Calculus II: Lectures on Differential Equations

Kimball Martin*

December 9, 2015

These are notes for three lectures on differential equations for my Calculus II course at the University of Oklahoma in Fall 2015. Please let me know if you find any errors.

While our main motivation for developing integral calculus is to be able to determine things like area, volume and lengths of simple geometric objects, some of the major modern uses of this subject come out of the theory of differential equations, which are important in many subject such as physics, chemistry, biology, engineering and economics.

A **differential equation** is simply an equation involving the derivatives of a quantity y. Often y will be a function of time, usually denoted by t. Some simple examples are

$$y' = cy, \tag{1}$$
$$y'' = c, \tag{2}$$
$$y' = ct, \tag{3}$$
$$y'' = c \sin y. \tag{4}$$

Here c is assumed to be a constant.

The first equation says the growth of y at some time t is proportional to the value of y. This arises, for instance, in a very simplified model of population growth. The second equation says that y'' is constant, and can be interpreted as saying an object with position y at time t has constant acceleration. The third equation, thinking again of y as position, says that y has velocity proportional to t, and we see the second and third equations are almost equivalent (they are equivalent if the initial velocity is 0). The last example arises from modeling the motion of a pendulum, where the acceleration at time t depends on the position y in an oscillatory manner.

The basic mathematical problem in differential equations is to solve for the function y, i.e., determine what are the possible functions y that satisfy our equation. For applications, preliminary to solving a differential equation is finding suitable differential equations to model our problem and understanding what they represent and their limitations. E.g., in first approaches to modeling motion of objects, one might ignore things like friction or the curvature of space-time. However, first solving these simplified models gives us some intuition for what's going on as well as approximate solutions to problem, which are hopefully reasonable on a suitable timescale (or spacetimescale). Then, if needed, we can try to refine our models to account for complications imposed by the real world.

*Department of Mathematics, University of Oklahoma, Norman, OK 73019

Of course one can also have multiple quantities involved, in particular quantities depending on several variables. This is the situation for most interesting problems, leading to the situations such as the famous Korteweg–de Vries (KdV) and Navier–Stokes equations for modelling waves and fluid dynamics, or the Black–Scholes equation for pricings in economics.[1] However, even writing these equations down requires some notions from multivariable calculus. We won't try to develop the theory you would learn at the beginning of a differential equations course, but rather focus on a few simple examples to indicate the utility of differential equations and give a taste of the subject.

1 Population growth without limits

Fibonnaci,[2] in his book *Liber Abaci* from 1202, posed the following problem, which might be the first account of a mathematical approach to "modeling" population growth.

> Say you start one pair of baby rabbits in month 1. Rabbits take 1 month to mature, and each pair of mature rabbits produces a pair of baby rabbits in another month. How many pairs of rabbits do you have the end of 1 year?

Let F_n be the number of pairs of rabbits in month n. We see $F_1 = 1$ (1 pair of baby rabbits), $F_2 = 1$ (1 pair of mature rabbits), $F_3 = 1 + 1 = 2$ (1 pair of mature rabbits + 1 pair of new baby rabbits), $F_4 = 2+1 = 3$ (2 pairs of mature rabbits, and 1 pair of new baby rabbit), $F_5 = 3 + 2 = 5$, and so on. By now, I'm sure you've recognized these numbers, named after our only *dramatis persona* so far, and it's easy to reason out that we get the recursively defined sequence

$$F_1 = 1, F_2 = 1, F_{n+2} = F_{n+1} + F_n \qquad (n \geq 1).$$

Then we see the answer to Fibonnaci's problem is $F_{12} = 144$, i.e., we have 288 rabbits after a year.

Note the Fibonnaci numbers display **exponential growth**: since we always have $F_{n+2} \geq F_n + F_n = 2F_n$, we have, for instance

$$F_{10} \geq 2F_8 \geq 2 \cdot 2F_6 \cdot 2 \cdot 2 \cdot 2 \cdot F_4 = 2^3 \cdot 3 > 2^4.$$

In general, we always get $F_{2n+2} > 2^n$. Similarly, $F_{n+2} \leq F_{n+1} + F_{n+1} = 2F_{n+1}$, so we get $F_n < 2^n$, and one can check that

$$2^{\frac{n}{2}-1} = \frac{1}{2}(\sqrt{2})^n < F_n < 2^n.$$

In fact, for $n > 6$, we have $(\sqrt{2})^n < F_n < 2^n$, so while F_n itself not an exponential function a^n for some $a > 1$, it is bounded between two exponential functions (for $n > 6$), which is

[1] One of the issues that led to the 2008 financial crisis was too many people not understanding the limitations of Black–Scholes type models and what its implicit assumptions are. Rule #5: Understand your model.

[2] The guy who brought the west (hindu-)arabic numerals. Thank him for making us not do calculus in roman numerals. Can you imagine? $\int_{10}^{500}(490x^{510}+10x^{520}+510)\,dx$ would be $\int_{\mathrm{X}}^{\mathrm{D}}(\mathrm{XD}\times x^{\mathrm{DX}}\times \mathrm{D}x^{\mathrm{DXX}}\mathrm{DX})\,dx$. I could not integrate that 5 times fast.

what we mean by exponential growth. (More colloquially, exponential growth just means F_n grows really fast—faster than any polynomial in n for n large.)

