THIRTEEN

# Local Stability and Phase Portraits of Autonomous Differential Equations

We temporarily break from the study of optimal control theory in this chapter in order to present some essential results from the theory of autonomous ordinary differential equations. These results form the foundation upon which we will build our understanding and the construction of phase portraits, and consequently, a qualitative understanding of the solution of such a class of differential equations. This is important in dynamic economic theory, for phase portraits are ubiquitous owing to the fact that economists characteristically specify only the qualitative properties of the integrand and transition functions in an optimal control problem, rather than their functional forms. Accordingly, one cannot usually derive an explicit solution of an optimal control problem. Hence all that one can typically expect is a qualitative characterization of the solution in such instances. This is where the phase portrait comes in, for this is what it provides.

To begin, consider the following system of $N$ autonomous first-order nonlinear ordinary differential equations:

$$\begin{aligned} \dot{x}_1 &= f^1(x_1, x_2, \ldots, x_N), \\ \dot{x}_2 &= f^2(x_1, x_2, \ldots, x_N), \\ &\vdots \\ \dot{x}_N &= f^N(x_1, x_2, \ldots, x_N), \end{aligned} \tag{1}$$

or in vector notation, $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, where $\mathbf{x} \in \Re^N$. We assume throughout this chapter that $f^n(\cdot) \in C^{(1)}$, $n = 1, 2, \ldots, N$, in some domain $D$. Then by the fundamental existence and uniqueness theorem, if $\mathbf{x}_0 \in D$, there exists a unique solution $\mathbf{x} = \phi(t; t_0, \mathbf{x}_0)$ of Eq. (1) satisfying the initial conditions

$$\mathbf{x}(t_0) = \mathbf{x}_0. \tag{2}$$

The solution is defined in some interval $\alpha < t < \beta$ that contains the point $t_0$. The system of ordinary differential equations in Eq. (1) is *autonomous* because the independent variable $t$ does not appear explicitly in the functions $f^n(\cdot)$, $n = 1, 2, \ldots, N$.

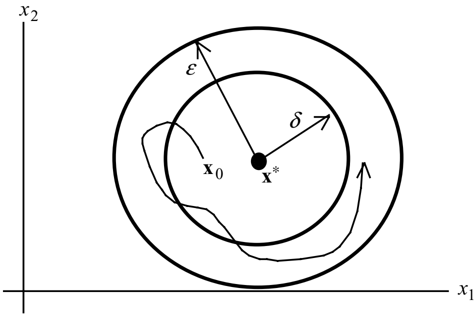


**Figure 13.1**

In economics, an autonomous system of differential equations is one in which the parameters of the system are *not* time dependent, nor is there exogenous technical change present in the system. Note that in what follows, we often elect to suppress the dependence of the solution on the parameters $t_0$ and $\mathbf{x}_0$ when they are not germane to the discussion, in which case, we write the solution of the Eq. (1) as $\mathbf{x} = \phi(t)$.

The first definition of this chapter introduces an important type of solution that will play a central role in our investigation of the qualitative properties of the solution to optimal control problems.

**Definition 13.1:** A constant value of $\mathbf{x}$, say, $\mathbf{x}^*$, is called a *fixed point* or *steady state* of the nonlinear differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ if $\mathbf{f}(\mathbf{x}^*) = 0$. Moreover, if there exists a neighborhood about a fixed point $\mathbf{x}^*$ in which there are no other fixed points, then $\mathbf{x}^*$ is called an *isolated* fixed point or steady state.

Definition 13.1 asserts that a fixed point or steady state $\mathbf{x}^*$ is a constant solution of the differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, since $\mathbf{f}(\mathbf{x}^*) = 0$ and thus $\dot{\mathbf{x}} = 0$ at such a point. Isolated fixed points are just that, since no other fixed points lie "close" to them. By way of a reminder, let us note that the Euclidean norm of a vector $\mathbf{z} \in \Re^N$ is defined as $\|\mathbf{z}\| \stackrel{\text{def}}{=} \sqrt{z_1^2 + z_2^2 + \cdots + z_N^2}$. With this in mind, we can now introduce the three notions of stability of fixed points that will prove most useful to us when we examine the qualitative properties of a solution to an optimal control problem.

**Definition 13.2:** An isolated fixed point $\mathbf{x}^*$ of the autonomous differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is said to be *stable* if, given any $\varepsilon > 0$, there exists a $\delta > 0$ such that every solution $\phi(t; t_0, \mathbf{x}_0)$ of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, which at $t = t_0$ satisfies $\|\phi(t_0; t_0, \mathbf{x}_0) - \mathbf{x}^*\| = \|\mathbf{x}_0 - \mathbf{x}^*\| < \delta$, implies that $\|\phi(t; t_0, \mathbf{x}_0) - \mathbf{x}^*\| < \varepsilon$ for all $t \geq t_0$.

This definition states that all solutions of the autonomous differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ that start "sufficiently close" to an isolated fixed point $\mathbf{x}^*$ stay "close" to $\mathbf{x}^*$. Note, however, that this definition of stability does not require that the solution approach the fixed point. Figure 13.1 gives an illustration of this definition in the plane, that is, for $N = 2$. Notice that the trajectory in Figure 13.1 starts

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

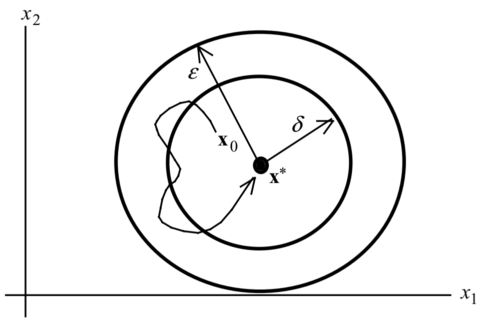


**Figure 13.2**

within the circle $(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2 = \delta^2$ at $t = t_0$, and although it eventually passes outside this circle, it remains within the larger radius circle $(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2 = \varepsilon^2$ for all $t \geq t_0$.

The next definition of stability builds upon that given in Definition 13.2. It turns out to be the type of stability that will be most prominent in the next five chapters.

**Definition 13.3:** An isolated fixed point $\mathbf{x}^*$ of the autonomous differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is said to be *locally asymptotically stable* if it is stable and if there exists a $\delta_0$, $0 < \delta_0 < \delta$, such that if a solution $\phi(t; t_0, \mathbf{x}_0)$ satisfies $\|\phi(t_0; t_0, \mathbf{x}_0) - \mathbf{x}^*\| = \|\mathbf{x}_0 - \mathbf{x}^*\| < \delta_0$, then $\lim_{t \to +\infty} \phi(t; t_0, \mathbf{x}_0) = \mathbf{x}^*$.

Definition 13.3 asserts that solutions of the autonomous differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ that start "sufficiently close" to an isolated fixed point $\mathbf{x}^*$ not only stay "close" to $\mathbf{x}^*$, but must eventually approach the fixed point $\mathbf{x}^*$ as $t$ approaches infinity. Note that local asymptotic stability is a stronger requirement than is stability, in view of the fact that a fixed point must be stable before one can even talk about whether or not it is locally asymptotically stable. On the other hand, the convergence requirement of local asymptotic stability does not, in and of itself, imply stability. This concept of stability is illustrated in Figure 13.2.

The third, and for our purposes the final, definition of stability is not one that we will make as much use of as Definition 13.3, but is nonetheless important given its prominence in dynamic economic theory.

**Definition 13.4:** An isolated fixed point $\mathbf{x}^*$ of the autonomous differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ is said to be *globally asymptotically stable* if it is stable and if the solution $\phi(t; t_0, \mathbf{x}_0)$ satisfies $\lim_{t \to +\infty} \phi(t; t_0, \mathbf{x}_0) = \mathbf{x}^*$ for *any* initial point $\mathbf{x}_0$.

This definition says that regardless of how close to or far from the initial point $\mathbf{x}_0$ is the fixed point $\mathbf{x}^*$, solutions of the autonomous differential equation system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ approach the fixed point $\mathbf{x}^*$ as $t$ approaches infinity. It should be clear, therefore, that if the fixed point $\mathbf{x}^*$ is globally asymptotically stable, then it is locally

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

asymptotically stable, but not vice versa. Moreover, because all solutions of the differential equation converge at the fixed point $\mathbf{x}^*$ regardless of their initial value $\mathbf{x}_0$, a globally asymptotically stable fixed point is unique. Finally, we note that a fixed point that is not stable is said to be *unstable*.

