Chapter One

Ordinary Differential Equations of Order One

1. Introduction to Ordinary Differential Equations

The first question one might ask in a course in differential equations is ”what are differential equations and where do they come from?” To begin to answer that question, consider the following example. A cup of hot coffee is brought into a room and set on the table. We know from experience that the coffee will begin to cool and continue cooling until the temperature of the coffee is reduced to the temperature of the room. We are going to appeal to known physical laws, namely Newton’s law of cooling, and translate the application of that law to the cooling cup of coffee into mathematical terms. This mathematical expression of Newton’s law as it applies to the coffee will be a differential equation whose solution will allow us to predict the temperature of the coffee at every time during the process of cooling.

In other words, a differential equation often arises when we translate into mathematical terms the physical laws that describe the behavior of some physical system. Solving the differential equation allows us to understand and to predict the behavior of the physical system. Of course not all differential equations arise in this way, but in any case the solution of a differential equation is not just a number as is the case when we solve an algebraic equation. The solution of a differential equation is a function having the property that when it is substituted into the differential equation, the equation is identically satisfied for all values of the independent variable. But let us demonstrate what all this means with some concrete examples.

1.1 A Simple Mathematical Model

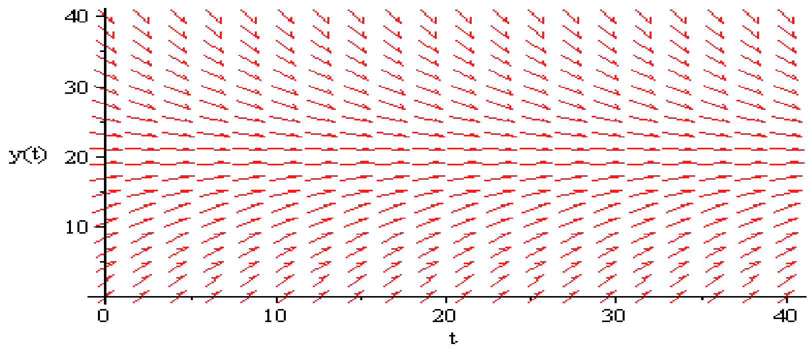
To describe the situation of the cooling cup of coffee analytically, we first define a function $T(t)$, whose value at each positive value of t is the temperature of the coffee at time t. We make the simplifying assumption that the temperature is the same at every point in the cup;i.e., that $T(t)$ depends only on t. We define $t = 0$ to be the initial time at which the coffee was set on the table and we assume t increases from zero. The time t is referred to as the **independent variable** in the problem, and $T(t)$ is the **dependent variable** (because T depends on t). If we denote the temperature of the room to be the constant T_R, (another simplifying assumption) then we can begin an analytic description of the cooling process for the coffee. For example, we can say that:

- if $T(t) > T_R$ then $T(t)$ *decreases* $\left(i.e.,\ T'(t) < 0\right)$
- if $T(t) < T_R$ then $T(t)$ *increases* $\left(i.e.,\ T'(t) > 0\right)$

Then we might be inclined to assume that $T'(t)$ and $(T(t) - T_R)$ are related, in fact, we might assume that $T'(t)$ is *proportional to* $(T(t) - T_R)$; i.e., the temperature changes rapidly when the difference between the temperature of the coffee and the temperature of the room is large, and it changes more slowly as this difference decreases. This would be expressed as follows

$$T'(t) = -k(T(t) - T_R) \qquad (1)$$

where k denotes a positive constant of proportionality and the negative sign appears so that $T'(t)$ is negative when $(T(t) - T_R)$ is positive in accordance with our previous observations. Then (1) is an example of a **differential equation**. It is called this because it is an equation involving an unknown function $T(t)$ and also its derivative $T'(t)$. More precisely, (1) is called an **ordinary differential equation** because the derivative which appears is an ordinary derivative (as opposed to a partial derivative which occurs when the unknown function depends on more than one independent variable). In addition, we say that (1) is a **first order** equation because there are no derivatives of order more than one, and finally we say that (1) is a **linear** first order ordinary differential equation (but we will wait until later to explain the meaning of the term, ”linear”).



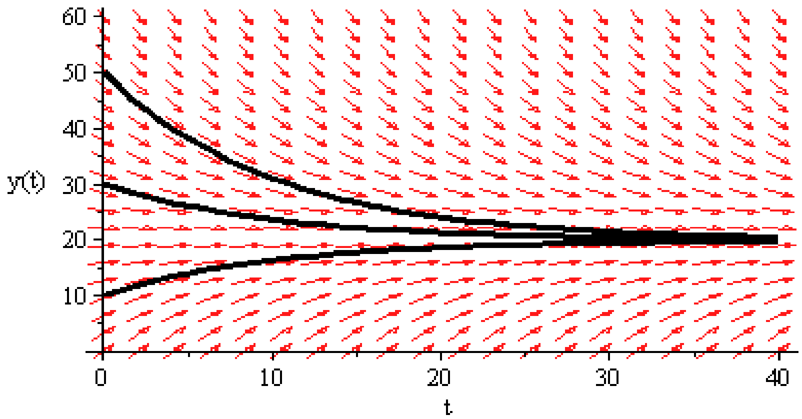
Direction Field

We can now think of the equation (1) as a mathematical model for the physical process of a cooling cup of coffee.

1.2 The Notion of a Solution

Once we have the mathematical model in the form of a differential equation, it will be our aim to solve the equation and to use the solution to predict and understand the physical system. Of course we need to understand what is meant by a solution for the equation. One meaning for the term solution might be a ”**direction field**” for the equation (1). Note that the equation defines a value of the derivative $T'(t)$ at each point in the T-t plane. Then at each point (t, T) in the plane, we could imagine a small segment whose slope equals $T'(t)$. A picture of this would look like the figure above.

The value of T_R in this plot is $T_R = 20$ and we can see from the plot that temperature of the coffee tends toward this limiting value T_R. If we define a **solution curve** for the equation to be a curve in the T-t plane that is tangent to the direction field at each point on the curve, then it is clear from the plot above that there are infinitely many such curves. Here is a picture of just three possible solution curves



Three Solution Curves

Evidently the differential equation (1) has an infinite family of solution curves, each of which is a possible description of how the temperature of the coffee varies with time. Notice that the three solution curves in the figure above correspond to different "initial values". That is, the uppermost curve starts from the value 50 at $t = 0$, while the other two curves begin from the values 30 and 10, respectively. Since 50 and 30 are greater than $T_R = 20$, these curves indicate that the temperatures decrease gradually to T_R, while the curve that begins from the value 10 gradually increases to T_R.

To better understand what this means, we can try to find the equations for the solution curves by doing the following. We can write (1) in the form

$$\frac{d}{dt}T(t) = -k(T(t) - T_R)$$

and then rewrite it as

$$\int \frac{dT}{T(t) - T_R} = -k \int dt.$$

Since

$$\int dt = t + C_0$$

and $$\int \frac{dT}{T(t) - T_R} = \ln(T(t) - T_R) + C_1,$$

we find

$$\ln(T(t) - T_R) + C_1 = -k(t + C_0)$$

or, equivalently

$$\ln(T(t) - T_R) = -kt + C_2.$$

Here C_0, C_1, C_2 all are arbitrary constants (you can solve for C_2 in terms of C_0, C_1if you want to but C_2 is still just an arbitrary constant). What we have at this point is an implicit functional relation between T and t. In some cases it will not be possible to resolve this expression to obtain an explicit formula for T in terms of t, but here we note that

$$\begin{aligned} e^{\ln(T(t)-T_R)} &= e^{-kt+C_2} \\ &= e^{-kt}e^{C_2} \\ &= C_3e^{-kt} \end{aligned}$$

and since $e^{\ln(T(t)-T_R)} = T(t) - T_R$, we obtain,

$$T(t) = T_R + C_3 e^{-kt}. \qquad (2)$$

The constant C_3 is another arbitrary constant of integration and for each choice of this constant, (2) is a solution of the equation (1). It is also true that for each choice of the constant, (2) is the equation of one of the solution curves for (1). We refer to (2) as a 1-parameter family of solutions for (1) where C_3 is the parameter that distinguishes between the infinitely many family members. Evidently an equation like (1) has an infinite family of solutions.

Finally, suppose that the initial temperature of the coffee is known; i.e., suppose $T(0) = T_0 > T_R$ where T_0 is given. Then, according to (2)

$$\begin{aligned} T(0) &= T_R + C_3 e^{-k0} \\ &= T_R + C_3 \\ &= T_0, \end{aligned}$$

hence $C_3 = T_0 - T_R$, and

$$T(t) = T_R + (T_0 - T_R)e^{-kt}. \qquad (3)$$

This is the unique solution that satisfies (1) as well as the **initial condition**, $T(0) = T_0$. Evidently, the problem consisting of the differential equation (1) and the initial condition has a unique solution. In the picture of solution curves seen above, the upper curve is the solution curve corresponding to $T_0 = 50$, while the lower curves correspond to $T_0 = 30$ and $T_0 = 10$, respectively. Here $T_R = 20$ so the upper two curves decrease toward T_R while the lowest curve increases asymptotically towards T_R.

Once the solution is known, then it may be used to analyze the physical system. For example, if the initial temperature of the coffee was unknown but the temperature of the coffee at some later time, say $t = 2$, is known to be equal to $T(2) = T_2$, then the initial temperature can be found. If we use (2) to write

$$T(2) = T_R + C_3 e^{-2k} = T_2,$$

then

$$C_3 e^{-2k} = T_2 - T_R$$

and

$$C_3 = (T_2 - T_R)e^{2k}.$$

Now we use this expression for C_3 in (2) to obtain

$$\begin{aligned} T(t) &= T_R + (T_2 - T_R)e^{2k}e^{-kt} \\ &= T_R + (T_2 - T_R)e^{-k(t-2)} \end{aligned}$$

from which it follows that the initial temperature of the coffee is equal to,

$$T(0) = T_R + (T_2 - T_R)e^{2k}.$$

We are assuming here that the constant k is a given positive constant.

