Math 18500 Week 1: First Order Differential Equations Part I

Introduction. A *differential equation* is a relationship between an unknown function and its derivatives.

In science and engineering, differential equations are used to model physical quantities which change over time. The prototypical example is *Newton's law*, which is a *second order* differential equation

$$F = ma = m\frac{d^2x}{dt^2}.$$

This equation models the *position* $x(t)$ of a moving object, as a function of time. Newton's law allows us to predict the future motion of any object, if we know all of the forces acting on it.

For example, a falling object near the surface of the Earth experiences a constant downward force, so its height $y(t)$ above the ground satisfies an equation of the form

$$m\frac{d^2y}{dt^2} = my''(t) = -mg.$$

For concreteness, let's take the particular value $g = 10$. Then the equation becomes

$$y''(t) = -10.$$

We can solve this equation by integrating both sides twice. Integrating once gives

$$y'(t) = -10t + c_1,$$

where c_1 is a constant of integration, and integrating a second time gives

$$y(t) = -5t^2 + c_1t + c_2,$$

where c_2 is a second constant of integration.

Notice that we do not obtain a *single* solution. Instead, the solution we have found depends on two unknown constants c_1 and c_2, whose values can be chosen arbitrarily. This is a general phenomenon - differential equations usually have *many* different solutions.

In our example, the ambiguity of our solution has an important physical meaning: the trajectory of a moving object depends on its initial position and velocity. For example, if we set $t = 0$, we find that

$$y'(0) = -0 + c_1 = c_1$$

and

$$y(0) = -0 + 0 + c_2 = c_2.$$

So, the constants of integration c_1 and c_2 precisely reflect the initial values of $y(t)$ and $y'(t)$.

In general, to single out a particular solution of a differential equation we must select some *initial conditions*. A differential equation together with a prescribed set of initial conditions is called an *initial value problem*.

For example, to find the motion of a ball which is dropped from rest at a height of 10 meters, we would need to solve the initial value problem

$$y'' = -1, \quad y(0) = 10, \quad y'(0) = 0.$$

Our computations above show that the solution of this initial value problem is

$$y(t) = 10 - 5t^2.$$

Differential equations are classified by their *order*, which is the highest order derivative of the unknown function $y(t)$ that appears in the equation. Generally speaking, solving for this highest order derivative yields an equation of the form

$$y^{(n)}(t) = F\left(t, y, y'(t), \ldots, y^{(n-1)}(t)\right).$$

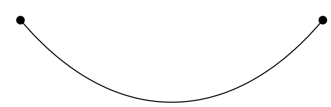
To determine a particular solution of an n^{th} order differential equation of this form, we must specify a total of n initial conditions,

$$y(t_0), y'(t_0), \ldots, y^{(n-1)}(t_0),$$

at an initial time t_0.

You will also encounter differential equations where the variable in the unknown function does not represent time, but instead represents some other quantity.

For example, consider a cable suspended between two points,



The shape of the cable can be modeled as the graph of a function $y = y(x)$. To find this unknown function, it is necessary to solve the following second order equation (which can be derived from mechanical principles):

$$y''(x) = a\sqrt{1 + y'(x)^2}.$$

Here a is a constant (determined by the total length of the cable and the locations of its endpoints).

You may not see how to solve an equation like this, but if someone tells you that they have a solution, it is easy to check whether or not they are correct! For example, in the case $a = 1$, one solution is

$$y(x) = \frac{e^x + e^{-x}}{2} = \cosh(x).$$

To verify this is a solution, we just need to calculate its derivatives:

$$y'(x) = \frac{e^x - e^{-x}}{2} = \sinh(x)$$

$$y''(x) = \frac{e^x + e^{-x}}{2} = \cosh(x)$$

We can then substitute the derivatives into our equation, and verify that the two sides are equal:

$$\sqrt{1 + y'(x)^2} = \sqrt{1 + \cosh^2(x)} = \sinh(x) = y''(x)$$

You will also encounter differential equations in which the unknown function is a function of more than one variable. These are called *partial differential equations*, because they involve partial derivatives of the unknown function.

For example, in 184 you encountered the following partial differential equation:

$$\vec{\nabla}^2 \phi = \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0$$

In this case the unknown function ϕ was a function of two variables:

$$\phi = \phi(x, y).$$

By contrast, differential equations in which the unknown function has only one dependent variable are called *ordinary differential equations*. With the exception of a few special cases, we will only be discussing ordinary differential equations in this course.

Slope Fields. The simplest differential equations are *first order equations*,

$$y'(x) = f(x, y(x)).$$

To write an equation like this, we will usually use the notation

$$\frac{dy}{dx} = f(x, y),$$

in which the dependence of y on x is suppressed.

Most first order equations can't be solved explicitly. However, it is always possible to visualize the solutions by drawing careful pictures (either by hand, or with the help of a computer).