With these basic definitions behind us, let's now turn to the task of establishing several fundamental properties of systems of autonomous differential equations in a deliberate and clear fashion. Before getting into the details of this undertaking, however, we first present some definitions of terms and concepts that are essential to our purpose, and then illustrate them with uncomplicated examples.

Our next definition is central to developing a qualitative understanding of the solution to a system of autonomous differential equations. That is to say, it is an essential ingredient in constructing the phase portrait of a nonlinear and autonomous system of differential equations, in that it gives the direction of motion, or forces acting on a point, in the phase portrait.

**Definition 13.5:** The *vector field* of the autonomous differential equation system (1) is defined as follows. Imagine the vector

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f^1(x_1, x_2, \ldots, x_N) \\ f^2(x_1, x_2, \ldots, x_N) \\ \vdots \\ f^N(x_1, x_2, \ldots, x_N) \end{bmatrix}$$

drawn at every point $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in D$. This vector determines the tangent vector to the solution $\phi(t) = (\phi_1(t), \phi_2(t), \ldots, \phi_N(t))$ at every point.

The vector field of an autonomous system of differential equations is best visualized in the case $N = 2$, for then we can rather easily graph the vectors in the plane. The next example shows how this is done for a simple autonomous and nonlinear system of differential equations in the plane.

**Example 13.1:** Consider the autonomous and nonlinear system of differential equations

$$\dot{x}_1 = x_2^2,$$
$$\dot{x}_2 = x_1,$$

where $D = \Re^2$ in this case. In order to draw the vector field of this system of differential equations, Definition 13.5 directs us to plot the vector

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f^1(x_1, x_2) \\ f^2(x_1, x_2) \end{bmatrix} = \begin{bmatrix} x_2^2 \\ x_1 \end{bmatrix}$$

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

TABLE 13.1.

| $(x_1, x_2)$ | $\mathbf{f}(\mathbf{x})' = (x_2^2, x_1)$ |
|---|---|
| $(0, 0)$ | $(0, 0)$ |
| $(1, 0)$ | $(0, 1)$ |
| $(0, 1)$ | $(1, 0)$ |
| $(-1, 0)$ | $(0, -1)$ |
| $(0, -1)$ | $(1, 0)$ |
| $(2, 1)$ | $(1, 2)$ |
| $(1, 1)$ | $(1, 1)$ |
| $(-1, 1)$ | $(1, -1)$ |
| $(-2, 1)$ | $(1, -2)$ |
| $(1, -2)$ | $(4, 1)$ |

in the $x_1x_2$-plane for each $\mathbf{x} = (x_1, x_2) \in D = \Re^2$. To facilitate this, it is quite often convenient to make up a table with the values of the point $(x_1, x_2)$ in one column, and the values of the vector $\mathbf{f}(\mathbf{x})' = (x_2^2, x_1)$ in the other. Doing just that for ten different points in the $x_1x_2$-plane yields Table 13.1.

In order to use the information contained in Table 13.1 to draw the vector field corresponding to the above system of differential equations, note that the vector given in column 2 has its tail at the position in the $x_1x_2$-plane given by the corresponding point in column 1. A sketch of the resulting vector field is given in Figure 13.3. Observe that we have been careful to draw the length and direction of the vectors reasonably accurately so as to impart a better feel to the vector field.

By filling in the vector field at each point $(x_1, x_2)$ in the $x_1x_2$-plane, it would be possible, in principle, to construct the phase portrait of the above system. This, however, is not generally the best way to construct the phase portrait of a nonlinear

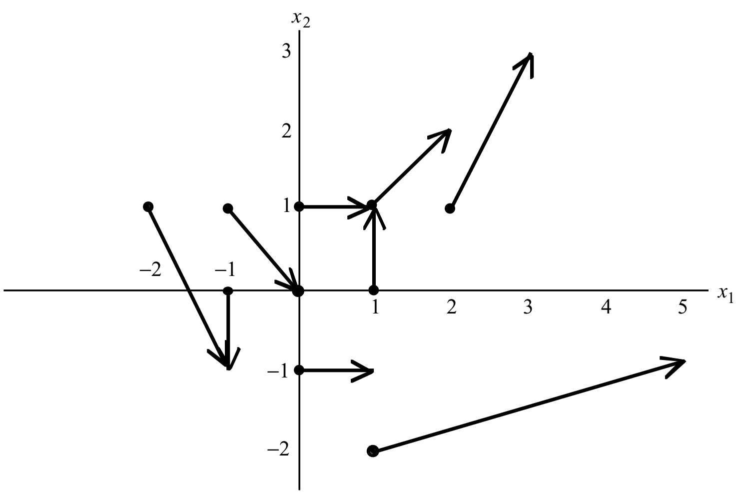


**Figure 13.3**

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

and autonomous system of differential equations, as we shall see. Doing so would not only be time consuming without the aid of a computer, but it would hopelessly clutter up the picture. Because we have chosen only ten points at which to compute the vector field, we have essentially presented only a partial picture of it.

Now consider the following two definitions.

**Definition 13.6:** A *trajectory* of an autonomous system of differential equations is the curve $\{(\phi_1(t), \phi_2(t), \ldots, \phi_N(t)) : \alpha < t < \beta\}$, where $\phi(t) = (\phi_1(t), \phi_2(t), \ldots, \phi_N(t))$ is the solution of the autonomous system of differential equations.

**Definition 13.7:** A *phase portrait* or *phase diagram* is the collection of all trajectories.

With these two definitions, we are now in a position to construct the phase portrait of an autonomous system of differential equations. Let's therefore return to the autonomous system of differential equations in Example 13.1 and draw its phase portrait. In the process of constructing the phase diagram, we will see how the vector field aids in this endeavor, as remarked above.

**Example 13.2:** Recall the autonomous and nonlinear system of differential equations

$$\dot{x}_1 = x_2^2,$$
$$\dot{x}_2 = x_1$$

from Example 13.1. Also recall that a partial sketch of the vector field is given in Figure 13.3. From it, we get some indication of the trajectories of this system, and thus a hint at the nature of the phase portrait. To complete the phase portrait, however, we must add to the information contained in the vector field. This is accomplished by dividing $\dot{x}_2 = x_1$ by $\dot{x}_1 = x_2^2$ to get the first-order differential equation for $x_2 = g(x_1)$:

$$\frac{\dot{x}_2}{\dot{x}_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{dx_2}{dx_1} = \frac{x_1}{x_2^2}.$$

The procedure just used to get a first-order differential equation for $x_2 = g(x_1)$ is quite useful for sketching the trajectories of autonomous systems. We will discuss it more generally and more rigorously shortly.

Separating the variables of the above first-order differential equation gives $x_2^2 dx_2 = x_1 dx_1$, and integrating yields $\frac{1}{3}x_2^3 = \frac{1}{2}x_1^2 + c_1$, where $c_1$ is an arbitrary constant. We can equivalently rewrite the general solution as $x_2 = [\frac{3}{2}x_1^2 + c]^{\frac{1}{3}}$, where $c \stackrel{\text{def}}{=} 3c_1$. Hence, every trajectory of the above system of autonomous differential equations must lie on the graph of the function $x_2 = [\frac{3}{2}x_1^2 + c]^{\frac{1}{3}}$, because

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

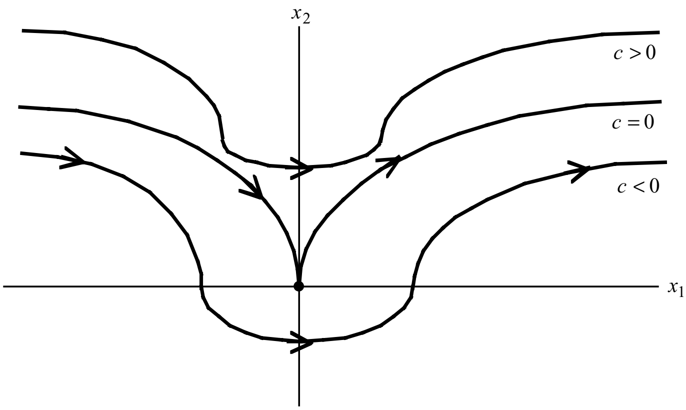


**Figure 13.4**

every point $(x_1, x_2) = (\phi_1(t), \phi_2(t))$ such that $(\phi_1(t), \phi_2(t))$ is a solution of the system of differential equations must satisfy this equation. The general solution $x_2 = [\frac{3}{2}x_1^2 + c]^{\frac{1}{3}}$ therefore permits us to rather easily construct the phase portrait of the system of differential equations, which we have done in Figure 13.4. Note that it is consistent with the vector field in Figure 13.3.

