4

Separable First-Order Equations

As we will see below, the notion of a differential equation being "separable" is a natural generalization of the notion of a first-order differential equation being directly integrable. What's more, a fairly natural modification of the method for solving directly integrable first-order equations gives us the basic approach to solving "separable" differential equations. However, it cannot be said that the theory of separable equations is just a trivial extension of the theory of directly integrable equations. Certain issues can arise that do not arise in solving directly integrable equations. Some of these issues are pertinent to even more general classes of first-order differential equations than those that are just separable, and may play a role later on in this text.

In this chapter we will, of course, learn how to identify and solve separable first-order differential equations. We will also see what sort of issues can arise, examine those issues, and discuss some ways to deal with them. Since many of these issues involve graphing, we will also draw a bunch of pictures.

4.1 Basic Notions

Separability

A first-order differential equation is said to be *separable* if, after solving it for the derivative,

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

the right-hand side can then be factored as "a formula of just x" times "a formula of just y",

$$F(x, y) = f(x)g(y) \quad .$$

If this factoring is not possible, the equation is not separable.

More concisely, a first-order differential equation is *separable* if and only if it can be written as

$$\frac{dy}{dx} = f(x)g(y) \tag{4.1}$$

where f and g are known functions.

!►*Example 4.1:* *Consider the differential equation*

$$\frac{dy}{dx} - x^2y^2 = x^2 \quad . \tag{4.2}$$

Solving for the derivative (by adding $x^2 y^2$ to both sides),

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

and then factoring out the x^2 on the right-hand side gives

$$\frac{dy}{dx} = x^2 \left(1 + y^2\right) \quad ,$$

which is in form

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

with

$$f(x) = \underbrace{x^2}_{\text{no } y\text{'s}} \qquad \text{and} \qquad g(y) = \underbrace{\left(1 + y^2\right)}_{\text{no } x\text{'s}} \quad .$$

So equation (4.2) is a separable differential equation.

!►Example 4.2: *On the other hand, consider*

$$\frac{dy}{dx} - x^2 y^2 = 4 \quad . \tag{4.3}$$

Solving for the derivative here yields

$$\frac{dy}{dx} = x^2 y^2 + 4 \quad .$$

The right-hand side of this clearly cannot be factored into a function of just x times a function of just y. Thus, equation (4.3) is not separable.

We should (briefly) note that any directly integrable first-order differential equation

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

can be viewed as also being the separable equation

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

with $g(y)$ being the constant 1. Likewise, a first-order autonomous differential equation

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

can also be viewed as being separable, this time with $f(x)$ being 1. Thus, both directly integrable and autonomous differential equations are all special cases of separable differential equations.

Integrating Separable Equations

Observe that a directly-integrable equation

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

can be viewed as the separable equation

$$\frac{dy}{dx} = f(x)g(y) \qquad \text{with} \quad g(y) = 1 \quad .$$

We note this because the method used to solve directly-integrable equations (integrating both sides with respect to x) is rather easily adapted to solving separable equations. Let us try to figure out this adaptation using the differential equation from the first example. Then, if we are successful, we can discuss its use more generally.

!►*Example 4.3:* *Consider the differential equation*

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

In example 4.1 we saw that this is a separable equation, and can be written as

$$\frac{dy}{dx} = x^2\left(1 + y^2\right) \quad .$$

If we simply try to integrate both sides with respect to x *, the right-hand side would become*

$$\int x^2\left(1 + y^2\right) dx \quad ,$$

Unfortunately, the y *here is really* $y(x)$ *, some unknown formula of* x *; so the above is just the integral of some unknown function of* x *— something we cannot effectively evaluate. To eliminate the* y*'s on the right-hand side, we could, before attempting the integration, divide through by* $1 + y^2$ *, obtaining*

$$\frac{1}{1 + y^2}\frac{dy}{dx} = x^2 \quad . \tag{4.4}$$

The right-hand side can now be integrated with respect to x *. What about the left-hand side? The integral of that with respect to* x *is*

$$\int \frac{1}{1 + y^2}\frac{dy}{dx}\,dx \quad .$$

Tempting as it is to simply "cancel out the dx*'s", let's not (at least, not yet). After all,* ${}^{dy}/_{dx}$ *is not a fraction; it denotes the derivative* $y'(x)$ *where* $y(x)$ *is some unknown formula of* x *. But* y *is also shorthand for that same unknown formula* $y(x)$ *. So this integral is more precisely written as*

$$\int \frac{1}{1 + [y(x)]^2}\,y'(x)\,dx \quad .$$

Fortunately, this is just the right form for applying the generic substitution $y = y(x)$ *to convert the integral with respect to* x *to an integral with respect to* y *. No matter what* $y(x)$ *might be (so long as it is differentiable), we know*

$$\int \underbrace{\frac{1}{1 + [y(x)]^2}}_{\frac{1}{1+y^2}}\,\underbrace{y'(x)\,dx}_{dy} = \int \frac{1}{1 + y^2}\,dy \quad .$$

Combining all this, we get

$$\int \frac{1}{1+y^2}\frac{dy}{dx}\,dx \;=\; \int \frac{1}{1+[y(x)]^2}\,y'(x)\,dx \;=\; \int \frac{1}{1+y^2}\,dy \quad ,$$

which, after cutting out the middle, reduces to

$$\int \frac{1}{1+y^2}\frac{dy}{dx}\,dx \;=\; \int \frac{1}{1+y^2}\,dy \quad ,$$

the very equation we would have obtained if we had yielded to temptation and naively "cancelled out the dx's".

Consequently, the equation obtained by integrating both sides of equation (4.4) with respect to x,

$$\int \frac{1}{1+y^2}\frac{dy}{dx}\,dx \;=\; \int x^2\,dx \quad ,$$

is the same as

$$\int \frac{1}{1+y^2}\,dy \;=\; \int x^2\,dx \quad .$$

Doing the indicated integration on both sides then yields

$$\arctan(y) \;=\; \frac{1}{3}x^3 \;+\; c \quad ,$$

which, in turn, tells us that

$$y \;=\; \tan\left(\frac{1}{3}x^3 + c\right) \quad .$$

This is the general solution to our differential equation.

Two generally useful ideas were illustrated in the last example. One is that, whenever we have an integral of the form

$$\int H(y)\,\frac{dy}{dx}\,dx$$

where y denotes some (differentiable) function of x, then this integral is more properly written as

$$\int H(y(x))\,y'(x)\,dx \quad ,$$

which reduces to

$$\int H(y)\,dy$$

via the substitution $y = y(x)$ (even though we don't yet know what $y(x)$ is). Thus, in general,

$$\int H(y)\,\frac{dy}{dx}\,dx \;=\; \int H(y)\,dy \quad , \tag{4.5}$$

This equation is true whether you derive it rigorously, as we have, or obtain it naively by mechanically canceling out the dx's.[1]

The other idea seen in the example was that, if we divide an equation of the form

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

[1] One of the reasons our notation is so useful is that naive manipulations of the differentials often do lead to valid equations. Just don't be too naive and cancel out the d's in $^{dy}/_{dx}$.

by $g(y)$, then (with the help of equation (4.5)) we can compute the integral with respect to x of each side of the resulting equation,

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

This leads us to a *basic procedure for solving separable first-order differential equations* :

1. Get the differential equation into the form

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

2. Divide through by $g(y)$ to get

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

(Note: At this point we've "separated the variables", getting all the y's and its derivatives on one side, and all the x's on the other.)

3. Integrate both sides with respect to x , making use of the fact that

$$\int \frac{1}{g(y)}\frac{dy}{dx}\,dx = \int \frac{1}{g(y)}\,dy \quad .$$

4. Solve the resulting equation for y .

There are a few issues that can arise in some of these steps, and we will have to slightly refine this procedure to address those issues. Before doing that, though, let us practice with another differential equation for which the above approach can be applied without any difficulty.