To plot the solutions of an equation, we must first understand what the equation means geometrically. Recall that the first derivative $y'(x)$ represents the *slope* of the graph of $y(x)$. So, if $y = y_1(x)$ is one particular solution of the equation

$$\frac{dy}{dx} = f(x, y),$$

we can conclude that

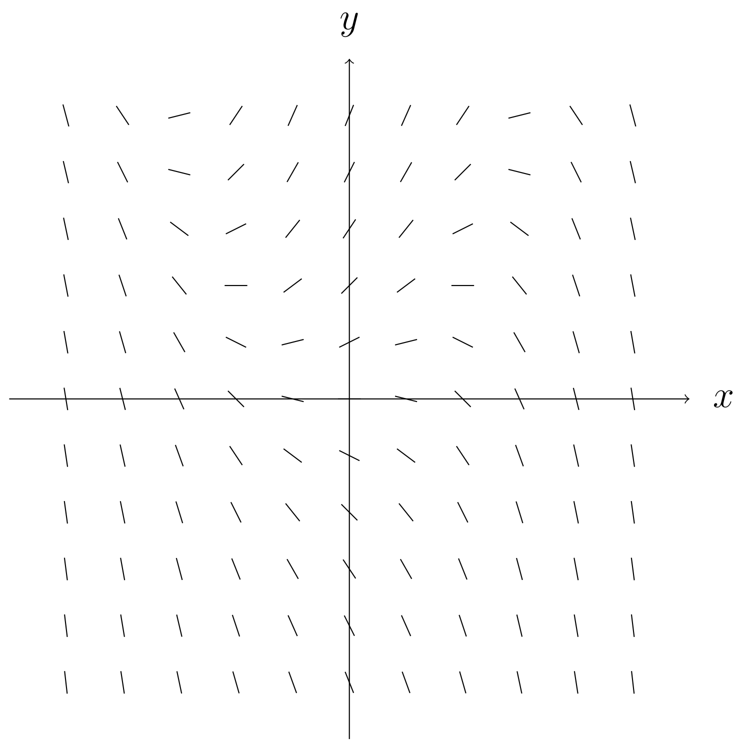
The slope of the graph of $y_1(x)$ at the point $(x, y_1(x))$ is $f(x, y_1(x))$.

To see what this looks like, the first step is to draw a line segment of slope $f(x, y)$ at a large number of points (x, y). The diagram which results from this process is called a *slope field*.

For example, here is a slope field representing the equation

$$\frac{dy}{dx} = y^2 - x,$$

drawn with the help of a computer:



What the computer has done here is plugged the values

$$x = -2.5, -2, \ldots, 2, 2.5$$

$$y = -2.5, -2, \ldots, 2, 2.5$$

into the function

$$f(x, y) = y^2 - x.$$

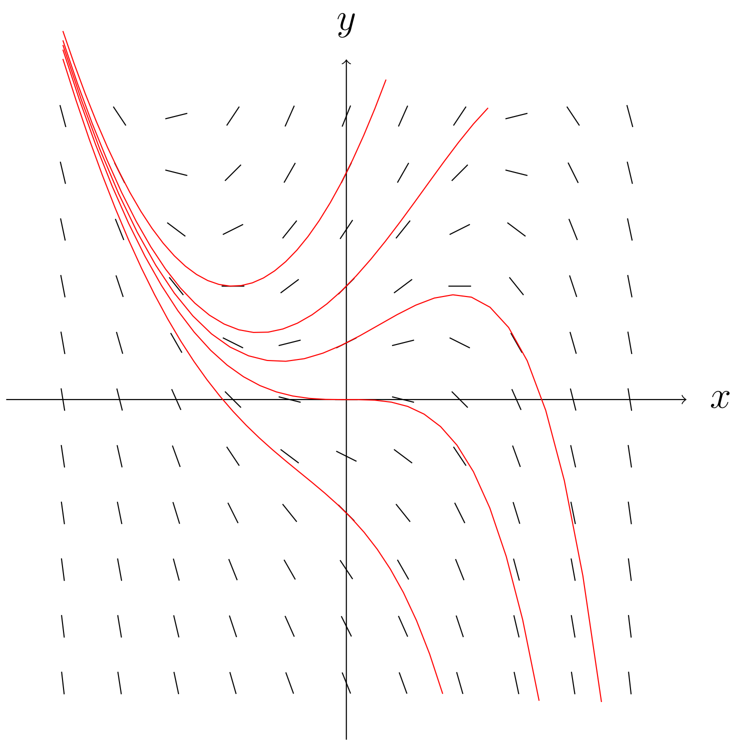
For each of those values it drew a line segment of slope $f(x, y)$. For example, at the point $(1, 1)$ we have

$$f(1, 1) = 1^2 - 1 = 0,$$

so the computer drew a horizontal line segment at that point.

Once we have drawn the slope field corresponding to our differential equation, we can begin to visualize the solutions. To do this we must draw the *integral curves* for our slope field. By definition, these are curves which are *everywhere tangent* to the slope field.

For example, here are the graphs of a few different integral curves for the equation $\frac{dy}{dx} = y - x^2$:



Notice that at each point where a integral curve passes through one of the line segments we have drawn, the curve and the line segment are tangent to one another. If we were to pick any point along any of the integral curves and evaluate the slope function at the point, the result would be the slope of the integral curve.