A question of some interest in this case is this: When does the solution escape to $+\infty$? The answer is almost always, as inspection of Figure 13.4 reveals. More precisely, the solution of the above system escapes to $+\infty$ except when the initial condition satisfies the equation $x_2 = [\frac{3}{2}x_1^2]^{\frac{1}{3}}$ for $x_1 \leq 0$, as this is the only trajectory that approaches the origin.

Before presenting the five fundamental results for autonomous systems of differential equations, let's consider one more example and derive its phase portrait.

**Example 13.3:** Consider the linear and autonomous system of differential equations

$$\dot{x}_1 = x_2,$$
$$\dot{x}_2 = -x_1.$$

In order to derive its phase portrait, first note that the origin is the only fixed point of this system. Proceeding as we did in Example 13.2, form the quotient

$$\frac{\dot{x}_2}{\dot{x}_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{dx_2}{dx_1} = \frac{-x_1}{x_2}.$$

Separating the variables gives $x_2 dx_2 = -x_1 dx_1$, and integrating yields $\frac{1}{2}x_2^2 = -\frac{1}{2}x_1^2 + c_1$, where $c_1$ is an arbitrary constant. Alternatively, we can rewrite the general solution equivalently as $x_1^2 + x_2^2 = c$, where $c \stackrel{\text{def}}{=} 2c_1$. This is the equation of a circle, which any trajectory other than the fixed point (0, 0) must lie on for $c \neq 0$.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

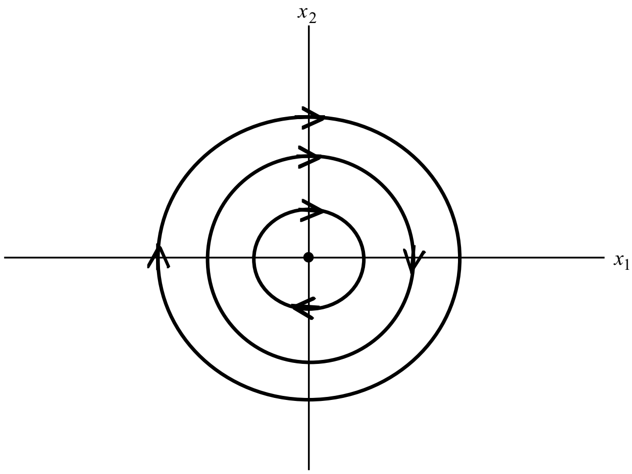


**Figure 13.5**

The fixed point of the above linear system is therefore called a *center*, and its phase portrait is given in Figure 13.5.

To determine whether the motion on the trajectories is clockwise or counter-clockwise, all we need to do is compute the vector field along one of the axes. For example, at the point $(x_1, x_2) = (1, 0)$, the vector field is given by $(\dot{x}_1, \dot{x}_2) = (x_2, -x_1) = (0, -1)$, hence the arrow pointing vertically downward along the positive part of the $x_1$-axis. Similarly, at the point $(x_1, x_2) = (0, 1)$, the vector field is given by $(\dot{x}_1, \dot{x}_2) = (x_2, -x_1) = (1, 0)$, hence the arrow pointing horizontally rightward along the positive part of the $x_2$-axis. We therefore conclude that the motion is clockwise on the trajectories.

Let's now turn to our first result for system (1). Geometrically, it says that we get another solution to an autonomous system of differential equations by translating the solution curve in the $t$-direction.

**Theorem 13.1:** *If $\phi(t)$ is a solution to the autonomous system of differential equations (1) for $t \in (\alpha, \beta)$, then so is $\phi(t + c)$ for $t \in (\alpha - c, \beta - c)$, where $c$ is a constant.*

***Proof:*** Differentiating $\phi(t + c)$ with respect to the independent variable $t$, using the chain rule and the fact that $c$ is a constant, gives

$$\frac{\partial}{\partial t}\phi(t+c) = \frac{d}{d(t+c)}\phi(t+c)\frac{\partial}{\partial t}(t+c) = \frac{d}{d(t+c)}\phi(t+c) = \dot{\phi}(t+c),$$

where we have used the definition $\dot{\phi}(\tau) \stackrel{\text{def}}{=} \frac{d}{d\tau}\phi(\tau)$. This equation asserts that increments to the independent variable $t$ have the same effect on $\phi(t + c)$ as do increments in the variable $t + c$. Because $\dot{\phi}(t) \equiv \mathbf{f}(\phi(t))$ for all $t \in (\alpha, \beta)$ by its

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

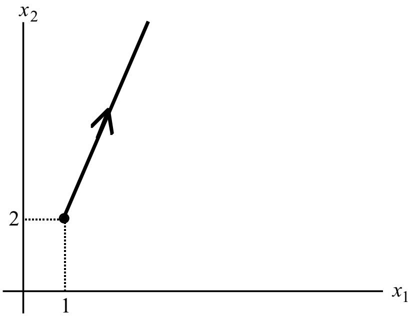


**Figure 13.6**

definition as a solution of system (1), it follows that $\dot{\phi}(t+c) \equiv \mathbf{f}(\phi(t+c))$ for all $t \in (\alpha - c, \beta - c)$, therefore proving that $\phi(t+c)$ is a solution to system (1), for all $t \in (\alpha - c, \beta - c)$. Q.E.D.

This property of autonomous systems of differential equations is of such fundamental importance that we pause for a moment and examine it in more detail by way of the following example. The example is intended to help build intuition about autonomous systems, whether linear or nonlinear.

**Example 13.4:** Suppose that particles are being continuously emitted at the point $x_1 = 1$ and $x_2 = 2$, and then move in the $x_1x_2$-phase according to the law

$$\dot{x}_1 = x_1, \ \dot{x}_2 = x_2. \tag{3}$$

The particle emitted at the time $t = s$ is specified by the initial conditions $x_1(s) = 1$ and $x_2(s) = 2$. Separating the variables of the differential equations in Eq. (3), integrating, and then applying the initial conditions yields the solution

$$x_1 = \phi_1(t; s) = e^{t-s}, \ x_2 = \phi_2(t; s) = 2e^{t-s}, \ t \geq s. \tag{4}$$

The trajectories or paths followed by the particles emitted at different times $s$ can be best visualized by eliminating $t$ from Eq. (4). Doing just that yields the straight line

$$x_2 = 2x_1. \tag{5}$$

It should be clear from Eq. (4) that $x_1 \geq 1$ and $x_2 \geq 2$ for all $t \geq s$. The portion of the straight line determined by Eq. (5) corresponding to Eq. (4), with the direction of motion indicated by an arrow, is given in Figure 13.6.

By far the most important feature of Eq. (5) is that the initial time $s$ does not appear in it. This means that no matter when a particle is emitted from the point $(x_1, x_2) = (1, 2)$, it *always* moves along the same curve, videlicet, the straight line given by Eq. (5).

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

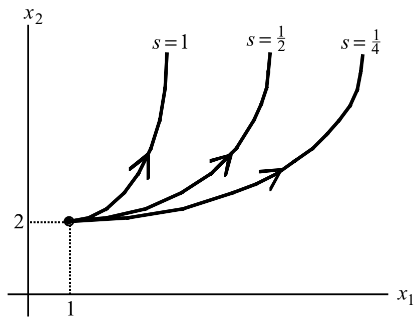


**Figure 13.7**

Now consider a similar pair of differential equations, namely,

$$\dot{x}_1 = t^{-1}x_1,\ \ \dot{x}_2 = x_2. \tag{6}$$

In contrast to Eq. (3), Eq. (6) is a pair of *nonautonomous* differential equations because of the explicit appearance of the independent variable $t$ on the right-hand side. Separating the variables of the differential equations in Eq. (6), integrating, and then applying the initial conditions $x_1(s) = 1$ and $x_2(s) = 2$ yields the solution

$$x_1 = \phi_1(t; s) = s^{-1}t,\ \ x_2 = \phi_2(t; s) = 2e^{t-s},\ \ t \geq s. \tag{7}$$

Solving the first of these equations for $t$ yields $t = sx_1$. Substituting this into the second equation yields a result analogous to Eq. (5), namely,

$$x_2 = 2e^{s(x_1-1)}. \tag{8}$$

It should be clear from Eq. (7) that $x_1 \geq 1$ and $x_2 \geq 2$ for all $t \geq s$. Because the initial time $s$ appears in Eq. (8), the path that a particle follows depends on the time at which it is emitted, that is, the time at which the initial condition is applied. Figure 13.7 highlights this feature of nonautonomous differential equations by showing the paths followed by particles emitted at several different initial times.