!►Example 4.4: *Consider solving the initial-value problem*

$$\frac{dy}{dx} = -\frac{x}{y-3} \qquad \textit{with} \quad y(0) = 1 \quad .$$

Here,

$$\frac{dy}{dx} = f(x)g(y) \qquad \textit{with} \quad f(x) = -x \quad \textit{and} \quad g(y) = \frac{1}{y-3} \quad ,$$

and "dividing through by $g(y)$ *" is the same as multiplying through by* $y-3$ *. Doing so, and then integrating both sides with respect to* x *, we get the following:*

$$[y-3]\frac{dy}{dx} = -x$$

$$\hookrightarrow \qquad \int [y-3]\frac{dy}{dx}\,dx = -\int x\,dx$$

$$\hookrightarrow \qquad \int [y-3]\,dy = -\int x\,dx$$

$$\hookrightarrow \qquad \frac{1}{2}y^2 - 3y = -\frac{1}{2}x^2 + c \quad .$$

Though hardly necessary, we can multiply through by 2 *, obtaining the slightly simpler expression*

$$y^2 - 6y = -x^2 + 2c \quad .$$

We are now faced with the less-than-trivial task of solving the last equation for y *in terms of* x *. Since the left-hand side looks something like a quadratic for* y*, let us rewrite this equation as*

$$y^2 - 6y + \left[x^2 - 2c\right] = 0$$

so that we can apply the quadratic formula to solve for y *. Applying that venerable formula, we get*

$$y = \frac{-(-6) \pm \sqrt{(-6)^2 - 4\left[x^2 - 2c\right]}}{2} = 3 \pm \sqrt{9 - x^2 + 2c} \quad ,$$

which, since $9 + 2c$ *is just another unknown constant, can written a little more simply as*

$$y = 3 \pm \sqrt{a - x^2} \quad . \tag{4.6}$$

This is the general solution to our differential equation.

Now for the initial-value problem: Combining the general solution just derived with the given initial value at $x = 0$ *yields*

$$1 = y(0) = 3 \pm \sqrt{a - 0^2} = 3 \pm \sqrt{a} \quad .$$

So

$$\pm\sqrt{a} = -2 \quad .$$

This means that $a = 4$*, and that we must use the* negative *root in formula (4.6) for* y *. Thus, the solution to our initial-value problem is*

$$y = 3 - \sqrt{4 - x^2} \quad .$$

4.2 Constant Solutions
Avoiding Division by Zero

In the above procedure for solving

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

we divided both sides by $g(y)$. This requires, of course, that $g(y)$ not be zero — which is often *not* the case.

!►*Example 4.5:* *Consider solving*

$$\frac{dy}{dx} = 2x(y - 5) \quad .$$

As long as $y \neq 5$*, we can divide through by* $y - 5$ *and follow our basic procedure:*

$$\frac{1}{y-5}\frac{dy}{dx} = 2x$$

$$\hookrightarrow \qquad \int \frac{1}{y-5}\frac{dy}{dx}\,dx = \int 2x\,dx$$

$$\hookrightarrow \qquad \int \frac{1}{y-5}\,dy = \int 2x\,dx$$

$$\hookrightarrow \qquad \ln|y-5| = x^2 + c$$

$$\hookrightarrow \qquad |y-5| = e^{x^2+c} = e^{x^2}e^c$$

$$\hookrightarrow \qquad y-5 = \pm e^{x^2}e^c \quad .$$

So, assuming $y \neq 5$, we get

$$y = 5 \pm e^c e^{x^2} \quad .$$

Notice that, because $e^c \neq 0$ for every real value c, this formula for y never gives us $y = 5$ for any real choice of c and x.

But what about the case where $y = 5$?

Well, suppose $y = 5$. To be more specific, let y be the constant function

$$y(x) = 5 \qquad \text{for every} \quad x \quad ,$$

and plug this constant function into our differential equation

$$\frac{dy}{dx} = 2x(y-5) \quad .$$

Recalling (again) that derivatives of constants are zero, we get

$$0 = 2x(5-5) \quad ,$$

which is certainly a true equation. So $y = 5$ is a solution. In fact, it is one of those "constant" solutions we discussed in the previous chapter.

Combining all the above, we see that the "general solution" to the given differential equation is actually the set consisting of the solutions

$$y(x) = 5 \qquad \text{and} \qquad y(x) = 5 \pm e^c e^{x^2} \quad .$$

Now consider the general case, where we seek all possible solutions to

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

If y_0 is any single value for which

$$g(y_0) = 0 \quad ,$$

then plugging the corresponding constant function

$$y(x) = y_0 \qquad \text{for all} \quad x$$

into the differential equation gives, after a trivial bit of computation,

$$0 = 0 \quad ,$$

showing that

$$y(x) = y_0 \quad \text{is a constant solution to} \quad \frac{dy}{dx} = f(x)g(y) \quad ,$$

just as we saw (in the above example) that

$$y(x) = 5 \quad \text{is a constant solution to} \quad \frac{dy}{dx} = 2x(y-5) \quad .$$

Conversely, suppose $y = y_0$ is a constant solution to

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

(and f is not the zero function). Then the equation is valid with y replaced by the constant y_0, giving us

$$0 = f(x)g(y_0) \quad ,$$

which, in turn, means that y_0 must be a constant such that

$$g(y_0) = 0 \quad .$$

What all this shows is that our basic method for solving separable equations may miss the constant solutions because those solutions correspond to a division by zero in our basic method.[2]

Because constant solutions are often important in understanding the physical process the differential equation might be modeling, let us be careful to not miss them. Accordingly, we will insert the following step into our procedure on page 77 for solving separable equations:

- Identify all constant solutions by finding all values y_0, y_1, y_2, ... such that

  $$g(y_k) = 0 \quad ,$$

  and then write down

  $$y(x) = y_0 \quad , \quad y(x) = y_1 \quad , \quad y(x) = y_2 \quad , \quad \ldots \quad .$$

  (These are the constant solutions.)

(And we will renumber the other steps as appropriate.)

Sometimes, the formula obtained by our basic procedure for solving can be 'tweaked' to also account for the constant solutions. A standard 'tweak' can be seen by reconsidering the general solution obtained in our last example.

!►Example 4.6: *The general solution obtained in the previous example was the set containing*

$$y(x) = 5 \qquad \text{and} \qquad y(x) = 5 \pm e^c e^{x^2} \quad ,$$

If we let $A = \pm e^c$, the second equation reduces to

$$y(x) = 5 + Ae^{x^2} \quad .$$

Remember, though, $A = \pm e^c$ can be any positive or negative number, but cannot be zero (because of the nature of the exponential function). So, by our definition of A, our general solution is

$$y(x) = 5 \tag{4.7a}$$

[2] Because $g(y_0) = 0$ is a 'singular' value for division, many authors refer to constant solutions of separable equations as *singular* solutions.

and

$$y(x) = 5 + Ae^{x^2} \quad \text{where } A \text{ can be any nonzero real number.} \tag{4.7b}$$

However, if we allow A to be zero, then equation (4.7b) reduces to equation (4.7a),

$$y(x) = 5 + 0 \cdot e^{x^2} = 5 \quad ,$$

which means the entire set of possible solutions can be expressed more simply as

$$y(x) = 5 + Ae^{x^2}$$

where A is an arbitrary constant with no restrictions on its possible values.

In the future, we will usually express our general solutions as simply as practical, with the trick of letting

$$A = \pm e^c \text{ or } 0$$

often being used without comment. Keep in mind, though, that the sort of tweaking just described is not always possible.

?►*Exercise 4.1:* *Verify that the general solution to*

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

is given by the set consisting of

$$y(x) = 0 \qquad \text{and} \qquad y(x) = \frac{1}{x+c} \quad .$$

Is there anyway to rewrite these two formula for $y(x)$ as a single formula using just one arbitrary constant?

