

FIRST-ORDER ORDINARY DIFFERENTIAL EQUATIONS II: Graphical Methods and Applications

David Levermore
Department of Mathematics
University of Maryland

14 September 2012

Because the presentation of this material in lecture will differ from that in the book, I felt that notes that closely follow the lecture presentation might be appreciated.

CONTENTS

5. First-Order Equations: General Theory	
5.1. Well-Posed Problems	2
5.2. Existence and Uniqueness	3
6. First-Order Equations: Graphical Methods	
6.1. Phase-Line Portraits for Autonomous Equations	5
6.2. Plots of Explicit Solutions	7
6.3. Contour Plots of Implicit Solutions	9
6.4. Direction Fields	12
7. First-Order Equations: Applications	
7.1. General Guidelines	14
7.2. Tanks and Mixtures	14
7.3. Loans	17
7.4. Motion	18
7.5. Population Dynamics	21

5. FIRST-ORDER EQUATIONS: GENERAL THEORY

So far we have used analytical methods to construct solutions of first-order differential equations and solve initial-value problems. These methods have required the equation to be either linear or separable. More generally, they can be applied to any equation that can be transformed into either of those forms. They also have required finding certain primitives. We will learn other analytic methods later, but they also will require the equation to satisfy certain restrictions and finding certain primitives. However, most first-order differential equations either will not satisfy any of these restrictions or will lead to a primitive that cannot be found analytically. Moreover, even when analytic methods can work, they can produce complicated expressions which are not easy to understand. Therefore it is helpful to have other methods that can be applied to a broad class of first-order equations.

Subsequent sections will develop graphical and numerical methods that can be applied to a broad class of first-order equations. These methods require a little theoretical groundwork so that you have a clear understanding of when they can be applied.

5.1. Well-Posed Problems. The notion of a well-posed problem is central to science and engineering. It motivated by the idea that mathematical problems in science and engineering are used to predict or explain something. A problem is called well-posed if

- (i) the problem has a solution,
- (ii) the solution is unique,
- (iii) the solution depends continuously upon the problem.

The motivations for the first two points are fairly clear: a problem with no solution will not give a prediction, and a problem with many solutions gives too many. The third point is crucial. It recognizes that a mathematical problem is always a model of reality. Some nearby mathematical problems will be equally valid models. To have predictive value, we should be confident that the solutions of these other models lie close the solution of our model. For example, if our model is an initial-value problem associated with a differential equation then we would like to know that its solution would not change much if the initial value was a bit different or if a coefficient in the differential equation were a bit different. This is what is meant by saying the solution depends continuously upon the problem.

The solution of a well-posed problem can be approximated accurately by a wealth of techniques. The solution of a problem that is not well-posed is very difficult, if not impossible to approximate accurately. This is why scientists and engineers want to know which problems are well-posed and which are not.

In this section we consider initial-value problems of the form

$$(5.1) \quad \frac{dy}{dt} = f(t, y), \quad y(t_I) = y_I.$$

We will give conditions on $f(t, y)$ that insure this problem has a unique solution. In subsequent sections we will use this theory to develop methods by which we can study the solution to this problem when analytical methods either do not apply or become complicated.

5.2. Existence and Uniqueness. Here we will address only the existence and uniqueness of solutions. We begin with a definition that a picture should help clarify.

Definition 5.1. Let S be a set in the ty -plane. A point (t_o, y_o) is said to be in the interior of S if there exists a rectangle $(t_L, t_R) \times (y_L, y_R)$ that contains the point (t_o, y_o) and also lies within the set S .

Our basic existence and uniqueness theorem is the following.

Theorem 5.1. Let $f(t, y)$ be a function defined over a set S in the ty -plane such that

- f is continuous over S ,
- f is differentiable with respect to y over S ,
- $\partial_y f$ is continuous over S .

Then for every initial time t_I and every initial value y_I such that (t_I, y_I) is in the interior of S there exists a unique solution $y = Y(t)$ to initial-value problem (5.1) that is defined over some time interval (a, b) such that

- t_I is in (a, b) ,
- $\{(t, Y(t)) : t \in (a, b)\}$ lies within the interior of S .

Moreover, $Y(t)$ extends to the largest such interval and $Y'(t)$ is continuous over that interval.

Remark. This is not the most general theorem we could state, but it is one that applies to most equations you will face in this course, and is easy to apply. It asserts that $Y(t)$ will exist until $(t, Y(t))$ leaves S .

As the following examples show, applying this theorem is often a matter of simply checking that $f(t, y)$ and $\partial_y f(t, y)$ are continuous over the set S where they are defined. When this is the case, Theorem 5.1 insures that for every initial data (t_I, y_I) in the interior of S the initial-value problem has a unique solution $y = Y(t)$ that is defined over some time interval (a, b) that contains t_I .

Example. Determine (t_I, y_I) for which a unique solution exists to the initial-value problem

$$\frac{dy}{dt} = \frac{\log(|y|)}{1+t^2-y^2}, \quad y(t_I) = y_I.$$

Solution. Because $f(t, y) = \frac{\log(|y|)}{1+t^2-y^2}$ is defined everywhere except where $1+t^2-y^2=0$ or $y=0$, we try taking S to be all points in the ty -plane except those points on the hyperbola $1+t^2-y^2=0$ (where $y = \pm\sqrt{1+t^2}$) and those points on the t -axis (where $y=0$). Clearly, f is continuous over S , f is differentiable with respect to y over S with

$$\partial_y f(t, y) = \frac{1}{1+t^2-y^2} \frac{1}{y} + \frac{\log(|y|)}{(1+t^2-y^2)^2} 2y,$$

and $\partial_y f$ is continuous over S . Every point in S is also in the interior of S . Therefore Theorem 5.1 insures that for every (t_I, y_I) in S the initial-value problem has a unique solution $y = Y(t)$ that is defined over some time interval (a, b) that contains t_I . The solution extends to the largest interval (a, b) for which $(t, Y(t))$ remains within S .

Example. Determine (t_I, y_I) for which a unique solution exists to the initial-value problem

$$\frac{dy}{dt} = \sqrt{1+t^2+y^2}, \quad y(t_I) = y_I.$$

Solution. Because $f(t, y) = \sqrt{1 + t^2 + y^2}$ is defined over $(-\infty, \infty) \times (-\infty, \infty)$, we try taking S to be the entire ty -plane. Clearly, f is continuous over S , f is differentiable with respect to y over S with

$$\partial_y f(t, y) = \frac{y}{\sqrt{1 + t^2 + y^2}},$$

and $\partial_y f$ is continuous over S . Every point in S is also in the interior of S . Therefore Theorem 5.1 insures that for every (t_I, y_I) in the ty -plane the initial-value problem has a unique solution $y = Y(t)$ that is defined over some time interval (a, b) that contains t_I . Either the solution extends to all t or $Y(t)$ blows up in finite time because those are the only ways for $(t, Y(t))$ to leave S . (In fact, it extends to all t .)

Remark. The initial-value problem in the example above cannot be solved by analytic methods. However, Theorem 5.1 insures that its solution exists and is unique. In the next sections we will see how to visualize and approximate it.

6. FIRST-ORDER EQUATIONS: GRAPHICAL METHODS

Sometimes the best way to understand the solution of a differential equation is by graphical methods. Of course sometimes these methods can be applied when analytic methods fail to yield explicit solutions. But even when analytic methods can yield explicit solutions, it is often better to gain some understanding of the solutions through a graphical method. Often you can find out everything you need to know graphically, thereby saving yourself from a complicated analytic calculation.

6.1. Phase-Line Portraits for Autonomous Equations. This method can be applied to autonomous equations of the form

$$(6.1) \quad \frac{dy}{dt} = g(y).$$

It has the virtue that it can often be carried out quickly without the aid of a calculator or computer. It requires that g be continuous over an interval (y_L, y_R) , that g be differentiable at each of its zeros. Then by Theorem 4.1 every point of (y_L, y_R) has a unique solution of (6.1) passing through it.