We are now in a position to compare our results for the differential equation systems given in Eqs. (3) and (6). In the case of the autonomous system (3), when we eliminated the independent variable $t$ in deriving Eq. (5), we automatically eliminated the initial time $s$. This is a result of the fact that $t$ and $s$ appear only in the form $t - s$ in Eq. (4). For the nonautonomous system (6), on the other hand, when we eliminated the independent variable $t$ in deriving Eq. (8), we were not able to eliminate the initial time $s$.

This example illustrates the essence of Theorem 13.4. In particular, it shows that autonomous systems of differential equations exhibit a special property not shared by nonautonomous systems of differential equations, to wit, that all particles passing through a given point follow the same trajectory in the phase plane. In other words, the same trajectory in the phase plane is represented parametrically by many

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

different solutions differing from one another by a translation of the independent variable $t$.

In determining the trajectories of an autonomous system of differential equations in the plane, that is, the case of $N = 2$ for system (1), it is often quite useful to eliminate the independent variable $t$ from the solution $x_n = \phi_n(t)$, $n = 1, 2$. Doing so results in a relation (and sometimes a function) between $x_1$ and $x_2$ describing the trajectory in the phase plane. We essentially did this when we derived Eq. (5) from the autonomous system (3) in Example 13.4. Another way to eliminate the independent variable $t$ from the solution in Example 13.4 is to write down the ratio of $\dot{x}_2$ to $\dot{x}_1$ using system (3). This yields

$$\frac{\dot{x}_2}{\dot{x}_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{dx_2}{dx_1} = \frac{x_2}{x_1}, \tag{9}$$

which is a first-order differential equation for $x_2 = g(x_1)$. Separating the variables in Eq. (9) and integrating using the initial conditions $x_1(s) = 1$ and $x_2(s) = 2$ yields Eq. (5), just as anticipated. Note that in order to form Eq. (9), it must be that $x_1 \neq 0$, which, as noted in Example 13.4, holds in this instance.

More generally, for a pair of autonomous differential equations $\dot{x}_n = f^n(x_1, x_2)$, $n = 1, 2$, we can proceed to eliminate the independent variable $t$ from the solution $x_n = \phi_n(t)$, $n = 1, 2$, as follows. In a region where $f^1(x_1, x_2) \neq 0$, we can form the ratio

$$\frac{\dot{x}_2}{\dot{x}_1} = \frac{dx_2/dt}{dx_1/dt} = \frac{dx_2}{dt}\frac{dt}{dx_1} = \frac{dx_2}{dx_1} = \frac{f^2(x_1, x_2)}{f^1(x_1, x_2)}, \tag{10}$$

which is a first-order differential equation for $x_2 = g(x_1)$. The requirement that $f^1(x_1, x_2) \neq 0$ is needed so that we can divide by $f^1(x_1, x_2)$ to form Eq. (10). Moreover, because $\dot{x}_1 = f^1(x_1, x_2)$, the condition $f^1(x_1, x_2) \neq 0$ is equivalent to $\dot{x}_1 = dx_1/dt \neq 0$. By the implicit function theorem, this then implies that the solution $x_1 = \phi_1(t)$ can be solved for $t$ as a function of $x_1$. This, in turn, implies that $x_2$ becomes a function of $x_1$, as in, for example, Eq. (5). The one-parameter family of solutions of Eq. (10) is the set of trajectories of the system $\dot{x}_n = f^n(x_1, x_2)$, $n = 1, 2$. Equation (10) is especially convenient for determining the slope of a trajectory at a point in the phase plane, since it is an explicit representation of it. In a region where $f^2(x_1, x_2) \neq 0$, one can alternatively form the analogous differential equation to Eq. (10), that is to say,

$$\frac{\dot{x}_1}{\dot{x}_2} = \frac{dx_1/dt}{dx_2/dt} = \frac{dx_1}{dx_2} = \frac{f^1(x_1, x_2)}{f^2(x_1, x_2)}. \tag{11}$$

If both $f^1(x_1, x_2) \neq 0$ and $f^2(x_1, x_2) \neq 0$, then one can form either Eq. (10) or Eq. (11) to determine the trajectories of the system of autonomous differential equations.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

If, however, there is a point such that $f^1(x_1, x_2) = 0$ *and* $f^2(x_1, x_2) = 0$, then we cannot solve for either $dx_2/dx_1$ or $dx_1/dx_2$. Recall that a point such that $f^1(x_1, x_2) = 0$ and $f^2(x_1, x_2) = 0$ is defined as a fixed point or steady state. Thus we see that fixed points have special significance in the study of differential equations. Also recall that if $\mathbf{x}^*$ is a fixed point of system (1), then $\mathbf{x} = \mathbf{x}^*$ is a solution of system (1), albeit a constant one. Moreover, it follows from the existence and uniqueness theorem that the only solution of system (1) passing through the fixed point $\mathbf{x}^*$ is the constant solution $\mathbf{x} = \mathbf{x}^*$ itself. The trajectory of this solution is just the fixed point $\mathbf{x}^*$. A particle at the point $\mathbf{x}^*$ is thus often said to be at rest, or in equilibrium. Let us also remark that a trajectory of system (1), which is represented by the solution $\mathbf{x} = \phi(t)$, $t \geq \alpha$, is said to approach the fixed point $\mathbf{x}^*$ as $t \to +\infty$ if $\phi(t) \to \mathbf{x}^*$ as $t \to +\infty$.

The importance of fixed points in the analysis of autonomous systems of differential equations is a result of the fact that (a) they are constant solutions of the system, and (b) the qualitative behavior of all trajectories in the phase plane is determined to a considerable degree by the location of the fixed points and the behavior of the trajectories near them.

Let us now turn to the four remaining fundamental properties of autonomous systems. Seeing as their proofs are relatively simple and direct, we provide them. After presenting and proving these results, we provide a summary of their implications for the study of autonomous systems of differential equations.

**Theorem 13.2:** *The trajectories of autonomous systems of differential equations do not intersect.*

***Proof:*** Assume that there are two solutions of system (1), say $\phi^1(t)$ and $\phi^2(t)$, that is,

$$\dot{\phi}^1(t) = \mathbf{f}(\phi^1(t)),$$
$$\dot{\phi}^2(t) = \mathbf{f}(\phi^2(t)).$$

For the trajectories to intersect, there must exist a point $\mathbf{x}_0$ such that

$$\phi^1(0) = \mathbf{x}_0,$$
$$\phi^2(t_0) = \mathbf{x}_0.$$

By Theorem 13.1, we know that $\mathbf{y}(t) \stackrel{\text{def}}{=} \phi^2(t + t_0)$ is also a solution. Hence, we may conclude that $\phi^1(0) = \mathbf{x}_0$ and $\mathbf{y}(0) \stackrel{\text{def}}{=} \mathbf{x}_0$. By uniqueness, $\phi^1(t) = \mathbf{y}(t)$ for every $t$, which, by the definition of $\mathbf{y}(\cdot)$, is equivalent to $\phi^1(t) = \phi^2(t + t_0)$ for every $t$. Thus the trajectory determined by $\phi^1(\cdot)$ is the same as the trajectory determined by $\phi^2(\cdot)$. Q.E.D.

This theorem is essentially a corollary to the fundamental existence and uniqueness theorem. Its content is beautifully straightforward: *different trajectories never*

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

*intersect*. If two trajectories did intersect, then there would be two solutions starting from the same point, scilicet, the point at which the two trajectories cross. This, however, violates the uniqueness part of the fundamental existence and uniqueness theorem. This argument, in essence, constitutes the proof of Theorem 13.2. Said in a more geometrical manner, Theorem 13.2 asserts that a trajectory can't move in two directions at once. Because trajectories for autonomous systems can't intersect, the associated phase portraits always have a neat and tidy appearance. Otherwise, they might degenerate into an entanglement of curves. The fundamental existence and uniqueness theorem prevents such a mess from happening.