The Importance of Constant Solutions

Even if we can use the same general formula to describe all the solutions (constant and otherwise), it is often worthwhile to explicitly identify any constant solutions. To see this, let us now solve the differential equation from chapter 1 describing a falling object when we take into account air resistance.

!►*Example 4.7:* *Let $v = v(t)$ be the velocity (in meters per second) at time t of some object of mass m plummeting towards the ground. In chapter 1, we decided that F_{air}, the force of air resistance acting on the falling body, could be described by*

$$F_{\text{air}} = -\gamma v$$

where γ was some positive constant dependent on the size and shape of the object (and probably determined by experiment). Using this, we obtained the differential equation

$$\frac{dv}{dt} = -9.8 - \kappa v \qquad \text{where} \quad \kappa = \frac{\gamma}{m} \quad .$$

This is a relatively simple separable equation. Assuming v *equals a constant* v_0 *yields*

$$0 = -9.8 - \kappa v_0 \quad \Longrightarrow \quad v_0 = -\frac{9.8}{\kappa} = -\frac{9.8m}{\gamma} \quad .$$

So, we have one constant solution,

$$v(t) = v_0 \qquad \text{for all} \quad t$$

where

$$v_0 = -\frac{9.8}{\kappa} = -\frac{9.8m}{\gamma} \quad .$$

For reasons that will soon become clear, v_0 *is called the* terminal velocity *of the object that is falling.*

To find the other possible solutions, we assume $v \neq v_0$ *and proceed:*

$$\frac{dv}{dt} = -9.8 - \kappa v$$

$$\hookrightarrow \qquad \frac{1}{9.8 + \kappa v}\frac{dv}{dt} = -1$$

$$\hookrightarrow \qquad \int \frac{1}{9.8 + \kappa v}\frac{dv}{dt}\,dt = -\int 1\,dt$$

$$\hookrightarrow \qquad \int \frac{1}{9.8 + \kappa v}\,dv = -\int dt$$

$$\hookrightarrow \qquad \frac{1}{\kappa}\ln|9.8 + \kappa v| = -t + c$$

$$\hookrightarrow \qquad \ln|9.8 + \kappa v| = -\kappa t + \kappa c$$

$$\hookrightarrow \qquad 9.8 + \kappa v = \pm e^{-\kappa t + \kappa c}$$

$$\hookrightarrow \qquad v(t) = \frac{1}{\kappa}\left[-9.8 \pm e^{\kappa c} e^{-\kappa t}\right] \quad .$$

Since $v_0 = -9.8\kappa^{-1}$ *, the last equation reduces to*

$$v(t) = v_0 + Ae^{-\kappa t} \qquad \text{where} \quad A = \pm\frac{1}{\kappa}e^{\kappa c} \quad .$$

This formula for $v(t)$ *yields the constant solution,* $v = v_0$ *, if we allow* $A = 0$ *. Thus, letting* A *be a completely arbitrary constant, we have that*

$$v(t) = v_0 + Ae^{-\kappa t} \tag{4.8a}$$

where

$$v_0 = -\frac{9.8m}{\gamma} \qquad \text{and} \qquad \kappa = \frac{\gamma}{m} \tag{4.8b}$$

describes all possible solutions to the differential equation of interest here. The graphs of some possible solutions (assuming a terminal velocity of -10 meters/second) are sketched in figure 4.1.

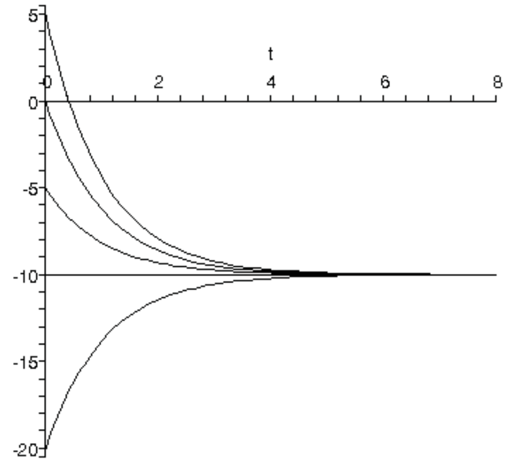

Figure 4.1: Graphs of the velocity of a falling object during the first 8 seconds of its fall assuming a terminal velocity of -10 meters per second. Each graph corresponds to a different initial velocity.

Notice how the constant in the constant solution, v_0*, appears in the general solution (equation (4.8a)). More importantly, notice that the exponential term in this solution rapidly goes to zero as* t *increases, so*

$$v(t) = v_0 + Ae^{-\kappa t} \quad \rightarrow \quad v(t) = v_0 \qquad \text{as} \quad t \rightarrow \infty \quad .$$

This is graphically obvious in figure 4.1. Consequently, no matter what the initial velocity and initial height were, eventually the velocity of this falling object will be very close to v_0 *(provided it doesn't hit the ground first). That is why* v_0 *is called the terminal velocity. That is also why that constant solution is so important here (and is appropriately also called the equilibrium solution). It accurately predicts the final velocity of any object falling from a sufficiently high height. And if you are that falling object, then that velocity*[3] *is probably a major concern.*

4.3 Explicit Versus Implicit Solutions

Thus far, we have been able to find explicit formulas for all of our solutions; that is, we have been able to carry out the last step in our basic procedure — that of solving the resulting (integrated) equation for y in terms of x — obtaining

$$y = y(x) \qquad \text{where } y(x) \text{ is some formula of } x \text{ (with no } y\text{'s).}$$

For example, as the general solution to

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

[3] between 120 and 150 miles per hour for a typical human body

we obtained (in example 4.3)

$$y = \underbrace{\tan\left(\frac{1}{3}x^3 + c\right)}_{y(x)} \quad .$$

Unfortunately, this is not always possible.

!►Example 4.8: *Consider*

$$\frac{dy}{dx} = \frac{x+1}{8 + 2\pi\sin(\pi y)} \quad .$$

In this case,

$$g(y) = \frac{1}{8 + 2\pi\sin(\pi y)} \quad ,$$

which can never be zero. So there are no constant solutions, and we can blithely proceed with our procedure. Doing so:

$$\frac{dy}{dx} = \frac{x+1}{8 + 2\pi\sin(\pi y)}$$

$$\hookrightarrow \quad [8 + 2\pi\sin(\pi y)]\frac{dy}{dx} = x + 1$$

$$\hookrightarrow \quad \int [8 + 2\pi\sin(\pi y)]\frac{dy}{dx}\,dx = \int x + 1\,dx$$

$$\hookrightarrow \quad \int [8 + 2\pi\sin(\pi y)]\,dy = \int x + 1\,dx$$

$$\hookrightarrow \quad 8y - 2\cos(\pi y) = \frac{1}{2}x^2 + x + c \quad .$$

The next step would be to solve the last equation for y in terms of x. But look at that last equation. Can you solve it for y as a formula of x? Neither can anyone else. So we are not able to obtain an explicit formula for y. At best, we can say that $y = y(x)$ satisfies the equation

$$8y - 2\cos(\pi y) = \frac{1}{2}x^2 + x + c \quad .$$

Still, this equation is not without value. It does implicitly describe the possible relations between x and y. In particular, the graphs of this equation can be sketched for different values of c (we'll do this later on in this chapter). These graphs, in turn, give you the graphs you would obtain for $y(x)$ if you could actually find the formula for $y(x)$.