Exercises-

1. If the constant k is unknown, but it is given that $T(0) = T_0$, $T(2) = T_2 < T_0$, then use this information to find k in terms of T_R, T_0, and T_2 .
2. If T_R, T_0, and k are known, with $T_0 > T_R$, find the time $t_* > 0$ at which $T(t_*) = T_R + T_0/2$. (t_*

will be expressed in terms of $T_R, T_0,$ and k.)

3. If $T_R, T_0,$ and k are known, with $T_0 > T_R,$ can you find a time $t_* > 0$ at which $T(t_*) = T_R$? Can you find a time $t_* > 0$ at which $T(t_*) < T_R$? For any $\varepsilon > 0$ can you find a time $t_* > 0$ at which $T(t_*) = T_R + \varepsilon$?

4. If $T_R, T_0,$ and k are known, with $T_0 < T_R,$ find the time $t_* > 0$ at which $T(t_*) = T_R - T_0/4$. (t_* will be expressed in terms of $T_R, T_0,$ and k.)

1.3 A Population Model

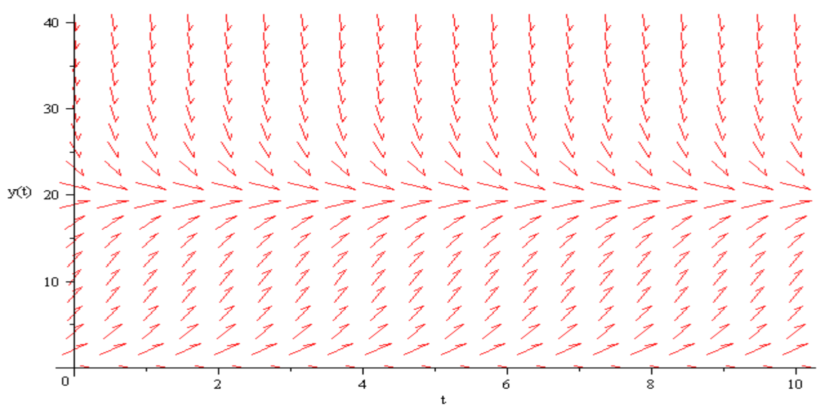
As second example of a mathematical model leading to a differential equation, consider an insect population in a confined environment. Let $P(t)$ denote the population size at time t and note that we can assert that

$$P'(t) = bP(t) - dP(t)$$

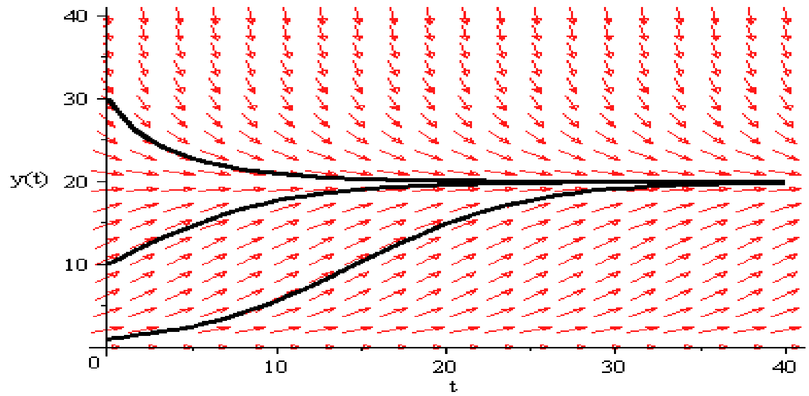
where b and d denote the (constant) birth and death rates for the population. Then, letting $k = b - d,$ we have $P'(t) = kP(t),$ for which the solution is $P(t) = Ce^{kt}$. Evidently, if $b > d$ then $k > 0$ and $P'(t) > 0$ for all time so the population grows without bound. If $b < d$ then $P'(t) < 0$ for all time and the population decreases steadily to zero (assuming $P(0) > 0$). Neither of these scenarios is a realistic description of what we expect to happen with this insect population and so we try to modify the model to obtain a more accurate description of what goes on. We expect that $k,$ which we can think of as the effective growth rate for the population, will be positive as long as the population does not get too large. If the population exceeds some critical value we might expect that, possibly due to limited food resources, conditions will be such that the population would tend to decrease. The simplest such dependence of k on P would be to have $k(P) = -r(P(t) - P_\infty),$ where r denotes a positive constant and P_∞ represents some critical population level that can be thought of as the maximum sustainable population level in the given environment. With this assumption the model for $P(t)$ becomes

$$P'(t) = -r(P(t) - P_\infty)P(t). \qquad (4)$$

The direction field and solution curves for this equation are as follows:



Direction Field



Three Solution Curves

This differential equation for the unknown function $P(t)$ differs from the equation in the previous example in the following way. On the right side of equation (1) we have $F_1(T) = -kT + kT_R$, while on the right side of (4) we have $F_2(P) = rP_\infty P - rP^2$. The function $F_1(T)$ is a linear function of T, meaning that if we plotted F_1 *versus* T, we would see a straight line graph. On the other hand, the function $F_2(P)$ is a nonlinear function of P, meaning that if we plotted F_2 *versus* P, we would see a graph that is not a straight line. As we shall see, linear differential equations are much easier to deal with than nonlinear equations. While the equation (4) can be solved, it is not hard to write down other nonlinear examples which we would not be able to solve. For the time being therefore, we will concentrate largely on linear equations and the solution techniques required to solve them.

Exercises

1. Sketch some solution curves for: $\frac{du}{dt} = -3u$

- **a.** If $u(0) > 0$, is there any $t > 0$ where $u(t) < 0$? Why or why not?
- **b.** Is there any finite $t > 0$ where $u(t) = 0$? Why or why not?
- **c.** Is it possible for the solution to this equation to oscillate? Why or why not?

2. Sketch some solution curves and answer questions a), b), c) for: $\frac{du}{dt} = -3u^3$

3. Repeat for: $\frac{du}{dt} = -3u^4$

4. Does changing p have much affect on the qualitative behavior of the solution to $\frac{du}{dt} = -3u^p$? In other words, do the answers to questions a,b,c change as the value of p changes?

5. How do the sketches of the solution curves change when the equation changes to: $\frac{du}{dt} = +3u^p$?

6. Consider the equation $\frac{dT}{dt} = -k(T(t) - T_A)$. Let $u(t) = T(t) - T_A$. Show that this implies $\frac{du}{dt} = -k\,u(t)$.

7. If $T(0) = T_A + 100$, then what is the value of $u(0)$? As $t \to \infty$, what happens to the values of $T(t)$ and $u(t)$?

8. Integrate to solve:

a. $\frac{du}{dt} = -u(t), \quad u(0) = 5.$

b. $\frac{du}{dt} = -u^2(t), \quad u(0) = 5.$

2. Solution Methods For First Order Ordinary Differential Equations

We are going to discuss a number of methods for constructing solutions for first order ordinary differential equations but our discussion will be confined to a few methods that can be applied to the equations that occur most frequently in applications.

2.1 Separable equations

Perhaps the simplest case in which the solution can be constructed is when the differential equation has the form

$$\frac{dy}{dt} = P(y)\,Q(t)$$

We say the equation is *separable* and proceed to integrate by writing

$$\int \frac{dy}{P(y)} = \int Q(t)dt$$

and integrating. Note that it appears as if we are treating $\frac{dy}{dt}$ as if it is a fraction and cross multiplying. In fact, what we are doing is the following:

$$\int Q(t)dt = \int \frac{1}{P(y)}\frac{dy}{dt}dt = \int \frac{dy}{P(y)}$$

Examples

1. Consider the equation

$$\frac{dy}{dt} = ty(t).$$

Then

$$\int \frac{dy}{y} = \int tdt.$$

Integrating the two sides of this last equation leads to,

$$\ln y(t) = t^2/2 + C_1$$

and since $exp[\ln y(t)] = y(t)$,

$$y(t) = C_2 e^{t^2/2}.$$

2. Sometimes the integration is more difficult as in the following equation,

$$\frac{dy}{dt} = -y(t)(1 - y(t))$$

Then

$$\int \frac{dy}{y(1-y)} = -\int dt$$

We must employ partial fractions to write

$$\frac{1}{y(1-y)} = \frac{1}{y} - \frac{1}{y-1}.$$

Then

$$\begin{aligned}\int \frac{dy}{y(1-y)} &= \int (\frac{1}{y} - \frac{1}{y-1})dy \\ &= \ln y - \ln(y-1) \\ &= \ln \frac{y}{y-1} = -t + C_1.\end{aligned}$$

Then, as in the previous example,

$$\frac{y(t)}{y(t)-1} = C_2 e^{-t}$$

or

$$y(t) = C_2 e^{-t}[y(t) - 1].$$

These last two expressions are called "implicit solutions" since they do not give $y(t)$ explicitly as a function of t. After a little algebra, the explicit solution can be written as,

$$y(t) = \frac{C_2 e^{-t}}{C_2 e^{-t} - 1}.$$

Exercises

Find the general solution for each of the following equations by separating and integrating. If there is an initial condition, find the unique solution to the initial value problem.