In particular, the graph of any solution to our equation will be (a piece of) an integral curve, because the slope of its graph at any point (x, y)

will be the slope prescibed by the slope function - that is exactly what is meant by the equation

$$\frac{dy}{dx} = f(x, y).$$

So, the curves we have drawn above are actually *graphs* of several solutions of our equation.

To draw pictures like this by hand, it is important to be more clever than a computer (otherwise your picture will look terrible). One method is to draw what are called **isoclines**. These are curves where the slope has a particular constant value C:

$$f(x, y) = C.$$

In other words, the isoclines are *level curves* of the slope function.

For example, suppose we wanted to visualize the slope field corresponding to the equation

$$\frac{dy}{dx} = -\frac{x}{y}.$$

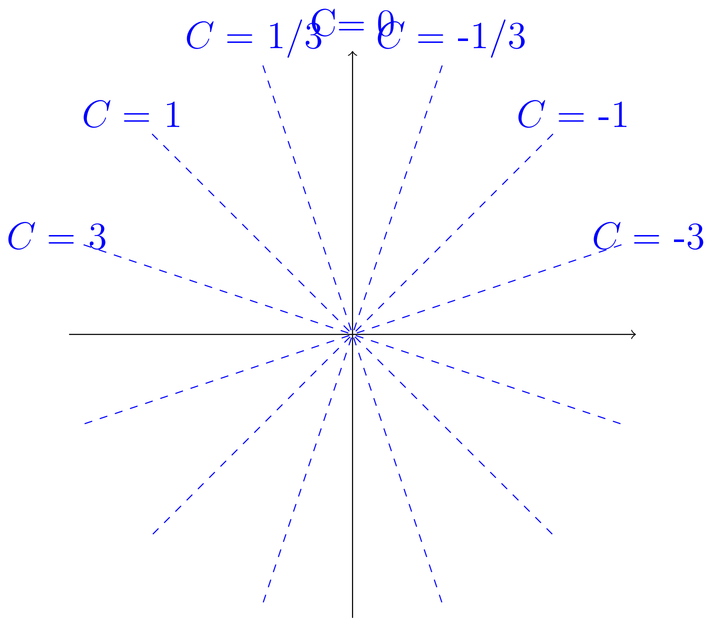
Then we would start by drawing curves of the form

$$-\frac{x}{y} = C$$

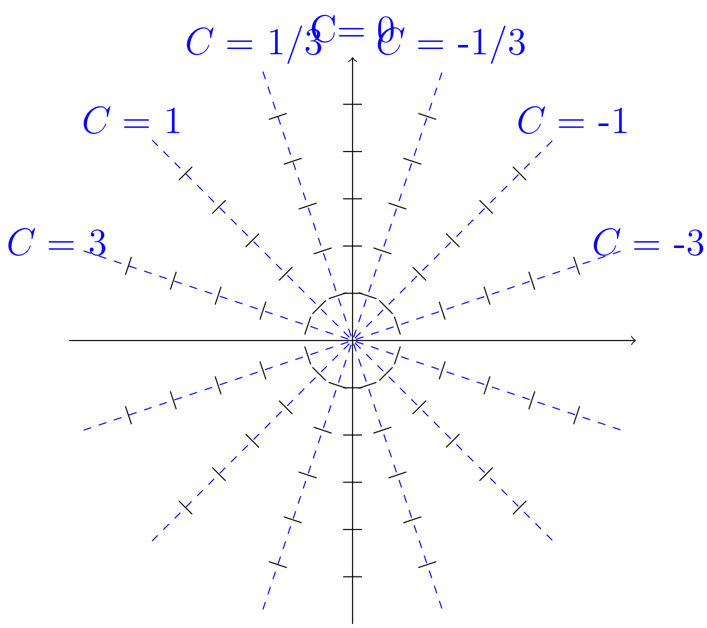
for different values of C. These are lines through the origin,

$$y = -\frac{1}{C}x \text{ or } x = -Cy$$

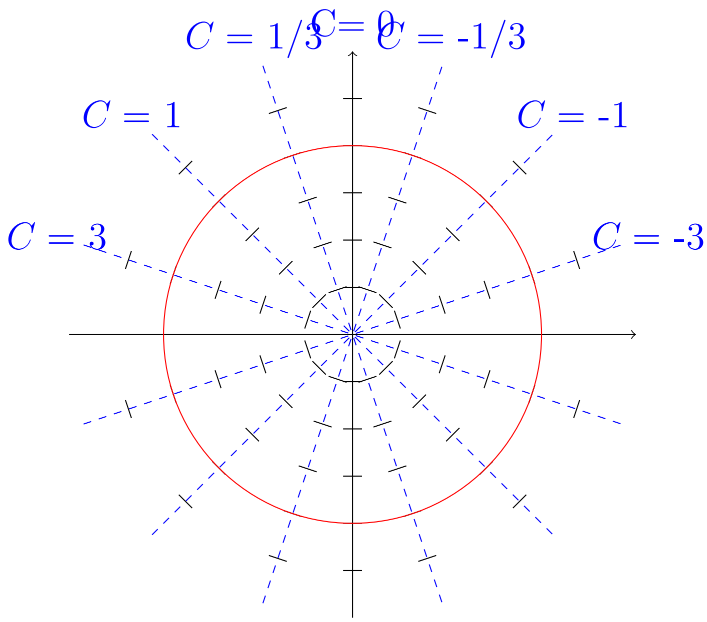
so they look like this:



To get the slope field, we draw a series of line segments with slope C, on each of these lines:



From the picture, we can see that the integral curves are *circles*:



Each circle corresponds to two solutions of our differential equation,

$$y_1(x) = \sqrt{R^2 - x^2} \text{ and } y_2(x) = -\sqrt{R^2 - x^2},$$

where R is the radius of the circle. Indeed, we can verify algebraically that y_1 is a solution of the equation:

$$y_1'(x) = \frac{d}{dx}\left[\sqrt{R^2 - x^2}\right] = \frac{-2x}{2\sqrt{R^2 - x^2}} = -\frac{x}{\sqrt{R^2 - x^2}} = -\frac{x}{y_1(x)}$$

You can check for yourself that $y_2(x)$ is also a solution (for any value of R).

One thing to observe from this example is that it is not always possible to extend our solutions to *all* values of x. The solutions we have graphed only exist on an interval $[-R, R]$, and cannot be extended any further, because the develop vertical tangent lines at $x = \pm R$.

In a mathematically rigorous treatment of differential equations, we would prove the following theorem:

Existence of Solutions to First Order Differential Equations. Let $f(x,y)$ be a function which is continuous in the vicinity of a point (x_0, y_0). Then the initial value problem

$$\frac{dy}{dx} = f(x,y)\ ,\quad y(x_0) = y_0$$

has a solution $y = y(x)$, which can be defined on a small interval $(x_0 - \epsilon, x_0 + \epsilon)$, but which cannot necessarily be extended beyond this interval.

For example, the initial value problem

$$\frac{dy}{dx} = -\frac{x}{y}\ ,\quad y(1) = 0$$

does not have a solution on any interval of the form $(1-\epsilon, 1+\epsilon)$. This is consistent with the theorem, because the function $-\frac{x}{y}$ is not continuous at the point $(x_0, y_0) = (1, 0)$ - its value approaches infinity at this point (and this results in an infinite slope for the integral curve).

Next week you will see how power series can be used to find solutions of most differential equations which come up in practice. Once you learn this method, it will become clear to you why all of the important aspects of the theorem are true (i.e. why solutions always exist, and why they cannot necessarily be extended beyond a certain interval).

Autonomous Equations. Often one needs to model the behavior of a physical systems whose governing laws are fixed and not dependent on time. This leads to differential equations of the following special form:

$$\frac{dy}{dt} = f(y).$$

Equations of this form, where the right hand side depends only on y and not on t, are said to be *autonomous*.[1]

As a first example of an autonomous equation, suppose we take a hot piece of metal and submerge it in a large pool of water. Over time the metal will cool down, and its temperature will approach the temperature of the water.[2] This can be modeled using *Newton's law of cooling*, which states that the rate of change of the temperature of the metal is proportional to the temperature difference between the metal and the water.

More explicitly, if we write

$$T(t) = \text{temperature of the metal at time } t,$$

then Newton's law of cooling can be expressed as follows:

$$\frac{dT}{dt} = -k(T - T_{water}).$$

Notice that the right hand side of this equation does not depend on t - the equation is autonomous.

The constant k which appears in Newton's law must be determined empirically - it will depend on the particular type of metal we use. For the purposes of illustration, let's choose our units of measurement so that $k = T_{water} = 1$ - then the equation simplifies to

$$\frac{dT}{dt} = 1 - T.$$

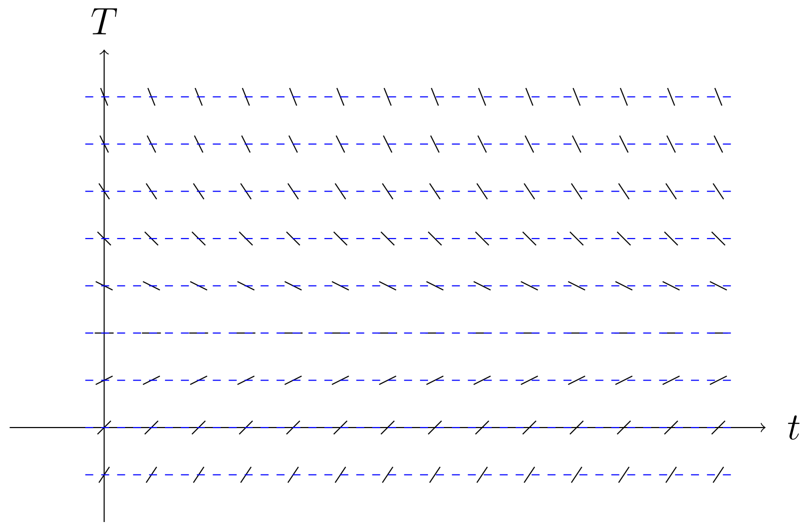
To visualize the solutions of this equation, we must make a slope field. This can be done using the simple observation that the isoclines

$$1 - T = \text{constant}$$

[1] The word autonomous comes from ancient Greek, and means "independent" or "subject to one's own laws".