This was probably not an attempt by Fibonacci to get a numerically accurate answer for how many rabbits can you get in a year from a single pair of rabbits, but rather a mathematical puzzle motivated by the rapid reproductive capabilities of rabbits, and it shows their potential for exponential growth, and illustrates the surprising (for those unfamiliar) speed of exponential growth.

If you're curious, Wikipedia tells me rabbits get to reproductive age in about 3–4 months, and take about a month to reproduce, typically having 2–12 babies per litter, with a limit of about 4–7 litters per year. A pair of mature rabbits can produce 30–40 children in a year (this does not count grandchildren). So while Fibonacci's model could use some tweaking, it's actually not so bad. A calculation indicates that using these numbers it's reasonable that 1 pair of rabbits could turn into 200 rabbits within a year, assuming no seasonal or other restrictions on reproduction (I assumed each pair of mature rabbits produces about 6 babies every 2 months).

For a more serious population model, one should consider rabbits' lifespan and the effects of aging on reproduction, as well as external influences like limited resources and predators, but for short-term modeling in an ample environment, these other factors will not be significant. More significant will be that rabbits don't actually reproduce at a constant rate—there is some randomness involved—but the hope is to have a model that gives a reasonable rough picture of population dynamics—it would be absurd to expect a deterministic formula that give exact predictions for such a complicated scenario.[3]

While we can compute any given Fibonnaci number exactly, an exact expression for F_n is not entirely obvious. There is a exact formula, discovered in the early 1700's de Moivre:

$$F_n = \frac{\varphi^n - (-\varphi)^{-n}}{\sqrt{5}},$$

where $\varphi = \frac{1+\sqrt{5}}{2} \approx 1.618$ is the golden ratio. However, even with this formula, it's still somewhat complicated to compute F_n, and it's not clear this formula actually makes it any easier to compute F_n.

We can get a nicer expression if we take the following *continuous* model of population growth. Let $y(t)$ be the number of rabbits at time t. Suppose the number of new rabbits at time t is proportional to the number of rabbits at time t, i.e.,

$$y'(t) = cy(t), \tag{5}$$

for some constant c, called the **rate of growth**. Note that this is a bit of a simplification from Fibonacci's model: we don't explicitly take into account the time it takes for rabbits to mature. In Fibonacci's model, the change in y is the number of mature rabbits. However, the number of mature rabbits is approximately a fixed proportion c of the total number of rabbits (at a given time), so the change in y is approximately $cy(t)$ at time t. Precisely, from the formula for F_n, one can compute

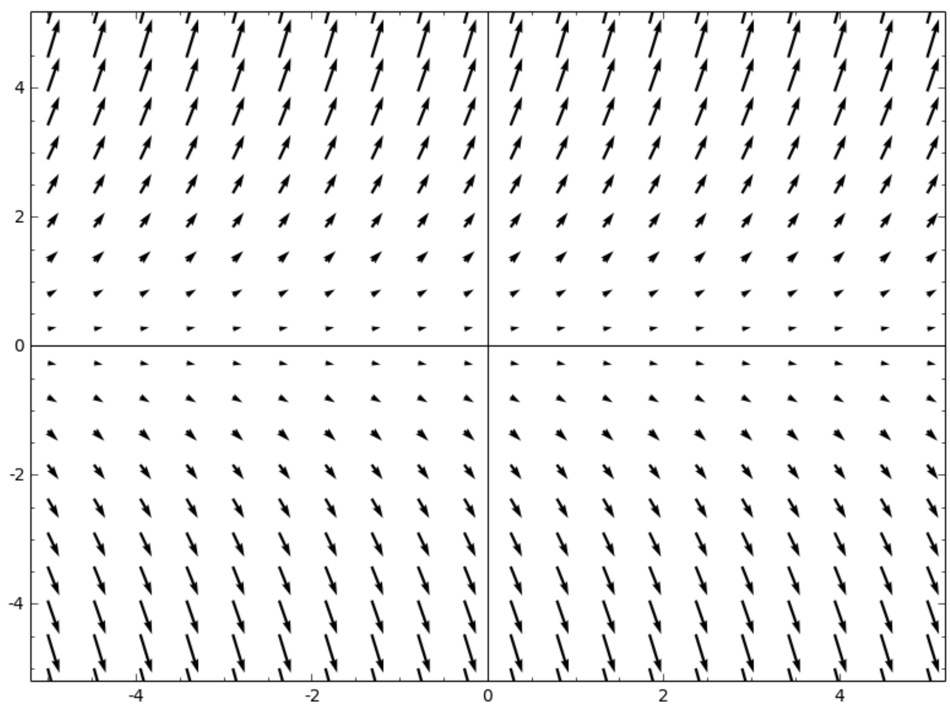
$$\lim_{n\to\infty} \frac{F_{n-1}}{F_n} = \frac{1}{\varphi} = \frac{\sqrt{5}-1}{2} = \varphi - 1 \approx 0.618.$$

[3]Of course many people expect such models to give very precise predictions—often the kind of people who make financial boo-boos or think they won't get cancer if they eat XLVII blueberries a day.

Since $\frac{F_{n-1}}{F_n}$ is the proportion of mature rabbits among all rabbits at month n, we can take $c = \varphi - 1$, and we can check F_n indeed increases at about a rate of $\varphi - 1$ (i.e., $F_{n+1} \approx F_n + (\varphi - 1)F_n = \varphi F_n$) for, say, $n > 5$ (e.g., $\frac{F_6}{F_5} = \frac{8}{5} = 1.6$, $\frac{F_7}{F_6} = \frac{13}{8} = 1.625$, $\frac{F_8}{F_7} = \frac{21}{13} = 1.615...$).