Any solution $y = Y(t)$ of (6.1) can be viewed as giving the position of a point moving along the interval (y_L, y_R) as a function of time. We can determine the direction that this point moves as time increases from the sign of $g(y)$:

- where $g(y) = 0$ the point does not move because $Y'(t) = \frac{dy}{dt} = 0$,
- where $g(y) > 0$ the point moves to the right because $Y'(t) = \frac{dy}{dt} > 0$,
- where $g(y) < 0$ the point moves to the left because $Y'(t) = \frac{dy}{dt} < 0$.

We can present the sign analysis of $g(y)$ on a graph of the interval (y_L, y_R) as follows.

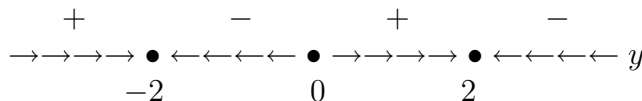
1. Find all the zeros of $g(y)$. These are the stationary points of the equation. They usually are isolated. Plot these points on the interval (y_L, y_R) . They partition (y_L, y_R) into subintervals.
2. Determine the sign of $g(y)$ on each of these subintervals. Plot right arrows on each subinterval where $g(y)$ is positive and left arrows on each subinterval where $g(y)$ is negative. These arrows indicate the direction that solutions of the equation will move along (y_L, y_R) as time increases.

The resulting graph is called the *phase portrait* for equation (6.1) along the interval (y_L, y_R) of the *phase-line*. It gives a rather complete picture of how all solutions of (6.1) behave when they take values in (y_L, y_R) .

Example. Describe the behavior of all solutions of the equation

$$\frac{dy}{dt} = 4y - y^3.$$

Solution. Because $4y - y^3 = y(2 + y)(2 - y)$, the stationary points of this equation are $y = -2$, $y = 0$, and $y = 2$. To get a complete picture you should sketch a phase portrait over an interval that includes all of these points, say $(-4, 4)$. Clearly $g(y)$ is positive over $(-4, -2)$, negative over $(-2, 0)$, positive over $(0, 2)$ and negative over $(2, 4)$. Therefore the phase portrait is



It is clear from this portrait that:

- every solution $y(t)$ that initially lies within $(-\infty, -2)$ will move towards -2 as t increases with $y(t) \rightarrow -2$ as $t \rightarrow \infty$;
- every solution $y(t)$ that initially lies within $(-2, 0)$ will move towards -2 as t increases with $y(t) \rightarrow -2$ as $t \rightarrow \infty$;
- every solution $y(t)$ that initially lies within $(0, 2)$ will move towards 2 as t increases with $y(t) \rightarrow 2$ as $t \rightarrow \infty$;
- every solution $y(t)$ that initially lies within $(2, \infty)$ will move towards 2 as t increases with $y(t) \rightarrow 2$ as $t \rightarrow \infty$.

Of course, all of this information can be read off from the analytic general solution we had worked out earlier. But if this is all you wanted to know, sketching the phase portrait is certainly a faster way to get it.

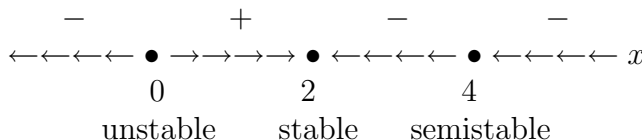
The above portrait shows that all solutions near -2 and 2 move towards them while all solutions near 0 move away from it. We say that the stationary points -2 and 2 are *stable*, while the stationary point 0 is *unstable*. A stationary point that has some solutions that move towards it and others that move away from it are called *semistable*. Phase portraits allow you to quickly classify the stability of every stationary point.

Example. Classify the stability of all the stationary solutions of

$$\frac{dx}{dt} = x(2-x)(4-x)^2.$$

Describe the behavior of all solutions.

Solution. The stationary points of this equation are $x = 0$, $x = 2$, and $x = 4$. A sign analysis of $x(2-x)(4-x)^2$ shows that the phase portrait for this equation is



We thereby classify the stability of the stationary as indicated above. Moreover:

- every solution $x(t)$ that initially lies within $(-\infty, 0)$ will move towards $-\infty$ as t increases with $x(t) \rightarrow -\infty$ as $t \rightarrow t_*$, where t_* is some finite “blow-up” time;
- every solution $x(t)$ that initially lies within $(0, 2)$ will move towards 2 as t increases with $x(t) \rightarrow 2$ as $t \rightarrow \infty$;
- every solution $x(t)$ that initially lies within $(2, 4)$ will move towards 2 as t increases with $x(t) \rightarrow 2$ as $t \rightarrow \infty$;
- every solution $x(t)$ that initially lies within $(4, \infty)$ will move towards 4 as t increases with $y(t) \rightarrow 4$ as $t \rightarrow \infty$.

Perhaps the only one of these statements that might be a bit surprising is the last one. By just looking at the arrows on the phase-line you might have thought that such solutions would move past 4 and continue down to 2 . But you have to remember that Theorem 4.1 tells us that nonstationary solutions never hit the stationary solutions. They merely approach them as $t \rightarrow \infty$.

6.2. Plots of Explicit Solutions. This method can be applied whenever you can obtain an explicit formula for a solution or a family of solutions. For example, if you want to see what a solution $y = Y(t)$ looks like over the time interval $[t_L, t_R]$ then the simplest thing to do when you already have an explicit expression for $Y(t)$ is use the MATLAB command **ezplot**. In general, you plot $Y(t)$ over a time interval $[t_L, t_R]$ by

```
>> ezplot('Y(x)', [t_L t_R])
>> xlabel 't', ylabel 'y'
>> title 'Plot of y = Y(t)'
```

You should get into the habit of labeling each axis and titling each graph.

Example. Consider the linear initial-value problem

$$\frac{dy}{dt} + y = 5 \sin(2t), \quad y(0) = 1.$$

You have found that its solution is

$$y = 3e^{-t} - 2 \cos(2t) + \sin(2t).$$

Plot this solution over the time interval $[0, 5]$.

Solution. You can input the explicit expression directly into **ezplot** as a symbolic expression. For example,

```
>> ezplot('3*exp(-x) - 2*cos(2*x) + sin(2*x)', [0 5])
>> xlabel 't', ylabel 'y'
>> title 'Plot of y = 3 exp(-t) - 2 cos(2t) + sin(2t)'
```

Remark. The first argument of **ezplot** must be a symbolic expression with one variable. Here we used x as the variable.

Alternative Solution. You can also input the explicit expression indirectly into **ezplot** as a symbolic expression. For example,

```
>> syms t; sol = 3*exp(-t) - 2*cos(2*t) + sin(2*t);
>> ezplot(sol, [0 5])
>> xlabel 't', ylabel 'y'
>> title 'Plot of y = 3 exp(-t) - 2 cos(2t) + sin(2t)'
```

Remark. We will take the second approach when plotting members of a family of solutions.

You can also use **ezplot** when the input is a symbolic solution generated by the MATLAB command **dsolve**.

Example. Consider the linear initial-value problem

$$\frac{dy}{dt} + y = 5 \sin(2t), \quad y(0) = 1.$$

Plot its solution over the time interval $[0, 5]$.

Solution. You can input the initial-value problem into **dsolve** as symbolic strings. For example,

```
>> sol = dsolve('Dy + y = 5*sin(2*t)', 'y(0) = 1', 't');
>> ezplot(sol, [0 5])
>> xlabel 't', ylabel 'y'
>> title 'Solution of Dy + y = 5 sin(2t), y(0) = 1'
```

Remark. The semicolon after the **dsolve** command suppresses its explicit output.

If $y = Y(t, c)$ is a family of solutions to a differential equation and you want to see how these solutions look over the time interval $[t_L, t_R]$ for several different values of c then you can change the values of c in `ezplot` by using the MATLAB command `subs` inside a loop.