[2] Obviously, the water will also be heated by the metal, but if we assume that the volume of water is sufficiently large then we can neglect this effect.

are horizontal lines:



When you see this picture, one thing should immediately pop out: the horizontal slopes along the isocline $T = 1$. In this case, the isocline is actually a integral curve, corresponding to the *constant* solution

$$T = 1.$$

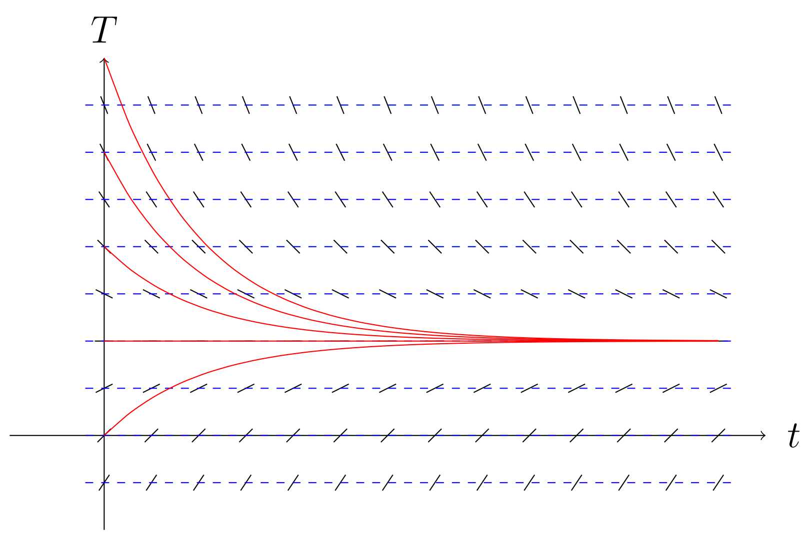
It's easy to verify that this in fact a solution:

$$\frac{dT}{dt} = 0 = 1 - 1 = 1 - T.$$

Physically, this makes sense - if the metal is already the same temperature as the water when we drop it in, then its temperature won't change.

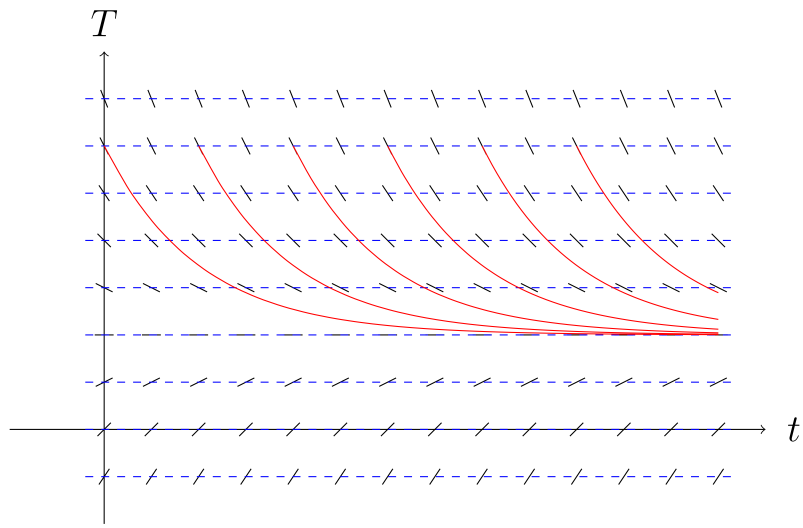
A second thing which should pop out at you is that the slopes above the line $T = 1$ are all negative, whereas the slopes below the line $T = 1$ are all positive. This makes physical sense: if the initial temperature of the metal is greater than the temperature of the water, then water will cool the metal off, and its temperature will decrease. If the initial temperature of the metal is less than the temperature of the water, then the water will heat the metal up, and its temperature will increase.

If we plot several solutions with different initial values at $t = 0$, here is what we see:



As you can see, the solutions all approach the value $T = 1$ as $t \to \infty$. We will soon be able to see this directly, by explicitly solving the equation.

We can see another interesting feature of this equation by plotting solutions which start at the same value of T, but at different times t:



The solutions are exactly the same, except they are shifted in time. This is a reflection of the fact that the equation is autonomous - the solutions obey the same rules, no matter what time they start, so their time evolution must be the same.

All three of the observations we have made above can be applied to arbitrary autonomous equations. In general, the following principles can be applied to any autonomous equation:

1. If $f(y_0) = 0$, then the constant function $y(t) = y_0$ is a solution of the autonomous equation $\frac{dy}{dt} = f(y)$. Solutions like this are referred to as **equilibrium solutions**, and the values y_0 are referred to as **equilibrium values**.

2. If $f(y_0) > 0$, then any solution of the equation $\frac{dy}{dt} = f(y)$ will increase as $t \to \infty$, eventually approaching an equilibrium value (or diverging to infinity).

3. If $f(y_0) < 0$, then any solution of the equation $\frac{dy}{dt} = f(y)$ will decrease as $t \to \infty$, eventually approaching an equilibrium value (or diverging to negative infinity).