Now let's solve (5). We need to find a function y, whose derivative is a constant multiple of itself. We already know one such function: e^{ct} works. Now you might ask: are there other such functions? Since constants get pulled out of derivatives, we see that Ae^{ct} works for any constant A. On the other hand, if we add a constant to get $Ae^{ct} + B$, then the derivative kills B, so (5) is not satisfied if $B \neq 0$. Similarly, you can check that no nonzero polynomial will satisfy (5), nor will any other exponential function a^{kt} unless $a = e$ and $k = c$. You might try a few other things and fail to find other solutions, but how you do know some other crazy combination of rational functions, trig functions and log or exponential functions won't work?

Here the introduction of a new conceptual ideal will show us why nothing else should work. Given the equation (5), we don't know anything about the values of y *a priori*, so we can't graph y, but we can draw a graph of its derivative y' in terms of y and t. Namely, draw a cartesian plane with the horizontal axis representing t and the vertical axis representing y. Since $y'(t)$ represents the slope of y, it makes sense to "graph" y' in the t-y plane by drawing an arrow at (t, y) with slope $y'(t)$. For $c = 1$, we get the following picture:



Note there is no dependence of y' on t in this case, only on y, so each "column of arrows" is identical, i.e. shifting our picture horizontally doesn't change the picture. This kind of picture is called a **direction field**, because at any value of y and t, it shows you the direction (slope) the graph of $y(t)$ is going.

Consequently, if we know a specific value of y at some point t, say $y = 1$ at $t = 0$, we can draw an approximation to $y(t)$ near $t = 0$, namely a little line segment at $(0, 1)$ with slope $y' = cy = c$. Let's say $c = 2$. Now we follow this line segment for a little bit, say to the point $(0.1, 1.2)$, and look at the direction field there. At this point, we have $y' = 2y = 2.4$, and we can draw another little line segment from $(0.1, 1.2)$ with slope 2.4. Then we follow this a little ways, say to $(0.2, 1.44)$ and then we extend our graph with a little line segment there with slope $y' = 2y = 2.88$. Continuing in this manner, will give an approximation to $y(t)$, and we can make successively better approximations by using shorter and shorter line segments. If we take the limit of this process, we should get a smooth graph which really is our function y, which in this case will be $y(t) = e^{2t}$.

At a crude level, we can say that a solution to (5) is obtained by connecting the arrows in the direction field. On the other hand, any solution to (5) should have a graph which is tangent to the direction field at every point. The point is this limiting process should give a unique graph, i.e., a unique solution $y(t)$ to (5), and this solution only depends upon the choice of *initial value*, which in our case was $y(0) = 1$. Thus I hope I have convinced you of the following:

Theorem 1. *Given the differential equation* (5) *and any initial value condition $y(t_0) = y_0$, there is a unique solution $y(t)$. If we take our initial value condition to be $y(0) = A$, then the solution is just $y(t) = Ae^{ct}$.*

For similar differential equations of the form $y'(t) = F(t, y)$ for some two-variable function F, this hand-wavy argument suggests there is a unique solution given any initial condition $y(t_0) = y_0$—this is essentially true (look up the *Picard–Lindelöf theorem*), but to be correct one needs some conditions on F stronger than just continuity like you might expect.[4] For our particular, differential equation, a proper proof is not hard—it just makes use of a little trick.

Proof. As remarked above, clearly $y(t) = Ae^{ct}$ satisfies (5). It is easy to check that given any initial condition $y(t_0) = y_0$, there exist some A satisfying this. The only part that's not clear is why the solution is unique.

Suppose $y = f(t)$ is a solution to (5). We want to show $f(t)$ is a constant multiple of e^{ct}, that is, we want to show the auxiliary function $g(t) = f(t)e^{-ct}$ is constant. (Looking at $g(t)$ and taking its derivative is the trick.) By the product rule, we have

$$g'(t) = f'(t)e^{-ct} - cf(t)e^{-ct} = (f'(t) - cf(t))e^{-ct} = 0,$$

using (5) at the last step. As we know from Calc I, this means $g(t)$ must be constant. Hence $y(t) = Ae^{ct}$ for some A, and it is easy to see the initial value condition determines A uniquely. □

In particular, this theorem says that $f(x) = e^x$ is the unique function $f(x)$ such that $f' = f$ and $f(0) = 1$. Sometimes this is taken as the *definition* of e^x.

[4] Many people move their hands around trying to persuade others: Obi Wan Kenobi, David Copperfield, that guy selling "new underwear" in a gas station parking lot. Rule #9: Whenever someone starts waving their hands in the air, it's always a con.

Example 1

From our above discussion, we can look at a continuous analogue of the Fibonnaci problem by considering the differential equation (5) with $c = \varphi - 1 \approx 0.618$, so $y(t) = Ae^{ct}$ for some A. Then, with an appropriate initial value condition $y(t_0) = y_0$, the number of pairs of rabbits after a year in Fibonacci's problem should be approximately $y(12)$. The simplest thing to do would be to take $y(1) = 1$—Fibonacci's first initial condition. This condition means $Ae^c = 1$, i.e., $A = e^{-c} \approx 0.539$. But this yields $y(12) = e^{-c}e^{12c} = e^{11c} \approx 896$. The problem is, as we remarked above, our simplified continuous approximation is only a decent approximation for, say $n > 5$.