1. $y'(t) = \dfrac{ty(t)}{1+y(t)^2}$

2. $y'(t) = (t+1)(y+1)$

3. $y'(t) = t\,(y^2 - 4y + 3) \qquad y(0) = 1$

4. $y'(t) = t\dfrac{1}{y+2} \qquad y(0) = 2$

5. $(t^2+1)y'(t) = 2t\,y(t)$

6. $2y'(t) = t\,y(t)^3 \qquad y(1) = 1$

7. $y'(t) = 3t^2\,y(t) \qquad y(0) = \frac{1}{10}$

8. $y'(t) = 3t^2\,y(t)^2 \qquad y(0) = \frac{1}{10}$

9. $y'(t) = e^{t-y}$

10. $2y'(t) = 1 - y(t)^2$

2.2 Exact Equations

Another type of equation where constructing the solution is rather simple are the equations which take the form of an exact differential. To see what this means suppose $F(t, y(t))$ is a smooth function of two variables. Then the total differential of F is written,

$$dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial y}dy.$$

Since F depends explicitly on both t and $y,$ then any change in either t or y will contribute to a change in $F.$ The total differential expresses the relationship between the changes in t and y and the corresponding change in $F.$ If F is constant for all t and y then $dF = 0$ so it is clear that the differential equation

$$\frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial y}dy = 0 \qquad or \qquad \frac{\partial F}{\partial y}\frac{dy}{dt} = -\frac{\partial F}{\partial t}$$

is equivalent to $F(t,y) =$ constant. In general, $F(t,y) = C$ implicitly defines curves $y = y(t)$ called level curves of the function F and these curves are then also the solution curves of the differential equation $dF = 0.$

For example, $F(t,y) = y^p t^q$ and $dF = 0$ leads to the equation

$$qt^{q-1}y^p\, dt + py^{p-1}t^q\, dy = 0$$

$$or \;\; \frac{dy}{dt} = -\frac{qt^{q-1}y^p}{py^{p-1}t^q} = -\frac{qy}{pt}$$

whose solution is $y^p t^q = C$ or $y(t) = Ct^{-q/p}.$

In a case like this, we say the differential equation is *exact*. In general, the equation

$$P(y,t)dy + Q(y,t)dt = 0$$

is exact if $\quad P = \dfrac{\partial F}{\partial y} \quad$ and $\quad Q = \dfrac{\partial F}{\partial t} \quad$ for some smooth function $F(t,y).$ If such an F exists then it is necessary that

$$\frac{\partial}{\partial t}\left(\frac{\partial F}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial F}{\partial t}\right)$$

or

$$\frac{\partial P}{\partial t} = \frac{\partial Q}{\partial y}.$$

If this last condition is satisfied then a function $F(t,y)$ exists such that the solution curves of the equation are the so called "level curves" of $F;$ i.e. the curves such that $F(t,y) =$ constant.

For example, consider the differential equation

$$(2t-2y)dy - (4t-2y+5)dt = 0.$$

Then

$$P(y,t) = 2t - 2y \qquad \text{and} \qquad \frac{\partial P}{\partial t} = 2,$$

$$Q(y,t) = -4t + 2y - 5 \qquad \text{and} \qquad \frac{\partial Q}{\partial y} = 2.$$

This tells us there exists a smooth $F(t,y)$ for which

$$P = \frac{\partial F}{\partial y} = 2t - 2y$$

$$Q = \frac{\partial F}{\partial t} = -(4t - 2y + 5)$$

Integrating the first of these equalities with respect to y implies

$$F(t,y) = 2ty - y^2 + \phi(t).$$

Note that integrating $P(t,y)$ with respect to y leads to an expression for F that includes $\phi(t),$ an unknown function of t rather than an arbitrary constant of integration. This is the most general antiderivative of P with respect the variable y. Similarly, integrating $Q(t,y)$ with respect to t leads to

$$F(t,y) = -2t^2 + 2ty - 5t + \psi(y),$$

which is the most general antiderivative of Q with respect to the variable t. By inspection, it would appear from comparing the two results for $F(t,y)$, that $\phi(t) = -5t - 2t^2$, and $\psi(y) = -y^2$.

A more rigorous approach is to use the first expression for F to compute

$$\frac{\partial F}{\partial t} = Q$$
$$2y + \phi'(t) = -4t + 2y - 5.$$

Then

$$\phi'(t) = -4t - 5$$

and $$\phi(t) = -2t^2 - 5t + C_1.$$

Similarly, the second expression derived for F implies

$$\frac{\partial F}{\partial y} = P$$
$$2t + \psi'(y) = 2t - 2y$$

Then

$$\psi'(y) = -2y,$$

or $$\psi(y) = -y^2 + C_2.$$

Combining these two results leads to,

$$F(t,y) = 2ty - y^2 - 2t^2 - 5t + C_1$$
$$= -2t^2 + 2ty - 5t + -y^2 + C_2$$

and the implicit solution to the differential equation can therefore be written

$$y^2 - 2ty + 2t^2 + 5t = C_3.$$

The explicit solution in this case is found to be,

$$y(t) = t \pm \sqrt{C - 5t - t^2}$$

Often exact equations lead to implicit solutions for which it is not possible to solve explicitly for $y = y(t)$.

Exercises

1. Show that the following differential equation is not exact

$$(xy - 2)dx + (x^2 - xy)dy = 0$$

Now find a function $p(x)$ such that

$$p(x)(xy - 2)dx + p(x)(x^2 - xy)dy = 0$$

is exact and find a function $F(x,y)$ whose level curves are the solution curves for the equation.

2. Repeat this procedure for the differential equation

$$(2x^2 + y)dx + (x^2y - x)dy = 0$$

3. Show that the following differential equation is not exact

$$y\,dx + (2x - ye^y)dy = 0$$

Now find a function $q(y)$ such that

$$q(y)\,y\,dx + q(y)(2x - ye^y)dy = 0$$

is exact and find a function $F(x,y)$ whose level curves are the solution curves for the equation.

4. Show that

$$\frac{dy}{dt} = \frac{t^3 - y}{y^3 + t}$$

is exact and solve.

5. Show that

$$\frac{dy}{dt} = \frac{y^2 + yt}{t^2}$$

is not exact but

$$\frac{dy}{dt} = \frac{(y^2 + yt)/y^2t}{t^2/y^2t} = \frac{1/t + 1/y}{t/y^2}$$

is exact. Solve this equation.

2.3 Linear Equations

Linear equations provide a mathematical description of a large number of physical phenomena and are therefore an important class of differential equation. For example, the nonlinear problem

$$y'(t) = F(y(t)) \qquad y(0) = y_0$$

can be approximated by the following linear problem,

$$y'(t) = F(y_0) + F'(y_0)(y(t) - y_0) \qquad y(0) = y_0.$$

In general the behavior of solutions to linear equations is more restricted than the behavior of solutions to nonlinear differential equations but the linear problem is usually easier to solve.

The most general first order linear ODE is an equation of the form

$$p(t)\,\frac{dy}{dt} + q(t)\,y(t) = f(t). \qquad (1)$$

Here p and q are called **coefficients** and f is referred to as the **forcing term** in the equation. When $f = 0$, we say the equation is **homogeneous** and when f is not identically zero, we say the equation is **inhomogeneous**.

It is usually customary to divide equation (1) by the coefficient of the derivative so as to have the equation in so called **standard form**, i.e., the coefficient of the derivative term is a 1,

$$y'(t) + \frac{q(t)}{p(t)}\,y(t) = \frac{f(t)}{p(t)}.$$

More simply, we write

$$y'(t) + a(t)y(t) = F(t). \qquad (2)$$

If we introduce the notation

$$L[y(t)] = y'(t) + a(t)y(t)$$

then it is easy to see that for any constant, $C,$ and any function $y(t)$ having a derivative,

$$\begin{aligned} L[Cy(t)] &= (Cy(t))' + a(t)\,Cy(t) \\ &= C\,y'(t) + Ca(t)y(t) \\ &= CL[y(t)] \end{aligned}$$

In the same way, it is clear that

$$\begin{aligned} L[y_1(t) + y_2(t)] &= (y_1(t) + y_2(t))' + a(t)(y_1(t) + y_2(t)) \\ &= y_1'(t) + a(t)y_1(t) + y_2'(t) + a(t)y_2(t) \\ &= L[y_1(t)] + L[y_2(t)]. \end{aligned}$$

Combining these two results leads to the result that for all functions y_1 and y_2 and all constants C_1 and C_2,

$$L[C_1y_1(t) + C_2y_2(t)] = C_1L[y_1(t)] + C_2L[y_2(t)].$$

This property is referred to as the property of **linearity** and $L[\circ]$ here is called a **linear operator**. The differential equation $L[y(t)] = F(t)$ is called a linear differential equation. Since the operator L involves no derivative of order higher than one, L is called a first order linear differential operator.

2.3.1 Integrating Factors

The most direct way to solve the linear equation, $(2),$ is to suppose $A(t)$ is an anti-derivative of $a(t),$ that is, $A'(t) = a(t)$ and to note that

$$\begin{aligned} \frac{d}{dt}(e^{A(t)}y(t)) &= e^{A(t)}y'(t) + A'(t)e^{A(t)}y(t) \\ &= e^{A(t)}[y'(t) + a(t)y(t)]. \end{aligned}$$

Then (2) is equivalent to,

$$\frac{d}{dt}(e^{A(t)}y(t)) = e^{A(t)}F(t)$$

and

$$e^{A(t)}y(t) = \int^t e^{A(s)}F(s)ds$$

$$or \qquad y(t) = e^{-A(t)}\int^t e^{A(s)}F(s)ds.$$

The function $e^{A(t)}$ is called an integrating factor for this equation and while this is certainly an efficient approach to solving the differential equation, it conceals some of the structure of the solution. For this reason we will consider some other approaches, less efficient in the short run but more illuminating in the long run.

Examples

1. Consider the linear equation:

$$y'(t) + \cos t\, y(t) = \cos t$$

The anti-derivative of the coefficient $a(t) = \cos t$ is $A(t) = \sin t$ and the integrating factor is $e^{\sin t}.$ Then

$$\frac{d}{dt}(e^{\sin t}y(t)) = e^{\sin t}\left[y'(t) + \cos t\, y(t)\right]$$
$$= e^{\sin t}\cos t$$

and

$$e^{\sin t}y(t) = \int^t e^{\sin s}\cos s\, ds$$
$$= e^{\sin t} + C$$

Finally, the general solution is,

$$y(t) = 1 + Ce^{-\sin t}$$

2. Consider

$$y'(t) + \frac{1}{t}y(t) = t^4$$

Here, the antiderivative of the coefficient $a(t) = 1/t$ is $A(t) = \ln t$ and the integrating factor is $e^{A(t)} = t.$ Then multiplying by the integrating factor reduces the equation to

$$ty'(t) + y(t) = t^5$$
$$\frac{d}{dt}(ty(t)) = t^5$$

and

$$ty(t) = \frac{1}{6}t^6 + C$$
$$\text{or} \quad y(t) = \frac{1}{6}t^5 + C/t$$

Note that in both of these examples it is necessary to include the constant, $C,$ of integration. Omitting this constant would overlook an important component of the general solution.

