often worth doing. Just plug in the function to both sides, and also check that it satisfies the initial condition.

In (1) only one initial condition was needed, since only one parameter $c$ needed to be recovered.

1.3. **Classification of differential equations.** There are two kinds:

- **Ordinary differential equation (ODE):** involves derivatives of a function of *only one* variable.
- **Partial differential equation (PDE):** involves *partial derivatives* of a *multivariable* function.

Notation for higher derivatives of a function $y(t)$:

- first derivative: $\dot{y}$, $y'$, $\frac{dy}{dt}$
- second derivative: $\ddot{y}$, $y''$, $\frac{d^2y}{dt^2}$
- third derivative: $\dddot{y}$, $y^{(3)}$, $\frac{d^3y}{dt^3}$
  
  $\vdots$
- $n^{th}$ derivative: $y^{(n)}$

**Order** of a DE: the highest $n$ such that the $n^{th}$ derivative of the function appears.
Example 1.1. Is

\[ 707099375 \cos(t^5) \dot{y}^4 + 3487980982(y + t^3)^7 \dddot{y} - 389750387y^{(3)}y^{(4)} + 2ty^{(5)} + 8453723054723985730987 \\
= 80970874y^6 - 2809754087 \sin(t/y) + 8957092 \ln(1 - t^7) \\
+ 64893745723786e^{y^8 - t^3} + 987t^6 + 543y^2 + 18.03 \]

an ODE or a PDE? ODE.

Flashcard question: What is its order?
The order is 5, because the highest derivative that appears is the 5th derivative, \( y^{(5)} \).

2. Modeling

There are two kinds of modeling. I’m not going to talk about the kind in which I dress up in fancy clothes and get photographed. The other kind, mathematical modeling, is converting a real-world problem into mathematical equations.

Guidelines:
1. Identify relevant quantities, both known and unknown, and give them symbols. Find the units for each.
2. Identify the independent variable(s). The other quantities will be functions of them, or constants. Often time is the only independent variable.
3. Write down equations expressing how the functions change in response to small changes in the independent variable(s). Also write down any “laws of nature” relating the variables. As a check, make sure that each summand in an equation has the same units.

Often simplifying assumptions need to be made; the challenge is to simplify the equations so that they can be solved but so that they still describe the real-world system well.

2.1. Example: savings account.

Problem 2.1. I have a savings account earning interest compounded daily, and I make frequent deposits or withdrawals into the account. Find an ODE with initial condition to model the balance.

Simplifying assumptions: Daily compounding is almost the same as continuous compounding, so let’s assume that interest is paid continuously instead of at the end of each day. Similarly, let’s assume that my deposits/withdrawals are frequent enough that they can approximated by a continuous money flow at a certain rate, the savings rate (which is negative when I am withdrawing). Finally, let’s assume that the interest rate and savings rate vary continuously with time, but do not depend on the balance.
Variables and functions (with units): Define the following:

- $P$: the principal, the initial amount that the account starts with (dollars)
- $t$: time from the start (years)
- $x$: balance (dollars)
- $I$: the interest rate (year$^{-1}$) (e.g., 4%/year = 0.04 year$^{-1}$)
- $q$: the savings rate (dollars/year).

Here $t$ is the independent variable, $P$ is a constant, and $x, I, q$ are functions of $t$.

Equations: During a time interval $[t, t + dt]$ for an “infinitesimally small” increment $dt$, the following hold (technically speaking, $dt$ is a differential; if $dt$ were replaced by a positive number $\Delta t$, then the equations below would be only approximations, but when we divide by $\Delta t$ and take a limit, the end result is the same):

- interest earned per dollar = $I(t) \, dt$
- interest earned = $I(t) x(t) \, dt$ (asked as flashcard question)
- amount deposited into the account = $q(t) \, dt$

so

$$dx = \text{change in balance} = I(t)x(t) \, dt + q(t) \, dt$$

$$\frac{dx}{dt} = I(t)x(t) + q(t).$$

(Check: the units in each of the three terms are dollars/year.) Also, there is the initial condition $x(0) = P$. Thus we have an ODE with initial condition:

$$\dot{x} = I(t)x + q(t), \quad x(0) = P. \quad (2)$$

Now that the modeling is done, the next step might be to solve (2), but we won’t do that yet.

2.2. Systems and signals. Maybe for financial planning I am interested in testing different saving strategies (different functions $q$) to see what balances $x$ they result in. To help with this, rewrite the ODE as

$$\dot{x} - I(t)x = q(t).$$

In the “systems and signals” language of engineering, $q$ is called the input signal, the bank is the system, and $x$ is the output signal. These terms do not have a mathematical meaning dictated by the DE alone; their interpretation is guided by what is being modeled. But the general picture is this:
• The input signal is a function of the independent variable alone, a function that enters into the DE somehow (usually the right side of the DE, or part of the right side).
• The system processes the input signal by solving the DE with the given initial condition.
• The output signal (also called system response) is the solution to the DE.

3. Separation of variables for first-order ODEs

(To be done in recitation.)

Separation of variables is a technique that quickly solves some simple first-order ODEs. Here is how it works:

1. Check that the DE is a first-order ODE. (If not, give up and try another method.) Suppose that the function to be solved for is \( y = y(t) \).
2. Rewrite \( \dot{y} \) as \( \frac{dy}{dt} \).
3. Add and/or subtract to move terms to the other side of the DE so that the term with \( \frac{dy}{dt} \) is on the left and all other terms are on the right.
4. Try to separate the \( y \)'s and \( t \)'s. Specifically, try to multiply and/or divide (and in particular move the \( dt \) to the right side) so that it ends up as an equality of differentials of the form
   \[
   f(y) \, dy = g(t) \, dt.
   \]

   Note: If there are factors involving both variables, such as \( y + t \), then it is impossible to separate variables; in this case, give up and try a different method.

   Warning: Dividing the equation by an expression invalidates the calculation if that expression is 0, so at the end, check what happens if the expression is 0; this may add to the list of solutions.
5. Integrate both sides to get an equation of the form
   \[
   F(y) = G(t) + C.
   \]

   These are implicit equations for the solutions, in terms of a parameter \( C \).
6. If possible (and if desired), solve for \( y \) in terms of \( t \).
Problem 3.1. Solve $\dot{y} - 2ty = 0$.

Solution:

Step 1. This involves only the first derivative of a one-variable function $y(t)$, so it is a first-order ODE. Thus we can attempt separation of variables.

Step 2. Rewrite as $\frac{dy}{dt} - 2ty = 0$.

Step 3. Isolate the $\frac{dy}{dt}$ term: $\frac{dy}{dt} = 2ty$.

Step 4. We can separate variables! Namely, $\frac{1}{y} \, dy = 2t \, dt$. (Warning: We divided by $y$, so at some point we will have to check $y = 0$ as a potential solution.)

Step 5. Integrate: $\ln |y| = t^2 + C$.

Step 6. Solve for $y$:

$$|y| = e^{t^2+C}$$
$$y = \pm e^C e^{t^2}.$$ 

As $C$ runs over all real numbers, and as the $\pm$ sign varies, the coefficient $\pm e^C$ runs over all nonzero real numbers. Thus these solutions are $y = ce^{t^2}$ for all nonzero $c$.

Step 7. Because of Step 4, we need to check also the constant function $y = 0$; it turns out that it is a solution too. It can be considered as the function $ce^{t^2}$ for $c = 0$.

Conclusion: The general solution to $\dot{y} - 2ty = 0$ is

$$y = ce^{t^2}, \text{ where } c \text{ is an arbitrary real number.} \, \square$$

Step 8. Plugging in $y = ce^{t^2}$ to $\dot{y} - 2ty = 0$ gives $ce^{t^2}(2t) - 2tce^{t^2} = 0$, which is true, as it should be.

4. Linear ODEs vs. Nonlinear ODEs

4.1. Linear ODEs.

4.1.1. Homogeneous. A homogeneous linear ODE is a differential equation such as

$$e^t \dot{y} + 5 \dot{y} + t^9 y = 0$$

in which each summand is a function of $t$ times one of $y$, $\dot{y}$, $\ddot{y}$, . . . .

Most general $n^{\text{th}}$ order homogeneous linear ODE:

$$p_n(t) \dot{y}^{(n)} + \cdots + p_1(t) \dot{y} + p_0(t) y = 0$$
for some functions \( p_n(t), \ldots, p_0(t) \) called the **coefficients**.

4.1.2. *Inhomogeneous*. An inhomogeneous linear ODE is the same except that it has also one term that is a function of \( t \) only; this function is usually moved to the right hand side. For example,

\[
e^t \ddot{y} + 5 \dot{y} + t^9 y = 7 \sin t
\]

is a second-order inhomogeneous linear ODE.

**Most general \( n \)th order inhomogeneous linear ODE:**

\[
p_n(t) y^{(n)} + \cdots + p_1(t) \dot{y} + p_0(t) y = q(t)
\]

for some functions \( p_n(t), \ldots, p_0(t), q(t) \). Imagine feeding different “input signals” \( q(t) \) into the right hand side of an inhomogeneous linear ODE to see what “output signals” \( y(t) \) the system responds with.

4.1.3. *Both kinds together*. A *linear ODE* is an ODE that can be rearranged into either of the two types above.

To simplify slightly, divide the whole equation by the leading coefficient \( p_n(t) \) to get rid of it:

\[
y^{(n)} + p_{n-1}(t) y^{(n-1)} + \cdots + p_1(t) \dot{y} + p_0(t) y = q(t)
\]

for some functions \( p_{n-1}(t), \ldots, p_0(t), q(t) \) (not the same ones as before).

We always assume that we are looking for a solution \( y(t) \) defined on an open interval \( I \), and that the functions \( p_{n-1}(t), \ldots, p_0(t), q(t) \) are continuous (or at least piecewise continuous) on \( I \). **Open interval** means a connected set of real numbers without endpoints, i.e., one of the following: \((a, b)\), \((-\infty, b)\), \((a, \infty)\), or \((-\infty, \infty) = \mathbb{R}\).

**Remark 4.1.** If you already know that an ODE is linear, there is an easy test to decide if it is homogeneous or not: Plug in the constant function \( y = 0 \).

- If \( y = 0 \) is a solution, the ODE is homogeneous.
- If \( y = 0 \) is not a solution, the ODE is inhomogeneous.

4.2. *Nonlinear ODEs*. For an ODE to be nonlinear, the functions \( y, \dot{y}, \ldots \) must enter the equation in a more complicated way: raised to powers, multiplied by each other, or with nonlinear functions applied to them.

**Flashcard question:** Which of the following ODEs is linear?

\[
\begin{align*}
\dot{y} - 7t y \ddot{y} & = 0 \\
\dot{y} & = e^t (y + t^2) \\
\dot{y} - y^2 & = 0 \\
\dot{y}^2 - ty & = \sin t \\
\dot{y} & = \cos(y + t).
\end{align*}
\]
Answer: The second one is linear since it can be rearranged into

\[ \dot{y} + (-e^t)y = t^2 e^t. \]

The others are nonlinear (the nonlinear portion is highlighted in red).

---

**February 7**

5. **Solving a first-order linear ODE**

Every *first-order* linear ODE in simplified form is as follows:

- **Homogeneous:** \( \dot{y} + p(t) y = 0 \)
- **Inhomogeneous:** \( \dot{y} + p(t) y = q(t) \).

5.1. **Homogeneous equations: separation of variables.** Homogeneous first-order linear ODEs can always be solved by separation of variables:

\[
\begin{align*}
\dot{y} + p(t) y &= 0 \\
\frac{dy}{dt} + p(t) y &= 0 \\
\frac{dy}{dt} &= -p(t) y \\
\frac{dy}{y} &= -p(t) dt
\end{align*}
\]

Choose *any* antiderivative \( P(t) \) of \( p(t) \). Then

\[
\ln |y| = -P(t) + C \\
|y| = \pm e^{-P(t)+C} \\
y = ce^{-P(t)},
\]

where \( c \) is any number (we brought back the solution \( y = 0 \) corresponding to \( c = 0 \)).

If you choose a different antiderivative, it will have the form \( P(t) + d \) for some constant \( d \), and then the new \( e^{-P(t)} \) is just a constant \( e^{-d} \) times the old one, so the set of *all* scalar multiples of the function \( e^{-P(t)} \) is the same as before.

**Conclusion:**

**Theorem 5.1** (General solution to first-order homogeneous linear ODE). Let \( p(t) \) be a continuous function on an open interval \( I \) (this ensures that \( p(t) \) has an antiderivative). Let \( P(t) \) be any antiderivative of \( p(t) \). Then the general solution to \( \dot{y} + p(t) y = 0 \) is \( y = ce^{-P(t)} \), where \( c \) is a parameter.
5.2. **Inhomogeneous equations: variation of parameters.** Variation of parameters is a method for solving inhomogeneous linear ODEs. Given a first-order inhomogeneous linear ODE

\[ \dot{y} + p(t) y = q(t), \]  

(3)

follow these steps:

1. Find a nonzero solution, say \( y_h \), of the associated homogeneous ODE

\[ \dot{y} + p(t) y = 0. \]

2. For an undetermined function \( u(t) \), substitute

\[ y = u(t) y_h(t) \]  

(4)

into the inhomogeneous equation (3) to find out which choices of \( u(t) \) make this \( y \) a solution to the inhomogeneous equation.

3. Once the general \( u(t) \) is found, plug it back into (4) to get the general solution to the inhomogeneous equation.

The idea is that the functions of the form \( cy_h \) are solutions to the homogeneous equation; maybe we can get solutions to the inhomogeneous equation by allowing the parameter \( c \) to vary, i.e., if we replace it by a nonconstant function \( u(t) \).

**Problem 5.2.** Solve \( t \dot{y} + 2y = t^5 \) on the interval \((0, \infty)\).

**Solution:**

*Step 1.* The associated homogeneous equation is \( t \dot{y} + 2y = 0 \), or equivalently, \( \dot{y} + \frac{2}{t} y = 0 \). Solve by separation of variables:

\[
\begin{align*}
\frac{dy}{dt} &= -\frac{2}{t} y \\
\frac{dy}{y} &= -\frac{2}{t} dt \\
\ln |y| &= -2 \ln t + C \quad \text{(since } t > 0) \\
y &= ce^{-2\ln t} \\
y &= ct^{-2}.
\end{align*}
\]

Choose one nonzero solution, say \( y_h = t^{-2} \).

*Step 2.* Substitute \( y = ut^{-2} \) into the inhomogeneous equation: the left side is

\[
t \dot{y} + 2y = t(\dot{u}t^{-2} + u(-2t^{-3})) + 2ut^{-2} = t^{-1} \dot{u},
\]
so the inhomogeneous equation becomes

\[ t^{-1} \dot{u} = t^5 \]
\[ \dot{u} = t^6 \]
\[ u = \frac{t^7}{7} + c. \]

**Step 3.** The general solution to the inhomogeneous equation is

\[ y = ut^{-2} = \left( \frac{t^7}{7} + c \right) t^{-2} = \frac{t^5}{7} + ct^{-2}. \]

(If you want, check by direct substitution that this really is a solution.) □

### 5.3. Inhomogeneous equations: integrating factor. (To be done in recitation.)

Another approach to solving

\[ \dot{y} + p(t)y = q(t) \] \hspace{1cm} (5)

is to use an **integrating factor**:

1. Find an antiderivative \( P(t) \) of \( p(t) \).
2. Multiply both sides of the ODE by the integrating factor \( e^{P(t)} \) in order to make the left side the derivative of something:

\[ e^{P(t)} \dot{y} + e^{P(t)} p(t)y = q(t)e^{P(t)} \]
\[ \frac{d}{dt} \left( e^{P(t)}y \right) = q(t)e^{P(t)} \]
\[ e^{P(t)}y = \int q(t)e^{P(t)} dt \]
\[ y = e^{-P(t)} \int q(t)e^{P(t)} dt. \]

Here \( \int q(t)e^{P(t)} dt \) represents all possible antiderivatives of \( q(t)e^{P(t)} \), so there are infinitely many solutions.

If you fix one antiderivative, say \( R(t) \), then the others are \( R(t) + c \) for a constant \( c \), so the general solution is

\[ y = R(t)e^{-P(t)} + ce^{-P(t)}. \] □

### 5.4. Linear combinations. A **linear combination** of a list of functions is any function that can be built from them by scalar multiplication and addition.

- **Linear combinations of** \( f(t) \): the functions \( cf(t) \), where \( c \) is any number
- **Linear combinations of** \( f_1(t) \) and \( f_2(t) \): the functions of the form \( c_1f_1(t) + c_2f_2(t) \), where \( c_1 \) and \( c_2 \) are any numbers.

::
Examples:

- $2 \cos t + 3 \sin t$ is a linear combination of the functions $\cos t$ and $\sin t$.
- $9t^5 + 3$ is a linear combination of the functions $t^5$ and 1.

**Flashcard question:** One of the functions below is not a linear combination of $\cos^2 t$ and 1. Which one?

1. $3 \cos^2 t - 4$
2. $\sin^2 t$
3. $\sin(2t)$
4. $\cos(2t)$
5. 5
6. 0

**Answer:** 3.

All the others are linear combinations:

\[
3 \cos^2 t - 4 = 3 \cos^2 t + (-4) \cdot 1
\]
\[
\sin^2 t = (-1) \cos^2 t + 1 \cdot 1
\]
\[
\sin(2t) = ???
\]
\[
\cos(2t) = 2 \cos^2 t + (-1) \cdot 1
\]
\[
5 = 0 \cos^2 t + 5 \cdot 1
\]
\[
0 = 0 \cos^2 t + 0 \cdot 1.
\]

Could there be some fancy identity that expresses $\sin(2t)$ as a linear combination of $\cos^2 t$ and 1? No; here’s one way to see this: Every linear combination of $\cos^2 t$ and 1 has the form

\[c_1 \cos^2 t + c_2\]

for some numbers $c_1$ and $c_2$. All such functions are *even* functions, but $\sin(2t)$ is an *odd* function. (Warning: This trick might not work in other situations.)
5.5. **Superposition.** Let’s compare the solutions to a homogeneous equation and some inhomogeneous equations with the same left hand side:

- **general solution to** \( t \ddot{y} + 2y = 0 \): \( ct^{-2} \)
- **one solution to** \( t \ddot{y} + 2y = t^5 \): \( \frac{t^5}{7} \)
- **general solution to** \( t \ddot{y} + 2y = t^5 \): \( \frac{t^5}{7} + ct^{-2} \)
- **one solution to** \( t \ddot{y} + 2y = 1 \): \( \frac{1}{2} \)
- **general solution to** \( t \ddot{y} + 2y = 1 \): \( \frac{1}{2} + ct^{-2} \).

From each “one solution” above, scalar-multiply to get

- **one solution to** \( t \ddot{y} + 2y = 9t^5 \): \( \frac{9t^5}{7} \)
- **one solution to** \( t \ddot{y} + 2y = 3 \): \( \frac{3}{2} \)

and add to get

- **one solution to** \( t \ddot{y} + 2y = 9t^5 + 3 \): \( \frac{9t^5}{7} + \frac{3}{2} \).

The general principle, which works for all linear ODEs, is this:

**Superposition principle.**

1. Multiplying a solution to \( p_n(t) y^{(n)} + \cdots + p_0(t) y = q(t) \) by a number \( a \) gives a solution to \( p_n(t) y^{(n)} + \cdots + p_0(t) y = aq(t) \).

2. Adding a solution of \( p_n(t) y^{(n)} + \cdots + p_0(t) y = q_1(t) \) to a solution of \( p_n(t) y^{(n)} + \cdots + p_0(t) y = q_2(t) \) gives a solution of \( p_n(t) y^{(n)} + \cdots + p_0(t) y = q_1(t) + q_2(t) \).

Using both parts shows that linear combinations of \( y \)'s solve the ODE with the corresponding linear combination of the \( q \)'s.
5.5.1. **Consequence of superposition for a homogeneous linear ODE.** Apply superposition with right hand sides equal to 0.

**Conclusion:** The set $S$ of solutions to a homogeneous linear ODE has the following properties:

0. The zero function 0 is in $S$.
1. Multiplying any one function in $S$ by a scalar gives another function in $S$.
2. Adding any two functions in $S$ gives another function in $S$.

A set of functions with these properties is called a **vector space** of functions, since these properties say that you can scalar-multiply and add such functions, as you can with vectors. (One can also talk about vector spaces of vectors, or vector spaces of matrices. There is a more abstract notion of vector space that includes all these as special cases.)

Thus:

**Theorem 5.3.** For any homogeneous linear ODE, the set of all solutions is a vector space.

Theorem 5.3 is why homogeneous linear ODEs are so nice. It says that if you know some solutions, you can form linear combinations to build new solutions, with no extra work! This is the key point of linearity in the homogeneous case. We will use it over and over again in applications throughout the course.

5.5.2. **Consequence of superposition for an inhomogeneous linear ODE.** To understand the general solution $y_i$ to an inhomogeneous linear ODE

$$\text{inhomogeneous: } p_n(t) y^{(n)} + \cdots + p_0(t) y = q(t),$$

do the following:

1. List all solutions to the associated homogeneous equation

$$\text{homogeneous: } p_n(t) y^{(n)} + \cdots + p_0(t) y = 0;$$

i.e., write down the general solution $y_h$.
2. Find (in some way) any one particular solution $y_p$ to the inhomogeneous ODE.
3. Add $y_p$ to all the solutions of the homogeneous ODE to get all the solutions to the inhomogeneous ODE.

**Summary:**

$$y_i = y_p + y_h$$

Why does this work? Superposition says that adding $y_p$ to a solution with right hand side 0 gives a solution with right hand side $q(t)$. All solutions with right hand side $q(t)$ arise this way, since subtracting $y_p$ from one of them gives a solution with right hand side 0.
The result of this section is the key point of linearity in the inhomogeneous case. It lets you build all the solutions to a inhomogeneous DEs out of one (provided that you have already solved the associated homogeneous ODE).

February 10

5.6. Newton’s law of cooling.

Problem 5.4. My minestrone soup is in an insulating thermos. Model its temperature as a function of time.

Simplifying assumptions:
- The insulating ability of the thermos does not change with time.
- The rate of cooling depends only on the difference between the soup temperature and the external temperature.

Variables and functions (with units): Define the following:
- $t$: time (minutes)
- $x$: external temperature ($^\circ$C)
- $y$: soup temperature ($^\circ$C)

Here $t$ is the independent variable, and $x$ and $y$ are functions of $t$.

Equation:

$$\dot{y} = f(y - x)$$

for some function $f$. Another simplifying assumption: $f(z) = -kz + \ell$ for some constants $k$ and $\ell$ (any reasonable function be approximated on small inputs by its linearization at 0); this leads to

$$\dot{y} = -k(y - x) + \ell.$$

Common sense says
- If $y = x$, then $\dot{y} = 0$. Thus $\ell$ should be 0.
- If $y > x$, then $y$ is decreasing. This is why we wrote $-k$ instead of just $k$.

So the equation becomes

$$\dot{y} = -k(y - x).$$

This is Newton’s law of cooling: the rate of cooling of an object is proportional to the difference between its temperature and the external temperature. The (positive) constant $k$ is called the coupling constant, in units of minutes$^{-1}$; smaller $k$ means better insulation, and $k = 0$ is perfect insulation. This ODE can be rearranged into standard form:

$$\dot{y} + ky = kx.$$
It's a first-order inhomogeneous linear ODE! The input signal is $x$, the system is the thermos, and the output signal is $y$.

If the external temperature $x$ is constant, then

- Particular solution to the inhomogeneous ODE: $y_p = x$
  (the solution in which the soup is already in equilibrium with the exterior)
- General solution to the homogeneous ODE: $y_h = ce^{-kt}$.
- General solution to the inhomogeneous ODE: $y = x + ce^{-kt}$.
  (As $t \to \infty$, the soup temperature approaches $x$; this makes sense.)

6. Existence and uniqueness of solutions

Using separation of variables (in the homogeneous case) and variation of parameters (in the inhomogeneous case), we showed that every first-order linear ODE has a 1-parameter family of solutions. To nail down a specific solution in this family, we need one initial condition, such as $y(0)$.

It will turn out that every second-order linear ODE has a 2-parameter family of solutions. To nail down a specific solution, we need two initial conditions at the same starting time, $y(0)$ and $\dot{y}(0)$. The starting time could also be some number $a$ other than 0.

Here is the general result:

**Existence and uniqueness theorem for a linear ODE.** Let $p_{n-1}(t), \ldots, p_0(t), q(t)$ be continuous functions on an open interval $I$. Let $a \in I$, and let $b_0, \ldots, b_{n-1}$ be given numbers. Then there exists a unique solution to the $n$th order linear ODE

$$y^{(n)} + p_{n-1}(t) y^{(n-1)} + \cdots + p_1(t) \dot{y} + p_0(t) y = q(t)$$

satisfying the $n$ initial conditions

$$y(a) = b_0, \quad \dot{y}(a) = b_1, \quad \ldots, \quad y^{(n-1)}(a) = b_{n-1}.$$

Existence means that there is at least one solution.

Uniqueness means that there is only one solution.

*Remark* 6.1. For a linear ODE as above, the solution $y(t)$ is defined on the whole interval $I$ where the functions $p_{n-1}(t), \ldots, p_0(t), q(t)$ are continuous. In particular, if $p_{n-1}(t), \ldots, p_0(t), q(t)$ are continuous on all of $\mathbb{R}$, then the solution $y(t)$ will be defined on all of $\mathbb{R}$.

7. Complex numbers

Complex numbers are expressions of the form $x + yi$, where $x$ and $y$ are real numbers, and $i$ is a new symbol. Multiplication of complex numbers will eventually be defined so that $i^2 = -1$. (Electrical engineers sometimes write $j$ instead of $i$, because they want to reserve $i$
for current, but everybody else thinks that’s weird.) Just as the set of all real numbers is denoted \( \mathbb{R} \), the set of all complex numbers is denoted \( \mathbb{C} \).

**Flashcard question:** Is 9 a real number or a complex number?

**Possible answers:**

1. real number
2. complex number
3. both
4. neither

**Answer:** Both, because 9 can be identified with 9 + 0i.

### 7.1. Operations on complex numbers.

<table>
<thead>
<tr>
<th>real part</th>
<th>( \text{Re}(x + yi) := x )</th>
</tr>
</thead>
<tbody>
<tr>
<td>imaginary part</td>
<td>( \text{Im}(x + yi) := y ) (Note: It is ( y ), not ( yi ), so ( \text{Im}(x + yi) ) is real)</td>
</tr>
</tbody>
</table>
| complex conjugate | \( 
\bar{x} + yi := x - yi \) (negate the imaginary component) |

One can add, subtract, multiply, and divide complex numbers (except for division by 0). Addition, subtraction, and multiplication are as for polynomials, except that after multiplication one should simplify by using \( i^2 = -1 \); for example,

\[
(2 + 3i)(1 - 5i) = 2 - 7i - 15i^2 = 2 - 7i + 15 = 17 - 7i.
\]

To divide \( z \) by \( w \), multiply \( z/w \) by \( \bar{w}/\bar{w} \) so that the denominator becomes real; for example,

\[
\frac{2 + 3i}{1 - 5i} = \frac{2 + 3i}{1 - 5i} \cdot \frac{1 + 5i}{1 + 5i} = \frac{2 + 13i + 15i^2}{1 - 25i^2} = \frac{-13 + 13i}{26} = -\frac{1}{2} + \frac{1}{2}i.
\]

The arithmetic operations on complex numbers satisfy the same properties as for real numbers (\( zw = wz \) and so on). The mathematical jargon for this is that \( \mathbb{C} \), like \( \mathbb{R} \), is a **field**. In particular, for any complex number \( z \) and integer \( n \), the \( n \text{th} \) **power** \( z^n \) can be defined in the usual way (need \( z \neq 0 \) if \( n < 0 \)); e.g., \( z^3 := zzz \), \( z^0 := 1 \), \( z^{-3} := 1/z^3 \). (Warning: Although there is a way to define \( z^n \) also for a complex number \( n \), when \( z \neq 0 \), it turns out that \( z^n \) has more than one possible value for non-integral \( n \), so it is ambiguous notation. Anyway, the most important cases are \( e^z \), and \( z^n \) for integers \( n \); the other cases won’t even come up in this class.)

If you change every \( i \) in the universe to \( -i \) (that is, take the complex conjugate everywhere), then all true statements remain true. For example, \( i^2 = -1 \) becomes \( (-i)^2 = -1 \). Another example: If \( z = vw \), then \( \overline{z} = \overline{v} \overline{w} \).
7.2. **The complex plane.** Just as real numbers can be visualized as points on a line, complex numbers can be visualized as points in a plane: plot $x + yi$ at the point $(x, y)$.

Addition and subtraction of complex numbers has the same geometric interpretation as for vectors. The same holds for scalar multiplication of a complex number by a real number. (The geometric interpretation of multiplication by a complex number is different; we’ll explain it soon.) Complex conjugation reflects a complex number in the real axis.

The **absolute value** (or magnitude or modulus) $|z|$ of a complex number $z = x + iy$ is its distance to the origin:

$$|x + yi| := \sqrt{x^2 + y^2} \quad \text{(this is a real number)}.$$  

For a complex number $z$, inequalities like $z < 3$ do not make sense, but inequalities like $|z| < 3$ do, because $|z|$ is a real number. The complex numbers satisfying $|z| < 3$ are those in the open disk of radius 3 centered at 0 in the complex plane. (Open disk means the disk without its boundary.)
7.3. **Some useful identities.** The following are true for all complex numbers \( z \):

\[
\text{Re} z = \frac{z + \overline{z}}{2}, \quad \text{Im} z = \frac{z - \overline{z}}{2i}, \quad \overline{z} = z, \quad z\overline{z} = |z|^2.
\]

Also, for any real number \( a \) and complex number \( z \),

\[
\text{Re}(az) = a \text{Re} z, \quad \text{Im}(az) = a \text{Im} z.
\]

(These can fail if \( a \) is not real.)

**Proof of the first identity:** Write \( z \) as \( x + yi \). Then \( \text{Re} z = x \) and \( \frac{z + \overline{z}}{2} = \frac{(x+yi)+(x-yi)}{2} = x \) too.

The proofs of the others are similar.

7.4. **Complex roots of polynomials.**

**real polynomial**: polynomial with real coefficients

**complex polynomial**: polynomial with complex coefficients

*Example 7.1.* How many roots does the polynomial \( z^3 - 3z^2 + 4 \) have? It factors as \((z - 2)(z - 2)(z + 1)\), so it has only two distinct roots (2 and \(-1\)). But if we count 2 twice, then the number of roots *counted with multiplicity* is 3, equal to the degree of the polynomial.

Some real polynomials, like \( z^2 + 9 \), cannot be factored completely into degree 1 real polynomials, but do factor into degree 1 complex polynomials: \((z + 3i)(z - 3i)\). In fact, *every* complex polynomial factors completely into degree 1 complex polynomials — this is proved in advanced courses in complex analysis. This implies the following:

**Fundamental theorem of algebra.** *Every degree \( n \) complex polynomial \( f(z) \) has exactly \( n \) complex roots, if counted with multiplicity.*
Since real polynomials are special cases of complex polynomials, the fundamental theorem of algebra applies to them too. For real polynomials, the non-real roots can be paired off with their complex conjugates.

**Example 7.2.** The degree 3 polynomial \( z^3 + z^2 - z + 15 \) factors as \((z+3)(z-1-2i)(z-1+2i)\), so it has three distinct roots: \(-3, 1+2i, \) and \(1-2i\). Of these roots, \(-3\) is real, and \(1+2i\) and \(1-2i\) form a complex conjugate pair.

**Example 7.3.** Want a fourth root of \(i\)? The fundamental theorem of algebra guarantees that \( z^4 - i = 0 \) has a complex solution (in fact, four of them). We’ll soon learn how to find them.

The fundamental theorem of algebra will be useful for constructing solutions to higher order linear ODEs with constant coefficients, and for discussing eigenvalues.

---

**February 12**

7.5. **Real and imaginary parts of complex-valued functions.** Suppose that \( y(t) \) is a complex-valued function of a real variable \( t \). Then

\[
y(t) = f(t) + i\, g(t)
\]

for some real-valued functions of \( t \). Here \( f(t) := \text{Re} \, y(t) \) and \( g(t) := \text{Im} \, y(t) \). Differentiation and integration can be done component-wise:

\[
y'(t) = f'(t) + i\, g'(t)
\]

\[
\int y(t) \, dt = \int f(t) \, dt + i \int g(t) \, dt.
\]

**Example 7.4.** Suppose that \( y(t) = \frac{2+3i}{1+it} \). Then

\[
y(t) = \frac{2+3i}{1+it} = \frac{2+3i}{1-it} = \frac{(2+3t) + i(3-2t)}{1+t^2} = \left(\frac{2+3t}{1+t^2}\right) f(t) + i \left(\frac{3-2t}{1+t^2}\right) g(t).
\]

The functions in parentheses labelled \( f(t) \) and \( g(t) \) are real-valued, so these are the real and imaginary parts of the function \( y(t) \).

7.6. **The complex exponential function.** Derivatives and DEs make sense for complex-valued functions of a complex variable \( z \), and work in a similar way. In particular, the existence and uniqueness theorem shows that there is a unique such function \( f(z) \) satisfying

\[
f'(z) = f(z), \quad f(0) = 1.
\]

This function is called the complex exponential function \( e^z \).
The number $e$ is defined as the value of $e^z$ at $z = 1$. But it is the function $e^z$, not the number $e$, that is truly important. Defining $e$ without defining $e^z$ first is a little unnatural. And even if $e$ were defined first, one could not use it to define $e^z$, because “$e$ raised to a complex number” has no a priori meaning.

**Theorem 7.5.** The complex exponential function $e^z$ has the following properties:

(a) The derivative of $e^z$ is $e^z$.

(b) $e^0 = 1$.

(c) $e^{z+w} = e^z e^w$ for all complex numbers $z$ and $w$.

(d) $(e^z)^n = e^{nz}$ for every complex number $z$ and integer $n$. The $n = -1$ case says $\frac{1}{e^z} = (e^z)^{-1} = e^{-z}$.

(e) Euler’s identity:

$$e^{it} = \cos t + i \sin t$$
for every real number $t$.

(f) More generally,

$$e^{x+yi} = e^x (\cos y + i \sin y)$$
for all real numbers $x$ and $y$. (6)

(g) $e^{-it} = \overline{e^{it}} = \cos t - i \sin t$ for every real number $t$.

(h) $|e^{it}| = 1$ for every real number $t$.

Of lesser importance is the power series representation

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \cdots.$$ (7)

This formula can be deduced by using Taylor’s theorem with remainder, or by showing that the right hand side satisfies the DE and initial condition. Some books use (6) or (7) as the definition of the complex exponential function, but the DE definition we gave is less contrived and focuses on what makes the function useful.

**Proof of Theorem 7.5.**

(a) True by definition.

(b) True by definition.

(c) As a warm-up, consider the special case in which $w = 3$. By the chain rule, $e^{z+3}$ is the solution to the DE with initial condition

$$f'(z) = f(z), \quad f(0) = e^3.$$

The function $e^z e^3$ satisfies the same DE with initial condition. By uniqueness, the two functions are the same: $e^{z+3} = e^z e^3$. The same argument works for any other complex constant $w$ in place of 3, so $e^{z+w} = e^z e^w$.

(d) If $n = 0$, then this is $1 = 1$ by definition. If $n > 0$,

$$(e^z)^n = \underbrace{e^z e^z \cdots e^z}_{\text{n copies}} = (e^z)^n = \underbrace{e^z + e^z + \cdots + e^z}_{\text{n copies}} = e^{nz}.$$
If \( n = -m < 0 \), then
\[
(e^z)^{-m} = \frac{1}{(e^z)^m} = \frac{1}{e^{mz}} = e^{-mz}
\]
since \( e^{mz}e^{-mz} = e^{mz+(-mz)} = e^0 = 1 \).

(e) The calculation
\[
\frac{d}{dt} (\cos t + i \sin t) = -\sin t + i \cos t
\]
\[
= i(\cos t + i \sin t)
\]
shows that the function \( \cos t + i \sin t \) is the solution to the DE with initial condition
\[
f'(t) = if(t), \quad f(0) = 1.
\]

But \( e^{it} \) is a solution too, by the chain rule. By uniqueness, the two functions are the same
(the existence and uniqueness theorem applies also to complex-valued functions of a real variable \( t \)).

(f) By (c) and (e), \( e^{x+yi} = e^x e^{iy} = e^x (\cos y + i \sin y) \).

(g) Changing every \( i \) in the universe to \( -i \) transforms \( e^{it} = \cos t + i \sin t \) into \( e^{-it} = \cos t - i \sin t \).
(Substituting \(-t \) for \( t \) would do it too.) On the other hand, applying complex conjugation
to both sides of \( e^{it} = \cos t + i \sin t \) gives \( \overline{e^{it}} = \cos t - i \sin t \).

(h) By (e), \( |e^{it}| = \sqrt{\cos^2 t + \sin^2 t} = \sqrt{1} = 1 \). \( \square \)

Remark 7.6. Some older books use the awful abbreviation \( \text{cis } t := \cos t + i \sin t \), but this belongs in a cispool [sic], since \( e^{it} \) is a more useful expression for the same thing.

As \( t \) increases, the complex number \( e^{it} = \cos t + i \sin t \) travels counterclockwise around the unit circle.

7.7. Polar form of a complex number. Given a nonzero complex number \( z = x + yi \), we can express the point \((x, y)\) in polar coordinates \( r \) and \( \theta \):
\[
x = r \cos \theta, \quad y = r \sin \theta.
\]
Then

\[ x + yi = (r \cos \theta) + (r \sin \theta)i = r(\cos \theta + i \sin \theta). \]

In other words,

\[ z = re^{i\theta}. \]

Here \( re^{i\theta} \) is called a polar form of the complex number \( z \). One has \( r = |z| \); here \( r \) must be a positive real number (assuming \( z \neq 0 \)).

Any possible \( \theta \) for \( z \) (a possible value for the angle or argument of \( z \)) may be called \( \text{arg} \, z \), but this is dangerously ambiguous notation since there are many values of \( \theta \) for the same \( z \): this means that \( \text{arg} \, z \) is not a function.

**Example 7.7.** Suppose that \( z = -3i \). So \( z \) corresponds to the point \((0, -3)\). Then \( r = |z| = 3 \), but there are infinitely many possibilities for the angle \( \theta \). One possibility is \(-\pi/2\); all the others are obtained by adding integer multiples of \( 2\pi \):

\[ \text{arg} \, z = \ldots, -5\pi/2, -\pi/2, 3\pi/2, 7\pi/2, \ldots \]
So \( z \) has many polar forms:

\[
\ldots = 3e^{i(-5\pi/2)} = 3e^{-i\pi/2} = 3e^{i(3\pi/2)} = 3e^{i(7\pi/2)} = \ldots \quad \square
\]

To specify a unique polar form, we would have to restrict the range for \( \theta \) to some interval of width 2\( \pi \). The most common choice is to require \(-\pi < \theta \leq \pi\). This special \( \theta \) is called the principal value of the argument, and is denoted in various ways:

\[
\theta = \text{Arg} z = \text{Arg}[z] = \text{ArcTan}[x,y] = \text{atan2}(y,x) \quad \text{Mathematica} \\
\text{Mathematica} = \text{atan2}(y,x) \quad \text{MATLAB}
\]

Warning: The supplementary notes require \( 0 \leq \theta < 2\pi \) instead. Warning: In MATLAB, be careful to use \((y,x)\) and not \((x,y)\). Warning: Although the principal value \( \theta \) satisfies the “slope formula” \( \tan \theta = y/x \) whenever \( x \neq 0 \), the formula \( \theta = \tan^{-1}(y/x) \) is true only half the time. The problem is that there are two angles in the range \((-\pi,\pi]\) with the same value of \( \tan \) (there are two values of \( \theta \) for each line through the origin, differing by \( \pi \)), and \( \tan^{-1}(y/x) \) evaluated at \((1,1)\) and \((-1,-1)\) produces the same value \( \pi/4 \), but \((-1,-1)\) is actually at angle \(-3\pi/4\). The “2-variable arctangent function” used above fixes this.

**Test for equality** of two nonzero complex numbers in polar form:

\[
\left| r_1e^{i\theta_1}\right| = \left| r_2e^{i\theta_2}\right| \iff r_1 = r_2 \quad \text{and} \quad \theta_1 = \theta_2 + 2\pi k \quad \text{for some integer} \ k.
\]

(This assumes that \( r_1 \) and \( r_2 \) are positive real numbers, and that \( \theta_1 \) and \( \theta_2 \) are real numbers, as you would expect for polar coordinates.)

**7.8. Operations in polar form.** Some arithmetic operations on complex numbers are easy in polar form:

- **multiplication:** \((r_1e^{i\theta_1})(r_2e^{i\theta_2}) = r_1r_2e^{i(\theta_1+\theta_2)}\) (multiply absolute values, add angles)
- **reciprocal:** \(\frac{1}{re^{i\theta}} = \frac{1}{r}e^{-i\theta}\)
- **division:** \(\frac{r_1e^{i\theta_1}}{r_2e^{i\theta_2}} = \frac{r_1}{r_2}e^{i(\theta_1-\theta_2)}\) (divide absolute values, subtract angles)
- **\(n\)th power:** \((re^{i\theta})^n = r^ne^{in\theta}\) for any integer \(n\)
- **complex conjugation:** \(\overline{re^{i\theta}} = re^{-i\theta}\).

Taking absolute values gives identities:

\[
\left| z_1z_2 \right| = \left| z_1 \right| \left| z_2 \right|, \quad \left| \frac{1}{z} \right| = \left| \frac{1}{z} \right|, \quad \left| \frac{z_1}{z_2} \right| = \frac{\left| z_1 \right|}{\left| z_2 \right|}, \quad \left| z^n \right| = \left| z \right|^n, \quad \left| \overline{z} \right| = \left| z \right|.
\]

**Question 7.8.** What happens if you take a smiley in the complex plane and multiply each of its points by 3\(i\)?

**Solution:** Since \(i = e^{i\pi/2}\), multiplying by \(i\) adds \(\pi/2\) to the angle of each point; that is, it rotates counterclockwise by \(90^\circ\) (around the origin). Next, multiplying by 3 does what you would expect: dilate by a factor of 3. Doing both leads to...
For example, the nose was originally on the real line, a little less than 2, so multiplying it by $3i$ produces a big nose close to $(3i)^2 = 6i$. □
Question 7.9. How do you trap a lion?

Answer: Build a cage in the shape of the unit circle \( |z| = 1 \). Get inside the cage. Make sure that the lion is outside the cage. Apply the function \( 1/z \) to the whole plane. Voilà! The lion is now inside the cage, and you are outside it. (Only problem: There’s a lot of other stuff inside the cage too. Also, don’t stand too close to \( z = 0 \) when you apply \( 1/z \).)

Question 7.10. Why not always write complex numbers in polar form?

Answer: Because addition and subtraction are difficult in polar form!

7.9. The function \( e^{(a+bi)t} \). Fix a nonzero complex number \( a + bi \). As the real number \( t \) increases, the complex number \( (a + bi)t \) moves along a line through 0, and \( e^{(a+bi)t} \) moves along part of a line, a circle, or a spiral, depending on the value of \( a + bi \). Try the “Complex Exponential” mathlet

\[ \text{http://mathlets.org/mathlets/complex-exponential/} \]
to see this.

Example 7.11. Consider \( e^{(-5-2i)t} = e^{-5t}e^{(-2it)} \) as \( t \to \infty \). Its absolute value is \( e^{-5t} \), which tends to 0, so the point is moving inward. Its angle is \(-2t\), which is decreasing, so the point is moving clockwise. It’s spiraling inwards clockwise.

7.10. Finding \( n \)th roots.

7.10.1. An example.

Problem 7.12. What are the complex solutions to \( z^5 = -32 \)?

Solution: Rewrite the equation in polar form, using \( z = re^{i\theta} \):

\[
(r e^{i\theta})^5 = 32 e^{i\pi}
\]

\[
r^5 e^{i(5\theta)} = 32 e^{i\pi}
\]

for some integer \( k \)

\[
r = 2 \quad \text{and} \quad \theta = \frac{\pi}{5} + \frac{2\pi k}{5}
\]

These are numbers on a circle of radius 2; to get from one to the next (increasing \( k \) by 1), rotate by \( 2\pi/5 \). Increasing \( k \) five times brings the number back to its original position. So it’s enough to take \( k = 0, 1, 2, 3, 4 \). Answer:

\[
2e^{i(\pi/5)}, \ 2e^{i(3\pi/5)}, \ 2e^{i(5\pi/5)}, \ 2e^{i(7\pi/5)}, \ 2e^{i(9\pi/5)}.
\]

□
Remark 7.13. The fundamental theorem of algebra predicts that the polynomial $z^5 + 32$ has 5 roots when counted with multiplicity. We found 5 roots, so each must have multiplicity 1.

February 14

7.10.2. Roots of unity.

The same method shows that the $n^{\text{th}}$ roots of unity (the solutions to $z^n = 1$) are the numbers $e^{i(2\pi k/n)}$ for $k = 0, 1, 2, \ldots, n - 1$. Taking $k = 1$ gives the number $\zeta := e^{2\pi i/n}$. In terms of $\zeta$, the complete list of $n^{\text{th}}$ roots of unity is

\[1, \zeta, \zeta^2, \ldots, \zeta^{n-1}\]

(after that they start to repeat: $\zeta^n = 1$).
7.10.3. Finding $n^{th}$ roots in general.

**Problem 7.14.** Given $\alpha \in \mathbb{C}$ with $\alpha \neq 0$, what are the solutions to $z^n = \alpha$?

Write $\alpha$ as $re^{i\theta}$. Then the solutions to $z^n = \alpha$ are $\sqrt[n]{r}e^{i\left(\frac{\theta}{n} + \frac{2\pi k}{n}\right)}$ for $k = 0, 1, 2, \ldots, n-1$.

Another way to list the solutions: If $z_0$ is any particular solution, such as $\sqrt[n]{r}e^{i\theta/n}$, then the complete list of solutions is $z_0, \zeta z_0, \zeta^2 z_0, \ldots, \zeta^{n-1} z_0$. □


8. Linear combinations of cosine and sine; sinusoidal functions

8.1. $e^{it}$ and $e^{-it}$ as linear combinations of $\cos t$ and $\sin t$, and vice versa.

**Example 8.1.** The functions $e^{it}$ and $e^{-it}$ are linear combinations of the functions $\cos t$ and $\sin t$:

\[
e^{it} = \cos t + i \sin t
\]
\[
e^{-it} = \cos t - i \sin t.
\]

If we view $e^{it}$ and $e^{-it}$ as known, and $\cos t$ and $\sin t$ as unknown, then this is a system of two linear equations in two unknowns, and can be solved for $\cos t$ and $\sin t$. This gives

\[
\cos t = \frac{e^{it} + e^{-it}}{2}, \quad \sin t = \frac{e^{it} - e^{-it}}{2i}.
\]

Thus $\cos t$ and $\sin t$ are linear combinations of $e^{it}$ and $e^{-it}$. (Explicitly, $\sin t = \frac{1}{2i}(e^{it} + e^{-it})$.)

**Important:** The function $e^z$ has nicer properties than $\cos t$ and $\sin t$, so it is often a good idea to use these formulas to replace $\cos t$ and $\sin t$ by these combinations of $e^{it}$ and $e^{-it}$, or to view $\cos t$ and $\sin t$ as the real and imaginary parts of $e^{it}$.

Replacing $t$ by $\omega t$ in the identities above leads to

\[
e^{i\omega t} = \cos \omega t + i \sin \omega t
\]
\[
e^{-i\omega t} = \cos \omega t - i \sin \omega t.
\]

and

\[
\cos \omega t = \frac{e^{i\omega t} + e^{-i\omega t}}{2}, \quad \sin \omega t = \frac{e^{i\omega t} - e^{-i\omega t}}{2i}.
\]

8.2. Sinusoidal functions.
8.2.1. *Construction.* Start with the curve $y = \cos x$. Then

1. Shift the graph $\phi$ units to the right ($\phi$ is *phase lag*, measured in radians). (For example, shifting by $\phi = \pi/2$ gives the graph of $\sin x$, which reaches its maximum $\pi/2$ radians after $\cos x$ does.)
2. Compress the result horizontally by *dividing* by a scale factor $\omega$ (*angular frequency*, measured in radians/s).
3. Amplify (stretch vertically) by a factor of $A$ (*amplitude*).