Instead, let's take for our initial condition $y(6) = F_6 = 8$, so $Ae^{6c} = 8$, i.e., $A \approx 0.196$. This yield $y(12) \approx 326$. This is still not so close to $F_{12} = 144$. To get a better sense of what's going on, let's compare approximate values of $y(n)$ and F_n:

n	1	2	3	4	5	6	7	8	9	10	11	12
F_n	1	1	2	3	5	8	13	21	34	55	89	144
$\approx y(n)$	0.36	0.67	1.2	2.3	4.3	8	14.8	27.5	51.0	94.7	175.8	326.2

We see that the values of $y(n)$ are pretty close to F_n when n is close to 6, but as n gets further and further away, $y(n)$ and F_n get further and further apart. So quantitatively, $y(n)$ is an okay model for F_n for n close to our initial condition, but gets bad quite quickly. This is because these are both functions of exponential growth and small differences become quite big when exponentiated. This also illustrates the phenomenon known as *sensitivity to initial conditions*—that small changes in initial conditions (i.e., the choice of the value of A) may result in huge differences after a moderate amount of time. So our simplified continuous model (5) is only a good approximation to the Fibonacci problem on a reasonably small timescale. (This actually isn't a terrible drawback, because Fibonacci's setup is also only reasonable for small timescales.) For the same reason, weather simulation models provide useful predictions for a couple of days into the future, but are useless as indicators of what will happen weeks from now.[5]

While one could argue that the continuous exponential model for population growth is not as good in some ways as Fibonacci's discrete model, they are qualitatively not so different—they both exhibit exponential growth, so their graphs will be about the same shape. Moreover, the continuous exponential is generally easier to work with and analyze, so if we want to revise our models to incorporate other factors like death and aging or environmental constraints, it's probably better to try this in the continuous setting of differential equations. (The discrete analogue of differential equations are called *difference equations*.) This is part of a more general principle: continuous processes are often easier to analyze than discrete ones, so in modelling physical, economic or social phenomena, one looks for a continuous approximation so one can apply calculus. A somewhat more sophisticated model incorporating environmental constraints is presented in the exercises.

It's also natural to consider *systems of differential equations*, where one has several vari-

[5] Rule #23: Whenever you hear: "Quantity X is expected to reach amount Y by the year Z," don't take this literally. Unless it's about the level of zombie outbreaks in 2020. That's an exact science.

ables and several differential equations to solve simultaneously. With systems of differential equations, one can model populations of several different species and how they interact. For instance, there is the famous **predator-prey model** of Lotka–Volterra proposed by Lotka in 1910 (in the context of chemical reactions, and later a plant and animal species) and independently by Volterra in 1926 (who went straight for predators—kill! kill! kill!):

$$x' = ax - by$$
$$y' = cx + dy.$$

Here x and y represent populations where the y-population is the predator and the x-population is the prey. The first equation says x has some natural growth rate a but the larger y is the slower the growth rate will be, and the second equation say y has some natural growth rate c, but will grow at a faster rate in the presence of more prey x. These equations are also used in economics. The general solutions to the predator-prey equations are not elementary functions your are familiar with, but can still be analyzed.

2 Compound interest

A completely unrelated context providing similar examples of differential equations is modeling compound interest.

Example 2

Suppose you put \$10,000 in a 24-month bank CD that has a *nominal* annual interest rate of 1%. (By nominal interest, I mean the amount of interest on the original deposit—not accounting interest on interest.) How much money you actually have at the end of 24 months depends upon how often interest is applied. If the interest is only compounded (applied) yearly, after year 1, you have \$10,100, and after year 2 you have \$10,201. So here compound interest gets you an extra dollar in year 2, because you earn interest on interest.

On the other hand, if the interest is compounded monthly (which is pretty common), you will get more money. Compounding monthly with the above nominal interest rate means you get $\frac{1}{12}$% interest each month. Precisely, if $f(n)$ be the amount of money in your CD account after n months, then

$$f(n+1) = (1 + \frac{1}{1200})f(n).$$

So we see

$$f(n) = (1 + \frac{1}{1200})f(n-1) = (1 + \frac{1}{1200})^2 f(n-2) = \cdots = (1 + \frac{1}{1200})^n f(0)$$
$$= (1 + \frac{1}{1200})^n \times \$10,000.$$

(Okay, here I'm assuming the bank carries over fractions of cents in your account—but it might round down, cheating you out of invaluable fractions of pennies.) That means that after a year, you have $f(12) = \$10,100.45$ in your account and after 2 years, $f(24) = \$10,201.92$ in your account. So compounding interest monthly gives you almost twice the compound interest over 2 years. Still, it's not even \$2, so who cares.

The reason this isn't so impressive is that I wanted to use realistic numbers and interest rates are very low now. At one point, interest rates were around 5%. With this nominal rate, compounding annually leaves you with \$10,500 after 1 year and \$11,025 after 2 years. On the other hand, compounding monthly leaves you with \$10,511.61 after 1 year and \$11,049.41 after 2 years, almost a \$25 difference, which is at least something. However, the compound interest will be much more significant in the next example.