Exercises

1. A pill in the shape of a cube with sidelength s is dropped into a container of solvent. If the pill retains its cubical shape as it dissolves then find an expression for s as a function of $t,$ given that the initial sidelength is 4 cm and after 5 minutes, the sidelength has decreased to 3 cm. **Hint**: Use the fact that the volume of the cube decreases at a rate that is proportional to the surface area of the pill to write a differential equation.

Classify each of the following ODE’s as separable or linear and solve for $u(t)$

2. $\frac{du}{dt} + \frac{1}{t}u(t) = t^4$

3. $\frac{du}{dt} - 4u(t) = te^{-4t}$

4. $\frac{du}{dt} + 2u(t) = 2\sin 3t$

5. $\frac{du}{dt} = e^{t-u}$

6. $2\dfrac{du}{dt} = (1-u^2)$

2.3.2 Variation of Parameters

An alternative method for solving linear equations consists of two steps. In order to solve equation (2) in the general inhomogeneous case (i.e., $F \neq 0$) we will first solve the corresponding homogeneous version of (2), that is,

$$y'(t) + a(t)\,y(t) = 0.$$

This equation is separable and we integrate as follows

$$\begin{aligned}\ln y &= \int \frac{dy}{y} = -\int^t a(s)\,ds + C_0\\ &= -A(t) + C_0\end{aligned}$$

Then the solution to the homogeneous equation is $y_H(t) = C_1 e^{-A(t)}$ where $A'(t) = a(t)$.

Here $y_H(t)$ is referred to as the general solution for the **homogeneous** equation. This means that for any choice of the constant C_1, the function $y_H(t)$ solves the homogeneous ODE, and every solution of the homogeneous equation must be given by this formula for some choice of C_1. The first half of this statement has been proved by the construction we have just performed. The second half of the statement is true but has not been proved yet. At any rate, we will now use the homogeneous solution to find a solution for the inhomogeneous equation. This is accomplished by supposing the inhomogeneous equation (2) has a solution of the form

$$y_p(t) = C(t)\,e^{-A(t)}.$$

We refer to this as a **particular solution**. Here $C(t)$ denotes an unknown function which we will now find. Note that

$$\begin{aligned}y_p'(t) &= C'(t)\,e^{-A(t)} + C(t)\,e^{-A(t)}(-A'(t))\\ &= C'(t)\,e^{-A(t)} - C(t)\,e^{-A(t)}a(t).\end{aligned}$$

Then

$$\begin{aligned}y_p{}'(t) + a(t)\,y_p(t) &= C'(t)\,e^{-A(t)} - C(t)\,e^{-A(t)}a(t) + a(t)\,C(t)\,e^{-A(t)}\\ &= C'(t)\,e^{-A(t)},\end{aligned}$$

and the inhomogeneous equation now reduces to,

$$\begin{aligned}y_p{}'(t) + a(t)\,y_p(t) &= C'(t)\,e^{-A(t)}\\ &= F(t)\end{aligned}$$

This is not a coincidence. If the differential equation is written in standard form, then the assumption $y_p(t) = C(t)\,e^{-A(t)}$ will always reduce the equation to $C'(t)\,e^{-A(t)} = F(t)$. Then

$$C'(t) = F(t)e^{A(t)}.$$

and we integrate to get $C(t)$,

$$C(t) = \int^t F(\tau)e^{A(\tau)}\,d\tau,$$

This is the solution for $C(t)$. The solution to the ODE is given by,

$$
\begin{aligned}
y_p(t) &= C(t)\, e^{-A(t)} \\
&= e^{-A(t)} \int^t F(\tau) e^{A(\tau)}\, d\tau.
\end{aligned}
$$

This method of finding a particular solution is called the method of **variation of parameters**. The assumption that the particular solution could be written in the form $y_p(t) = C(t) y_H(t)$ had the effect of reducing the inhomogeneous equation to a simple integration to find $C(t)$.

The general solution for the inhomogeneous equation is defined to be the sum of the general homogeneous solution and a particular solution,

$$
\begin{aligned}
y(t) &= y_H(t) + y_p(t) \\
&= C_1 e^{-A(t)} + e^{-A(t)} \int^t F(\tau) e^{A(\tau)}\, d\tau.
\end{aligned}
$$

Note that if we add any homogeneous solution to a particular solution, we obtain a new particular solution. It is for this reason that we refer to $y_p(t)$ as A particular solution and not THE particular solution. It is sometimes useful to think of the homogeneous solution to the equation as the response of the system to the initial conditions, while the particular solution is the response to the forcing term. The system we are referring to is the physical system the differential equation models although we may also think of the equation itself as being a mathematical operator that produces responses to both the initial state and the forcing.

Examples

1. Consider

$$y'(t) + 6y(t) = t$$

We easily find the homogeneous solution to be $y_H(t) = Ce^{-6t}$.
Now we suppose the particular solution has the following form

$$y_p(t) = C(t)\, e^{-6t}$$

If we substitute this into the original equation, we arrive at

$$C'(t)\, e^{-kt} = F(t) = t$$

Then

$$C'(t) = F(t) e^{6t} = t e^{6t},$$

and

$$
\begin{aligned}
C(t) &= \int t e^{6t} dt \\
&= \frac{1}{36} e^{6t} (6t - 1),
\end{aligned}
$$

Then our particular solution is

$$
\begin{aligned}
y_p(t) &= C(t)\, e^{-6t} \\
&= \frac{1}{36} (6t - 1)
\end{aligned}
$$

and the general solution for this example is

$$y(t) = Ce^{-6t} + \frac{1}{36} (6t - 1).$$

Note that the part of the solution containing the arbitrary constant of integration is the

homogeneous solution while the particular solution contains no arbitrary constant.
2. Consider

$$y'(t) + \frac{1}{t}y(t) = t^4$$

Here the homogeneous solution is easily found to be $y_H(t) = C/t$ and if we assume the particular solution has the form $y_p(t) = C(t)t^{-1}$, then

$$C'(t)t^{-1} = t^4$$

$$\text{or} \quad C'(t) = t^5$$

Then $C(t) = t^6/6$ and the particular solution is $y_p(t) = t^5/6$, which leads to the following general solution

$$y(t) = \frac{1}{6}t^5 + C/t.$$

This is the result found in the example in the previous section but we recognize now that the part of the solution with the arbitrary constant is the homogeneous solution while the remainder is a particular solution.

2.3.3 Undetermined Coefficients

Still another way to obtain a particular solution to an inhomogeneous equation is by guessing (or we could call it the method of undetermined coefficients). For example, in solving the equation

$$y_p'(t) + ky_p(t) = t,$$

we note that it would be reasonable to assume that $y_p(t) = at + b$ where the constants a, b are to be determined. Substituting this guess into the equation, we find

$$(at+b)' + k(at+b) = a + kb + akt = t.$$

Then, equating the coefficients of like powers of t on the two sides of this last equation,

$$ak = 1 \qquad \text{and} \qquad a + kb = 0,$$

hence

$$a = \frac{1}{k} \qquad \text{and} \qquad b = -\frac{a}{k} = -\frac{1}{k^2}.$$

This leads to $\quad y_p(t) = at + b = \frac{1}{k}t - \frac{1}{k^2}$, which agrees with the previous result.

As a second example, consider the inhomogeneous equation,

$$y'(t) + ky(t) = F_0 \cos\Omega t$$

where k, Ω, F_0 all denote given constants. If we were to use the method of variation of parameters we suppose $y_p(t) = C(t)e^{-kt}$ where

$$C'(t)e^{-kt} = F(t) = F_0 \cos\Omega t$$

Then

$$C(t) = F_0 \int e^{kt} \cos\Omega t\, dt$$

$$=: F_0 \frac{e^{kt}}{k^2 + \Omega^2}[k \cos\Omega t + \Omega \sin\Omega t],$$

and

$$y_p(t) = C(t)\,e^{-kt} = \frac{F_0}{k^2+\Omega^2}[k\cos\Omega t + \Omega\sin\Omega t].$$

This particular solution could as easily have been obtained by guessing. Since the forcing term involves $\cos\Omega t,$ it is logical to suppose that the particular solution could only be composed of some combination of $\cos\Omega t$ and $\sin\Omega t;$ i.e.,

$$y_p(t) = a\cos\Omega t + b\sin\Omega t$$

for some constants a,b.
Then

$$\begin{aligned} y_p'(t) + ky_p(t) &= \frac{d}{dt}(a\cos\Omega t + b\sin\Omega t) + k(a\cos\Omega t + b\sin\Omega t) \\ &= (b\Omega + ka)\cos\Omega t + (bk - a\Omega)\sin\Omega t \\ &= F_0\cos\Omega t. \end{aligned}$$

Equating coefficients of $\cos\Omega t$ and $\sin\Omega t$ on the two sides of this last equation leads to

$$b\Omega + ka = F_0 \qquad \text{and} \qquad bk - a\Omega = 0$$

Then

$$a = \frac{F_0 k}{k^2+\Omega^2}, \qquad b = \frac{F_0\Omega}{k^2+\Omega^2}$$

This agrees with the result obtained by variation of parameters. The method of undetermined coefficients is most useful when the linear equation has constant coefficients and the forcing term is fairly simple. In more complicated examples it is generally more effective to use one of the other methods of finding the solution.