(Here $A, \omega > 0$, but $\phi$ can be any real number.)

Result? The graph of a new function $f(t)$, called a *sinusoid function*.

8.2.2. *Formula.* What is the formula for $f(t)$? According to the instructions, each point $(x, y)$ on $y = \cos x$ is related to a point $(t, f(t))$ on the graph of $f$ by

$$t = \frac{x + \phi}{\omega}, \quad f = Ay.$$ 

Solving for $x$ gives $x = \omega t - \phi$; substituting into $f = Ay = A \cos x$ gives

$$f(t) = A \cos(\omega t - \phi).$$
8.2.3. **Alternative geometric description.** Alternatively, the graph of \( f(t) \) can be described geometrically in terms of

- **\( A \):** its **amplitude**, as above, how high the graph rises above the \( t \)-axis at its maximum
- **\( \tau \):** its **time lag**, also sometimes called \( t_0 \), a \( t \)-value at which a maximum is attained (s)
- **\( P \):** its **period**, the time for one complete oscillation (= width between successive maxima) (s or s/cycle)

How do \( \tau \) and \( P \) relate to \( \omega \) and \( \phi \)?

- \( \tau = \phi/\omega \), since this is the \( t \)-value for which the angle \( \omega t - \phi \) becomes 0.
- \( P = 2\pi/\omega \), since adding \( 2\pi/\omega \) to \( t \) increases the angle \( \omega t - \phi \) by \( 2\pi \).

There is also **frequency** \( \nu := 1/P \), measured in Hz = cycles/s. It is the number of complete oscillations per second. To convert from frequency \( \nu \) to angular frequency \( \omega \), multiply by \( 2\pi \) radians per cycle; thus \( \omega = 2\pi \nu = 2\pi/P \), which is consistent with the formula \( P = 2\pi/\omega \) above.

---

February 18

8.2.4. **Three forms.** There are three ways to write a sinusoid function:

- **amplitude-phase form:** \( A \cos(\omega t - \phi) \)
- **complex form:** \( \text{Re}(ce^{i\omega t}) \), where \( c \) is a complex number
- **linear combination:** \( a \cos \omega t + b \sin \omega t \), where \( a \) and \( b \) are real numbers

Different forms are useful in different contexts, so we’ll need to know how to convert between them. The following proposition explains how.

**Proposition 8.2.**

If \( Ae^{i\phi} \)

then \( A \cos(\omega t - \phi) \)

\( \text{Re}(ce^{i\omega t}) \)

\( a \cos \omega t + b \sin \omega t \).

Use these key equations to convert between the three forms!

(The little numbers refer to the proofs below.)

**Warning:** Don’t forget that it is \( \overline{c} \) and not \( c \) itself that appears in the key equations.

An equivalent form of the key equations (obtained by taking complex conjugates) is

If you ever forget the key equations above, you can do the conversion manually by going through the steps in the proof below.
Proof of Proposition 8.2.

1. 

\[ \text{Re} \left( ce^{i\omega t} \right) = \text{Re} \left( Ae^{-i\phi} e^{i\omega t} \right) \]
\[ = \text{Re} \left( Ae^{i(\omega t-\phi)} \right) \]
\[ = A \cos(\omega t - \phi). \]

2. 

\[ \text{Re} \left( ce^{i\omega t} \right) = \text{Re} \left( (a - bi)(\cos \omega t + i \sin \omega t) \right) \]
\[ = \text{Re} \left( a \cos \omega t + b \sin \omega t + i(\cdot\cdot\cdot) \right) \]
\[ = a \cos \omega t + b \sin \omega t \]

since \( a = A \cos \phi \) and \( b = A \sin \phi \) (when \( A, \phi \) are the polar coordinates of \((a, b)\)).

3. Using \( \cos(x - y) = \cos x \cos y + \sin x \sin y \) shows that

\[ A \cos(\omega t - \phi) = A \cos \omega t \cos \phi + A \sin \omega t \sin \phi \]
\[ = a \cos \omega t + b \sin \omega t. \]

4. 

\[ a \cos \omega t + b \sin \omega t = (a, b) \cdot (\cos \omega t, \sin \omega t) \]
\[ = |(a, b)| \left| (\cos \omega t, \sin \omega t) \right| \cos(\text{angle between the vectors}), \]
\[ \quad \text{(by the geometric interpretation of the dot product)} \]
\[ = A \cos(\omega t - \phi) \]

(Actually, it would have been enough to prove equality on two sides of the triangle.) \( \square \)

Problem 8.3. Convert \(-\cos(5t) - \sqrt{3}\sin(5t)\) to amplitude-phase form.

Solution: Given: \( a = -1, \ b = -\sqrt{3}, \ \omega = 5 \). Wanted: \( A, \omega, \phi \). So we use \( Ae^{i\phi} = a + bi \).
First, \( A = \sqrt{a^2 + b^2} = 2 \). The equation \( a = A \cos \phi \) says that \( \cos \phi = -1/2 \), so the angle of \((-1, -\sqrt{3})\) is \( \phi = -2\pi/3 \) (or this plus \( 2\pi k \) for any integer \( k \)). Thus the answer is

\[ A \cos(\omega t - \phi) = 2 \cos(5t + 2\pi/3). \] \( \square \)

Try the “Trigonometric Identity” mathlet

http://mathlets.org/mathlets/trigonometric-id/
8.3. **Complex gain, gain, and phase lag.** Consider a system with sinusoidal input signal and sinusoidal output signal of the same frequency. How can we compare the two sinusoids?

Write each sinusoid in complex form:

- input signal: \( \text{Re}(ce^{i\omega t}) \)
- output signal: \( \text{Re}(Ce^{i\omega t}) \).

Imagine feeding a corresponding “complex replacement” signal into the system and getting a complex output signal (this probably makes no physical sense, but do it anyway):

- complex input: \( ce^{i\omega t} \)
- complex output: \( Ce^{i\omega t} \).

Define **complex gain** as the factor by which the complex input signal has gotten “bigger” (in some weird complex sense):

\[
G := \frac{\text{complex output}}{\text{complex input}} = \frac{Ce^{i\omega t}}{ce^{i\omega t}} = \frac{C}{c}.
\]

Complex gain is a complex number.

**Question 8.4.** What is the physical interpretation of the complex gain \( G \), in terms of amplitudes and phases of the real signals?

The answers are in the two boxes below. The conversion \( \text{Re}(ce^{i\omega t}) = A \cos(\omega t - \phi) \) uses the key equation

\[
c = Ae^{-i\phi},
\]
so multiplying \( c \) by the complex scale factor \( G \) to get \( C \) amounts to

- multiplying the amplitude \( A \) by \( |G| \), and
- increasing \( \phi \) by \( -\arg G \).

The amplitude scale factor is called **gain**:

\[
\text{gain} := \frac{\text{output amplitude}}{\text{input amplitude}} = |G|.
\]

Gain is a nonnegative real number.

The increase in \( \phi \) is called the **phase lag** (of the output relative to the input):

\[
\text{phase lag} := \phi_{\text{output}} - \phi_{\text{input}} = -\arg G.
\]

Phase lag is a real number, measured in radians. **Warning:** This is a relative phase lag, different from the absolute phase lag defined earlier comparing a sinusoid to the standard sinusoid \( \cos x \).

If the phase lag is \( \pi/2 \), that means that the maximum of the output sinusoid occurs \( \pi/2 \) radians after the maximum of the input signal. (To get instead the relative time lag, divide the phase lag by \( \omega \).)
8.4. **Beats.** Bonus section! Not covered in lecture.

Try the “Beats” mathlet


**Beats** occur when two very nearby pitches are sounded simultaneously.

**Problem 8.5.** Consider two sinusoid sound waves of angular frequencies \( \omega + \epsilon \) and \( \omega - \epsilon \), say \( \cos((\omega + \epsilon)t) \) and \( \cos((\omega - \epsilon)t) \), where \( \epsilon \) is much smaller than \( \omega \). What happens when they are superimposed?

**Solution:** The sum is

\[
\cos((\omega + \epsilon)t) + \cos((\omega - \epsilon)t) = \text{Re}(e^{i(\omega+\epsilon)t}) + \text{Re}(e^{i(\omega-\epsilon)t})
\]

\[
= \text{Re}(e^{i\omega t}(e^{i\epsilon t} + e^{-i\epsilon t}))
\]

\[
= \text{Re}(e^{i\omega t}(2 \cos \epsilon t))
\]

\[
= (2 \cos \epsilon t) \text{Re}(e^{i\omega t})
\]

\[
= 2(\cos \epsilon t)(\cos \omega t).
\]

The function \( \cos \omega t \) oscillates rapidly between \( \pm 1 \). Multiplying it by the slowly varying function \( 2 \cos \epsilon t \) produces a rapid oscillation between \( \pm 2 \cos \epsilon t \), so one hears a sound wave of angular frequency \( \omega \) whose amplitude is the slowly varying function \( |2 \cos \epsilon t| \). □

---

9. **Higher-order linear ODEs with constant coefficients**

9.1. **Modeling: a spring-mass-dashpot system.**

**Problem 9.1.** A cart is attached to a spring attached to a wall. The cart is attached also to a dashpot, a damping device. (A dashpot could be a cylinder filled with oil that a piston moves through. Door dampers and car shock absorbers often actually work this way.) Also, there is an external force acting on the cart. Model the motion of the cart.
Solution: Define variables

\[ t \]: time (s)
\[ x \]: position of the cart (m), with \( x = 0 \) being where the spring exerts no force
\[ m \]: mass of the cart (kg)
\[ F_{\text{spring}} \]: force exerted by the spring on the cart (N)
\[ F_{\text{dashpot}} \]: force exerted by the dashpot on the cart (N)
\[ F_{\text{external}} \]: external force on the cart (N)
\[ F \]: total force on the cart (N).

The independent variable is \( t \); everything else is a function of \( t \) (well, maybe \( m \) is constant).

Physics tells us that

- \( F_{\text{spring}} \) is a function of \( x \) (of opposite sign), and
- \( F_{\text{dashpot}} \) is a function of \( \dot{x} \) (again of opposite sign).

To simplify, approximate these by \textit{linear} functions (probably OK if \( x \) and \( \dot{x} \) are small):

\[ F_{\text{spring}} = -kx, \quad F_{\text{dashpot}} = -b\dot{x}; \]

where \( k \) is the \textit{spring constant} (in units N/m) and \( b \) is the \textit{damping constant} (in units Ns/m); here \( k, b > 0 \). Substituting this and Newton’s second law \( F = m\ddot{x} \) into

\[ F = F_{\text{spring}} + F_{\text{dashpot}} + F_{\text{external}} \]

gives

\[ m\ddot{x} = -kx - b\dot{x} + F_{\text{external}}, \]

a second order linear ODE, which we would usually write as

\[ m\ddot{x} + b\dot{x} + kx = F_{\text{external}}(t). \]
All this works even if $m, b, k$ are functions of time, but we’ll assume from now on that they are constants.

- input signal: $F_{\text{external}}(t)$
- system: spring, mass, and dashpot
- output signal: $x(t)$.

Carts attached to springs are not necessarily what interest us. But oscillatory systems arising in all the sciences are governed by the same math, and this physical system lets us visualize their behavior.

9.2. A homogeneous example. To see how to solve ODEs like this, we’ll first consider a simpler example.

Problem 9.2. What are the solutions to

$$\ddot{y} + \dot{y} - 6y = 0? \quad (8)$$

Solution: Try $y = e^{rt}$, where $r$ is a constant to be determined. Then $\dot{y} = re^{rt}$ and $\ddot{y} = r^2e^{rt}$, so (8) becomes

$$r^2e^{rt} + re^{rt} - 6e^{rt} = 0$$

$$(r^2 + r - 6)e^{rt} = 0.$$ 

This holds as an equality of functions if and only if

$$r^2 + r - 6 = 0$$

$$(r - 2)(r + 3) = 0$$

$$r = 2 \quad \text{or} \quad r = -3.$$ 

So $e^{2t}$ and $e^{-3t}$ are solutions.

Now for the magic of linearity: Since the equation is a homogeneous linear ODE, the set of solutions is a vector space. This means that we can multiply by scalars and add, to get more solutions: $5e^{2t}$ is a solution, $7e^{-3t}$ is a solution, $5e^{2t} + 7e^{-3t}$ is a solution — any linear combination

$$c_1e^{2t} + c_2e^{-3t},$$

for any numbers $c_1$ and $c_2$, is a solution. It turns out that this is the general solution. □

(In the future, we’ll go straight from the ODE to the equation $r^2 - r - 6 = 0$, now that we know how this works.)
9.3. **Span.** To make it easier to understand more complicated examples, let’s reformulate the answer in terms of linear algebra.

**Definition 9.3.** Suppose that $f_1, \ldots, f_n$ are functions. The span of $f_1, \ldots, f_n$ is the set of all linear combinations of $f_1, \ldots, f_n$: 

\[
\text{Span}(f_1, \ldots, f_n) := \{\text{all functions } c_1 f_1 + \cdots + c_n f_n, \text{ where } c_1, \ldots, c_n \text{ are numbers}\}.
\]

(If not otherwise specified, we allow the numbers $c_1, \ldots, c_n$ to be complex.)

The span is always a vector space.

**Example 9.4.** \(\text{Span}(t^2)\) is the set of all functions of the form \(ct^2\), where \(c\) is a number. It is an infinite set of functions:

\[
\text{Span}(t^2) = \{\ldots, -3t^2, 4t^2, 0, t^2, -t^2, \pi t^2, \ldots\}
\]

**Flashcard question:** Is it true that

\[
\text{Span}(2t^2) = \text{Span}(t^2) ?
\]

**Possible answers:**
1. Yes, these sets are equal.
2. No, there is a function in \(\text{Span}(2t^2)\) that is not in \(\text{Span}(t^2)\).
3. No, there is a function in \(\text{Span}(t^2)\) that is not in \(\text{Span}(2t^2)\).

**Answer:** 1. They are equal. Every function that is a number times \(2t^2\) is also a (different) number times \(t^2\), and vice versa. □

**Example 9.5.** \(\text{Span}(e^{2t}, e^{-3t})\) is the set of all solutions to \(\ddot{y} + \dot{y} - 6y = 0\). It is an infinite set of functions:

\[
\text{Span}(e^{2t}, e^{-3t}) = \{\ldots, 5e^{2t} + (-7)e^{-3t}, 0e^{2t} + 1e^{-3t}, \pi e^{2t} + 9e^{-3t}, 0e^{2t} + 0e^{-3t}, \ldots\}.
\]

**Flashcard question:** Is it true that

\[
\text{Span}(e^{2t}, e^{-3t}, e^{2t} + e^{-3t}) = \text{Span}(e^{2t}, e^{-3t}) ?
\]

**Possible answers:**
1. Yes, these sets are equal.
2. No, there are more possible linear combinations on the left.
3. No, there are more possible linear combinations on the right.
Answer: 1. They are equal. Any linear combination in the left hand side, such as
\[ 5e^{2t} + 6e^{-3t} + 2(e^{2t} + e^{-3t}) = 7e^{2t} + 8e^{-3t}. \]
is also a linear combination in the right hand side (and vice versa). □

So technically it is correct to say that
\[ c_1e^{2t} + c_2e^{-3t} + c_3(e^{2t} + e^{-3t}) \]
is the general solution to (8), but this is silly since the third function \( e^{2t} + e^{-3t} \) is redundant, because it is a linear combination of the other two.

February 19

9.4. Linearly dependent functions.

Definition 9.6. Functions \( f_1, \ldots, f_n \) are linearly dependent (think redundant) if at least one of them is a linear combination of the others. Otherwise, call them linearly independent.

Example 9.7. The functions \( e^{2t}, e^{-3t}, e^{2t} + e^{-3t} \) are linearly dependent, because the third function \( e^{2t} + e^{-3t} \) is a linear combination of the first two. (In fact, each function is a linear combination of the other two.)

Example 9.8. The functions \( e^{2t}, e^{-3t} \) are linearly independent.

Equivalent definition: Functions \( f_1, \ldots, f_n \) are linearly dependent if there exist numbers \( c_1, \ldots, c_n \) not all zero such that \( c_1f_1 + \cdots + c_nf_n = 0 \).

Example 9.9. The three functions \( e^{2t}, e^{-3t}, 5e^{2t} + 7e^{-3t} \) are linearly dependent, because
\[ (-5)e^{2t} + (-7)e^{-3t} + (1)(5e^{2t} + 7e^{-3t}) = 0. \]

Moral: When describing the space of solutions to a homogeneous linear ODE, it is most efficient to give it as the span of linearly independent functions.

9.5. Basis.

Definition 9.10. A basis of a vector space \( S \) is a list of functions \( f_1, f_2, \ldots \) such that
1. \( \text{Span}(f_1, f_2, \ldots) = S \), and
2. The functions \( f_1, f_2, \ldots \) are linearly independent.

The plural of basis is bases, pronounced BAY-sees.

Think of the functions \( f_1, f_2, \ldots \) in the basis as the “basic building blocks”: condition 1 says that every function in \( S \) can be built from \( f_1, f_2, \ldots \), and condition 2 says that there is no redundancy in the list (no building block could have been built from the others).
Example 9.11. The functions $e^{2t}$, $e^{-3t}$ form a basis for the space of solutions to (8).

Flashcard question: What is a basis for the space of solutions to $\dot{y} = 3y$?

Possible answers:
1. This is not a homogeneous linear ODE, so the solutions don’t form a vector space. It’s a trick question.
2. The function $e^{3t}$ by itself is a basis.
3. The function $2e^{3t}$ by itself is a basis.
4. The basis is the set of all functions of the form $ce^{3t}$.

Answer: 2 or 3! First, answer 1 is wrong: each term is a (constant) function of $t$ times either $y$ or $\dot{y}$, so this is a homogeneous linear ODE, and the solutions do form a vector space. The basis is supposed to consist of linearly independent functions such that all the solutions can be built from them. Answer 4 is wrong since the functions in the basis are supposed to be linearly independent; if $e^{3t}$ is in a basis, then $5e^{3t}$ should not be in the same basis since it is a linear combination of $e^{3t}$ by itself. Answer 2 is correct since the solutions are exactly the functions $ce^{3t}$ for all numbers $c$. Answer 3 is correct too since the functions $c(2e^{3t})$ also run through all solutions as $c$ ranges over all numbers.

Key point: A vector space usually has infinitely many functions. To describe it compactly, give a basis of the vector space.

9.6. Dimension. It turns out that, although a vector space can have different bases, each basis has the same number of functions in it.

Definition 9.12. The dimension of a vector space is the number of functions in any basis.

Example 9.13. The space of solutions to (8) is 2-dimensional, since the basis $e^{2t}, e^{-3t}$ (or any other basis for the same vector space) has 2 elements.

Example 9.14. The space of solutions to $\dot{y} = 3y$ is 1-dimensional.

In these two examples, the dimension equals the order of the homogeneous linear ODE. It turns out that this holds in general:

Dimension theorem for a homogeneous linear ODE. The dimension of the space of solutions to an $n$th order homogeneous linear ODE is $n$.

In other words, the number of parameters needed in the general solution to an $n$th order homogeneous linear ODE is $n$. This is compatible with the existence and uniqueness theorem, which says that, given a $n$th order linear ODE, a particular solution can be singled out by giving $n$ numbers, namely the values of $y(a)$, $y'(a)$, $\ldots$, $y^{(n-1)}(a)$ (the initial conditions).
9.7. **Solving a homogeneous linear ODE with constant coefficients.** The solution method we used for \( \ddot{y} - \dot{y} + 6y = 0 \) generalizes as follows:

Given

\[ a_n y^{(n)} + \cdots + a_1 \dot{y} + a_0 y = 0, \tag{9} \]

where \( a_n, \ldots, a_0 \) are constants, do the following:

1. Write down the **characteristic equation**

   \[ a_n r^n + \cdots + a_1 r + a_0 = 0, \]

   in which the coefficient of \( r^i \) is the coefficient of \( y^{(i)} \) from the ODE. The left hand side is called the **characteristic polynomial** \( p(r) \). (For example, \( \ddot{y} + 5\dot{y} = 0 \) has characteristic polynomial \( r^2 + 5 \).)

2. Factor \( p(r) \) as

   \[ a_n (r - r_1)(r - r_2) \cdots (r - r_n) \]

   where \( r_1, \ldots, r_n \) are (possibly complex) numbers.

3. If \( r_1, \ldots, r_n \) are **distinct**, then the functions \( e^{r_1 t}, \ldots, e^{r_n t} \) form a basis for the space of solutions to the ODE \[ \text{[9]} \]. In other words, the general solution is

   \[ c_1 e^{r_1 t} + \cdots + c_n e^{r_n t}. \]

4. If \( r_1, \ldots, r_n \) are **not** distinct, then \( e^{r_1 t}, \ldots, e^{r_n t} \) cannot be a basis since some of these functions are redundant (definitely not linearly independent!) If a particular root \( r \) is repeated \( m \) times, then

   replace \( m \) copies

   \[ e^{rt}, e^{rt}, e^{rt}, \ldots, e^{rt} \]

   by \( e^{rt}, te^{rt}, t^2 e^{rt}, \ldots, t^{m-1} e^{rt} \).

   (We’ll explain later on why this works.) \( \square \)

In all cases,

\[ \# \text{ functions in basis} = \# \text{ roots of } p(r) \text{ counted with multiplicity} = \text{ order of ODE}, \]

as predicted by the dimension theorem.

**Problem 9.15.** Find the general solution to

\[ y^{(6)} + 6y^{(5)} + 9y^{(4)} = 0. \]

**Solution:** The characteristic polynomial is

\[ r^6 + 6r^5 + 9r^4 = r^4(r + 3)^2, \]
whose roots listed with multiplicity are

\[ 0, 0, 0, 0, -3, -3. \]

Since the roots are not distinct, the basis is not

\[ e^0t, e^0t, e^0t, e^0t, e^{-3t}, e^{-3t}. \]

We need to replace the first block of four functions, and also the last block of two functions. So the correct basis is

\[ e^0t, te^0t, t^2e^0t, t^3e^0t, e^{-3t}, te^{-3t}, \]

which simplifies to

\[ 1, t, t^2, t^3, e^{-3t}, te^{-3t}. \]

Thus the general solution is

\[ c_1 + c_2 t + c_3 t^2 + c_4 t^3 + c_5 e^{-3t} + c_6 t e^{-3t}. \]

(As expected, there is a 6-dimensional space of solutions to this 6th order ODE.) \( \Box \)

**Problem 9.16.** Find the simplest constant-coefficient homogeneous linear ODE having \((5t + 7)e^{-t} - 9e^{2t}\) as one of its solutions.

**Solution:** The given function is a linear combination of

\[ e^{-t}, te^{-t}, e^{2t} \]

so the roots of the characteristic polynomial (with multiplicity) should include \(-1, -1, 2\). So the simplest characteristic polynomial is

\[(r + 1)(r + 1)(r - 2) = r^3 - 3r - 2\]

and the corresponding ODE is

\[ y^{(3)} - 3\dot{y} - 2y = 0. \] \( \Box \)

Midterm 1 covers everything up to here.

### 9.8. **Complex roots.** The general method in the previous section works even if some of the roots are not real.

The simplest example is the ODE

\[ \dot{y} + y = 0. \]

The characteristic polynomial is \( r^2 + 1 \), which factors as \((r - i)(r + i)\). Its roots are \( i \) and \(-i\). Thus

\[ e^{it}, e^{-it} \text{ form a basis for the space of solutions}. \]
In other words, the general solution is

\[ c_1 e^{it} + c_2 e^{-it}, \quad \text{where } c_1, c_2 \in \mathbb{C}. \]

Substituting \( e^{it} = \cos t + i \sin t \) and \( e^{-it} = \cos t - i \sin t \) and regrouping leads to

\[ (c_1 + c_2) \cos t + (c_1 i - c_2 i) \sin t, \]

which is of the form

\[ d_1 \cos t + d_2 \sin t, \quad \text{where } d_1, d_2 \in \mathbb{C}. \]

This process is reversible, by substituting \( \cos t = \frac{e^{it} + e^{-it}}{2} \) and \( \sin t = \frac{e^{it} - e^{-it}}{2i} \). So linear combinations of \( e^{it}, e^{-it} \) are linear combinations of \( \cos t, \sin t \), and vice versa:

\[ \text{Span}(e^{it}, e^{-it}) = \text{Span}(\cos t, \sin t) \]

(assuming that complex numbers are allowed as coefficients in the linear combinations). Moreover, the functions \( \cos t, \sin t \) are linearly independent since neither is a scalar multiple of the other. Conclusion:

\[ \cos t, \sin t \text{ form another basis for the same space of solutions.} \]

**Question 9.17.** Which basis should be used?

**Answer:** It depends:

- The basis \( e^{it}, e^{-it} \) is easier to calculate with, but it’s not immediately obvious which linear combinations of these functions are real-valued.
- The basis \( \cos t, \sin t \) consisting of real-valued functions is useful for interpreting solutions in a physical system. The general real-valued solution is \( d_1 \cos t + d_2 \sin t \) where \( d_1, d_2 \) are real numbers.

So we will be converting back and forth.

**February 21**

Similar remarks apply to more complicated ODEs:

**Complex basis vs. real-valued basis.** Let \( y(t) \) be a complex-valued function of a real-valued variable \( t \). If \( y \) and \( \overline{y} \) are part of a basis of solutions to a homogeneous linear system of ODEs with real coefficients, then

\[ \text{replacing } y, \overline{y} \text{ by } \text{Re}(y), \text{Im}(y) \]

gives a new basis.

**Sketch of proof.** To know that the new list (including \( \text{Re}(y) \) and \( \text{Im}(y) \)) is a basis, we need to check two things:
1. The span of the new list is the set of all solutions.
   (True, because any solution is a linear combination of the old basis, and can be converted to a linear combination of the new list by substituting
   \[ y = \text{Re}(y) + i \text{Im}(y), \quad \overline{y} = \text{Re}(y) - i \text{Im}(y). \]

2. The new list is linearly independent.
   (If not, say
   \[ c_1 \text{Re}(y) + c_2 \text{Im}(y) + \cdots = 0, \]
   then substituting
   \[ \text{Re}(y) = \frac{y + \overline{y}}{2}, \quad \text{Im}(y) = \frac{y - \overline{y}}{2i} \]
   would show that the old basis was linearly dependent, which is impossible for a basis.) □

**Question 9.18.** Would it be OK to replace \( y, \overline{y} \) instead by \( \text{Re}(y), i \text{Im}(y) \)?

**Answer:** Yes, but this would be less useful, because the whole point is to obtain a basis consisting of real-valued functions.

**Question 9.19.** Would it be OK to replace \( y, \overline{y} \) instead by \( \text{Re}(y), \text{Re}(\overline{y}) \)?

**Answer:** No, because if \( y = f + ig \), then \( \overline{y} = f - ig \), so \( \text{Re}(y), \text{Re}(\overline{y}) \) are both \( f \)! They are linearly independent, so they can’t be part of a basis.

Here is an example of how this is used in practice:

**Problem 9.20.** Find a basis of solutions to
\[ y^{(3)} + 3\overline{y} + 9\dot{y} - 13y = 0 \]
consisting of real-valued functions.

**Solution:** The characteristic polynomial is \( p(r) := r^3 + 3r^2 + 9r - 13 \). Checking the divisors of \(-13\) (as instructed by the rational root test), we find that 1 is a root, so \( r - 1 \) is a factor. Long division (or solving for unknown coefficients) produces the other factor:
\[ p(r) = (r - 1)(r^2 + 4r + 13). \]
The quadratic formula, or completing the square (rewriting the second factor as \((r + 2)^2 + 9\)), shows that the roots of \( p(r) \) are 1, \(-2 + 3i\), \(-2 - 3i\). Thus \( e^t, e^{(-2+3i)t}, e^{(-2-3i)t} \) form a basis of solutions. But the last two are not real-valued!
So instead replace $y = e^{(-2+3i)t}$ and $\overline{y} = e^{(-2-3i)t}$ by $\text{Re}(y)$ and $\text{Im}(y)$, found by expanding
\[
e^{(-2+3i)t} = e^{-2t}e^{i(3t)}
= e^{-2t} (\cos(3t) + i\sin(3t))
= e^{-2t} \cos(3t) + i e^{-2t} \sin(3t).
\]
Thus
\[
\begin{bmatrix}
  e^t, & e^{-2t} \cos(3t), & e^{-2t} \sin(3t)
\end{bmatrix}
\]
is another basis, this time consisting of real-valued functions. □

9.9. **Harmonic oscillators and damped frequency.** Let’s apply all this to the spring-mass-dashpot system, assuming no external force.

9.9.1. **Undamped case.** If there is no damping, the DE is
\[
m\dddot{x} + kx = 0.
\]
Characteristic polynomial: $p(r) = mr^2 + k$.
Roots: $\pm \sqrt{-k/m} = \pm i\omega$, where $\omega := \sqrt{k/m}$.
Basis of solution space: $e^{i\omega t}, e^{-i\omega t}$.
Real-valued basis: $\cos \omega t, \sin \omega t$.
General real solution: $a \cos \omega t + b \sin \omega t$, where $a, b$ are real constants.

In other words, the real-valued solutions are all the sinusoid functions of angular frequency $\omega$. They could also be written as $A \cos(\omega t - \phi)$, where $A$ and $\phi$ are real constants.

This system, or any other system governed by the same DE, is also called a simple harmonic oscillator. The angular frequency $\omega$ is also called the natural frequency (or resonant frequency) of the oscillator.

9.9.2. **Damped case.** If there is damping, the DE is
\[
m\dddot{x} + b\dot{x} + kx = 0.
\]
Characteristic polynomial: $p(r) = mr^2 + br + k$.
Roots: $\frac{-b \pm \sqrt{b^2 - 4mk}}{2m}$ (by the quadratic formula)

There are three cases, depending on the sign of $b^2 - 4mk$.

**Case 1:** $b^2 < 4mk$ (underdamped).
Then there are two complex roots $-s \pm i\omega_d$, where

$$s := \frac{b}{2m} \quad \text{(positive)}$$

**Damped frequency**

$$\omega_d := \frac{\sqrt{4mk - b^2}}{2m} \quad \text{(positive)}$$

Basis of solution space: $e^{(-s+i\omega_d)t}$, $e^{(-s-i\omega_d)t}$

Real-valued basis: $e^{-st}\cos(\omega_d t)$, $e^{-st}\sin(\omega_d t)$.

General real solution: $e^{-st}(a\cos(\omega_d t) + b\sin(\omega_d t))$, where $a, b$ are real constants.

This is a sinusoid multiplied by a decaying exponential. It can also be written as $e^{-st}(A\cos(\omega_d t - \phi))$ for some $A$ and $\phi$. Each nonzero solution tends to 0, but changes sign infinitely many times along the way. The system is called **underdamped**, because there was not enough damping to eliminate the oscillation completely.

The damping not only causes the solution to decay exponentially, but also *changes the frequency of the sinusoid*. The new angular frequency, $\omega_d$, is called damped frequency. It is less than the undamped frequency we computed earlier (same formula, but with $b = 0$):

$$\omega_d = \frac{\sqrt{4mk - b^2}}{2m} < \sqrt{\frac{4mk}{2m}} = \sqrt{\frac{k}{m}} = \omega.$$

**Warning:** The damped solutions are not actually periodic (they don’t repeat exactly, because of the decay). Sometimes $2\pi/\omega_d$ is called the **pseudo-period**.

**Case 2:** $b^2 = 4mk$ (**critically damped**).

There there is a repeated real root: $-\frac{b}{2m}, -\frac{b}{2m}$. Call it $-s$.

Basis of solution space: $e^{-st}$, $te^{-st}$.

General real solution: $[e^{-st}(a + bt)]$, where $a, b$ are real constants.

What happens to the solutions as $t \to +\infty$? The solution $e^{-st}$ tends to 0. So does $te^{-st} = \frac{t}{e^{st}}$: even though the numerator $t$ is tending to $+\infty$, the denominator $e^{st}$ is tending to $+\infty$ faster (in a contest between exponentials and polynomials, exponentials always win). Thus all solutions eventually decay.

This case is when there is just enough damping to eliminate oscillation. The system is called **critically damped**.

**Case 3:** $b^2 > 4mk$ (**overdamped**).

In this case, the roots $\frac{-b \pm \sqrt{b^2 - 4mk}}{2m}$ are real and distinct. Both roots are negative, since $\sqrt{b^2 - 4mk} < b$. Call them $-s_1$ and $-s_2$.

General real solution: $[ae^{-s_1t} + be^{-s_2t}]$, where $a, b$ are real constants.
As in all the other damped cases, all solutions tend to 0 as \( t \to +\infty \). The term corresponding to the less negative root eventually controls the rate of return to equilibrium. The system is called \textit{overdamped}; there is so much damping that it is slowing the return to equilibrium.

Summary:

<table>
<thead>
<tr>
<th>Case</th>
<th>Roots</th>
<th>Situation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b = 0 )</td>
<td>two complex roots ( \pm i\omega )</td>
<td>undamped (simple harmonic oscillator)</td>
</tr>
<tr>
<td>( b^2 &lt; 4mk )</td>
<td>two complex roots ( -s \pm i\omega_d )</td>
<td>underdamped (damped oscillator)</td>
</tr>
<tr>
<td>( b^2 = 4mk )</td>
<td>repeated real root ( -s, -s )</td>
<td>critically damped</td>
</tr>
<tr>
<td>( b^2 &gt; 4mk )</td>
<td>distinct real roots ( -s_1, -s_2 )</td>
<td>overdamped</td>
</tr>
</tbody>
</table>

Flashcard question: Any nonzero solution \( x = e^{-t(a \cos(\sqrt{3}t) + b \sin(\sqrt{3}t))} \) crosses the equilibrium position \( x = 0 \) infinitely many times. How much time elapses between two consecutive crossings?

Possible answers:
1. \( \pi\sqrt{3} \)
2. \( \pi/\sqrt{3} \)
3. \( 2\pi\sqrt{3} \)
4. \( 2\pi/\sqrt{3} \)
5. \( \sqrt{3}/\pi \)
6. \( \sqrt{3}/(2\pi) \)
7. None of the above
8. Don’t know

Answer: \( \pi/\sqrt{3} \). Why? The solution has the same zeros as the sinusoid \( a \cos(\sqrt{3}t) + b \sin(\sqrt{3}t) \) of angular frequency \( \sqrt{3} \), period \( 2\pi/\sqrt{3} \). But a sinusoid crosses 0 twice within each period, so the answer is half a period, \( \pi/\sqrt{3} \).

Try the “Damped Vibrations” mathlet 

9.10. **Operator notation.**

9.10.1. The operator \( D \).

- A function (e.g., \( f(t) = t^2 \)) takes an input number and returns another number.
- An operator takes an input function and returns another function.

For example, the differential operator \( \frac{d}{dt} \) takes an input function \( y(t) \) and returns \( \frac{dy}{dt} \). This operator is also called \( D \). So \( De^{3t} = 3e^{3t} \), for instance (chain rule). The operator \( D \) is linear, which means that

\[
D(f + g) = Df + Dg, \quad D(af) = aDf
\]

for any functions \( f \) and \( g \), and any number \( a \). Because of this, \( D \) respects linear combinations, meaning that

\[
D(c_1f_1 + \cdots + c_nf_n) = c_1Df_1 + \cdots + c_nDf_n
\]

for any numbers \( c_1, \ldots, c_n \) and functions \( f_1, \ldots, f_n \).

9.10.2. Multiplying and adding operators. To apply a product of two operators, apply each operator in succession. For instance, \( DDy \) means take the derivative of \( y \), and then take the derivative of the result; therefore we write \( D^2y = \ddot{y} \).

To apply a sum of two operators, apply each operator to the function and add the results. For instance, \( (D^2 + D)y = D^2y + Dy = \ddot{y} + \dot{y} \).

Any number can be viewed as the “multiply-by-the-number” operator: for instance, the operator 5 transforms the function \( \sin t \) into the function \( 5 \sin t \). **Warning:** In the MITx reading, this operator is denoted \( 5I \) instead of 5.

**Example 9.22.** The ODE

\[
2\ddot{y} + 3\dot{y} + 5y = 0,
\]

whose characteristic polynomial is \( p(r) := 2r^2 + 3r + 5 \), can be rewritten as

\[
(2D^2 + 3D + 5)y = 0
\]

\[
p(D)y = 0.
\]
The same argument shows that every constant-coefficient homogeneous linear ODE

\[ a_n y^{(n)} + \cdots + a_0 y = 0 \]

can be written simply as

\[ p(D)y = 0, \]

where \( p \) is the characteristic polynomial.

9.10.3. **Shortcut for applying an operator to an exponential function.**

**Warm-up problem:** If \( r \) is a number, what is \((2D^2 + 3D + 5)e^{rt}\)?

**Solution:** First, \( De^{rt} = re^{rt} \) and \( D^2 e^{rt} = r^2 e^{rt} \) (keep applying the chain rule). Thus

\[
(2D^2 + 3D + 5)e^{rt} = 2r^2 e^{rt} + 3re^{rt} + 5e^{rt} \\
= (2r^2 + 3r + 5)e^{rt}. \quad \square
\]

The same calculation, but with an arbitrary polynomial, proves the general rule:

**Theorem 9.23.** For any polynomial \( p \) and any number \( r \),

\[
[ p(D)e^{rt} = p(r)e^{rt}].
\]

9.10.4. **Basis of solutions for a constant-coefficient homogeneous linear ODE.** Probably it would be better to skip much of this.

Remember the method for finding a basis of solutions to a constant-coefficient homogeneous linear ODE, now written as \( p(D)y = 0 \)? Using operators, we can explain why it works.

**Example 9.24.** Consider the ODE

\[ p(D)y = 0, \]

where \( p(r) := (r - 2)(r - 3)(r - 5) \) (so \( p(r) \) is the characteristic polynomial of the ODE). The order is 3, so the dimension of the vector space of solutions is 3.

Now

- \( e^{2t} \) is a solution since \( p(D)e^{2t} = p(2)e^{2t} = 0e^{2t} = 0 \),
- \( e^{3t} \) is a solution since \( p(D)e^{3t} = p(3)e^{3t} = 0e^{3t} = 0 \), and
- \( e^{5t} \) is a solution since \( p(D)e^{5t} = p(5)e^{5t} = 0e^{5t} = 0 \).

Just because we wrote down 3 solutions does not mean that they form a basis: If we had written down \( e^{2t}, e^{3t}, 4e^{2t} + 6e^{3t} \), then they would not have been a basis, because they are linearly dependent (and their span is only 2-dimensional).

To know that \( e^{2t}, e^{3t}, e^{5t} \) really form a basis, we need to know that they are linearly independent. Could it instead be that

\[
e^{5t} = c_1 e^{2t} + c_2 e^{3t} \quad \text{as functions}\]

(as functions)
for some numbers $c_1, c_2$? If so, then applying $(D - 2)(D - 3)$ to both sides would give

$$ (5 - 2)(5 - 3)e^{5t} = 0 + 0, $$

which is ridiculous. So $e^{5t}$ is not a linear combination of $e^{2t}$ and $e^{3t}$. Similarly, no one of $e^{2t}$, $e^{3t}$, $e^{5t}$ is a linear combination of the other two. So $e^{2t}$, $e^{3t}$, $e^{5t}$ are linearly independent, so they form a basis for a 3-dimensional space (which must be the space of all solutions, since that too is 3-dimensional). □

What about the case of repeated roots?

**Example 9.25.** Find a basis of solutions to $D^3y = 0$.

(The characteristic polynomial is $r^3$, whose roots with multiplicity are 0, 0, 0.)

**Solution:** Integrate three times:

\[
\begin{align*}
D^2y &= c_1 \\
Dy &= c_1 t + c_2 \\
y &= c_1 \frac{t^2}{2} + c_2 t + c_3 \\
&= C_1 t^2 + c_2 t + c_3. 
\end{align*}
\]

for some numbers $C_1 = c_1/2$, $c_2$, and $c_3$. Since $t^2$, $t$, 1 are linearly independent, they form a basis for the space of solutions. □

**Example 9.26.** Find a basis of solutions to $(D - 5)^3y = 0$.

The characteristic polynomial is $(r - 5)^3$, which has a repeated root 5, 5, 5, so I told you earlier that $e^{5t}$, $te^{5t}$, $t^2e^{5t}$ should be a basis. Let’s now explain why these are solutions.

We know that $D - 5$ sends $e^{5t}$ to 0. What does $D - 5$ do to $ue^{5t}$, if $u$ is a function of $t$? Let’s find out:

\[
(D - 5)ue^{5t} = \left(\frac{d}{dt}e^{5t} + u(5e^{5t})\right) - 5ue^{5t} \\
= \frac{du}{dt}e^{5t}.
\]

Similarly,

\[
\begin{align*}
(D - 5)^2ue^{5t} &= \frac{d}{dt}ue^{5t} \\
(D - 5)^3ue^{5t} &= \frac{d^2}{dt^2}ue^{5t}.
\end{align*}
\]
In order for $ue^{5t}$ to be a solution to $(D−5)^2y = 0$ the function $\ddot{u}$ must be 0; i.e., $u = a+bt+ct^2$ for some numbers $a, b, c$, so the solutions are

$$ue^{5t} = ae^{5t} + bte^{5t} + ct^2e^{5t},$$

as expected. □

Since 1, $t, t^2$ are linearly independent, so are (any relation between the last three functions could be divided by $e^{5t}$ to get a relation between the first three). Thus $e^{5t}, te^{5t}, t^2e^{5t}$ form a basis. □.

A similar approach handles more complicated characteristic polynomials involving many repeated roots.

(You don’t have to go through the discussion of this section each time you want to solve $p(D)y = 0$; this section is just a theoretical explanation for why the method given earlier works.)

9.11. **Exponential response.** For any polynomial $p$ and number $r$,

$$p(D)e^{rt} = p(r)e^{rt},$$

so

$$e^{rt} \text{ output signal} \text{ is a particular solution to } p(D)y = p(r)e^{rt} \text{ input signal}.$$  

\[ \text{New problem: } \text{What if the input signal is just } e^{rt}? \]

\[ \text{Answer (superposition): Multiply by } \frac{1}{p(r)} \text{ to get...} \]

**Exponential response formula (ERF).**

For any polynomial $p$ and any number $r$ such that $p(r) \neq 0$,

$$\frac{1}{p(r)}e^{rt} \text{ output signal} \text{ is a particular solution to } p(D)y = e^{rt} \text{ input signal}.$$

In other words, multiply the input signal by the number $\frac{1}{p(r)}$ to get an output signal.

**Problem 9.27.** Find the general solution to $\ddot{y} + 7\dot{y} + 12y = -5e^{2t}$.

**Solution:**

Characteristic polynomial: $p(r) = r^2 + 7r + 12 = (r + 3)(r + 4)$.

Roots: $−3, −4$.

General solution to homogeneous equation: $y_h := c_1e^{-3t} + c_2e^{-4t}$.

ERF says:

$$\frac{1}{p(2)}e^{2t} \text{ is a particular solution to } p(D)y = e^{2t};$$
i.e.,
\[ \frac{1}{30} e^{2t} \] is a particular solution to \[ \ddot{y} + 7\dot{y} + 12y = e^{2t}, \]
so
\[ -\frac{1}{6} e^{2t} \] is a particular solution to \[ \ddot{y} + 7\dot{y} + 12y = -5e^{2t}. \]

General solution to inhomogeneous equation:
\[ y = y_p + y_h \]
\[ = -\frac{1}{6} e^{2t} + c_1 e^{-3t} + c_2 e^{-4t}. \]

February 28

The existence and uniqueness theorem says that
\[ p(D)y = e^{rt} \]
should have a solution even if \( p(r) = 0 \) (when ERF does not apply). Here is how to find a particular solution in this bad case:

**Generalized exponential response formula.**

*If \( p \) is a polynomial having \( r \) as a root with multiplicity \( s \), then
\[
z_p = \frac{1}{p^{(s)}(r)} t^s e^{rt} \quad \text{is a particular solution to} \quad p(D)z = \frac{e^{rt}}{\text{input signal}}.
\]*

In other words, multiply the input signal by \( t^s \), and then multiply by the number \( \frac{1}{p^{(s)}(r)} \), where \( p^{(s)} \) is the \( s \)th derivative of \( p \).

(The proof that this works uses rules for operators as in Supplementary Notes, Section O. We’ll skip it.)

Generalized ERF comes up less often than regular ERF, since in most applications, \( p(r) \neq 0 \).

9.12. **Sinusoidal response (complex replacement).** The **complex replacement** method (for an input function \( \cos \omega t \)) is for finding one particular solution to an inhomogeneous linear ODE
\[ p(D)x = \cos \omega t, \]
where \( p \) is a real polynomial, and \( \omega \) is a real number.

1. Replace \( \cos \omega t \) by a complex exponential \( e^{i\omega t} \) whose real part is \( \cos \omega t \) (and use a different letter \( z \) for the unknown function). This produces the **complex replacement** ODE
\[ p(D)z = e^{i\omega t}. \]
2. Find a particular solution $z_p$ to the complex replacement ODE. (Use ERF for this, provided that $p(i\omega) \neq 0$.)

3. Compute $x_p := \text{Re}(z_p)$. Then $x_p$ is a particular solution to the original ODE.

**Problem 9.28.** Find a particular solution $x_p$ to

$$\ddot{x} + \dot{x} + 2x = \cos(2t).$$

**Solution:** The characteristic polynomial is $p(r) := r^2 + r + 2$.

**Step 1.** Since $\cos(2t)$ is the real part of $e^{2it}$, replace $\cos(2t)$ by $e^{2it}$.

$$\ddot{z} + \dot{z} + 2z = e^{2it}.$$

**Step 2.** ERF says that one particular solution to this new ODE is

$$z_p := \frac{1}{p(2i)} e^{2it} = \frac{1}{-2 + 2i} e^{2it}.$$

**Step 3.** A particular solution to the original ODE is

$$x_p := \text{Re}(z_p) = \text{Re} \left( \frac{1}{-2 + 2i} e^{2it} \right).$$

This is a sinusoid expressed in complex form.

It might be more useful to have the answer in amplitude-phase form or as a linear combination of cos and sin, but we are given $x_p = \text{Re}(ce^{2it})$ with $c := \frac{1}{-2 + 2i}$. To convert, we need to rewrite $c$ as $Ae^{i\phi}$ or $a + bi$, using

$$c = \frac{1}{-2 + 2i} = \frac{1}{2\sqrt{2}} e^{i(-3\pi/4)} = \frac{1}{2\sqrt{2}} e^{i(3\pi/4)},$$

which is supposed to be $Ae^{i\phi}$. Thus the amplitude is $A = \frac{1}{2\sqrt{2}} = \frac{\sqrt{2}}{4}$ and the phase lag is $\phi = 3\pi/4$. **Conclusion:** In amplitude-phase form,

$$x_p = \frac{\sqrt{2}}{4} \cos(2t - 3\pi/4).$$
Converting to a linear combination of \( \cos \) and \( \sin \). We have
\[
c = \frac{1}{-2 + 2i} = \frac{1}{-2 + 2i} \left( \frac{-2 - 2i}{-2 - 2i} \right) = \frac{-2 - 2i}{8} = -\frac{1}{4} - \frac{1}{4}i
\]
\[
\bar{c} = -\frac{1}{4} + \frac{1}{4}i,
\]
which is supposed to be \( a + bi \), so \( a = -1/4 \) and \( b = 1/4 \). Conclusion:
\[
\boxed{x_p = -\frac{1}{4} \cos(2t) + \frac{1}{4} \sin(2t)}
\]

Try the “Amplitude and Phase: Second Order IV” mathlet
\[
\text{http://mathlets.org/mathlets/amplitude-and-phase-second-order-iv/}
\]
with \( m = 1, b = 1, k = 2, \omega = 2 \) to see the input signal \( \cos(2t) \), and output signal (in yellow). Can you see the amplitude and phase lag of the output signal? The red segment indicates the time lag \( t_0 = \phi/\omega = (3\pi/4)/2 = 3\pi/8 \approx 1.18 \).