In practice, we must deal with both "explicit" and "implicit" solutions to differential equations. When we have an explicit formula for the solution in terms of the variable, that is, we have something of the form

$$y = y(x) \qquad \text{where } y(x) \text{ is some formula of } x \text{ (with no } y\text{'s)} \quad , \tag{4.9}$$

then we say that we have an *explicit solution* to our differential equation. Technically, it is that "formula of x" in equation (4.9) which is the *explicit solution*. In practice, though, it is common to refer to the entire equation as "an explicit solution". For example, we found that the solution to

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

is explicitly given by

$$y = \tan\left(\frac{1}{3}x^3 + c\right) \quad .$$

Strictly speaking, the explicit solution here is the formula

$$\tan\left(\frac{1}{3}x^3 + c\right) \quad .$$

That, of course, is what is really meant when someone answers the question

What is the explicit solution to $\quad \frac{dy}{dx} - x^2y^2 = x^2 \quad$?

with the equation

$$y = \tan\left(\frac{1}{3}x^3 + c\right) \quad .$$

If, on the other hand, we have an equation (other than something like (4.9)) involving the solution and the variable, then that *equation* is called an *implicit solution*. In trying to solve the differential equation in example 4.8,

$$\frac{dy}{dx} = \frac{x+1}{8 + 2\pi \sin(\pi y)} \quad ,$$

we derived the equation

$$8y - 2\cos(\pi y) = \frac{1}{2}x^2 + x + c \quad .$$

This equation is an implicit solution for the given differential equation.[4]

Differential equations — be they separable or not — can have both implicit and explicit solutions. Indeed, implicit solutions often arise in the process of deriving an explicit solution. For example, in solving

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

we first obtained

$$\arctan(y) = \frac{1}{3}x^3 + c \quad .$$

This is an implicit solution. Fortunately, it could be easily solved for y, giving us the explicit solution

$$y = \tan\left(\frac{1}{3}x^3 + c\right) \quad .$$

As a general rule, explicit solutions are preferred over implicit solutions. Explicit solutions usually give more information about the solutions, and are easier to use than implicit solutions (even when you have sophisticated computer math packages). So, whenever you solve a differential equation,

FIND AN EXPLICIT SOLUTION IF AT ALL PRACTICAL.

Do not be surprised, however, if you encounter a differential equation for which an explicit solution is not obtainable. This is not a disaster, it just means a little more work may be needed to extract useful information about the possible solutions.

---

[4] The fact that an explicit solution is a formula while an implicit solution is an equation may be a little confusing at first. If it helps, think of the phrase "implicit solution" as being shorthand for "an equation implicitly defining the solution $y = y(x)$".

4.4 The Full Procedure for Solving Separable Equations

In light of the possibility of singular solutions and the possibility of not finding explicit solutions, we should refine our procedure for solving a separable differential equation to:

1. Get the differential equation into the form

$$\frac{dy}{dx} = f(x)g(y) \quad . \tag{4.10}$$

2. Identify all constant solutions by finding all values y_0 , y_1 , y_2 , ... such that

$$g(y_k) = 0 \quad ,$$

and then write down

$$y(x) = y_0 \quad , \quad y(x) = y_1 \quad , \quad y(x) = y_2 \quad , \quad \ldots \quad .$$

(These are the constant solutions.)

3a. Divide equation (4.10) through by $g(y)$ to get

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

(assuming y is not one of the constant solutions just found).

b. Integrate both sides of the equation just obtained with respect to x .

c. Solve the resulting equation for y , if practical (thus obtaining an explicit solution). If not practical, use that resulting equation as an implicit solution, possibly rearranged or simplified if appropriate.

4. If constant solutions were found, see if the formulas obtained for the other solutions can be tweaked to also describe the constant solutions. In any case, be sure to write out all solution(s) obtained.

The above yields the general solution. If initial values are also given, then use those initial conditions with the general solution just obtained to derive the particular solutions satisfying the given initial-value problems.

4.5 Existence, Uniqueness, and False Solutions

On the Existence and Uniqueness of Solutions

Let's consider a generic initial-value problem involving a separable differential equation,

$$\frac{dy}{dx} = f(x)g(y) \qquad \text{with} \quad y(x_0) = y_0 \quad .$$

Letting $F(x, y) = f(x)g(y)$ this is

$$\frac{dy}{dx} = F(x, y) \qquad \text{with} \quad y(x_0) = y_0 \quad ,$$

which was the initial-value problem considered in theorem 3.1 on page 48. That theorem assures us that there is exactly one solution to our initial-value problem on some interval (a, b) containing x_0 provided

$$F(x, y) = f(x)g(y)$$

and

$$\frac{\partial F}{\partial y} = \frac{\partial}{\partial y}[f(x)g(y)] = f(x)g'(y)$$

are continuous in some open region containing the point (x_0, y_0). This means our initial-value problem will have exactly one solution on some interval (a, b) provided $f(x)$ is continuous on some open interval containing x_0, and both $g(y)$ and $g'(y)$ are continuous on some open interval containing y_0. In practice, this is typically what we have.

Typically, also, one rarely worries about the existence and uniqueness of the solution to an initial-value problem with a separable differential equation, at least not when one can carry out the integration and algebra required by our procedure. After all, doesn't our refined procedure for solving separable differential equations always lead us to "the solution"? Well, here are two reasons to have at least a little concern about existence and uniqueness:

1. After the integration in step 3, the resulting equation may involve a nontrivial formula of y. After applying the initial condition and solving for y, it is possible to end up with more than one formula for $y(x)$. But as long as f, g and g' are sufficiently continuous, the above tells us that there is only one solution. Thus, only one of these formulas for $y(x)$ can be correct. The others are "false solutions" that should be identified and eliminated. (An example is given in the next subsection.)

2. Suppose $g(y_0) = 0$. Our refined procedure tells us that the constant function $y = y_0$, which certainly satisfies the initial condition, is also a solution to the differential equation. So $y = y_0$ is immediately seen to be a solution to our initial-value problem. Do we then need to go through the rest of our procedure to see if any other solutions to the differential equation satisfy $y(x_0) = y_0$? The answer is *No, not if f is continuous on an open interval containing x_0, and both g and g' are continuous on an open interval containing y_0.* If that continuity holds, then the above analysis assures us that there is only one solution. Thus, if we find *a* solution, we have found *the* solution.

It is possible, to have an initial-value problem

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

in which the f or g or g' is not suitably continuous. The problem in exercise 3.5 on page 70,

$$\frac{dy}{dx} = 2\sqrt{y} \qquad \text{with} \quad y(1) = 0 \quad ,$$

is one such problem. Here,

$$y_0 = 0 \quad , \quad f(x) = 1 \quad , \quad g(y) = 2\sqrt{y} \quad \text{and} \quad g'(y) = \frac{1}{\sqrt{y}} \quad .$$

Clearly, g and, especially, g' are not continuous in any open interval containing $y_0 = 0$. So the above results on existence and uniqueness cannot be assumed. Indeed, in this case there is not just the one constant solution $y = 0$, but, as shown in that exercise, there are many different solutions, including

$$y(x) = \begin{cases} 0 & \text{if} \quad x < 1 \\ (x-1)^2 & \text{if} \quad 1 \leq x \end{cases} \qquad \text{and} \qquad y(x) = \begin{cases} 0 & \text{if} \quad x < 3 \\ (x-3)^2 & \text{if} \quad 3 \leq x \end{cases} \quad .$$

A Caution on False Solutions

It is always a good idea to verify that any 'solution' obtained in solving a differential equation really is a solution. This is even more true when solving separable differential equations. Not only does the extra algebra involved naturally increase the likelihood of human error, this algebra can, as noted above, lead to 'false solutions' — formulas that are obtained as solutions, but do not actually satisfy the original problem.