2.3.4 Amplitude and Phase

In some situations, it is desirable to express the particular solution in a form that is consistent with the forcing term. For example, in the case of the equation,

$$y'(t) + ky(t) = F_0\cos\Omega t$$

we found

$$y_p(t) = \frac{F_0}{k^2+\Omega^2}[k\cos\Omega t + \Omega\sin\Omega t].$$

We can choose to write the solution in the form,

$$y_p(t) = \frac{F_0}{\sqrt{k^2+\Omega^2}}\left[\frac{k}{\sqrt{k^2+\Omega^2}}\cos\Omega t + \frac{\Omega}{\sqrt{k^2+\Omega^2}}\sin\Omega t\right].$$

Now let

$$\alpha =. \frac{k}{\sqrt{k^2+\Omega^2}} \qquad \text{and} \qquad \beta =. \frac{\Omega}{\sqrt{k^2+\Omega^2}}$$

and note that

$$\alpha^2 + \beta^2 = \frac{k^2}{k^2+\Omega^2} + \frac{\Omega^2}{k^2+\Omega^2} = 1.$$

Since the sum of α^2 *and* β^2 equals one, it follows that we can always find an angle θ such that $\alpha = \cos\theta,$ *and* $\beta = \sin\theta;$ i.e.

$$\frac{\beta}{\alpha} = \frac{\cos\theta}{\sin\theta} = Tan\theta \qquad or \qquad \theta = Tan^{-1}\left(\frac{\beta}{\alpha}\right).$$

Then we can write the particular solution in the form

$$\begin{aligned} y_p(t) &= \frac{F_0}{\sqrt{k^2+\Omega^2}}\left[\frac{k}{\sqrt{k^2+\Omega^2}}\cos\Omega t + \frac{\Omega}{\sqrt{k^2+\Omega^2}}\sin\Omega t\right] \\ &= \frac{F_0}{\sqrt{k^2+\Omega^2}}[\cos\theta\cos\Omega t + \sin\theta\sin\Omega t] \\ &= \frac{F_0}{\sqrt{k^2+\Omega^2}}\cos(\Omega t - \theta) \end{aligned}$$

from which it is evident that $y_p(t)$ is a periodic function having the same period as the forcing term, $F(t)$ but with a phase shift equal to θ. In addition it is clear that the amplitudes of the forcing term (i.e., the input) and the particular solution (the output) are equal to F_0 *and* $\frac{F_0}{\sqrt{k^2+\Omega^2}}$, respectively. When the solution is expressed in this form, it is evident that the effect of system on the input is to modify the amplitude and introduce a shift in the phase.

3. Existence and Uniqueness

We have learned a few methods for constructing solutions for first order ordinary differential equations and, in particular, for finding solutions that satisfy a given initial condition. We have proceeded to do so, assuming that such solutions always exist and are unique. Unfortunately this is not always the case. A full discussion of the questions of existence and uniqueness of solutions to first order ordinary differential equations may be beyond the scope of this course but we should at least make a cursory attempt to understand the factors which determine the existence of solutions to differential equations and the uniqueness of solutions to initial value problems.

We will by considering some examples of a linear first order equation:

$$y'(t) + a(t)y(t) = f(t).$$

Recall that the general solution to a linear equation is of the form $y_G(t) = y_p(t) + y_H(t)$, consisting of the sum of a particular solution y_p and a homogeneous solution y_H where y_H contains an arbitrary constant. This general solution is, in fact, an infinite one parameter family of solutions where the arbitrary constant in y_H plays the role of the parameter. If an initial condition is added to the differential equation to form an initial value problem, then there is a unique value of the parameter for which the general solution satisfies the initial condition. It would appear that the questions of existence and uniqueness for a linear first order equation can be simply answered. It is, however, the case that the coefficient $a(t)$ and the forcing term, $f(t)$, in the equation play a role in determining the properties of the solution to the equation. We will illustrate with the following examples .

Linear Examples

1. Let $f(t) = 0$ and suppose $a(t)$ is defined for all values of t with $A(t)$ an antiderivative of $a(t)$; *i.e.*, $A'(t) = a(t)$. Then, as we have seen previously, the solution of our differential equation is equal to

$$y(t) = Ce^{A(t)}.$$

Then $y(t)$ is defined and satisfies the differential equation for all values of t. We say that $y(t)$ is a global solution for the differential equation. It is also easy to see that if there is an initial condition of the form, $y(t_0) = y_0$, then the unique value of C for which the initial condition is satisfied is given by $C = y_0 e^{-A(t_0)}$.

2. Again let $f(t) = 0$ and consider $a(t) = \sqrt{1-t^2}$, which is only defined on the interval, $-1 \leq t \leq 1$. In this case we have

$$y(t) = Ce^{-\arcsin t}$$

and we see that the solution to the equation exists but is, like $a(t)$, defined only on the interval $[-1,1]$.

3. Consider $f(t) = 0$ and $a(t) = \begin{cases} 2 & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$. Here, $a(t)$ is defined for all t but $a(t)$ is not continuous at $t = 0$. In this case the solution is found to be $y(t) = \begin{cases} C_1 e^{-2t} & \text{if } t > 0 \\ C_2 & \text{if } t \leq 0 \end{cases}$. If we choose $C_1 = C_2$ then $y(t)$ is defined and continuous for all t but $y'(t)$ is discontinuous at $t = 0$.

4. Finally, consider $a(t) = 1$ and $f(t) = \begin{cases} 2 & \text{if } t > 0 \\ 0 & \text{if } t \leq 0 \end{cases}$. Here, $a(t)$ is defined and continuous for all t but $f(t)$ is not continuous at $t = 0$. In this case the solution is found to be $y(t) = Ce^{-t} + \begin{cases} 2 & \text{if } t > 0 \\ 2e^{-t} & \text{if } t \leq 0 \end{cases}$. As in the previous example, $y(t)$ is defined and continuous for all t but $y'(t)$ is discontinuous at $t = 0$.

The point of these examples is to show that although existence and uniqueness are assured, the properties of the solution to the linear equation are dependent on the behavior of the coefficient and forcing term in the equation. Our reason for considering these examples here is to contrast the behavior of solutions to linear equations to the behavior of nonlinear equations. Solutions to linear differential equations will exist and be as smooth as the ingredients of the equation allow them to be but solutions to nonlinear equations can exhibit singular behavior even when there is nothing in the equation to suggest the solution may fail at some point.

Nonlinear Examples

1 Consider the initial value problem,

$$y'(t) = y(t)^2, \qquad y(0) = A > 0.$$

Rewrite the equation as

$$\int \frac{dy}{y^2} = \int dt.$$

Integrating leads to

$$-y(t)^{-1} = t - C_0,$$

or

$$y(t) = \frac{1}{C_0 - t}.$$

Then the initial condition, $y(0) = A$, implies $C_0 = 1/A$ and the solution of the initial value problem is given by,

$$y(t) = \frac{A}{1 - At}.$$

This solution becomes undefined at $t = 1/A$ even though there is nothing in the equation to suggest the solution should not exist for all values of t. We say that the interval $[0, 1/A]$ is the interval of existence for this solution, and since the solution does not exist for all values of t, it is a "local solution". Note also that the function $y(t) = 0$ satisfies the differential equation for all $t \geq 0$ but for no choice of C_0 does the "general" solution $y(t) = (C_0 - t)^{-1}$ equal this solution.

2 Consider the initial value problem,

$$y'(t) = 2\sqrt{y(t)}, \qquad y(t_0) = 0, \quad t_0 > 0.$$

Rewrite the equation as follows,

$$\int \frac{dy}{2\sqrt{y(t)}} = \int dt,$$

and integrate to find

$$\sqrt{y(t)} = t - C_0$$

or

$$y(t) = (t - C_0)^2.$$

The 1 parameter family of functions $y(t) = (t - C_0)^2$ solves the differential equation for each value of the constant C_0. However, this cannot be called the general solution of the equation since the zero function $y(t) = 0$ also solves the equation but the zero function does not equal $(t - C_0)^2$ for any value of C_0.

Notice also that

$$y(t) = (t - t_0)^2$$

solves the initial value problem, while at the same time, for every $t_1 > t_0$, the piecewise defined functions

$$y(t) = \left\{ \begin{array}{ccc} 0 & if & t_0 < t < t_1 \\ (t - t_1)^2 & if & t > t_1 \end{array} \right\}$$

also solve the initial value problem. Evidently, the initial value problem has infinitely many distinct solutions.

If we replace the initial condition above with the condition, $y(1) = 4$, then we find

$$y(1) = (1 - C_0)^2 = 4$$

$$\text{and} \qquad C_0 = 1 \pm 2 = 3, -1;$$

i.e., $y_1(t) = (t - 3)^2$ and $y_2(t) = (t + 1)^2$ are apparently both solutions to the new initial value problem. However, the equation asserts that $y'(1) = 2\sqrt{1} = 2$ (recall that $\sqrt{1}$ equals 1, not ± 1) and only y_2 satisfies this equation.

These examples illustrate that solutions to nonlinear differential equations may behave quite differently from solutions to linear problems. In particular, they illustrate the need for some way of deciding if an equation has or does not have a solution and if there is a unique solution to an associated initial value problem. The following theorem, which we state without proof, provides the answers to these questions.

Theorem *(Existence-Uniqueness theorem) Consider the differential equation*

$$y'(t) = F(t,y(t))$$

where $F(t,y)$ is defined and continuous on a rectangular region $R = I \times U,$ in the t-y plane. Then there exists a function $y(t)$ which is a solution for the equation for t in an interval J contained in I. If F is such that $\partial_y F$ is continuous on $R,$ then for any (t_0, y_0) in R there exists a unique function $y(t)$, which satisfies the initial condition $y(t_0) = y_0$ and is a solution for the equation on an interval $[t_0, T],$ contained in I.

Let us apply this theorem to the examples we have just considered. In example 1, $F(t,y) = y^2$ is defined and continuous along with $\partial_y F = 2y,$ for all values of y (and t). Then $R = I \times U = (-\infty,\infty) \times (-\infty,\infty)$ in this case.The solution

$$y(t) = \frac{1}{C_0 - t},$$

exists on an interval $J = (C_0,\infty)$ or, alternatively on $J = (-\infty, C_0).$ The initial condition $y(0) = A > 0$ is satisfied by the solution

$$y(t) = \frac{A}{1 - At},$$

on the interval $[0, 1/A]$ which is contained in $I.$ This is the unique solution to the initial value problem on the interval of existence, $[0, 1/A]$.