9.13. **Complex gain, gain, and phase lag for an ODE.** Here is what happens in general when solving
\[
p(D)x = \cos \omega t,
\]
assuming \( p(i\omega) \neq 0 \):
- The complex replacement ODE
\[
p(D)z = e^{i\omega t}
\]
has input signal \( e^{i\omega t} \) and output signal \( \frac{1}{p(i\omega)} e^{i\omega t} \) by ERF and superposition, so
\[
\text{complex gain } G = \frac{1}{p(i\omega)}
\]
(a complex number).
- The original ODE
\[
p(D)x = \cos \omega t
\]
has sinusoid output signal \( x_p := \text{Re}(G e^{i\omega t}) \).
- The angular frequency of the output signal is the same as the angular frequency of the input signal: \( \omega \).
- For this system,
\[
\text{gain } = |G| = \frac{1}{|p(i\omega)|}
\]
and
\[
\text{phase lag } = - \arg G = \arg p(i\omega)
\]
Question 9.29. Why does the complex replacement method work?

Answer: If \( z_p \) is a solution to the complex replacement ODE, i.e.,

\[
p(D)z_p = e^{i\omega t},
\]

then taking real parts of both sides (while remembering that \( p \) has real coefficients) gives

\[
p(D)x_p = \cos \omega t,
\]

which says that \( x_p \) is a solution to the original ODE. \( \square \)

Complex replacement is helpful also with other real input signals, with any real-valued function that can be written as the real part of a reasonably simple complex input signal. Here are some examples:

<table>
<thead>
<tr>
<th>Real input signal</th>
<th>Complex replacement</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \cos \omega t )</td>
<td>( e^{i\omega t} )</td>
</tr>
<tr>
<td>( A \cos(\omega t - \phi) )</td>
<td>( Ae^{-i\phi}e^{i\omega t} )</td>
</tr>
<tr>
<td>( a \cos \omega t + b \sin \omega t )</td>
<td>( (a - bi)e^{i\omega t} )</td>
</tr>
<tr>
<td>( e^{at} \cos \omega t )</td>
<td>( e^{(a+i\omega)t} )</td>
</tr>
</tbody>
</table>

Each function in the first column is the real part of the corresponding function in the second column. The nice thing about these examples is that the complex replacement is a constant times a complex exponential, so ERF (or generalized ERF) applies.

Remark 9.30. The second and third input signals in the table are different forms of the general sinusoid function of angular frequency \( \omega \). For these, the complex replacement input is just a scalar multiple of \( e^{i\omega t} \). Therefore the output signal for the complex replacement ODE will always be \( \frac{1}{p(i\omega)} \) times the input signal, by ERF (provided that \( p(i\omega) \neq 0 \)). This means that:

- The complex gain depends only on the system and the input angular frequency (that is, on \( p \) and \( \omega \)), not on the specific sinusoid used as input.
- The gain and phase lag depend only on the system and the input angular frequency. (This is because gain and phase lag are determined by complex gain.)
9.14. **Time invariance.** Here is another point of view on the effect of changing the input sinusoid.

Since the coefficients of $p$ are constants not depending on $t$, the linear operator $p(D)$ is called **time-invariant**. This means that if an input signal $F(t)$ is delayed in time by $a$ seconds, then the output signal is delayed by $a$ seconds. Mathematically: If $f(t)$ is a solution to $p(D)x = F(t)$ and $a$ is a number, then $f(t-a)$ is a solution to $p(D)x = F(t-a)$.

**Example 9.31.** If $p(D)x = \cos \omega t$ has $x_p = A \cos(\omega t - \phi)$ as a particular solution, then shifting time by $a = \alpha / \omega$ shows that $p(D)x = \cos(\omega t - \alpha)$ has $x_p = A \cos(\omega t - \alpha - \phi)$ as a particular solution. The gain and (relative) phase lag of the system are still $A$ and $\phi$.

9.15. **Stability.**

9.15.1. **Steady-state solution, transient.**

**Problem 9.32.** What is the general solution to $\ddot{x} + 7\dot{x} + 12x = \cos(2t)$?

**Solution:** The characteristic polynomial is $p(r) = r^2 + 7r + 12 = (r + 3)(r + 4)$.

The complex gain is

$$G = \frac{1}{p(2i)} = \frac{1}{(2i)^2 + 7(2i) + 12} = \frac{1}{8 + 14i}.$$ 

Complex replacement and ERF show that

$$x_p = \operatorname{Re}\left(\frac{1}{8 + 14i}e^{2it}\right)$$

is a particular solution.

On the other hand, the general solution to the associated homogeneous ODE is $x_h = c_1 e^{-3t} + c_2 e^{-4t}$.

Therefore the general solution to the original inhomogeneous ODE is

$$x = x_p + x_h = \operatorname{Re}\left(\frac{1}{8 + 14i}e^{2it}\right) + c_1 e^{-3t} + c_2 e^{-4t}. \quad \square$$
In general, for a forced damped oscillator, complex replacement and ERF will produce a periodic output signal, and that particular solution is called the steady-state solution. Every other solution is the steady-state solution plus a transient, where the transient is a function that decays to 0 as \( t \to +\infty \).

Changing the initial conditions changes only the \( c_1, c_2 \) above, so the steady-state solution is the same. A system like this, in which changes in the initial conditions have vanishing effect on the long-term behavior of the solution, is called stable.

Try the “Forced Damped Vibration” mathlet

http://mathlets.org/mathlets/forced-damped-vibration/

Notice that changing the initial conditions (dragging the yellow square on the left) does not change the long-term behavior of the output signal (yellow curve) much. So this is a stable system.

9.15.2. Testing a second-order system for stability in terms of roots. Stability depends on the shape of the solution to the associated homogeneous solution. In the problem above, this was \( c_1e^{-3t} + c_2e^{-4t} \), which decays as \( t \to +\infty \) no matter what \( c_1 \) and \( c_2 \) are, because \(-3\) and \(-4\) are negative. For a general 2nd-order constant-coefficient linear ODE, stability depends on the roots of the characteristic polynomial, as shown in the following table:

<table>
<thead>
<tr>
<th>Roots</th>
<th>General solution ( x_h )</th>
<th>Condition for stability</th>
<th>Characteristic poly.</th>
</tr>
</thead>
<tbody>
<tr>
<td>complex ( a \pm bi )</td>
<td>( e^{at}(c_1 \cos(bt) + c_2 \sin(bt)) )</td>
<td>( a &lt; 0 )</td>
<td>( r^2 - 2ar + (a^2 + b^2) )</td>
</tr>
<tr>
<td>repeated real ( s, s )</td>
<td>( e^{at}(c_1 + c_2 t) )</td>
<td>( s &lt; 0 )</td>
<td>( r^2 - 2sr + s^2 )</td>
</tr>
<tr>
<td>distinct real ( r_1, r_2 )</td>
<td>( c_1 e^{r_1 t} + c_2 e^{r_2 t} )</td>
<td>( r_1, r_2 &lt; 0 )</td>
<td>( r^2 - (r_1 + r_2)r + r_1r_2 )</td>
</tr>
</tbody>
</table>

More generally:

**Theorem 9.33** (Stability test in terms of roots). A constant-coefficient linear ODE of any order is stable if and only if every root of the characteristic polynomial has negative real part.

9.15.3. Testing a second-order system for stability in terms of coefficients. (To be covered in recitation)

In the 2nd order case, there is also a simple test directly in terms of the coefficients:

**Theorem 9.34** (Stability test in terms of coefficients, 2nd order case). Assume that \( a_0, a_1, a_2 \) are real numbers with \( a_0 > 0 \). The ODE

\[
(a_0D^2 + a_1D + a_2)x = F(t)
\]

is stable if and only if \( a_1 > 0 \) and \( a_2 > 0 \).

**Proof.** By dividing by \( a_0 \), we can assume that \( a_0 = 1 \). Break into cases according to the table above.

54
• When the roots are \( a \pm bi \), we have \( a < 0 \) if and only if the coefficients \(-2a\) and \( a^2 + b^2 \) are both positive.

• When the roots are \( s, s \), we have \( s < 0 \) if and only if the coefficients \(-2s\) and \( s^2 \) are both positive.

• When the roots are \( r_1, r_2 \), we have \( r_1, r_2 < 0 \) if and only if the coefficients \(-(r_1 + r_2)\) and \( r_1 r_2 \) are both positive. (Knowing that \(-(r_1 + r_2)\) is positive means that at least one of \( r_1, r_2 \) is negative; if moreover the product \( r_1 r_2 \) is positive, then the other root must be negative too.) □

There is a generalization of the coefficient test to higher-order ODEs, called the Routh–Hurwitz conditions for stability, but the conditions are much more complicated: see the MITx reading.

9.16. **Resonance.** Recall that a harmonic oscillator has a natural frequency. **Resonance** is a phenomenon that occurs when a harmonic oscillator is driven with an input sinusoid whose frequency is close to the natural frequency:

- the gain becomes larger and larger as the input frequency approaches the natural frequency, and
- when the input frequency equals the natural frequency, there is not any bounded solution at all.

We now explain all of this by solving ODEs explicitly.

9.16.1. **Warm-up: harmonic oscillator with no input signal.** A typical ODE modeling a harmonic oscillator is

\[
\ddot{x} + 9x = 0.
\]

Characteristic polynomial: \( r^2 + 9 \).
Roots: \( \pm 3i \).
Basis of solutions: \( e^{3it}, e^{-3it} \).
Real-valued basis: \( \cos 3t, \sin 3t \).
General real-valued solution: \( a \cos 3t + b \sin 3t \), for real numbers \( a, b \). These are all the sinusoids with angular frequency 3, the natural frequency.

9.16.2. **Near resonance.** Now let’s drive the harmonic oscillator with an input sinusoid. A typical ODE modeling this situation is

\[
\ddot{x} + 9x = \cos \omega t.
\]

The complex replacement ODE is

\[
\ddot{z} + 9z = e^{i\omega t}.
\]
Characteristic polynomial: \( p(r) = r^2 + 9. \)

Assume that \( \omega \neq 3, \) so that \( i\omega \) is not a root of \( p(r). \) Then ERF gives the particular solution
\[
 z_p := \frac{1}{p(i\omega)} e^{i\omega t} = \frac{1}{9 - \omega^2} e^{i\omega t}.
\]

Then a particular solution to the original ODE is
\[
 x_p := \frac{1}{9 - \omega^2} \cos \omega t.
\]

Complex gain: \( G = \frac{1}{p(i\omega)} = \frac{1}{9 - \omega^2}. \)

Gain: \( |G| = \frac{1}{|9 - \omega^2|}. \) This becomes very large as \( \omega \) approaches 3.

Phase lag: \( -\arg G, \) which is 0 or \( \pi \) depending on whether \( \omega < 3 \) or \( \omega > 3. \)

Try the “Harmonic Frequency Response: Variable Input Frequency” mathlet


to see this. (In this mathlet, the natural frequency is 1, and the frequency of the input signal is adjustable. RMS stands for root mean square, which for a sinusoid is amplitude/\( \sqrt{2}. \))

In engineering, the graph of gain as a function of \( \omega \) is called a Bode plot (Bode is pronounced Boh-dee). (Actually, engineers usually instead use a log-log plot: they plot \( \log(\text{gain}) \) as a function of \( \log \omega \).) On the other hand, a Nyquist plot shows the trajectory of the complex gain \( G \) as \( \omega \) varies.

Also try the “Harmonic Frequency Response: Variable Natural Frequency” mathlet


(In this one, the input signal is fixed to be \( \sin t, \) and the natural frequency is adjustable.)

9.16.3. Pure resonance.

**Question 9.35.** What happens if \( \omega = 3 \) exactly?

This time, the complex replacement ODE
\[
 \ddot{z} + 9z = e^{3it}
\]
cannot be solved by ERF, since \( 3i \) is a root of \( p(r) = r^2 + 9. \) This one requires generalized ERF. First, \( p(r) \) has distinct roots \( 3i \) and \( -3i, \) so \( s = 1, \) and \( p^{(s)}(r) = p'(r) = 2r = 6i \) at
$r = 3i$. Generalized ERF gives

$$z_p := \frac{1}{6i} t e^{3it}$$

$$= -\frac{i}{6} t (\cos(3t) + i \sin(3t))$$

$$= \frac{1}{6} t (-i \cos(3t) + \sin(3t)),$$

so

$$x_p := \frac{1}{6} t \sin(3t)$$

is a particular solution to the original ODE. This is not a sinusoid, but an oscillating function whose oscillations grow without bound as time progresses.

9.16.4. Resonance with damping. In a realistic physical situation, there is at least a tiny amount of damping, and this prevents the runaway growth of the previous section.

**Question 9.36.** What happens if $\omega = 3$ exactly, but there is a tiny amount of damping, so that the ODE is

$$\ddot{x} + b\dot{x} + 9x = \cos \omega t$$

for some small positive constant $b$?

New characteristic polynomial: $p(r) = r^2 + br + 9$. Since $3i$ is no longer a root, ERF applies.

Complex gain: $G = \frac{1}{p(3i)} = \frac{1}{3bi}$.

Gain: $|G| = \frac{1}{3b}$. This is large, but the oscillations are bounded; there is a steady-state solution.

Try the “Amplitude and Phase: First Order” mathlet

http://mathlets.org/mathlets/amplitude-and-phase-1st-order/

Try the “Amplitude and Phase: Second Order I” mathlet

http://mathlets.org/mathlets/amplitude-and-phase-2nd-order/

Try the “Amplitude and Phase: Second Order II” mathlet


Try the “Amplitude and Phase: Second Order III” mathlet

9.17. **RLC circuits.** Let’s model a circuit with a voltage source, resistor, inductor, and capacitor attached in series: an RLC circuit.

Variables and functions (with units):

- $t$: time (s)
- $R$: resistance of the resistor (ohms)
- $L$: inductance of the inductor (henries)
- $C$: capacitance of the capacitor (farads)
- $Q$: charge on the capacitor (coulombs)
- $I$: current (amperes)
- $V$: voltage source (volts)
- $V_R$: voltage drop across the resistor (volts)
- $V_L$: voltage drop across the inductor (volts)
- $V_C$: voltage drop across the capacitor (volts).

The independent variable is $t$. The quantities $R$, $L$, $C$ are constants. Everything else is a function of $t$.

**Equations:** Physics says

\[
I = \dot{Q} \\
V_R = RI \quad \text{Ohm's law} \\
V_L = L\dot{I} \quad \text{Faraday's law} \\
V_C = \frac{1}{C}Q \\
V = V_R + V_L + V_C \quad \text{Kirchoff's voltage law.}
\]
The last equation can be written as follows:

\[ V_L + V_R + V_C = V \]
\[ L\ddot{Q} + R\dot{Q} + \frac{1}{C}Q = V(t), \]

a second-order inhomogeneous linear ODE with unknown function \( Q(t) \). Mathematically, this is equivalent to the spring-mass-dashpot ODE

\[ m\ddot{x} + b\dot{x} + kx = F_{\text{external}}(t), \]

with the following table of analogies:

<table>
<thead>
<tr>
<th>Spring-mass-dashpot system</th>
<th>RLC circuit</th>
</tr>
</thead>
<tbody>
<tr>
<td>displacement ( x )</td>
<td>( Q ) charge</td>
</tr>
<tr>
<td>velocity ( \dot{x} )</td>
<td>( I ) current</td>
</tr>
<tr>
<td>mass ( m )</td>
<td>( L ) inductance</td>
</tr>
<tr>
<td>damping constant ( b )</td>
<td>( R ) resistance</td>
</tr>
<tr>
<td>spring constant ( k )</td>
<td>( 1/C ) 1/capacitance</td>
</tr>
<tr>
<td>external force ( F_{\text{external}}(t) )</td>
<td>( V(t) ) voltage source</td>
</tr>
</tbody>
</table>

Similarly, a harmonic oscillator (undamped) is analogous to an LC circuit (no resistor).

Try the “Series RLC Circuit” mathlet


9.18.1. Conservation of energy in the harmonic oscillator. Consider the harmonic oscillator described by \( m\ddot{x} + kx = 0 \). Let’s check conservation of energy.

Kinetic energy: \( KE = \frac{m\dot{x}^2}{2} \).

Potential energy \( PE \) is a function of \( x \), and

\[
\text{PE}(x) - \text{PE}(0) = -\int_0^x F_{\text{spring}}(X) dX = -\int_0^x -kX dX = \frac{kx^2}{2}.
\]

If we declare \( PE = 0 \) at position 0, then \( \text{PE}(x) = \frac{kx^2}{2} \).

Total energy is \( KE + PE \):

\[
E = \frac{m\dot{x}^2}{2} + \frac{kx^2}{2}. \tag{10}
\]

How does total energy change with time?

\[
\dot{E} = m\dddot{x} + k\ddot{x} = \ddot{x}(m\dddot{x} + k\ddot{x}) = 0.
\]

So energy is conserved.
9.18.2. *Phase plane.* Three ways to depict the harmonic oscillator:

- Movie showing the motion of the mass directly

- Graph of $x$ as a function of $t$

- Trajectory of $\langle x(t), \dot{x}(t) \rangle$ (a parametrized curve) in the phase plane, whose horizontal axis shows $x$ and whose vertical axis shows $\dot{x}$

Let’s start with the mass to the right of equilibrium, and then let go. At $t = 0$, we have $x > 0$ and $\dot{x} = 0$. At the first time the mass crosses equilibrium, $x = 0$ and $\dot{x} < 0$. When the mass reaches its leftmost point, $x < 0$ and $\dot{x} = 0$ again. These give three points on the phase plane trajectory.

Here are two ways to see that the whole trajectory is an ellipse:

1. We have $x = A \cos \omega t$ for some $A$ and $\omega$. Thus $\dot{x} = -A \omega \sin \omega t$. The parametrized curve

   $\langle A \cos \omega t, -A \omega \sin \omega t \rangle$
is an ellipse. (This is like the parametrization of a circle, but with axes stretched by different amounts.)

2. Rearrange (10) as

\[ \frac{x^2}{2E/k} + \frac{\dot{x}^2}{2E/m} = 1. \]

This is an ellipse with semi-axes \( \sqrt{2E/k} \) and \( \sqrt{2E/m} \).

**Flashcard question:** In which direction is the ellipse traversed?

Possible answers:
1. Clockwise.
2. Counterclockwise.
3. It depends on the initial conditions.

**Answer:** Clockwise. Above the horizontal axis, \( \dot{x} > 0 \), which means that \( x \) is increasing.

Changing the initial conditions changes \( E \), which changes the ellipse. The family of all such trajectories is a nested family of ellipses, called the phase diagram or phase portrait of the system.

9.18.3. **Energy loss in the damped oscillator.** Now consider a damped oscillator described by \( m\ddot{x} + b\dot{x} + kx = 0 \). Now

\[ \dot{E} = \dot{x}(m\ddot{x} + kx) = -b\dot{x}^2. \]

Energy is lost to friction. The dashpot heats up. The phase plane trajectory crosses through the equal-energy ellipses, inwards towards the origin.

It may spiral in (underdamped case), or approach the origin more directly (critically damped and overdamped cases).
10. Geometric interpretation of linear algebra concepts

For a while, we are going to assume that vectors have real numbers as coordinates, and that all scalars are real numbers. This is so we can describe things geometrically in \( \mathbb{R}^n \) more easily. But eventually, we will work with vectors in \( \mathbb{C}^n \) whose coordinates can be complex numbers, and will allow scalar multiplication by complex numbers.

10.1. Vector space. Up to now, we considered vector spaces whose elements were functions. There are also vector spaces whose elements are vectors!

**Definition 10.1.** Suppose that \( S \) is a set consisting of some of the vectors in \( \mathbb{R}^n \) (for some fixed value of \( n \)). Then \( S \) is a vector space if all of the following are true:

0. The zero vector \( \mathbf{0} \) is in \( S \).
1. Multiplying any one vector in \( S \) by a scalar gives another vector in \( S \).
2. Adding any two vectors in \( S \) gives another vector in \( S \).

**Example 10.2.** Such a set \( S \) is also called a subspace of \( \mathbb{R}^n \), because \( \mathbb{R}^n \) itself is a vector space, and \( S \) is a vector space contained in it.

**Flashcard question:** Which of the following subsets of \( \mathbb{R}^2 \) are vector spaces?

(a) The set of all vectors \((x, y)\) satisfying \( x^2 + y^2 = 1 \).
(b) The set of all vectors \((x, y)\) satisfying \( xy = 0 \).
(c) The set of all vectors \((x, y)\) satisfying \( 2x + 3y = 0 \).
**Answer:** Only (c) is a vector space.

**Explanation:** Let $S$ be the set. For $S$ to be a vector space, it must satisfy all three conditions.

Example (a) doesn’t even satisfy the first condition, because the zero vector $(0, 0)$ is not in $S$.

Example (b) satisfies the first condition: the zero vector is in $S$. It satisfies the second condition too: If $(x, y)$ is one vector in $S$ (so $xy = 0$) and $c$ is any scalar, then the vector $c(x, y) = (cx, cy)$ satisfies $(cx)(cy) = c^2xy = c^2(0) = 0$. But it does not satisfy the third condition for every pair of vectors in $S$: for example, $(2, 0)$ and $(0, 3)$ are in $S$, but their sum $(2, 3)$ is not in $S$.

Example (c) is a vector space, as we will now check. First, the zero vector is in $S$. Second, if $(x, y)$ is any element of $S$ (so $2x + 3y = 0$) and $c$ is any scalar, then multiplying the equation by $c$ gives $2(cx) + 3(cy) = 0$, which shows that the vector $c(x, y) = (cx, cy)$ is in $S$. Third, if $(x_1, y_1)$ and $(x_2, y_2)$ are in $S$ (so $2x_1 + 3y_1 = 0$ and $2x_2 + 3y_2 = 0$), then adding the equations shows that $2(x_1 + x_2) + 3(y_1 + y_2) = 0$, which says that the vector $(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$ is in $S$. Thus $S$ is a vector space. □

Subspaces of $\mathbb{R}^2$ (it turns out that this is the complete list):

- $\{0\}$ (the set containing only the origin)
- a line through the origin
- the whole plane $\mathbb{R}^2$.

Subspaces of $\mathbb{R}^3$ (again, the complete list):

- $\{0\}$
- a line through the origin
- a plane through the origin
- the whole space $\mathbb{R}^3$.

10.2. **Linear combinations.**
**Definition 10.3.** A **linear combination** of vectors $v_1, \ldots, v_n$ is a vector of the form $c_1v_1 + \cdots + c_nv_n$ for some scalars $c_1, \ldots, c_n$.

**10.3. Span.**

**Definition 10.4.** The **span** of $v_1, \ldots, v_n$ is the set of all linear combinations of $v_1, \ldots, v_n$:

$$\text{Span}(v_1, \ldots, v_n) := \{\text{all vectors } c_1v_1 + \cdots + c_nv_n, \text{ where } c_1, \ldots, c_n \text{ are scalars}\}.$$  

This set is always a vector space.

*Example 10.5.* If $v = (1, 1)$ in $\mathbb{R}^2$, then $\text{Span}(v)$ is the line $y = x$.

*Example 10.6.* If $i = (1, 0, 0)$ and $j = (0, 1, 0)$, then $\text{Span}(i, j)$ is the set of all vectors of the form

$$c_1i + c_2j = (c_1, c_2, 0).$$

These form the $xy$-plane in $\mathbb{R}^3$, whose equation is $z = 0$.

**10.4. Linearly dependent vectors.**

**Definition 10.7.** Vectors $v_1, \ldots, v_n$ are **linearly dependent** if one of them is a linear combination of the others.

*Equivalent definition:* $v_1, \ldots, v_n$ are **linearly dependent** if there exist scalars $c_1, \ldots, c_n$ not all zero such that $c_1v_1 + \cdots + c_nv_n = 0$.

**10.5. Basis.**

**Definition 10.8.** A **basis** of a vector space $S$ (of vectors) is a list of vectors $v_1, v_2, \ldots$ such that

1. $\text{Span}(v_1, v_2, \ldots) = S$, and
2. The functions $v_1, v_2, \ldots$ are linearly independent.

*Example 10.9.* If $S$ is the $xy$-plane in $\mathbb{R}^3$, then $i, j$ is a basis for $S$.

**10.6. Dimension.** Every basis for a vector space has the same number of vectors.

**Definition 10.10.** The **dimension** of a vector space is the number of vectors in any basis.

*Example 10.11.* The line $x + 3y = 0$ in $\mathbb{R}^2$ is a vector space $L$. The vector $(-3, 1)$ by itself is a basis for $L$, so the dimension of $L$ is 1. (Not a big surprise!)
11. Solved linear systems

11.1. Linear systems in matrix form. A linear system

\[\begin{align*}
2x + 5y - 7z &= 15 \\
x + z &= 1 \\
4x + 6y + 3z &= 0
\end{align*}\]

can be written in matrix form \(Ax = b:\)

\[
\begin{pmatrix}
2 & 5 & -7 \\
1 & 0 & 1 \\
4 & 6 & 3
\end{pmatrix}
\begin{pmatrix}
x \\
y \\
z
\end{pmatrix}
=
\begin{pmatrix}
15 \\
1 \\
0
\end{pmatrix}
\]

and can be represented by the augmented matrix

\[
\begin{pmatrix}
2 & 5 & -7 & 15 \\
1 & 0 & 1 & 1 \\
4 & 6 & 3 & 0
\end{pmatrix}
\]

(augmented with an extra column containing the right hand sides). Each row corresponds to an equation. Each column except the last one corresponds to a variable (and contains the coefficients of that variable).

A linear system is **homogeneous** if the right hand sides (the constants) are all zero, and **inhomogeneous** otherwise. So a linear system is homogeneous if and only if the zero vector is a solution.

A linear system is called **consistent** if it has at least one solution, and **inconsistent** if there are no solutions.

11.2. Equation operations. A good way to solve a linear system is to perform the following operations repeatedly, in some order:

- Multiply an equation by a nonzero number.
- Interchange two equations.
- Add a multiple of one equation to another equation.

The solution set is unchanged at each step.

11.3. Row operations. The equation operations correspond to operations on the augmented matrix, called **elementary row operations**:

- Multiply a row by a nonzero number.
- Interchange two rows.
- Add a multiple of one row to another row (while leaving the first row as it was).
11.4. **Row-echelon form.**

Steps to solve a linear system $Ax = b$:

1. Use row operations to convert the augmented matrix to a particularly simple form, called *row-echelon form*.
2. Solve the new system by *back-substitution*.

11.4.1. **Pivots.** Before explaining row-echelon form, we need a few preliminary definitions. A *zero row* of a matrix is a row consisting entirely of zeros. A *nonzero row* of a matrix is a row with at least one nonzero entry. In each nonzero row, the first nonzero entry is called the *pivot*.

*Example 11.1.* The following $4 \times 5$ matrix has one zero row, and three pivots (shown in red):

$$
\begin{pmatrix}
0 & -5 & 4 & 4 & 3 \\
2 & 0 & 0 & 1 & 7 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 3
\end{pmatrix}
$$

11.4.2. **Definition of row-echelon form.**

**Definition 11.2.** A matrix is in *row-echelon form* if it satisfies both of the following conditions:

1. All the zero rows (if any) are grouped at the bottom of the matrix.
2. Each pivot lies farther to the right than the pivots of higher rows.

**Warning:** Some books require also that each pivot be a 1. We are not going to require this for row-echelon form, but we will require it for *reduced* row-echelon form later on.

11.4.3. **Gaussian elimination.** This section was covered in March 6 recitation.

**Gaussian elimination** is an algorithm for converting any matrix into row-echelon form by performing row operations. Here are the steps:

1. Find the leftmost nonzero column, and the first nonzero entry in that column (read from the top down).
2. If that entry is not already in the first row, interchange its row with the first row.
3. Make all other entries of the column zero by adding suitable multiples of the first row to the others.
4. At this point, the first row is done, so ignore it, and repeat the steps above for the remaining submatrix (with one fewer row). In each iteration, ignore the rows already taken care of. Eventually the whole matrix will be in row-echelon form.
Problem 11.3. Convert the $4 \times 7$ matrix
\[
\begin{pmatrix}
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
6 & -9 & 0 & 11 & -19 & 3 & 0 \\
\end{pmatrix}
\]
to row-echelon form. (This example is taken from Hill, *Elementary linear algebra with applications*, p. 17.)

Solution:

*Step 1.* The leftmost nonzero column is the first one, and its first nonzero entry is the 2:
\[
\begin{pmatrix}
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
6 & -9 & 0 & 11 & -19 & 3 & 0 \\
\end{pmatrix}
\]

*Step 2.* The 2 is not in the first row, so interchange its row with the first row:
\[
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
6 & -9 & 0 & 11 & -19 & 3 & 0 \\
\end{pmatrix}
\]

*Step 3.* To make all other entries of the column zero, we need to add $-3$ times the first row to the last row (the other rows are OK already):
\[
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
0 & 0 & -3 & -1 & 2 & 0 & -6 \\
\end{pmatrix}
\]

*Step 4.* Now the first row is done. Start over with the $3 \times 7$ submatrix that remains beneath it:
\[
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
0 & 0 & -3 & -1 & 2 & 0 & -6 \\
\end{pmatrix}
\]

*Step 1.* The leftmost nonzero column is now the third column, and its first nonzero entry is the 3:
\[
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
0 & 0 & 6 & 2 & -4 & -8 & 8 \\
0 & 0 & -3 & -1 & 2 & 0 & -6 \\
\end{pmatrix}
\]
Step 2. The 3 is already in the first row of the submatrix (we are ignoring the first row of the whole matrix), so no interchange is necessary.

Step 3. To make all other entries of the column zero, add $-2$ times the (new) first row to the (new) second row, and 1 times the (new) first row to the (new) third row:

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 & 1 & 2 \\ 0 & 0 & 3 & 1 & -2 & -4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & -2 \end{pmatrix}.$$  

Step 4. Now the first and second row of the original matrix are done. Start over with the $2 \times 7$ submatrix beneath them:

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 & 1 & 2 \\ 0 & 0 & 3 & 1 & -2 & -4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & -2 \end{pmatrix}.$$  

Step 1. The leftmost nonzero column is now the penultimate column, and its first nonzero entry is the $-4$ at the bottom:

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 & 1 & 2 \\ 0 & 0 & 3 & 1 & -2 & -4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & -2 \end{pmatrix}.$$  

Step 2. The $-4$ is not in the first row of the submatrix, so interchange its row with the first row of the submatrix:

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 & 1 & 2 \\ 0 & 0 & 3 & 1 & -2 & -4 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -4 & -2 \end{pmatrix}.$$  

Step 3. The other entry in this column of the submatrix is already 0, so this step is not necessary.

The matrix is now in row-echelon form:

$$\begin{pmatrix} 2 & -3 & 1 & 4 & -7 & 1 & 2 \\ 0 & 0 & 3 & 1 & -2 & -4 & 4 \\ 0 & 0 & 0 & 0 & -4 & -2 \end{pmatrix}.$$  

11.5. **Reduced row-echelon form.** This section was covered in March 6 recitation.
11.6. **Definition of reduced row-echelon form.** With even more row operations, one can simplify a matrix in row-echelon form to an even more special form:

**Definition 11.4.** A matrix is in **reduced row-echelon form (RREF)** if it satisfies all of the following conditions:

1. It is in row-echelon form.
2. Each pivot is a 1.
3. In each pivot column, all the entries are 0 except for the pivot itself.

11.7. **Gauss–Jordan elimination.** **Gauss–Jordan elimination** is an algorithm for converting any matrix into reduced row-echelon form by performing row operations. Here are the steps:

1. Use Gaussian elimination to convert the matrix to row echelon form.
2. Divide the last nonzero row by its pivot, to make the pivot 1.
3. Make all entries in that pivot’s column 0 by adding suitable multiples of the pivot’s row to the rows above.
4. At this point, the row in question (and all rows below it) are done. Ignore them, and go back to Step 2, but now with the remaining submatrix, above the row just completed.

Eventually the whole matrix will be in reduced row-echelon form.

**Problem 11.5.** Convert the $4 \times 7$ matrix

$$
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
0 & 0 & 0 & 0 & -4 & -2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
$$

to reduced row-echelon form.

**Solution:**

**Step 1.** The matrix is already in row-echelon form.

**Step 2.** The last nonzero row is the third row, and its pivot is the $-4$, so divide the third row by $-4$:

$$
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 1 & 2 \\
0 & 0 & 3 & 1 & -2 & -4 & 4 \\
0 & 0 & 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
$$

**Step 3.** To make all other entries of that pivot’s column 0, add $-1$ times the third row to the first row, and add $4$ times the third row to the second row:

$$
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 0 & 3/2 \\
0 & 0 & 3 & 1 & -2 & 0 & 6 \\
0 & 0 & 0 & 0 & 1 & 1/2 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
$$
Step 4. Now the last two rows are done:
\[
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 0 & 3/2 \\
0 & 0 & 3 & 1 & -2 & 0 & 6 \\
0 & 0 & 0 & 0 & 0 & 1 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
\]

Go back to Step 2, but with the 2 × 7 submatrix above them.

Step 2. The last nonzero row of the new matrix (ignoring the bottom two rows of the original matrix) is the second row, and its pivot is the 3, so we divide the second row by 3:
\[
\begin{pmatrix}
2 & -3 & 1 & 4 & -7 & 0 & 3/2 \\
0 & 0 & 1 & 1/3 & -2/3 & 0 & 2 \\
0 & 0 & 0 & 0 & 0 & 1 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]

Step 3. To make the other entries of that pivot’s column 0, add \(-1\) times the second row to the first row:
\[
\begin{pmatrix}
2 & -3 & 0 & 11/3 & -19/3 & 0 & -1/2 \\
0 & 0 & 1 & 1/3 & -2/3 & 0 & 2 \\
0 & 0 & 0 & 0 & 1 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]

Step 4. Now the last three rows are done:
\[
\begin{pmatrix}
2 & -3 & 0 & 11/3 & -19/3 & 0 & -1/2 \\
0 & 0 & 1 & 1/3 & -2/3 & 0 & 2 \\
0 & 0 & 0 & 0 & 1 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]

Go back to Step 2, but with the 1 × 7 submatrix above them.

Step 2. The last nonzero row of the new matrix is the only remaining row (the first row), and its pivot is the initial 2, so we divide the first row by 2:
\[
\begin{pmatrix}
1 & -3/2 & 0 & 11/6 & -19/6 & 0 & -1/4 \\
0 & 0 & 1 & 1/3 & -2/3 & 0 & 2 \\
0 & 0 & 0 & 0 & 1 & 1/2 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
\]

The matrix is now in reduced row-echelon form. □

Remark 11.6. Performing row operations on \(A\) in a different order than specified by Gaussian elimination and Gauss–Jordan elimination can lead to different row-echelon forms. But it turns out that row operations leading to reduced row-echelon form always give the same result, a matrix that we will write as RREF(\(A\)).
11.8. **Back-substitution.**

Key point of row-echelon form: Matrices in row-echelon form correspond to systems that are ready to be solved immediately by **back-substitution**: solve for each variable in reverse order, while introducing a parameter for each variable not directly expressed in terms of later variables, and substitute values into earlier equations once they are known.

**Problem 11.7.** The inhomogeneous linear system

\[
\begin{align*}
x + 2y + 2v + 3w &= 4 \\
y - 2z + 3v + w &= 5 \\
2w &= 6
\end{align*}
\]

has augmented matrix

\[
\begin{pmatrix}
1 & 2 & 0 & 2 & 3 & 4 \\
0 & -1 & 2 & 3 & 1 & 5 \\
0 & 0 & 0 & 0 & 2 & 6
\end{pmatrix}
\]

in row-echelon form. Find the general solution to the system.

**Solution:** Use the equations in reverse order to solve for the variables in reverse order. Start by solving for the last variable, \(w\):

\[w = 3.\]

There is no equation for \(v\) in terms of the later variable \(w\), so \(v\) can be any number; set

\[v = c_1\] for a parameter \(c_1\).

There is no equation for \(z\) in terms of \(v, w\), so set

\[z = c_2\] for a parameter \(c_2\).

Substitute the values of \(w, v, z\) into the next-to-last equation, and solve for \(y\):

\[
\begin{align*}
-y + 2c_2 + 3c_1 + 3 &= 5 \\
y &= 3c_1 + 2c_2 - 2 \\
x + 2(3c_1 + 2c_2 - 2) + 2c_1 + 3(3) &= 4 \\
x &= -8c_1 - 4c_2 - 1.
\end{align*}
\]
Conclusion: The general solution is

\[
\begin{pmatrix}
    x \\
    y \\
    z \\
    v \\
    w
\end{pmatrix} =
\begin{pmatrix}
    (−8c_1 − 4c_2 − 1) \\
    3c_1 + 2c_2 − 2 \\
    c_2 \\
    c_1 \\
    3
\end{pmatrix}
+ c_1 \begin{pmatrix}
    −1 \\
    −2 \\
    0 \\
    0 \\
    0
\end{pmatrix}
+ c_2 \begin{pmatrix}
    −8 \\
    3 \\
    0 \\
    1 \\
    0
\end{pmatrix},
\]

where \(c_1, c_2\) are parameters. □

Suppose that a matrix is in row echelon form. Then any column that contains a pivot is called a pivot column. A variable whose corresponding column is a pivot column is called a dependent variable or pivot variable. The other variables are called free variables. (The augmented column does not correspond to any variable.)

In the problem above, \(x, y, w\) were dependent variables, and \(v, z\) were free variables.

**Problem 11.8.** Find the general solution to the system with augmented matrix

\[
\begin{pmatrix}
    2 & 3 & 5 & 7 \\
    0 & 1 & 4 & 6 \\
    0 & 0 & 0 & 9
\end{pmatrix}
\]

**Solution:** The last equation says

\[0x + 0y + 0z = 9,\]

i.e., \(0 = 9\), which cannot be satisfied. So there are no solutions! □

**11.9. Interpreting an augmented matrix in row-echelon form.** Consider a linear system of \(m\) equations in \(n\) variables. Let \(B\) be the \(m \times (n + 1)\) augmented matrix after being put in row-echelon form.

Case 1. The matrix \(B\) has a pivot in the augmented column. This means that one of the new equations has the form

\[0x_1 + \cdots + 0x_n = \frac{b}{\text{nonzero number}}.\]

Conclusion: The linear system is inconsistent.

Case 2. The matrix \(B\) has no pivot in the augmented column. In a solution, the free variables may take any values, but in terms of these one can solve for the dependent variables in reverse order. Conclusion: The linear system is consistent, and

\[\#\text{parameters in general solution} = \#\text{non-pivot columns excluding the augmented column},\]

\[\#\text{free variables}\]

**11.10. Homogeneous linear systems: theory and algorithms.** For a homogeneous system, there is no need to keep track of an augmented column, because it would consist of zeros, and would stay that way even after row operations.
11.10.1. An example.

**Problem 11.9.** The homogeneous linear system

\[
\begin{align*}
    x + 2y + 2v + 3w &= 0 \\
    -y + 2z + 3v + w &= 0 \\
    2w &= 0
\end{align*}
\]

has matrix

\[
\begin{pmatrix}
1 & 2 & 0 & 2 & 3 \\
0 & -1 & 2 & 3 & 1 \\
0 & 0 & 0 & 0 & 2
\end{pmatrix}
\]

in row-echelon form; it is the homogeneous linear system associated to the inhomogeneous system in Problem 11.7. Find the general solution to the system.

**Solution:** Back-substitution again. Start with

\[w = 0.\]

No equation for \(v\) in terms of the later variable \(w\), so set

\[v = c_1 \text{ for a parameter } c_1.\]

No equation for \(z\) in terms of \(v, w\), so set

\[z = c_2 \text{ for a parameter } c_2.\]

\[-y + 2c_2 + 3c_1 = 0\]

\[y = 3c_1 + 2c_2\]

\[x + 2(3c_1 + 2c_2) + 2c_1 = 0\]

\[x = -8c_1 - 4c_2.\]

General solution:

\[
\begin{pmatrix}
x \\ y \\ z \\ v \\ w
\end{pmatrix} = \begin{pmatrix}
-8c_1 - 4c_2 \\ 3c_1 + 2c_2 \\ c_2 \\ c_1 \\ 0
\end{pmatrix} = c_1 \begin{pmatrix}
-8 \\ 3 \\ 0 \\ 1 \\ 0
\end{pmatrix} + c_2 \begin{pmatrix}
-4 \\ 2 \\ 1 \\ 0 \\ 0
\end{pmatrix},
\]
where $c_1, c_2$ are parameters.

Set of all solutions:

$$\text{Span} \left\{ \begin{pmatrix} -8 \\ 3 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}.$$ 

Since the set of solutions is a span, it is a vector space. Moreover, the two vectors are linearly independent, since, in the equivalent definition of linearly dependent, the only $c_1, c_2$ such that

$$c_1 \begin{pmatrix} -8 \\ 3 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -4 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

are $c_1 = 0, c_2 = 0$ (this is obvious if you look at the green rows, the rows corresponding to the free variables).

**Conclusion:** The set of solutions to this homogeneous linear system is a vector space with basis

$$\begin{pmatrix} -8 \\ 3 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$ 

Dimension $= 2$. □

11.10.2. **Nullspace.** Here is what happens in general for homogeneous linear systems:

**Theorem 11.10.** For any homogeneous linear system $A\mathbf{x} = \mathbf{0}$, the set of all solutions is a vector space, called the **nullspace** of the matrix $A$, and denoted $\text{NS}(A)$.

Notice the similarity with Theorem 5.3!

**Theorem 11.11** (Formula for the dimension of the nullspace). Suppose that the result of putting a matrix $A$ in row-echelon form is $B$. Then $\text{NS}(A) = \text{NS}(B)$ (since row reductions do not change the solutions), and

$$\text{dim } \text{NS}(A) = \#\text{non-pivot columns of } B.$$ 

(The boxed formula holds since it is the same as $\text{dim } \text{NS}(B) = \#\text{free variables}$.)

In other words, here are the steps to find the dimension of the space of solutions to a homogeneous linear system $A\mathbf{x} = \mathbf{0}$:
1. Perform Gaussian elimination on \( A \) to convert it to a matrix \( B \) in row-echelon form.
2. Identify the pivots of \( B \).
3. Count the number of non-pivot columns of \( B \); that number is \( \dim \text{NS}(A) \).

Warning: You must put the matrix in row-echelon form before counting non-pivot columns!

And here are the steps to find a basis of the space of solutions to a homogeneous linear system \( Ax = 0 \):
1. Perform Gaussian elimination on \( A \) to convert it to a matrix \( B \) in row-echelon form.
2. Use back-substitution to find the general solution to \( Bx = 0 \).
3. The general solution will be expressed as the general linear combination of a list of vectors; that list is a basis of \( \dim \text{NS}(A) \).

11.11. **Inhomogeneous linear systems: theory and algorithms.** For an inhomogeneous linear system \( Ax = b \), there are two possibilities:

1. There are no solutions.
2. There exists a solution. In this case, if \( x_p \) is a particular solution to \( Ax = b \), and \( x_h \) is the general solution to the homogeneous system \( Ax = 0 \), then \( x := x_p + x_h \) is the general solution to \( Ax = b \).

Here is why: Suppose that a solution exists; let \( x_p \) be one, so \( Ax_p = b \). If \( x_h \) satisfies \( Ax_h = 0 \), adding the two equations gives \( A(x_p + x_h) = b \), so adding \( x_p \) to \( x_h \) produces a solution \( x \) to the inhomogeneous equation. Every solution \( x \) to \( Ax = b \) arises this way from some \( x_h \) (specifically, from \( x_h := x - x_p \), which satisfies \( Ax_h = Ax - Ax_p = b - b = 0 \)).

Remark 11.12. To solve \( Ax = b \), however, don’t use \( x = x_p + x_h \). Instead use Gaussian elimination and back-substitution. The above is just to describe the shape of the solution.

**Problem 11.13.** For which vectors \( b \in \mathbb{R}^2 \) does the inhomogeneous linear system

\[
\begin{pmatrix}
1 & 2 & 3 \\
2 & 4 & 6
\end{pmatrix}
\begin{pmatrix}
x_1 \\
x_2 \\
x_3
\end{pmatrix} = b
\]

have a solution?

**Answer:** The left hand side can be rewritten as

\[
\begin{pmatrix}
x_1 + 2x_2 + 3x_3 \\
2x_1 + 4x_2 + 6x_3
\end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 4 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 6 \end{pmatrix}
\]

Thus, saying that the system has a solution is the same as saying that \( b \) is a linear combination of \( \begin{pmatrix} 1 \\ 2 \end{pmatrix} \), \( \begin{pmatrix} 2 \\ 4 \end{pmatrix} \), \( \begin{pmatrix} 3 \\ 6 \end{pmatrix} \), \( \ldots \), \( \begin{pmatrix} 0 \\ 0 \end{pmatrix} \).
or equivalently, that
\[
\mathbf{b} \text{ is in the span of } \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{pmatrix}, \quad \square
\]

11.11.1. **Column space.**

**Definition 11.14.** The **column space** of a matrix \( A \) is the span of its columns. The notation for it is \( \text{CS}(A) \). (It is also called the **image** of \( A \), and written \( \text{im}(A) \); the reason will be clearer when we talk about the geometric interpretation.)

Since \( \text{CS}(A) \) is a span, it is a vector space.

Here is what happens in general for (possibly inhomogeneous) linear systems (the explanation is the same as in the example above):

**Theorem 11.15.** The linear system \( A\mathbf{x} = \mathbf{b} \) has a solution if and only if \( \mathbf{b} \) is in \( \text{CS}(A) \).

For the matrix \( A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{pmatrix} \) in the problem above,

\[
\text{CS}(A) = \text{the span of } \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 6 \end{pmatrix},
\]

which is the line \( y = 2x \) in \( \mathbb{R}^2 \), a 1-dimensional vector space.

**Steps to compute a basis for \( \text{CS}(A) \):**

1. Perform Gaussian elimination to convert \( A \) to a matrix \( B \) in row-echelon form.
2. Identify the pivot columns of \( B \).
3. The corresponding columns of \( A \) are a basis for \( \text{CS}(A) \).

Here is a summary of why this works (not discussed in lecture). Let \( C \) be the reduced row-echelon form of \( A \). If

\[
\text{fifth column} = 3(\text{first column}) + 7(\text{second column})
\]

is true for a matrix, it will remain true after any row operation. Similarly, any linear relation between columns is preserved by row operations. So the linear relations between columns of \( A \) are the same as the linear relations between columns of \( C \). The condition that certain numbered columns (say the first, second, and fourth) of a matrix form a basis is expressible in terms of which linear relations hold, so if certain columns form a basis for \( \text{CS}(C) \), the same numbered columns will form a basis for \( \text{CS}(A) \). Also, performing Gauss–Jordan reduction on \( B \) to obtain \( C \) in reduced row-echelon form does not change the pivot locations. Thus it will be enough to show that the pivot columns of \( C \) form a basis of \( \text{CS}(C) \). Since \( C \) is in reduced row-echelon form, the pivot columns of \( C \) are the first \( r \) of the \( m \) standard basis vectors for \( \mathbb{R}^m \), where \( r \) is
the number of nonzero rows of \( C \). These columns are linearly independent, and every other column is a linear combination of them, since the entries of \( C \) below the first \( r \) rows are all zeros. Thus the pivot columns of \( C \) form a basis of \( \text{CS}(C) \).

In particular, \( \dim \text{CS}(A) = \# \text{pivot columns of } B \).

**Warning:** Usually \( \text{CS}(A) \neq \text{CS}(B) \).

### 11.11.2. Row space.

**Definition 11.16.** The **row space** of a matrix \( A \) is the span of its rows. The notation for it is \( \text{RS}(A) \).

**Steps to compute a basis for \( \text{RS}(A) \):**

1. Perform Gaussian elimination to convert \( A \) to a matrix \( B \) in row-echelon form.
2. The nonzero rows of \( B \) are a basis for \( \text{RS}(A) = \text{RS}(B) \).

This works since row operations do not change the row space, and the nonzero rows of a matrix in row-echelon form are linearly independent.

In particular, \( \dim \text{RS}(A) = \# \text{nonzero rows of } B \).