Let's return to our task of laying out the fundamental properties of systems of autonomous differential equations. The next two theorems are especially important in dynamic economic theory, as we shall see in the next five chapters.

**Theorem 13.3:** *Let $\phi(t)$ be a solution to system (1). If $\lim_{t\to+\infty}\phi(t) = \mathbf{x}^*$, then $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}_N$, so that $\mathbf{x}^*$ is a fixed point or steady state of system (1).*

***Proof:*** Given that $\lim_{t\to+\infty}\phi(t) = \mathbf{x}^*$, it follows that

$$\lim_{t\to+\infty}[\phi(t+1) - \phi(t)] = \mathbf{0}_n. \tag{12}$$

On the other hand,

$$\lim_{t\to+\infty}[\phi(t+1) - \phi(t)] = \lim_{t\to+\infty}\left[\int_0^1 \dot{\phi}(t+\tau)d\tau\right] = \lim_{t\to+\infty}\left[\int_0^1 \mathbf{f}(\phi(t+\tau))\,d\tau\right]. \tag{13}$$

Using Eq. (13) and the facts that $t+\tau \to +\infty$ as $t \to +\infty$ and $\phi(t+\tau) \to \mathbf{x}^*$ as $t \to +\infty$ permits us to reach the conclusion that

$$\lim_{t\to+\infty}[\phi(t+1) - \phi(t)] = \mathbf{f}(\mathbf{x}^*). \tag{14}$$

Equations (12) and (14) then imply that $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}_N$, so that $\mathbf{x}^*$ is a fixed point. Q.E.D.

Theorem 13.3 is actually more general than it may appear at first glance. The reason is that uniqueness of the solution is not a prerequisite to its conclusion, as is evident from the statement of the theorem ($\phi(t)$ is *a* solution, not *the* solution) and the proof (it was not used in it). This theorem simply asserts that if a solution of system (1) converges to a constant value in the limit as $t \to +\infty$, then that constant value is a fixed point of system (1). The next theorem is of a similar character, but relies on uniqueness of the solution, and therefore is less general, but often more useful, in intertemporal economic problems.

**Theorem 13.4:** *If the unique solution $\phi(t)$ of system (1) begins at a point $\mathbf{x}_0$ that is not a fixed point of the system, then it cannot reach a fixed point $\mathbf{x}^*$ in a finite length of time.*

***Proof:*** We employ a contrapositive proof. To that end, assume that the unique solution $\phi(t)$ has reached the fixed point $\mathbf{x}^*$ in a finite length of time, say $\phi(\tau) = \mathbf{x}^*$ for some finite $\tau$. Now recall that the fixed point $\mathbf{x}^*$ is a (constant) solution of system (1) satisfying the initial condition $\mathbf{x}(\tau) = \mathbf{x}^*$. By uniqueness $\phi(t) = \mathbf{x}^*$ for all $t$ is the only solution of system (1) satisfying the initial condition $\mathbf{x}(\tau) = \mathbf{x}^*$. Thus the initial value of the solution $\phi(t)$ is the fixed point $\mathbf{x}^*$, contradicting the hypothesis of the theorem, and thereby completing its proof. Q.E.D.

Simply put, this theorem asserts that if $\mathbf{x}^*$ is a fixed point of an autonomous system and a solution of the system approaches $\mathbf{x}^*$, then necessarily $t \to +\infty$. The final fundamental result about autonomous systems concerns periodic solutions.

**Theorem 13.5:** *Let $\phi(t)$ be the solution to system (1). If $\phi(0) = \phi(T)$, then $\phi(t) = \phi(t+T)$ for every $t$, and thus the solution is periodic with period $T$.*

***Proof:*** By Theorem 13.1, $\mathbf{y}(t) \stackrel{\text{def}}{=} \phi(t+T)$ is a solution to system (1). Because $\phi(0) = \mathbf{y}(0)$, by uniqueness it then follows that $\phi(t) = \mathbf{y}(t)$, or $\phi(t) = \phi(t+T)$, for every $t$. Q.E.D.

The implications of Theorems 13.1 through 13.5 are of fundamental importance in the study of autonomous systems of differential equations. They essentially say that if a solution starts at a point that is not a fixed point, then it moves on the same trajectory no matter at what time it starts, it can never come back to its initial point unless the motion is periodic, it can never cross another trajectory, and it can only "reach" a fixed point or steady state in the limit as $t \to +\infty$. These theorems thus suggest that a solution of an autonomous system of differential equations either approaches a fixed point, moves on a closed trajectory or approaches a closed trajectory as $t \to +\infty$, or else goes off to infinity. Moreover, they demonstrate that for autonomous systems of differential equations, the study of the fixed points and periodic solutions is of fundamental importance. Actually, other much less common behaviors may occur as well, but this is not something we are equipped to explore.

Let us now turn to the analytical determination of the stability of fixed points for systems of nonlinear and autonomous ordinary differential equations. In doing so, we will also extend our understanding of phase portraits by making use of our knowledge about the fundamental properties of autonomous differential equations exposited in Theorems 13.1 through 13.5.

As is typical when one studies nonlinear functions, the calculus is often initially employed to glean information about the nonlinear function at a given point. This information is then used to deduce properties of the nonlinear function in a neighborhood of the given point. It is no different with nonlinear differential equations. The basic idea is to replace the rather complicated system of nonlinear differential equations with a linear system of differential equations, with the hope that the linear system is a good enough approximation so that we can infer the behavior of the nonlinear system from that of its linear approximation. We will see that in most instances, the study of the simpler linear system corresponding to the nonlinear one is of value in deducing the local properties of the nonlinear system. Thus we begin with the so-called method of linearization at a fixed point, a central method of local stability analysis of nonlinear differential equation systems.

The basic idea behind the linearization method is really quite simple: replace the nonlinear differential equation system with a linear approximation to it at a fixed point of the nonlinear system. Then essentially use theorems about linear systems to deduce the phase portrait of the nonlinear system in a neighborhood of the fixed point. If there is more than one fixed point of the nonlinear system, which is not at all unusual, then one applies the method of linearization at each of the fixed points. Because the idea of approximation is important in what follows, it is advisable at this juncture to review the multivariate version of Taylor's theorem. We will concentrate on the case $N = 2$ seeing as this is the most common case encountered in dynamic economic theory, at least initially, and the case in which the geometry is best developed. The definitions and theorems, however, may be stated for the general system (1) when such generality creates no additional burden over the case $N = 2$.

Let $(x_1^*, x_2^*) \in D$ be a fixed point of system (1) when $N = 2$. Given that $f^n(\cdot) \in C^{(1)}$ for all $(x_1, x_2) \in D$, $n = 1, 2$, we can apply Taylor's theorem to these functions at the fixed point $(x_1^*, x_2^*)$ to get

$$\begin{aligned} f^n(x_1, x_2) = f^n(x_1^*, x_2^*) &+ \frac{\partial f^n}{\partial x_1}(x_1^*, x_2^*)[x_1 - x_1^*] + \frac{\partial f^n}{\partial x_2}(x_1^*, x_2^*)[x_2 - x_2^*] \\ &+ R^n(x_1, x_2), \quad n = 1, 2. \end{aligned} \tag{15}$$

Recall that under our assumptions, the so-called remainder functions $R^n(\cdot)$ satisfy

$$\lim_{r \to 0} \frac{R^n(x_1, x_2)}{r} = 0, \quad n = 1, 2, \tag{16}$$

where $r \stackrel{\text{def}}{=} \sqrt{(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2}$ is the Euclidean distance of the point $(x_1, x_2) \in D$ from the fixed point $(x_1^*, x_2^*) \in D$. Essentially, the functions $R^n(\cdot)$ are expressions containing the quadratic or higher-order terms $(x_1 - x_1^*)^2$, $(x_2 - x_2^*)^2$, and $(x_1 - x_1^*)(x_2 - x_2^*)$; hence that is why they go to zero faster than does $r \stackrel{\text{def}}{=} \sqrt{(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2}$. Note that if $(x_1 - x_1^*)$ and $(x_2 - x_2^*)$ are small, then the quadratic terms are extremely small.