Example. Consider the general initial-value problem

$$\frac{dy}{dt} + y = 5 \sin(2t), \quad y(0) = c.$$

You have found that its solution is

$$y = (c + 2)e^{-t} - 2 \cos(2t) + \sin(2t).$$

Plot this solution over the time interval $[0, 5]$ for $c = -3, -1, 1,$ and 3 .

Solution. We can define the family as a symbolic expression outside the loop as

```
>> syms t c
>> gensol = (c + 2)*exp(-t) - 2*cos(2*t) + sin(2*t);
>> figure; hold on
>> for cval = -3:2:3
    ezplot(subs(gensol, 'c', cval), [0 5])
end
>> axis tight, xlabel 't', ylabel 'y'
>> title 'Plot of y = (c + 2) exp(-t) - 2 cos(2t) + sin(2t) for c = -3, -1, 1, 3'
>> hold off
```

Remark. Notice that we have plotted the general solution for only four values of c . Plotting the solution for many more values of c can make the resulting graph look cluttered.

You can do a similar thing when the input is a symbolic family of solutions generated by the MATLAB command `dsolve`.

Example. Consider the general linear initial-value problem

$$\frac{dy}{dt} + y = 5 \sin(2t), \quad y(0) = c.$$

Plot its solution over the time interval $[0, 5]$ for $c = -3, -1, 1,$ and 3 .

Solution. Simply modify the previous solution as

```
>> gensol = dsolve('Dy + y = 5*sin(2*t)', 'y(0) = c', 't');
>> figure; hold on
>> for cval = -3:2:3
    ezplot(subs(gensol, 'c', cval), [0 5])
end
>> axis tight, xlabel 't', ylabel 'y'
>> title 'Solution of Dy + y = 5 sin(2t), y(0) = c for c = -3, -1, 1, 3'
>> hold off
```

Remark. You can find more examples of how to use `ezplot` to graph explicit solutions of first-order differential equations in our MATLAB book, *Differential Equations with MATLAB* by Hunt, Lipsman, Osborn, and Rosenberg.

6.3. Contour Plots of Implicit Solutions. This method can be applied anytime the solutions of a differential equation are implicitly given by an equation of the form

$$(6.2) \quad H(x, y) = c.$$

This situation might arise when solving a separable equation for which you can analytically find the primitives $F(x)$ and $G(y)$, but $G(y)$ is so complicated that you cannot analytically compute the inverse function G^{-1} . In that case (6.2) takes on the special form

$$F(x) - G(y) = c.$$

We will soon learn other methods that can lead to implicit solutions of the form (6.1) for any function $H(x, y)$ with continuous second partial derivatives. We will assume that $H(x, y)$ has continuous second partial derivatives.

Recall from multivariable calculus that the *critical points* of H are those points in the xy -plane where the gradient of H vanishes — i.e. those points (x, y) where

$$\partial_x H(x, y) = \partial_y H(x, y) = 0.$$

The value of $H(x, y)$ at a critical point is called a *critical value* of H . We will assume that each of critical point of H is *nondegenerate*. This means that at each of critical point of H the Hessian matrix $\mathbf{H}(x, y)$ of second partial derivatives has a nonzero determinant — i.e. at every critical point (x, y) we have

$$\det(\mathbf{H}(x, y)) \neq 0, \quad \text{where} \quad \mathbf{H}(x, y) = \begin{pmatrix} \partial_{xx}H(x, y) & \partial_{xy}H(x, y) \\ \partial_{yx}H(x, y) & \partial_{yy}H(x, y) \end{pmatrix}.$$

Because they are nondegenerate, the critical points of H can be classified as follows.

- If $\det(\mathbf{H}(x, y)) > 0$ and $\partial_{xx}H(x, y) > 0$ then (x, y) is a local minimizer of H .
- If $\det(\mathbf{H}(x, y)) > 0$ and $\partial_{xx}H(x, y) < 0$ then (x, y) is a local maximizer of H .
- If $\det(\mathbf{H}(x, y)) < 0$ then (x, y) is a saddle point of H .

Remark. Nondegenerate critical points of H are *isolated*. This means that each of critical point of H is contained within a rectangle $(a, b) \times (c, d)$ that contains no other critical points. This is a consequence of the *Implicit Function Theorem* of multivariable calculus.

Given a value for c , solutions of the differential equation lie on the set in the xy -plane given by

$$\{(x, y) : H(x, y) = c\}.$$

This is the so-called *level set* of $H(x, y)$ associated with c . Whenever this set has at least one point in it, the *Implicit Function Theorem* of multivariable calculus tells us the following.

- If c is not a critical value then its level set will look like one or more curves in the xy -plane that never meet. These curves will either be loops that close on themselves or extend to infinity. They will not have endpoints.
- If c is a critical value then its level set will look like one or more local extremizers of H plus some curves in the xy -plane that might meet only at saddle points. These curves will either be loops that close on themselves, extend to infinity, or have an endpoint at a saddle point.

The idea is to plot one or more level sets of H within a bounded rectangle in the xy -plane that is of interest. When many level sets are used the result is called a *contour plot* of H .

Remark. If we consider $H(x, y)$ to give the height of the graph of H over the xy -plane then a contour plot shows the height of this graph in exactly the same way a contour map shows the elevation of topographical features.

We can produce a contour plot of a function $H(x, y)$ over the rectangle $[x_L, x_R] \times [y_L, y_R]$ by using the MATLAB commands **meshgrid** and **contour** as follows.

```
>> [X, Y] = meshgrid(x_L:h:x_R, y_L:k:y_R);
>> contour(X, Y, H(X, Y))
>> axis square, xlabel 'x', ylabel 'y'
>> title 'Contour Plot of H(x, y)'
```

Here h and k are the resolutions of the intervals $[x_L, x_R]$ and $[y_L, y_R]$ respectively, which should have values of the form

$$h = \frac{x_R - x_L}{m}, \quad k = \frac{y_R - y_L}{n}, \quad \text{where } m \text{ and } n \text{ are positive integers.}$$

The meshgrid command creates an array of *grid points* in the rectangle $[x_L, x_R] \times [y_L, y_R]$ given by (x_i, y_j) where

$$x_i = x_L + ih \text{ for } i = 0, 1, \dots, m, \quad y_j = y_L + jk \text{ for } j = 0, 1, \dots, n.$$

More precisely, meshgrid creates two arrays; the array X contains x_i in its ij^{th} -entry while the array Y contains y_j in its ij^{th} -entry.

The contour command computes $H(x_i, y_j)$ at each of these grid points, uses these values to construct an approximation to $H(x, y)$ over $[x_L, x_R] \times [y_L, y_R]$ by interpolation, selects nine values for c that are evenly spaced between the minimum and maximum values of this interpolation, and then constructs an approximation to the level set of each c based on this interpolation. Exactly how contour does all this is beyond the scope of this course, so we will not discuss it further. You can learn more about such algorithms in a numerical analysis course. Here all you need to know is that MATLAB produces an approximation to each level set. This approximation can be improved by making h and k smaller. The resulting contour plot will give you a good idea of what $H(x, y)$ looks like over the rectangle $[x_L, x_R] \times [y_L, y_R]$ provided h and k are small enough. Typical values for m and n run between 50 and 200.

Example. Produce a contour plot of $H(x, y) = x^2 + y^2$ in the rectangle $[-5, 5] \times [-5, 5]$.