### 11.11.3. Rank.

**Problem 11.17.** Let \( A \) be the \( 3 \times 5 \) matrix

\[
\begin{pmatrix}
1 & 2 & 3 & 4 & 5 \\
-1 & -2 & 9 & 10 & 11 \\
1 & 2 & 9 & 11 & 13
\end{pmatrix}
\]

(a) Find bases for \( \text{RS}(A) \) and \( \text{CS}(A) \).
(b) What are \( \dim \text{NS}(A) \), \( \dim \text{RS}(A) \), \( \dim \text{CS}(A) \)?

**Solution:**

(a) First we must find a row-echelon form. Add the first row to the second, and add \(-1\) times the first row to the third:

\[
\begin{pmatrix}
1 & 2 & 3 & 4 & 5 \\
0 & 0 & 12 & 14 & 16 \\
0 & 0 & 6 & 7 & 8
\end{pmatrix}
\]

Add \(-1/2\) times the second row to the third:

\[
B := \begin{pmatrix}
1 & 2 & 3 & 4 & 5 \\
0 & 0 & 12 & 14 & 16 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix}
\]

This is in row-echelon form.

Basis for \( \text{RS}(A) = \text{RS}(B) \): nonzero rows of \( B \), i.e., \((1, 2, 3, 4, 5)\) and \((0, 0, 12, 14, 16)\).
Basis for CS(B): first and third columns (the pivot columns) of \( B \), i.e., \[
\begin{pmatrix}
1 \\
0 \\
3 \\
0
\end{pmatrix},
\begin{pmatrix}
12 \\
0 \\
0 \\
0
\end{pmatrix}.
\]

Basis for CS(A): first and third columns of \( A \), i.e., \[
\begin{pmatrix}
1 \\
-1 \\
3
\end{pmatrix},
\begin{pmatrix}
3 \\
9
\end{pmatrix}.
\]

(b)

\[
\begin{align*}
dim \text{NS}(A) &= \# \text{non-pivot columns of } B = 3 \\
dim \text{RS}(A) &= \# \text{nonzero rows of } B = 2 \\
dim \text{CS}(A) &= \# \text{pivot columns of } B = 2.
\end{align*}
\]

For any \( B \) in row-echelon form,

\[
\# \text{nonzero rows of } B = \# \text{pivots of } B = \# \text{pivot columns of } B.
\]

Therefore \( \dim \text{RS}(A) = \dim \text{CS}(A) \) holds for every matrix \( A \! \)

**Definition 11.18.** The **rank** of \( A \) is defined by

\[
\text{rank}(A) := \dim \text{RS}(A) = \dim \text{CS}(A).
\]

**Theorem 11.19.** For any \( m \times n \) matrix \( A \),

\[
\text{dim NS}(A) + \text{rank}(A) = n.
\]

**Proof.**

\[
\begin{align*}
\dim \text{NS}(A) + \text{rank}(A) &= (\# \text{non-pivot columns of } B) + (\# \text{pivot columns of } B) \\
&= \# \text{columns of } B \\
&= n.
\end{align*}
\]

11.11.4. **Computing a basis for a span.**

**Problem 11.20.** Given vectors \( \mathbf{v}_1, \ldots, \mathbf{v}_n \in \mathbb{R}^m \), how can one compute a basis of \( \text{Span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) \)?

**Method 1:**

1. Form the matrix \( A \) whose rows are \( \mathbf{v}_1, \ldots, \mathbf{v}_n \).
2. Find a basis for \( \text{RS}(A) \) as above (nonzero rows of \( B \)).

**Method 2:**

1. Form the matrix \( A \) whose columns are \( \mathbf{v}_1, \ldots, \mathbf{v}_n \).
2. Find a basis for \( \text{CS}(A) \) as above (columns of \( A \) corresponding to pivot columns of \( B \)).
Method 1 is slightly easier. Method 2 has the feature that the basis it gives is a subset of the original list of vectors.

12. Geometric interpretation of solving linear systems

12.1. Linear transformations.

12.1.1. Evaluating a linear transformation. Each \( m \times n \) matrix \( A \) represents a function \( f : \mathbb{R}^n \to \mathbb{R}^m \), called a linear transformation: you plug in an input vector \( x \) of \( \mathbb{R}^n \), and the output vector is \( f(x) := Ax \) in \( \mathbb{R}^m \).

**Problem 12.1.** The matrix \( \begin{pmatrix} 1 & 2 \\ 4 & 7 \\ 8 & 9 \end{pmatrix} \) represents a linear transformation \( f \).

Evaluate \( f \) at \( \begin{pmatrix} 10 \\ 1 \end{pmatrix} \).

**Solution:** The value is

\[
f \begin{pmatrix} 10 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 4 & 7 \\ 8 & 9 \end{pmatrix} \begin{pmatrix} 10 \\ 1 \end{pmatrix} = \begin{pmatrix} 12 \\ 47 \\ 89 \end{pmatrix}.
\]

12.1.2. Depicting a linear transformation. Imagine evaluating \( f \) on every vector in the input space \( \mathbb{R}^n \), to get vectors in the output space \( \mathbb{R}^m \). To visualize it, draw a shape in the input space, apply \( f \) to every point in the shape, and plot the output points in the output space.

**Problem 12.2.** The matrix \( \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \) represents a linear transformation \( f \). Depict \( f \) by showing what it does to the standard basis vectors \( i, j \) of \( \mathbb{R}^2 \) and to the unit smiley. What is the area scaling factor?

**Solution:** We have

\[
f \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}
\]

\[
f \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\]

and the unit smiley is stretched horizontally into a fat smiley of the same height.
For a $2 \times 2$ matrix, the area scaling factor is the absolute value of the determinant:

$$\left| \det \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \right| = |2| = 2. \quad \square$$

Try the “Matrix Vector” mathlet


12.1.3. *Going from a linear transformation to a matrix.*

**Problem 12.3.** Given $\theta$, there is a $2 \times 2$ matrix $R$ whose associated linear transformation rotates each vector in $\mathbb{R}^2$ counterclockwise by the angle $\theta$. What is it?
Solution: The rotation maps $\langle 1, 0 \rangle$ to $\langle \cos \theta, \sin \theta \rangle$ and $\langle 0, 1 \rangle$ to $\langle -\sin \theta, \cos \theta \rangle$. Thus

(first column of $R$) = $R \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$

(second column of $R$) = $R \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$,

so

$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$.

12.2. Example: a projection. Let $f(x, y, z) := (x, y, 0)$. This $f$ is a projection from the input space $\mathbb{R}^3$ onto the $xy$-plane in the output space $\mathbb{R}^3$. Let $A$ be the matrix representing $f$. So $A$ is a $3 \times 3$ matrix such that

(first column of $A$) = $f \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

(second column of $A$) = $f \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

(third column of $A$) = $f \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$.

Thus $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

NS($A$) is a subspace of the input space:

NS($A$) = \{solutions to $Ax = 0$\}
= \{solutions to $f(x, y, z) = 0$\}
= \{(0, 0, z) : z \in \mathbb{R}\}
= the $z$-axis in the input space $\mathbb{R}^3$.

The image CS($A$) is a subspace of the output space:

CS($A$) = \{values of $Ax$\}
= \{values of $f(x, y, z)$\}
= \{(x, y, 0) : x, y \in \mathbb{R}\}
= the $xy$-plane in the output space $\mathbb{R}^3$. 

81
Here \( \text{rank}(A) = \dim \text{CS}(A) = 2 \).

The linear transformation \( f \) crushes \( \text{NS}(A) \) to the point \( 0 \) in the output space, and it flattens the whole input space \( \mathbb{R}^3 \) onto \( \text{CS}(A) \) in the output space. Of the 3 input dimensions, 1 is crushed, so \( 3 - 1 = 2 \) dimensions are left.

(Mathematically, “there are \( m \) crushed dimensions” means just that \( \text{NS}(A) \) is \( m \)-dimensional.)

In general, for a linear transformation \( f : \mathbb{R}^n \to \mathbb{R}^m \) represented by an \( m \times n \) matrix \( A \), of the \( n \) input dimensions, \( \dim \text{NS}(A) \) of them are crushed, leaving an image of dimension \( n - \dim \text{NS}(A) \). This explains geometrically why

\[
\dim \text{NS}(A) + \text{rank}(A) = n
\]

which is the same as the theorem

\[
\dim \text{NS}(A) + \text{rank}(A) = n
\]

we had earlier.

\[\text{March 12}\]

Back to the example: What does the solution set to \( Ax = b \) look like?
• If \( b \) is not in \( \text{CS}(A) \), then there are no solutions.
• If \( b \) is in \( \text{CS}(A) \), say \( b = (b_1, b_2, 0) \), then

\[
\{ \text{solutions to } Ax = b \} = \{ \text{solutions to } f(x, y, z) = (b_1, b_2, 0) \} = \{(b_1, b_2, z) : z \in \mathbb{R} \}
\]

= a vertical line parallel to the \( z \)-axis in the input space \( \mathbb{R}^3 \).

The general solution to the homogeneous system \( Ax = 0 \) is the line \( \text{NS}(A) \). To get from \( \text{NS}(A) \) to the general solution to \( Ax = b \), choose a particular solution vector to \( Ax = b \) and add it to every vector in \( \text{NS}(A) \).

**Problem 12.4.** What is the volume scaling factor?

**Solution 1:** It’s the absolute value of the determinant:

\[
\left| \begin{array}{ccc}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
\end{array} \right| = |0| = 0.
\]

**Solution 2:** The linear transformation \( f \) takes any unit cube of volume 1 to a flat object of volume 0, so volume is getting multiplied by 0.

13. **Square matrices**

13.1. **Determinants.** To each square matrix \( A \) is associated a number called the determinant:

\[
\det \left( \begin{array}{ccc}
a & b & c \\
0 & 1 & 0 \\
0 & 0 & 0 \\
\end{array} \right) = a
\]

\[
\det \left( \begin{array}{ccc}
a & b & c \\
c & d & e \\
0 & 0 & 0 \\
\end{array} \right) = ad - bc
\]

\[
\det \left( \begin{array}{ccc}
a_1 & a_2 & a_3 \\
b_1 & b_2 & b_3 \\
c_1 & c_2 & c_3 \\
\end{array} \right) = a_1b_2c_3 + a_2b_3c_1 + a_3b_1c_2 - c_1b_2a_3 - c_2b_3a_1 - c_3b_1a_2.
\]

Alternative notation for determinant: \( |A| = \left| \begin{array}{ccc}
a & b & c \\
0 & 1 & 0 \\
0 & 0 & 0 \\
\end{array} \right| \). (This is a scalar, not a matrix!)

Geometric meaning: The absolute value of \( \det A \) is the area scaling factor (or volume scaling factor or . . . ).

**Laplace expansion** (along the first row) for a \( 3 \times 3 \) determinant:

\[
\left| \begin{array}{ccc}
a_1 & a_2 & a_3 \\
b_1 & b_2 & b_3 \\
c_1 & c_2 & c_3 \\
\end{array} \right| = +a_1 \left| \begin{array}{cc}
b_2 & b_3 \\
c_2 & c_3 \\
\end{array} \right| - a_2 \left| \begin{array}{cc}
b_1 & b_3 \\
c_1 & c_3 \\
\end{array} \right| + a_3 \left| \begin{array}{cc}
b_1 & b_2 \\
c_1 & c_2 \\
\end{array} \right|.
\]
The general rule leading to the formula above is this:

1. Move your finger along the entries in a row.
2. At each position, compute the minor, defined as the smaller determinant obtained by crossing out the row and the column through your finger; then multiply the minor by the number you are pointing at, and adjust the sign according to the checkerboard pattern
   
   +  −  +  
   −  +  −  
   +  −  +  
   (the pattern always starts with + in the upper left corner).
3. Add up the results.

There is a similar expansion for a determinant of any size, computed along any row or column.

The diagonal of a matrix consists of the entries \( a_{ij} \) with \( i = j \).

A diagonal matrix is a matrix that has zeros everywhere outside the diagonal:

\[
\begin{pmatrix}
a_{11} & 0 & 0 \\
0 & a_{22} & 0 \\
0 & 0 & a_{33}
\end{pmatrix}
\]

(it may have some zeros along the diagonal too).

An upper triangular matrix is a matrix whose entries strictly below the diagonal are all 0:

\[
\begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
0 & a_{22} & a_{23} \\
0 & 0 & a_{33}
\end{pmatrix}
\]

(the entries on or above the diagonal may or may not be 0).

**Example 13.1.** Any square matrix in row-echelon form is upper triangular.

**Theorem 13.2.** The determinant of an upper triangular matrix equals the product of the diagonal entries.

For example,

\[
\det \begin{pmatrix}
a_{11} & a_{12} & a_{13} \\
0 & a_{22} & a_{23} \\
0 & 0 & a_{33}
\end{pmatrix} = a_{11} a_{22} a_{33}.
\]

Why is the theorem true? The Laplace expansion along the first column shows that the determinant is \( a_{11} \) times a upper triangular minor with diagonal entries \( a_{22}, \ldots, a_{nn} \).

Properties of determinants:

1. Interchanging two rows changes the sign of \( \det A \).
(2) Multiplying an entire row by a scalar \(c\) multiples \(\det A\) by \(c\).

(3) Adding a multiple of a row to another row does not change \(\det A\).

(4) If one of the rows is all 0, then \(\det A = 0\).

(5) \(\det(AB) = \det(A) \det(B)\) (assuming \(A, B\) are square matrices of the same size).

In particular, row operations multiply \(\det A\) by nonzero scalars, but do not change whether \(\det A = 0\).

**Question 13.3.** Suppose that \(A\) is a 3 \(\times\) 3 matrix such that \(\det A = 5\). Doubling every entry of \(A\) gives a matrix \(2A\). What is \(\det(2A)\)?

**Solution:** Each time we multiply a row by 2, the determinant gets multiplied by 2. We need to do this three times to double the whole matrix \(A\), so the determinant gets multiplied by \(2 \cdot 2 \cdot 2 = 8\). Thus \(\det(2A) = 8 \det(A) = 40\).

**13.2. Identity matrix.** The linear transformation \(f: \mathbb{R}^3 \to \mathbb{R}^3\) that does nothing to its input, \(f(x, y, z) := (x, y, z)\), is called the identity. The corresponding matrix, the 3 \(\times\) 3 identity matrix \(I\), has

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

so

\[
I := \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

(The \(n \times n\) identity matrix is similar, with 1s along the diagonal.)

It has the property that \(AI = A\) and \(IA = A\) whenever the matrix multiplication is defined.

**Example 13.4.** Suppose that \(A\) is a square matrix and \(\det A \neq 0\). Then \(\text{RREF}(A)\) has nonzero determinant too, but is now upper triangular, so its diagonal entries are nonzero. In fact, these diagonal entries are 1 since they are pivots of a RREF matrix. Moreover, all non-diagonal entries are 0, by definition of RREF. So \(\text{RREF}(A) = I\).

Now imagine solving \(Ax = b\). Gauss–Jordan elimination converts the augmented matrix \([A|b]\) to \([I|c]\) for some vector \(c\). Thus \(Ax = b\) has the same solutions as \(Ix = c\); the unique solution is \(c\).

What if instead we wanted to solve many equations with the same \(A\): \(Ax_1 = b_1, \ldots, Ax_p = b_p\)? Use many augmented columns! Gauss–Jordan elimination converts \([A|b_1 \ldots b_p]\) to \([I|c_1 \ldots c_p]\), and \(c_1, \ldots, c_p\) are the solutions.

In other words, to solve an equation \(AX = B\) to find the unknown matrix \(X\), convert \([A|B]\) to RREF \([I|C]\). Then \(C\) is the solution. \(\square\)
13.3. **Inverse matrices.**

**Definition 13.5.** The inverse of an \( n \times n \) matrix \( A \) is another \( n \times n \) matrix \( A^{-1} \) such that \( AA^{-1} = I \) and \( A^{-1}A = I \).

*It exists if and only if \( \det A \neq 0 \).*

Suppose that \( A \) represents the linear transformation \( f \). Then \( A^{-1} \) exists if and only if an inverse function \( f^{-1} \) exists; in that case, \( A^{-1} \) represents \( f^{-1} \).

**Problem 13.6.** Does the rotation matrix \( R := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \) have an inverse? If so, what is it?

**Solution:** The inverse linear transformation is rotation by \(-\theta\), so

\[
R^{-1} = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}.
\]

(As a check, try multiplying \( R \) by this matrix, in either order.) \( \Box \)

**Problem 13.7.** Does the projection matrix \( A := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \) have an inverse? If so, what is it?

**Solution:** The associated linear transformation \( f \) is not a 1-to-1 correspondence, because it maps more than one vector to \( \mathbf{0} \) (it maps the whole \( z \)-axis to \( \mathbf{0} \)). Thus \( f^{-1} \) does not exist, so \( A^{-1} \) does not exist. \( \Box \)

Suppose that \( \det A \neq 0 \). In 18.02, you learned one algorithm to compute \( A^{-1} \), using the cofactor matrix. Now that we know how to compute RREF, we can give a faster algorithm (faster for big matrices, at least):

**New algorithm to find \( A^{-1} \):**

1. Form the \( n \times 2n \) augmented matrix \([A|I]\).
2. Convert to RREF; the result will be \([I|B]\) for some \( n \times n \) matrix \( B \).
3. Then \( A^{-1} = B \).

This is a special case of Example 13.4 since \( A^{-1} \) is the solution to \( AX = I \).

13.4. **Conditions for invertibility.** There are two types of square matrices \( A \):

- those with \( \det A \neq 0 \) (called **nonsingular** or **invertible**), and
- those with \( \det A = 0 \) (called **singular**).
13.4.1. Nonsingular matrices.

**Theorem 13.8.** For a square matrix $A$, the following are equivalent:

1. $\det A \neq 0$ (scaling factor is positive)
2. $\text{NS}(A) = \{0\}$ (the only solution to $Ax = 0$ is $0$)
3. $\text{rank}(A) = n$ (image is $n$-dimensional)
4. $\text{CS}(A) = \mathbb{R}^n$ (image is the whole space $\mathbb{R}^n$)
5. For each vector $b$, the system $Ax = b$ has exactly one solution.
6. $A^{-1}$ exists.
7. $\text{RREF}(A) = I$

So if you have a matrix $A$ for which one of these conditions holds, then all of the conditions hold for $A$.

Let’s explain the consequences of $\det A \neq 0$. Suppose that $\det A \neq 0$. Then the volume scaling factor is not 0, so the input space $\mathbb{R}^n$ is not flattened by $A$. This means that there are no “crushed dimensions”, so $\text{NS}(A) = \{0\}$. Since no dimensions were crushed, the image $\text{CS}(A)$ has the same dimension as the input space, namely $n$. By definition, $\text{rank}(A) = \dim \text{CS}(A) = n$. (Alternatively, this follows from $\dim \text{NS}(A) + \text{rank}(A) = n$.) The only $n$-dimensional subspace of $\mathbb{R}^n$ is $\mathbb{R}^n$ itself, so $\text{CS}(A) = \mathbb{R}^n$. Thus every $b$ is in $\text{CS}(A)$, so $Ax = b$ has a solution for every $b$. The system $Ax = b$ has the same number of solutions as $Ax = 0$ (they are just shifted by adding a particular solution $x_p$); that number is 1 (the only solution to $Ax = 0$ is $0$). To say that $Ax = b$ has exactly one solution for each $b$ means that the associated linear transformation $f$ is a 1-to-1 correspondence, so $f^{-1}$ exists, so $A^{-1}$ exists. (Moreover, we showed how to find $A^{-1}$ by Gauss–Jordan elimination.) We have $\text{RREF}(A) = I$ as explained earlier, since $I$ is the only RREF square matrix with nonzero determinant.

Midterm 2 covers everything up to here.

---

March 14

---

13.4.2. Singular matrices. The same theorem can be stated in terms of the complementary conditions (it’s essentially the same theorem, so this is really just review):

**Theorem 13.9.** For a square matrix $A$, the following are equivalent:

1. $\det A = 0$ (scaling factor is 0)
2. $\text{NS}(A)$ is larger than $\{0\}$ (i.e., $Ax = 0$ has a nonzero solution)
3. $\text{rank}(A) < n$ (image has dimension less than $n$)
4. $\text{CS}(A)$ is smaller than $\mathbb{R}^n$ (image is not the whole space $\mathbb{R}^n$)
The system $Ax = b$ has no solutions for some vectors $b$, and infinitely many solutions for other vectors $b$.

A$^{-1}$ does not exist.

$\text{RREF}(A) \neq I$

Now let’s explain the consequences of $\det A = 0$.

Suppose that $\det A = 0$. Then the volume scaling factor is 0, so the input space is flattened by $A$. This means that some input dimensions are getting crushed, so $\text{NS}(A)$ is larger than $\{0\}$ (at least 1-dimensional), and the image is smaller than the $n$-dimensional input space: $\text{rank}(A) < n$. In particular, the image $\text{CS}(A)$ is not all of $\mathbb{R}^n$.

- If $b \notin \text{CS}(A)$, then $Ax = b$ has no solution.
- If $b \in \text{CS}(A)$, then $Ax = b$ has the same number of solutions as $Ax = 0$, i.e., infinitely many since $\dim \text{NS}(A) \geq 1$.

The associated linear transformation $f$ is not a 1-to-1 correspondence (it maps many vectors to 0); thus $f^{-1}$ does not exist, so $A^{-1}$ does not exist. Row operations do not change the condition $\det A = 0$, so $\det \text{RREF}(A) = 0$, so definitely $\text{RREF}(A) \neq I$. (In fact, $\text{RREF}(A)$ must have at least one 0 along the diagonal.)

**Problem 13.10.** Devise a test for deciding whether a homogeneous square system $Ax = 0$ has a nonzero solution.

**Solution:** Compute $\det A$. If $\det A = 0$, there exists a nonzero solution. If $\det A \neq 0$, then $Ax = 0$ has only the zero solution. □

13.5. **Review of RLC circuits.**

**Problem 13.11.** A resistor of resistance $R$ and an inductor of inductance $L$ are attached in parallel. A voltage source provides the combination with AC voltage of angular frequency $\omega$. Find the gain and phase lag of the resistor current relative to the total current (through the voltage source).

**Solution:**

Let $V(t)$ be the sinusoidal voltage provided.

**Unknown functions:** Let $I_1$ be the resistor current. Let $I_2$ be the inductor current. Let $I$ be the total current.

**Equations:** Physics says

\[
V = RI_1 \\
V = LI_2 \\
I = I_1 + I_2.
\]
The same relationships hold between the complex replacements:

\[ \tilde{V} = R\tilde{I}_1 \]
\[ \tilde{V} = L\dot{\tilde{I}}_2 \]
\[ \tilde{I} = \tilde{I}_1 + \dot{\tilde{I}}_2 \]

(because taking real parts is compatible with real scalar multiplication and with taking derivatives).

Suppose that \( \tilde{V} = e^{i\omega t} \). (In general, \( \tilde{V} = \gamma e^{i\omega t} \) for some \( \gamma \in \mathbb{C} \), but then everything will be multiplied by \( \gamma \), so when we take a ratio to get complex gain, \( \gamma \) will disappear.) How do we solve for the other three functions \( \tilde{I}_1, \tilde{I}_2, \tilde{I} \)?

In the steady-state solution,

\[ \tilde{I}_1 = \alpha_1 e^{i\omega t} \]
\[ \tilde{I}_2 = \alpha_2 e^{i\omega t} \]
\[ \tilde{I} = \beta e^{i\omega t} \]

for some unknown complex numbers \( \alpha_1, \alpha_2, \beta \). To find \( \alpha_1, \alpha_2, \beta \), substitute into the three complex replacement equations:

\[ e^{i\omega t} = R\alpha_1 e^{i\omega t} \]
\[ e^{i\omega t} = L\alpha_2 e^{i\omega t} i\omega \]
\[ \beta e^{i\omega t} = \alpha_1 e^{i\omega t} + \alpha_2 e^{i\omega t} \]

This simplifies to

\[ 1 = R\alpha_1 \]
\[ 1 = L\alpha_2 i\omega \]
\[ \beta = \alpha_1 + \alpha_2. \]

So \( \alpha_1 = 1/R, \alpha_2 = 1/(L\omega), \beta = 1/R + 1/(L\omega) \). The complex gain of \( I_1 \) relative to \( I \) is the complex constant

\[ G := \frac{\alpha_1}{\beta} = \frac{1/R}{1/R + 1/(L\omega)}. \]

The gain is \( |G| \), and the phase lag is \(-\arg G\). \( \square \)
Guest lecture by Professor Haynes Miller. I’m told it will involve a movie about how the math we’ve learned is used in a fluorescent light.

Notes for this lecture:

[http://math.mit.edu/~poonen/03/guest-19mar.pdf](http://math.mit.edu/~poonen/03/guest-19mar.pdf)

March 21

13.6. **Trace.** This section was covered in March 20 recitation.

**Definition 13.12.** The **trace** of a square matrix $A$ is the sum of the entries along the diagonal. It is denoted $\text{tr} A$.

**Example 13.13.** If $A = \begin{pmatrix} 4 & 6 & 9 \\ 1 & 7 & 8 \\ 2 & 3 & 5 \end{pmatrix}$, then $\text{tr} A = 4 + 7 + 5 = 16$.

**Warning:** Trace and determinant make sense only for **square** matrices.

**Problem 13.14.** For an $n \times n$ matrix $A$, how do $\text{tr}(-A)$ and $\det(-A)$ relate to $\text{tr} A$ and $\det A$?

**Solution:** Negating $A$ negates in particular all diagonal entries of $A$, so $\text{tr}(-A) = - \text{tr} A$.

On the other hand, negating $A$ amounts to multiplying every row by $-1$, which multiplies $\det A$ by $(-1)^n$ because there is one factor of $-1$ for each row. Thus

- If $n$ is even, then $\det(-A) = \det A$.
- If $n$ is odd, then $\det(-A) = - \det A$.

13.7. **Characteristic polynomial of a matrix.** This section was covered in March 20 recitation.

Use $\lambda$ to denote a scalar-valued variable.

**Definition 13.15.** The **characteristic polynomial** of a square matrix $A$ is $\det(\lambda I - A)$; this is a monic degree $n$ polynomial in the variable $\lambda$ (monic means that the leading coefficient is 1, so the polynomial looks like $\lambda^n + \ldots$).

**Warning:** Some authors use $\det(A - \lambda I)$ instead. This is the same, except negated when $n$ is odd.

**Warning:** This is not the same concept as the characteristic polynomial of a constant-coefficient linear ODE, but there is a connection, arising when such a DE is converted to a first-order system of linear ODEs.

**Steps to compute the characteristic polynomial of an $n \times n$ matrix $A$:**

1. Write down $A - \lambda I$. (This usually involves less writing than does $\lambda I - A$.)
2. Compute $\det(A - \lambda I)$. 

90
3. If \( n \) is odd, change the sign of the result to get \( \det(\lambda I - A) \). (The result should have the form \( \lambda^n + \cdots \).)

**Problem 13.16.** What is the characteristic polynomial of \( A := \begin{pmatrix} 7 & 2 \\ 3 & 5 \end{pmatrix} \)?

**Solution:** We have

\[
A - \lambda I = \begin{pmatrix} 7 - \lambda & 2 \\ 3 & 5 - \lambda \end{pmatrix}
\]

\[
\det(A - \lambda I) = (7 - \lambda)(5 - \lambda) - 2(3) = \lambda^2 - 12\lambda + 29
\]

\[
\det(\lambda I - A) = \lambda^2 - 12\lambda + 29.
\]

Here is a shortcut for \( 2 \times 2 \) matrices:

**Theorem 13.17.** If \( A \) is a \( 2 \times 2 \) matrix, then the characteristic polynomial of \( A \) is

\[
\lambda^2 - (\text{tr} A)\lambda + (\det A).
\]

**Proof.** Write \( A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \). Then the characteristic polynomial is

\[
\det(\lambda I - A) = \det(A - \lambda I)
\]

\[
= \det \begin{pmatrix} a - \lambda & b \\ c & d - \lambda \end{pmatrix}
\]

\[
= (a - \lambda)(d - \lambda) - bc
\]

\[
= \lambda^2 - (a + d)\lambda + (ad - bc)
\]

\[
= \lambda^2 - (\text{tr} A)\lambda + (\det A).
\]

We can solve Problem 13.16 again, using this shortcut: the matrix \( A := \begin{pmatrix} 7 & 2 \\ 3 & 5 \end{pmatrix} \) has \( \text{tr} A = 12 \) and \( \det A = 29 \), so the characteristic polynomial of \( A \) is \( \lambda^2 - 12\lambda + 29 \).

**Remark 13.18.** Suppose that \( n > 2 \). Then, for an \( n \times n \) matrix \( A \), the characteristic polynomial has the form

\[
\lambda^n - (\text{tr} A)\lambda^{n-1} + \cdots \pm \det A
\]

where the \( \pm \) is + if \( n \) is even, and \( - \) if \( n \) is odd. So knowing \( \text{tr} A \) and \( \det A \) determines some of the coefficients of the characteristic polynomial, but not all of them.

**Question 13.19.** What is the characteristic polynomial of an upper triangular matrix \( A \)?
Answer: If \( a_{11}, \ldots, a_{nn} \) are the diagonal entries of \( A \), then \( \lambda I - A \) is another upper triangular matrix, but with diagonal entries \( \lambda - a_{11}, \ldots, \lambda - a_{nn} \). Thus the characteristic polynomial of \( A \) is
\[
\det(\lambda I - A) = (\lambda - a_{11}) \cdots (\lambda - a_{nn}).
\]


13.8.1. Definition.

Definition 13.20. Suppose that \( A \) is an \( n \times n \) matrix.

- An eigenvector of \( A \) is a nonzero vector \( v \) such that \( Av = \lambda v \) for some scalar \( \lambda \).
  (Warning: Some authors consider \( 0 \) to be an eigenvector too, but we will not.)
- An eigenvalue of \( A \) is a scalar \( \lambda \) such that \( Av = \lambda v \) for some nonzero vector \( v \).

Try the “Matrix Vector” mathlet
http://mathlets.org/mathlets/matrix-vector/

Problem 13.21. Let \( A = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix} \) and let \( v = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \). Is \( v \) an eigenvector of \( A \)?

Solution: The calculation
\[
Av = \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 4 \\ -2 \end{pmatrix} = 2v,
\]
shows that \( v \) is an eigenvector, and that the associated eigenvalue is 2. \( \square \)

13.8.2. Zero as an eigenvalue. It is possible for 0 to be an eigenvalue. The following are equivalent for a square matrix \( A \):

- 0 is an eigenvalue.
- There exists a nonzero solution to \( Ax = 0 \).
- \( \det A = 0 \).
- ...

13.8.3. Five as an eigenvalue. The following are equivalent for a square matrix \( A \):

- 5 is an eigenvalue.
- There exists a nonzero solution to
\[
Ax = 5x
\]
\[
Ax = 5Ix
\]
\[
Ax - 5Ix = 0
\]
\[
(A - 5I)x = 0.
\]
• \( \det(A - 5I) = 0 \).
• \( \det(5I - A) = 0 \).

### 13.8.4. Finding all the eigenvalues.

**Theorem 13.22.** Let \( A \) be a square matrix, and let \( \lambda \) be a scalar. Then \( \lambda \) is an eigenvalue of \( A \) if and only if \( \det(\lambda I - A) = 0 \).

**Steps to find all the eigenvalues of a square matrix \( A \):**

1. Calculate the characteristic polynomial \( \det(\lambda I - A) \) or \( \det(A - \lambda I) \).
2. The roots of this polynomial are all the eigenvalues of \( A \).

**Problem 13.23.** Find all the eigenvalues of \( A := \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix} \).

**Solution:** We have \( \text{tr} \ A = 1 + 0 = 1 \) and \( \det \ A = 0 - 2 = -2 \), so the characteristic polynomial is

\[
p(\lambda) = \lambda^2 - (\text{tr} \ A)\lambda + (\det \ A) = \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1).
\]

Its roots are 2 and -1; these are the eigenvalues. \( \square \)

The **multiplicity** of an eigenvalue is just its multiplicity as a root of the characteristic polynomial.

**Problem 13.24.** Find the eigenvalues of the upper triangular matrix \( A := \begin{pmatrix} 2 & 3 & 5 \\ 0 & 2 & 7 \\ 0 & 0 & 6 \end{pmatrix} \).

**Solution:** The characteristic polynomial is

\[
\det(\lambda I - A) = \det \begin{pmatrix} \lambda - 2 & -3 & -5 \\ 0 & \lambda - 2 & -7 \\ 0 & 0 & \lambda - 6 \end{pmatrix} = (\lambda - 2)(\lambda - 2)(\lambda - 6),
\]

so the eigenvalues, listed with multiplicity, are 2, 2, 6. \( \square \)

In general, for any upper triangular or lower triangular matrix, the eigenvalues are the diagonal entries.
13.8.5. Eigenspaces.

**Definition 13.25.** The *eigenspace* of an eigenvalue \( \lambda \) of a square matrix \( A \) is the set of all eigenvectors having that eigenvalue, together with the zero vector \( \mathbf{0} \).

So each eigenspace is a set of vectors. In fact, each eigenspace is a *vector space*. Why? It is the set of all solutions to \( A\mathbf{x} = \lambda \mathbf{x} \) (including \( \mathbf{x} = \mathbf{0} \)), or equivalently to \( (A - \lambda I)\mathbf{x} = \mathbf{0} \). Thus the eigenspace of \( \lambda \) is the same as \( \text{NS}(A - \lambda I) \), which is a vector space.

**Steps to find all the eigenvectors associated to a given eigenvalue \( \lambda \) of a square matrix \( A \):**

1. Calculate \( A - \lambda I \).
2. Use Gaussian elimination and back-substitution to compute a basis of \( \text{NS}(A - \lambda I) \).
3. The eigenvectors having eigenvalue \( \lambda \) are all the linear combinations of those basis vectors (not including the zero vector).

**Problem 13.26.** Describe all the eigenvalues, eigenvectors, and eigenspaces of \( A := \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix} \).

**Solution:**

Eigenvalues: \( 2 \) and \( -1 \). (We already found these, by determining the roots of the characteristic polynomial).

**Eigenspace of 2:** this is \( \text{NS}(A - 2I) \), i.e., the set of all solutions to

\[
(A - 2I)\mathbf{x} = \mathbf{0} \\
\begin{pmatrix} -1 & -2 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{0}.
\]

In harder examples, one would need Gaussian elimination to proceed, but this linear system is easy to solve since the equations are the same:

\[-x - 2y = 0.\]

A very easy back-substitution shows that the general solution is \( \begin{pmatrix} -2c \\ c \end{pmatrix} = c \begin{pmatrix} -2 \\ 1 \end{pmatrix} \), where \( c \) is a parameter. Thus \( \begin{pmatrix} -2 \\ 1 \end{pmatrix} \) by itself is a basis for the eigenspace of 2.

**Eigenspace of -1:** this is \( \text{NS}(A + I) \). Since \( A + I = \begin{pmatrix} 2 & -2 \\ -1 & 1 \end{pmatrix} \), for which a row-echelon form is \( \begin{pmatrix} 2 & -2 \\ 0 & 0 \end{pmatrix} \), this eigenspace is the space of solutions to

\[
2x - 2y = 0.
\]
The general solution is \( \begin{pmatrix} c \\ c \end{pmatrix} = c \begin{pmatrix} 1 \\ 1 \end{pmatrix} \), where \( c \) is a parameter. Thus \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) by itself is a basis for the eigenspace of \(-1\).

**Conclusions:**

- The eigenvalues are 2 and \(-1\).
- The eigenvectors with eigenvalue 2 are the nonzero scalar multiples of \( \begin{pmatrix} -2 \\ 1 \end{pmatrix} \). These together with \( \mathbf{0} \) form the eigenspace of 2.
- The eigenvectors with eigenvalue \(-1\) are the nonzero scalar multiples of \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \). These together with \( \mathbf{0} \) form the eigenspace of \(-1\). \( \square \)

**Flashcard question:** Which of the following statements is true of the \( n \times n \) identity matrix \( I \)?

**Possible answers:**

- Every vector in \( \mathbb{R}^n \) is an eigenvector of \( I \).
- No vector in \( \mathbb{R}^n \) is an eigenvector of \( I \).
- Every eigenvalue of \( I \) is 0.
- None of the above.

**Answer:** None of the above, but only because of a technicality. It is true that \( Ix = x \) for every vector \( x \), so every nonzero vector is an eigenvector with eigenvalue 1. But our convention is that \( \mathbf{0} \) never qualifies as an eigenvector!


**Problem 13.27.** Find the eigenvalues and eigenvectors of the \( 90^\circ \) counterclockwise rotation matrix \( R = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \).

**Solution:** Since \( \text{tr} \ R = 0 \) and \( \det R = 1 \), the characteristic polynomial of \( R \) is \( \lambda^2 + 1 \). Its roots are \( i \) and \(-i\); these are the eigenvalues.

The eigenspace of \( i \) is \( \text{NS}(R - iI) \). Converting

\[
R - iI = \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}
\]

\[
= \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}
\begin{pmatrix} -i \\ 1 \end{pmatrix}
\]
to row-echelon form gives \( \begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix} \), so we solve \(-ix - y = 0\) by back-substitution and get the general solution \( c \begin{pmatrix} i \\ 1 \end{pmatrix} \). Thus the eigenvectors having eigenvalue \( i \) are the nonzero scalar multiples of \( \begin{pmatrix} i \\ 1 \end{pmatrix} \).

Applying complex conjugation to the entire universe shows that the eigenvectors having eigenvalue \(-i\) are the nonzero scalar multiples of \( \begin{pmatrix} -i \\ 1 \end{pmatrix} \).

For any \( n \times n \) matrix, the characteristic polynomial is of degree \( n \), so the fundamental theorem of algebra shows that the total number of complex eigenvalues counted with multiplicity is \( n \).

13.8.7. Dimension of an eigenspace.

**Theorem 13.28.** Let \( \lambda \) be an eigenvalue of an \( n \times n \) matrix \( A \). Suppose that the multiplicity of \( \lambda \) (as a root of the characteristic polynomial) is \( m \). Then

\[
1 \leq \text{(dimension of eigenspace of } \lambda \text{)} \leq m.
\]

Given \( \lambda \), the dimension of the eigenspace of \( \lambda \) is also the maximum number of linearly independent eigenvectors of eigenvalue \( \lambda \) that can be found. This dimension is at least 1 since \( A \) has at least one eigenvector of eigenvalue \( \lambda \) (otherwise \( \lambda \) would not have been an eigenvalue). That this dimension is at most \( m \) requires more work to prove, and we’re not going to do it in this class.

**Problem 13.29.** A \( 9 \times 9 \) matrix has characteristic polynomial \((\lambda - 2)^3(\lambda - 5)^6\). What are the possibilities for the dimension of the eigenspace of 2?

**Solution:** In this case, \( m = 3 \), so the dimension is 1, 2, or 3.

**Definition 13.30.** The eigenspace of \( \lambda \) is called **complete** if its dimension equals the multiplicity \( m \) of \( \lambda \), and **deficient** if its dimension is less than \( m \). **Warning:** Different authors use different terminology here.

**Example 13.31.** If the multiplicity is 1, then the dimension of the eigenspace is sandwiched between 1 and 1, so the eigenspace is not deficient.

**Definition 13.32.** A matrix is **deficient** if one of its eigenspaces is deficient.

For the application to solving linear systems of ODEs, given an \( n \times n \) matrix \( A \) we will want to find as many linearly independent eigenvectors as possible. To do this, we choose a
basis of each eigenspace, and concatenate these lists of eigenvectors; it turns out that the resulting list is linearly independent.

How many eigenvectors are in this list?

- If all the eigenspaces are complete, then the number of linearly independent eigenvectors from each eigenspace is the multiplicity of the eigenvalue, so the total number of eigenvectors in our list is the total number of eigenvalues counted with multiplicity, which is $n$. In this case, the $n$ eigenvectors form a basis of $\mathbb{C}^n$ (since their span is $n$-dimensional).
- If instead $A$ is deficient, then the number of linearly independent eigenvectors is less than $n$.

Why does concatenating the bases produce a linearly independent list? The vectors within each basis are linearly independent, and there are no linear relations involving eigenvectors from different eigenspaces because of the following:

**Theorem 13.33.** Fix a square matrix $A$. Eigenvectors corresponding to distinct eigenvalues are linearly independent.

**Proof.** Suppose that $v_1, \ldots, v_n$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \ldots, \lambda_n$. Suppose that there were a linear relation

$$c_1 v_1 + \cdots + c_n v_n = 0.$$

Apply $A - \lambda_1 I$ to both sides; this gets rid of the first summand on the left. Next apply $A - \lambda_2 I$, and so on, up to $A - \lambda_{n-1} I$. This shows that some nonzero number times $c_n v_n$ equals 0. But $v_n \neq 0$, so $c_n = 0$. Similarly each $c_i$ must be 0. Thus only the trivial relation between $v_1, \ldots, v_n$ exists, so they are linearly independent. □

13.8.8. Examples. Here are three examples showing all the situations that can arise for a $2 \times 2$ matrix.

**Example 13.34.** For $A := \begin{pmatrix} 1 & -2 \\ -1 & 0 \end{pmatrix}$, the characteristic polynomial is $(\lambda - 2)(\lambda + 1)$, so the eigenvalues have multiplicity 1, so the eigenspaces are automatically complete. The eigenvectors $\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ are linearly independent.

**Example 13.35.** Let $B = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$. Since $B$ is upper triangular (even diagonal), the eigenvalues are 5, 5. The eigenspace of 5 is $\text{NS}(B - 5I)$, which is the set of solutions to $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} x = 0$, which is the entire space $\mathbb{C}^2$. Its dimension (namely, 2) matches the multiplicity of the eigenvalue 5, so this eigenspace is complete. Every nonzero vector in $\mathbb{C}^2$ is an eigenvector...
with eigenvalue 5. So it is easy to find two linearly independent eigenvectors: for example, take \(
\begin{pmatrix} 1 \\ 0 \end{pmatrix}
\) and \(
\begin{pmatrix} 0 \\ 1 \end{pmatrix}
\).

Example 13.36. Let \( C = \begin{pmatrix} 5 & 3 \\ 0 & 5 \end{pmatrix} \). Again the eigenvalues are 5, 5. The eigenspace of 5 is \( \text{NS}(C - 5I) \), which is the set of solutions to \(
\begin{pmatrix} 0 & 3 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0
\). This system consists of a single nontrivial equation \( 3y = 0 \). Thus the eigenspace is the set of vectors of the form \( \begin{pmatrix} c \\ 0 \end{pmatrix} \); it is only 1-dimensional, even though the multiplicity of the eigenvalue 5 is still 2. This means that this eigenspace is deficient, and hence \( C \) is deficient. It is impossible to find two linearly independent eigenvectors.

March 31

14. Homogeneous linear systems of ODEs


Flashcard question: Consider the system

\[
\begin{align*}
\dot{x} &= 2t^2 x + 3y \\
\dot{y} &= 5x - 7e^t y
\end{align*}
\]

involving two unknown functions, \( x(t) \) and \( y(t) \). Which of the following describes this system?

Possible answers:

- first-order homogeneous linear system of ODEs
- second-order homogeneous linear system of ODEs
- first-order inhomogeneous linear system of ODEs
- second-order inhomogeneous linear system of ODEs
- first-order homogeneous linear system of PDEs
- second-order homogeneous linear system of PDEs
- first-order inhomogeneous linear system of PDEs
- second-order inhomogeneous linear system of PDEs

Answer: It’s a first-order homogeneous linear system of ODEs. The equations are ODEs since the functions are still functions of only one variable, \( t \). This is a homogeneous linear system since every summand is a function of \( t \) times one of \( x, \dot{x}, \ldots, y, \dot{y}, \ldots \). (If there were also terms that were functions of \( t \), then it would be an inhomogeneous linear system.) The system is first-order since it involves only the first derivatives of the unknown functions. □
The system can be written in matrix form,
\[ \dot{x} = A(t) x, \]
by defining
\[ x := \begin{pmatrix} x \\ y \end{pmatrix}, \quad A(t) := \begin{pmatrix} 2t^2 & 3 \\ 5 & -7e^t \end{pmatrix}. \]

14.2. Theory.

Existence and uniqueness theorem for a linear system of ODEs. Let \( A(t) \) be a matrix-valued function and let \( q(t) \) be a vector-valued function, both continuous on an open interval \( I \). Let \( a \in I \), and let \( b \) be a vector. Then there exists a unique solution \( x(t) \) to the system
\[ \dot{x} = A(t) x + q(t) \]
satisfying the initial condition \( x(a) = b \).

(Of course, the sizes of these matrices and vectors should match in order for this to make sense.)

If \( x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \), then the initial condition \( x(a) = b \) is really \( n \) initial conditions.

Dimension theorem for a homogeneous linear system of ODEs. For any first-order homogeneous linear system of \( n \) ODEs in \( n \) unknown functions
\[ \dot{x} = A(t) x, \]
the set of solutions is an \( n \)-dimensional vector space.

Once the system \( \dot{x} = A(t) x \) and the starting time \( a \) are fixed, the solutions to the system correspond to the possibilities for the initial condition vector \( b \); that is why the set of solutions is \( n \)-dimensional. Moreover, solutions are linearly independent if and only if their values at the starting time \( a \) are linearly independent vectors.

Complex basis vs. real-valued basis. If a complex vector-valued function \( x \) and its complex conjugate \( \bar{x} \) are part of a basis of solutions to a homogeneous linear system of ODEs with real coefficients, then

replacing \( x, \bar{x} \) by \( \text{Re}(x), \text{Im}(x) \)
gives a new basis.

Inhomogeneous principle for a linear system of ODEs. If \( x_p \) is one particular solution to the inhomogeneous linear system
\[ \dot{x} = A(t)x + q(t) \]

and $x_h$ is the general solution to the associated homogeneous linear system

$$\dot{x} = A(t)x$$

then $x_p + x_h$ is the general solution to the inhomogeneous linear system.

14.3. **Solving homogeneous linear systems of ODEs.** Consider a first-order homogeneous linear system of ODEs with constant coefficients:

$$\dot{x} = Ax,$$

where $A$ is an $n \times n$ matrix with constant entries.

**Question 14.1.** For which pairs $(\lambda, v)$ consisting of a scalar and a nonzero vector is the vector-valued function $x = e^{\lambda t}v$ a solution to the system $\dot{x} = Ax$?

**Solution:** Plug it in, to see what has to happen in order for it to be a solution:

$$\lambda e^{\lambda t}v = Ae^{\lambda t}v \quad (\text{for all } t).$$

This is equivalent to

$$\lambda v = Av$$

(to go forwards, set $t = 0$; to go back, multiply both sides by $e^{\lambda t}$).

**Conclusion:** $e^{\lambda t}v$ is a solution if and only if $v$ is an eigenvector of $A$ with eigenvalue $\lambda$. \qed

**Steps to find a basis of solutions to** $\dot{x} = Ax$, **given an** $n \times n$ **matrix** $A$:

1. Find the eigenvalues of $A$ (they are the roots of the characteristic polynomial $\det(\lambda I - A)$).
2. For each eigenvalue $\lambda$, find a basis for the corresponding eigenspace $\text{NS}(A - \lambda I)$; these basis vectors will be eigenvectors $v$.
3. For each eigenvalue $\lambda$ and each such $v$, the vector-valued function $e^{\lambda t}v$ is one solution.
4. If $n$ such solutions were found (i.e., the sum of the dimensions of the eigenspaces is $n$), then these are enough solutions to form a basis.

**Remark 14.2.** The solutions of this type will automatically be linearly independent, since their values at $t = 0$ are linearly independent (the chosen eigenvectors within each eigenspace are independent, and there are no linear dependences between eigenvectors with different eigenvalues).

**Remark 14.3.** The only thing that could go wrong is this: if there is a repeated eigenvalue $\lambda$, and the dimension of the eigenspace of $\lambda$ is less than the multiplicity of $\lambda$, then the method above does not produce enough solutions. In this case, more complicated functions are needed, such as $te^{\lambda t}u + e^{\lambda t}v$ and so on.
The simple solutions forming a basis are sometimes called normal modes. There is not a precise mathematical definition of normal mode, however, since what counts as simple is subjective.