Because $(x_1^*, x_2^*)$ is a fixed point of system (1), it follows from the very definition of a fixed point that $f^n(x_1^*, x_2^*) = 0$, $n = 1, 2$. Using this conclusion in Eq. (15), and then substituting Eq. (15) into system (1) yields the equivalent system

$$\dot{x}_1 = \frac{\partial f^1}{\partial x_1}(x_1^*, x_2^*)[x_1 - x_1^*] + \frac{\partial f^1}{\partial x_2}(x_1^*, x_2^*)[x_2 - x_2^*] + R^1(x_1, x_2),$$
$$\dot{x}_2 = \frac{\partial f^2}{\partial x_1}(x_1^*, x_2^*)[x_1 - x_1^*] + \frac{\partial f^2}{\partial x_2}(x_1^*, x_2^*)[x_2 - x_2^*] + R^2(x_1, x_2). \quad (17)$$

Note that because the partial derivatives of the functions $f^n(\cdot)$, $n = 1, 2$, are evaluated at the fixed point $(x_1^*, x_2^*)$, they are numbers in Eqs. (17), not functions. In going from system (1) to system (17), we have replaced the nonlinear functions $f^n(\cdot)$, $n = 1, 2$, with the sum of a linear expression, given by the first two terms on the right-hand side of system (17), and a residual nonlinear term that is extremely small compared with the linear part of the system.

By defining the new variables $y_n \stackrel{\text{def}}{=} x_n - x_n^*$, $n = 1, 2$, we move the fixed point from $(x_1^*, x_2^*)$ to the origin, as is straightforward to verify. Moreover, as $\dot{y}_n = \dot{x}_n$, $n = 1, 2$, we can rewrite system (17) in terms of the new coordinates as follows:

$$\dot{y}_1 = \frac{\partial f^1}{\partial x_1}(x_1^*, x_2^*)y_1 + \frac{\partial f^1}{\partial x_2}(x_1^*, x_2^*)y_2 + R^1(y_1 + x_1^*, y_2 + x_2^*),$$
$$\dot{y}_2 = \frac{\partial f^2}{\partial x_1}(x_1^*, x_2^*)y_1 + \frac{\partial f^2}{\partial x_2}(x_1^*, x_2^*)y_2 + R^2(y_1 + x_1^*, y_2 + x_2^*). \quad (18)$$

This is just system (17) with the fixed point moved to the origin.

Next, define the *Jacobian matrix* $\mathbf{J}(x_1^*, x_2^*)$ of the functions $f^n(\cdot)$, $n = 1, 2$, evaluated at the fixed point $(x_1^*, x_2^*)$ by

$$\mathbf{J}(x_1^*, x_2^*) \stackrel{\text{def}}{=} \begin{bmatrix} \frac{\partial f^1}{\partial x_1}(x_1^*, x_2^*) & \frac{\partial f^1}{\partial x_2}(x_1^*, x_2^*) \\ \frac{\partial f^2}{\partial x_1}(x_1^*, x_2^*) & \frac{\partial f^2}{\partial x_2}(x_1^*, x_2^*) \end{bmatrix}. \quad (19)$$

The Jacobian matrix $\mathbf{J}(x_1^*, x_2^*)$ is just the multivariate analog of the derivative $f'(x^*)$ corresponding to the single nonlinear ordinary differential equation $\dot{x} = f(x)$. In view of the fact that one may use the derivative $f'(x^*)$ to determine the stability of a fixed point $x^*$ in the scalar case, it is natural to conjecture that the Jacobian matrix $\mathbf{J}(x_1^*, x_2^*)$ may be similarly used to deduce the stability of a nonlinear system of differential equations. This conjecture is correct for the most part, as we shall shortly see.

Because the higher-order terms $R^n(x_1, x_2)$, $n = 1, 2$, are small, it would appear to be rather appealing to neglect them altogether. If in fact we do so, then we obtain the *linearized system* corresponding to the original nonlinear system (1), to wit,

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \end{bmatrix} = \begin{bmatrix} \frac{\partial f^1}{\partial x_1}(x_1^*, x_2^*) & \frac{\partial f^1}{\partial x_2}(x_1^*, x_2^*) \\ \frac{\partial f^2}{\partial x_1}(x_1^*, x_2^*) & \frac{\partial f^2}{\partial x_2}(x_1^*, x_2^*) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{20}$$

The goal at this juncture is to use our prior understanding of the behavior of the linearized system to come to some understanding about the behavior of the original nonlinear system in a neighborhood of a fixed point. Before presenting the first theorem on using the linearized system to study the qualitative properties of the original nonlinear system, however, we require a few definitions. We begin with one that should be somewhat familiar to you from your prior study of linear systems of ordinary differential equations.

**Definition 13.8:** A fixed point $\mathbf{x}^*$ of a nonlinear system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in D \subseteq \Re^N$, is said to be *simple* if the $N \times N$ Jacobian matrix $\mathbf{J}(\mathbf{x}^*)$ of its linearized system has no zero eigenvalues, where

$$\mathbf{J}(\mathbf{x}^*) \stackrel{\text{def}}{=} \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}^*) = \begin{bmatrix} \frac{\partial f^1}{\partial x_1}(\mathbf{x}^*) & \frac{\partial f^1}{\partial x_2}(\mathbf{x}^*) & \cdots & \frac{\partial f^1}{\partial x_N}(\mathbf{x}^*) \\ \frac{\partial f^2}{\partial x_1}(\mathbf{x}^*) & \frac{\partial f^2}{\partial x_2}(\mathbf{x}^*) & \cdots & \frac{\partial f^2}{\partial x_N}(\mathbf{x}^*) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f^N}{\partial x_1}(\mathbf{x}^*) & \frac{\partial f^N}{\partial x_2}(\mathbf{x}^*) & \cdots & \frac{\partial f^N}{\partial x_N}(\mathbf{x}^*) \end{bmatrix}.$$

This definition extends the idea of simplicity for linear systems, which requires that the coefficient matrix be nonsingular, to the fixed points of a nonlinear system. An equivalent statement of simplicity for nonlinear systems is that $|\mathbf{J}(\mathbf{x}^*)| \neq 0$, or equivalently, that $\mathbf{J}(\mathbf{x}^*)$ is nonsingular, seeing as $\prod_{n=1}^{N} \lambda_n = |\mathbf{J}(\mathbf{x}^*)|$, where $\lambda_n$, $n = 1, 2, \ldots, N$, are the eigenvalues of $\mathbf{J}(\mathbf{x}^*)$.

Because a great deal is known about the trajectories of the linearized system (20), it would be outstanding to find out that by using the linearized system instead of the original nonlinear system, we could come to some understanding of the behavior of the trajectories of the nonlinear system in a neighborhood of the fixed point. That is, how safe is it to ignore the higher-order terms $R^n(x_1, x_2)$, $n = 1, 2$, in the system (18)? In other words, does the linearized system give a qualitatively correct depiction of the phase portrait of the nonlinear system near the fixed point? The

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

following theorem, whose proof we omit, provides an affirmative answer to these questions, but with some important qualifications that we will subsequently explore.

**Theorem 13.6:** *Let* $(x_1^*, x_2^*) \in D$ *be an isolated and simple fixed point of the nonlinear system* $\dot{x}_n = f^n(x_1, x_2)$, $n = 1, 2$, *where* $f^n(\cdot) \in C^{(1)}$ *for all* $(x_1, x_2) \in D$, $n = 1, 2$, *and let* $\lambda_n$, $n = 1, 2$, *be the eigenvalues of the Jacobian matrix* $\mathbf{J}(x_1^*, x_2^*)$ *of the corresponding linearized system. If* $\lambda_n$, $n = 1, 2$, *are real and unequal, or complex conjugates with nonzero real parts, then the type of the fixed point is correctly predicted by the linearized system and so is its stability in a neighborhood of* $(x_1^*, x_2^*)$.