4. If $y(t)$ is any solution of the equation $\frac{dy}{dt} = f(y)$, then $y(t - t_0)$ is also a solution, for any value t_0.

We will not attempt to fully justify these principles (although you can easily verify 1 and 4 yourself). Instead, we will illustrate all of them in the context of a single autonomous equation,

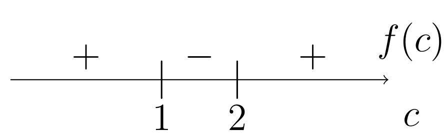
$$\frac{dc}{dt} = (c - 1)(c - 2).$$

At the end of this section you will see how an equation like this might come up in chemistry (Hint: c stands for concentration). But for the moment, let's try to understand what its solutions look like mathematically.

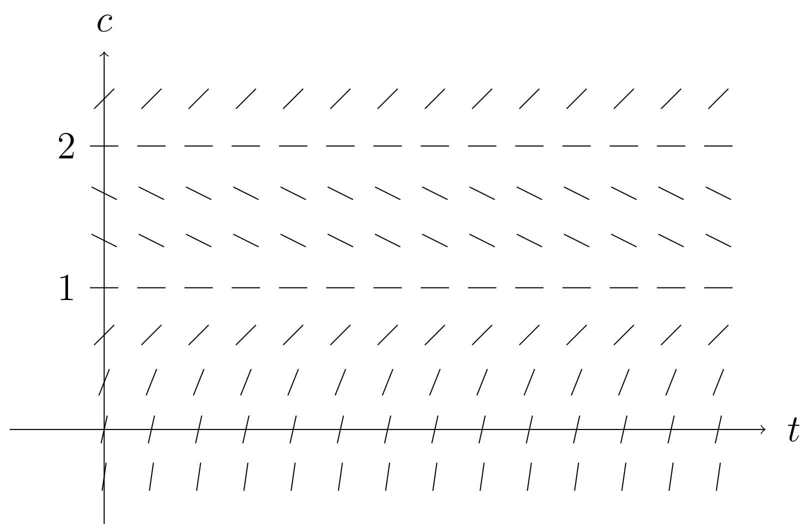
The first step in understanding this equation is to graph the function

$$f(c) = (c - 1)(c - 2).$$

For values of c that are between 1 and 2, it is negative (because it is the product of a positive term and a negative term). At $c = 1$ and $c = 2$, it is zero. For all other values, it is positive. To visualize this, we can make a *sign plot*:

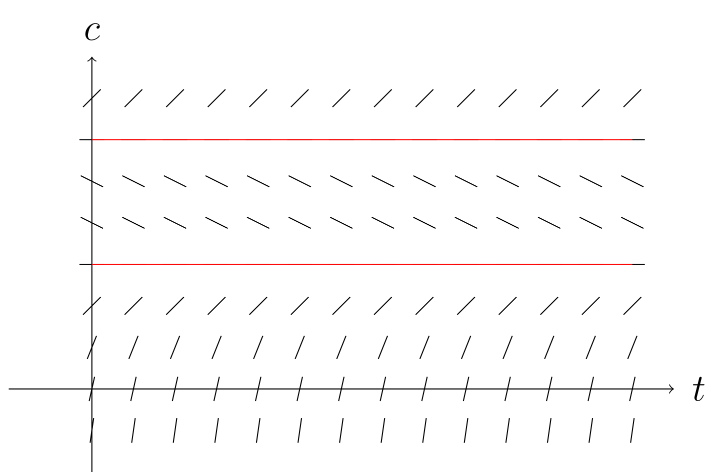


Based on the sign plot, we can draw the slope field:

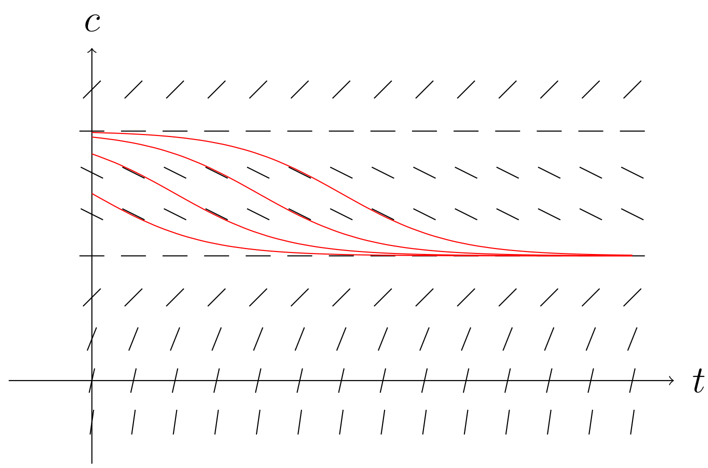


Notice that we drew a negative slope for values of c between 1 and 2, and a positive slope for all other values.