Solution. If we select $m = n = 100$, so that $h = k = .1$, then our MATLAB program becomes

```
>> [X, Y] = meshgrid(-5:0.1:5, -5:0.1:5);
>> contour(X, Y, X.^2 + Y.^2)
>> axis square, xlabel 'x', ylabel 'y'
>> title 'Contour Plot of x^2 + y^2'
```

Because the minimum and maximum of $x^2 + y^2$ over the rectangle $[-5, 5] \times [-5, 5]$ are 0 and 50 respectively, the values of c MATLAB will choose are $5l$ where $l = 0, 1, 2, \dots, 10$. The resulting contour plot should show a point at the origin plus the intersections of the rectangle $[-5, 5] \times [-5, 5]$ with ten concentric circles centered at the origin that have radii $\sqrt{5}, \sqrt{10}, \sqrt{15}, 2\sqrt{5}, 5, \sqrt{30}, \sqrt{35}, 2\sqrt{10}, 3\sqrt{5},$ and $5\sqrt{2}$.

Remark. The dots that appear in $X.^2 + Y.^2$ tell MATLAB that we want to square each entry in the arrays X and Y rather than square the arrays themselves by matrix multiplication.

If we are investigating an initial-value problem with initial data (x_I, y_I) inside a rectangle $[x_L, x_R] \times [y_L, y_R]$ then the only level set that we want to plot is the one corresponding to $c = H(x_I, y_I)$. This level set might show more solutions than the one we seek, so be careful! We can plot just this level set by the following.

```
>> c = H(x_I, y_I);
>> [X, Y] = meshgrid(x_L:h:x_R, y_L:k:y_R);
>> contour(X, Y, H(X,Y), [c c])
>> axis square, xlabel 'x', ylabel 'y'
>> title 'Level Set for H(x, y) = H(x_I, y_I)'
```

Example. Consider the initial-value problem

$$\frac{dx}{dt} = \frac{e^x \cos(t)}{1+x}, \quad x(0) = -2.$$

Graph its solution over the time interval $[-4, 4]$.

Solution. In Section 4.4 we used the fact this equation is separable to show that the solution of this initial-value problem satisfies

$$\sin(t) = -(2+x)e^{-x}.$$

The analysis in Section 4.4 showed moreover that there is a unique oscillatory solution $x = X(t) < -1$ to this equation with interval of definition $(-\infty, \infty)$. However, other solutions lie on this level set that have $x > -1$. We can plot the level set by the following.

```
>> [T, X] = meshgrid(-4:0.1:4,-4:0.1:4);
>> contour(T, X, sin(T) + (2 + X).*exp(-X), [0 0])
>> axis square, xlabel 't', ylabel 'x'
>> title 'Solution of dy/dt = exp(-x) cos(x) / (1 + x), x(0) = -2'
```

We can plot the level sets for the c values c_1 , c_2 , and c_3 on a single graph by the following.

```
>> [X, Y] = meshgrid(x_L:h:x_R, y_L:k:y_R);
>> contour(X, Y, H(X,Y), [c_1 c_2 c_3])
>> axis square, xlabel 'x', ylabel 'y'
>> title 'Level Sets for H(x, y) = c_1, c_2, c_3'
```

Example. Consider the general initial-value problem

$$\frac{dx}{dt} = \frac{e^x \cos(t)}{1+x}, \quad x(0) = x_I.$$

Graph its solutions over the time interval $[-4, 4]$ for $x_I = -2, 0, 1$, and 2 .

Solution. In Section 4.4 we showed that the solution of this initial-value problem satisfies

$$\sin(t) = (2+x_I)e^{-x_I} - (2+x)e^{-x}.$$

This equation has the form

$$\sin(t) + (2+x)e^{-x} = c,$$

where the values of c corresponding to $x_I = -2, 0, 1$, and 2 are $c = 0, 2, 3/e$ and $4/e^2$ respectively. We can plot these level sets by the following.

```
>> [T, X] = meshgrid(-4:0.1:4,-4:0.1:4);
>> contour(T, X, sin(T) + (2 + X).*exp(-X), [0 2 3/exp(1) 4/exp(2)])
>> axis square, xlabel 't', ylabel 'x'
>> title 'Solutions of dy/dt = exp(-x) cos(x) / (1 + x), x(0) = x_I'
```

6.4. Direction Fields. This is the crudest tool in your toolbox. Its virtue is that it can be applied to almost any first-order equation

$$(6.3) \quad \frac{dy}{dt} = f(t, y).$$

We assume that the function $f(t, y)$ is defined over a set S in the ty -plane such that

- f is continuous over S ,
- f is differentiable with respect to y over S ,
- $\partial_y f$ is continuous over S .

Moreover, we assume that every point in a rectangle $[t_L, t_R] \times [y_L, y_R]$ is in the interior of S . Then by Theorem 5.1 every point (t_I, y_I) in $[t_L, t_R] \times [y_L, y_R]$ has a unique curve $(t, Y(t))$ passing through it such that $y = Y(t)$ is a solution of (6.3). This curve can be extended to the largest time interval $[a, b]$ such that $(t, Y(t))$ remains within the rectangle $[t_L, t_R] \times [y_L, y_R]$.

If we cannot find an explicit or implicit solution of (6.3) then we cannot plot the curve $(t, Y(t))$ by the methods of the previous two sections. However, if $f(t, y)$ that meets the criteria given above then we know by Theorem 5.1 that the curve exists and has a tangent vector given by

$$\frac{d}{dt}(1, Y(t)) = (1, Y'(t)) = (1, f(t, Y(t))).$$

In other words, the unique solution that goes through any point (t, y) in the rectangle $[t_L, t_R] \times [y_L, y_R]$ has the tangent vector $(1, f(t, y))$. A *direction field* for equation (6.3) over the rectangle $[t_L, t_R] \times [y_L, y_R]$ is a plot that shows the direction of this tangent vector with an arrow at each point of a grid in the rectangle $[t_L, t_R] \times [y_L, y_R]$. The idea is that these arrows might give us a correct picture of how the orbits move inside the rectangle.

We can produce such a direction field by using the MATLAB commands **meshgrid** and **quiver** as follows.

```
>> [T, Y] = meshgrid(t_L:h:t_R, y_L:k:y_R);
>> S = f(T, Y);
>> L = sqrt(1 + S.^2);
>> quiver(T, Y, 1./L, S./L, l)
>> axis tight, xlabel 't', ylabel 'y'
>> title 'Direction Field for dy/dt = f(t, y)'
```

Here h and k are the grid spacings for the intervals $[t_L, t_R]$ and $[y_L, y_R]$ respectively, which should have values of the form

$$h = \frac{t_R - t_L}{m}, \quad k = \frac{y_R - y_L}{n}, \quad \text{where } m \text{ and } n \text{ are positive integers.}$$

The meshgrid command creates an array of *grid points* in the rectangle $[t_L, t_R] \times [y_L, y_R]$ given by (t_i, y_j) where

$$t_i = t_L + ih \text{ for } i = 0, 1, \dots, m, \quad y_j = y_L + jk \text{ for } j = 0, 1, \dots, n.$$

More precisely, meshgrid creates two arrays; the array T contains t_i in its ij^{th} -entry while the array Y contains y_j in its ij^{th} -entry. Next, an array S is computed that contains the slope $f(t_i, y_j)$ in its ij^{th} -entry. Then an array L is computed that contains the length of the tangent vector $(1, f(t_i, y_j))$ in its ij^{th} -entry.

Finally, the quiver command plots an array of arrows of length ℓ so that the ij^{th} -arrow is centered at the grid point (t_i, y_j) and is pointing in the direction of the unit tangent vector

$$\left(\frac{1}{\sqrt{1 + f(t_i, y_j)^2}}, \frac{f(t_i, y_j)}{\sqrt{1 + f(t_i, y_j)^2}} \right).$$

The length ℓ should be smaller than h or k so that the plotted arrows will not overlap. Typically m and n will be about 20 to insure there will be enough arrows to give a complete picture of the direction field, but not so many that the plot becomes cluttered.

Remark. Often it is hard to figure out how the orbits move from the arrows in a direction field. Therefore they should be used only as a tool of last resort. They should never be used for autonomous equations because phase-line portraits are much easier to use.