**Problem 14.4.** Find the general solution \((x(t), y(t), z(t))\) to the system

\[
\begin{align*}
\dot{x} &= 2x \\
\dot{y} &= -6x + 8y + 3z \\
\dot{z} &= 18x - 18y - 7z.
\end{align*}
\]

**Solution:** In matrix form, this is \(\dot{x} = Ax\), where

\[
A = \begin{pmatrix} 2 & 0 & 0 \\ -6 & 8 & 3 \\ 18 & -18 & -7 \end{pmatrix}.
\]

**Step 1. Find the eigenvalues.** To do this, compute

\[
\det(A - \lambda I) = \det \begin{pmatrix} 2 - \lambda & 0 & 0 \\ -6 & 8 - \lambda & 3 \\ 18 & -18 & -7 - \lambda \end{pmatrix}.
\]

Use Laplace expansion along the first row to get

\[
(2 - \lambda)((8 - \lambda)(-7 - \lambda) - (-18)3) = (2 - \lambda)(-2 - \lambda + \lambda^2) = (2 - \lambda)(\lambda - 2)(\lambda + 1),
\]

so the eigenvalues are \([2, 2, -1]\).

**Step 2. Find a basis of each eigenspace.**

**Eigenspace at 2:** This is the nullspace of

\[
A - 2I = \begin{pmatrix} 0 & 0 & 0 \\ -6 & 6 & 3 \\ 18 & -18 & -9 \end{pmatrix}.
\]

Converting to row-echelon form gives

\[
\begin{pmatrix} -6 & 6 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},
\]

which corresponds to the single equation

\[-6p + 6q + 3r = 0\]
(the letters $x, y, z$ are already taken). Solve by back-substitution: \( r = c_1, \quad q = c_2, \quad p = q + r/2 = c_2 + c_1/2, \) so
\[
\begin{pmatrix} p \\ q \\ r \end{pmatrix} = c_1 \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}
\]
so \( \begin{pmatrix} 1/2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \) form a basis for the eigenspace at 2. (We were lucky here that the number of basis eigenvectors is as large as the multiplicity of the eigenvalue, so that the eigenspace of 2 was not deficient.)

**Eigenspace at -1:** This is the nullspace of
\[
A + I = \begin{pmatrix} 3 & 0 & 0 \\ -6 & 9 & 3 \\ 18 & -18 & -6 \end{pmatrix}
\]
Converting to row-echelon form gives
\[
\begin{pmatrix} 3 & 0 & 0 \\ 0 & 9 & 3 \\ 0 & 0 & 0 \end{pmatrix},
\]
which corresponds to the system
\[
3p = 0 \\
9q + 3r = 0.
\]
Back-substitution leads to
\[
\begin{pmatrix} p \\ q \\ r \end{pmatrix} = c \begin{pmatrix} 0 \\ -1/3 \\ 1 \end{pmatrix},
\]
so \( \begin{pmatrix} 0 \\ -1/3 \\ 1 \end{pmatrix} \) by itself is a basis for the eigenspace at -1.

**Steps 3 and 4. Build solutions from the eigenvalue-eigenvector pairs, and check whether there are enough to form a basis.**

We have three independent solutions,
\[
e^{2t} \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix}, \quad e^{2t} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad e^{-t} \begin{pmatrix} 0 \\ -1/3 \\ 1 \end{pmatrix},
\]
so they are a basis of solutions. The general solution is
\[
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = c_1 e^{2t} \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix} + c_2 e^{2t} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c_3 e^{-t} \begin{pmatrix} 0 \\ -1/3 \\ 1 \end{pmatrix}. \]

If there were initial conditions, we could solve for \( c_1, c_2, c_3 \) to get a specific solution.

### 14.4. Higher order ODEs vs. systems of ODEs.

An \( n^{\text{th}} \) order ODE can be converted to a first-order system of ODEs by introducing new function variables for the derivatives.

**Problem 14.5.** Convert \( \ddot{x} + 5\dot{x} + 6x = 0 \) to a first-order system of ODEs.

**Solution:** Define \( y := \dot{x} \). Then
\[
\dot{x} = y, \\
\dot{y} = \ddot{x} = -5\dot{x} - 6x = -6x - 5y.
\]

In matrix form, this is \( \dot{x} = Ax \) with \( A = \begin{pmatrix} 0 & 1 \\ -6 & -5 \end{pmatrix} \).

(The matrix \( \begin{pmatrix} 0 & 1 \\ -6 & -5 \end{pmatrix} \) arising this way is called the *companion matrix* of the polynomial \( x^2 + 5x + 6 \).)

Conversely, given a first-order system, one can eliminate function variables to find a higher-order ODE satisfied by one of the functions. But usually it is better just to leave it as a system.

**Problem 14.6.** Given that
\[
\dot{x} = 2x - y, \\
\dot{y} = 5x + 7y,
\]
find a higher-order ODE involving only \( x \).

**Solution:** Solve for \( y \) in the first equation \( y = 2x - \dot{x} \) and substitute into the second:
\[
2\dot{x} - \ddot{x} = 5x + 7(2x - \dot{x}).
\]
This simplifies to
\[
\ddot{x} - 9\dot{x} + 19x = 0. \]

**Remark 14.7.** First-order systems with more than two equations can be converted too, but the conversion is not so easy.
Remark 14.8. For constant-coefficient ODEs, the characteristic polynomial of the higher-order ODE (scaled, if necessary, to have leading coefficient 1) equals the characteristic polynomial of the matrix of the first-order system.

Remark 14.9. One can also convert higher-order systems of ODEs to first-order systems. For example, a system of 4 fifth-order ODEs can be converted to a system of 20 first-order ODEs. That’s why it’s enough to study first-order systems.

April 2

14.5. Two-dimensional dynamics.

Question 14.10. How can you visualize a real-valued solution \( \langle x(t), y(t) \rangle \) to \( \dot{x} = A x \) for a constant \( 2 \times 2 \) real matrix \( A \)?

Answer: Make two plots, the first showing \( x(t) \) as a function of \( t \), and the second showing \( y(t) \) as a function of \( t \).

Better answer: Draw the solution as a parametrized curve \( \langle x(t), y(t) \rangle \) in the phase plane with axes \( x \) and \( y \). In other words, plot the point \( (x(t), y(t)) \) for every real number \( t \) (including negative \( t \)). The ODE specifies, in terms of the current position, which direction the phase plane point will move next (and how fast).

The phase plane trajectory by itself does not describe a solution fully, however, since it does not show at what time each point is reached. The trajectory contains no information about speed, though one can specify the direction by drawing an arrow on the trajectory.

Flashcard question: One of the solutions to \( \dot{x} = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} x \) is

\[
\begin{pmatrix} x \\ y \end{pmatrix} = e^{-t} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + e^{-2t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{-t} \\ e^{-2t} \end{pmatrix}.
\]

Which of the following describes the motion in the phase plane with axes \( x \) and \( y \) as \( t \to +\infty \)?

Possible answers:

- approaching infinity, along a curve asymptotic to the \( x \)-axis
- approaching infinity, along a curve asymptotic to the \( y \)-axis
- approaching infinity, along a straight line
- approaching the origin, along a curve tangent to the \( x \)-axis
- approaching the origin, along a curve tangent to the \( y \)-axis
- approaching the origin, along a straight line
- spiraling
- none of the above
Answer: Approaching the origin, along a curve tangent to the x-axis. As $t \to +\infty$, both $x = e^{-t}$ and $y = e^{-2t}$ tend to $(0,0)$, but the $y$-coordinate tends to 0 faster than the $x$-coordinate, so the trajectory is tangent to the $x$-axis. (In fact, the $y$-coordinate is always the square of the $x$-coordinate, so the trajectory is part of the parabola $y = x^2$.) □

**Question 14.11.** If $v \in \mathbb{R}^2$ is an eigenvector of $A$, with real eigenvalue $\lambda$, then $e^{\lambda t}v$ is a solution to $\dot{x} = Ax$. Each value of this solution is a positive scalar multiple of $v$, so the trajectory is contained in the ray through $v$. What is the direction of the trajectory?

**Answer:**
- If $\lambda > 0$, the phase point tends to infinity (repelled from $(0,0)$).
- If $\lambda < 0$, the phase point tends to $(0,0)$ (attracted to $(0,0)$).
- If $\lambda = 0$, the phase point is stationary at $v$! The point $v$ is called an critical point since $\dot{x} = 0$ there. □

The phase portrait is the diagram showing all the trajectories in the phase plane. We are now ready to classify all possibilities for the phase portrait in terms of the eigenvalue behavior. The most common ones are indicated in green; the others are degenerate cases.

Try the “Linear Phase Portraits: Matrix Entry” mathlet

http://mathlets.org/mathlets/linear-phase-portraits-matrix-entry/

See if you can get every possibility listed below.

14.5.1. **Distinct real eigenvalues.** Suppose that the eigenvalues $\lambda_1, \lambda_2$ are real and distinct. Let $v_1, v_2$ be corresponding eigenvectors. Each eigenspace is 1-dimensional, spanned by $v_1$ or $v_2$, and will be called an eigenline. General solution:

$$c_1 e^{\lambda_1 t}v_1 + c_2 e^{\lambda_2 t}v_2.$$

**Opposite sign:** $\lambda_1 > 0, \lambda_2 < 0$. This is called a saddle. Trajectories flow outward along the positive eigenline (the eigenspace of $\lambda_1$) and inward along the negative eigenline (the
eigenspace of $\lambda_2$). Other trajectories are asymptotic to both eigenlines, tending to infinity towards the positive eigenline. (Typical solution: $x = e^{2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^{-3t} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. When $t = +1000$, this is approximately a large positive multiple of $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. When $t = -1000$, this is approximately a large positive multiple of $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.)

In the next two cases, in which the eigenvalues have the same sign, we’ll want to know which is bigger. If $|\lambda_1| > |\lambda_2|$, call $\lambda_1$ the fast eigenvalue and $\lambda_2$ the slow eigenvalue; use the same adjectives for the eigenlines.

**Both positive:** $\lambda_1, \lambda_2 > 0$. This is called a repelling node (or node source). All nonzero trajectories flow from $(0, 0)$ towards infinity. Trajectories not contained in the eigenlines are tangent to the slow eigenline at $(0, 0)$, and far from $(0, 0)$ have direction approximately parallel to the fast eigenline.

**Both negative:** $\lambda_1, \lambda_2 < 0$. This is called an attracting node (or node sink). All nonzero trajectories flow from infinity towards $(0, 0)$. Trajectories not contained in the eigenlines are tangent to the slow eigenline at $(0, 0)$, and far from $(0, 0)$ have direction approximately parallel to the fast eigenline.

One eigenvalue is zero: $\lambda_1 \neq 0, \lambda_2 = 0$. General solution: $c_1 e^{\lambda_1 t} v_1 + c_2 v_2$. This is a degenerate case called a comb. It could also be described by the words nonisolated critical points, since every point on the 0 eigenline is stationary. Other trajectories are along lines parallel to the other eigenline, tending to infinity if $\lambda_1 > 0$, and approaching the 0 eigenline if $\lambda_1 < 0$.

14.5.2. **Complex eigenvalues.** Suppose that the eigenvalues $\lambda_1, \lambda_2$ are not real. Then they are $a \pm bi$ for some real numbers $a, b$. In $e^{(a+bi)t}$, the number $a$ controls repulsion/attraction, while $b$ controls rotation (angular frequency).

**Zero real part:** $a = 0$. This is called a center. The nonzero trajectories are concentric ellipses. Solutions are periodic with period $2\pi/b$.

(Typical solution: $x = \text{Re} \left[ e^{it} \begin{pmatrix} 2 \\ -i \end{pmatrix} \right] = \begin{pmatrix} 2 \cos t \\ \sin t \end{pmatrix}$, a parametrization of a fat ellipse.)

**Positive real part:** $a > 0$. This is called a repelling spiral (or spiral source). All nonzero trajectories spiral outward.
Negative real part: $a < 0$. This is called an **attracting spiral** (or **spiral sink**). All nonzero trajectories spiral inward.

In these spiralling or rotating cases, how can one determine whether trajectories go clockwise or counterclockwise? It’s complicated to see this in terms of eigenvalues and eigenvectors, but easy to see by testing a single velocity vector. The velocity vector at $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ is $\dot{\mathbf{x}} = A\mathbf{x} = A \begin{pmatrix} 1 \\ 0 \end{pmatrix}$; trajectories go counterclockwise if and only if this velocity vector has positive $y$-coordinate.

14.5.3. **Repeated real eigenvalue.** Suppose that there is a repeated real eigenvalue, say $\lambda, \lambda$. The eigenspace of $\lambda$ could be either 1-dimensional or 2-dimensional.

- $\lambda \neq 0$ and $A \neq \lambda I$ (1-dimensional eigenspace): This is a called a **degenerate node** (or **improper node** or **defective node**). There is just one eigenline. Every trajectory not contained in it is tangent to it at $(0,0)$, and approximately parallel to it when far from $(0,0)$. Such trajectories are repelled from $(0,0)$ if $\lambda > 0$, and attracted to $(0,0)$ if $\lambda < 0$. (This is a borderline case between a node and a spiral.)

- $\lambda \neq 0$ and $A = \lambda I$ (2-dimensional eigenspace): This is a called a **star node**. Every nonzero vector is an eigenvector. Nonzero trajectories are along rays, repelled from $(0,0)$ if $\lambda > 0$, and attracted to $(0,0)$ if $\lambda < 0$.

- $\lambda = 0$ and $A \neq 0$ (1-dimensional eigenspace): This could be called **parallel lines**. Points on the eigenline are stationary. All other trajectories are lines parallel to the eigenline.

- $\lambda = 0$ and $A = 0$ (2-dimensional eigenspace): This could be called **stationary**. Every point is stationary.

14.5.4. **Summary.** Although all of the above may be needed for homework problems, for exams you are expected to know only the main cases listed in green above and also the case of a center, not the other “borderline” cases.

**Steps to sketch a phase portrait of $\dot{\mathbf{x}} = A\mathbf{x}$ (in most cases):**

1. Find the eigenvalues of $A$.
2. If the eigenvalues are distinct real numbers ($(\text{tr} A)^2 - 4 \det A > 0$) and are nonzero, find and draw the two eigenlines, and indicate the direction of motion along each (repelling/attracting according to eigenvalue being $+/−$).
   - If opposite sign, saddle. Other trajectories are asymptotic to both eigenlines, in the direction matching that of the nearby eigenline.
   - If same sign, then repelling/attracting node. Other trajectories are tangent to the slow eigenline at $(0,0)$.
3. If the eigenvalues are complex, say $a \pm bi$, check the sign of $a$:
   - If $+$, repelling spiral.
• If $-$, attracting spiral.
• If 0, center.

14.6. **Trace-determinant plane.** The possibilities above are determined by the eigenvalues $\lambda_1, \lambda_2$ (except in the case of a repeated eigenvalue, when one needs to know also the dimension of the eigenspace). And the eigenvalues are determined by the characteristic polynomial

$$\det(\lambda I - A) = \lambda^2 - (\text{tr} A)\lambda + (\det A) = (\lambda - \lambda_1)(\lambda - \lambda_2).$$

(Comparing coefficients shows that $\text{tr} A = \lambda_1 + \lambda_2$ and $\det A = \lambda_1\lambda_2$.)

Therefore the classification can be re-expressed in terms of $\text{tr} A$ and $\det A$. First, by the quadratic formula, the number of real eigenvalues is determined by the sign of the discriminant $(\text{tr} A)^2 - 4\det A$.

The details of the cases below were not covered in lecture.

14.6.1. **Distinct real eigenvalues.** Suppose that $(\text{tr} A)^2 - 4\det A > 0$.
Then the eigenvalues are real and distinct.

• If $\det A < 0$, then $\lambda_1\lambda_2 < 0$, so the eigenvalues have opposite sign: saddle.
• If $\det A > 0$, then the eigenvalues have the same sign.
  – If $\text{tr} A > 0$, repelling node.
  – If $\text{tr} A < 0$, attracting node.
• If $\det A = 0$, then one eigenvalue is 0; comb.

14.6.2. **Complex eigenvalues.** Suppose that $(\text{tr} A)^2 - 4\det A < 0$.
Then the eigenvalues are $a \pm bi$, and their sum is $\text{tr} A = 2a$.

• If $\text{tr} A = 0$, center.
• If $\text{tr} A > 0$, repelling spiral.
• If $\text{tr} A < 0$, attracting spiral.

14.6.3. **Repeated real eigenvalues.** Suppose that $(\text{tr} A)^2 - 4\det A = 0$.
Then we get a repeated real eigenvalue $\lambda, \lambda$, and $\text{tr} A = 2\lambda$.

• If $\text{tr} A \neq 0$, degenerate node or star node.
• If $\text{tr} A = 0$, parallel lines or the stationary case.
The trace-determinant plane is the plane with axes tr and det. This is completely different from the phase plane (because the axes are different).

Whereas the phase portrait shows all possible trajectories for a system $\dot{x} = Ax$, the trace-determinant plane has just one point for the system. The position of that point contains information about the kind of phase portrait.

Above the parabola $\det = \frac{1}{4}tr^2$, the eigenvalues are complex. Below the parabola, the eigenvalues are real and distinct.

April 4

14.7. Stability. Consider a system $\dot{x} = Ax$.

- If all trajectories tend to 0 as $t \to +\infty$, the system is called stable.
- If some trajectories are unbounded as $t \to +\infty$, then the system is called unstable.
- In the borderline case in which all solutions are bounded, but do not all tend to 0, the system is called semistable or neutrally stable. Example: a center.
The tests for stability are the same as for a single higher-order ODE, in terms of the roots or coefficients of the characteristic polynomial:

stable $\iff$ all eigenvalues have negative real part

(that makes each $e^{\lambda t}$ in the general solution tend to 0)

$\iff$ the characteristic polynomial has positive coefficients

( equivalently, $\text{tr} A < 0$ and $\det A > 0$).

(The green test is for the $2 \times 2$ case only.)

14.8. **Structural stability.** Stability is a question of what happens to solutions of a fixed system of ODEs. What happens if the system of ODEs itself is changed, by changing the matrix $A$? There is an unrelated definition to describe this:

**Definition 14.12.** If the phase portrait type is robust in the sense that small perturbations in the entries of $A$ cannot change the type of the phase portrait, then the system is called **structurally stable.**

The structurally stable cases are those corresponding to the large regions in the trace-determinant plane, not the borderline cases. For a $2 \times 2$ matrix $A$, the system $\dot{x} = Ax$ is structurally stable if and only if $A$ has either

- distinct nonzero real eigenvalues (saddle, repelling node, or attracting node), or
- complex eigenvalues with nonzero real part (spiral).

14.9. **Fundamental matrix.**

14.9.1. **Definition.** Consider a homogeneous linear system of $n$ ODEs $\dot{x} = Ax$. (We’ll assume that $A$ is constant, but everything in this section remains true even if $A$ is replaced by a matrix-valued function
We know that the set of solutions is an \( n \)-dimensional vector space. Let \( \mathbf{x}_1, \ldots, \mathbf{x}_n \) be a basis of solutions. Write \( \mathbf{x}_1, \ldots, \mathbf{x}_n \) as column vectors side-by-side to form a matrix

\[
X(t) := \begin{pmatrix} \vdots & \vdots \\ \mathbf{x}_1 & \mathbf{x}_n \\ \vdots & \vdots \end{pmatrix}.
\]

(It’s really a matrix-valued function, since each \( \mathbf{x}_i \) is a vector-valued function of \( t \).) Any such \( X(t) \) is called a fundamental matrix for \( \dot{\mathbf{x}} = A\mathbf{x} \). (There are many possible bases, so there are many possible fundamental matrices.)

### 14.9.2. General solution in terms of a fundamental matrix.

What is the point of putting the solutions in a fundamental matrix?

The general solution to \( \dot{\mathbf{x}} = A\mathbf{x} \) is \( c_1 \mathbf{x}_1 + \cdots + c_n \mathbf{x}_n = \begin{pmatrix} \vdots & \vdots \\ \mathbf{x}_1 & \mathbf{x}_n \\ \vdots & \vdots \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \).

**Conclusion:** If \( X(t) \) is a fundamental matrix, then the general solution is \( X(t)c \), where \( c \) ranges over constant vectors.

### 14.9.3. Solving a homogeneous system of ODEs with initial conditions.

**Problem 14.13.** The matrix \( A = \begin{pmatrix} 1 & 2 \\ -1 & 4 \end{pmatrix} \) has

- an eigenvector \( \begin{pmatrix} 2 \\ 1 \end{pmatrix} \) with eigenvalue 2 and
- an eigenvector \( \begin{pmatrix} 1 \\ 1 \end{pmatrix} \) with eigenvalue 3.

(a) Find a fundamental matrix for \( \dot{\mathbf{x}} = A\mathbf{x} \).

(b) Use it to find the solution to \( \dot{\mathbf{x}} = A\mathbf{x} \) satisfying the initial condition \( \mathbf{x}(0) = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \).

**Solution:**

(a) The functions

\[
e^{2t} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2e^{2t} \\ e^{2t} \end{pmatrix} \quad \text{and} \quad e^{3t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{3t} \\ e^{3t} \end{pmatrix}
\]

are a basis of solutions, so one fundamental matrix is

\[
X(t) = \begin{pmatrix} 2e^{2t} & e^{3t} \\ e^{2t} & e^{3t} \end{pmatrix}.
\]
(b) The solution will be \( X(t)c \) for some constant vector \( c \). Thus
\[
x = \begin{pmatrix} 2e^{2t} & e^{3t} \\ e^{2t} & e^{3t} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}
\]
for some \( c_1, c_2 \) to be determined. Set \( t = 0 \) and use the initial condition to get
\[
\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.
\]
In other words,
\[
2c_1 + c_2 = 4
\]
\[
c_1 + c_2 = 5.
\]
Solving leads to \( c_1 = -1 \) and \( c_2 = 6 \), so
\[
x = \begin{pmatrix} 2e^{2t} & e^{3t} \\ e^{2t} & e^{3t} \end{pmatrix} \begin{pmatrix} -1 \\ 6 \end{pmatrix} = \begin{pmatrix} -2e^{2t} + 6e^{3t} \\ -e^{2t} + 6e^{3t} \end{pmatrix}.
\]

14.9.4. Criterion for a matrix to be a fundamental matrix. To say that each column of \( X(t) \) is a solution is the same as saying that \( \dot{X} = AX \), because the matrix multiplication can be done column-by-column.

For a \( n \times n \) matrix whose columns are solutions, to say that the columns form a basis is equivalent to saying that they are linearly independent (the space of solutions is \( n \)-dimensional, so if \( n \) solutions are linearly independent, their span is the entire space). By the existence and uniqueness theorem, linear independence of solutions is equivalent to linear independence of their initial values at \( t = 0 \), i.e., to linear independence of the columns of \( X(0) \). So it is equivalent to say that \( X(0) \) is a nonsingular matrix.

Conclusion:

**Theorem 14.14.** A matrix-valued function \( X(t) \) is a fundamental matrix for \( \dot{x} = Ax \) if and only if

- \( \dot{X} = AX \) and
- the matrix \( X(0) \) is nonsingular.


14.10.1. **Definition.** For a real (or complex) number \( x \),
\[
e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots.
\]

**Definition 14.15.** For any square matrix \( A \),
\[
e^A := I + A + \frac{A^2}{2!} + \frac{A^3}{3!} + \cdots.
\]
So \( e^A \) is another matrix of the same size as \( A \).
14.10.2. Properties.

- \( e^0 = I \) (here 0 is the zero matrix)
  (Proof: \( e^0 = I + 0 + \frac{0^2}{2!} + \cdots = I \).)

- \( \frac{d}{dt} e^{At} = Ae^{At} \)
  (Proof: Take the derivative of \( e^{At} \) term by term.)

- If \( AB = BA \), then \( e^{A+B} = e^A e^B \). (Warning: This can fail if \( AB \neq BA \).)
  (Proof: Skipped.)

- If \( A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \), then \( e^A = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} \).
  (Proof: \( A^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} \), \( A^3 = \begin{pmatrix} \lambda_1^3 & 0 \\ 0 & \lambda_2^3 \end{pmatrix} \), and so on. Thus

\[
e^A = I + A + \frac{A^2}{2!} + \cdots = \begin{pmatrix} 1 + \lambda_1 + \frac{\lambda_1^2}{2!} + \cdots & 0 \\ 0 & 1 + \lambda_2 + \frac{\lambda_2^2}{2!} + \cdots \end{pmatrix} = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix}.
\]

A similar statement holds for diagonal matrices of any size.)

14.10.3. Matrix exponential and systems of ODEs.

**Theorem 14.16.** The function \( e^{At} \) is a fundamental matrix for the system \( \dot{x} = Ax \).

**Proof.** The function \( e^{At} \) satisfies \( \dot{X} = AX \) and its value at 0 is nonsingular. \( \square \)

Consequence: The general solution to \( \dot{x} = Ax \) is \( e^{At} c \).

**Compare:**
The solution to \( \dot{x} = ax \) satisfying the initial condition \( x(0) = c \) is \( e^{at} c \).
The solution to \( \dot{x} = Ax \) satisfying the initial condition \( x(0) = c \) is \( e^{At} c \).

**Question 14.17.** If the solution is as simple as \( e^{At} c \), why did we bother with the method involving eigenvalues and eigenvectors?

**Answer:** Because computing \( e^{At} \) is usually hard! (In fact, the standard method for computing it involves finding the eigenvalues and eigenvectors of \( A \).)

**Problem 14.18.** Use the matrix exponential to find the solution to the system

\[
\begin{align*}
\dot{x} &= 2x + y \\
\dot{y} &= 2y
\end{align*}
\]

satisfying \( x(0) = 5 \) and \( y(0) = 7 \).
Solution: This is $\dot{x} = Ax$ with $A := \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Write $A = D + N$ with $D := \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ (diagonal) and $N := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $N^2 = 0$, so

$$e^{Nt} = I + Nt = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$ 

Then $Dt$ and $Nt$ commute (a scalar times $I$ commutes with any matrix of the same size), so

$$e^{At} = e^{Dt+Nt} = e^{Dt}e^{Nt} = \begin{pmatrix} e^{2t} & 0 \\ 0 & e^{2t} \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{2t} & te^{2t} \\ 0 & e^{2t} \end{pmatrix},$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = e^{At} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} e^{2t} & te^{2t} \\ 0 & e^{2t} \end{pmatrix} \begin{pmatrix} 5 \\ 7 \end{pmatrix} = \begin{pmatrix} 5e^{2t} + 7te^{2t} \\ 7e^{2t} \end{pmatrix}. \quad \square$$

15. Inhomogeneous linear systems of ODEs

15.1. Diagonalization and decoupling.

15.1.1. Solving a decoupled system. The system $\dot{x} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix} x$ is the same as

$$\dot{x} = 3x \quad \dot{y} = 2y.$$ 

This is a decoupled system, consisting of two ODEs that can be solved separately. More generally, if $D$ is a diagonal matrix of any size, the inhomogeneous system

$$\dot{x} = Dx + q(t)$$

consists of first-order linear ODEs that can be solved separately.

Plan: develop a method to transform other systems into this form.
15.1.2. *Diagonalization.* Suppose that $A$ is a $2 \times 2$ matrix with a basis of eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ having eigenvalues $\lambda_1, \lambda_2$ (in other words, we are assuming that $A$ is not deficient). (Warning: For what we are about to do, the eigenvectors must be listed in the same order as their eigenvalues.)

Use the eigenvalues to define a diagonal matrix

$$D := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$ 

The matrix mapping \( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) to $\mathbf{v}_1, \mathbf{v}_2$, respectively, is the matrix

$$S := \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix}$$

whose columns are the eigenvectors $\mathbf{v}_1, \mathbf{v}_2$.

**Question 15.1.** What is the $2 \times 2$ matrix that maps \( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) to $\lambda_1 \mathbf{v}_1$, $\lambda_2 \mathbf{v}_2$, respectively?
Answer 1: \( AS \), because applying \( AS \) means that \( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) are mapped by \( S \) to \( v_1, v_2 \), which are then mapped by \( A \) to \( \lambda_1 v_1, \lambda_2 v_2 \).

Answer 2: \( SD \), because \( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) are mapped by \( D \) to \( \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix} \) which are then mapped by \( S \) to \( \lambda_1 v_1, \lambda_2 v_2 \).

Conclusion: \( AS = SD \) (Memory aid: Look at where \( A, S, D \) are on your keyboard.)

Multiply by \( S^{-1} \) on the right to get another way to write it: \( A = SDS^{-1} \).

Writing the matrix \( A \) like this is called diagonalizing \( A \). Think of \( S \) as a “coordinate-change matrix” or “change-of-basis matrix” that

- relates the standard basis \( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) to the basis of eigenvectors of \( A \), and
- relates the easy matrix \( D \) scaling the standard basis vectors to the original matrix \( A \) scaling the original eigenvectors.

Remark 15.2. Diagonalization of an \( n \times n \) matrix \( A \) is possible if and only if \( A \) has \( n \) independent eigenvectors (this fails if there is a repeated eigenvalue and the dimension of its eigenspace is less than its multiplicity).

15.1.3. Applications of diagonalization. Suppose that \( A = SDS^{-1} \). Then

\[
A^3 = S D S^{-1} S D S^{-1} S D S^{-1} = S D^3 S^{-1}.
\]

More generally, for any integer \( n \geq 0 \),

\[
A^n = SD^n S^{-1}.
\]  \( \text{(11)} \)

The MITx reading explains how to use equation (11) to get an explicit formula for Fibonacci numbers.

Using (11) to evaluate each term in the power series definition of \( e^A \) leads to

\[
e^A = S e^D S^{-1}.
\]

This lets one compute \( e^A \) for any diagonalizable matrix \( A \)!

15.1.4. Decoupling. Here is a slightly silly way to solve

\[
\dot{x} = Ax.
\]

Substitute \( x = Sy \), and rewrite the system in terms of \( y \):

\[
S \dot{y} = ASy \quad \text{(since } S \text{ is constant)}
\]
\[
S \dot{y} = SDy \quad \text{(since } AS = SD) 
\]
\[
\dot{y} = Dy \quad \text{(we multiplied by } S^{-1} \text{ on the left)}.
\]
This is decoupled! So solve for each coordinate function of \( y \), and then compute \( x = S y \).

Why is that silly? Because we already know how to solve \( \dot{x} = Ax \) when we have a basis of eigenvectors (and their eigenvalues).

But... the same decoupling method also lets us solve an *inhomogeneous* linear system, and that’s not silly:

**Steps to solve \( \dot{x} = Ax + q(t) \) by decoupling:**

1. Find the eigenvalues of \( A \) (with multiplicity), and put them in a diagonal matrix \( D \).
2. Find a basis of each eigenspace. If the total number of independent eigenvectors found is less than \( n \), then a more complicated method (not discussed here) is required. Put the eigenvectors as columns of a matrix \( S \).
3. Substitute \( x = Sy \) to get
   
   \[
   \begin{align*}
   S\ddot{y} &= ASy + q(t) \\
   S\dot{y} &= SDy + q(t) \\
   \dot{y} &= Dy + S^{-1}q(t).
   \end{align*}
   \]
   
   (You may skip to the last of these equations.) This is a decoupled system of inhomogeneous linear ODEs.
4. Solve for each coordinate function of \( y \).
5. Compute \( Sy \); the result is \( x \).

The following problem was not actually solved in lecture.

**Problem 15.3.** Find a particular solution to \( \dot{x} = Ax + q \) where \( A := \begin{pmatrix} -4 & -3 \\ 6 & 5 \end{pmatrix} \) and \( q = \begin{pmatrix} 0 \\ \cos t \end{pmatrix} \).

**Solution:** We will solve it instead with \( q = \begin{pmatrix} 0 \\ e^{it} \end{pmatrix} \) (complex replacement), and take the real part of the solution at the very end.

*Step 1.* We have \( \text{tr} \ A = 1 \) and \( \det A = -20 - (-18) = -2 \).

Characteristic polynomial: \( \lambda^2 - \lambda - 2 = (\lambda - 2)(\lambda + 1) \).

Eigenvalues: 2, -1. Therefore define

\[
D := \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}.
\]
Step 2. Calculating eigenspaces in the usual way leads to corresponding eigenvectors
\[
\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix},
\]
so define
\[
S := \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix}.
\]
Now \( A = SDS^{-1} \).

Step 3. The result of substituting \( x = Sy \) is
\[
\dot{y} = Dy + S^{-1}q.
\]
We have
\[
S^{-1}q = \frac{1}{\det S} \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ e^{it} \end{pmatrix}
= \begin{pmatrix} -e^{it} \\ e^{it} \end{pmatrix}
\]
so the decoupled system is
\[
\begin{align*}
\dot{y}_1 &= 2y_1 - e^{it} \\
\dot{y}_2 &= -y_2 + e^{it}.
\end{align*}
\]

Step 4. Solving with ERF gives particular solutions
\[
\begin{align*}
y_1 &= \frac{-1}{i - 2} e^{it} = \left( \frac{2}{5} + \frac{1}{5}i \right) e^{it} \\
y_2 &= \frac{1}{i + 1} e^{it} = \left( \frac{1}{2} - \frac{1}{2}i \right) e^{it}.
\end{align*}
\]

Step 5.
\[
x = Sy
= \begin{pmatrix} 1 & 1 \\ -2 & -1 \end{pmatrix} \begin{pmatrix} \frac{2}{5} + \frac{1}{5}i \\ \frac{1}{2} - \frac{1}{2}i \end{pmatrix} e^{it}
= \begin{pmatrix} \frac{9}{10} - \frac{3}{10}i \\ -\frac{13}{10} + \frac{1}{10}i \end{pmatrix} (\cos t + i \sin t).
\]

Final step: Take the real part to get a particular solution to the original system:
\[
x = \begin{pmatrix} \frac{9}{10} \cos t + \frac{3}{10} \sin t \\ -\frac{13}{10} \cos t - \frac{1}{10} \sin t \end{pmatrix}.
\]
15.2. Variation of parameters.

Long ago we learned how to use variation of parameters to solve inhomogeneous linear ODEs

\[ \dot{y} + p(t)y = q(t). \]

Now we're going to use the same idea to solve an inhomogeneous linear system of ODEs such as

\[ \dot{x} = Ax + q, \]

where \( q \) is a vector-valued function of \( t \). First find a basis of solutions to the corresponding homogeneous system

\[ \dot{x} = Ax, \]

and put them together to form a fundamental matrix \( X \) (a matrix-valued function of \( t \)). We know that \( Xc \), where \( c \) ranges over constant vectors, is the general solution to the homogeneous equation. Replace \( c \) by a vector-valued function \( u \): try \( x = Xu \) in the original system:

\[
\begin{align*}
\dot{x} &= Ax + q \\
\dot{X}u + X\dot{u} &= AXu + q \\
AXu + X\dot{u} &= AXu + q \\
X\dot{u} &= q \\
\dot{u} &= X^{-1}q
\end{align*}
\]

Steps to solve \( \dot{x} = Ax + q \) by variation of parameters:

1. Find a fundamental matrix \( X \) for the homogeneous system \( \dot{x} = Ax \) (e.g., by using eigenvalues and eigenvectors to find a basis of solutions).
2. Substitute \( x = Xu \) for a vector-valued function \( u \); this eventually leads to

\[ \dot{u} = X^{-1}q \]

(and you may jump right to this if you want).
3. Compute the right hand side and integrate each component function to find \( u \).
   (The indefinite integral will have a +c.)
4. Then \( x = Xu \) is the general solution to the inhomogeneous equation.
   (It is a family of vector-valued functions because of the +c in \( u \).)

Remark 15.4. One choice of \( X \) is \( e^{At} \), in which case \( \dot{u} = e^{-At}q \) and

\[ x = e^{At}u = e^{At} \int e^{-At}q \, dt. \]
16. Coordinates

16.1. Coordinates with respect to a basis.

**Problem 16.1.** The vectors \( \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \) form a basis for \( \mathbb{R}^2 \), so any vector in \( \mathbb{R}^2 \) is a linear combination of them. Find \( c_1 \) and \( c_2 \) such that

\[
c_1 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}
\]

(These \( c_1, c_2 \) are called the *coordinates* of \( \begin{pmatrix} 2 \\ 4 \end{pmatrix} \) with respect to the basis. There is only one solution, since if there were two different linear combinations giving \( \begin{pmatrix} 2 \\ 4 \end{pmatrix} \), subtracting them would give a nontrivial linear combination giving \( \mathbf{0} \), which is impossible since the basis vectors are linearly independent.)

**Solution:** Multiply it out to get

\[
2c_1 - c_2 = 2 \\
c_1 + c_2 = 4
\]

or equivalently, in matrix form,

\[
\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}.
\]

Solving gives \( c_1 = 2 \) and \( c_2 = 2 \).

“Coordinates with respect to a basis” make sense also in vector spaces of functions.

**Problem 16.2.** Let \( V \) be the vector space with basis consisting of the three functions \( 1, t - 3, (t - 3)^2 \). Find the coordinates of the function \( t^2 \) with respect to this basis.

**Solution:** We need to find \( c_1, c_2, c_3 \) such that

\[
t^2 = c_1(1) + c_2(t - 3) + c_3(t - 3)^2.
\]

Equating constant terms gives

\[
0 = c_1 - 3c_2 + 9c_3.
\]

Equating coefficients of \( t \) gives

\[
0 = c_2 - 6c_3.
\]

Equating coefficients of \( t^2 \) gives

\[
1 = c_3.
\]

Solving this system of three equations leads to \( (c_1, c_2, c_3) = (9, 6, 1) \). \( \square \)
16.2. **Orthogonal basis and orthonormal basis.**

Consider a basis \( \mathbf{v}_1, \ldots, \mathbf{v}_n \) of \( \mathbb{R}^n \).

- If each \( \mathbf{v}_i \) is orthogonal (perpendicular) to every other \( \mathbf{v}_j \), then the basis is called an **orthogonal basis**.
- If in addition each \( \mathbf{v}_i \) has length 1, then the basis is called an **orthonormal basis**.

(The vectors in it are neither abnormally long nor abnormally short!)

**Example 16.3.** The vectors \( \mathbf{i}, \mathbf{j}, \mathbf{k} \) form an orthonormal basis for \( \mathbb{R}^3 \).

**Flashcard question:** What is true of the list of vectors \( \mathbf{v}_1 := \begin{pmatrix} 2 \\ 2 \end{pmatrix} \) and \( \mathbf{v}_2 := \begin{pmatrix} -1 \\ 1 \end{pmatrix} \) in \( \mathbb{R}^2 \)?

**Possible answers:**

- This is not a basis of \( \mathbb{R}^2 \).
- This is a basis, but not an orthogonal basis.
- This is an orthogonal basis, but not an orthonormal basis.
- This is an orthonormal basis.

**Answer:** It is an orthogonal basis, but not an orthonormal basis. To test, use dot products: \( \mathbf{v}_1 \cdot \mathbf{v}_2 = 0 \), so they are orthogonal. But \( \mathbf{v}_1 \cdot \mathbf{v}_1 \neq 1 \), so \( \mathbf{v}_1 \) does not have length 1. □

16.3. **Shortcuts for finding coordinates.**

**Question 16.4.** Suppose that \( \mathbf{v}_1, \ldots, \mathbf{v}_n \) is an **orthonormal** basis of \( \mathbb{R}^n \). How can we find the coordinates \( c_1, \ldots, c_n \) of a vector \( \mathbf{w} \) with respect to this basis?

**Answer:** We need to solve

\[
\mathbf{w} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n
\]

for \( c_1, \ldots, c_n \). Trick: dot both sides with \( \mathbf{v}_1 \) to get

\[
\mathbf{w} \cdot \mathbf{v}_1 = c_1 (1) + 0 + \cdots + 0.
\]

Get

\[
c_1 = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}, \quad \cdots , \quad c_n = \frac{\mathbf{w} \cdot \mathbf{v}_n}{\mathbf{v}_n \cdot \mathbf{v}_n}.
\]

**Question 16.5.** Suppose that \( \mathbf{v}_1, \ldots, \mathbf{v}_n \) is only an **orthogonal** basis of \( \mathbb{R}^n \). How can we find the coordinates \( c_1, \ldots, c_n \) of a vector \( \mathbf{w} \) with respect to this basis?

**Answer:** The same trick leads to

\[
\mathbf{w} \cdot \mathbf{v}_1 = c_1 \mathbf{v}_1 \cdot \mathbf{v}_1,
\]

so we get

\[
c_1 = \frac{\mathbf{w} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}, \quad \cdots , \quad c_n = \frac{\mathbf{w} \cdot \mathbf{v}_n}{\mathbf{v}_n \cdot \mathbf{v}_n}.
\]
17. Introduction to Fourier series

17.1. Periodic functions. Because

$$\sin(t + 2\pi) = \sin t, \quad \cos(t + 2\pi) = \cos t,$$

hold for all \( t \), the functions \( \sin t \) and \( \cos t \) are called periodic with period \( 2\pi \). In general, “\( f(t) \) is periodic of period \( P \)” means that \( f(t + P) = f(t) \) for all \( t \) (or at least all \( t \) for which either side is defined).

There are many such functions beyond the sinusoidal functions. To construct one, divide the real line into intervals of length \( P \), start with any function defined on one such interval \([t_0, t_0 + P)\), and then copy its values in the other intervals. The entire graph consists of horizontally shifted copies of the width \( P \) graph.

Today: \( P = 2\pi \), interval \([-\pi, \pi)\).

Question 17.1. Is \( \sin 3t \) periodic of period \( 2\pi \)?

Answer: The shortest period is \( 2\pi/3 \), but \( \sin 3t \) is also periodic with period any positive integer multiple of \( 2\pi/3 \), including \( 3(2\pi/3) = 2\pi \):

$$\sin(3(t + 2\pi)) = \sin(3t + 6\pi) = \sin 3t.$$

So the answer is yes.

17.2. Square wave. Define

$$\text{Sq}(t) := \begin{cases} 1, & \text{if } 0 < t < \pi, \\ -1 & \text{if } -\pi < t < 0. \end{cases}$$

and extend it to a periodic function of period \( 2\pi \), called a square wave. The function \( \text{Sq}(t) \) has jump discontinuities, for example at \( t = 0 \). If you must define \( \text{Sq}(0) \), compromise between the upper and lower values: \( \text{Sq}(0) := 0 \). The graph is usually drawn with vertical segments at the jumps (even though this violates the vertical line test).

It turns out that

$$\text{Sq}(t) = \frac{4}{\pi} \left( \sin t + \frac{\sin 3t}{3} + \frac{\sin 5t}{5} + \cdots \right).$$

We’ll explain later today where this comes from.

Try the “Fourier Coefficients” mathlet

[http://mathlets.org/mathlets/fourier-coefficients/]
17.3. **Fourier series.** A linear combination like $2 \sin 3t - 4 \sin 7t$ is another periodic function of period $2\pi$.

**Definition 17.2.** A **Fourier series** is a linear combination of the infinitely many functions $\cos nt$ and $\sin nt$ as $n$ ranges over integers:

$$f(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + a_3 \cos 3t + \cdots + b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t + \cdots$$

(Terms like $\cos(-2t)$ are redundant since $\cos(-2t) = \cos 2t$. Also $\sin 0t = 0$ produces nothing new. But $\cos 0t = 1$ is included. We’ll explain later why we write $a_0$ times $1/2$ instead of times 1.)

Written using sigma-notation:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt + \sum_{n=1}^{\infty} b_n \sin nt.$$ 

Recall that, for example, $\sum_{n=1}^{\infty} b_n \sin nt$ means the sum of the series whose $n^{th}$ term is obtained by plugging in the positive integer $n$ into the expression $b_n \sin nt$, so

$$\sum_{n \geq 1} b_n \sin nt = b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t + \cdots.$$ 

Any Fourier series as above is periodic of period $2\pi$. (Later we’ll extend to the definition of Fourier series to include functions of other periods.) The numbers $a_n$ and $b_n$ are called the **Fourier coefficients** of $f$. Each summand ($a_0/2$, $a_n \cos nt$, or $b_n \sin nt$) is called a **Fourier component** of $f$.

**Fourier’s theorem.** “Every” periodic function $f$ of period $2\pi$ “is” a Fourier series, and the Fourier coefficients are uniquely determined by $f$.

(The word “Every” has to be taken with a grain of salt: The function has to be “reasonable”. Piecewise differentiable functions with jump discontinuities are reasonable, as are virtually all other functions that arise in physical applications.

The word “is” has to be taken with a grain of salt: If $f$ has a jump discontinuity at $\tau$, then the Fourier series might disagree with $f$ there; the value of the Fourier series at $\tau$ is always the average of the left limit $f(\tau^-)$ and the right limit $f(\tau^+)$, regardless of the actual value of $f(\tau)$.

In other words, the functions

$$1, \cos t, \cos 2t, \cos 3t, \ldots, \sin t, \sin 2t, \sin 3t, \ldots$$

form a basis for the vector space of “all” periodic functions of period $2\pi$.

**Question 17.3.** Given $f$, how do you find the Fourier coefficients $a_n$ and $b_n$?

In other words, how do you find the coordinates of $f$ with respect to the basis of cosines and sines?
17.4. A “dot product” for functions. If \( \mathbf{v} \) and \( \mathbf{w} \) are vectors in \( \mathbb{R}^n \), then

\[
\mathbf{v} \cdot \mathbf{w} := \sum_{i=1}^{n} v_i w_i.
\]

Can one define the dot product of two functions? Sort of.

**Definition 17.4.** If \( f \) and \( g \) are real-valued periodic functions with period \( 2\pi \), then their inner product is

\[
\langle f, g \rangle := \int_{-\pi}^{\pi} f(t)g(t) \, dt
\]

(It acts like a dot product \( f \cdot g \), but don’t write it that way because \( \cdot \) could be misinterpreted as multiplication.)

**Example 17.5.** By definition,

\[
\langle 1, \cos t \rangle = \int_{-\pi}^{\pi} \cos t \, dt = 0.
\]

Thus the functions 1 and \( \cos t \) are orthogonal.

In fact, calculating all the inner products shows that

\[
1, \cos t, \cos 2t, \cos 3t, \ldots, \sin t, \sin 2t, \sin 3t, \ldots
\]

is an orthogonal basis!

**Question 17.6.** Is it an orthonormal basis?

**Answer:** No, since \( \langle 1, 1 \rangle = \int_{-\pi}^{\pi} 1 \, dt = 2\pi \neq 1. \) □

**Example 17.7.**

\[
\langle \sin t, \sin t \rangle = \int_{-\pi}^{\pi} \sin^2 t \, dt = ?
\]

\[
\langle \cos t, \cos t \rangle = \int_{-\pi}^{\pi} \cos^2 t \, dt = ?
\]

Since \( \cos t \) is just a shift of \( \sin t \), the answers are going to be the same. Also, the two answers add up to

\[
\int_{-\pi}^{\pi} (\sin^2 t + \cos^2 t) \, dt = 2\pi,
\]

so each is \( \pi \).

The same idea works to show that

\[
\langle \cos nt, \cos nt \rangle = \pi \quad \text{and} \quad \langle \sin nt, \sin nt \rangle = \pi
\]

for each positive integer \( n \).
17.5. **Fourier coefficient formulas.** Given $f$, how do you find the $a_n$ and $b_n$ such that

$$f(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + a_3 \cos 3t + \cdots + b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t + \cdots?$$

By the shortcut formulas in Section 16.3,

$$a_n = \frac{\langle f, \cos nt \rangle}{\langle \cos nt, \cos nt \rangle} = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt,$$

and the coefficient of 1 is

$$a_0 = \frac{\langle f, 1 \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, dt.$$

so

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos 0t \, dt.$$

(Using $a_0/2$ in the series ensures that the formula for $a_n$ for $n > 0$ works also for $n = 0$.) A similar formula holds for $b_n$.

**Conclusion:** Given $f$, its Fourier coefficients can be calculated as follows:

<table>
<thead>
<tr>
<th>$a_n$</th>
<th>$b_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt , dt$</td>
<td>$\frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt , dt$</td>
</tr>
</tbody>
</table>

for all $n \geq 0$, for all $n \geq 1$.

17.6. **Meaning of the constant term.** The constant term of the Fourier series of $f$ is

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, dt,$$

which is the average value of $f$ on $(-\pi, \pi)$.

17.7. **Even and odd symmetry.**

- A function $f(t)$ is **even** if $f(-t) = f(t)$ for all $t$.
- A function $f(t)$ is **odd** if $f(-t) = -f(t)$ for all $t$.