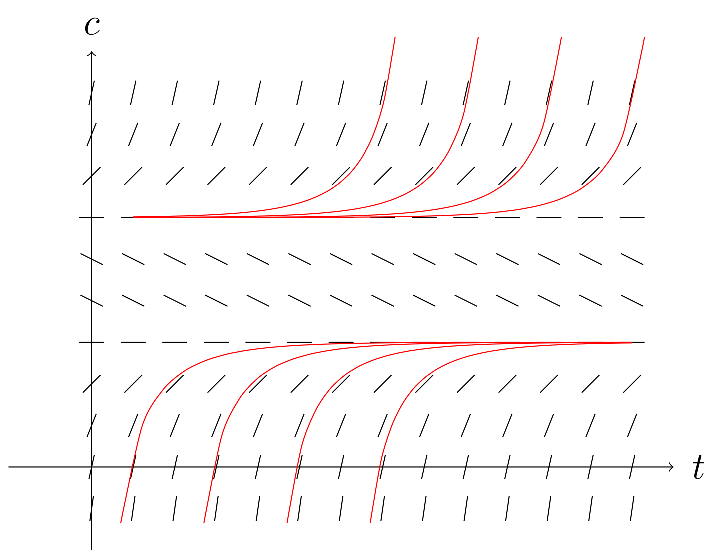
If $c = 1$ or $c = 2$, the slope is 0. This gives us two equilibrium solutions:



If $1 < c < 2$, the solutions decrease until they approach the equilibrium value $c = 1$:



If $c < 1$, the solutions increase until they approach the equilibrium value $c = 1$. If $c > 2$, they increase to infinity:



To explain how we might come upon an equation like this one, consider a chemical reaction of the following form:

$$A + B \to C$$

For a reaction like this it is reasonable to use the following model:

$$\frac{dc}{dt} = kab.$$

Here k is a constant (the *reaction rate constant*), and $a = [A]$, $b = [B]$, and $c = [C]$ represent the concentrations of the three chemicals. Since the reaction produces one molecule of C for each pair of molecules of A and B, we must have

$$a_0 - a = b_0 - b = c - c_0$$

where $[A]_0$, $[B]_0$, and $[C]_0$ are the initial concentrations. This leads to an initial value problem:

$$\frac{dc}{dt} = k(a_0 + c_0 - c)(b_0 + c_0 - c)\ , \quad c(0) = c_0.$$

If we take the particular values $k = 1$, $a_0 = 1$, $b_0 = 2$, and $c_0 = 0$, then the initial value problem becomes

$$\frac{dc}{dt} = (c - 1)(c - 2)\ , \quad c(0) = 0$$

Looking at the slope field above, we can see what happens - the concentration of C increases, approaching the limiting value 1 as $t \to \infty$. This makes sense physically: the concentration of C can't increase beyond 1, because at that point all of the molecules of A would have been used up in the reaction.

Separable Equations. A first order equation is said to be **separable** if it can be written in the form

$$\frac{dy}{dx} = f(x)g(y),$$

where f and g are arbitrary functions.

Note that autonomous equations are a special case ($f(x) = 1$). In fact, we can borrow an idea from that context and notice that separable equations often have *equilibrium* solution, which are solutions of the form

$$y = y_0 = \text{constant},$$

where

$$g(y_0) = 0.$$

The first thing we should always do with a separable equation is find all solutions of this form.

Once we've found the equilibrium solutions, we attempt to find all solutions $y = y(x)$ which satisfy $g(y) \neq 0$. To do this, we just divide both sides of the equation by $g(y)$:

$$\frac{1}{g(y)}\frac{dy}{dx} = f(x)$$

This operation is valid *precisely* because $g(y) \neq 0$.

We can then do something cute and multiply both sides of the equation by dx: [3]

$$\frac{1}{g(y)}dy = f(x)dx$$

[3]You might find this notation a bit unclear. To understand it, you have to keep in mind that y is a function of x, so the left hand side should really be

$$\frac{1}{g(y(x))}y'(x)dx$$

We can then integrate both sides of the equation: [4]

$$\int \frac{1}{g(y)}dy = \int f(x)dx + C$$

Evaluating the integrals leads to an equation of the form

$$G(y) = F(x) + C,$$

which we can hope to solve, obtaining a formula for y as a function of x.

For example, consider the equation

$$\frac{dy}{dx} = -\frac{x}{y}$$

from the "slope fields" section of this week's notes. To solve this equation we first separate the variables,

$$ydy = -xdx,$$

and then we integrate both sides,

$$\frac{y^2}{2} = -\frac{x^2}{2} + C$$

and solve for y,

$$y = \pm\sqrt{2C - x^2} = \pm\sqrt{R^2 - x^2}\ , \quad R = \sqrt{2C}.$$

When we integrate this, we immediately apply the "substitution" technique, with $y = y(x)$ and $dy = y'(x)dx$. This gives

$$\int \frac{1}{g(y(x))}y'(x)dx = \int \frac{1}{g(y)}dy.$$

[4]Technically, we need to include a constant of integration on both sides. But we can always subtract the constant on the left hand side and absorb it into the constant on the right hand side.

The solution takes the exact form we predicted, based on the slope field picture!

We can use this same technique to solve the autonomous equations we studeied in the previous section. For example, consider the equation

$$\frac{dT}{dt} = 1 - T,$$

where $T(t)$ represents the temperature of a iece of metal which has been submerged in a water bath.