Example. Describe the solutions of the equation

$$\frac{dy}{dt} = \sqrt{1 + t^2 + y^2}.$$

Solution. This equation is not linear or separable, so it cannot be solved by the analytic methods we have studied. Earlier we showed that this equation meets the criteria of Theorem 5.1, so there is a unique solution that passes through every point in the ty -plane. Therefore the only method we have discussed that applies to this equation is direction fields.

Before applying the method of direction fields, we should see what information can be seen directly from the equation. Its right-hand side satisfies $\sqrt{1 + t^2 + y^2} \geq 1$. This means its solutions will all be increasing functions of t . If our direction fields are not consistent with this observation then we will know we have made a mistake in our MATLAB program.

We can produce a direction field for this equation in the rectangle $[-5, 5] \times [-5, 5]$ with a grid spacing of 0.5 (by taking $m = n = 20$) and arrows of length 0.35 as follows.

```
>> [T, Y] = meshgrid(-5:0.5:5,-5:0.5:5);
>> S = sqrt(1 + T.^2 + Y.^2);
>> L = sqrt(1 + S.^2);
>> quiver(T, Y, 1./L, S./L, 0.35)
>> axis tight, xlabel 't', ylabel 'y'
>> title 'Direction Field for dy/dt = sqrt(1 + t^2 + y^2)'
```

In this case it is easy to figure out how the orbits move from the direction field.

7. FIRST-ORDER EQUATIONS: APPLICATIONS

The subject of differential equations was invented along with calculus by Newton and Leibniz in order to solve problems in geometry and physics. It played a central role in the development of Newtonian physics by the Bernoulli family, Euler, and others. It rapidly found applications in biology, finance, and engineering. It now is applied in almost every discipline that has been quantified. You benefit from these applications every time you use your cell phone, drive your car, fly in an airplane, listen to a weather report, use the internet, take a modern medicine, or do many other daily activities.

7.1. General Guidelines. Mathematical problems that arise in applications are usually stated as word problems. The problems considered in this section will ask questions whose answers require analyzing an initial-value problem. These problems typically will not hand you an initial-value problem. Rather, you will have to construct an initial-value problem from the information given in a word problem. In other words, you must transform a problem stated verbally into a problem that you can analyze by the methods we have been studying. Often this can be done more than one way. The more ways you see to approach a problem, the better you understand it. Here we address how to go about seeking ways to approach these problems.

There is no set of magic steps by which you can approach every word problem. Rather, you should be ready to consider several possible approaches when you are faced with one. To guide your efforts, it is helpful to keep in mind the following three objectives.

- (1) *Identify the variables in the problem and the relationships between them.* An appropriately labeled picture is often helpful in this regard. You should note carefully any relations between or constraints on the variables you have introduced. You should be prepared to rethink your choice of variables to help achieve the next objective.
- (2) *Reduce the problem analyzing a an initial-value problem.* You should used relations between the various variables to reduce the number of parameters in the differential equation as much as possible. These relations may or may not be stated explicitly in the problem. Similarly, constraints may or may not be stated explicitly in the problem (like the fact that populations should not be negative).
- (3) *Solve the resulting problem.* You should pick the method to suit the problem. What method is best will depend on the question being asked. Sometimes a graphical method will be the fastest route to the solution. A numerical method might be best when analytical and graphical methods prove to be difficult.

These are the so-called **IRS** guidelines, which apply to many kinds of word problems: **I**dentify the problem; **R**educe the problem to one you can analyze; **S**olve the reduced problem.

The remainder of this section will illustrate how these guidelines can help you approach word problems in the context of several applications.

7.2. Tanks and Mixtures. These represent a broad class of problems in which one asks a question about the transport of some quantity into and out of a tank or some other volume. The quantity might be a fluid like water, oil, or air, or it might be something carried by a fluid like a pollutant or solute. The tank might be any well-defined volume like a pond, lake, or room in a building. These problems generalize to ones involving networks of interconnected tanks, which lie at the heart of many numerical simulations of fluids.

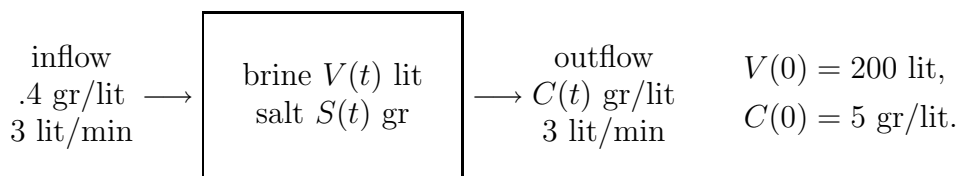
In the following problems we will construct an initial-value problem for the amount Q of some quantity in the tank. The associated ordinary differential equation will have the form

$$\frac{dQ}{dt} = \text{RATE IN} - \text{RATE OUT},$$

where RATE IN is the rate the quantity enters the tank while RATE OUT is the rate the quantity exits the tank. Sometimes RATE IN and RATE OUT will be given explicitly in the problem. At other times they will be given in terms of other variables in the problem.

Example. A tank initially contains 200 liters of brine (salt solution) with a salt concentration of 5 grams per liter. At some instant brine with a salt concentration of .4 grams per liter begins to flow into the tank at a rate of 3 liters per minute, while the well-stirred mixture flows out at the same rate. How long will it take for the salt concentration in the tank to be reduced to .7 grams per liter?

Solution. Let $V(t)$ be the volume (lit) of brine in the tank at time t minutes. Let $S(t)$ be the mass (gr) of salt in the tank at time t minutes. Because the mixture is assumed to be well-stirred, the salt concentration of the brine in the tank at time t is $C(t) = S(t)/V(t)$. In particular, this will be the concentration of the brine that flows out of the tank. We have the following picture.



We want to find the time T at which $C(T) = .7$.

Because brine flows in and out of the tank at the same rate, we see that $V(t) = V(0) = 200$. Hence, $C(t) = S(t)/200$. Therefore $S(t)$ satisfies

$$\frac{dS}{dt} = \text{RATE IN} - \text{RATE OUT} = .4 \cdot 3 - \frac{S}{200} \cdot 3 = 1.2 - \frac{3}{200}S.$$

Because $S(0) = 200C(0) = 200 \cdot 5 = 1000$, the initial-value problem that governs $S(t)$ is

$$\frac{dS}{dt} = 1.2 - .015S, \quad S(0) = 1000.$$

We want to find T such that $S(T) = 200C(T) = 200 \cdot .7 = 140$.

This is a nonhomogeneous linear differential equation with normal form

$$\frac{dS}{dt} + .015S = 1.2.$$

Its integrating factor form is

$$\frac{d}{dt} (e^{.015t} S) = 1.2e^{.015t}.$$

Upon integrating this equation while using the initial condition $S(0) = 1000$ we obtain

$$e^{.015t} S - 1000 = \frac{1.2}{.015} (e^{.015t} - 1) = 80 (e^{.015t} - 1),$$

whereby

$$S(t) = 920e^{-.015t} + 80.$$

By setting $S(T) = 140$ we find that T satisfies

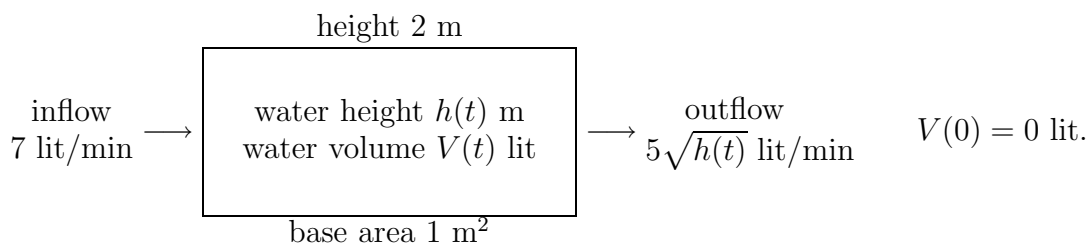
$$920e^{-.015T} + 80 = 140.$$

By solving for T we obtain

$$T = \frac{1}{.015} \log\left(\frac{920}{60}\right) = \frac{200}{3} \log\left(\frac{46}{3}\right) \text{ minutes.}$$

Example. A tank with an open top has a base of 1 square meter and a height of 2 meters. The tank is initially empty when water begins to pour into it at a rate of 7 liters per minute. The water also drains from the tank through a hole in its bottom at a rate of $5\sqrt{h}$ liters per minute where h is the height of the water in the tank in meters. Will the tank overflow? If not, how high will it fill? If so, how long does it take to happen?