If

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt + \sum_{n=1}^{\infty} b_n \sin nt,$$

then substituting $-t$ for $t$ gives

$$f(-t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt + \sum_{n=1}^{\infty} (-b_n) \sin nt.$$

The right hand sides match if and only if $b_n = 0$ for all $n$. 

---

125
Conclusion: The Fourier series of an even function $f$ has only cosine terms (including the constant term):

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt.$$

Similarly, the Fourier series of an odd function $f$ has only sine terms:

$$f(t) = \sum_{n=1}^{\infty} b_n \sin nt.$$

Example 17.8. The square wave $\text{Sq}(t)$ is an odd function, so

$$\text{Sq}(t) = \sum_{n=1}^{\infty} b_n \sin nt$$

for some numbers $b_n$. The Fourier coefficient formula says

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \text{Sq}(t) \sin nt \, dt$$

$$= \frac{2}{\pi} \int_{0}^{\pi} \text{Sq}(t) \sin nt \, dt \quad \text{(the two halves of the integral are equal, by symmetry)}$$

$$= \frac{2}{\pi} \int_{0}^{\pi} \sin nt \, dt \quad \text{(since $\text{Sq}(t) = 1$ whenever $0 < t < \pi$)}$$

$$= \left[ \frac{2(-\cos nt)}{\pi n} \right]_{0}^{\pi}$$

$$= \frac{2}{\pi n} (-\cos n\pi + \cos 0)$$

$$= \begin{cases} \frac{4}{\pi n}, & \text{if } n \text{ is odd,} \\ 0, & \text{if } n \text{ is even.} \end{cases}$$

Thus

$$b_1 = \frac{4}{\pi}, \quad b_3 = \frac{4}{3\pi}, \quad b_5 = \frac{4}{5\pi}, \ldots$$

and all other Fourier coefficients are 0.

Conclusion:

$$\text{Sq}(t) = \frac{4}{\pi} \left( \sin t + \frac{\sin 3t}{3} + \frac{\sin 5t}{5} + \cdots \right). \quad \square$$

17.8. Finding a Fourier series representing a function on an interval.

Problem 17.9. Suppose that $f(t)$ is a (reasonable) function defined only on the interval $(0, \pi)$. Find numbers $a_0, a_1, \ldots$ such that

$$f(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + \cdots$$

for all $t \in (0, \pi)$. 

126
Solution: For any $a_i$, the right hand side will define an even periodic function of period $2\pi$ (if the series converges). So begin by extending $f(t)$ to a function of the same type:

- Extend $f(t)$ to an even function on $(-\pi, \pi)$ by defining $f(-t) := f(t)$ for all $t \in (-\pi, 0)$ (and then define $f(0)$ and $f(-\pi)$ arbitrarily).
- Shift the graph of $f$ horizontally by integer multiples of $2\pi$ to get a period $2\pi$ function defined on all of $\mathbb{R}$.

Define

$$a_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt = \frac{2}{\pi} \int_{0}^{\pi} f(t) \cos nt \, dt.$$  

Then

$$f(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + \cdots$$

holds for all $t \in \mathbb{R}$, so in particular it holds for $t \in (0, \pi)$ (possibly excluding points of discontinuity). □

Remark 17.10. The same function $f(t)$ on $(0, \pi)$ can be extended to an odd periodic function of period $2\pi$, in order to obtain

$$f(t) = b_1 \sin t + b_2 \sin 2t + \cdots,$$

where

$$b_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt = \frac{2}{\pi} \int_{0}^{\pi} f(t) \sin nt \, dt.$$  

Midterm 3 covers everything up to here.

April 11

Eigenvalues and eigenvectors are used in many ways in science and engineering, not just for solving DEs.

17.9. How Google search uses an eigenvector. This can be skipped.

Google claims that the heart of its software is PageRank: this is the algorithm for deciding how to order search results. The core idea involves an eigenvector, as we’ll now explain. (The details of the algorithm are more complicated and proprietary.)

The web consists of webpages linking to each other. Number them.
Let \( v_i \) be the “importance” of webpage \( i \).

**Idea:** A webpage is important if important webpages link to it. Each webpage “shares” its importance equally with all the webpages it links to.

In the example above, page 2 inherits \( \frac{1}{2} \) the importance of page 1, \( \frac{1}{2} \) the importance of page 3, and \( \frac{1}{3} \) the importance of page 4:

\[
v_2 = \frac{1}{2} v_1 + \frac{1}{2} v_3 + \frac{1}{3} v_4.
\]

Yes, this is self-referential, but still it makes sense. All eight equations are encapsulated in

\[
\begin{pmatrix}
  v_1 \\
  v_2 \\
  v_3 \\
  v_4 \\
  v_5 \\
  v_6 \\
  v_7 \\
  v_8 \\
\end{pmatrix} = 
\begin{pmatrix}
  0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 0 \\
  1/2 & 0 & 1/2 & 1/3 & 0 & 0 & 0 & 0 \\
  1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
  0 & 0 & 1/2 & 1/3 & 0 & 0 & 1/3 & 0 \\
  0 & 0 & 0 & 1/3 & 1/3 & 0 & 0 & 1/2 \\
  0 & 0 & 0 & 0 & 1/3 & 0 & 0 & 1/2 \\
  0 & 0 & 0 & 0 & 1/3 & 1 & 1/3 & 0 \\
\end{pmatrix}
\begin{pmatrix}
  v_1 \\
  v_2 \\
  v_3 \\
  v_4 \\
  v_5 \\
  v_6 \\
  v_7 \\
  v_8 \\
\end{pmatrix},
\]

which is of the form \( \mathbf{v} = A \mathbf{v} \). In other words, \( \mathbf{v} \) should be an eigenvector with eigenvalue 1.

**Question 17.11.** How do we know that a matrix like \( A \) has an eigenvector with eigenvalue 1? Could it be that 1 is just not an eigenvalue?

**Trick:** use the transpose \( A^T \).

\[
\det A = \det A^T
\]

\[
\det(A - \lambda I) = \det(A^T - \lambda I)
\]

eigenvalues of \( A \) = eigenvalues of \( A^T \).
The equation

\[
\begin{pmatrix}
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 1/2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\
0 & 0 & 0 & 1/3 & 0 & 0 & 1/3 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
1/3 & 0 & 0 & 0 & 1/3 & 0 & 0 & 1/3 \\
0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 & 0
\end{pmatrix}
\begin{pmatrix}
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1
\end{pmatrix}
= 
\begin{pmatrix}
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1
\end{pmatrix}
\begin{pmatrix}
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1 \\
1
\end{pmatrix}
, 
\]

shows that 1 is an eigenvalue of \(A^T\), so 1 is an eigenvalue of \(A\).

In the example above, the unique solution (up to multiplying by a scalar) is

\[
\begin{pmatrix}
0.0600 \\
0.0675 \\
0.0300 \\
0.0675 \\
0.0975 \\
0.2025 \\
0.1800 \\
0.2950
\end{pmatrix}
. 
\]

Google finds the eigenvector of a 50,000,000,000 \(\times\) 50,000,000,000 matrix.

See

http://www.ams.org/samplings/feature-column/fcarc-pagerank

for more details.

17.10. Review.

17.10.1. How to check your answers. To check...

- that \(\lambda\) is an eigenvalue of \(A\): check that \(\text{det}(A - \lambda I) = 0\).
- that \(v\) is an eigenvector of \(A\): check that \(Av\) is a scalar multiple of \(v\).
- a diagonalization \(A = SDS^{-1}\): compute \(S^{-1}\) and multiply out \(SDS^{-1}\) to check that it gives \(A\). Or, even better: multiply out both sides of \(AS = SD\).
- a solution to a DE: plug it in.
- that \(c_1, c_2\) are the coordinates of \(w\) with respect to a basis \(v_1, v_2\): check that \(c_1v_1 + c_2v_2\) really gives \(w\).
- a phase portrait for \(\dot{x} = Ax\): compute the velocity vector at a point or two by evaluating \(Ax\) at specific points \(x\) (see problem below for an example).
**Problem 17.12.** Given \( A := \begin{pmatrix} 1 & -4 \\ 1 & 1 \end{pmatrix} \), sketch the phase portrait for \( \dot{x} = A \dot{x} \).

**Solution:** We have \( \text{tr} \, A = 1 + 1 = 2 \) and \( \det \, A = 1 - (-4) = 5 \), so the characteristic polynomial is \( \lambda^2 - 2\lambda + 5 = (\lambda - 1)^2 + 4 \), and the eigenvalues are \( 1 \pm 2i \). Since the eigenvalues are not real, and have positive real part, the phase portrait is a repelling spiral.

Do the trajectories go clockwise or counterclockwise? It’s complicated to see this in terms of eigenvalues and eigenvectors, but the velocity vector at \( x = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) is

\[
\dot{x} = A \dot{x} = \begin{pmatrix} 1 & -4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},
\]

so the trajectories are counterclockwise.

17.10.2. **Fourier series.** Key points:

- Fourier’s theorem: “Every” periodic function \( f \) of period \( 2\pi \) is a Fourier series

\[
f(t) = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + a_3 \cos 3t + \cdots \\
+ b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t + \cdots.
\]

- Given \( f \), the Fourier coefficients \( a_n \) and \( b_n \) can be computed using:

\[
a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos nt \, dt \quad \text{for all } n \geq 0,
\]

\[
b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin nt \, dt \quad \text{for all } n \geq 1.
\]

- If \( f \) is even, then only the cosine terms (including the \( a_0/2 \) term) appear.
- If \( f \) is odd, then only the sine terms appear.

**Problem 17.13.** Let \( f(t) \) be a periodic function of period \( 2\pi \) such that

\[
f(t) = \begin{cases} 
0, & \text{if } -\pi < t < 0; \\
1, & \text{if } 0 < t < \pi/2; \\
0, & \text{if } \pi/2 < t < \pi.
\end{cases}
\]

What is the Fourier coefficient \( b_1 \) for this function?
Solution: If \( f \) were an even function, then \( b_1 \) would have to be 0. But \( f \) is not even, so we’ll have to use the Fourier coefficient formula:

\[
b_1 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin t \, dt
\]

\[
= \frac{1}{\pi} \int_{0}^{\pi/2} \sin t \, dt
\]

\[
= \frac{1}{\pi} \left( -\cos t \right)_{0}^{\pi/2}
\]

\[
= \frac{1}{\pi} \left( -\cos \frac{\pi}{2} + \cos 0 \right)
\]

\[
= \frac{1}{\pi}.
\]

17.10.3. Matrix exponential. How do you compute \( e^A \)?

- If \( D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \), then \( e^D = \begin{pmatrix} e^{\lambda_1} & 0 \\ 0 & e^{\lambda_2} \end{pmatrix} \), and the same idea works for diagonal matrices of any size.
- If \( A = SDS^{-1} \) for some diagonal matrix \( D \) and nonsingular matrix \( S \) (so \( A \) is diagonalizable), then \( e^A = Se^DS^{-1} \) (and we just saw how to compute \( e^D \), so this lets you compute \( e^A \)).
- If \( A^2 = 0 \), then \( e^A = I + A \) (and \( e^{At} = I + At \)).
- If \( A^3 = 0 \), then \( e^A = I + A + \frac{A^2}{2!} \) (and \( e^{At} = I + At + \frac{A^2}{2!} t^2 \)).

The same idea works for all nilpotent matrices, i.e., matrices having a power that is 0.
- If \( A = B + N \) where \( BN = NB \), then \( e^A = e^B e^N \).

It turns out that every square matrix can be written as \( B + N \) where \( B \) is diagonalizable and \( N \) is nilpotent and \( BN = NB \); so in principle, \( e^A \) can always be computed.

Other facts about \( e^A \) and the matrix-valued function \( e^{At} \):

- The derivative of \( e^{At} \) is \( Ae^{At} \).
- The matrix-valued function \( e^{At} \) is the fundamental matrix (for \( \dot{x} = Ax \)) whose value at \( t = 0 \) is \( I \).
- The solution to \( \dot{x} = Ax \) satisfying the initial condition \( x(0) = c \) is \( e^{At} c \).

17.10.4. Solving an inhomogeneous system of ODEs. What are the methods to solve an inhomogeneous system \( \dot{x} = Ax + q \)?

- Convert to a higher-order ODE involving only one unknown function.
- Diagonalize \( A \) to decouple the system: Solve \( \dot{y} = Dy + S^{-1}q \), then compute \( x = Sy \).
- Variation of parameters: Substitute \( x = Xu \) to get \( \dot{u} = X^{-1}q \), solve for the general \( u \), compute \( x = Xu \).
• Find one particular solution somehow, and add it to the general solution to the homogeneous system.

If there are initial conditions, first find the general solution to the inhomogeneous system, and then use the initial conditions to solve for the unknown parameters (and plug them back in at the end).

17.10.5. **Phase portraits.** Assume that $A$ has distinct nonzero eigenvalues.

1. If the eigenvalues are real, draw the two eigenlines, and indicate the direction of motion along each (repelling/attracting according to eigenvalue being $+/-$).
   - If opposite sign, **saddle**. Other trajectories are asymptotic to both eigenlines, in the direction matching that of the nearby eigenline.
   - If same sign, then repelling/attracting **node**. Other trajectories are tangent to the slow eigenline at $(0,0)$.

2. If the eigenvalues are $a \pm bi$, check the sign of $a$:
   - If $+$, repelling spiral.
   - If $-$, attracting spiral.
   - If 0, center.

---

**April 14**

Midterm 3

**April 16**

17.11. **Solving ODEs with Fourier series.**

**Problem 17.14.** Suppose that $f(t)$ is an odd periodic function of period $2\pi$. Find the periodic function $x(t)$ of period $2\pi$ that is a solution to

$$\ddot{x} + 50x = f(t)$$

Think of $f(t)$ as the input signal, and the solution $x(t)$ as the system response (output signal).

17.11.1. **Warm-up: system response to sinusoidal input (review).**

**Special case:** What is the system response to the input signal $\sin nt$? In other words, what is a solution to

$$\ddot{x} + 50x = \sin nt$$

with the same (smallest) period as $\sin nt$?
**Solution:** First find the response to $e^{int}$, and then take the imaginary part. In other words, we first solve

$$\ddot{z} + 50z = e^{int}.$$  

The characteristic polynomial is $p(r) = r^2 + 50$, so by ERF, the system response to $e^{int}$ is

$$z = \frac{1}{p(in)}e^{int} = \frac{1}{50 - n^2}e^{int}$$

(this is the solution we want since it has the right period). The complex gain is $\frac{1}{50 - n^2}$. Then

$$x = \text{Im} \left( \frac{1}{50 - n^2}e^{int} \right) = \frac{1}{50 - n^2} \sin nt$$

is the system response to $\sin nt$. This explains all the rows of the table below except the last row.

<table>
<thead>
<tr>
<th>input signal</th>
<th>system response</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e^{int}$</td>
<td>$\frac{1}{50 - n^2}e^{int}$</td>
</tr>
<tr>
<td>$\sin nt$</td>
<td>$\frac{1}{50 - n^2}\sin nt$</td>
</tr>
<tr>
<td>$\sin t$</td>
<td>$\frac{1}{49}\sin t$</td>
</tr>
<tr>
<td>$\sin 2t$</td>
<td>$\frac{1}{46}\sin 2t$</td>
</tr>
<tr>
<td>$\sin 3t$</td>
<td>$\frac{1}{41}\sin 3t$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$\sum_{n \geq 1} b_n \sin nt$</td>
<td>$\sum_{n \geq 1} \frac{1}{50 - n^2}b_n \sin nt$</td>
</tr>
</tbody>
</table>

17.11.2. **System response to Fourier series input.**

Now let’s return to the original problem. Suppose that the input signal $f$ is an odd periodic function of period $2\pi$. Since $f$ is odd, the Fourier series of $f$ is a linear combination of the shape

$$f(t) = b_1 \sin t + b_2 \sin 2t + b_3 \sin 3t + \cdots.$$  

By the superposition principle, the system response to $f(t)$ is

$$x(t) = b_1 \frac{1}{49} \sin t + b_2 \frac{1}{46} \sin 2t + b_3 \frac{1}{41} \sin 3t + \cdots.$$
Note that each Fourier component \( \sin nt \) has a different gain: the gain depends on the frequency.

One could write the answer using sigma-notation:

\[
x(t) = \sum_{n \geq 1} \frac{1}{50 - n^2} b_n \sin nt.
\]

This is better since it shows precisely what every term in the series is (no need to “guess the pattern”). \( \square \)

17.11.3. Near resonance.

**Problem 17.15.** For which input signal \( \sin nt \) is the gain the largest?

**Solution:** The gain is \( \frac{1}{50 - n^2} \), which is largest when \(|50 - n^2|\) is smallest. This happens for \( n = 7 \). \( \square \)

The gain for \( \sin 7t \) is 1, and the next largest gain, occurring for \( \sin 6t \) and \( \sin 8t \), is \( \frac{1}{14} \). Thus the system approximately filters out all the Fourier components of \( f(t) \) except for the \( \sin 7t \) term.

**Problem 17.16.** Let \( x(t) \) be the periodic solution to

\[
\ddot{x} + 50x = \frac{\pi}{4} \text{Sq}(t).
\]

Which Fourier coefficient of \( x(t) \) is largest? Which is second largest?

**Solution:** The input signal

\[
\frac{\pi}{4} \text{Sq}(t) = \sum_{n \geq 1, \text{odd}} \frac{\sin nt}{n}
\]

elicits the system response

\[
x(t) = \sum_{n \geq 1, \text{odd}} \left( \frac{1}{50 - n^2} \right) \frac{\sin nt}{n}
\approx 0.020 \sin t + 0.008 \sin 3t + 0.008 \sin 5t + 0.143 \sin 7t - 0.003 \sin 9t - (\text{even smaller terms})
\]

so the coefficient of \( \sin 7t \) is largest, and the coefficient of \( \sin t \) is second largest. (This makes sense since the Fourier coefficient \( \frac{1}{(50 - n^2)n} \) is large only when one of \( n \) or \( 50 - n^2 \) is small.) \( \square \)

**Remark 17.17.** Even though the system response is a complicated Fourier series, with infinitely many terms, only one or two are significant, and the rest are negligible.
17.11.4. Pure resonance. What happens if we change 50 to 49 in the ODE?

Flashcard question: Which of the following is true of the ODE
\[ \ddot{x} + 49x = \frac{\pi}{4} \text{Sq}(t) \]?

Possible answers:
- There are no solutions.
- There is exactly one solution, but it is not periodic.
- There is exactly one solution, and it is periodic.
- There are infinitely many solutions, but none of them are periodic.
- There are infinitely many solutions, but only one of them is periodic.
- There are infinitely many solutions, and all of them are periodic.

Answer: There are infinitely many solutions, but none of them are periodic. Here is why: For \( n \neq 7 \), we can solve \( \ddot{x} + 49x = \sin nt \) using complex replacement and ERF since \( in \) is not a root of \( r^2 + 49 \). For \( n = 7 \), we can still solve \( \ddot{x} + 49x = \sin 7t \) (the existence and uniqueness theorem guarantees this), but the solution requires generalized ERF, and involves \( t \), and hence is not periodic: it turns out that one solution is \(-\frac{t}{14} \cos 7t\).

For the input signal \( \text{Sq}(t) \), we can find a solution \( x_p \) by superposition: most of the terms will be periodic, but one of them will be \( \frac{1}{7} \left(-\frac{t}{14} \cos 7t\right) \), and this makes the whole solution \( x_p \) non-periodic.

There are infinitely many other solutions, namely \( x_p + c_1 \cos 7t + c_2 \sin 7t \) for any \( c_1 \) and \( c_2 \), but these solutions still include the \( \frac{1}{7} \left(-\frac{t}{14} \cos 7t\right) \) term and hence are not periodic. \( \square \)

Remark 17.18. If the ODE had been
\[ \ddot{x} + 36x = \frac{\pi}{4} \text{Sq}(t) \]
then all solutions would have been periodic, because \( \frac{\pi}{4} \text{Sq}(t) \) has no sin 6\( t \) term in its Fourier series.

In general, for a periodic function \( f \), the ODE \( p(D)x = f(t) \) has a periodic solution if and only if for each term \( \cos \omega t \) or \( \sin \omega t \) appearing with a nonzero coefficient in the Fourier series of \( f \), the number \( i\omega \) is not a root of \( p(r) \).

17.12. Resonance with damping. In real life, there is always damping, and this prevents the runaway growth in the pure resonance scenario of the previous section.

Problem 17.19. Describe the steady-state solution to
\[ \ddot{x} + 0.1\dot{x} + 49x = \frac{\pi}{4} \text{Sq}(t). \]

Remark 17.20. The term 0.1\( \dot{x} \) is the damping term.
Recall: The steady-state solution is the periodic solution. (Other solutions will be a sum of the steady-state solution with a transient solution solving the homogeneous ODE
\[ \ddot{x} + 0.1\dot{x} + 49x = 0; \]
these transient solutions tend to 0 as \( t \to \infty \), because the coefficients of the characteristic polynomial are positive (in fact, this is an underdamped system).

Solution: First let’s solve
\[ \ddot{x} + 0.1\dot{x} + 49x = \sin nt. \]
Before doing that, solve the complex replacement ODE
\[ \ddot{z} + 0.1\dot{z} + 49z = e^{int}. \]
The characteristic polynomial is \( p(r) = r^2 + 0.1r + 49 \), so ERF gives
\[ z = \frac{1}{p(in)} e^{int} = \frac{1}{(49 - n^2) + (0.1n)i} e^{int}, \]
with complex gain \( \frac{1}{(49 - n^2) + (0.1n)i} \) and gain
\[ g_n := \frac{1}{|(49 - n^2) + (0.1n)i|}. \]
Thus
\[ x = \text{Im} \left( \frac{1}{(49 - n^2) + (0.1n)i} e^{int} \right); \]
this is a sinusoid of amplitude \( g_n \), so \( x = g_n \cos(nt - \phi_n) \) for some \( \phi_n \).

The input signal
\[ \frac{\pi}{4} \text{Sq}(t) = \sum_{n \geq 1, \text{odd}} \frac{\sin nt}{n}, \]
elicits the system response
\[ x(t) = \sum_{n \geq 1, \text{odd}} g_n \frac{\cos(nt - \phi_n)}{n} \]
\[ \approx 0.020 \cos(t - \phi_1) + 0.008 \cos(3t - \phi_3) + 0.008 \cos(5t - \phi_5) + 0.204 \cos(7t - \phi_7) + 0.003 \cos(9t - \phi_9) + \text{(even smaller terms)}. \]

Conclusion: The system response is almost indistinguishable from a pure sinusoid of angular frequency 7.
17.13. **Listening to Fourier series.** Try the “Fourier Coefficients: Complex with Sound” mathlet


If using headphones, start with a low volume, since pure sine waves carry more energy than they seem to, and can damage your hearing after sustained listening.

Your ear is capable of decomposing a sound wave into its Fourier components of different frequencies. Each frequency corresponds to a certain pitch. Increasing the frequency produces a higher pitch. More precisely, multiplying the frequency by a number greater than 1 increases the pitch by what in music theory is called an **interval**. For example, multiplying the frequency by 2 raises the pitch by an octave, and multiplying by 3 raises the pitch an octave plus a perfect fifth.

When an instrument plays a note, it is producing a periodic sound wave in which typically many of the Fourier coefficients are nonzero. In a general Fourier series, the combination of the first two nonconstant terms \((a_1 \cos t + b_1 \sin t, \text{ if the period is } 2\pi)\) is a sinusoid of some frequency \(\nu\), and the next combination (e.g., \(a_2 \cos 2t + b_2 \sin 2t\)) has frequency \(2\nu\), and so on: the frequencies are the positive integer multiples of the lowest frequency \(\nu\). The note corresponding to the frequency \(\nu\) is called the **fundamental**, and the notes corresponding to frequencies \(2\nu, 3\nu, \ldots\) are called the **overtones**.

The musical staffs below show these for \(\nu \approx 131\) Hz (the C below middle C), with the integer multiplier shown in green.

![Musical Staffs](image)

**Question 17.21.** Can you guess what note corresponds to \(9\nu\)? 
Can you hear the phases of the sinusoids? No.

17.14. **Fourier series of arbitrary period.** Everything we did with periodic functions of period $2\pi$ can be generalized to periodic functions of other periods.

**Problem 17.22.** Define

$$f(t) := \begin{cases} 1, & \text{if } 0 < t < L, \\ -1 & \text{if } -L < t < 0. \end{cases}$$

and extend it to a periodic function of period $2L$. Express this new square wave $f(t)$ in terms of $\text{Sq}$.

**Solution:** To avoid confusion, let’s use $u$ as the variable for $\text{Sq}$. Stretching the graph of $\text{Sq}(u)$ horizontally by a factor $L/\pi$ produces the graph of $f(t)$.

In other words, if $t$ and $u$ are related by $t = \frac{L}{\pi}u$ (so that $u = \pi$ corresponds to $t = L$), then $f(t) = \text{Sq}(u)$. In other words, $u = \frac{\pi t}{L}$, so

$$f(t) = \text{Sq}\left(\frac{\pi t}{L}\right). \quad \Box$$

Similarly we can stretch any function of period $2\pi$ to get a function of different period. Let $L$ be a positive real number. Start with “any” periodic function

$$g(u) = \frac{a_0}{2} + \sum_{n \geq 1} a_n \cos nu + \sum_{n \geq 1} b_n \sin nu,$$

of period $2\pi$. Stretching horizontally by a factor $L/\pi$ gives a periodic function $f(t)$ of period $2L$, and “every” $f$ of period $2L$ arises this way. By the same calculation as above,

$$f(t) = g\left(\frac{\pi t}{L}\right)$$

$$= \frac{a_0}{2} + \sum_{n \geq 1} a_n \cos \frac{n\pi t}{L} + \sum_{n \geq 1} b_n \sin \frac{n\pi t}{L}.$$
The substitution $u = \frac{\pi t}{L}$ (and $du = \frac{\pi}{L} dt$) also leads to Fourier coefficient formulas for period $2L$:

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} g(u) \cos nu \, du$$
$$= \frac{1}{\pi} \int_{-L}^{L} g\left(\frac{\pi t}{L}\right) \cos \left(\frac{n\pi t}{L}\right) \frac{\pi}{L} \, dt$$
$$= \frac{1}{L} \int_{-L}^{L} f(t) \cos \left(\frac{n\pi t}{L}\right) \, dt.$$ 

A similar formula gives $b_n$ in terms of $f$.

17.14.1. *The inner product for periodic functions of period $2L$.* Adapt the definition of the inner product to the case of functions $f$ and $g$ of period $2L$:

$$\langle f, g \rangle := \int_{-L}^{L} f(t)g(t) \, dt.$$ 

(This conflicts with the earlier definition of $\langle f, g \rangle$, for functions for which both make sense, so perhaps it would be better to write $\langle f, g \rangle_L$ for the new inner product, but we won’t bother to do so.)

The same calculations as before show that the functions

$$1, \cos \frac{\pi t}{L}, \cos \frac{2\pi t}{L}, \cos \frac{3\pi t}{L}, \ldots, \sin \frac{\pi t}{L}, \sin \frac{2\pi t}{L}, \sin \frac{3\pi t}{L}, \ldots$$

form an orthogonal basis for the vector space of “all” periodic functions of period $2L$, with

$$\langle 1, 1 \rangle = 2L$$
$$\langle \cos \frac{n\pi t}{L}, \cos \frac{n\pi t}{L} \rangle = L$$
$$\langle \sin \frac{n\pi t}{L}, \sin \frac{n\pi t}{L} \rangle = L$$

(the average value of $\cos^2 \omega t$ is $1/2$ for any $\omega$, and the average value of $\sin^2 \omega t$ is $1/2$ too).

This gives another way to derive the Fourier coefficient formulas for functions of period $2L$.


- Fourier’s theorem: “Every” periodic function $f$ of period $2L$ is a Fourier series

$$f(t) = \frac{a_0}{2} + \sum_{n \geq 1} a_n \cos \frac{n\pi t}{L} + \sum_{n \geq 1} b_n \sin \frac{n\pi t}{L}.$$
Given \( f \), the Fourier coefficients \( a_n \) and \( b_n \) can be computed using:

\[
a_n = \frac{1}{L} \int_{-L}^{L} f(t) \cos \frac{n\pi t}{L} \, dt \quad \text{for all } n \geq 0,
\]

\[
b_n = \frac{1}{L} \int_{-L}^{L} f(t) \sin \frac{n\pi t}{L} \, dt \quad \text{for all } n \geq 1.
\]

- If \( f \) is even, then only the cosine terms (including the \( a_0/2 \) term) appear.
- If \( f \) is odd, then only the sine terms appear.

**Problem 17.23.** Define

\[
s(t) := \begin{cases} 8, & \text{if } 0 < t < 5, \\ 2, & \text{if } -5 < t < 0. \end{cases}
\]

and extend it to a periodic function of period 10. Find the Fourier series for \( s(t) \).

**Solution:** One way would be to use the Fourier coefficient formulas directly. But we will instead obtain the Fourier series for \( s(t) \) from the Fourier series for \( \text{Sq}(t) \), by stretching and shifting.

First, stretch horizontally by a factor of \( 5/\pi \) to get

\[
\text{Sq} \left( \frac{\pi t}{5} \right) = \begin{cases} 1, & \text{if } 0 < t < 5, \\ -1, & \text{if } -5 < t < 0. \end{cases}
\]

Here the difference between the upper and lower values is 2, but for \( s(t) \) we want a difference of 6, so multiply by 3:

\[
3 \text{Sq} \left( \frac{\pi t}{5} \right) = \begin{cases} 3, & \text{if } 0 < t < 5, \\ -3, & \text{if } -5 < t < 0. \end{cases}
\]

Finally add 5:

\[
5 + 3 \text{Sq} \left( \frac{\pi t}{5} \right) = \begin{cases} 8, & \text{if } 0 < t < 5, \\ 2, & \text{if } -5 < t < 0. \end{cases}
\]

Since

\[
\text{Sq}(t) = \frac{4}{\pi} \sum_{n \geq 1, \text{ odd}} \frac{1}{n} \sin nt,
\]
we get

$$s(t) = 5 + 3 \text{Sq} \left( \frac{\pi t}{5} \right)$$

$$= 5 + 3 \left( \frac{4}{\pi} \right) \sum_{n \geq 1, \text{odd}} \frac{1}{n} \sin \frac{n\pi t}{5}$$

$$= 5 + \sum_{n \geq 1, \text{odd}} \frac{12}{n\pi} \sin \frac{n\pi t}{5}.$$

17.15. Convergence of a Fourier series.

**Definition 17.24.** A periodic function $f$ of period $2L$ is called **piecewise differentiable** if

- there are at most finitely many points in $[-L, L)$ where $f'(t)$ does not exist, and
- at each such point $\tau$, the left limit $f(\tau^-) := \lim_{t \to \tau^-} f(t)$ and right limit $f(\tau^+) := \lim_{t \to \tau^+} f(t)$ exist (although they might be unequal, in which case we say that $f$ has a **jump discontinuity** at $\tau$).

**Theorem 17.25.** If $f$ is a piecewise differentiable periodic function, then the Fourier series of $f$ (with the $a_n$ and $b_n$ defined by the Fourier coefficient formulas)

- converges to $f(t)$ at values of $t$ where $f$ is continuous, and
- converges to $\frac{f(t^-) + f(t^+)}{2}$ where $f$ has a jump discontinuity.

**Example 17.26.** The left limit $\text{Sq}(0^-) = -1$ and right limit $\text{Sq}(0^+) = 1$ average to 0. The Fourier series

$$\frac{4}{\pi} \left( \sin t + \frac{\sin 3t}{3} + \frac{\sin 5t}{5} + \cdots \right)$$

evaluated at $t = 0$ converges to 0 too.

17.16. **Antiderivative of a Fourier series.** Suppose that $f$ is a piecewise differentiable periodic function, and that $F$ is an antiderivative of $f$. (If $f$ has jump discontinuities, one can still define $F(t) := \int_0^t f(\tau) d\tau + C$, but at the jump discontinuities $F$ will be only continuous, not differentiable.)

The function $F$ might not be periodic. For example, if $f$ is a function of period 2 such that

$$f(t) := \begin{cases} 
2, & \text{if } 0 < t < 1, \\
-1, & \text{if } -1 < t < 0,
\end{cases}$$

then $F(t)$ creeps upward over time.
An even easier example: if \( f(t) = 1 \), then \( F(t) = t + C \) for some \( C \), so \( F(t) \) is not periodic.

But if the constant term \( a_0/2 \) in the Fourier series of \( f \) is 0, then \( F \) is periodic, and its Fourier series can be obtained by taking the simplest antiderivative of each cosine and sine term, and adding an overall \( +C \), where \( C \) is the average value of \( F \).

**Problem 17.27.** Let \( T(t) \) be the periodic function of period 2 such that \( T(t) = |t| \) for \(-1 \leq t \leq 1\); this is called a **triangle wave**. Find the Fourier series of \( T(t) \).

**Solution:** We could use the Fourier coefficient formula. But instead, notice that \( T(t) \) has slope \(-1\) on \((-1,0)\) and slope \(1\) on \((0,1)\), so \( T(t) \) is an antiderivative of the period 2 square wave

\[
\text{Sq}(\pi t) = \sum_{n \geq 1, \text{odd}} \frac{4}{n\pi} \sin n\pi t.
\]
Taking an antiderivative termwise (and using that the average value of $T(t)$ is $1/2$) gives

$$T(t) = \frac{1}{2} + \sum_{n \geq 1, \text{ odd}} \frac{4}{n\pi} \left( -\cos \frac{n\pi t}{n\pi} \right)$$

$$= \frac{1}{2} - \sum_{n \geq 1, \text{ odd}} \frac{4}{n^{2}\pi^{2}} \cos n\pi t. \quad \square$$

**Warning:** If a periodic function $f$ is not continuous, it will not be an antiderivative of any piecewise differentiable function, so you cannot find the Fourier series of $f$ by integration.

**Remark 17.28.** A Fourier series of a piecewise differentiable periodic function $f$ can also be differentiated termwise, but the result will often fail to converge. For example, the termwise derivative of the Fourier series $S_{c}(t)$ gives a nonsensical value at $t = 0$. (Here is one good case, however: If $f$ is continuous and piecewise twice differentiable, then the derivative series converges.)

---

**April 23**

18. Boundary value problems

**Problem 18.1.** Find all nonzero functions $v(x)$ on $[0, \pi]$ satisfying $v''(x) = \lambda v(x)$ for a constant $\lambda$ and satisfying the boundary conditions $v(0) = 0$ and $v(\pi) = 0$.

18.1. **Failure of existence and uniqueness.** Is the following argument valid?

“For each fixed constant $\lambda$, this is a second-order linear ODE, and there are two conditions, so the existence and uniqueness theorem says that there is exactly one solution. The function $v = 0$ is a solution, so the only solution is $v = 0$.”

**Answer:** No. The two conditions are not *initial conditions* at the same point, so the existence and uniqueness theorem does not apply. In fact, we’re going to see that for certain values of $c$, there is more than one solution.

18.2. **Solving a boundary value problem.**

**Solution to Problem 18.1** This is a homogeneous linear ODE with characteristic polynomial $r^{2} - \lambda$.

Case 1: $\lambda > 0$. Then the general solution is $ae^{\sqrt{\lambda}x} + be^{-\sqrt{\lambda}x}$, and the boundary conditions say

$$a + b = 0$$

$$ae^{\sqrt{\lambda}\pi} + be^{-\sqrt{\lambda}\pi} = 0.$$
Since \( \det \begin{pmatrix} 1 & 1 \\ e^{\sqrt{\lambda} \pi} & e^{-\sqrt{\lambda} \pi} \end{pmatrix} \neq 0 \), the only solution to this linear system is \((a, b) = (0, 0)\). Thus there are no nonzero solutions \(v\).

**Case 2:** \(\lambda = 0\). Then the general solution is \(a + bx\), and the boundary conditions say

\[
a = 0 \quad \quad a + b\pi = 0.
\]

Again the only solution to this linear system is \((a, b) = (0, 0)\). Thus there are no nonzero solutions \(v\).

**Case 3:** \(\lambda < 0\). We can write \(\lambda = -\omega^2\) for some \(\omega > 0\). Then the roots of the characteristic polynomial are \(\pm i\omega\), and the general solution is \(a\cos \omega x + b\sin \omega x\). The first boundary condition says \(a = 0\), so \(v = b\sin \omega x\). The second boundary condition then says \(b\sin \omega \pi = 0\). We are looking for nonzero solutions \(v\), so we can assume that \(b \neq 0\). Then \(\sin \omega \pi = 0\), so \(\omega\) is an integer \(n\); also \(n > 0\), since \(\omega > 0\).

**Conclusion:** There exist nonzero solutions if and only if \(\lambda = -n^2\) for some positive integer \(n\); in that case, all solutions are of the form \(b\sin nx\). \(\Box\)

We will use this conclusion as one step in the solution of the heat equation.

18.3. **Analogy with eigenvalue-eigenvector problems.** To describe a function \(v(x)\), one needs to give infinitely many numbers, namely its values at all the different input \(x\)-values. Thus \(v(x)\) is like a vector of infinite length.

The linear differential operator \(\frac{d^2}{dx^2}\) maps each function to a function, just as a \(2 \times 2\) matrix defines a linear transformation mapping each vector in \(\mathbb{R}^2\) to another vector in \(\mathbb{R}^2\). Thus \(\frac{d^2}{dx^2}\) is like an \(\infty \times \infty\) matrix.

The ODE \(\frac{d^2}{dx^2}v = \lambda v\) (with boundary conditions) amounts to an infinite system of equations: the ODE consists of one equality of numbers at each \(x \in (0, \pi)\), and boundary conditions are equalities at the endpoints. Thus the ODE with boundary conditions is like a system of equations \(Av = \lambda v\). Nonzero solutions \(v(x)\) to \(\frac{d^2}{dx^2}v = \lambda v\) exist only for special values of \(\lambda\), namely

\[
\lambda = -1, -4, -9, \ldots,
\]

just as \(Av = \lambda v\) has a nonzero solution \(v\) only for special values of \(\lambda\), namely the eigenvalues of \(\lambda\). But the differential operator \(\frac{d^2}{dx^2}\) has infinitely many eigenvalues, as one would expect for an \(\infty \times \infty\) matrix.

The nonzero solutions \(v(x)\) to \(\frac{d^2}{dx^2}v = \lambda v\) satisfying the boundary conditions are called **eigenfunctions**, since they act like eigenvectors.
Summary of the analogies:

<table>
<thead>
<tr>
<th>vector ( \mathbf{v} )</th>
<th>function ( v(x) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>the linear operator ( \frac{d^2}{dx^2} )</td>
</tr>
<tr>
<td>eigenvalue-eigenvector problem</td>
<td>boundary value problem</td>
</tr>
<tr>
<td>( Av = \lambda v )</td>
<td>( \frac{d^2}{dx^2}v = \lambda v, \ v(0) = 0, \ v(\pi) = 0 )</td>
</tr>
<tr>
<td>eigenvalues ( \lambda )</td>
<td>eigenvalues ( \lambda = -1, -4, -9, \ldots )</td>
</tr>
<tr>
<td>eigenvectors ( \mathbf{v} )</td>
<td>eigenfunctions ( v(x) = \sin nx )</td>
</tr>
</tbody>
</table>

18.4. Another little fact to be used in solution of the heat equation.

**Lemma 18.2.** Suppose that \( f(x) \) and \( g(t) \) are functions of independent variables \( x \) and \( t \), respectively. If \( f(x) = g(t) \) for all values of \( x \) and \( t \), then there is a constant \( \lambda \) such that \( f(x) = \lambda \) for all \( x \) and \( g(t) = \lambda \) for all \( t \).

**Proof.** Both sides of \( f(x) = g(t) \) equal the same function; because it equals the left hand side, it does not depend on \( x \); because it equals the right hand side, it does not depend on \( t \) either. Thus both sides equal a constant function, which may be called \( \lambda \). \( \square \)

19. Heat equation


Probably much of the physics should be skipped in lecture.

**Problem 19.1.** An insulated uniform metal rod with exposed ends starts at a constant temperature, but then its ends are held in ice. Model its temperature.

**Variables and functions:** Define

- \( L \): length of the rod
- \( A \): cross-sectional area of the rod
- \( u_0 \): initial temperature of the rod
- \( t \): time
- \( x \): position along the rod (from 0 to \( L \))
- \( u \): temperature at a point of the rod at a given time
- \( q \): heat flux density at a point of the rod at a given time (to be explained).

Here

- \( L \), \( A \), and \( u_0 \) are constants;
- \( t \) and \( x \) are independent variables; and
- \( u = u(x, t) \) and \( q = q(x, t) \) are functions defined for \( x \in [0, L] \) and \( t \geq 0 \).
Physics: Each bit of the rod contains internal energy, consisting of the microscopic kinetic energy of particles (and the potential energy associated with microscopic forces). This energy can be transferred from point to point, via atoms colliding with nearby atoms. Heat flux density measures such heat transfer from left to right across a cross-section of the rod, per unit area, per unit time.

We will use three laws of physics:

1. The **first law of thermodynamics** (conservation of energy), in the special case in which no work is being done, states that for any bit of the rod,

   \[
   \text{(increase in internal energy)} = \text{(net amount of heat flowing in)}.
   \]

2. For any bit of the rod,

   \[
   \frac{\text{(increase in internal energy)}}{\text{volume}} \sim \text{(increase in temperature)}
   \]

   (Today we’ll use $\sim$ to denote “proportional to”.) The constant of proportionality depends on the material.

3. **Fourier’s law of heat transfer:**

   \[
   q \sim -\frac{\partial u}{\partial x}.
   \]

   This makes sense: If $u(x + dx, t)$ is greater than $u(x, t)$, then the heat flow at $x$ is to the left (negative), and the rate of heat flow is proportional to the (infinitesimal) difference of temperature $u(x + dx, t) - u(x, t)$, just as in Newton’s law of cooling.

**Deducing the PDE:** For any interior bit of rod defined by the interval $[x, x + dx]$, during a time interval $[t, t + dt]$, the first law of thermodynamics states

\[
\text{(increase in internal energy)} = \text{(heat flowing in)} - \text{(heat flowing out)}
\]

\[
\text{(increase in temperature)}(\text{volume}) \sim \text{(heat flowing in)} - \text{(heat flowing out)}
\]

\[
(u(x, t + dt) - u(x, t)) A dx \sim q(x, t) A dt - q(x + dx, t) A dt.
\]

Divide by $A dx dt$ to get

\[
\frac{\partial u}{\partial t} \sim -\frac{\partial q}{\partial x}.
\]

(More correct would be to use $\Delta x$, $\Delta t$, and so on, and to take a limit, but the end result is the same.) Finally, substitute Fourier’s law \( q \sim -\frac{\partial u}{\partial x} \) into the right hand side to get the **heat equation**

\[
\frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2},
\]

for some constant $\alpha > 0$ (called thermal diffusivity) that depends only on the material. The heat equation is a second-order homogeneous linear partial differential equation involving the unknown function $u(x, t)$. 

146
Remark 19.2. This PDE makes physical sense, since if the temperature profile (graph of $u(x, t)$ versus $x$ at a fixed time) is curving upward at a point ($\frac{\partial^2 u}{\partial x^2} > 0$), then the average of the point’s neighbors is warmer than the point, so the point’s temperature should increase.

Boundary conditions: $u(0, t) = 0$ and $u(L, t) = 0$ for all $t \geq 0$ (for $u$ in degrees Celsius).

Initial condition: $u(x, 0) = u_0$ for all $x \in (0, L)$.

Try the “Heat Equation” mathlet http://mathlets.org/mathlets/heat-equation/

19.2. **Solving the PDE with homogeneous boundary conditions: separation of variables; normal modes.** Let’s now try to solve the PDE. For simplicity, suppose that $L = \pi$, $u_0 = 1$, and $\alpha = 1$. (The general case is similar. In fact, one could reduce to this special case by changes of variable.)

So now we are solving

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}$$

$$u(0, t) = 0 \quad \text{for } t \geq 0$$

$$u(\pi, t) = 0 \quad \text{for } t \geq 0$$

$$u(x, 0) = 1 \quad \text{for } x \in (0, \pi).$$

**Idea (separation of variables):** Forget about the initial condition $u(x, 0) = 1$ for now, but look for nonzero solutions of the form

$$u(x, t) = w(t) v(x)$$

Substituting into the PDE gives

$$\dot{w}(t) v(x) = w(t) v''(x)$$

$$\frac{\dot{w}(t)}{w(t)} = \frac{v''(x)}{v(x)}.$$ 

(at least where $w(t)$ and $v(x)$ are nonzero). By Lemma 18.2, there is a constant $\lambda$ such that

$$\frac{v''(x)}{v(x)} = \lambda \quad \text{and} \quad \frac{\dot{w}(t)}{w(t)} = \lambda,$$

or in other words,

$$v''(x) = \lambda v(x) \quad \text{and} \quad \dot{w}(t) = \lambda w(t).$$

Substituting $u(x, t) = w(t) v(x)$ into the first boundary condition $u(0, t) = 0$ gives $w(t)v(0) = 0$ for all $t$, but $w(t)$ is not the zero function, so this translates into $v(0) = 0$. Similarly, the second boundary condition $u(\pi, t) = 0$ translates into $v(\pi) = 0$. 


We already solved \( v''(x) = \lambda v(x) \) subject to the boundary conditions \( v(0) = 0 \) and \( v(\pi) = 0 \): nonzero solutions \( v(x) \) exist only if \( \lambda = -n^2 \) for some positive integer \( n \), and in that case \( v(x) \) is a scalar times \( \sin nx \).

For \( \lambda = -n^2 \), what is a matching possibility for \( w \)? Since \( \dot{w} = -n^2w \), the function \( w \) is a scalar times \( e^{-n^2t} \).

This gives rise to one solution
\[
u(x, t) = e^{-n^2t} \sin nx
\]
(and its scalar multiples) for each positive integer \( n \), to the PDE with boundary conditions. Each such solution is called a normal mode.

The PDE and boundary conditions are homogeneous, so we can get other solutions by taking linear combinations:
\[
u(x, t) = b_1 e^{-t} \sin x + b_2 e^{-4t} \sin 2x + b_3 e^{-9t} \sin 3x + \cdots
\]
(12)
This turns out to be the general solution to the PDE with the boundary conditions.

---

April 25

Summary of last lecture:

- We modeled an insulated metal rod with exposed ends held at 0°C.
- Using physics, we found that its temperature \( u(x, t) \) was governed by the PDE
  \[
  \frac{\partial u}{\partial t} = \alpha \frac{\partial^2 u}{\partial x^2}
  \]
  (the heat equation).
  For simplicity, we specialized to the case \( \alpha = 1 \), length \( \pi \), and initial temperature \( u(x, 0) = 1 \).
- Trying \( u = w(t)v(x) \) led to separate ODEs for \( v \) and \( w \), leading to solutions \( e^{-n^2t} \sin nx \) for \( n = 1, 2, \ldots \) to the PDE with boundary conditions.
- We took linear combinations to get the general solution
  \[
u(x, t) = b_1 e^{-t} \sin x + b_2 e^{-4t} \sin 2x + b_3 e^{-9t} \sin 3x + \cdots
\]
to the PDE with boundary conditions.

19.3. **Initial condition.** As usual, we postponed imposing the initial condition, but now it is time to impose it.