Going back to our 1% interest rate, what if if the bank compounded daily, or every minute, or every second? Then you'll get even more compound interest (well, relatively a lot more). Rather than doing these computations, let's compare with compounding *continuously*. For simplicity, let's measure t in year rather than months now. Continuously compounding at a nominalized annual rate of 1% means

$$f'(t) = 0.01f(t)$$

which is just the same as (5) above with $c = 0.01$. Here our initial condition is $f(0) = 10000$. Note at $t = 0$, we get $f'(0) = 100$, which is indeed the amount of nominal interest earned in a year. I.e., the tangent slope at 0 is the annual nominal interest rate, so the height of the tangent line at $t = 1$ is the amount of money in your account after one year counting only nominal interest. From Theorem 1, we see $f(t) = 10000e^{0.01t}$ so after 1 year we have $f(1) \approx \$10{,}100.50$ and after 2 years, $f(2) \approx \$10{,}202.01$.

We see that to determine how much money you'll earn in different accounts means not just knowing the nominal interest rate, but also how often interest gets compounded. The amount of interest you'd earn in a year (taking into account compounding) is called the annual percentage yield, or APY, and is what banks advertise.

Example 3

Let's say you get a job and are able to start saving for the future. Maybe but you can afford to invest about \$1,000 per year. Since you're a rational investor, you invest in an index fund, maybe something like the S&P 500 index. (You might want to diversify at some point down the road, and maybe you'll have more money to invest later, but let's not worry about that here.) Over the last 10 years, the annual return averaged about 7%, but around 10% over the last 20 or 30 years. Let's assume an average annual return of 10%. Using a continuous model, we can estimate your investment funds after t years by $f(t)$, where $f(t)$ satisfies the differential equation

$$f'(t) = 0.1f(t) + 1000, \qquad f(0) = 0. \tag{6}$$

If we didn't have the +1000 (the amount you're actively contributing to your investments each year), this would just be (5) again and $e^{0.1t}$ would be a solution. Now you might think to start with $e^{0.1t}$ and tweak it. Your first thought might be to try adding a factor. E.g., if $y(t) = e^{0.1t} + g(t)$, then

$$y'(t) = 0.1e^{0.1t} + g'(t) = 0.1(e^{0.1t} + g(t) - g(t)) + g'(t) = 0.1y(t) + g'(t) - 0.1g(t).$$

So we want $g(t)$ to satisfy $g'(t) - 0.1g(t) = 1000$, which means g satisfies the same differential equation as f (though not necessarily the same initial value condition). If you try the same thing with g, then you'll be stuck in the same situation, so maybe we need another idea. Well, the above at least tells us that if f is $e^{0.1t}$ plus something, that something might also contain $e^{0.1t}$ in it.

This suggests it might be better to tweak $e^{0.1t}$ by *multiplying* it by another factor. Namely, suppose $f(t) = e^{0.1t}g(t)$ for some $g(t)$. Then, by the product rule,

$$f'(t) = 0.1e^{0.1t}g(t) + e^{0.1t}g'(t) = 0.1f(t) + e^{0.1t}g'(t).$$

This means we need $e^{0.1t}g'(t) = 1000$, i.e. $g'(t) = 1000e^{-0.1t}$, i.e.,

$$g(t) = \int 1000e^{-0.1t}\,dt = -10000e^{-0.1t} + C,$$

so

$$f(t) = e^{0.1t}(C - 10000e^{-0.1t}) = Ce^{0.1t} - 10000 = 10000(e^{0.1t} - 1),$$

where we solved for C using the initial value $f(0) = 0$. In hindsight we can see that the product rule means $f(t) = e^{0.1t}g(t)$ was a good idea. In this particular case, we also see that guessing a solution of the form $f(t) = Ae^{0.1t} + B$ (or $f(t) = Ae^{0.1t} + g(t)$) would've worked, but I wanted to illustrate the use of the product rule as this technique comes up a lot in differential equations.

This leads to the following estimates

t years	cont. model	disc. model	amount invested
1	\$1,051	\$1,000	\$1,000
5	\$6,487	\$6,105	\$5,000
10	\$17,182	\$15,937	\$10,000
15	\$34,816	\$31,772	\$15,000
20	\$63,890	\$57,274	\$20,000
25	\$111,824	\$98,347	\$25,000
30	\$190,855	\$164,494	\$30,000

The second column is $f(t)$ and the third column comes from a discrete model of compounding at a rate of 10% annually. Note the discrete model fails to take into account interest on the added \$1,000 during the year it is deposited, which is not accurate if you are gradually adding it over the year (but would be if you deposit the \$1,000 as a lump sum at the end of the year). On the other hand, if you gradually deposit the \$1,000 each year, this is automatically taken into account in the continuous model by virtue of continuous compounding.

This shows the power of compound interest and the value of investing early. Of course, these numbers are just estimates to give you a rough idea of what you might have after investing *assuming* markets behave comparably in the next 30 years as they have in the past 30 years. (I also didn't account for any investment fund fees, though these are usually pretty small for index funds, or taxes, but these can be avoided if you use a Roth IRA.)