Solution. Let $V(t)$ be the volume (lit) of water in the tank at time t minutes. We have the following picture.



It is clear that $h(t)$ is an increasing function. We want to determine if $h(T) = 2$ for some T . If not, we want to determine the value $h(t)$ approaches as $t \rightarrow \infty$.

Because $1 \text{ m}^3 = 1000 \text{ lit}$, $V(t) = 1000 \cdot 1 \cdot h(t)$. Because $V(t)$ satisfies

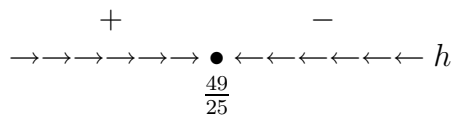
$$\frac{dV}{dt} = \text{RATE IN} - \text{RATE OUT} = 7 - 5\sqrt{h},$$

the initial-value problem that governs $h(t)$ is

$$1000 \frac{dh}{dt} = 7 - 5\sqrt{h}, \quad h(0) = 0.$$

We want to determine if $h(T) = 2$ for some T . If not, we want to determine the value $h(t)$ approaches as $t \rightarrow \infty$.

This equation is autonomous. It has one stationary point at $h = \frac{49}{25}$. Its phase portrait is



This portrait shows that if $h(0) = 0$ then $h(t) \rightarrow \frac{49}{25}$ as $t \rightarrow \infty$. Because $\frac{49}{25} = 1.96 < 2$, the tank does not overflow. Rather, the water approaches a height of 1.96 meters as $t \rightarrow \infty$.

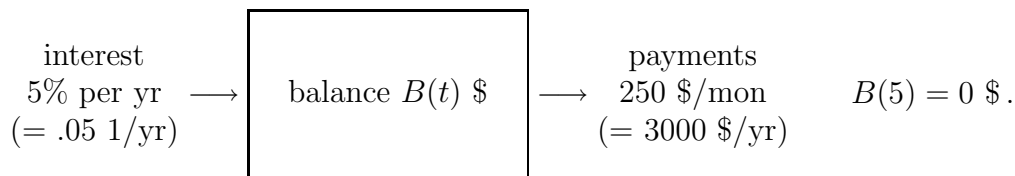
7.3. Loans. A loan problem can be viewed as a tank problem where the loan balance is being drained away by the borrower. To keep things simple, we consider fixed-rate loans with continuously compounded interest and with continuous payments by the borrower at a constant rate. The term of the loan is the time period over which the loan is paid off. If we let $B(t)$ be the balance of the loan at time t years then $B(t)$ will satisfy

$$\frac{dB}{dt} = rB - P,$$

where r is the per annum interest rate and P is the per annum payment rate. If the initial loan amount is known to be B_I then we would impose the initial condition $B(0) = B_I$. However, if you are trying to find B_I or P , given r and the term T of the loan in years then you would impose the initial condition $B(T) = 0$.

Example. A car buyer has 4000\$ for a down payment and can afford to make continuous payments for a loan at a constant rate of no more than 250\$ per month. If five-year fixed-rate loans are available at an interest rate of 5% per year compounded continuously, what is the price of the most expensive car that the buyer can afford?

Solution. Let $B(t)$ be the balance of the loan at time t years. Because 1 year = 12 months, we see $P = 12 \cdot 250 = 3000$ \$/yr. Because 1 = 100%, we see $r = 5/100 = .05$ 1/yr. The term $T = 5$ yr. We have the following picture.



The price of the most expensive car the buyer can afford will be $4000 + B(0)$ dollars.

The initial-value problem satisfied by $B(t)$ is

$$\frac{dB}{dt} = .05B - 3000, \quad B(5) = 0.$$

We want to find $4000 + B(0)$.

This is a nonhomogeneous linear differential equation with normal form

$$\frac{dB}{dt} - .05B = -3000.$$

Its integrating factor form is

$$\frac{d}{dt} (e^{-.05t} B) = -3000e^{-.05t}.$$

Upon integrating this equation while using the initial condition $B(5) = 0$ we obtain

$$e^{-.05t} B(t) = -3000 \int_5^t e^{-.05s} ds = \frac{3000}{.05} (e^{-.05t} - e^{-.05 \cdot 5}) = 60000 \left(e^{-\frac{1}{20}t} - e^{-\frac{1}{4}} \right).$$

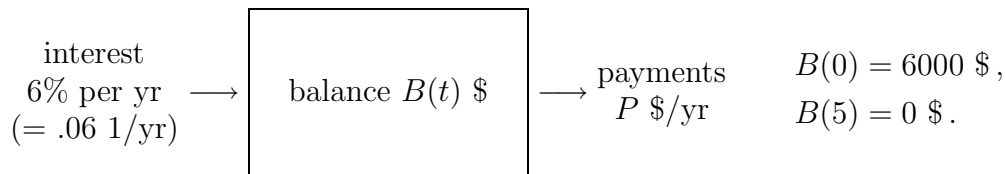
By setting $t = 0$ above we find that the car buyer can afford a car costing at most

$$4000 + 60000 \left(1 - e^{-\frac{1}{4}} \right) \text{ dollars.}$$

This is about 17,272\$, but you should leave the answer in the exact form above.

Example. A student borrows 6000\$ at an interest rate of 6% compounded continuously. The student wants to pay off the loan in five years by making payments continuously at a constant rate of P dollars per year. What should P be?

Solution. Let $B(t)$ be the balance of the loan at time t years. Because $1 = 100\%$, we see $r = 6/100 = .06$ 1/yr. The term $T = 5$ yr. We have the following picture.



One of the last two conditions will determine P .

The initial-value problem satisfied by $B(t)$ is

$$\frac{dB}{dt} = .06B - P, \quad B(5) = 0.$$

We want to find P such that $B(0) = 6000$.

This is a nonhomogeneous linear differential equation with normal form

$$\frac{dB}{dt} - .06B = -P.$$

Its integrating factor form is

$$\frac{d}{dt} (e^{-.06t} B) = -P e^{-.06t}.$$

Upon integrating this equation while using the initial condition $B(5) = 0$ we obtain

$$e^{-.06t} B(t) = \frac{P}{.06} (e^{-.06t} - e^{-.06 \cdot 5}) = -P \int_5^t e^{-\frac{3}{50}s} ds = \frac{50P}{3} \left(e^{-\frac{3}{50}t} - e^{-\frac{3}{10}} \right).$$

By setting $t = 0$ we see that $B(0) = \frac{50}{3}(1 - e^{-\frac{3}{10}})$. Then $B(0) = 6000$ implies that the student has to pay off the loan at a rate of

$$P = \frac{6000}{\frac{50}{3}(1 - e^{-\frac{3}{10}})} = \frac{360}{1 - e^{-\frac{3}{10}}} \text{ dollars per year.}$$

This is about 1,389\$ per year, but you should leave your answer in the exact form above.

7.4. Motion. A falling object with fixed mass m is governed by the Newton law of motion $ma = F$, where a is its acceleration and F is the net force acting on it. The net force is the sum of the gravitational force and the drag force. If the object has a downward velocity $v(t)$ then the Newton law of motion takes the form

$$m \frac{dv}{dt} = \text{gravitational force} + \text{drag force} = F_{\text{grav}} + F_{\text{drag}}.$$

For objects near the surface of the Earth that fall distances that are short compared to the radius of the Earth, we can approximate the gravitational force F_{grav} as the constant mg where g is the so-called gravitational acceleration given by $g = 9.8$ m/sec².