To solve this equation, we first look for its equilibrium solutions:

$$1 - T = 0 \implies T = 1.$$

Next, we find the non-equilibrium solutions, by assuming that $T \neq 1$ and separating the variables:

$$\frac{dT}{T-1} = -dt$$

Integrating both sides, we obtain

$$\ln|T-1| = -t + C_1.$$

Here we have to be careful to include the absolute value sign, to avoid missing some solutions!

Solving for T, we find that

$$|T-1| = e^{-t+C_1}.$$

$$T = 1 \pm e^{-t+C_1}$$

The form of the solution reflects the fact that the equation is *autonomous*: its solutions can be translated in time.

At this point it is helpful to introduce a new constant C_2 and write

$$T = 1 + C_2 e^{-t}$$

Something nice happens when we do this, which is that although $C_2 = \pm e^{C_1}$ had to be a nonzero number, we can recover our equilibrium solution by setting $C_2 = 0$! So, the most general solution of the equation is

$$T = 1 + Ce^{-t},$$

where C is a completely arbitrary constant (either zero or nonzero).

Observe that all solutions approach the equilibrium value $T = 1$ in the limit as $t \to \infty$:

$$\lim_{t\to\infty} T = \lim_{t\to\infty} 1 + Ce^{-t} = 1 + 0 = 1.$$

This confirms our geometric intuition: $T = 1$ is a *stable* equilibrium point.

Finally, let's solve the initial value problem which we considered in the second example:

$$\frac{dc}{dt} = (c-1)(c-2)\ , \quad c(0) = 0.$$

Here we have *two* equilibrium solutions, $c = 1$ and $c = 2$. Neither of these is consistent with our initial condition $c(0) = 0$, so we can ignore them here.

Since $c \neq 1$ and $c \neq 2$, we can proceed:

$$\frac{dc}{(c-1)(c-2)} = dt.$$

To integrate both sides of this equation, we need a special trick: partial fractions. We write

$$\frac{1}{(c-1)(c-2)} = \frac{A}{c-1} + \frac{B}{c-2}$$

and try to determine the appropriate values of A and B. One method of doing this is to clear the denominators on both sides:

$$1 = A(c-2) + B(c-1)$$

and then simplify:

$$1 = (A+B)c + (-2A - B).$$

In order for this equation to be valid for all values of c (which is what we're after), we must have

$$A + B = 0\ , \quad -2A - B = 1$$

It's not terribly difficult to solve this system of equations and obtain values for A and B.

However, there is a quicker method of obtaining the values of A and B, since we know that we will be able to solve for them! It's called the *Heaviside cover-up method*, and it goes like this. We take the equation

$$1 = A(c-2) + B(c-1)$$

and we substitute values of c which cause one or the other term on the right hand side to disappear (if you do this on paper, you can "cover up" that term with your fingers, hence the name of the method). For example, if we substitute $c = 1$, we obtain:

$$1 = A(1-2) + B(1-1) = -A$$

so we must have $A = -1$. Similarly if we substitute $c = 2$, we obtain the value of B:

$$1 = A(2-2) + B(2-1) = B.$$

Putting it together, we obtain the desired partial fractions decomposition:

$$\frac{1}{(c-1)(c-2)} = \frac{A}{c-1} + \frac{B}{c-2} = \frac{-1}{c-1} + \frac{1}{c-2} = \frac{1}{1-c} - \frac{1}{2-c}.$$

You should practice applying this method quickly and accurately - you'll need it at the end of the course when we do Laplace transforms.

Returning to our problem, we now have the equation

$$\frac{dc}{1-c} - \frac{dc}{2-c} = dt.$$

Integrating both sides, we obtain

$$-\ln|1-c| + \ln|2-c| = t + K$$

where K is a constant of integration.[5]

At this point, we can take account of the initial condition we have chosen, $c(0) = 0$. Because we are starting with this value, our solution will satisfy $1 - c(t) > 0$ and $2 - c(t) > 0$, at least for small values of t. Therefore, we can eliminate the absolute values:

$$-\ln(1-c) + \ln(2-c) = t + K$$

We can also obtain the correct value of K by substituting $t = 0$ and $c = 0$:

$$-\ln(1) + \ln(2) = 0 + K \implies K = \ln(2).$$

Finally, we can solve for c:

$$-\ln(1-c) + \ln(2-c) = t + \ln(2)$$

$$\ln\left(\frac{2-c}{1-c}\right) = t + \ln(2)$$

$$\frac{2-c}{1-c} = 2e^t$$

$$2 - c = (2 - 2c)e^t$$

[5] I guess this is a disadvantage of using c as the dependent variable - we need a different letter for the constant of integration!

$$2 - 2e^t = c - 2ce^t$$
$$c = \frac{2 - 2e^t}{1 - 2e^t} = \frac{2e^{-t} - 2}{e^{-t} - 2}$$

If we take the limit as $t \to \infty$, we see that the solution approaches the equilibrium value $c = 1$, as predicted:

$$\lim_{t\to\infty} c = \lim_{t\to\infty} \frac{2e^{-t} - 2}{e^{-t} - 2} = \frac{-2}{-2} = 1.$$