3 Mechanics

We've already seen a very simple kind of differential equation in motion problems, along the following lines. Suppose an object X, initially at rest, falls to the ground from some initial height h at time 0. If $y(t)$ is the distance traveled by time t, then $f(t)$ satisfies

$$y'' = g, \quad y(0) = 0, \quad y'(0) = 0. \tag{7}$$

where g is the acceleration due to gravity (about $9.8m/s^2$ or $32ft/s^2$). Without the initial conditions, integrating twice show the solutions are precisely the functions of the form $y(t) = \frac{g}{2}t^2 + C_1 t + C_2$. Knowing only the initial condition $y(0) = 0$ just says $C_2 = 0$, but the added initial condition $y'(0) = 0$ allows us to conclude $y(t) = \frac{g}{2}t^2$. Of course (7) is only a valid model until you hit the ground, i.e., until $y(t) = h$, i.e., for $0 \le t \le \sqrt{\frac{2h}{g}}$.

In general, if we have a differential equation involving the n-th derivative but no higher derivatives, called an **n-th order differential equation**), we need more than one initial condition to determine y. (Note all of our previous examples were first-order differential equations, so we just needed one initial condition.)

For instance, if we just have the simple equation

$$y^{(n)} = F(t),$$

then need to integrate n times to get y, and therefore there are n constants involved in y. For each initial condition we have, we can solve for one of the constants (at least in terms of the others), and we need n initial conditions to determine y uniquely. There are many ways to choose the initial conditions, for example in (7) we could have two conditions in terms of $y(t)$, such $y(0) = y(1) = 100$ (this implies the object has some initial velocity upward so and gravity pulls it back to the same position by time $t = 1$) or a condition on y and one on y' at different times, such as $y(10) = y'(0) = 0$. However we can't make both initial conditions in terms of y', because y' doesn't see C_2.

Recall from Newtonian mechanics $F = ma$, where F denotes force, m the mass of an object, and a its acceleration. So if a force acts on an object, it's acceleration (due to that force) will be $\frac{F}{m}$. Consequently, problems in mechanics often lead to second-order differential equations, because one has some forces acting and one relates them to acceleration.

Now let's look at some slightly more sophisticated examples from mechanics.

First we consider the free fall of an object X mentioned above, but now with air resistance.

Example 4

Suppose an object X initially at rest is dropped from height h. Let $y(t)$ be the distance traveled at time t. It's acceleration due to gravity is (approximately) the gravitational constant g. One model for the decelerating force due to air resistance a constant time the velocity squared. Hence we can consider the model

$$y'' = g - k(y')^2, \quad y(0) = 0, \quad y'(0) = 0. \tag{8}$$

Here k is a constant that depends on the shape (drag and area), mass and density of the object, as well as the density of the air. (Of course the density of air changes with altitude—and in fact the force of gravity does too—but this should be reasonably accurate for relatively short distance drops.)

To solve this, we can express this (without initial conditions) in terms of a first-differential equation

$$u' = \frac{du}{dt} = g - ku^2, \quad u = y'. \tag{9}$$

Up until now, all of our examples of (first-order) equations we've seen were **linear differential equations**, i.e., of the form $f' = a(t)f + b(t)$. This equation is **nonlinear** because of the presence of the u^2 term. As with usual equations, nonlinear differential equations are harder to solve than linear ones, and not solvable in general (at least not in an elementary way, but see the remarks on series at the end). However, we can solve certain classes of non-linear equations.

This differential equation (9) is an example of a **separable equation** (one of the form $f' = p(t)q(f)$ for some functions p and q), and we can solve such equations in the following manner (which I won't justify now, but trust me, it works [hand wave]). Treating du and dt as separate objects like we do in integration by parts, we can write

$$\frac{du}{g - ku^2} = dt.$$

Integrating both sides gives

$$\int \frac{du}{g - ku^2} = \int dt = t + C.$$

Now we know how to compute the left-hand integral. (Finally, we can make use apply our integration techniques to differential equations!) Note the denominator is $k(\frac{g}{k} - u^2)$, so we put $u = \sqrt{\frac{g}{k}}x$. Then

$$\begin{aligned}\int \frac{du}{g - ku^2} &= \sqrt{\frac{g}{k}} \int \frac{dx}{g(1 - x^2)} = \frac{1}{\sqrt{gk}} \int \frac{dx}{1 - x^2} \\ &= \frac{1}{\sqrt{gk}} \int \frac{1}{2}\left(\frac{1}{1 - x} + \frac{1}{1 + x}\right) dx = \frac{1}{2\sqrt{gk}} \ln\left|\frac{1 + x}{1 - x}\right| + C.\end{aligned}$$

Let's leave things in terms of $x = x(t)$ for now for simplicity, so our differential equation says

$$\frac{1}{2\sqrt{gk}} \ln\left|\frac{1 + x}{1 - x}\right| = t + C.$$

The initial condition $y'(0) = 0$ means $u(0) = x(0) = 0$, which means $C = 0$. So multiplying by $2\sqrt{gk}$ and exponentiating gives

$$e^{2\sqrt{gk}t} = \frac{1+x}{1-x} = \frac{-(1-x)+2}{1-x} = \frac{2}{1-x} - 1.$$

(To drop absolute values, one should check $\frac{1+x}{1-x} > 0$, but this can be justified *post facto* when we find x. There is also a solution with $\frac{1+x}{1-x} < 0$ if one drops the initial condition $x(0) = 0$, but for our physical model, we should have $x < 1$ when t is small, which means we should have $\frac{1+x}{1-x} > 0$.) Adding 1 from both sides and taking reciprocals leads to

$$x(t) = 1 - \frac{2}{e^{2\sqrt{gk}t} + 1},$$

i.e., the velocity $u = y'$ is given by

$$u(t) = \sqrt{\frac{g}{k}} x(t) = \sqrt{\frac{g}{k}} \left(1 - \frac{2}{e^{2\sqrt{gk}t} + 1}\right).$$

Indeed this satisfies $u(0) = 0$, and we note that

$$\lim_{t\to\infty} u(t) = \sqrt{\frac{g}{k}},$$

so there is a "maximum" (technically supremum, since it's never quite attained) velocity, which is better known as *terminal velocity*.