In example 2, $F(t,y) = 2\sqrt{y(t)}$ is defined and continuous for $y \geq 0$ and all $t,$ while the derivative $\partial_y F = 1/\sqrt{y(t)}$ is undefined at $y = 0$ but is continuous for $y > 0.$ Then $R = (-\infty,\infty) \times (0,\infty).$ We have seen that the 1 parameter family of functions $y(t) = (t - C_0)^2$ solves the differential equation for each value of the constant C_0 and according to the theorem, this solution is valid for t in some interval, $0 \leq C_0 < t < T,$ or $-T < t < C_0 \leq 0.$ As we saw in the example, there is not a unique solution which satisfies the initial condition, $y(t_0) = 0,$ but this does not contradict the theorem since the initial value $y_0 = 0$ does not belong to R. An initial value problem with a non-zero initial condition $y(t_0) = y_0 > 0$ would have a unique solution according to the theorem since (t_0, y_0) does belong to R in this case.

For any linear problem, $F(t,y) = a(t)y(t) + f(t)$ and $\partial_y F(t,y) = a(t),$ so existence and uniqueness of solutions is determined by the properties of $a(t)$ and $f(t)$ as we have seen previously. In particular, existence and uniqueness can be inferred from the theorem on any rectangle $R = I \times U$ where $a(t)$ and $f(t)$ are defined and continuous on I.

Exercises

Discuss the existence, uniqueness and interval of existence for the following initial value problems. Consider any special values for y_0, t_0.

1. $y'(t) = \dfrac{y(t)}{\sqrt{t}}, \qquad y(t_0) = y_0.$
2. $y'(t) = \sqrt{\dfrac{y(t)}{t}}, \qquad y(t_0) = y_0.$
3. $y'(t) = y(t)^3, \qquad y(t_0) = y_0.$
4. $y'(t) = \dfrac{1}{y(t) + 3}, \qquad y(t_0) = y_0.$

4. First Order Nonlinear Equations

4.1 Some preliminary remarks

The most general first order ordinary differential equation we could imagine would be of the form

$$F(t,y(t),y'(t)) = 0. \qquad (1)$$

For example, (1) could take the form

$$\frac{a(t) + y(t)}{\sqrt{1 + b(t)(y'(t))^2}} = f(t).$$

In general we would have no hope of solving such an equation analytically and our only recourse would be to resort to some numerical solution scheme. A less general but still nonlinear equation would be one of the form

$$y'(t) = F(t,y(t)), \qquad (2)$$

but even this less general equation is often too difficult to solve. We will consider then, even simpler equations of the form

$$y'(t) = F(y(t)). \qquad (3)$$

Equation (3) is said to be an **autonomous** differential equation, meaning that the nonlinear function F depends on $y(t)$ but does not depend explicitly on t. The equation (2) is nonautonomous because F does contain explicit t dependence. The equations,

$$y'(t) = y(t)^2 \qquad and \qquad y'(t) = y(t)^2 + t^2,$$

are examples of autonomous and nonautonomous equations, respectively.

We will consider some examples of nonlinear first order equations first and then state some general principles that will make it clear why autonomous equations are easier to deal with than nonautonomous ones. In particular, it is often possible to extract considerable information about the solution of an autonomous equation even when it is impossible to construct the solution.

4.2 Autonomous First Order Equations- A population model

The simplest possible model for population growth is the equation

$$P'(t) = kP(t), \qquad P(0) = P_0$$

where the constant k denotes the growth rate of the population. If k is positive, the population described by this equation grows rapidly to infinity while, if k is negative, it decays steadily to zero.

In an effort to make a more realistic model for growth of population size, we supposed in section 1.3 that k is not a constant but depends on P. In particular, we supposed $k(P) = r(P_\infty - P(t))$ for some positive constants r, P_∞. Note that if $P(t) < P_\infty$ then $k(P) > 0$ and $P(t)$ increases, while $P(t)$ will decrease if $P(t) > P_\infty$ since then $k(P) < 0$. With this choice of $k(P)$, we obtain the autonomous equation

$$P'(t) = r(P_\infty - P(t))P(t), \qquad P(0) = P_0, \qquad (4)$$

where $F(P) = r(P_\infty - P(t))P(t)$ does not depend explicitly on t.

It is not difficult to solve this differential equation but we are going to see that it is possible to completely understand the behavior this equation predicts without actually solving the equation. In fact, it is generally harder to see the predicted behavior from the analytic solution than from the qualitative analysis we are going to describe.

We begin by determining if the equation has any ***critical points***. These are values P_* of P such that $F(P_*) = 0$. Critical points are important since if there is any time $t = T_*$ at which $P(T_*) = P_*$, then $P(t) = P_*$ for all $t \geq T_*$. Evidently, equation (4) has two critical points, namely $P = 0 \; and \; P = P_\infty$ and if the population $P(t)$ starts with an initial value P_0 equal to either of these values, then the population will remain constantly equal to that value for all $t > 0$.

For any solution curve which starts at an initial value, P_0, with $0 < P_0 < P_\infty$, we can show that P increases along this solution curve for all $t > 0$, and approaches the value P_∞ asymptotically. The argument goes like this:

At $t = 0$ we have $P'(0) = r(P_\infty - P_0)P_0 > 0$, which means the curve starts out with P increasing. If there is a point on this solution curve where $P(t)$ is not increasing, then there has to be a time $t_0 > 0$ where $P'(t_0) = 0$ (i.e., in order for $P(t)$ to be decreasing, it has to first stop increasing). Since this is a point on a solution curve where $P'(t_0) = 0$ then (4) requires that either $P(t_0) = 0$ or else $P(t_0) = P_\infty$. Clearly $P(t_0) = 0$ is not possible since P started at P_0 with $0 < P_0 < P_\infty$, and P has been increasing since then. On the other hand, $P'(t_0) = 0$ with $0 < P(t_0) < P_\infty$ is not possible either since this contradicts equation (4) which asserts that $P'(t_0) = 0$ only when $P(t_0) = 0$ or $P(t_0) = P_\infty$. The only remaining possibility is that a solution curve which originates at P_0 with $0 < P_0 < P_\infty$, must approach the horizontal asymptote $P = P_\infty$, increasing steadily as t increases. In the same way, we can argue that any solution curve that originates at P_0 with $P_0 > P_\infty$, must approach the horizontal asymptote $P = P_\infty$, from above, decreasing steadily as t increases. A plot showing three different solution curves is shown below.

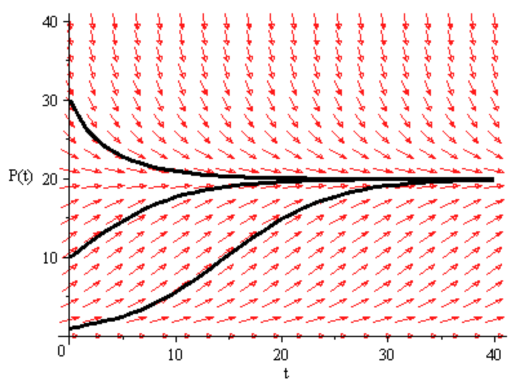


Figure 1 Solution Curves

The arrows on this plot indicate the slope of the $P(t)$ versus t at each point in the plane. We can evaluate the slope, $P'(t)$, at each point $(t, P(t))$ by just evaluating $F(P(t))$. The collection of arrows is called a "direction field" for this equation and any curve that is tangent to the direction field at each point is called a solution curve or trajectory for the equation. The term "orbit" may also be used to refer to a solution curve.

4.3 Critical Points for Nonlinear Equations

Consider the autonomous nonlinear equation

$$P'(t) = F(P(t)).$$

We have defined a critical point for this equation to be any value $P = P_*$ such that

$F(P_*) = 0.$ Then we define an **asymptotically stable critical point** to be a value $P = P_*$ such that $F(P_*) = 0,$ with the additional condition that for any solution curve, $P(t),$ originating at a point $P(0)$ near $P_*,$ it must follow that $P(t) \to P_*$ *as* $t \to \infty.$ That is, P_* is asymptotically stable if any trajectory that begins near P_* must converge to $P_*.$ On the other hand, a critical point P_* for which the distance between $P(t)$ and P_* *increases* as $t \to \infty,$ even for $P(0)$ arbitrarily near $P_*,$ is said to be **unstable**. Equation (4) above has critical points at $P = 0$ *and* $P = P_\infty.$ By examining the direction field in the figure above, we see that the critical point at $P = 0$ is unstable, while the critical point at $P = P_\infty$ is asymptotically stable. That is, all the solution curves that start near $P = 0$ move away from $P = 0$ and are attracted toward $P = P_*.$ Solution curves that begin either above or below the value $P = P_*$ tend toward $P_*.$

It is possible for a critical point to be such that solution curves that begin on one side of the critical point are attracted back to the critical point while solution curves that begin on the opposite side of the critical point are repelled and move away from the critical point. A critical point with this kind of behavior is said to be "neutrally stable.

Now we state some general results about autonomous nonlinear equations. If we consider the equation

$$y'(t) = F(y(t)), \qquad (5)$$

then

1. the critical points of (5) are the values y_* for which $F(y_*) = 0$
2. the critical point, y_* is **stable** if $F'(y_*) < 0$
3. the critical point, y_* is **unstable** if $F'(y_*) > 0$
4. distinct solution curves of (5) can never cross

Here point 1 is just the definition of critical point.