!►Example 4.9: *Consider the initial-value problem*

$$\frac{dy}{dx} = 2\sqrt{y} \qquad \text{with} \quad y(0) = 4 \quad .$$

The differential equation does have one constant solution, $y = 0$, but since that doesn't satisfy the initial condition, it hardly seems relevant. To find the other solutions, let's divide the differential equation by $\sqrt{y}$ and proceed with the basic procedure:

$$\frac{1}{\sqrt{y}}\frac{dy}{dx} = 2$$

$$\hookrightarrow \qquad \int \frac{1}{\sqrt{y}}\frac{dy}{dx}\,dx = \int 2\,dx$$

$$\hookrightarrow \qquad \int y^{-1/2}\,dy = \int 2\,dx$$

$$\hookrightarrow \qquad 2y^{1/2} = 2x + c \quad .$$

Dividing by 2 and squaring (and letting $a = c/2$), we get

$$y = (x+a)^2 \quad . \tag{4.11}$$

Plugging this into the initial condition, we obtain

$$4 = y(0) = (0+a)^2 = a^2 \quad ,$$

which means that

$$a = \pm 2 \quad .$$

Hence, we have two formulas for the solution to our initial-value problem,

$$y_+(x) = (x+2)^2 \qquad \text{and} \qquad y_-(x) = (x-2)^2 \quad .$$

Both satisfy the initial condition. Do both satisfy the differential equation

$$\frac{dy}{dx} = 2\sqrt{y} \quad ?$$

Well, plugging

$$y = y_{\pm}(x) = (x \pm 2)^2$$

into the differential equation yields

$$\frac{d}{dx}(x \pm 2)^2 = 2\sqrt{(x \pm 2)^2}$$

$$\hookrightarrow \qquad 2(x \pm 2) = 2\sqrt{(x \pm 2)^2} \quad .$$

So, for $y = y_{\pm}(x)$ *to be solutions to our differential equation, we must have*

$$x \pm 2 = \sqrt{(x \pm 2)^2} \tag{4.12}$$

for all values of x *'of interest'. In particular, this equation must be valid at the initial point* $x = 0$ *.*

So, consider what happens to equation (4.12) at the initial point $x = 0$ *. With* $y = y_{+}(x)$ *and* $x = 0$ *equation (4.12) becomes*

$$0 + 2 = \sqrt{(0+2)^2} = \sqrt{4} \quad ,$$

which, of course, simplifies to the perfectly acceptable equation

$$2 = 2.$$

But with $y = y_{-}(x)$ *and* $x = 0$ *we get*

$$0 - 2 = \sqrt{(0-2)^2} = \sqrt{4} = 2 \quad ,$$

which, of course, simplifies to

$$-2 = 2 \quad ,$$

which is not acceptable. So we can not accept $y = y_{-}(x)$ *as a solution to our initial-value problem. It was a false solution.*

While we are at it, let's look a little more closely at equation (4.12) with $y = y_{+}(x)$ *,*

$$x + 2 = \sqrt{(x+2)^2} \quad .$$

Remember, if A *is any real number, then*

$$\sqrt{A^2} = |A| \quad .$$

So equation (4.12) with $y = y_{+}$ *can be written as*

$$x + 2 = |x + 2| \quad ,$$

which is true if and only if $x + 2 \geq 0$ *(i.e.,* $x \geq -2$ *). This means that our solution,* $y = y_{+}(x)$ *, is not valid for all values of* x *, only for those greater than or equal to* -2 *. Thus, the actual solution that we have is*

$$y = y_{+}(x) = (x+2)^2 \qquad \text{for} \quad -2 \leq x \quad .$$

There was a lot of analysis done in the last example after obtaining the apparent solutions

$$y = (x \pm 2)^2 \quad .$$

Don't be alarmed. In most of the problems you will be given, verifying that your formula is a solution should be fairly easy. Still, take the moral of this example to heart: It is a good idea to verify that any formulas derived as solutions truly are solutions.

By the way, in a later chapter we will develop some graphical technics that would have simplified our work in the above example.

4.6 On the Nature of Solutions to Differential Equations

When we solve a first-order directly integrable differential equation,

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

we get something of the form

$$y = F(x) + c$$

where F is any antiderivative of f and c is an arbitrary constant. Computationally, all we have to do is find a single antiderivative F for f and then add an arbitrary constant. Thus, also, the graph of any possible solution is nothing more than the graph of $F(x)$ shifted vertically by the value of c (up if $c > 0$, down if $c < 0$). What's more, the interval for x over which

$$y = F(x) + c$$

is a valid solution depends only on the one function F. If $F(x)$ is continuous for all x in an interval (a, b), then (a, b) is a valid interval for our solution. This interval does not depend on the choice for c.

The situation can be much more complicated if our differential equation is not directly integrable. First of all, finding an explicit solution can be impossible. And consider those explicit general solutions we have found,

$$y = \tan\left(\frac{1}{3}x^3 + c\right) \qquad \text{(from example 4.3 on page 75)}$$

and

$$y = 3 \pm \sqrt{a - x^2} \qquad \text{(from example 4.4 on page 77)} \quad .$$

In both of these, the arbitrary constants are not simply "added on" to some formula of x. Instead, each solution formula combines the variable, x, with the arbitrary constant, c or a, in a very nontrivial manner. There are two immediate consequences of this:

1. The graphs of the solutions are no longer simply vertically shifted copies of some single function.

2. The possible intervals over which any solution is valid may depend on the arbitrary constant. And since the value of that constant can be determined by the initial condition, the interval of validity for our solutions may depend on the initial condition.

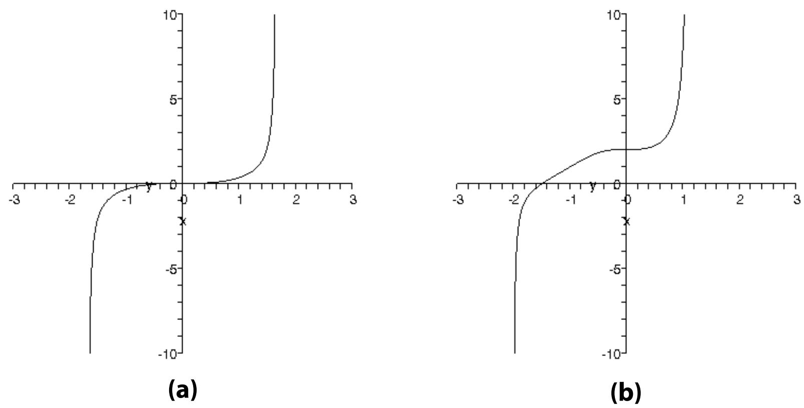

Figure 4.2: The graph of $y = \tan\left(\frac{1}{3}x^3 + c\right)$ **(a)** when $y(0) = 0$ and **(b)** when $y(0) = 2$.

Both of these consequences are illustrated in figure 4.2, in which the graphs of two solutions to the differential equation in example 4.3 have been sketched corresponding to two different initial values (namely, $y(0) = 0$ and $y(0) = 2$). In these figures you can see how changing the initial condition from $y(0) = 0$ to $y(0) = 2$ changes the interval over which the solution exists. Even more apparent is that the graph corresponding to $y(0) = 2$ is not merely a 'shift' of the graph corresponding to $y(0) = 0$; there is also a small but clear distortion in the shape of the graph.

The possible dependence of a solution's interval of validity is even better illustrated by the solutions obtained in example 4.4. There, the differential equation was

$$\frac{dy}{dx} = -\frac{x}{y-3}$$

and the general solution was found to be

$$y = 3 \pm \sqrt{a - x^2} \quad .$$

The arbitrary constant here, a , occurs in the square root. For this square root to be real, we must have

$$a - x^2 \geq 0 \quad .$$

That is,

$$-\sqrt{a} \leq x \leq \sqrt{a}$$

is the maximal interval over which

$$y = 3 + \sqrt{a - x^2} \qquad \text{and} \qquad y = 3 - \sqrt{a - x^2}$$

are valid solutions.

To properly indicate this dependence of the solution's possible domain on the arbitrary constant or the initial value, we should state the maximal interval of validity along with any formula or equation describing our solution(s). For example 4.4, that would mean writing the general solution as

$$y = 3 \pm \sqrt{a - x^2} \qquad \text{for all} \quad -\sqrt{a} \leq x \leq \sqrt{a} \quad .$$

When this is particularly convenient or noteworthy, we will attempt to remember to do so. Even when we don't, keep in mind that there may be limits as to the possible values of x, and that these limits may depend on the values assumed by the arbitrary constants.