**Question 19.3.** Which choices of \( b_1, b_2, \ldots \) make the general solution above also satisfy the initial condition \( u(x, 0) = 1 \) for \( x \in (0, \pi) \)?
Set $t = 0$ in (12) and use the initial condition on the left to get

$$1 = b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots \quad \text{for } x \in (0, \pi),$$

which must be solved for $b_1, b_2, \ldots$. Section 17.8 showed how to find such $b_i$: the left hand side extends to an odd period $2\pi$ function, namely $\text{Sq}(x)$, so we need to solve

$$\text{Sq}(x) = b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots \quad \text{for all } x \in \mathbb{R}.$$

We already know the answer:

$$\text{Sq}(x) = \frac{4}{\pi} \sin x + \frac{4}{3\pi} \sin 3x + \frac{4}{5\pi} \sin 5x + \cdots.$$ 

In other words $b_n = 0$ for even $n$, and $b_n = \frac{4}{n\pi}$ for odd $n$. Substituting these $b_n$ back into (12) gives

$$u(x, t) = \frac{4}{\pi} e^{-t} \sin x + \frac{4}{3\pi} e^{-9t} \sin 3x + \frac{4}{5\pi} e^{-25t} \sin 5x + \cdots. \quad \Box$$

**Question 19.4.** What does the temperature profile look like when $t$ is large?

**Answer:** All the Fourier components are decaying, so $u(x, t) \to 0$ as $t \to +\infty$ at every position. Thus the temperature profile approaches a horizontal segment, the graph of the zero function. But the Fourier components of higher frequency decay much faster than the first Fourier component, so when $t$ is large, the formula

$$u(x, t) \approx \frac{4}{\pi} e^{-t} \sin x$$

is a very good approximation. Eventually, the temperature profile is indistinguishable from a sinusoid of angular frequency 1 whose amplitude is decaying to 0. This is what was observed in the mathlet. \Box

19.4. **Analogy between a linear system of ODEs and the heat equation.** We can continue the table of analogies from Section 18.3.
<table>
<thead>
<tr>
<th>vector $\mathbf{v}$</th>
<th>function $v(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>the linear operator $\frac{d^2}{dx^2}$</td>
</tr>
<tr>
<td>eigenvalue-eigenvector problem</td>
<td>boundary value problem</td>
</tr>
<tr>
<td>$Av = \lambda v$</td>
<td>$\frac{d^2}{dx^2}v = \lambda v$, $v(0) = 0$, $v(\pi) = 0$</td>
</tr>
<tr>
<td>eigenvalues $\lambda$</td>
<td>eigenvalues $\lambda = -1, -4, -9, \ldots$</td>
</tr>
<tr>
<td>eigenvectors $\mathbf{v}$</td>
<td>eigenfunctions $v(x) = \sin nx$</td>
</tr>
<tr>
<td>linear system of ODEs</td>
<td>heat equation with boundary conditions</td>
</tr>
<tr>
<td>$\dot{x} = Ax$</td>
<td>$\dot{u} = \frac{\partial^2}{\partial x^2}u$, $u(0,t) = 0$, $u(\pi,t) = 0$</td>
</tr>
<tr>
<td>normal modes: $e^{\lambda t}\mathbf{v}$</td>
<td>normal modes: $e^{\lambda t}v(x) = e^{-n^2t}\sin nx$</td>
</tr>
<tr>
<td>for an eigenvector $\mathbf{v}$ with eigenvalue $\lambda$</td>
<td>for eigenfunction $v(x) = \sin nx$, eigenvalue $\lambda = -n^2$</td>
</tr>
<tr>
<td>General solution: $\mathbf{u} = \sum c_n e^{\lambda_n t}\mathbf{v}_n$</td>
<td>General solution: $u = \sum b_n e^{-n^2t}\sin nx$</td>
</tr>
<tr>
<td>Solve $\mathbf{u}(0) = \sum c_n\mathbf{v}_n$ to get the $c_n$</td>
<td>Solve $u(x,0) = \sum b_n\sin nx$ to get the $b_n$</td>
</tr>
</tbody>
</table>

19.5. Solving the PDE with inhomogeneous boundary conditions.

Steps to solve a linear PDE with inhomogeneous boundary conditions:

1. Find a particular solution $u_p$ to the PDE with the inhomogeneous boundary conditions (but without initial conditions). If the boundary conditions do not depend on $t$, try to find the steady-state solution $u_p(x,t)$, i.e., the solution that does not depend on $t$.
2. Find the general solution $u_h$ to the PDE with the homogeneous boundary conditions.
3. Then $u := u_p + u_h$ is the general solution to the PDE with the inhomogeneous boundary conditions.
4. If initial conditions are given, use them to find the specific solution to the PDE with the inhomogeneous boundary conditions. (This often involves finding Fourier coefficients.)

**Problem 19.5.** Consider the same insulated uniform metal rod as before ($\alpha = 1$, length $\pi$, initial temperature $1^\circ C$), but now suppose that the left end is held at $0^\circ C$ while the right end is held at $20^\circ C$. Now what is $u(x,t)$?

**Solution:**

1. Forget the initial condition for now, and look for a solution $u = u(x)$ that does not depend on $t$. Plugging this into the heat equation PDE gives $0 = \frac{\partial^2 u}{\partial x^2}$. The general solution to this simplified DE is $u(x) = ax + b$. Imposing the boundary conditions $u(0) = 0$ and $u(\pi) = 20$ leads to $b = 0$ and $a = 20/\pi$, so $u_p = \frac{20}{\pi}x$.

2. The PDE with the homogeneous boundary conditions is what we solved earlier; the general solution is

$$u_h = b_1e^{-t}\sin x + b_2e^{-4t}\sin 2x + b_3e^{-9t}\sin 3x + \cdots.$$
3. The general solution to the PDE with inhomogeneous boundary conditions is
\[ u(x,t) = u_p + u_h = \frac{20}{\pi} x + b_1 e^{-t} \sin x + b_2 e^{-4t} \sin 2x + b_3 e^{-9t} \sin 3x + \cdots. \quad (13) \]

4. To find the \( b_n \), set \( t = 0 \) and use the initial condition on the left:
\[ 1 = \frac{20}{\pi} x + b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots \quad \text{for all } x \in (0, \pi). \]
\[ 1 - \frac{20}{\pi} x = b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots \quad \text{for all } x \in (0, \pi). \]

Extend \( 1 - \frac{20}{\pi} x \) on \( (0, \pi) \) to an odd periodic function \( f(x) \) of period \( 2\pi \). Then use the Fourier coefficient formulas to find the \( b_n \) such that
\[ f(x) = b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \cdots; \]
alternatively, find the Fourier series for the odd periodic extensions of \( x \) separately, and take a linear combination to get \( 1 - \frac{20}{\pi} x \). (We skipped this part in lecture.) Once the \( b_n \) are found, plug them back into (13). \( \square \)

19.6. Insulated ends.

**Problem 19.6.** Consider the same insulated uniform metal rod as before \((\alpha = 1, \text{length } \pi)\), but now assume that the ends are insulated too (instead of exposed and held in ice), and that the initial temperature is given by \( u(x,0) = x \) for \( x \in (0, \pi) \). Now what is \( u(x,t) \)?

**Solution:** As usual, we temporarily forget the initial condition, and use it only at the end.

“Insulated ends” means that there is zero heat flow through the ends, so the heat flux density function \( q \sim -\frac{\partial u}{\partial x} \) is 0 when \( x = 0 \) or \( x = \pi \). In other words, “insulated ends” means that the boundary conditions are
\[ \frac{\partial u}{\partial x}(0,t) = 0, \quad \frac{\partial u}{\partial x}(\pi,t) = 0 \quad \text{for all } t > 0, \quad (14) \]
instead of \( u(0,t) = 0 \) and \( u(\pi,t) = 0 \). So we need to solve the heat equation
\[ \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \]
with the boundary conditions (14). Separation of variables \( u(x,t) = v(x) w(t) \) leads to
\[ v''(x) = \lambda v(x) \quad \text{with } v'(0) = 0 \text{ and } v'(\pi) = 0 \]
\[ w(t) = \lambda w(t). \]
for a constant \( \lambda \). Looking at the cases \( \lambda > 0, \lambda = 0, \lambda < 0 \), we find that
\[ \lambda = -n^2 \quad \text{and} \quad v(x) = \cos nx \text{ (times a scalar)} \]

}\[ v(x) = \cos nx \text{ (times a scalar)} \]
where $n$ is one of 0, 1, 2, \ldots. (This time $n$ starts at 0 since $\cos 0x$ is a nonzero function.) For each such $v(x)$, the corresponding $w$ is $w(t) = e^{-n^2 t}$ (times a scalar), and the normal mode is

$$u = e^{-n^2 t} \cos nx.$$ 

The case $n = 0$ is the constant function 1, so the general solution is

$$u(x, t) = \frac{a_0}{2} + a_1 e^{-t} \cos x + a_2 e^{-4t} \cos 2x + a_3 e^{-9t} \cos 3x + \cdots.$$ 

Lecture actually ended here.

Finally, we bring back the initial condition: substitute $t = 0$ and use the initial condition on the left to get

$$x = \frac{a_0}{2} + a_1 \cos x + a_2 \cos 2x + a_3 \cos 3x + \cdots$$

for all $x \in (0, \pi)$. The right hand side is a period $2\pi$ even function, so extend the left hand side to a period $2\pi$ even function $T(x)$, a triangle wave, which is an antiderivative of

$$\text{Sq}(x) = \frac{4}{\pi} \left( \sin x + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \cdots \right).$$

Integration gives

$$T(x) = \frac{a_0}{2} - \frac{4}{\pi} \left( \cos x + \frac{\cos 3x}{9} + \frac{\cos 5x}{25} + \cdots \right),$$

and the constant term $a_0/2$ is the average value of $T(x)$, which is $\pi/2$. Thus

$$T(x) = \frac{\pi}{2} - \frac{4}{\pi} \left( \cos x + \frac{\cos 3x}{9} + \frac{\cos 5x}{25} + \cdots \right)$$

$$u(x, t) = \frac{\pi}{2} - \frac{4}{\pi} \left( e^{-t} \cos x + e^{-9t} \cos 3x + \frac{e^{-25t} \cos 5x}{25} + \cdots \right).$$

This answer makes physical sense: when the entire bar is insulated, its temperature tends to a constant equal to the average of the initial temperature. \[\square\]

April 28

20. Wave equation

The wave equation is a PDE that models light waves, sound waves, waves along a string, etc.
20.1. **Modeling: vibrating string.**

Probably much of the physics should be skipped in lecture.

**Problem 20.1.** Model a vibrating guitar string.

**Variables and functions:** Define

\[
\begin{align*}
L &: \text{ length of the string} \\
\rho &: \text{ mass per unit length} \\
T &: \text{ magnitude of the tension force} \\
t &: \text{ time} \\
x &: \text{ position along the string (from 0 to } L) \\
u &: \text{ vertical displacement of a point on the string}
\end{align*}
\]

Here

- \(L, \rho, T\) are constants;
- \(t, x\) are independent variables; and
- \(u = u(x, t)\) is a function defined for \(x \in [0, L]\) and \(t \geq 0\). The vertical displacement is measured relative to the equilibrium position in which the string makes a straight line.

At any given time \(t\), the string is in the shape of the graph of \(u(x, t)\) as a function of \(x\).

**Assumption:** The string is taut, so the vertical displacement of the string is small, and the slope of the string at any point is small.

Consider the piece of string between positions \(x\) and \(x + dx\). Let \(\theta\) be the (small) angle formed by the string and the horizontal line at position \(x\), and let \(\theta + d\theta\) be the same angle at position \(x + dx\).
Newton’s second law says that $ma = F$. Taking the vertical component of each side gives

$$\rho\,dx \cdot \frac{\partial^2 u}{\partial t^2} = T \sin(\theta + d\theta) - T \sin \theta = T \, d(\sin \theta).$$

Side calculation:

$$d(\sin \theta) = \cos \theta \, d\theta$$
$$d(\tan \theta) = \frac{1}{\cos^2 \theta} \, d\theta,$$

but $\cos \theta = 1 - \frac{\theta^2}{2!} + \cdots \approx 1$, so up to a factor that is very close to 1 we get

$$d(\sin \theta) \approx d(\underbrace{\tan \theta}_{\text{slope of string}}) = d \left( \frac{\partial u}{\partial x} \right).$$

Substituting this in gives

$$\rho\,dx \cdot \frac{\partial^2 u}{\partial t^2} \approx T \left( \frac{\partial u}{\partial x} \right).$$
Divide by $\rho \, dx$ to get

$$\frac{\partial^2 u}{\partial t^2} \approx T\rho^{-1} d\left(\frac{\partial u}{\partial x}\right) \approx T\rho^{-1} \frac{\partial^2 u}{\partial x^2}. $$

If we define a new constant $c := \sqrt{T\rho^{-1}}$, then this becomes the wave equation:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}. $$

This makes sense intuitively, since at places where the graph of the string is concave up ($\frac{\partial^2 u}{\partial x^2} > 0$) the tension pulling on both sides should combine to produce an upward force, and hence an upward acceleration.

Comparing units of both sides of the wave equation shows that the units for $c$ are m/s. The physical meaning of $c$ as a velocity will be explained later.

The ends of a guitar string are fixed, so we have boundary conditions

$$u(0, t) = 0 \text{ for all } t \geq 0,$$

$$u(L, t) = 0 \text{ for all } t \geq 0.$$

20.2. Separation of variables in PDEs; normal modes. For simplicity, suppose that $c = 1$ and $L = \pi$. So now we are solving the PDE with boundary conditions

$$\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2},$$

$$u(0, t) = 0,$$

$$u(\pi, t) = 0.$$

As with the heat equation, we try separation of variables. In other words, try to find normal modes of the form

$$u(x, t) = v(x)w(t),$$

for some nonzero functions $v(x)$ and $w(t)$. Substituting this into the PDE gives

$$v(x)\ddot{w}(t) = v''(x)w(t)$$

$$\frac{\ddot{w}(t)}{w(t)} = \frac{v''(x)}{v(x)}.$$

As usual, a function of $t$ can equal a function of $x$ only if both are equal to the same constant, say $\lambda$, so this breaks into two ODEs:

$$\ddot{w}(t) = \lambda w(t),$$

$$v''(x) = \lambda v(x).$$

Moreover, the boundary conditions become $v(0) = 0$ and $v(\pi) = 0.$
We already solved the eigenfunction equation \( v''(x) = \lambda v(x) \) with the boundary conditions \( v(0) = 0 \) and \( v(\pi) = 0 \): nonzero solutions exist only when \( \lambda = -n^2 \) for some positive integer \( n \), and in this case \( v = \sin nx \) (times a scalar). What is different this time is that \( w \) satisfies a second-order ODE

\[
\ddot{w}(t) = -n^2 w(t).
\]

The characteristic polynomial is \( r^2 + n^2 \), which has roots \( \pm in \), so

\[
w(t) := \cos nt \quad \text{and} \quad w(t) := \sin nt
\]

are possibilities (and all the others are linear combinations). Multiplying each by the \( v(x) \) with the matching \( \lambda \) gives the normal modes

\[
\cos nt \sin nx, \quad \sin nt \sin nx.
\]

Any linear combination

\[
u(x, t) = \sum_{n \geq 1} a_n \cos nt \sin nx + \sum_{n \geq 1} b_n \sin nt \sin nx
\]

is a solution to the PDE with boundary conditions, and this turns out to be the general solution.

20.3. Initial conditions. To specify a unique solution, give two initial conditions: not only the initial position \( u(x, 0) \), but also the initial velocity \( \frac{\partial u}{\partial t}(x, 0) \), at each position of the string. (That two initial conditions are needed is related to the fact that the PDE is second-order in the \( t \) variable.)

For a plucked string, it is reasonable to assume that the initial velocity is 0, so one initial condition is \( \frac{\partial u}{\partial t}(x, 0) = 0 \). What condition does this impose on the \( a_n \) and \( b_n \)? Well, for the general solution above,

\[
\frac{\partial u}{\partial t} = \sum_{n \geq 1} -na_n \sin nt \sin nx + \sum_{n \geq 1} nb_n \cos nt \sin nx
\]

\[
\frac{\partial u}{\partial t}(x, 0) = \sum_{n \geq 1} nb_n \sin nx,
\]

so the initial condition says that \( b_n = 0 \) for every \( n \); in other words,

\[
u(x, t) = \sum_{n \geq 1} a_n \cos nt \sin nx.
\]

If we also knew the initial position \( u(x, 0) \), we could solve for the \( a_n \) by extending to an odd, period \( 2\pi \) function of \( x \) and using the Fourier coefficient formula.

Try the “Wave equation” mathlet

http://math.mit.edu/~jmc/18.03/waveEquation.html
20.4. D’Alembert’s solution: traveling waves. D’Alembert figured out another way to write down solutions, in the case when \( u(x, t) \) is defined for all real numbers \( x \) instead of just \( x \in [0, L] \). Then, for any reasonable function \( f \),

\[
  u(x, t) := f(x - ct)
\]

is a solution to the PDE, as shown by the following calculations:

\[
\begin{align*}
  \frac{\partial u}{\partial t} &= f'(x - ct) \\
  \frac{\partial^2 u}{\partial t^2} &= (-c)^2 f''(x - ct)
\end{align*}
\]

\[
\begin{align*}
  \frac{\partial u}{\partial x} &= (-c)f'(x - ct) \\
  \frac{\partial^2 u}{\partial x^2} &= f''(x - ct),
\end{align*}
\]

so

\[
\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}.
\]

What is the physical meaning of this solution? At \( t = 0 \), we have \( u(x, 0) = f(x) \), so \( f(x) \) is the initial position. For any number \( t \), the position of the wave at time \( t \) is the graph of \( f(x - ct) \), which is the graph of \( f \) shifted \( ct \) units to the right. Thus the wave travels at constant speed \( c \) to the right, maintaining its shape.

The function \( u(x, t) := g(x + ct) \) (for any reasonable function \( g(x) \)) is a solution too, a wave moving to the left. It turns out that the general solution is a superposition

\[
  u(x, t) = f(x + ct) + g(x - ct).
\]

There is a tiny bit of redundancy: one can add a constant to \( f \) and subtract the same constant from \( g \) without changing \( u \).

Try the “Waves” mathlet


**Problem 20.2.** Suppose that \( c = 1 \), that the initial position is \( I(x) \), and that the initial velocity is 0. What does the wave look like?

**Solution:** The initial conditions \( u(x, 0) = I(x) \) and \( \frac{\partial u}{\partial t}(x, 0) = 0 \) become

\[
\begin{align*}
  f(x) + g(x) &= I(x) \\
  -f'(x) + g'(x) &= 0.
\end{align*}
\]

The second equation says that \( g(x) = f(x) + C \) for some constant \( C \), and we can adjust \( f \) and \( g \) by constants to assume that \( C = 0 \). Then \( f(x) = I(x)/2 \) and \( g(x) = I(x)/2 \). So the wave

\[
  u(x, t) = I(x - t)/2 + I(x + t)/2
\]

consists of two equal waveforms, one traveling to the right and one traveling to the left. □
20.5. **Wave fronts.** Define the step function

\[ s(x) := \begin{cases} 
1, & \text{if } x < 0 \\
0, & \text{if } x > 0,
\end{cases} \]

and consider the solution \( u(x, t) = s(x - t) \). This is a “cliff-shaped” wave traveling to the right. (You would be right to complain that this function is not differentiable and therefore cannot satisfy the PDE in the usual sense, but you can imagine replacing \( s(x) \) with a smooth approximation, a function with very steep slope. The smooth approximation also makes more sense physically: a physical wave would not actually have a jump discontinuity.)

Another way to plot the behavior is to use a **space-time diagram**, in a plane with axes \( x \) (space) and \( t \) (time). (Usually one draws only the part with \( t \geq 0 \).) Divide the \((x, t)\)-plane into regions according to the value of \( u \). The boundary between the regions is called the **wave front**.

In the example above, \( u(x, t) = 1 \) for points to the left of the line \( x - t = 0 \), and \( u(x, t) = 0 \) for points to the right of the line \( x - t = 0 \). So the wave front is the line \( x - t = 0 \).

A different example:

**Flashcard question:** Suppose that the initial position is \( s(x) \), but the initial velocity is 0. Into how many regions is the \( t \geq 0 \) part of the space-time diagram divided?

**Answer:** 3. According to the previous problem,

\[ u(x, t) = s(x - t)/2 + s(x + t)/2. \]

Consider \( t \geq 0 \).

- If \( x < -t \), then \( u(x, t) = 1/2 + 1/2 = 1 \).
- If \( -t < x < t \), then \( u(x, t) = 1/2 + 0 = 1/2 \).
- If \( x > t \), then \( u(x, t) = 0 + 0 = 0 \).

So the upper half of the plane is divided by a V-shaped wave front (the graph of \(|x|\)) into three regions, with values 1 on the left, 1/2 in the middle, and 0 on the right. \( \square \)
Remark 20.3. We have talked about waves moving in one space dimension, but waves exist in higher dimensions too.

- In one dimension, a disturbance created wave fronts moving to the left and right, and the space-time diagram of the wave front was shaped like a V.
- In two dimensions, the disturbance caused by a pebble dropped in a still pond creates a circular wave front that moves outward in all directions. The space-time diagram of this wave front is shaped like an ice cream cone (without the ice cream).
- In three dimensions, the wave front created by a disturbance at a point is an expanding sphere.

20.6. **Real-life waves.** To be covered in recitation on April 29.

In real life, there is always damping. This introduces a new term into the wave equation:

\[
\frac{\partial^2 u}{\partial t^2} + b \frac{\partial u}{\partial t} = c^2 \frac{\partial^2 u}{\partial x^2}
\]

Separation of variables still works, but in each normal mode, the \(w(t)\) is a damped sinusoid involving a factor \(e^{-bt/2}\) (in the underdamped case).

For a movie of a real-life wave, see

[http://www.acoustics.salford.ac.uk/feschools/waves/quicktime/elastic2512K_Stream.mov](http://www.acoustics.salford.ac.uk/feschools/waves/quicktime/elastic2512K_Stream.mov)

---

April 30

**21. Graphical methods**

The final part of this course concerns nonlinear DEs. The sad fact is that we can hardly ever find formulas for the solutions to nonlinear DEs. Instead we try to understand the qualitative behavior of solutions, or use approximations.
21.1. **Solution curves.** We are going to consider nonlinear ODEs

\[ \dot{y} = f(t, y) \]

where \( f \) is a given function and we are to solve for the unknown function \( y(t) \). A **solution curve** (or **integral curve**) is the graph of one such solution in the \((t, y)\)-plane.

**Problem 21.1.** Draw the solution curves to \( \dot{y} = y^2 \). (This is the special case \( f(t, y) := y^2 \).)

**Solution:** Even though this DE is nonlinear, it can be solved exactly by separation of variables:

\[
\frac{dy}{dt} = y^2 \\
y^{-2} \, dy = dt \\
\int y^{-2} \, dy = \int dt \\
y^{-1} - 1 = t + c \quad \text{(for some constant } c) \\
y = \frac{-1}{t + c} \quad \text{(for some constant } c). 
\]

When \( c = 0 \), the solution curve is a hyperbola \( ty = -1 \). For other values of \( c \), the solution curve is the same hyperbola except shifted \( c \) units to the left.

**Oops:** We divided by \( y^2 \), which is not valid if \( y = 0 \). The constant function \( y = 0 \) is a solution too. So in addition to the hyperbolas above, there is one more solution curve: the \( t \)-axis.

Solution curves are graphs of functions, so they must satisfy the vertical line test (at most one point on each vertical line).

**Problem 21.2.** Consider the solution to \( \dot{y} = y^2 \) satisfying the initial condition \( y(0) = 1 \). Is there a solution \( y(t) \) defined for all real numbers \( t \)?

**Solution:** If this were a **linear** ODE, then the existence and uniqueness theorem would guarantee a YES answer.

But here the answer is NO, as we’ll now explain. Setting \( t = 0 \) in the general solution above and using the initial condition leads to

\[
1 = \frac{-1}{0 + c} \\
c = -1, \\
y = \frac{-1}{t - 1} = \frac{1}{1 - t}. 
\]
The largest open interval containing 0 on which a solution exists is \((-\infty, 1)\). One says that the solution blows up in finite time. □

21.2. **Existence and uniqueness.** For nonlinear ODEs, there is still an existence and uniqueness theorem, but the solutions it provides are not necessarily defined for all \(t\).

**Existence and uniqueness theorem for a nonlinear ODE.** Consider a nonlinear ODE

\[
\dot{y} = f(t, y) \quad \text{with initial condition } y(t_0) = y_0
\]

Assume that \(f\) and \(\frac{\partial f}{\partial y}\) are continuous on the entire \((t, y)\)-plane. Then

(a) There exists a solution \(y(t)\) defined on some open interval containing \(t_0\). The largest such open interval is called the **domain of validity** of the solution; call it \(I\).

(b) The solution on \(I\) is unique.

(c) If \(I = (a, b)\) and \(b\) is finite, then as \(t \to b^-\), the function \(y(t)\) becomes unbounded. (A similar statement holds as \(t\) approaches the left endpoint of \(I\).)

**Remark 21.3.** If there are points in the \((t, y)\)-plane where \(f\) or \(\frac{\partial f}{\partial y}\) fails to be continuous, changes are needed. Let \(U\) be the largest open region in the \((t, y)\)-plane on which \(f\) and \(\frac{\partial f}{\partial y}\) are continuous.

(a) If \((t_0, y_0) \in U\), then a solution exists on some open interval containing \(t_0\). There is a largest such interval \(I\) such that the solution curve stays in \(U\).

(b) On that \(I\), the solution is unique.

(c) If \(I = (a, b)\) and \(b\) is finite, then as \(t \to b^-\), either \(y(t)\) becomes unbounded or else \((t, y(t))\) reaches points arbitrarily close to the boundary of \(U\).

What does the theorem mean graphically?

(a) Through each point \((t_0, y_0)\) there is exactly one solution curve. If you ever draw two solution curves that cross or even touch at a point, you are in big trouble! (Exception: They might meet at a point where \(f\) or \(\frac{\partial f}{\partial y}\) fails to be continuous, because the theorem does not apply there.)

(b) The solution curve keeps going (both to left and right) unless it becomes unbounded or approaches a point outside \(U\).

To see these principles in action, try the “Solution Targets” mathlet

[http://mathlets.org/mathlets/solution-targets/](http://mathlets.org/mathlets/solution-targets/)

Here is an example where the hypotheses of the theorem fail.

**Problem 21.4.** Draw the solution curves for \(\dot{y} = \frac{2y}{t}\).
Solution: Here \( f(t,y) = \frac{2y}{t} \), which is undefined when \( t = 0 \), so things might go wrong along the vertical line \( t = 0 \), and in fact they do go wrong.

Solve the ODE by separation of variables:

\[
\frac{dy}{dt} = \frac{2y}{t} \\
\frac{dy}{y} = \frac{2 \, dt}{t} \quad \text{(assuming that } y \text{ is not 0)}
\]

\[
\int \frac{dy}{y} = \int \frac{2 \, dt}{t} \\
\ln |y| = 2 \ln |t| + C \quad \text{(for some constant } C) \\
y = \pm e^{2 \ln |t| + C} \\
y = \pm |t|^2 e^C \\
y = c t^2,
\]

where \( c := \pm e^C \), which can be any nonzero real number. To bring back the solution \( y = 0 \), we allow \( c = 0 \) too. The solution curves are parabolas.

Weird behavior happens along \( t = 0 \) (where the theorem does not apply):

- Through \((0,0)\), there are infinitely many solution curves.
- Through \((0,1)\), there is no solution curve. (Same for \((0,b)\) for any nonzero \(b\).)

But the rest of the plane is covered with good solution curves, one through each point, none touching or crossing the others.

21.3. **Slope field.** We are now going to introduce concepts to help with drawing solution curves to an ODE \( \dot{y} = f(t,y) \). The **slope field** is a diagram in which at each point \((t,y)\), you draw a short segment whose slope is the value \( f(t,y) \).

**Problem 21.5.** Sketch the slope field for \( \dot{y} = y^2 - t \).

**Solution:** Let \( f(t,y) := y^2 - t \). Then

\[
f(1, 2) = 3, \text{ so at } (1, 2) \text{ draw a short segment of slope } 3; \\
f(0, 0) = 0, \text{ so at } (0, 0) \text{ draw a short segment of slope } 0; \\
f(1, 0) = -1, \text{ so at } (1, 0) \text{ draw a short segment of slope } -1; \\
f(0, 1) = 1, \text{ so at } (0, 1) \text{ draw a short segment of slope } 1;
\]

... 

The diagram of all these short segments is the slope field. \( \square \)

A computer can do the job more quickly: try the “Isoclines” mathlet
Warning: The slope field is not the same as the graph of $f$: in drawing the graph of $f$, the value of $f$ is used as a height, but in drawing a slope field, the value of $f$ is used as the slope of a little segment.

Why draw a slope field? The ODE is telling us that the slope of the solution curve at each point is the value of $f(t,y)$, so the short segment there is, to first approximation, a little piece of the solution curve. To get an entire solution curve, follow the segments!

21.4. Isoclines. Even with the computer display, it’s hard to tell what is going on. To understand better, we introduce a new concept: If $m$ is a number, the $m$-isocline is the set of points in the $(t,y)$-plane such that the solution curve through that point has slope $m$. (Isocline means “same incline”, or “same slope”.)

Question 21.6. What is the equation for the $m$-isocline?

Solution: The ODE says that the slope of the solution curve through a point $(t,y)$ is $f(t,y)$, so the equation of the $m$-isocline is $f(t,y) = m$. □

Finding the isoclines will help organize the slope field. The 0-isocline is especially helpful.

Problem 21.7. For $\dot{y} = y^2 - t$, what is the 0-isocline?

Solution: Here $f(t,y) := y^2 - t$, so the 0-isocline is the curve $y^2 - t = 0$, which is a parabola concave to the right. At every point of this parabola, the slope of the solution curve is 0. □
Problem 21.8. For $\dot{y} = y^2 - t$, where are all the points at which the slope of the solution curve is positive?

Solution: This will be the region in which $f(t, y) > 0$. The 0-isocline $f(t, y) = 0$ divides the plane into regions, and $f(t, y)$ has constant sign on each region. To test the sign, just check one point in each region. For $f(t, y) := y^2 - t$, we have $f(t, y) > 0$ in the region to the left of the parabola (since $f(0, 1) > 0$), and $f(t, y) < 0$ in the region to the right of the parabola (since $f(1, 0) < 0$). On the left region, solution curves slope upward; on the right region, solution curves slope downward. The answer is: in the region to the left of the parabola. \( \square \)
The solution curve through \((0, 0)\) increases for \(t < 0\) and decreases for \(t > 0\), so it reaches its maximum at \((0, 0)\). How did we know that the solution for \(t > 0\) does not cross the lower part of the parabola, \(y = -\sqrt{t}\), back into the upward sloping region? Answer: If it crossed somewhere, its slope would have to be negative there, but the DE says that the slope is 0 everywhere along \(y = -\sqrt{t}\). Thus \(y = -\sqrt{t}\) acts as a fence that solution curves already inside the parabola cannot cross.

21.5. Example: The logistic equation. The simplest model for population \(x(t)\) is the ODE \(\dot{x} = ax\) for a positive growth constant \(a\): the rate of population growth is proportional to the current population. But realistically, if \(x(t)\) gets too large, then because of competition for food and space, the population will grow less quickly. One way to model this is with the logistic equation

\[
\dot{x} = ax - bx^2,
\]

where \(a\) and \(b\) are positive constants. This is a nonlinear ODE.

Let’s consider the simplest case, in which \(a = 1\) and \(b = 1\):

Problem 21.9. Draw the solution curves for \(\dot{x} = x - x^2\) in the \((t, x)\)-plane.

Solution: The first step is always to find the 0-isocline. Here \(f(t, x) := x - x^2\), so the 0-isocline is \(x - x^2 = 0\), which consists of the horizontal lines \(x = 0\) and \(x = 1\). Each of these two lines has slope 0, matching the slope specified for the solution curve at each point of the line, so each line itself is a solution curve! (Warning: This is not typical. An isocline is not usually a solution curve.)

The 0-isocline divides the \((t, x)\)-plane into three regions: in the horizontal strip \(0 < x < 1\), we have \(f(t, x) = x - x^2 = x(1 - x) > 0\), so solutions slope upward. In the regions below and above, solutions slope downward.

The diagram below shows the slope field (gray segments), the 0-isocline (yellow line), and the solution curve with initial condition \(x(0) = 1/2\) (blue curve).
22. Autonomous equations

An autonomous equation is a differential equation that is time-invariant: $\dot{x} = f(x)$ instead of $\dot{x} = f(x, t)$. 
(Why is this called autonomous? In ordinary English, a machine or robot is called autonomous if it operates without human input. A differential equation is called autonomous if the coefficients of the problem are not changed over time, such as might happen if a human adjusted a dial on a machine.)

22.1. **Properties.** For an autonomous equation,

- If $x(t)$ is a solution, then so is $x(t - a)$ for any constant $a$.
  
  (Proof: If $x'(t) = f(x(t))$ holds for all $t$, then it holds also with $t$ replaced by $t - a$, so $x'(t - a) = f(x(t - a))$, and the left hand side is the same as the derivative of $x(t - a)$, so this says that $x(t - a)$ is a solution.)

- Each isocline (in the $(t, x)$-plane) consists of horizontal lines.
- For the 0-isocline, these horizontal lines are also solution curves, corresponding to constant solutions.

22.2. **Phase line.**

**Problem 22.1.** Describe the solutions to $\dot{x} = 3x - x^2$.

(This is a special case of the logistic equation $\dot{x} = ax - bx^2$.)

**Solution:** Let $f(x) := 3x - x^2$. First find the 0-isocline by solving $3x - x^2 = 0$. This leads to $x(3 - x) = 0$, so $x = 0$ or $x = 3$. These are horizontal lines. Moreover, they are solution curves corresponding to the constant functions $x(t) = 0$ and $x(t) = 3$.

As in last lecture, the 0-isocline divides the plane into “up” regions and “down” regions. These are the region $x < 0$, the region $0 < x < 3$, and the region $x > 3$. To find out which are up and which are down, test one point in each:

- Since $f(-1) < 0$, the region $x < 0$ is a down region.
- Since $f(1) > 0$, the region $0 < x < 3$ is an up region.
- Since $f(4) < 0$, the region $x > 3$ is a down region.

The phase line is a plot of the $x$-axis that summarizes this information:

$$
-\infty \quad \longleftrightarrow \quad 0 \quad \overset{\text{unstable}}{\rightarrow} \quad 3 \quad \underset{\text{stable}}{\leftarrow} \quad +\infty
$$

(The labels unstable and stable will be explained later. Sometimes the phase line is drawn vertically instead, with $+\infty$ at the top.)

What happens to solutions as time passes?

- If $x(0) = 0$, then the solution will be $x(t) = 0$ for all $t$. (We said this already.)
- If $x(0) = 3$, the solution will be $x(t) = 3$ for all $t$.
- Suppose that the initial condition is that $x(0)$ is a number strictly between 0 and 3. Then $x(t)$ will increase. But it will never reach 3, because the solution curve cannot cross or touch the solution curve at height 3. Could it be that $x(t)$ tends to a limit less than 3? No, because then $\dot{x}(t) = 3x - x^2$ would tend to a positive limit, but $\dot{x}(t)$
must tend to 0 as the solution curve levels off. Conclusion: \( x(t) \) increases, tending to 3 as \( t \to +\infty \) (but never actually reaching 3).

- Similarly, if \( x(0) > 3 \), then \( x(t) \) decreases, tending to 3 without actually reaching 3.
- Finally, if \( x(0) < 0 \), then \( x(t) \) decreases, and \( x(t) \to -\infty \) as \( t \) grows. (With more work, one could show that it tends to \(-\infty\) in finite time.)

**Flashcard question:** If \( x(0) \) is between 0 and 3, what is \( \lim_{t \to -\infty} x(t) \)?

*Answer:* To run time backwards, reverse the arrows in the phase line. As \( t \to -\infty \), we have \( x(t) \to 0 \).

**Warning:** Using a phase line makes sense only if the DE is autonomous!

Try the “Phase Lines” mathlet


22.3. **Stability.** In general, for \( \dot{x} = f(x) \), the real \( x \)-values such that \( f(x) = 0 \) are called **critical points**. **Warning:** Only real numbers can qualify as critical points.

A critical point is called

- **stable** if solutions starting near it move towards it,
- **unstable** if solutions starting near it move away from it,
- **semistable** if the behavior depends on which side of the critical point the solution starts.

In the case of the differential equation \( \dot{x} = 3x - x^2 \) studied above, the critical points are 0 and 3. The phase line shows that 0 is unstable, and 3 is stable.

**Remark** 22.2. An unstable critical point is also called a **separatrix** because it separates solutions having very different fates.

**Example** 22.3. For \( \dot{x} = 3x - x^2 \), a solution starting just **below** 0 tends to \(-\infty\), while a solution starting just **above** 0 tends to 3: very different fates! □

To summarize:

**Steps for understanding solutions to \( \dot{x} = f(x) \) qualitatively:**

1. Solve \( f(x) = 0 \) to find the critical points; these divide the \( x \)-axis into open intervals.
2. Evaluate \( f(x) \) at one point in each interval to find out whether solutions starting there are increasing or decreasing; use this to draw the phase line with \(-\infty\), the critical points, \(+\infty\), and arrows between them.
3. • Solutions starting at a critical point are constant.
• Solutions starting elsewhere tend, as $t$ increases, to the limit that the arrow points to. (To run time backwards, to see the behavior of the solution as $t$ decreases, reverse the arrows.)

Solutions exist for all $t \in \mathbb{R}$, except that if the limit in either direction is $\pm \infty$, then that limit might be reached in finite time.

22.4. Harvesting models and bifurcation diagrams.

Problem 22.4. Frogs grow in a pond according to the logistic equation with growth constant 3, and the population reaches an equilibrium of 3000 frogs, but then the frogs are harvested at a constant rate. Model the population of frogs.

Variables and functions:

$t$ : time (months)

$x$ : size of population (kilofrogs)

$h$ : harvest rate (kilofrogs/month)

Equation: Without harvesting,

$$\dot{x} = 3x - bx^2$$

for some constant $b > 0$. Since the population settles at $x = 3$ (three thousand frogs), $\dot{x} = 3x - bx^2 = 0$ at $x = 3$; thus $b = 1$.

With harvesting, $x(0) = 3$ and

$$\dot{x} = 3x - x^2 - h.$$  \(\square\)

This is an infinite family of autonomous equations, one for each value of $h$, and each has its own phase line. If in the $(h,x)$-plane, we draw each phase line vertically in the vertical line corresponding to a given value of $h$, and plot the critical points for each $h$, then we get a diagram called a bifurcation diagram. In this diagram, color the critical points according to whether they are stable, unstable, or semistable.

Example 22.5. If $h = 2$, then $\dot{x} = 3x - x^2 - 2$. Since $3x - x^2 - 2 = -(x - 2)(x - 1)$, the critical points are 1 and 2, and the phase line is

$$-\infty \quad \leftarrow \quad \text{unstable} \quad \rightarrow \quad 2 \quad \text{stable} \quad \leftarrow \quad +\infty. \quad \square$$

For each other value of $h$, the critical points are the real roots of $3x - x^2 - h$. We could use the quadratic formula to find these roots

$$r_1(h) = \frac{3 - \sqrt{9 - 4h}}{2}, \quad r_2(h) = \frac{3 + \sqrt{9 - 4h}}{2}$$

(assuming that $9 - 4h \geq 0$), and then graph both functions to get the bifurcation diagram.
But we don’t need to do this! The equation \(3x - x^2 - h = 0\) is the same as \(h = 3x - x^2\). The graph of this in the \((x, h)\)-plane is a downward parabola; to get the bifurcation diagram in the \((h, x)\)-plane, interchange the axes by reflecting in the line \(h = x\).

Checking one point inside the parabola (like \((h, x) = (0, 1)\)) shows that \(3x - x^2 - h\) is positive there, and similarly \(3x - x^2 - h\) is negative outside the parabola. Thus the upper branch \(x = r_2(h)\) consists of stable critical points, and the lower branch \(x = r_1(h)\) consists of unstable critical points, at least when \(9 - 4h > 0\).

**Question 22.6.** What happens when \(9 - 4h = 0\), i.e., when \(h = 9/4\)?

**Answer:** Then \(3x - x^2 - 9/4 = -(x - 3/2)^2\), so the phase line is

\[
-\infty \quad \leftarrow \quad \frac{3}{2} \quad \leftarrow \quad +\infty.
\]

Does this mean that a solution \(x(t)\) can go all the way from \(+\infty\) through \(3/2\) to \(-\infty\)? No, because it can’t cross the constant solution \(x = 3/2\). Instead there are three possible behaviors:

- If \(x(0) > 3/2\), then \(x(t) \to 3/2\) as \(t \to +\infty\).
- If \(x(0) = 3/2\), then \(x(t) = 3/2\) for all \(t\).
- If \(x(0) < 3/2\), then \(x(t)\) tends to \(-\infty\) (we interpret this as a *population crash*: the frog population reaches 0 in finite time; the part of the trajectory with \(x < 0\) is not part of the population model).

**Problem 22.7.** What is the maximum sustainable harvest rate?
(Sustainable means that the harvesting does not cause the population to crash to 0, but that instead $\lim_{t \to +\infty} x(t)$ is positive, so that the harvesting can continue indefinitely.)

**Solution:** $h = 9/4$, i.e., 2250 frogs/month. Why?

- For $h < 9/4$, the phase line is
  \[\begin{array}{c}
  -\infty \leftarrow r_1(h) \rightarrow r_2(h) \leftarrow +\infty
  \end{array}\]
  and $x(0) = 3 > r_2(h)$, so $x(t) \to r_2(h)$.

- For $h = 9/4$, the phase line is
  \[\begin{array}{c}
  -\infty \leftarrow 3/2 \leftarrow +\infty
  \end{array}\]
  and $x(0) = 3 > 3/2$, so $x(t) \to 3/2$.

- For $h > 9/4$, the phase line is
  \[\begin{array}{c}
  -\infty \leftarrow +\infty
  \end{array}\]
  so a population crash is inevitable (overharvesting). □

**Remark 22.8.** Harvesting at exactly the maximum rate is a little dangerous, however, because if after a while $x$ becomes very close to $3/2$, and a little kid comes along and takes one more frog out of the pond, the whole frog population will crash!

One of you suggested the following, which seems appropriate:


---

**May 5**

22.5. **Warm-up: 1D linear approximation.** Let’s return to $\dot{x} = 3x - x^2$, and study the solutions with $0 < x < 3$. (This particular ODE could be solved exactly, but for more complicated ODEs one cannot hope to find an exact formula, so we’ll want to illustrate the general method.)

Numerical results of a computer simulation can give us a clear picture of the part of the solutions in the range $0.1 < x < 2.9$, but not enough detail when $x$ is near the critical points 0 and 3 (as happens as $t \to -\infty$ or $t \to +\infty$, respectively). Linear approximations will show us what happens near the critical points.

- Consider $x \approx 0$, which is of interest when $t \to -\infty$, that is, when studying the origins of the population. Then
  \[\dot{x} = 3x - x^2 \approx 3x.\]

  Thus we can expect
  \[x \approx ae^{3t}/171\]
for some constant \( a \). That is, when the population is getting started, solutions to the logistic equation obey approximately exponential growth, until the competition for food or space implicit in the \(-x^2\) term becomes too large to ignore.

- Consider \( x \approx 3 \), which is of interest when \( t \to +\infty \). To measure deviations from 3, define \( X := x - 3 \approx 0 \), so \( x = 3 + X \). The best linear approximation to \( f(x) \) for \( x \approx 3 \) is

\[
 f(x) \approx f(3) + f'(3)(x - 3) \\
 = 0 + (-3)(x - 3) \\
= -3X
\]

so

\[
 \dot{X} = \dot{x} = f(x) \approx -3X.
\]

Thus we can expect

\[
 X \approx be^{-3t}
\]

and

\[
 x = 3 + X \approx 3 + be^{-3t}
\]

for some constant \( b \), as \( t \to +\infty \). (Since we are looking at solutions with \( 0 < x(t) < 3 \), we must have \( b < 0 \).)

The “big picture” combines numerical results for \( 0.1 < x < 2.9 \) with linear approximations near 0 and 3.

23. Autonomous systems

Now we study a system of two autonomous equations in two unknown functions \( x(t) \) and \( y(t) \):

\[
 \dot{x} = f(x, y) \\
 \dot{y} = g(x, y)
\]

for some functions \( f \) and \( g \) that do not depend on \( t \).

Example 23.1. If \( x(t) \) is deer population (in thousands), and \( y(t) \) is wolf population (in hundreds), then the system

\[
 \dot{x} = 3x - x^2 - xy \\
 \dot{y} = y - y^2 + xy
\]

is a reasonable model: each population obeys the logistic equation, except that there is an adjustment depending on \( xy \), which is proportional to the number of deer-wolf encounters. Such encounters are bad for the deer, but good for the wolves!
23.1. **Phase plane.** Solution curves would now exist in 3-dimensional \((t, x, y)\)-space, so they are hard to draw. Instead, forget \(t\), and draw the motion in the \((x, y)\) **phase plane**. At each point \((x, y)\), the system says that the velocity vector there is the value of \(\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}\).

**Problem 23.2.** In the deer-wolf example above, what is the velocity vector at \((x, y) = (3, 2)\)?

**Solution:**
\[
\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 9 - 9 - 6 \\ 2 - 4 + 6 \end{pmatrix} = \begin{pmatrix} -6 \\ 4 \end{pmatrix}.
\]
Draw this velocity vector with its foot at \((3, 2)\). □

The velocity vectors at all points together make a vector field. If you draw them all to scale, you will wreck your picture! Mostly what we care about is the direction, so it is OK to shorten them. Or better yet, don’t draw them at all, and instead just draw arrowheads along the phase plane trajectories in the direction of motion.

There is an existence and uniqueness theorem for systems of nonlinear ODEs similar to that for a single ODEs. For an autonomous system it implies that there is a unique trajectory through each point (in a region in which the partial derivatives of \(f\) and \(g\) are continuous):

**Trajectories never cross or touch!**

(But see the “exception” in Remark 23.4.)

23.2. **Critical points.** A **critical point** for an autonomous system is a point in the \((x, y)\)-plane where the velocity vector is \(0\). To find all the critical points, solve
\[
\begin{align*}
 f(x, y) &= 0 \\
 g(x, y) &= 0.
\end{align*}
\]

**Problem 23.3.** Find the critical points for the deer-wolf system.

**Solution:** We need to solve
\[
\begin{align*}
 3x - x^2 - xy &= 0 \\
  y - y^2 + xy &= 0.
\end{align*}
\]
Each polynomial factors, so we get
\[
\begin{align*}
  x &= 0 \quad \text{or} \quad 3 - x - y = 0 \\
  y &= 0 \quad \text{or} \quad 1 - y + x = 0.
\end{align*}
\]
Intersecting each of the first two lines with each of the last two lines gives the four points
\[(0, 0), \quad (0, 1), \quad (3, 0), \quad (1, 2). \quad \square\]
Critical points are also called **stationary points**, because each such point corresponds to a solution in which \( x(t) \) and \( y(t) \) are constant.

**Remark 23.4.** We said earlier that trajectories never cross. While it is true that no two trajectories can have a point in common, it *is* possible for two trajectories to have the same limit as \( t \to +\infty \) or \( t \to -\infty \), so they can *appear* to come together. For a trajectory to have a finite limiting position, the velocity must tend to 0, so the limiting position must be a critical point.

![Critical point](image)

**Conclusion:** It is only at a critical point that trajectories can appear to come together.

### 23.3. Linear approximation in 2D.

If you remember nothing else from 18.01, remember this:

If a problem you are trying to solve is too difficult because it involves a nonlinear function \( f(x) \), use the best linear approximation near the most relevant \( x \)-value \( a \): that approximation is

\[
 f(a) + f'(a) (x - a)
\]

since this linear polynomial has the same value and same derivative at \( a \) as \( f(x) \).

If you remember nothing else from 18.02, remember this:

If a problem you are trying to solve is too difficult because it involves a nonlinear function \( f(x, y) \), use the best linear approximation near the most relevant point \((a, b)\): that approximation is

\[
 f(a, b) + \frac{\partial f}{\partial x}(a, b) (x - a) + \frac{\partial f}{\partial y}(a, b) (y - b)
\]

since this linear polynomial has the same value and same partial derivatives at \((a, b)\) as \( f(x) \).

(We used green for numbers here.)

### 23.3.1. Warm-up: linear approximation at \((0, 0)\).

To understand the behavior of the deer-wolf system near \((0, 0)\), use

\[
 \dot{x} = 3x - x^2 - xy \approx 3x \\
 \dot{y} = y - y^2 + xy \approx y.
\]
In matrix form,
\[
\begin{pmatrix}
\dot{x} \\
\dot{y}
\end{pmatrix}
\approx
\begin{pmatrix}
3 & 0 \\
0 & 1
\end{pmatrix}
\begin{pmatrix}
x \\
y
\end{pmatrix}.
\]
The eigenvalues are 3 and 1, so this describes a repelling node at (0, 0).