Points 2 and 3 assert that a critical point, $y_*,$ for (5) can be classified as stable or unstable by noting the sign of $F'(y_*).$ To see why the sign of $F'(y)$ controls the stability of the critical point consider the example $F(y) = (1-y)(y-4),$ which has critical points at $y = 1,4.$ The figure below shows the plot of F versus $y.$ Consider first the critical point at $y = 1.$ Note that for y slightly larger than $y = 1,$ we have $F > 0$ which means $y'(t) > 0$ and y is increasing (i.e., moving away from $y = 1$). If y is slightly less than $y = 1,$ then $F < 0$ so $y'(t) < 0$ there and y is decreasing. In either case, whether y is slightly less than or slightly larger than $y = 1,$ the value of y tends to move away from $y = 1.$ At the critical point $y = 4,$ the behavior is just the opposite. If $y < 4,$ then $F(y) > 0$ and y increases and if $y > 4$ then $F(y) < 0$ so y decreases. In either case, when y is close to 4 it will tend to move closer to $y = 4.$

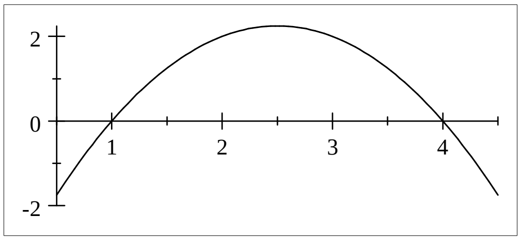


Figure 2 F(y) versus y

A more analytic argument is based on expanding $F(y)$ about a critical point y_*. That is, suppose $F(y_*) = 0$ and write the first couple of terms of the Taylor series for F about $y = y_*$

$$F(y) = F(y_*) + F'(y_*)(y - y_*) + \tfrac{1}{2}F''(z)(y - y_*)^2.$$

If y is close to y_* then $\frac{1}{2}F''(z)(y - y_*)^2$ is small relative to $F'(y_*)(y - y_*)$, and since $F(y_*) = 0$ we have

$$y'(t) \approx F'(y_*)(y - y_*)$$

But y_* is a constant so $y'(t) = \frac{d}{dt}(y - y_*)$ and $y - y_*$ satisfies $\frac{d}{dt}(y - y_*) = F'(y_*)(y - y_*)$. Clearly if $F'(y_*) > 0$, then $y - y_*$ will increase, while if $F'(y_*) < 0$, then $y - y_*$ will decrease. It follows that the stability of the critical point $y = y_*$ is determined by the sign of $F'(y_*)$.

To see that point 4 must be true suppose that $y = y_1(t)$ and $y = y_2(t)$ are two different solutions of (5) whose graphs cross at some time t_0. To say the graphs "cross" at $t = t_0$ means that $y_1(t_0) = y_2(t_0),\ \textit{and}\ y_1'(t_0) \neq y_2'(t_0).$; i.e., the graphs go through the same point but have different slopes there. But if $y_1(t_0) = y_2(t_0)$, then $F(y_1(t_0)) = F(y_2(t_0))$ and this implies $y_1'(t_0) = y_2'(t_0)$; i.e.,$y_1(t_0) = y_2(t_0)$

$$y_1'(t_0) = \overbrace{F(y_1(t_0)) = F(y_2(t_0))}^{y_1(t_0) = y_2(t_0)} = y_2'(t_0),$$

Clearly at a point where $y_1(t_0) = y_2(t_0)$ the slopes cannot be different. This contradiction shows that solution curves are either identical or else they never cross. We say that the family of solution curves is **coherent**.

Using these observations about the behavior of autonomous first order equations it is possible to sketch solution curves as in Figure 1. Such a sketch is called a "**phase plane portrait**" for the autonomous equation. The steps in making such a sketch are: i) find all the critical points where $F(y) = 0$, ii) Classify the critical points as stable or unstable by evaluating $F'(y_*)$, iii) Plot a selection of solution curves starting on either side of each critical point.

Exercises

For each of the following autonomous equations, find and classify the stability of all the critical points and then sketch the phase plane portrait for each. In doing this, it will be helpful to plot $F(y)$ versus y as in Figure 2. What will the plot of $F(y)$ versus y look like when there is a neutrally stable fixed point?

1. $y'(t) = (y + 2)(6 - y)$

2. $y'(t) = 9y - y^3$

3. $y'(t) = \cos 2y(t)$
4. $y'(t) = \sqrt{1 - y(t)^2}$
5. $y'(t) = y(4 - y^2)$
6. $y'(t) = y^2(4 - y)^2$

5. Applications of First Order Equations

In this section we are going to discuss a number of simple applications of first order ordinary differential equations.

5.1. Viscous Friction

Consider a small mass that has been dropped into a thin vertical tube of viscous fluid like oil. The mass falls, due to the force of gravity, but falls more slowly than it would in a fluid like water because the oil is thicker than water, (or, as we say, the oil is more viscous than the water). One of Newton's laws asserts that the time rate of change of the momentum of the mass is equal to the sum of the external forces acting on the mass,

$$\frac{d}{dt}(mv(t)) = F$$

where m is the mass of the object and $v(t)$ is the mass velocity as a function of time. The forces acting on the mass are the force of gravity, F_g and the friction or viscous force of the oil, F_f. Then

$$F_g = mg$$

$$\text{and} \qquad F_f = -kv(t),$$

where we have assumed that the friction force is proportional to the velocity (i.e., the faster the object moves, the more the oil retards it). The constant of proportionality, k, is assumed to be constant and then the negative sign appears in the definition of F_f so that the force acts in the direction opposite to the direction of the velocity. Then the differential equation becomes

$$mv'(t) = mg - kv(t), \qquad v(0) = v_0, \qquad (1.1)$$

where v_0 denotes the initial velocity of the mass. We can rewrite the equation in standard form as

$$v'(t) = g - \frac{k}{m}v(t). \qquad (1.2)$$

Before solving this equation we will first carry out a qualitative analysis on the equation. We note that

- if $g - \frac{k}{m}v(t) < 0,\ i.e.,\ v(t) > \frac{mg}{k}$ then $v'(t) < 0$ so $v(t)$ is decreasing
- if $g - \frac{k}{m}v(t) > 0,\ i.e.,\ v(t) < \frac{mg}{k}$ then $v'(t) > 0$ so $v(t)$ is increasing
- if $v(t) = \frac{mg}{k}$ then $v'(t) = 0$ so $v(t)$ is constant

Then a solution curve that begins at a point above the value $\frac{mg}{k}$ will decrease toward this value while a solution curve starting out below this value will increase towards $\frac{mg}{k}$. In addition, we can differentiate the equation (1.2) to obtain

$$v''(t) = -\frac{k}{m}v'(t) = -\frac{k}{m}\left(g - \frac{k}{m}v(t)\right),$$

from which we conclude that

$$\text{if } v(t) > \frac{mg}{k} \quad \text{then} \quad v''(t) > 0$$
$$\text{and if } v(t) < \frac{mg}{k} \quad \text{then} \quad v''(t) < 0.$$

Then the solution curves above the value $\frac{mg}{k}$ all have positive curvature and decrease toward the limiting value $v_\infty = \frac{mg}{k}$. Those curves that begin below the limiting value have negative curvature and increase toward v_∞. This constant could be interpreted as the terminal velocity of the projectile falling in the fluid..

The curve lying above the constant $v_\infty = \frac{mg}{k}$ is consistent with an object that is fired into the oil with a large initial velocity in which case the projectile gradually slows down as it continues to fall only under the effect of gravity and the viscous drag of the oil. The curve lying below the constant represents the case where the object is introduced into the tube with a small or zero initial velocity. In this case, the velocity increases toward the terminal velocity $\frac{mg}{k}$ due to gravity and friction.

In order to now solve for $v(t)$, we note that equation (1.2) is equivalent to

$$\int \frac{dv}{g - \frac{k}{m} v} = \int dt,$$

from which it follows that

$$-\frac{k}{m} \ln\left(g - \frac{k}{m} v\right) = t + C_0$$
$$\text{or,} \quad \ln\left(g - \frac{k}{m} v\right) = -\frac{mt}{k} + C_1.$$

Then

$$g - \frac{k}{m} v = C_2\, e^{-\frac{mt}{k}},$$

and

$$v(t) = \frac{mg}{k} - C_3 e^{-\frac{mt}{k}}.$$

The initial condition is satisfied if $\frac{mg}{k} - C_3 = v_0$ so, finally,

$$v(t) = \frac{mg}{k} - \left(\frac{mg}{k} - v_0\right) e^{-\frac{mt}{k}} = v_\infty - (v_\infty - v_0) e^{-\frac{mt}{k}},$$

where we introduce the notation, $v_\infty = \frac{mg}{k}$ for the terminal velocity. This is an explicit formula for the solution curves that were described above. Note that $v(t) = x'(t)$ where $x(t)$ denotes the position as a function of time. Then this position function can be obtained by integrating the function $v(t)$ with respect to t.

Exercises-

1. Find $x(t)$ if the initial position is $x(0) = 0.$.
2. Discuss how you might solve for k if the parameters m, g are assumed known. Can you think of an experiment for finding v_∞? If so, then k can be found from the knowledge of v_∞?
3. Choose simple numerical values for $\frac{m}{k}, v_\infty,$ and v_0 and plot $v(t)$ versus t. Try both $v_\infty > v_0$ and $v_\infty < v_0$ and imagine what the curve is describing.

5.2. A Dissolving Pill

A spherical pill is dropped into a fluid where it begins to dissolve at a rate that is proportional to the surface area exposed to the fluid. We can express this last statement in mathematical terms by writing

$$\frac{d}{dt}\left(\text{Volume}\right) = -K\left(\text{Surface Area}\right) \qquad (2.1)$$

where K denotes a positive constant of proportionality and the minus sign indicates that the volume decreases as long as the surface area is positive. For a sphere, we have

$$V = \frac{4}{3}\pi R(t)^3 \qquad \text{and} \qquad SA = 4\pi R(t)^2$$

where $R(t)$ denotes the radius of the sphere (which varies with t). Then

$$\frac{d}{dt}\left(\text{Volume}\right) = \frac{d}{dt}\left(\frac{4}{3}\pi R(t)^3\right) = 4\pi R(t)^2 R'(t)$$

and

$$4\pi R(t)^2 R'(t) = -K\left(4\pi R(t)^2\right),$$

$$\text{i.e.,} \qquad R'(t) = -K.$$

This is easily solved to obtain, $R(t) = -Kt + C$ and then $R(t) = R_0 - Kt,$ where R_0 denotes the initial radius of the pill. Then it is easy to find the time for the pill to completely dissolve if K is known. How might you go about finding K from an experiment?