The drag force F_{drag} is harder to approximate. It acts in the direction opposite to that of the velocity. If the object is going up then the drag force acts downward while if the object is going down the drag force acts upward. One simple model sets $F_{\text{drag}} = -\rho_{\text{air}} A |v|v$ where ρ_{air} is the density of the air and A is the aerodynamic cross-section, which has units of area.

The value of ρ_{air} varies significantly with temperature, pressure, and humidity; for dry air at 20 °C and one atmosphere pressure ρ_{air} is about 1.2 kg/m³. The value of A depends upon the shape of the object, the air density, and the velocity v in an extremely complicated way. When the speed of the falling object is far below the speed of sound then the dominant dependence is on the shape of the object. If values for ρ_{air} and A are needed for a problem then they will be given to you.

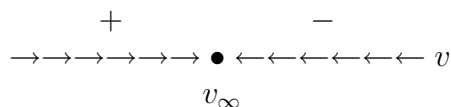
The resulting equation of motion is a first-order differential of the form

$$(7.1) \quad \frac{dv}{dt} = g - k|v|v, \quad \text{where } k = \frac{\rho_{\text{air}}A}{m} > 0.$$

This equation is autonomous. Its right-hand side is differentiable with respect to v because $\partial_v(|v|v) = 2|v|$, so that Theorem 4.1 applies to it. Its only stationary solution is

$$(7.2) \quad v_{\infty} = \sqrt{\frac{g}{k}}.$$

Its phase-line portrait is

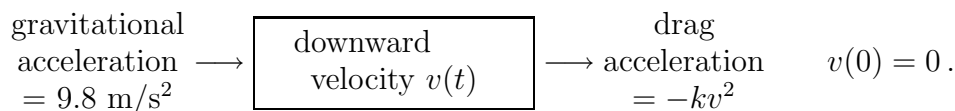


This portrait shows that the stationary point v_{∞} is attracting, and that $v(t) \rightarrow v_{\infty}$ as $t \rightarrow \infty$ for every nonstationary solution. The velocity v_{∞} is called the terminal velocity because it is the velocity being approached by the falling object until it hits the ground.

Remark. The terminal velocity depends upon the aerodynamic cross-section A as $A^{-\frac{1}{2}}$. Skydivers control the rate of their descent by changing their aerodynamic cross-section. They are close to terminal velocity after falling about ten seconds. They can reach speeds of 90 m/s in a bullet-like position or slow to about 50 m/s by spreading their arms and legs. With a parachute deployed they slow to 3-4 m/s, depending on the design of the parachute.

Example. A skydiver of mass 60 kg jumps from an airplane and assumes a position with an aerodynamic cross-section of 0.1 m² in air with a density of 1.2 kg/m³. What is her terminal velocity? What fraction of her terminal velocity does she reach after 10 s? How far has she fallen after 10 s?

Solution. Let $v(t)$ be her downward velocity at t seconds. Because she is always falling during the ten seconds, we know that $v(t) \geq 0$ and have the following picture.



We first seek v_{∞} and $v(10)/v_{\infty}$, where v_{∞} is her terminal velocity.

The initial-value problem satisfied by $v(t)$ is

$$\frac{dv}{dt} = g - kv^2, \quad v(0) = 0,$$

where $g = 9.8 \text{ m/s}^2$ and

$$k = \frac{\rho_{\text{air}}A}{m} = \frac{1.2 \cdot 0.1}{60} = \frac{1}{500} = .002 \text{ 1/m}.$$

Therefore her terminal velocity is

$$v_{\infty} = \sqrt{\frac{g}{k}} = \sqrt{9.8 \cdot 500} = \sqrt{4900} = 70 \text{ m/s}.$$

The next step is to compute $v(t)$. The initial-value problem can be expressed as

$$\frac{dv}{dt} = k(v_{\infty}^2 - v^2), \quad v(0) = 0,$$

where $k = .002$ and $v_{\infty} = 70$. This differential equation is autonomous, so an implicit solution can be found by using a partial fraction decomposition as

$$\begin{aligned} kt &= \int \frac{1}{v_{\infty}^2 - v^2} dv = \int \frac{1}{(v_{\infty} + v)(v_{\infty} - v)} dv = \int \frac{\frac{1}{2v_{\infty}}}{v_{\infty} + v} + \frac{\frac{1}{2v_{\infty}}}{v_{\infty} - v} dv \\ &= \frac{1}{2v_{\infty}} \log(v_{\infty} + v) - \frac{1}{2v_{\infty}} \log(v_{\infty} - v) + c = \frac{1}{2v_{\infty}} \log\left(\frac{v_{\infty} + v}{v_{\infty} - v}\right) + c. \end{aligned}$$

Here we do not need absolute values inside the log because we know from the phase-line portrait that $v(t)$ will increase from 0 to the terminal velocity v_{∞} . The initial condition $v(0) = 0$ implies that

$$k \cdot 0 = \frac{1}{2v_{\infty}} \log\left(\frac{v_{\infty} + 0}{v_{\infty} - 0}\right) + c = \frac{1}{2v_{\infty}} \log(1) + c = 0 + c,$$

whereby $c = 0$. By exponentiating the implicit solution we find that

$$e^{2v_{\infty}kt} = \frac{v_{\infty} + v}{v_{\infty} - v},$$

which can be solved for v to obtain the explicit solution

$$v(t) = v_{\infty} \frac{e^{2v_{\infty}kt} - 1}{e^{2v_{\infty}kt} + 1} = 70 \frac{e^{2.8t} - 1}{e^{2.8t} + 1}.$$

Therefore the fraction of her terminal velocity reached after ten seconds is

$$\frac{v(10)}{70} = \frac{e^{2.8} - 1}{e^{2.8} + 1}.$$

This is about .88535 of her terminal velocity, but you should leave your answer in the exact form given above.

If we let $y(t)$ be distance she has fallen after t seconds then

$$\begin{aligned} y(10) &= \int_0^{10} v(t) dt = \int_0^{10} 70 \frac{e^{2.8t} - 1}{e^{2.8t} + 1} dt = 70 \int_0^{10} \frac{2e^{2.8t}}{e^{2.8t} + 1} - 1 dt \\ &= 70 \left[\frac{2}{.28} \log(e^{2.8t} + 1) - t \right] \Big|_0^{10} = 70 \left[\frac{1}{.14} \log\left(\frac{e^{2.8} + 1}{2}\right) - 10 \right] \\ &= 500 \log\left(\frac{e^{2.8} + 1}{2}\right) - 700. \end{aligned}$$

This is about 383 meters, but you should leave your answer in the exact form given above.

7.5. Population Dynamics. We now consider models of population dynamics governed by first-order equations of the form

$$\frac{dp}{dt} = R(p)p - h(t).$$

where $p(t)$ is the size of the population as a function of time, $R(p)$ models the growth rate of the population as a function of p , and $h(t)$ is a harvest rate due to predators. For a population in a closed ecosystem (with $h(t) = 0$) the growth rate is simply the birth rate minus the mortality rate. In more complicated situations the growth rate must also account for migration in to and out of the ecosystem. The harvest rate can model the catch in fish populations, the hunting limit in deer populations, or the introduction of any other predator that will reduce the population at a rate that is independent of the size of the population. Obviously such a harvest model will break down if the population is reduced too much.