23.3.2. **Linear approximation via change of coordinates (method 1).** To understand the deer-wolf system near the critical point (1, 2), reduce to the previously solved case of (0, 0) by making the change of variable
\[
x = 1 + X
\]
\[
y = 2 + Y
\]
so that \((x, y) = (1, 2)\) is \((X, Y) = (0, 0)\) in the new coordinate system. Then
\[
\dot{X} = \dot{x} = 3(1 + X) - (1 + X)^2 - (1 + X)(2 + Y) = -X - Y - X^2 - XY \approx -X - Y
\]
\[
\dot{Y} = \dot{y} = (2 + Y) - (2 + Y)^2 + (1 + X)(2 + Y) = 2X - 2Y - Y^2 + XY \approx 2X - 2Y
\]
when \((X, Y)\) is close to \((0, 0)\). In matrix form,
\[
\begin{pmatrix}
\dot{X} \\
\dot{Y}
\end{pmatrix}
\approx
\begin{pmatrix}
-1 & -1 \\
2 & -2
\end{pmatrix}
\begin{pmatrix}
X \\
Y
\end{pmatrix}.
\]

23.3.3. **Linear approximation via Jacobian matrix (method 2).**

**Definition 23.5.** The **Jacobian matrix** of the vector-valued function \(\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}\) is the matrix-valued function
\[
J(x, y) := \begin{pmatrix}
\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\
\frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}
\end{pmatrix}.
\]
The **Jacobian determinant** is the determinant of the Jacobian matrix. In 18.02, you learned that the absolute value of the Jacobian determinant is the area scaling factor when doing a change of variable in a double integral.

The Jacobian matrix is also called the **derivative** of the multivariable function \(\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}\). The function has 2-variable input \(\begin{pmatrix} x \\ y \end{pmatrix}\) and 2-variable output \(\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix}\); this leads to each value of the Jacobian matrix being a \(2 \times 2\) matrix.

The best linear approximations to \(f\) and \(g\) at \((a, b)\) are
\[
f(x, y) \approx f(a, b) + \frac{\partial f}{\partial x}(a, b) (x - a) + \frac{\partial f}{\partial y}(a, b) (y - b)
\]
\[
g(x, y) \approx g(a, b) + \frac{\partial g}{\partial x}(a, b) (x - a) + \frac{\partial g}{\partial y}(a, b) (y - b).
\]

175
These can be combined into one equation:

\[
\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} \approx \begin{pmatrix} f(a, b) \\ g(a, b) \end{pmatrix} + \begin{pmatrix} J(a, b) \end{pmatrix} \begin{pmatrix} x - a \\ y - b \end{pmatrix},
\]

value at \((a, b)\)
derivative at \((a, b)\)

If \((a, b)\) is a critical point for the system, then \(f(a, b) = 0\) and \(g(a, b) = 0\), so this linear approximation simplifies to

\[
\begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} \approx \begin{pmatrix} J(a, b) \end{pmatrix} \begin{pmatrix} x - a \\ y - b \end{pmatrix}.
\]

If we make the change of variable \(X := x - a\) and \(Y := y - b\), this becomes

\[
\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} f(x, y) \\ g(x, y) \end{pmatrix} \approx \begin{pmatrix} J(a, b) \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.
\]

**Conclusion:** At a critical point \((a, b)\), if \(X := x - a\) and \(Y := y - b\), then

\[
\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} \approx \begin{pmatrix} J(a, b) \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.
\]

**Problem 23.6.** Find the behavior of the deer-wolf system near the critical point \((1, 2)\).

**Solution:** We have

\[
J(x, y) := \begin{pmatrix} \partial f/\partial x & \partial f/\partial y \\ \partial g/\partial x & \partial g/\partial y \end{pmatrix} = \begin{pmatrix} 3 - 2x - y & -x \\ y & 1 - 2y + x \end{pmatrix}.
\]

Plug in \(x = 1\) and \(y = 2\) to get

\[
J(1, 2) = \begin{pmatrix} -1 & -1 \\ 2 & -2 \end{pmatrix}.
\]

Thus, if we measure deviations from the critical point by defining \(X := x - 1\) and \(Y := y - 2\), we have

\[
\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} \approx \begin{pmatrix} -1 & -1 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}.
\]

The matrix has trace \(-3\) and determinant \(4\), so the characteristic polynomial is \(\lambda^2 + 3\lambda + 4\), and the eigenvalues are \(\frac{-3 \pm \sqrt{-7}}{2}\). These are complex numbers with negative real part, so this describes an attracting spiral. \(\square\)
23.4. **Structural stability.** Recall: We were studying an autonomous system
\[
\begin{align*}
\dot{x} &= f(x, y) \\
\dot{y} &= g(x, y).
\end{align*}
\]

To understand the behavior near a critical point \((a, b)\), we made a change of variable
\[
\begin{align*}
x &= a + X \\
y &= b + Y
\end{align*}
\]
to move the critical point to \((0, 0)\), and we replaced \(f(x, y)\) and \(g(x, y)\) by their best linear approximations to get the linear system
\[
\begin{pmatrix}
\dot{X} \\
\dot{Y}
\end{pmatrix} \approx A \begin{pmatrix} X \\ Y \end{pmatrix}.
\]

**Question 23.7.** When is it OK to say that the original system behaves like the linear system?

**Approximation principle.** If an autonomous system is approximated near a critical point \((a, b)\) by
\[
\begin{pmatrix}
\dot{X} \\
\dot{Y}
\end{pmatrix} \approx A \begin{pmatrix} X \\ Y \end{pmatrix},
\]
and if the system \(\dot{x} = Ax\) is structurally stable (saddle, repelling/attracting node, repelling/attracting spiral), then the phase portrait for the original system looks near \((a, b)\) like the phase portrait for
\[
\begin{pmatrix}
\dot{X} \\
\dot{Y}
\end{pmatrix} = A \begin{pmatrix} X \\ Y \end{pmatrix}
\]
near \((0, 0)\). (We aren’t going to prove this, or even make precise what “looks like” means.) The phase portrait may become more and more warped as one moves away from the critical point.

These cases, as opposed to the borderline cases in which \(A\) lies on the boundary between regions in the trace-determinant plane, are called **structurally stable**.

**Warning:** Stability and structural stability are different concepts:
- **Stable** means that all nearby solutions tend to the critical point.
- **Structurally stable** means that the phase portrait type is robust, unaffected by small changes in the matrix entries.

**Example 23.8.** The phase portrait for
\[
\begin{pmatrix}
\dot{X} \\
\dot{Y}
\end{pmatrix} \approx \begin{pmatrix} 0 & 4 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix}
\]

177
is a center, so trajectories are periodic. But if an autonomous system has this as its linear approximation at a critical point, it is not guaranteed that trajectories are periodic, because the slight warping might make the trajectories no longer come back to exactly the initial position after going around once.

23.5. **Big picture.**

Steps for drawing the phase portrait for an autonomous system \( \dot{x} = f(x, y) \), \( \dot{y} = g(x, y) \):

1. Solve the system

   \[
   f(x, y) = 0 \\
   g(x, y) = 0
   \]

   to find all the critical points in the \((x, y)\)-phase plane. There is a stationary trajectory at each critical point.

2. Calculate the Jacobian matrix

   \[
   J(x, y) := \begin{pmatrix}
   \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\
   \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}
   \end{pmatrix}
   \]

   This will be a \(2 \times 2\) matrix of functions of \(x\) and \(y\).

3. At each critical point \((a, b)\),

   (a) Compute the numerical \(2 \times 2\) matrix \(A := J(a, b)\), by evaluating \(J(x, y)\) at \((a, b)\).

   (b) Determine whether the critical point is stable (attracting) or not:

   \[
   \text{stable} \iff \text{tr} \, A < 0 \text{ and } \det A > 0.
   \]

   Or, for a more detailed picture, find the eigenvalues of \(A\) to classify the phase portrait for the “linear approximation system” \(\begin{pmatrix} \dot{X} \\ \dot{Y} \end{pmatrix} \approx A \begin{pmatrix} X \\ Y \end{pmatrix}\). For further details:

   • If the eigenvalues are real, find the eigenlines. If, moreover, the eigenvalues have the same sign, also determine the slow eigenline since trajectories in the \((X, Y)\)-plane will be tangent to that line.

   • If the eigenvalues are complex (and not real), compute a velocity vector to determine whether the rotation is clockwise or counterclockwise.

   (c) Mark the critical point \((a, b)\) in the \((x, y)\)-plane, and draw a miniature copy of the linear approximation’s phase portrait shifted so that it is centered at \((a, b)\); this is justified in the structurally stable cases (saddle, repelling node, attracting node, or spiral). Indicate with arrowheads the direction of motion on the trajectories near the critical point.

4. (Optional) Find the velocity vector at a few other points, or use a computer.
5. (Optional) Solve \( f(x, y) = 0 \) to find all the points where the velocity vector is vertical or \( 0 \). Similarly, one could solve \( g(x, y) = 0 \) to find all the points where the velocity vector is horizontal or \( 0 \).

6. Connect trajectories emanating from or approaching critical points, keeping in mind that trajectories never cross or touch.

**Problem 23.9.** Sketch the phase portrait for the deer-wolf system

\[
\begin{align*}
\dot{x} &= 3x - x^2 - xy \\
\dot{y} &= y - y^2 + xy.
\end{align*}
\]

**Solution:** We already found the critical points

\[(0, 0), \quad (0, 1), \quad (3, 0), \quad (1, 2).\]

We already found the Jacobian matrix

\[
J(x, y) = \begin{pmatrix}
3 - 2x - y & -x \\
y & 1 - 2y + x
\end{pmatrix}.
\]

*Critical point \((1, 2)\):* We already calculated

\[
J(1, 2) = \begin{pmatrix}
-1 & -1 \\
2 & -2
\end{pmatrix}.
\]

This has trace \(-3\) and determinant \(4\), so this critical point is stable.

The characteristic polynomial is \(\lambda^2 + 3\lambda + 4\), and the eigenvalues are \(\frac{-3 \pm \sqrt{-7}}{2}\). These are complex numbers with negative real part, so this describes an attracting spiral. The velocity vector at \(\begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\) is \(\begin{pmatrix} -1 \\ 2 \end{pmatrix}\), so the spiral is counterclockwise. This is a structurally stable case, so the phase portrait for the original system near \((1, 2)\) will be a counterclockwise attracting spiral too.

*Critical point \((0, 0)\):

\[
J(0, 0) = \begin{pmatrix}
3 & 0 \\
0 & 1
\end{pmatrix}.
\]

This has trace \(4\), so this critical point is unstable. Since the matrix is diagonal, its eigenvalues are the diagonal entries \(3\) and \(1\), and the vectors \(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\) and \(\begin{pmatrix} 0 \\ 1 \end{pmatrix}\) are corresponding eigenvectors.

The eigenvalues are distinct positive real numbers, so this describes a repelling node. The slow eigenline is the \(Y\)-axis, so most trajectories emanating from \((0, 0)\) are tangent to the \(Y\)-axis. This is a structurally stable case, so the phase portrait for the original system near
(0, 0) too will be a repelling node, and most trajectories emanating from (0, 0) are tangent to the y-axis.

**Critical point** (0, 1):

\[ J(0, 1) = \begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix}. \]

This has trace 1, so this critical point is unstable. Since the matrix is lower triangular, its eigenvalues are the diagonal entries 2 and −1. The eigenvalues are real numbers of opposite sign, so this describes a saddle. The eigenlines for the eigenvalues 2 and −1 are \( Y = \frac{1}{2}X \) and \( X = 0 \). This is a structurally stable case, so the phase portrait for the original system near (0, 1) is a saddle too.

**Critical point** (3, 0):

\[ J(3, 0) = \begin{pmatrix} -3 & -3 \\ 0 & 4 \end{pmatrix}. \]

This has trace 1, so this critical point is unstable. Since the matrix is upper triangular, its eigenvalues are the diagonal entries −3 and 4. The eigenvalues are real numbers of opposite sign, so this describes a saddle. The eigenlines for the eigenvalues −3 and 4 are \( Y = 0 \) and \( Y = -\frac{3}{4}X \). This is a structurally stable case, so the phase portrait for the original system near (3, 0) is a saddle too.

**At which points are the trajectories vertical?**

These are the points at which the x-coordinate of the velocity vector is 0, i.e., the points where

\[ 3x - x^2 - xy = 0. \]

Factoring shows that these are the points on the lines \( x = 0 \) and \( 3 - x - y = 0 \). So in the phase portrait we draw little vertical segments at points on these lines. In particular, there will be trajectories along \( x = 0 \), and we can plot them using the 1-dimensional phase line methods, by sampling the velocity vector at one point in each interval created by the critical points. The line \( 3 - x - y = 0 \) does not contain trajectories, however, since that line has slope −1, while trajectories are vertical as they pass through these points.

**At which points are the trajectories horizontal?**

These are points at which

\[ y - y^2 + xy = 0. \]

These are the lines \( y = 0 \) and \( 1 - y + x = 0 \), so draw little horizontal segments at points on these lines. Again we can study trajectories along \( y = 0 \) using 1-dimensional phase line methods.

**Big picture:**
Try the “Vector Fields” mathlet

http://mathlets.org/mathlets/vector-fields/

23.6. Changing the parameters of the system. The big picture suggests that all trajectories in the first quadrant tend to (1, 2) as \( t \to +\infty \). In other words, as long as there were some deer and some wolves to begin with, eventually the populations stabilize at about 1000 deer and 200 wolves.

Problem 23.10. Suppose that we start feeding the deer so that the system becomes

\[
\begin{align*}
\dot{x} &= ax - x^2 - xy \\
\dot{y} &= y - y^2 + xy
\end{align*}
\]

for some number \( a \) slightly larger than 3. What happens?

Solution: The critical points will move slightly, but they won’t change their stability. The populations will end up at the stable critical point, which is the one near (1, 2). To find it,
solve

\[
\begin{align*}
0 &= ax - x^2 - xy \\
0 &= y - y^2 + xy.
\end{align*}
\]

Since we’re looking for a solution with \(x > 0\) and \(y > 0\), it is OK to divide the equations by \(x\) and \(y\), respectively:

\[
\begin{align*}
0 &= a - x - y \\
0 &= 1 - y + x.
\end{align*}
\]

Solving gives

\[
x = \frac{a - 1}{2}, \quad y = \frac{a + 1}{2}.
\]

For \(a = 3\), this is \(x = 1\) and \(y = 2\). For \(a > 3\), the deer population increases, but we also see an increase in the wolf population. By feeding the deer we have provided more food for the wolves as well.

23.7. **Fences.** In the original deer-wolf system, how can you *prove* that all trajectories tend to \((1, 2)\)?

**Steps to prove** that all trajectories approach the stable critical point:

1. Find a window into which all trajectories must enter and never leave.
2. Do a numerical simulation within the window.

Let’s do step 1 for the deer-wolf system. A trajectory could escape in four ways: up, down, left, and right. We need to rule out all four.

**Bottom:** A trajectory that starts in the first quadrant cannot cross the nonnegative part of the \(x\)-axis, because the trajectories along the \(x\)-axis act as fences. A trajectory cannot even tend to a point on the \(x\)-axis, because such a point would be a critical point, and the phase portrait types at \((0, 0)\) and \((3, 0)\) make such an approach impossible.

**Left:** By the same argument, the nonnegative part of the \(y\)-axis is a fence that cannot be approached.

**Right:** We have

\[
\dot{x} = 3x - x^2 - xy \leq 3x - x^2 < 0
\]

whenever \(x > 3\) (if \(3x - x^2\) is negative, then \(3x - x^2 - xy\) is even more negative since it has something subtracted). So all the vertical lines \(x = c\) for \(c > 3\) are fences that prevent trajectories from moving to the right across them. All trajectories move leftward if \(x > 3\), and they can settle down only in the range \(0 \leq x \leq 3\).
Top: Assuming $x \leq 3$, we have

$$\dot{y} = y - y^2 + xy \leq y - y^2 + 3y = 4y - y^2 < 0$$

whenever $y > 4$. Thus for $c > 4$, the horizontal segments $y = c$, $0 \leq x \leq 3$ are fences preventing trajectories from moving up through them.

Conclusion: All trajectories starting with $x > 0$, $y > 0$ (the only ones we care about) eventually enter the window $0 \leq x \leq 3$, $0 \leq y \leq 4$ and stay there. This is small enough that a numerical simulation can now show that all these points tend to $(1, 2)$ (step 2).

The final exam covers everything up to here, in the sense that you are not required to know anything specific below. On the other hand, the topics below serve partially as review of earlier topics that are covered.

May 9

23.8. **Nonlinear centers, limit cycles, etc.** Consider an autonomous system. Suppose that $P$ is a critical point. Suppose that the linear approximation system at $P$ is a center. What is the behavior of the original system near $P$? It’s not necessarily a center. (This is not a structurally stable case.) In fact, there are many possibilities:

- **nonlinear center**, in which the trajectories are periodic (but not necessarily exact ellipses)
- repelling spiral
- attracting spiral
- hybrid situation containing a **limit cycle**: a periodic trajectory with an outward spiral approaching it from within and an inward spiral approaching it from outside!

For an example of a limit cycle (called the van der Pol limit cycle), set $a = 0.1$ in the system

$$\begin{align*}
\dot{x} &= y \\
\dot{y} &= a(1 - x^2)y - x
\end{align*}$$

in the “Vector Fields” mathlet

24. Pendulums

24.1. Modeling.

**Problem 24.1.** Model a pendulum, consisting of a weight attached to a rod hanging from a pivot at the top.

**Variables and functions:** Define

\[ t : \text{ time} \]
\[ \theta : \text{ angle measured counterclockwise from the rest position} \]

Here \( t \) is the independent variable, and \( \theta \) is a function of \( t \).

**Simplifying assumptions:**
- The rod has length 1, so \( \theta \) equals arc length and \( \dot{\theta} \) is velocity.
- The rod has negligible mass.
- The rod does not bend or stretch.
- The weight has mass 1.
- The pivot is such that the motion is in a plane (no Coriolis effect).
- The local gravitational field \( g \) is a constant (the pendulum is not thousands of kilometers tall).
Equation: When the weight is at a certain position, let \( \hat{\theta} \) be the unit vector in the direction that the weight moves as \( \theta \) starts to increase. The \( \hat{\theta} \)-component of the weight’s acceleration is

\[
\ddot{\theta} = -g \sin \theta. \tag{15}
\]

More realistic (with friction, assumed for simplicity to be proportional to \( \dot{\theta} \)):

\[
\ddot{\theta} = \underbrace{-b\dot{\theta}}_{\text{friction}} - g \sin \theta \underbrace{\text{gravity}}_{\text{gravity}}.
\]

The \( \ddot{\theta} \) and \( b\dot{\theta} \) terms are linear, but the \( g \sin \theta \) makes the whole DE nonlinear.

Remark 24.2. If \( \theta \) is very small, then it is reasonable to replace the nonlinear term by its best linear approximation at \( \theta = 0 \), namely \( \sin \theta \approx \theta \), which leads to

\[
\ddot{\theta} + b\dot{\theta} + g\theta = 0,
\]

a damped harmonic oscillator.

24.2. Converting to a first-order system. But to get an accurate understanding even when \( \theta \) is not so small, we need to analyze equation (15) in its full nonlinear glory. It is a second-order nonlinear ODE; we haven’t developed tools for those. So instead convert it to a (still nonlinear) system of first-order ODEs, by introducing a new function \( v := \dot{\theta} \) (velocity):

\[
\dot{\theta} = v \\
\dot{v} = -bv - g \sin \theta.
\]

This is an autonomous system! So we can use all the methods we’ve been developing.

24.3. Critical points. The critical points are given by

\[
v = 0 \\
-bv - g \sin \theta = 0.
\]

Substituting \( v = 0 \) into the second equation leads to \( \sin \theta = 0 \), so \( \theta = \ldots, -2\pi, -\pi, 0, \pi, 2\pi, \ldots \). Thus there are infinitely many critical points:

\[
\ldots, (-2\pi, 0), (-\pi, 0), (0, 0), (\pi, 0), (2\pi, 0), \ldots
\]

But these represent only two distinct physical situations, since adding \( 2\pi \) to \( \theta \) does not change the position of the weight.
24.4. **Phase portrait of the frictionless pendulum; energy levels.** Let’s draw the phase portrait in the \((\theta, v)\)-plane when \(b = 0\) and \(g = 1\). Now the system is

\[
\begin{align*}
\dot{\theta} &= v \\
\dot{v} &= -\sin \theta.
\end{align*}
\]

**Flashcard question:** In the frictionless case, are the critical points \((0, 0)\) and \((\pi, 0)\) stable?

**Answer:** Neither is stable.

- The point \((\pi, 0)\) corresponds to a vertical rod with the weight precariously balanced at the top. If the weight is moved slightly away, the trajectory goes far from \((\pi, 0)\).
- The point \((0, 0)\) corresponds to a vertical rod with the weight at the bottom. If the weight is moved slightly away, the trajectory does not tend to \((0, 0)\) in the limit because the pendulum oscillates forever in the frictionless case. \(\Box\)

To analyze the behavior near each critical point, use a linear approximation. The Jacobian matrix is

\[
J(\theta, v) = \begin{pmatrix} 0 & 1 \\ -\cos \theta & 0 \end{pmatrix}.
\]

---

**Critical point \((\pi, 0)\):**

\[
J(\pi, 0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\]

(This is not a diagonal matrix: wrong diagonal.)

Eigenvalues: \(1, -1\)

Eigenvectors: \(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\)

Type of the linear approximation: Saddle, with outgoing trajectories of slope 1, incoming trajectories of slope \(-1\).

Type of the original nonlinear system: Same.

---

**Critical point \((0, 0)\):**

\[
J(0, 0) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.
\]

Eigenvalues: \(\pm i\).

Type of the linear approximation: Center.

Type of the original nonlinear system: ??? (not a structurally stable case, so we can’t tell yet whether trajectories near \((0, 0)\) are periodic).
To figure out what happens near (0, 0), use conservation of energy! The energy function is

\[ E(\theta, v) := \frac{1}{2}v^2 + 1 - \cos \theta. \]

(Remember that all the constants were set to 1, so potential energy equals height, which we choose to measure relative to the rest position.)

Let’s check conservation of energy:

\[
\dot{E} = v \dot{v} + (\sin \theta) \dot{\theta} \\
= v(-\sin \theta) + (\sin \theta)v \\
= 0.
\]

This means that along each trajectory, \( E \) is constant. In other words, each trajectory is contained in a level curve of \( E \).

**Energy level \( E = 0 \):**

\[
\frac{1}{2}v^2 + (1 - \cos \theta) = 0.
\]

Both terms on the left are nonnegative, so their sum can be 0 only if both are 0, which happens only at \((\theta, v) = (0, 0)\) (and the similar points with some \(2\pi n\) added to \(\theta\)). The energy level \( E = 0 \) consists of the stationary trajectory at \((0, 0)\).

**Energy level \( E = \epsilon \) for small \( \epsilon > 0 \):**

\[
\frac{1}{2}v^2 + (1 - \cos \theta) = \epsilon.
\]

Both kinetic energy and potential energy must be small, so the height is small, so \(\theta\) is small, so \(\cos \theta \approx 1 - \frac{\theta^2}{2}\), so the energy level is very close to

\[
\frac{v^2}{2} + \frac{\theta^2}{2} = \epsilon,
\]

a small circle. The trajectory goes clockwise along it since \(\dot{\theta} > 0\), and decreasing when \(\dot{\theta} < 0\). So trajectories near \((0, 0)\) are periodic ovals (approximately circles); these represent a small oscillation near the bottom. The critical point is a nonlinear center.
Energy level $E = 2$:

\[
\frac{1}{2} v^2 + (1 - \cos \theta) = 2
\]

\[
\frac{1}{2} v^2 = 1 + \cos \theta
\]

\[
= 1 + (2 \cos^2 \frac{\theta}{2} - 1)
\]

\[
= 2 \cos^2 \frac{\theta}{2}
\]

\[
v = \pm 2 \cos \frac{\theta}{2}.
\]

Does this mean that the motion is periodic, going around and around? No. This energy level contains three physical trajectories:

- one in which the weight is stationary at the top
- one in which the weight does one clockwise loop as $t$ goes from $-\infty$ to $\infty$, slowing down as it approaches the top, taking infinitely long to get there (and infinitely long to come from there),
- the same, except counterclockwise.

In the last two cases, the weight can’t actually reach the top, since its phase plane trajectory can’t touch the stationary trajectory.

Energy level $E = 3$:

\[
\frac{1}{2} v^2 + (1 - \cos \theta) = 3
\]

\[
v = \pm \sqrt{4 + 2 \cos \theta}.
\]

The possibility $v = \sqrt{4 + 2 \cos \theta}$ is a periodic function of $\theta$, varying between $\sqrt{2}$ and $\sqrt{6}$. The energy level consists of two trajectories: in each, the weight makes it to the top still having some kinetic energy, so that it keeps going around (either clockwise or counterclockwise).
24.5. Phase portrait of the damped pendulum. Next let’s draw the phase portrait when $b > 0$ (so there is friction) and $g = 1$. The system is

$$\dot{\theta} = v$$

$$\dot{v} = -bv - \sin \theta.$$ 

This time,

$$\dot{E} = v\dot{v} + (\sin \theta)\dot{\theta}$$

$$= v(-bv - \sin \theta) + (\sin \theta)v$$

$$= -bv^2,$$

so energy is lost to friction whenever the weight is moving.

- There are still the stationary trajectories at critical points.
- All other trajectories cross the energy levels, moving to lower energy: the direction field always points “inwards” towards lower energy; the energy levels serve as fences preventing phase plane motion to higher energy; the trajectory tends to a limit that must be a critical point. one of

$$\ldots, (-2\pi, 0), (-\pi, 0), (0, 0), (\pi, 0), (2\pi, 0), \ldots$$
25. Numerical methods

25.1. Euler’s method. Consider a nonlinear ODE \( \dot{y} = f(t, y) \). It specifies a slope field in the \((t, y)\)-plane, and solution curves follow the slope field.

Suppose that we are given a starting point \((t_0, y_0)\) (here \(t_0\) and \(y_0\) are given numbers), and that we are trying to approximate the solution curve through it.

Question 25.1. Where, approximately, will be the point on the solution curve at a time \(h\) seconds later?

Solution: Pretend that the solution curve is a straight line segment between times \(t_0\) and \(t_0 + h\), with slope as specified by the ODE at \((t_0, y_0)\). The ODE says that the slope at \((t_0, y_0)\) is \(f(t_0, y_0)\), so the estimated answer is \((t_1, y_1)\) with

\[
\begin{align*}
t_1 &:= t_0 + h \\
y_1 &:= y_0 + f(t_0, y_0) \cdot h.
\end{align*}
\]

\(\square\)

Question 25.2. Where, approximately, will be the point on the solution curve at time \(t_0 + 3h\)?

Solution: The stupidest answer would be to take 3 steps each using the initial slope \(f(t_0, y_0)\) (or equivalently, one big step of width \(3h\)). The slightly less stupid answer is called Euler’s
method: take 3 steps, but reassess the slope after each step, using the slope field at each successive position:

\[
\begin{align*}
t_1 &:= t_0 + h & y_1 &= y_0 + f(t_0, y_0) h \\
t_2 &:= t_1 + h & y_2 &= y_1 + f(t_1, y_1) h \\
t_3 &:= t_2 + h & y_3 &= y_2 + f(t_2, y_2) h.
\end{align*}
\]

The sequence of line segments from \((t_0, y_0)\) to \((t_1, y_1)\) to \((t_2, y_2)\) to \((t_3, y_3)\) is an approximation to the solution curve. The answer to the question is approximately \((t_3, y_3)\). □

Usually these calculations are done by computer, and there are round-off errors in calculations. But even if there are no round-off errors, Euler’s method usually does not give the exact answer. The problem is that the actual slope of the solution curve changes between \(t_0\) and \(t_0 + h\), so following a segment of slope \(f(t_0, y_0)\) for this entire time interval is not exactly correct.

To improve the approximation, use a smaller step size \(h\), so that the slope is reassessed more frequently. The cost of this, however, is that in order to increase \(t\) by a fixed amount, more steps will be needed.

Under reasonable hypotheses on \(f\), one can prove that as \(h \to 0\), this process converges and produces an exact solution curve in the limit. This is how the existence theorem for ODEs is proved.

Try the “Euler’s Method” mathlet

http://mathlets.org/mathlets/eulers-method/

25.2. Euler’s method for systems. A first-order system of ODEs can be written in vector form \(\dot{x} = f(t, x)\), where \(f\) is a vector-valued function. Euler’s method works the same way.

Starting from \((t_0, x_0)\), define

\[
\begin{align*}
t_1 &:= t_0 + h & x_1 &= x_0 + f(t_0, x_0) h \\
t_2 &:= t_1 + h & x_2 &= x_1 + f(t_1, x_1) h \\
t_3 &:= t_2 + h & x_3 &= x_2 + f(t_2, x_2) h.
\end{align*}
\]

25.3. Tests for reliability.

Question 25.3. How can we decide whether answers obtained numerically can be trusted?

Here are some heuristic tests. (“Heuristic” means that these tests seem to work in practice, but they are not proved to work always.)

- **Self-consistency**: Solution curves should not cross! If numerically computed solution curves appear to cross, a smaller step size is needed. (E.g., try the mathlet “Euler’s Method” with \(\dot{y} = y^2 - x\), step size 1, and starting points \((0, 0)\) and \((0, 1/2)\).)
• **Convergence as** $h \to 0$: The estimate for $y(t)$ at a fixed later time $t$ should converge to the true value as $h \to 0$. If shrinking $h$ causes the estimate to change a lot, then $h$ is probably not small enough yet. (E.g., try the mathlet “Euler’s Method” with $y’ = y^2 - x$ with starting point $(0,0)$ and various step sizes.)

• **Structural stability**: If small changes in the DE’s parameters or initial conditions change the outcome completely, the answer probably should not be trusted. One reason for this could be a separatrix, a curve such that nearby starting points on different sides lead to qualitatively different outcomes; this is not a fault of the numerical method, but is an instability in the answer nevertheless. (E.g., try the mathlet “Euler’s Method” with $y’ = y^2 - x$, starting point $(-1,0)$ or $(-1, -0.1)$, and step size $0.125$ or actual solution.)

25.4. **Change of variable.** Euler’s method generally can’t be trusted to give reasonable values when $(t,y)$ strays very far from the starting point. In particular, the solutions it produces usually deviate from the truth as $t \to \pm \infty$, or in situations in which $y \to \pm \infty$ in finite time. Anything that goes off the screen can’t be trusted.

*Example* 25.4. The solution to $\dot{y} = y^2 - t$ starting at $(-2,0)$ seems to go to $+\infty$ in finite time. But Euler’s method never produces a value of $+\infty$.

To see what is really happening in this example, try the change of variable $u = 1/y$. To rewrite the DE in terms of $u$, substitute $y = 1/u$ and $\dot{y} = -\dot{u}/u^2$:

\[
\frac{-\dot{u}}{u^2} = \frac{1}{u^2} - t \\
\dot{u} = -1 + tu^2.
\]

This is equivalent to the original DE, but now, when $y$ is large, $u$ is small, and Euler’s method can be used to find the time when $u$ crosses 0, which is when $y$ blows up.

25.5. **Runge–Kutta methods.** This section was covered only briefly in lecture.

When computing $\int_a^b f(t) \, dt$ numerically, the most primitive method is to use the left Riemann sum: divide the range of integration into subintervals of width $h$, and estimate the value of $f(t)$ on each subinterval as being the value at the left endpoint. More sophisticated methods are the trapezoid rule and Simpson’s rule, which have smaller errors.

There are analogous improvements to Euler’s method.

<table>
<thead>
<tr>
<th>Integration</th>
<th>Differential equation</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>left Riemann sum</td>
<td>Euler’s method</td>
<td>$O(h)$</td>
</tr>
<tr>
<td>trapezoid rule</td>
<td>second-order Runge–Kutta method (RK2)</td>
<td>$O(h^2)$</td>
</tr>
<tr>
<td>Simpson’s rule</td>
<td>fourth-order Runge–Kutta method (RK4)</td>
<td>$O(h^4)$</td>
</tr>
</tbody>
</table>
The big-$O$ notation $O(h^4)$ means that there is a constant $C$ (depending on everything except for $h$) such that the error is at most $C h^4$, assuming that $h$ is small. The error estimates in the table are valid for reasonable functions.

The Runge–Kutta methods “look ahead” to get a better estimate of what happens to the slope over the course of the interval $[t_0, t_0 + h]$.

Here is how one step of the second-order Runge–Kutta method (RK2) goes

1. Starting from $(t_0, y_0)$, look ahead to see where one step of Euler’s method would land, say $(t_1, y_1)$, but do not go there!
2. Instead sample the slope at the midpoint $(\frac{t_0 + t_1}{2}, \frac{y_0 + y_1}{2})$.
3. Now move along the segment of that slope: the new point is

$$\left(t_0 + h, y_0 + f\left(\frac{t_0 + t_1}{2}, \frac{y_0 + y_1}{2}\right) h\right).$$

Repeat, reassessing the slope after each step. (RK2 is also called midpoint Euler.)

The fourth-order Runge–Kutta method (RK4) is similar, but more elaborate, averaging several slopes. It is probably the most commonly used method for solving DEs numerically. Some people simply call it the Runge–Kutta method. The mathlets use RK4 with a small step size to compute the “actual” solution to a DE.

May 14

26. Review

26.1. Check your answers! On the final exam, there is no excuse for getting an eigenvector wrong, since you will have plenty of time to check it! You can also check solutions to linear systems, or solutions to DEs.

26.2. Resonance involving Fourier series.

Problem 26.1. A voltage source providing a square wave voltage

$$V(t) = \frac{\pi}{4} \text{Sq}(t) = \sin t + \frac{\sin 3t}{3} + \frac{\sin 5t}{5} + \cdots$$

is attached in series to

- an inductor of inductance 1,
- a resistor of unknown resistance $R$, and
- a capacitor of capacitance $1/99^2$. 

193
It is observed that the steady-state charge \( Q(t) \) on the capacitor is very close to \( \cos(99t - \phi) \) for some \( \phi \), that is, a pure sinusoidal wave of amplitude 1 and angular frequency 99. What is the order of magnitude of \( R \)? In other words, find an integer \( m \) such that \( R \approx 10^m \).

**Hint:** The DE for such an RLC-circuit is

\[
L\ddot{Q} + R\dot{Q} + \frac{1}{C}Q = V(t).
\]

**Solution:** Plugging in the given data leads to the DE

\[
\ddot{Q} + R\dot{Q} + 99^2Q = \sum_{n\geq 1, \text{odd}} \frac{\sin nt}{n}
\]

The characteristic polynomial is \( p(r) = r^2 + Rr + 99^2 \).

The actual input signal (the right hand side) is complicated, so we’ll build up the solution in stages, starting with easier input signals, using ERF, complex replacement, and superposition:

<table>
<thead>
<tr>
<th>input signal ( V(t) )</th>
<th>steady-state output signal ( Q(t) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( e^{nit} )</td>
<td>( \frac{1}{p(ni)} e^{nit} )</td>
</tr>
<tr>
<td>( \sin nt )</td>
<td>( \frac{1}{n} \Im \left( \frac{1}{p(ni)} e^{nit} \right) )</td>
</tr>
<tr>
<td>( \frac{1}{n} \sin nt )</td>
<td>( \frac{1}{n} \Im \left( \frac{1}{p(ni)} e^{nit} \right) )</td>
</tr>
<tr>
<td>( \sum_{n\geq 1, \text{odd}} \frac{1}{n} \sin nt )</td>
<td>( \sum_{n\geq 1, \text{odd}} \frac{1}{n} \Im \left( \frac{1}{p(ni)} e^{nit} \right) )</td>
</tr>
</tbody>
</table>

The term indexed by \( n \) in the output signal is a sinusoid of angular frequency \( n \). If we convert \( \frac{1}{p(ni)} \) to polar form, we see that that sinusoid has amplitude

\[
\frac{1}{n} \left| \frac{1}{p(ni)} \right|.
\]

The problem tells us that the amplitude of the term with angular frequency 99 (the \( n = 99 \) term) is approximately 1, so

\[
1 \approx \frac{1}{99} \left| \frac{1}{p(99i)} \right|
\]

\[
\approx \frac{1}{99|p(99i)|}
\]

\[
\approx \frac{1}{99|p(99i)|^2 + R(99i) + 99^2|}
\]

\[
\approx \frac{1}{99|R(99i)|}
\]

\[
\approx \frac{1}{99^2R}
\]
and

\[ R \approx \frac{1}{99^2} \approx \frac{1}{100^2} \approx 10^{-4}. \]

**Remark 26.2.** We have

\[ \left| \frac{1}{n} \left| \frac{1}{p(ni)} \right| \right| \approx \frac{1}{n|99^2 - n^2 + 10^{-4}ni|}, \]

and if \( n \neq 99 \), this is much less than 1, so it is really true that for \( R \approx 10^{-4} \), the steady-state charge \( Q(t) \) is approximately a sinusoid of amplitude 1 and angular frequency 99.

Why is (near) resonance occurring? Because 99\( i \) is very close to a root of \( p(r) \).

26.3. **Phase plane trajectories.**

**Problem 26.3.** Suppose that \( x(t) = ce^{-kt} \cos \omega t \) for some positive constants \( c, k, \) and \( \omega \).
Which of the pictures below is most likely to represent the trajectory in the phase plane with horizontal axis \( x \) and vertical axis \( \dot{x} \)?

![Phase plane graphs](image)

**Solution:** First, \( x \to 0 \) as \( t \to +\infty \); this rules out all the graphs in the first row. Next, the trajectory should go to the right when \( \dot{x} > 0 \) and to the left when \( \dot{x} < 0 \). This rules out the
bottom right graph, and the bottom middle graph too because the motion would have to be inward if \( x \to 0 \). Thus the answer is the bottom left graph. \( \square \)


**Problem 26.4.** Let \( f \) be a periodic function of period 4 such that

\[
f(t) = \begin{cases} 
    t, & \text{if } 1 \leq t < 3 \\
    12 - t, & \text{if } 7 \leq t < 9.
\end{cases}
\]

Let \( g(t) \) be the Fourier series of \( f \). What are \( f(5) \) and \( g(5) \)?

**Solution:** We have \( f(5) = f(1) = 1 \). On the other hand, \( f(5^-) = f(9^-) = 3 \) and \( f(5^+) = f(1^+) = 1 \), so

\[
g(5) = \frac{f(5^-) + f(5^+)}{2} = \frac{3 + 1}{2} = 2. \quad \square
\]

26.5. D’Alembert’s solution to the wave equation. The general solution to the wave equation

\[
\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}
\]

(without boundary conditions) is

\[
u(x, t) = f(x - ct) + g(x + ct),
\]

where \( f \) and \( g \) are any functions. The \( f \) and \( g \) are like the parameters \( c_1 \) and \( c_2 \) in the general solution to an ODE: to find them, one must use initial conditions.


The following problem is more time-consuming than any problem that would actually appear on an exam. We only started it in lecture.

**Problem 26.5.** An insulated metal rod of length \( \pi/2 \) and thermal diffusivity 3 has exposed ends. Initially it is at a constant temperature 5, but then its ends are held at temperature 0 and 20, respectively. What is its temperature as a function of position and time?

**Solution:** Modeling it as usual leads to

\[
\frac{\partial u}{\partial t} = 3 \frac{\partial^2 u}{\partial x^2}
\]

\[
u(0, t) = 0 \quad \text{for all } t > 0 \quad \text{(left boundary condition)}
\]

\[
u(\pi/2, t) = 20 \quad \text{for all } t > 0 \quad \text{(right boundary condition)}
\]

\[
u(x, 0) = 5 \quad \text{for all } x \in (0, \pi/2) \quad \text{(initial condition)}.
\]

Temporarily forget the initial condition!
Now we are solving the PDE only with boundary conditions. One of the boundary conditions is inhomogeneous, so first find the general solution \( u_h \) to the PDE with homogeneous boundary conditions:

\[
\frac{\partial u}{\partial t} = 3 \frac{\partial^2 u}{\partial x^2} \]

\( u(0, t) = 0 \) for all \( t > 0 \) (left boundary condition)

\( u(\pi/2, t) = 0 \) for all \( t > 0 \) (right boundary condition).

The general solution will be a linear combination of normal modes of the form \( u(x, t) = v(x)w(t) \) for some nonzero functions \( v \) and \( w \). Substituting into the PDE and (homogeneous) boundary conditions leads to

\[
v(x)\dot{w}(t) = 3v''(x)w(t) \]

\( v(0)w(t) = 0 \) for all \( t > 0 \)

\( v(\pi/2)w(t) = 0 \) for all \( t > 0. \)

Since \( w(t) \) is not identically 0, the last two equations are equivalent to \( v(0) = 0 \) and \( v(\pi/2) = 0 \). Separating variables in the first equation leads to

\[
\frac{v''(x)}{v(x)} = \lambda = \frac{\dot{w}(t)}{3w(t)}
\]

for some constant \( \lambda \). Thus we need to solve

\[
v''(x) = \lambda v(x) \]

\( v(0) = 0 \)

\( v(\pi/2) = 0 \)

\( \dot{w}(t) = 3\lambda w(t). \)

To solve the first three equations, break into cases according to the sign of \( \lambda \), and use that the characteristic polynomial is \( r^2 - \lambda \).

Case 1: \( \lambda > 0 \). Then \( v(x) = ae^{\sqrt{\lambda}x} + be^{-\sqrt{\lambda}x} \) for some constants \( a \) and \( b \). Substituting \( x = 0 \) or \( x = \pi/2 \) and using the boundary conditions leads to the system

\[
\begin{pmatrix}
1 & 1 \\
e^{\sqrt{\lambda}\pi/2} & e^{-\sqrt{\lambda}\pi/2}
\end{pmatrix}
\begin{pmatrix}
a \\
b
\end{pmatrix} =
\begin{pmatrix}
0 \\
0
\end{pmatrix},
\]

but

\[
\det\begin{pmatrix}
1 & 1 \\
e^{\sqrt{\lambda}\pi/2} & e^{-\sqrt{\lambda}\pi/2}
\end{pmatrix} = e^{-\sqrt{\lambda}\pi/2} - e^{\sqrt{\lambda}\pi/2} < 0,
\]

so the only solution is \((a, b) = (0, 0)\), so \( v(x) = 0 \) for all \( x \). But we are looking for a nonzero solution.
Case 2: \( \lambda = 0 \). This time \( v(x) = a + bx \) for some constants \( a \) and \( b \). Substituting \( x = 0 \) or \( x = \pi/2 \) and using the boundary conditions leads to \( a = 0 \) and \( a + b\pi/2 = 0 \), so \( (a, b) = (0, 0) \) again, so \( v(x) = 0 \) for all \( x \).

Case 3: \( \lambda < 0 \). Write \( \lambda = -\omega^2 \) for some positive real number \( \omega \), so that the roots of the characteristic polynomial are \( \pm i\omega \). Then \( e^{i\omega x}, e^{-i\omega x} \) is a basis, and \( \cos \omega x, \sin \omega x \) is a real-valued basis. In other words, \( v(x) = a \cos \omega x + b \sin \omega x \) for some constants \( a \) and \( b \). Substituting \( x = 0 \) and using the first boundary condition leads to \( 0 = a \). Thus \( v(x) = b \sin \omega x \). Substituting \( x = \pi/2 \) and using the second boundary condition leads to \( 0 = b \sin \omega\pi/2 \). If we want a nonzero solution \( v(x) \), then \( b \) must be nonzero, so \( \omega\pi/2 \) must be an integer multiple of \( \pi \), say

\[ \omega\pi/2 = n\pi, \]

so \( \omega = 2n \) for some integer \( n \) (and \( n \) is positive since \( \omega \) was positive). In this case, \( \lambda = -4n^2 \) and we get \( v(x) = \sin 2nx \) and \( w(t) = e^{-12n^2t} \) (up to scalar multiples), so

\[ u(x, t) = e^{-12n^2t} \sin 2nx. \]

Hence the general solution to the PDE with homogeneous boundary conditions is

\[ u_h(x, t) := \sum_{n \geq 1} b_n e^{-12n^2t} \sin 2nx. \]

Next we need one solution to the PDE with inhomogeneous boundary conditions. Since the boundary conditions are constant in time, we look for a solution \( u(x, t) \) that does not depend on \( t \). In this case, the PDE becomes \( \frac{\partial^2 u}{\partial x^2} = 0 \), so \( u(x, t) = a + bx \) for some constants \( a \) and \( b \). The boundary condition \( u(0, t) = 0 \) forces \( a = 0 \), and then \( u(\pi/2, t) = 0 \) forces \( b\pi/2 = 20 \), so \( b = 40/\pi \). Thus

\[ u_p(x, t) := \frac{40}{\pi} x \]

is a solution to the PDE with inhomogeneous boundary conditions.

By the inhomogeneous principle, the function

\[ u = u_p + u_h = \frac{40}{\pi} x + \sum_{n \geq 1} b_n e^{-12n^2t} \sin 2nx \]

is the general solution to the PDE with inhomogeneous boundary conditions.

Finally, to determine the \( b_n \), we bring back the initial condition. Substituting \( t = 0 \) leads to

\[ 5 = \frac{40}{\pi} x + \sum_{n \geq 1} b_n \sin 2nx \quad \text{for} \ x \in (0, \pi/2) \]

\[ 5 - \frac{40}{\pi} x = \sum_{n \geq 1} b_n \sin 2nx \quad \text{for} \ x \in (0, \pi/2). \]
The right hand side is an odd periodic function of period $\pi$, so extend the left hand side to an odd periodic function of period $\pi$, and use the Fourier coefficient formula (with $L = \pi/2$) to find $b_n$:

$$b_n = \frac{2}{\pi/2} \int_0^{\pi/2} \left( 5 - \frac{40}{\pi} x \right) \sin 2nx \, dx.$$ 

Integration by parts leads to

$$b_n = \begin{cases} 
\frac{40}{\pi n}, & \text{if } n \text{ is even}, \\
-\frac{20}{\pi n}, & \text{if } n \text{ is odd}.
\end{cases}$$

Substituting these values into the general solution gives the final answer:

$$u(x,t) = \frac{40}{\pi} x + \sum_{n \geq 1} \frac{40}{\pi n} e^{-12n^2 t} \sin 2nx - \sum_{n \geq 1} \frac{20}{\pi n} e^{-12n^2 t} \sin 2nx.$$ 

27. What math subject to take next?

Now that you have finished 18.02 and 18.03, there are many options open to you. Here are some of them:

- Linear algebra: 18.06, 18.700, or 18.701. Of these, 18.701 is for students who are already very comfortable with writing proofs.
- Probability and statistics: 18.05 (spring) or 6.041, or 18.440.
- Discrete math: 18.062/6.042, or 18.310A or 18.310.
- Real analysis: 18.100A, 18.100B, or 18.100C. The A version is less abstract.
- Complex analysis: 18.04 (spring).
- Continuous applied math: 18.311 (spring).

Special advice for potential math majors/minors, or double majors involving math:

- A good starting point is 18.700 or 18.100C or 18.310.
- \textbf{18.100C} and \textbf{18.310} are 15 units instead of 12, and give CI-M credit in math because they include practice in written and oral communication; this feedback is important if you are learning to write proofs for the first time.
- In general, you’re probably better off taking the versions with first decimal digit 1 or higher.
- Instead of taking 18.04, wait until you’ve finished 18.100 so that you can take the more advanced complex analysis 18.112.

28. Thank you

Thank you to Jennifer French of MITx, to Galina Lastovkina and the rest of the staff in the Mathematics Academic Services office, to Jean-Michel Claus and the rest of the team that developed the mathlets, to Professors Jerison and Miller for sharing their advice and
materials from previous semesters, to the MIT audio-visual staff on hand at each lecture, to
the students erasing the boards before each lecture, and to all 16 recitation instructors for
their hard work over the semester!

And thank you to all the 18.03 students for making this class fun to teach! I hope that
you all ace the final!