By the way, notice also that the above a cannot be negative (otherwise, $\sqrt{a}$ will not be a real number). This points out that, in general, the 'arbitrary' constants appearing in general solutions are not always completely arbitrary.

4.7 Using and Graphing Implicit Solutions

Outside of courses specifically geared towards learning about differential equations, the main reason to solve an initial-value problem such as

$$\frac{dy}{dx} = \frac{x+1}{8+2\pi\sin(y\pi)} \qquad \text{with} \quad y(0) = 2$$

is so that we can predict what values $y(x)$ will assume when x has values other than 0. In practice, of course, $y(x)$ will represent something of interest (position, velocity, promises made, number of ducks, etc.) that varies with whatever x represents (time, position, money invested, food available, etc.). When the solution y is given explicitly by some formula $y(x)$, then those values are relatively easily obtained by just computing that formula for different values of x, and a picture of how $y(x)$ varies with x is easily obtained by graphing $y = y(x)$. If, instead, the solution is given implicitly by some equation, then the possible values of $y(x)$ for different x's, along with any graph of $y(x)$, must be extracted from that equation. It may be necessary to use advanced numerical methods to extract the desired information, but that should not be a significant problem — these methods are probably already incorporated into your favorite computer math package.

!►*Example 4.10:* *Let's consider the initial-value problem*

$$\frac{dy}{dx} = \frac{x+1}{8+2\pi\sin(y\pi)} \qquad \textit{with} \quad y(0) = 2 \quad .$$

In example 4.8, we saw that the general solution to the differential equation is given implicitly by

$$8y - 2\cos(y\pi) = \frac{1}{2}x^2 + x + c \quad . \tag{4.13}$$

The initial condition $y(0) = 2$ *tells us that* $y = 2$ *when* $x = 0$. *With this assumed, our implicit solution reduces to*

$$8\cdot 2 - 2\cos(2\pi) = \frac{1}{2}\left[0^2\right] + 0 + c \quad .$$

So

$$c = 8\cdot 2 - 2\cos(2\pi) - \frac{1}{2}\left[0^2\right] - 0 = 16 - 2 = 14 \quad .$$

Plugging this back into equation (4.13) gives

$$8y - 2\cos(y\pi) = \frac{1}{2}x^2 + x + 14 \tag{4.14}$$

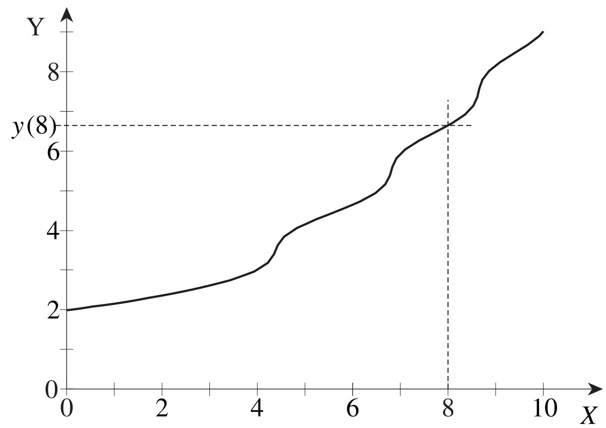

Figure 4.3: Graph of the implicit solution to the initial-value problem of example 4.10. (Graph created using Maple.)

as an implicit solution for our initial-value problem.

Replacing c *with* 14 *does not make it any easier for us to convert this equation relating* y *and* x *into a formula* $y(x)$ *for* y *. Still,* $y = y(x)$ *must satisfy equation (4.14), and the graph of that equation can be generated by invoking the appropriate command(s) in a suitable computer math package. That is how the graph in figure 4.3 was created. From this graph, we see that the value of* $y(8)$ *is between* 6 *and* 7 *. For a more precise determination of* $y(8)$ *, set* $x = 8$ *in equation (4.14). This gives us*

$$8y - 2\cos(y\pi) = \frac{1}{2}8^2 + 8 + 14 \quad ,$$

which, after a little arithmetic, reduces to

$$8y - 2\cos(y\pi) = 54 \quad .$$

Now apply some numerical method (such as the Newton-Raphson method for finding roots[5]*) to find, approximately, the corresponding value of* y *. Again, we need not do the tedious computations ourselves; we can go to our favorite computer math package, look up the appropriate commands, and let it compute that value for* y *. Doing so, we find that* $y(8) \approx 6.642079594$ *.*

Any curve that is at least part of the graph of an implicit solution for a differential equation is called an *integral curve* for the differential equation. Remember, this is the graph of an *equation*. If a *function* $y(x)$ is a solution to that differential equation, then $y = y(x)$ must also satisfy any equation serving as an implicit solution, and, thus, the graph of that $y(x)$ (which we will call a *solution curve*) must be at least a portion of one of the integral curves for that differential equation. Sometimes an integral curve will be a solution curve. That is "clearly" the case in figure 4.3, because that curve is "clearly" the graph of a function (more on that later).

Sometimes though, there are two (or more) different functions y_1 and y_2 such that both $y = y_1(x)$ and $y = y_2(x)$ satisfy the same equation for the same values of x . If that equation is an implicit solution to some differential equation, then its graph (the integral curve) will contain the graphs of both $y = y_1(x)$ and $y = y_2(x)$. In such a case, the integral curve is not a solution curve, but contains two or more solution curves.

[5] It should be in your calculus text.

To illuminate these comments, let us look at the solution curves and integral curves for one equation we've already solved. At the same time, we will discover that, at least occasionally, the use of implicit solutions can simplify our work, even when explicit solutions are available.

!►Example 4.11: *Consider graphing both all the solutions to*

$$\frac{dy}{dx} = -\frac{x}{y-3}$$

and the particular solution satisfying

$$y(0) = 1 \quad .$$

In example 4.4 (starting on page 77), we "separated and integrated" this differential equation to get the implicit solution

$$y^2 - 6y = -x^2 + c \quad . \tag{4.15}$$

We were then able to solve this equation for y in terms of x by using the quadratic formula.

This time, rather than attempting to solve for y, let's simply move the x^2 to the left, obtaining

$$x^2 + y^2 - 6y = 2c \quad .$$

This looks suspiciously like an equation for a circle. Writing 6 as $2 \cdot 3$ and adding 3^2 to both sides (to complete the square in the y terms) makes it look even more so:

$$x^2 + \underbrace{y^2 - 2 \cdot 3y + 3^2}_{(y-3)^2} = 2c + 3^2$$

$$\hookrightarrow \qquad (x-0)^2 + (y-3)^2 = 2c + 9 \quad .$$

Since the left-hand side is the sum of squares, it cannot be negative; hence, neither can the right-hand side. So we can let $R = \sqrt{2c + 9}$ and write our equation as

$$(x-0)^2 + (y-3)^2 = R^2 \quad . \tag{4.16}$$

You should recognize this implicit solution for our differential equation as also being the equation for a circle of radius R centered at $(0,3)$. One such circle (with $R = 2$) is sketched in figure 4.4a. These circles are integral curves for our differential equation. In this case, we can find the solution curves by solving our last equation for the explicit solutions

$$y = 3 \pm \sqrt{R^2 - x^2} \quad .$$

The solution curves, then, are the graphs of $y = y_-(x)$ and $y = y_+(x)$ where

$$y_+(x) = 3 + \sqrt{R^2 - x^2} \qquad \text{and} \qquad y_-(x) = 3 - \sqrt{R^2 - x^2} \quad .$$