The simplest such models take $h(t) = 0$ and $R(p) = r$ for some constant r . This is the so-called exponential model because it has the solution $p(t) = p_I e^{rt}$ when $p(0) = p_I$. This solution grows exponentially when $r > 0$ and decays exponentially when $r < 0$. Sometimes you will have to figure out the value of r from other information in the problem. For example, if you are told that a population triples every five years then you are being told that $p(t+5) = 3p(t)$. Then you can figure out r by setting

$$3 = \frac{p(t+5)}{p(t)} = \frac{p_I e^{r(t+5)}}{p_I e^{rt}} = e^{r5},$$

whereby the growth rate is $r = \frac{1}{5} \log(3)$ per year. Alternatively, if you are told that a population triples every five years then you are being told that $p(t) = p_I 3^{\frac{1}{5}t}$. Because $3^{\frac{1}{5}t} = e^{\frac{1}{5} \log(3)t}$, you can read off that the growth rate is $r = \frac{1}{5} \log(3)$ per year.

If a harvest $h(t) > 0$ is introduced into an exponential model then it takes the form

$$\frac{dp}{dt} = rp - h(t).$$

This is a nonhomogeneous linear equation that we know how to solve.

Example. In the absence of predators the population of mosquitoes in a certain area would increase at a rate proportional to its current population and its population would double every three weeks. There are 250,000 mosquitoes in the area initially when a flock of birds arrives that eats 80,000 mosquitoes per week. How many mosquitoes remain after two weeks?

Solution. Let $M(t)$ be the number of mosquitoes at time t weeks. Doubling every three weeks means the population grows like $2^{\frac{1}{3}t} = e^{\frac{1}{3} \log(2)t}$, which implies a growth rate of $\frac{1}{3} \log(2)$ per mosquito. The rate at which the mosquitoes reproduce thereby is $\frac{1}{3} \log(2)M(t)$ while the rate at which they are eaten is 80,000. Therefore the initial-value problem that M satisfies is

$$\frac{dM}{dt} = \frac{1}{3} \log(2)M - 80,000, \quad M(0) = 250,000.$$

We are asked to find $M(2)$.

This is a nonhomogeneous linear differential equation with normal form

$$\frac{dM}{dt} - \frac{1}{3} \log(2)M = -80,000.$$

Its integrating factor form is

$$\frac{d}{dt} \left(e^{-\frac{1}{3} \log(2)t} M \right) = -80,000 e^{-\frac{1}{3} \log(2)t}.$$

Upon integrating this equation while using the initial condition $M(0) = 250,000$ we obtain

$$e^{-\frac{1}{3} \log(2)t} M(t) - 250,000 = -80,000 \int_0^t e^{-\frac{1}{3} \log(2)s} ds = \frac{80,000}{\frac{1}{3} \log(2)} \left(e^{-\frac{1}{3} \log(2)t} - 1 \right).$$

Hence, we find that

$$\begin{aligned} M(t) &= \frac{80,000}{\frac{1}{3} \log(2)} + \left(250,000 - \frac{80,000}{\frac{1}{3} \log(2)} \right) e^{\frac{1}{3} \log(2)t} \\ &= \frac{240,000}{\log(2)} + \left(250,000 - \frac{240,000}{\log(2)} \right) 2^{\frac{1}{3}t}. \end{aligned}$$

Therefore after two weeks there are

$$M(2) = \frac{240,000}{\log(2)} + \left(250,000 - \frac{240,000}{\log(2)} \right) 4^{\frac{1}{3}} \text{ mosquitoes.}$$

This is about 193,500 mosquitoes, but you should leave the answer in the exact form above. It would be nice if the birds stuck around for several more weeks.

The ability of a individual to survive and reproduce depends upon the environment experienced by that individual. The assumption in the exponential model that $R(p)$ is constant assumes that this environment is unaffected by the number of individuals present. However once the population grows large enough there might be competition between its individuals for resources, which should reduce the growth rate. The simplest model that captures this effect takes $R(p)$ to be a linear function of p ,

$$R(p) = r - ap, \quad \text{for some positive constants } a \text{ and } r.$$

When there is no harvesting then the model becomes

$$\frac{dp}{dt} = (r - ap)p.$$

This is called the *logistic model*.

The logistic model is autonomous and so can be solved analytically. However, there are many questions you can address with a phase-line portrait. The stationary points of the equation are $p = 0$ and $p = \frac{r}{a}$. The phase-line portrait is as follows.

$$\begin{array}{ccccccc} & - & & + & & - & \\ \leftarrow \leftarrow \leftarrow \leftarrow \leftarrow & \bullet & \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow & \bullet & \leftarrow \leftarrow \leftarrow \leftarrow \leftarrow & h & \\ & 0 & & \frac{r}{a} & & & \end{array}$$

This portrait shows that if $p(t)$ is any solution with $p(0) > 0$ then $p(t) \rightarrow \frac{r}{a}$ as $t \rightarrow \infty$. In other words, the stationary point $\frac{r}{a}$ is attracting (asymptotically stable). This stationary point is called the *carrying capacity* of the ecosystem. It is the size of the population that the ecosystem naturally supports. Notice that any solution with $0 < p(0) < k$ has the interval of definition $(-\infty, \infty)$, while we expect that any solution with $k < p(0)$ an interval of definition of the form (t_L, ∞) for some $t_L < 0$.

If we let $k = r/a$ denote the carrying capacity of the ecosystem then the logistic model can be expressed as

$$\frac{dp}{dt} = r \left(1 - \frac{p}{k}\right) p.$$

Its stationary solutions are $p = 0$ and $p = k$. Its nonstationary solutions can be found by our recipe for autonomous equations. The phase-line portrait showed that solutions cannot cross the stationary point at $p = 0$, so solutions that are initially positive will remain positive. Such a solution is governed implicitly by

$$\begin{aligned} rt &= \int \frac{1}{\left(1 - \frac{p}{k}\right) p} dp = \int \frac{k}{(k-p)p} dp = \int \frac{1}{p} + \frac{1}{k-p} dp \\ &= \log(p) - \log(|k-p|) + c = \log\left(\frac{p}{|k-p|}\right) + c. \end{aligned}$$

Here we do not need absolute values on p because we are considering positive solutions.

If we impose the initial condition $p(0) = p_I > 0$ then we find that

$$0 = \log\left(\frac{p_I}{|k-p_I|}\right) + c,$$

so that $c = -\log(p_I/|k-p_I|)$. The implicit solution then becomes

$$rt = \log\left(\frac{p}{|k-p|}\right) - \log\left(\frac{p_I}{|k-p_I|}\right) = \log\left(\frac{p(k-p_I)}{p_I(k-p)}\right),$$

where we have dropped the absolute values because we know from the phase-line portrait that $k-p$ and $k-p_I$ will have the same sign. This can be exponentiated to find

$$e^{rt} = \frac{p(k-p_I)}{p_I(k-p)},$$

whereby $p_I(k-p)e^{rt} = p(k-p_I)$, which can be solved to obtain the explicit solution

$$p(t) = \frac{k e^{rt} p_I}{k + (e^{rt} - 1)p_I}.$$

Notice that if $0 < p_I < k$ then the interval of definition is $(-\infty, \infty)$ while if $k < p_I$ then the interval of definition is $(\log(1 - k/p_I)/r, \infty)$. This is consistent with what we expected from our earlier analysis of the phase-line portrait.

Remark. The logistic model is the simplest model for which the growth rate $R(p)$ is a decreasing function of p . However, as populations get smaller their growth rate might also decline. This can be due to the fact that a smaller herd might offer easier prey to predators, or that individuals in a more dispersed population might have a harder time finding mates. The simplest model that also captures this effect takes $R(p)$ to be a quadratic function of p ,

$$R(p) = r + bp - ap^2, \quad \text{for some constants } r, b \text{ and } a \text{ with } a > 0.$$

When there is no harvesting then the model becomes

$$\frac{dp}{dt} = (r + bp - ap^2) p.$$

Provided $R(p)$ is positive for some $p > 0$ then this equation has one positive stationary point when $r \geq 0$ and two positive stationary points when $r < 0$.