As you may recall, we can compute $\int \frac{dt}{e^{ct}+1}$ with a substitution. Then we can calculate the distance traveled as

$$\begin{aligned} y(t) = \int u(t)\,dt &= \sqrt{\frac{g}{k}} \left(t - 2\left(t - \frac{\ln(e^{2\sqrt{gk}t} + 1)}{2\sqrt{gk}}\right)\right) + C \\ &= \frac{1}{k}\ln(1 + e^{2\sqrt{gk}t}) - \sqrt{\frac{g}{k}}t - \frac{\ln 2}{k}, \end{aligned}$$

using the initial condition $y(0) = 0$ to solve for C. (Again, this model is only good until impact, i.e., when $0 \le y \le h$.)

For a general separable equation $y' = p(t)q(y)$, the analogous technique would be to rewrite the equation as

$$\frac{dy}{q(y)} = p(t)\,dt, \tag{10}$$

integrate both sides (where the left is dy and the right is dt):

$$\int \frac{dy}{q(y)} = \int p(t)\,dt \tag{11}$$

and solve for y. This method is known as **separation of variables**. While our manipulation of dy and dt was purely formal pushing around of symbols and not something mathematically justified, it is not hard to check this works: Let $F(t)$ be the left hand side of (11). Then by the chain rule and FTC, differentiating both sides of (11) gives

$$F'(t) = \frac{dF}{dy}\frac{dy}{dt} = \frac{1}{q(y)}y' = p(t),$$

which is equivalent to our original differential equation (provided $q(y) \neq 0$).

Note we can also apply this to (5). Namely, if $\frac{dy}{dt} = y' = cy$, we rewrite this as

$$\frac{dy}{y} = c\,dt.$$

Then we integrate both sides

$$\int \frac{dy}{y} = \int c\,dt$$

to get

$$\ln|y| = ct + C.$$

Exponentiating yields

$$|y| = e^{\ln y} = e^{ct+C} = e^C e^{ct}.$$

Putting $A = \pm e^C$, we get back the solution

$$y = Ae^{ct}.$$

Note $e^C > 0$ for any C, whereas A is allowed to be positive or negative, and one can check these are all solutions. There is one minor technicality—we didn't recover the case of $A = 0$, i.e., $y = 0$, which is also a solution—this is because we inherently assumed $y \neq 0$ when we wrote $\frac{dy}{y}$.

Our previous examples of differential equations were so simple that we didn't really need to know any integration techniques to solve them (the solution in Example 3 was not obvious, but it did not require any sophisticated integration—note one can also do Example 3 using separation of variables). However, the above example shows how being able to compute integrals is important in differential equations.

Here's another example with gravity—now in outer space!

Example 5

Say you have two bodies X_1 and X_2 in space with masses m_1 and m_2, far away from other bodies which are relatively at rest with respect to each other. Let $y = f(t)$ be their distance at time t. In Netwon's classical theory of gravitation, they will be attracted with force $F = G\frac{m_1 m_2}{y^2}$, where G is the universal gravitational constant. So the acceleration of X_1 will be $\frac{F}{m_1} = G\frac{m_2}{y^2}$ and the acceleration of X_2 will be $\frac{F}{m_2} = G\frac{m_1}{y^2}$. So this leads to the model

$$y'' = -\left(\frac{F}{m_1} + \frac{F}{m_2}\right) = -G\frac{m_1 + m_2}{y^2},$$

or equivalently

$$y^2 y'' = c,$$

where $c = -G(m_1 + m_2)$. I won't solve this, but just remark $y(t) = C\sqrt{t}$ is a solution if $c = -\frac{1}{4}$.

The above example is a special case of the n**-body problem**—how gravitation affects the movement of n bodies in space. The two-body problem is solvable, but the three-body problem (or the problem for $n > 3$) has elementary solution in general. To properly describe the n-body problem (or even the general setup of the two-body problem) requires multivariable calculus.

Here are a couple of other classical examples of differential equations, one with elementary solutions and one without.

Example 6

Suppose you hang an object from a spring, stretch it and release. Let $y(t)$ denote the position at time t. The simple model for motion here is

$$y'' = -ky,$$

where $k > 0$ is a constant associated to the spring. Clearly $\sin\sqrt{k}t$ and $\cos\sqrt{k}t$ are solutions, and the motion will be oscillatory. In reality, there are damping forces (resistance of material in the spring, etc) that slow down the oscillation and hinder perpetual motion. One can take into account damping with a model of the form

$$y'' = -ky - cy',$$

where c is a damping constant. This has a solution of the form an exponential function (with negative exponent, which represents the damping) times a certain combination of a sine and cosine function.

Example 7

One can model the motion of a pendulum (without damping) via the equation

$$y'' = \frac{k}{\sin t}$$

where k is a constant associated to the system. This is another separable equation, but there are no elementary solutions. Instead solutions will involve the so-called *elliptic integrals* (which you may remember came up in the arc length of ellipses and hyperbolas).