Since we must have $R^2 - x^2 \geq 0$ for the square roots, each of these functions can only be defined for

$$-R \leq x \leq R \quad .$$

Observe that the graphs of these functions are not the entire circles of the integral curves, but are semicircles, with the graph of

$$y = 3 + \sqrt{R^2 - x^2} \qquad \text{with} \quad -R \leq x \leq R$$

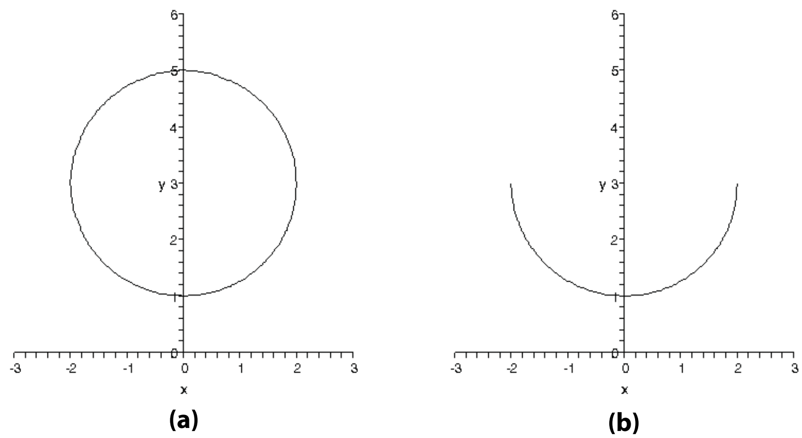

Figure 4.4: **(a)** The integral curve and **(b)** the solution curve for the differential equation in example 4.11 containing the point $(x, y) = (0, 1)$.

being the upper half of the circle of radius R *about* $(0, 3)$ *, and the graph of*

$$y = 3 - \sqrt{R^2 - x^2} \qquad \text{with} \qquad -R \leq x \leq R$$

being the lower half of that same circle.

If we further require that $y(0) = 1$ *, then implicit solution (4.16) becomes*

$$(0 - 0)^2 + (1 - 3)^2 = R^2 \quad .$$

So $R = 2$ *, and* $y = y(x)$ *must satisfy*

$$(x - 0)^2 + (y - 3)^2 = 2^2 \quad . \tag{4.17}$$

Solving this for y *in terms of* x *, we get the two functions*

$$y_+(x) = 3 + \sqrt{2^2 - x^2} \qquad \text{with} \qquad -2 \leq x \leq 2$$

and

$$y_-(x) = 3 - \sqrt{2^2 - x^2} \qquad \text{with} \qquad -2 \leq x \leq 2 \quad .$$

The graph of equation (4.17) (an integral curve) is a circle of radius 2 *about* $(0, 3)$ *(see figure 4.4a). It contains the point* $(x, y) = (0, 1)$ *corresponding to the initial value* $y(0) = 1$ *. To be specific, this point,* $(0, 1)$ *, is on the lower half of that circle (the solution curve for* $y_-(x)$ *) and not on the upper half (the solution curve for* $y_+(x)$ *). Thus, the (explicit) solution to our initial-value problem is*

$$y = y_-(x) = 3 - \sqrt{2^2 - x^2} \qquad \text{with} \qquad -2 \leq x \leq 2 \quad .$$

This is the solution curve sketched in figure 4.4b.

Let us now consider things more generally, and assume that we have any first-order differential equation that can be put into derivative formula form. Since what follows does not require "separability", let us simply assume we've managed to get the differential equation into the form

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

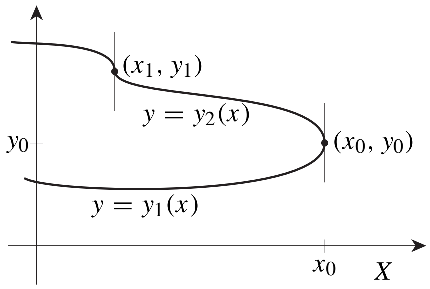

Figure 4.5: An integral curve containing two solution curves, with the portion above y_0 being the graph of $y_2(x)$ and the portion below y_0 being the graph of $y_1(x)$.

where $F(x, y)$ is some formula of x and y. This equation might be a separable differential equation such as

$$\frac{dy}{dx} = -\frac{x}{y-3} \quad ,$$

or it might be one like

$$\frac{dy}{dx} = x^2y^2 + 4 \quad ,$$

which is not separable. Suppose further that, either using methods developed in this chapter or methods that will be developed later, we have found an integral curve for this differential equation.

If no two distinct points on this integral curve have the same x-coordinate, then this curve is the graph of a function $y = y(x)$ that satisfies the differential equation (whether or not we can solve for the formula $y(x)$). So the entire integral curve is a solution curve.

On the other hand, if there are two points on this curve with the same x-coordinate, then the curve has to 'bend back' on itself at some point (x_0, y_0). At this point, the curve changes from being the graph of one solution $y = y_1(x)$ to being the graph of another solution $y = y_2(x)$. Also, at this point, the tangent line to the integral curve must be vertical (i.e., have "infinite slope"), provided that tangent line exists (see figure 4.5). But the slope of the tangent line to the graph of a differential equation's solution at any point (x, y) is simply the derivative ${}^{dy}/_{dx}$ of the solution at that point, and that value can be computed directly from the differential equation

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

Thus, (x_0, y_0), a point at which the integral curve 'bends back on itself', must be a point at which $F(x, y)$ becomes infinite (or, otherwise fails to exist).

Mind you, we cannot say that a curve 'bends back on itself' at a point just because the derivative becomes infinite there. Many functions have isolated points at which their derivative becomes infinite or otherwise fails to exist. Just look at point (x_1, y_1) in figure 4.5. Or consider

$$y(x) = 3x^{1/3} \quad .$$

This is a well-defined function on the entire real line whose derivative, $y'(x) = x^{-2/3}$, blows up at $x = 0$. So all we can say is that, if the curve does 'bend back on itself' then it can only do so at points where its derivative either becomes infinite or otherwise fails to exist.

Here is a little theorem summarizing some what we have just discussed.

Theorem 4.1
Let $\mathcal{C}$ be a curve contained in the graph of an implicit solution for some first-order differential equation

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

If $F(x, y)$ is a finite number for each point (x, y) in $\mathcal{C}$, then $\mathcal{C}$ is the graph of a function satisfying the given differential equation (i.e., $\mathcal{C}$ is a solution curve).

?►*Exercise 4.2:* *Explain why the integral curve graphed in figure 4.3 is "clearly" a solution curve.*

4.8 On Using Definite Integrals with Separable Equations

Just as with any directly integrable differential equation, a separable differential equation

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

once separated to the form

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

can be integrated using definite integrals instead of the indefinite integrals we've been using. The basic ideas are pretty much the same as for directly integrable differential equations:

1. Pick a convenient value for the lower limit of integration, a. In particular, if the value of $y(x_0)$ is given for some point x_0, set $a = x_0$.

2. Rewrite the differential equation with s denoting the variable instead of x. This means that we rewrite our separable equation as

$$\frac{dy}{ds} = f(s)g(y) \quad ,$$

   which 'separates' to

$$\frac{1}{g(y)}\frac{dy}{ds} = f(s) \quad .$$

3. Then integrate each side with respect to s from $s = a$ to $s = x$.

The integral on the left-hand side will be of the form

$$\int_{s=a}^{x} \frac{1}{g(y)}\frac{dy}{ds}\,ds \quad .$$

Keep in mind that, here, y is some unknown function of s, and that the limits in the integral are limits on s. Using the substitution $y = y(s)$, we see that

$$\int_{s=a}^{x} \frac{1}{g(y)}\frac{dy}{ds}\,ds = \int_{y=y(a)}^{y(x)} \frac{1}{g(y)}\,dy \quad .$$

Do not forget to convert the limits to being the corresponding limits on y , instead of s .

Once the integration is done, we attempt to solve the resulting equation for $y(x)$ just as before.