To fully grasp this theorem, let us begin by recalling what it means for the type and stability of a fixed point in a linear system when the eigenvalues are structured as in Theorem 13.6. If the eigenvalues are real, negative, and unequal, then the fixed point is a globally asymptotically stable node of the linearized system, whereas if the eigenvalues are real, positive, and unequal, then the fixed point is an unstable node of the linearized system. If, however, the eigenvalues are complex conjugates with negative real parts, then the fixed point is a globally asymptotically stable spiral node of the linearized system, whereas if the eigenvalues are complex conjugates with positive real parts, then the fixed point is an unstable spiral node of the linearized system. Finally, if the eigenvalues are real and of the opposite sign, then the fixed point is an unstable saddle point of the linearized system. Thus Theorem 13.6 states that if the linearized system determines that a fixed point of the nonlinear system is a globally asymptotically stable or unstable node or spiral node, or an unstable saddle point, then the fixed point *really is* of the type so determined by the linearized system, and its stability is accurately predicted in a neighborhood of the fixed point. The reason we cannot claim that the stability property carries over globally is that there may be more than one fixed point of the nonlinear system. In this case, it is entirely possible that the linearized system predicts that one of the fixed points is a globally asymptotically stable node of the linearized system, but the type and stability of another fixed point in effect forces this node to be only locally asymptotically stable. We will in fact see this in the next example.

Let us now turn to the limitations of Theorem 13.6. Essentially, Theorem 13.6 asserts that near the fixed point $(x_1^*, x_2^*)$ the higher-order terms $R^n(x_1, x_2)$, $n = 1, 2$, of system (18) are very small and do not affect the type and stability of the fixed point as determined by the linearized system *except* in two sensitive cases, scilicet, $\lambda_1$ and $\lambda_2$ pure imaginary (a center for the linearized system), and $\lambda_1$ and $\lambda_2$ real and equal (an improper node or star node for the linearized system). This should not be too surprising given that it is well known that small perturbations in the coefficients of a linear system, and thus in the eigenvalues $\lambda_1$ and $\lambda_2$, can alter the type and stability of a fixed point only in these two borderline cases. In particular, when $\lambda_1$ and $\lambda_2$ are pure imaginary, a small perturbation in the coefficients of a linear system can change the stable center into a globally asymptotically stable or an unstable spiral node, or even leave it as a center. When $\lambda_1$ and $\lambda_2$ are real and equal, a small

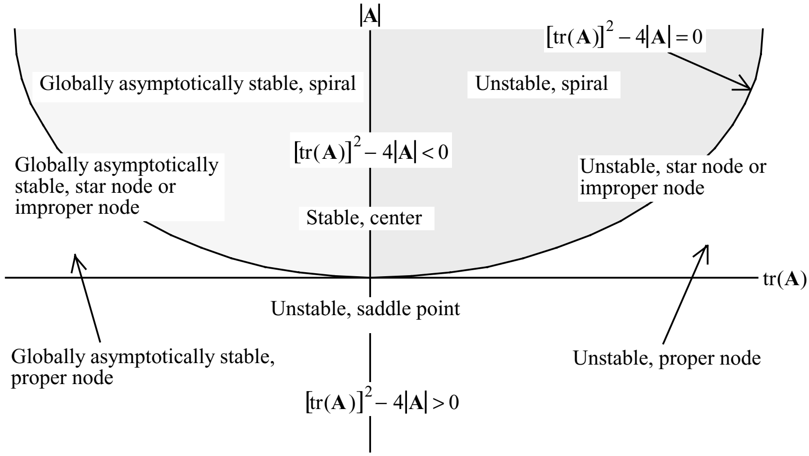


**Figure 13.8**

perturbation in the coefficients of the linear system do not affect the stability of the fixed point, but may change the star node or improper node into a spiral node. Thus, given these observations about linear systems in the two sensitive cases, it is not at all surprising that Theorem 13.6 asserts that the small nonlinear terms $R^n(x_1, x_2)$, $n = 1, 2$, of system (18) exhibit similar effects in the two sensitive cases.

That Theorem 13.6 is not that unexpected may also be seen from Figure 13.8. This figure, which you should be familiar with from a prior introductory course in differential equations, summarizes what is known about linear systems of autonomous ordinary differential equations with constant coefficients in the plane, in terms of the trace [tr(**A**)] and determinant [|**A**|] of the $2 \times 2$ coefficient matrix **A** of the linear system. Figure 13.8 shows that improper nodes, star nodes, and centers all “live” on curves in (as opposed to regions of) the trace-determinant plane, and as such, any slight perturbation of them will in general move them off the curve on which they once lived and thus change their type. The main significance of Theorem 13.6, however, is that in *all other cases*, the small nonlinear terms $R^n(x_1, x_2)$, $n = 1, 2$, do not alter the type of the fixed point or its local stability.

Let’s turn to an example to see how one would use Theorem 13.6 to classify the fixed points and draw the phase portrait of a nonlinear and autonomous system.

**Example 13.5:** Consider the following nonlinear system:

$$\dot{x}_1 = -x_1 + x_1^3,$$
$$\dot{x}_2 = -2x_2.$$

Our first step is to determine the fixed points. Recall that fixed points are the solution to the above system when $\dot{x}_1 = 0$ and $\dot{x}_2 = 0$, that is, they are the solution to the algebraic equations

$$-x_1 + x_1^3 = 0,$$
$$-2x_2 = 0.$$

It should not be too hard to see that we have three fixed points in this nonlinear system, namely, $(0, 0)$, $(1, 0)$, and $(-1, 0)$. Notice that unlike simple linear systems, simple nonlinear systems may have multiple isolated fixed points. At a general point $(x_1, x_2)$, the Jacobian matrix of the linearized system is given by

$$\mathbf{J}(x_1, x_2) \stackrel{\text{def}}{=} \begin{bmatrix} \dfrac{\partial \dot{x}_1}{\partial x_1} & \dfrac{\partial \dot{x}_1}{\partial x_2} \\ \dfrac{\partial \dot{x}_2}{\partial x_1} & \dfrac{\partial \dot{x}_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -1 + 3x_1^2 & 0 \\ 0 & -2 \end{bmatrix}.$$

Now recall that the eigenvalues of a diagonal matrix are the diagonal elements themselves. With this in mind, we therefore find that

$$\mathbf{J}(0, 0) = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}, (\lambda_1, \lambda_2) = (-1, -2),$$

$$\mathbf{J}(1, 0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, (\lambda_1, \lambda_2) = (2, -2),$$

$$\mathbf{J}(-1, 0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, (\lambda_1, \lambda_2) = (2, -2).$$

Thus the linearized system shows that the fixed point $(0, 0)$ is a globally asymptotically stable node, and that the fixed points $(1, 0)$ and $(-1, 0)$ are unstable saddle points.

Inasmuch as nodes and saddle points are not borderlines cases, that is to say, they are covered by Theorem 13.6, we are certain that the type of the three fixed points is predicted correctly for the nonlinear system, and that the stability property holds in a neighborhood of the fixed point. We will see shortly that the fixed point $(0, 0)$ is actually locally asymptotically stable for the nonlinear system. This shows that fixed points that are predicted as globally asymptotically stable by the linearized system may be only locally asymptotically stable for the nonlinear system when more than one isolated fixed point exists, just as Theorem 13.6 allows.

The above conclusions based on the linearized system and Theorem 13.6 can be readily verified for the above nonlinear system because the two differential equations are *uncoupled*, that is, the differential equation for $x_1$ is only a function of $x_1$ and the differential equation for $x_2$ is only a function of $x_2$. Hence the nonlinear system is essentially two independent equations at right angles to each other. This permits us

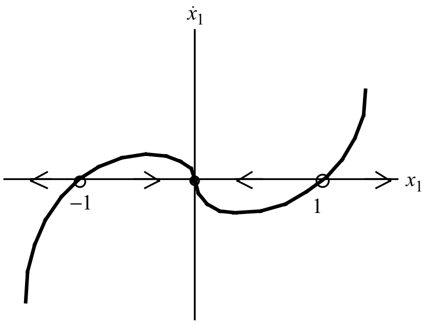


**Figure 13.9**

to draw the phase portrait for the nonlinear system quite easily. Let us now proceed to do just that.

First, observe that because $\dot{x}_2 = -2x_2$, $\dot{x}_2 < 0$ when $x_2 > 0$ and $\dot{x}_2 > 0$ when $x_2 < 0$. This implies that in the $x_2$-direction, all trajectories decay exponentially to $x_2 = 0$. Next, we plot the one-dimensional phase portrait for $\dot{x}_1 = -x_1 + x_1^3$, as in Figure 13.9. From this phase portrait, we see that the fixed points $x_1 = \pm 1$ are unstable whereas the origin is locally asymptotically stable. It is important to observe that the origin is *not* globally asymptotically stable, because if the initial condition for $x_1$ places it at a value greater than 1 or a value less than $-1$, then the trajectory will not converge to the origin, as is evident from Figure 13.9.