At some point, you realize that most of the differential equations you want to solve for some modeling problem don't have elementary solutions. At this point, there are a couple of things to do.

One is to start becoming comfortable with *defined* in terms of definite integrals (or by their differential equation). One famous example, besides elliptic integrals (but closely related) are the *J-Bessel functions* (or Bessel functions of the first kind)

$$J_n(t) = \frac{1}{\pi}\int_0^\pi \cos(nx - t\sin t)\,dx$$

which satisfies the differential equation

$$t^2y'' + ty' + (t^2 - n^2)y = 0.$$

There are also Y-Bessel functions Y_n, or Bessel functions of the second kind, which are different solutions to the same differential equation, and can be defined by different integrals. Bessel functions of the first kind go to 0 as $t \to \infty$, where as Bessel functions of the second kind go to ∞ as $t \to \infty$, so we see different functions satisfying the same differential equation can have very different behaviours. Bessel functions come up all over the place in physics and engineering.

Another example of a function defined by an integral is the *Gamma function*:

$$\Gamma(t) = \int_0^\infty x^{t-1}e^{-x}\,dx,$$

which converges for $t > 0$. The Gamma function does not come from any simple differential equation, as far as we know. Still, it's pretty cool: it interpolates the factorial function as $\Gamma(n+1) = n!$ for an integer $n \geq 0$ and in general satisfies the functional equation $\Gamma(t+1) = t\Gamma(t)$. So using this you can "differentiate factorials." Another fun fact: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. The Gamma function is important in physics, probability, combinatorics and number theory. There are also some relations of the Gamma function with elliptic integrals, and at least for certain ellipses, one can express the arc length in terms of values of the Gamma function.

The point is, at first, you may think defining functions in terms of integrals is some weird or unnatural thing. But a lot of phenomena you encounter in nature will be new, and not expressible in terms of any functions you already know, and you have to define them in terms of an integral or some other process. However, naming these functions makes you feel like you have some control over them (like knowing someone's name gives you power over them), and once you start using them enough, you begin to think of them as some basic functions on their own, like sine and cosine. Really, if you think about it, to what extent do you understand sine and cosine? You know their geometric definition, some properties they satisfy, what their graphs look like and a few special values. But if I ask you what is sin(1)—what the hell is that? It's some random irrational number whose only claim to fame is that it is the value of sine at 1. Nevertheless, you can approximate it. In this sense, one can understand things like Bessel functions and Gamma functions in a similar, but more complicated, manner. (Note defining functions in terms of integrals is a geometric definition—one in terms of areas.)

The other thing people do with differential equations you can't solve is work on computing numerical solutions (i.e., functions that approximate solutions to the differential equations). In fact, you probably can solve your differential equation, just not in the sense you were thinking. In Calc III, you learn about representing functions by power series: which is sort of analogous to writing a function as an integral—you write it as an infinite sum (which is a limit, sort of like a Riemann sum), e.g.,

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots.$$

(Check that if you formally take the term-by-term derivative of the infinite sum on the right, you get the back the same infinite sum, which reflects the fact that $\frac{d}{dx}e^x = e^x$.) You can also think of this as approximating a function by a polynomial of arbitrarily large degree. Any "reasonable" function can be represented as a power series.[6] Similarly, you can solve any reasonable differential equation in terms of power series, but then you may not be able to translate the solution back in terms of functions you know. Nevertheless, you can compute a bunch of terms (say up to the x^{20} term), which will be a good approximation of e^x if x is not too large. This provides one way to find a numerical solution to $y' = y$ for t not too big. Another method (also only good for t not too large) is to use the more elementary approach of direction fields that we explained earlier. The main problem in numerical analysis is how to get good approximations for t large.

Exercises

Exercise 1. Let's say you amass \$1,000 in credit card loans and just pay back the minimum amount \$25 every month, but never use this card again. The annual interest rate (APR, not APY, meaning nominal interest) the company charges you is 15%, which, since they're assholes, they compound daily.

(a) Write down a differential equation with an initial condition for a continuous model for your amount of credit card debt.

(b) Solve your equation from (a). (You can either guess a solution in a certain form as in Example 3 or use separation of variables.)

(c) Using this, estimate the amount of time until your loan is paid off.

(d) From (c), estimate the total amount you will have paid on your loan.

Exercise 2. Check the details of Example 4—specifically justify (a) the dropping of absolute values (i.e., check the solution $x(t)$ satisfies $\frac{1+x}{1-x} > 0$, or if you prefer, check u satisfies $u' = g - ku^2$), and (b) determining y from u, as I omitted the details for that integral. (Though you should go through all the details on your own to make sure you understand it completely.)

Exercise 3 (The logistic model)**.** For long-term analysis, a better way to model population growth than (5) is to impose a maximum capacity K for the population and force the rate of growth y' of the population to go to 0 as the population y approaches K. This idea led Verhulst to the following logistic differential equation in 1838,

$$y'(t) = cy(t)(1 - \frac{y(t)}{K}). \tag{12}$$

Note the factor $1 - \frac{y(t)}{K}$ is that it is close to 1 when y is small and close to 0 when y is close to K. Using the method of separation of variables as in Example 4 (cf. (10), (11)), find a solution to (12).

[6] Actually, my definition of reasonable is something having a power series expansion. I know, how perverted.