!►Example 4.12: *Let us solve*

$$\frac{dy}{dx} = \frac{1}{2y}e^{-x^2} \qquad \text{with} \quad y(0) = 3$$

using definite integrals. Proceeding as described above:

$$\frac{dy}{dx} = \frac{1}{2y}e^{-x^2}$$

$$\hookrightarrow \qquad 2y\frac{dy}{dx} = e^{-x^2}$$

$$\hookrightarrow \qquad 2y\frac{dy}{ds} = e^{-s^2}$$

$$\hookrightarrow \qquad \int_{s=0}^{x} 2y\frac{dy}{ds}\,ds = \int_{s=0}^{x} e^{-s^2}\,ds \quad .$$

Since $y(0) = 3$ *, we can rewrite the last equation as*

$$\int_{y=3}^{y(x)} 2y\,dy = \int_{s=0}^{x} e^{-s^2}\,ds \quad .$$

The integral on the left is easily evaluated; the one on the right is not. Doing the easy integration and solving for y *, we get*

$$y^2\Big|_{y=3}^{y(x)} = \int_{s=0}^{x} e^{-s^2}\,ds$$

$$\hookrightarrow \qquad [y(x)]^2 - 3^2 = \int_{s=0}^{x} e^{-s^2}\,ds$$

$$\hookrightarrow \qquad [y(x)]^2 = 9 + \int_{s=0}^{x} e^{-s^2}\,ds \quad .$$

So

$$y(x) = \pm\left[9 + \int_{s=0}^{x} e^{-s^2}\,ds\right]^{1/2} \quad .$$

Plugging in the initial value again,

$$3 = y(0) = \pm\left[9 + \int_{s=0}^{0} e^{-s^2}\,ds\right]^{1/2} = \pm[9 + 0]^{1/2} \quad ,$$

we clearly see that the $\pm$ *should be* $+$ *, not* $-$ *. Thus, the solution to our initial-value problem is*

$$y = \left[9 + \int_{s=0}^{x} e^{-s^2}\,ds\right]^{1/2} \quad .$$

Going back to the section on "named integrals" in chapter 2 (see page 30), we see that we can also express this as

$$y = \left[9 + \frac{\sqrt{\pi}}{2}\operatorname{erf}(x)\right]^{1/2} \quad .$$

The advantages of using definite integrals in solving a separable differential equation

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

are the same as in solving a directly integrable differential equation:

1. The solution directly involves the initial value instead of a constant to be determined from the initial value, and

2. Even if a 'nice' formula for
$$\int_a^x f(s)\,ds$$
cannot be found, the value of this integral can be closely approximated for specific values of x using standard methods (which are already in many computer math packages). Using these values for this integral, it is then often possible to find the corresponding values for $y(x)$ for specific values of x.

Unfortunately, we still have a serious problem if we cannot find a usable formula for

$$\int_{y(a)}^{y(x)} \frac{1}{g(y)}\,dy$$

since the numerical methods for computing this integral require knowing the value of $y(x)$ for the desired choice of x, and that $y(x)$ is exactly what we do *not* know.

Additional Exercises

4.3. *Determine whether each of the following differential equations is or is not separable, and, if it is separable, rewrite the equation in the form*

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

a. $\frac{dy}{dx} = 3y^2 - y^2\sin(x)$

b. $\frac{dy}{dx} = 3x - y\sin(x)$

c. $x\frac{dy}{dx} = (x-y)^2$

d. $\frac{dy}{dx} = \sqrt{1+x^2}$

e. $\frac{dy}{dx} + 4y = 8$

f. $\frac{dy}{dx} + xy = 4x$

g. $\frac{dy}{dx} + 4y = x^2$

h. $\frac{dy}{dx} = xy - 3x - 2y + 6$

i. $\frac{dy}{dx} = \sin(x+y)$

j. $y\frac{dy}{dx} = e^{x-3y^2}$

4.4. *Using the basic procedure, find the general solution to each of the following separable equations:*

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

b. $\frac{dy}{dx} = y^2 + 9$

c. $xy\frac{dy}{dx} = y^2 + 9$

d. $\frac{dy}{dx} = \frac{y^2+1}{x^2+1}$

e. $\cos(y)\frac{dy}{dx} = \sin(x)$

f. $\frac{dy}{dx} = e^{2x-3y}$

4.5. *Using the basic procedure, find the solution to each of the following initial-value problems:*

a. $\frac{dy}{dx} = \frac{x}{y}$ *with* $y(1) = 3$

b. $\frac{dy}{dx} = 2x - 1 + 2xy - y$ *with* $y(0) = 2$

c. $y\frac{dy}{dx} = xy^2 + x$ *with* $y(0) = -2$

d. $y\frac{dy}{dx} = 3\sqrt{xy^2 + 9x}$ *with* $y(1) = 4$

4.6. *Find all the constant solutions — and only the constant solutions — to each of the following. If no constant solution exists, say so.*

a. $\frac{dy}{dx} = xy - 4x$

b. $\frac{dy}{dx} - 4y = 2$

c. $y\frac{dy}{dx} = xy^2 - 9x$

d. $\frac{dy}{dx} = \sin(y)$

e. $\frac{dy}{dx} = e^{x+y^2}$

f. $\frac{dy}{dx} = 200y - 2y^2$

4.7. *Find the general solution for each of the following. Where possible, write your answer as an explicit solution.*

a. $\frac{dy}{dx} = xy - 4x$

b. $\frac{dy}{dx} = 3y^2 - y^2\sin(x)$

c. $\frac{dy}{dx} = xy - 3x - 2y + 6$

d. $\frac{dy}{dx} = \tan(y)$

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

f. $\frac{dy}{dx} = \frac{6x^2+4}{3y^2-4y}$

g. $\left(x^2+1\right)\frac{dy}{dx} = y^2 + 1$

h. $\left(y^2-1\right)\frac{dy}{dx} = 4xy^2$

i. $\frac{dy}{dx} = e^{-y}$

j. $\frac{dy}{dx} = e^{-y} + 1$

k. $\frac{dy}{dx} = 3xy^3$

l. $\frac{dy}{dx} = \frac{2+\sqrt{x}}{2+\sqrt{y}}$

m. $\frac{dy}{dx} - 3x^2y^2 = -3x^2$

n. $\frac{dy}{dx} - 3x^2y^2 = 3x^2$

o. $\frac{dy}{dx} = 200y - 2y^2$

4.8. *Solve each of the following initial-value problems. If possible, express each solution as an explicit solution.*

a. $\frac{dy}{dx} - 2y = -10$ *with* $y(0) = 8$

b. $y\frac{dy}{dx} = \sin(x)$ *with* $y(0) = -4$

c. $x\frac{dy}{dx} = y^2 - y$ *with* $y(1) = 2$

d. $x\frac{dy}{dx} = y^2 - y$ *with* $y(1) = 0$

e. $\left(y^2 - 1\right)\frac{dy}{dx} = 4xy$ *with* $y(0) = 1$

f. $\frac{dy}{dx} = \frac{y^2 - 1}{xy}$ *with* $y(1) = -2$

4.9. *In chapter 10, when studying population growth, we will obtain the "logistic equation"*

$$\frac{dy}{dx} = \beta y - \gamma y^2$$

with β *and* γ *being positive constants.*

a. *What are the constant solutions to this equation?*

b. *Find the general solution to this equation.*

4.10. *For each of the following initial-value problems, find the largest interval over which the solution is valid. (Note: You've already solved these initial-value problems in exercise set 4.8 or at least found the general solution to the differential equation in 4.7.)*

a. $\frac{dy}{dx} - 2y = -10$ *with* $y(0) = 8$

b. $x\frac{dy}{dx} = y^2 - y$ *with* $y(1) = 2$

c. $x\frac{dy}{dx} = y^2 - y$ *with* $y(1) = 0$

d. $\frac{dy}{dx} = e^{-y}$ *with* $y(0) = 1$

e. $\frac{dy}{dx} = 3xy^3$ *with* $y(0) = \frac{1}{2}$