5.3. Water Heating Strategy

Suppose a cylindrical water heater contains a heating element that electrically heats the water in the tank . We have already seen this model, which we can write in the form

$$\frac{d}{dt}T(t) = k(T_f - T(t)),$$

where T_f denotes the temperature to which the water will be eventually heated if the heating element is left on constantly, and k is a positive constant that reflects the proportionality between the time rate of change in the water temperature and the difference between T_f and $T(t).$ Note that as long as $T(t)$ is less than $T_f,$ $T'(t)$ is positive so the temperature of the water increases. We must also account for the fact that there is heat lost to the surroundings which we will assume is at a uniform temperature we will denote by $S.$ This loss will produce a rate of change of temperature proportional to the difference between the tank temperature and the temperature of the surroundings. That is,

$$\frac{d}{dt}T(t) = K(S - T(t))$$

Here K denotes a positive constant so this equation asserts that as long as $T(t)$ is greater than $S,$ the temperature of the water will decrease. Now the water in the tank is being simultaneously heated by the heating element and cooled by loss of heat to the surroundings so our equation that reflects this must have the form,

$$\begin{aligned}\frac{d}{dt}T(t) &= K(S - T(t)) + k(T_f - T(t)) \\ &= KS + kT_f - (k+K)T(t).\end{aligned}$$

Initially we have, $T(0) = T_{in},$ where T_{in} is the temperature at which the water enters the tank. It will be more convenient if we write this equation in the form

$$\frac{d}{dt}T(t) = (k+K)(M - T(t)) \qquad (3.1)$$

where

$$M = \frac{KS + kT_f}{k + K}.$$

We know that we can solve this equation by writing

$$\int \frac{dT}{M - T} = (k + K)\int dt$$

which leads to

$$-\ln(M - T(t)) = (k + K)t + C_0$$

and

$$T(t) = M + C_1 e^{-(k+K)t}.$$

It follows from the initial condition that

$$T(t) = M + (T_{in} - M)e^{-(k+K)t}. \qquad (3.2)$$

This solution predicts that the temperature will increase from T_{in} toward the limiting value M at an exponential rate that is determined by the constants k and K. Note that M is a temperature between T_f and S and that M is close to T_f if K is very small, while M is closer to S if k is small compared to $K.$ A small value of K corresponds to a tank that is well insulated to there is little heat loss to the surroundings.

Now suppose that when the water temperature reaches some value $T_{\max} < M,$ we turn off the water heater. Note that we can use the solution for $T(t)$ to determine the time t_a at which the water temperature reaches $T_{\max};$ i.e. we solve

$$T(t_a) = M + (T_{in} - M)e^{-(k+K)t_a} = T_{\max}.$$

for $t_a.$ This gives

$$e^{-(k+K)t_a} = \frac{T_{\max} - M}{T_{in} - M}$$

and

$$\begin{aligned} t_a &= \frac{-1}{k + K}\ln\left(\frac{T_{\max} - M}{T_{in} - M}\right) \\ &= \frac{1}{k + K}\ln\left(\frac{M - T_{in}}{M - T_{\max}}\right). \end{aligned}$$

Since the heater is turned off, the previous differential equation (3.1) is replaced by the equation

$$\frac{d}{dt}T(t) = K(S - T(t)) \qquad \text{with} \qquad T(t_a) = T_{\max} \qquad (3.3)$$

This describes how the water will gradually cool from a temperature of $T_{\max}$, as a result of heat loss to the exterior. The solution of this new equation is found in the usual way to be,

$$T(t) = S + (T_{\max} - S)e^{-K(t - t_a)}$$

We can now suppose that when the temperature of the water reaches some selected temperature $T_{\min}$ with $S < T_{\min} < T_{\max},$ we turn the heater back on again. The time t_b when the temperature reaches $T_{\min}$ is obtained by solving

$$T(t_b) = S + (T_{\max} - S)e^{-K(t_b - t_a)} = T_{\min}$$

to get

$$e^{-K(t_b - t_a)} = \frac{T_{\min} - S}{T_{\max} - S}$$

$$or \qquad t_b = t_a - \frac{1}{K} \ln\left(\frac{T_{\min} - S}{T_{\max} - S} \right)$$

We now solve (3.1) on the interval $t > t_b$ with the initial condition, $T(t_b) = T_{\min}$. The water temperature will increase toward M and we can then turn the heater off again when the temperature reaches $T_{\max}$ at some time $t_c > t_b$. In this way, the temperature is maintained between limits $T_{\min}$ and $T_{\max}$ without the heater having to be on constantly.

5.4. Absorption of Medications

When you take a pill to obtain medication, the pill first goes into your stomach and the medication passes into your GI tract. From there the medication is absorbed into your bloodstream and circulated through your body before being eliminated from the blood by the kidneys and other organs. If we let $x(t)$ denote the amount of medication in your GI tract at time t, then we can model the movement of the medication out of the GI tract with the equation

$$x'(t) = -k_1 x(t), \qquad x(0) = A. \qquad (4.1)$$

This is the assertion that after taking the pill, an amount A of medication is in the GI tract and it decreases at a rate proportional to the amount currently present in the GI tract. If the amount of medication in the bloodstream at time t is denoted by $y(t)$, then

$$y'(t) = k_1 x(t) - k_2 y(t), \qquad y(0) = 0, \quad (4.2)$$

expresses the fact that medication is coming into the bloodstream at exactly the rate it is leaving the GI tract and it is leaving the bloodstream at some rate expressed by the proportionality constant k_2. Also, we are assuming that there is no medication in the bloodstream initially. Now this is two equations for the two unknown functions $x(t)$ *and* $y(t)$, but the first equation can be solved independently and the solution substituted into the equation (4.2).

The solution of (4.1) is easily found to be

$$x(t) = A e^{-k_1 t},$$

and then

$$y'(t) + k_2 y(t) = k_1 x(t) = k_1 A e^{-k_1 t}.$$

We will find a particular solution for the y-equation by the method of undetermined coefficients. We guess that $y_p(t) = a e^{-k_1 t}$ and substituting this guess into the differential equation, we find

$$\begin{aligned} y_p'(t) + k_2 y_p(t) &= -k_1 a e^{-k_1 t} + k_2 a e^{-k_1 t} \\ &= k_1 A e^{-k_1 t}. \end{aligned}$$

This leads to

$$a = \frac{k_1 A}{k_2 - k_1} \qquad and \qquad y_p(t) = \frac{k_1 A}{k_2 - k_1} e^{-k_1 t}.$$

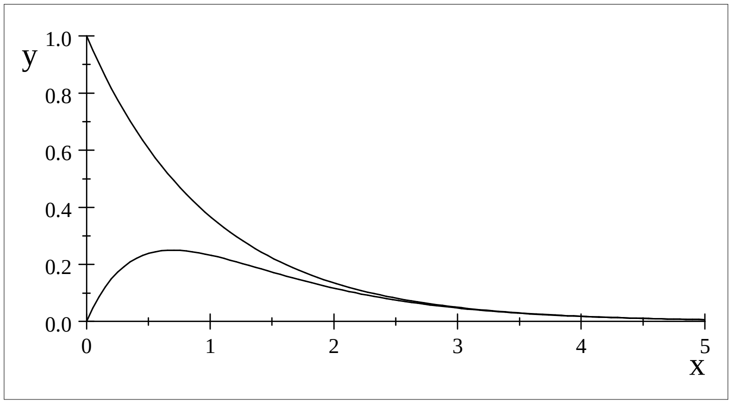
Then

$$y(t) = C e^{-k_2 t} + \frac{k_1 A}{k_2 - k_1} e^{-k_1 t}$$

and, using the initial condition to evaluate C, we get

$$y(t) = \frac{k_1 A}{k_2 - k_1}[\, e^{-k_1 t} - e^{-k_2 t}]$$

Plotting $x(t)$and $y(t)$ versus t for some representative values of the constants gives the following figure



$x(t)$ and $y(t)$ vs t

If we wish to consider a model that represents taking a continuously acting pill, (a pill that releases medication continuously so as to maintain a constant level of medication in the GI tract for a sustained period of time), we might modify the previous model to read

$$x'(t) = X_0 - k_1 x(t), \qquad x(0) = 0.$$

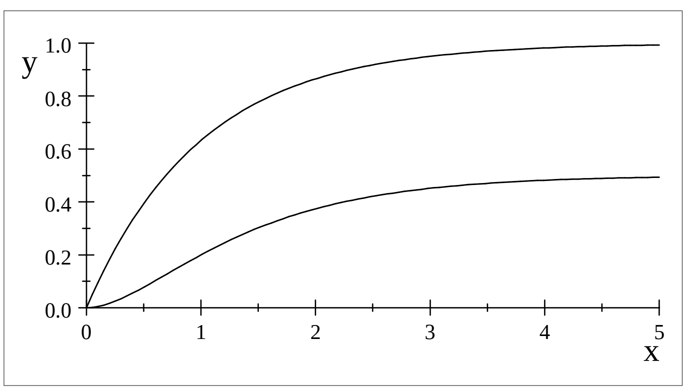
$$y'(t) = k_1 x(t) - k_2 y(t), \qquad y(0) = 0.$$

In this case, we find

$$x(t) = \frac{X_0}{k_1}(1 - e^{-k_1 t})$$

$$y(t) = \frac{1}{k_2}\left[1 + \frac{1}{k_2 - k_1}(k_1 e^{-k_2 t} - k_2 e^{-k_1 t})\right]$$

Plotting these solutions gives,



$x(t)$ and $y(t)$ vs t

Clearly this produces a longer constant level of medication in the bloodstream. Of course, eventually, the level of medication in the GI tract will go to zero and the level in the bloodstream will then also decrease to zero. A model which could describe the periodic taking of a sequence of time release pills would look like

$$x'(t) = X_0(t) - k_1 x(t), \qquad x(0) = 0,$$
$$y'(t) = k_1 x(t) - k_2 y(t), \qquad y(0) = 0,$$

where $X_0(t)$ denotes a piecewise constant function that alternates between a positive value and zero. We will discuss an effective way to solve equations involving such terms when we discuss the Laplace transform.