Next, observe that the vertical lines $x_1 = 0$ and $x_1 = \pm 1$ are *invariant* because $\dot{x}_1 = 0$ on them. This implies that any trajectory that starts on these lines can never leave them. Similarly, the line $x_2 = 0$ is an invariant horizontal line seeing as $\dot{x}_2 = 0$ when $x_2 = 0$, thus implying that any trajectory that starts on this line stays on it forever.

Finally, note that the phase portrait of the nonlinear system must be symmetric with respect to the $x_1$-axis and $x_2$-axis because the nonlinear system is invariant under the linear transformations $x_1 \to -x_1$ and $x_2 \to -x_2$. By bringing all of the information we have gathered about the nonlinear system together, we therefore arrive at its phase portrait, which we have depicted in Figure 13.10.

The phase portrait confirms that the fixed point (0, 0) is a locally asymptotically stable node, whereas the fixed points $(1, 0)$ and $(-1, 0)$ are unstable saddle points, just as we anticipated based on the linearized system.

The next example, a classical one, shows that small nonlinear terms can change a center into a spiral node. A mental exercise asks you to consider another case in which the nonlinear terms can change a center into a spiral node.

**Example 13.6:** Consider the nonlinear system

$$\dot{x}_1 = -x_2 + ax_1\left(x_1^2 + x_2^2\right),$$
$$\dot{x}_2 = x_1 + ax_2\left(x_1^2 + x_2^2\right)$$

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

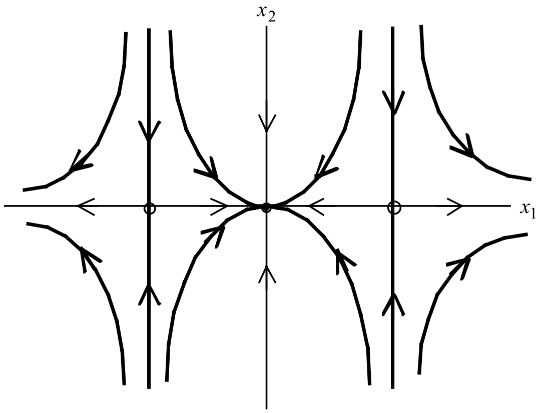


**Figure 13.10**

where $a$ is a parameter. The point of this example is to show that the linearized system *incorrectly* predicts that the origin, the fixed point of the system, is a center for all values of the parameter $a$, whereas the origin is in fact a globally asymptotically stable spiral node for $a < 0$ and an unstable spiral node for $a > 0$. We begin by analyzing the linearized system.

The Jacobian matrix of the linearized system evaluated at the origin is given by

$$\mathbf{J}(0,0) \overset{\text{def}}{=} \left.\begin{bmatrix} \dfrac{\partial \dot{x}_1}{\partial x_1} & \dfrac{\partial \dot{x}_1}{\partial x_2} \\ \dfrac{\partial \dot{x}_2}{\partial x_1} & \dfrac{\partial \dot{x}_2}{\partial x_2} \end{bmatrix}\right|_{(x_1,x_2)=(0,0)} = \left.\begin{bmatrix} 3ax_1^2 + ax_2^2 & -1 + 2ax_1x_2 \\ 1 + 2ax_1x_2 & ax_1^2 + 3ax_2^2 \end{bmatrix}\right|_{(x_1,x_2)=(0,0)}$$

$$= \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Because $\operatorname{tr}(\mathbf{J}(0,0)) = 0$ and $|\mathbf{J}(0,0)| = 1 > 0$, the origin is always a center, at least according to linearization. Equivalently, the eigenvalues of $\mathbf{J}(0,0)$ are readily found to be $\lambda = \pm i$, thus confirming that the origin is a center for the linearized system.

To analyze the original nonlinear system, we make a change of variables to *polar coordinates*. That is, let $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$, where $r$ is the radius and $\theta$ is the angle. Using the trigonometric identity $\cos^2\theta + \sin^2\theta \equiv 1$, we have the corresponding identity $x_1^2 + x_2^2 \equiv r^2$. Differentiating this latter identity with respect to the independent variable $t$, we thus arrive at the differential equation $2x_1\dot{x}_1 + 2x_2\dot{x}_2 = 2r\dot{r}$. Substituting in for $\dot{x}_1$ and $\dot{x}_2$ yields

$$\begin{aligned} r\dot{r} &= x_1\left[-x_2 + ax_1\left(x_1^2 + x_2^2\right)\right] + x_2\left[x_1 + ax_2\left(x_1^2 + x_2^2\right)\right] \\ &= ax_1^2\left(x_1^2 + x_2^2\right) + ax_2^2\left(x_1^2 + x_2^2\right) = a\left(x_1^2 + x_2^2\right)^2 \\ &= ar^4. \end{aligned}$$

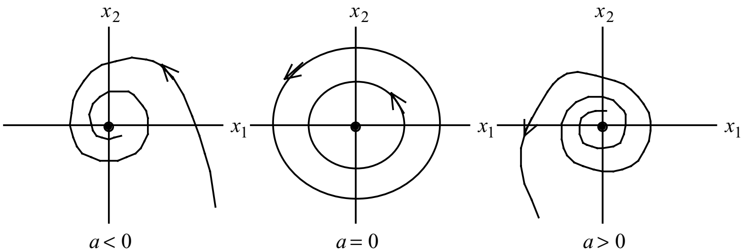


**Figure 13.11**

Therefore the differential equation for $r$ is $\dot{r} = ar^3$. We now turn to the derivation of the differential equation for $\theta$.

Given that $\tan\theta \stackrel{\text{def}}{=} \sin\theta/\cos\theta$, it follows from $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$ that $\tan\theta = x_2/x_1$. Differentiating this latter equation with respect to the independent variable $t$ and using the identity $x_1^2 + x_2^2 \equiv r^2$ gives

$$\frac{1}{\cos^2\theta}\dot{\theta} = \frac{x_1\dot{x}_2 - x_2\dot{x}_1}{x_1^2} = \frac{x_1^2 + ax_1x_2r^2 + x_2^2 - ax_1x_2r^2}{r^2\cos^2\theta}$$
$$= \frac{x_1^2 + x_2^2}{r^2\cos^2\theta} = \frac{r^2}{r^2\cos^2\theta} = \frac{1}{\cos^2\theta}.$$

Thus the differential equation for $\theta$ is simply $\dot{\theta} = 1$.

Summing up our results so far, we have shown that the original nonlinear differential equation system can be replaced by the following decoupled pair of differential equations in polar coordinates:

$$\dot{r} = ar^3, \ \dot{\theta} = 1. \tag{21}$$

This form of the nonlinear system is easy to analyze because the radial and angular motions are independent. As is plainly evident from inspection of these two differential equations, all the trajectories rotate about the origin with constant angular velocity $\dot{\theta} = 1$. The radial motion obviously depends on the sign of the parameter $a$. Figure 13.11 presents the phase diagram of the above system for the three qualitative values of the parameter $a$.

If $a < 0$, then Eq. (21) shows that $\dot{r} < 0$ for all $r > 0$, which implies that $r(t) \to 0$ monotonically as $t \to +\infty$, that is, the trajectories spiral inward toward the origin in view of the fact that the radius is shrinking. In this case, the origin is a globally asymptotically stable spiral node. This conclusion can also be verified by drawing the one-dimensional phase portrait for $\dot{r} = ar^3$ when $a < 0$. If $a = 0$, then $r(t) = r_0$ for all $t$ and the origin is a stable center as the radius is constant on any given trajectory. Finally, if $a > 0$, then Eq. (21) shows that $\dot{r} > 0$ for all $r > 0$, thereby implying that $r(t) \to +\infty$ monotonically as $t \to +\infty$, that is, the trajectories spiral outward from the origin because the radius is growing without bound. Hence the origin is

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

an unstable spiral node in this instance. As remarked above, this conclusion can be verified by drawing the one-dimensional phase portrait for $\dot{r} = ar^3$ when $a > 0$.

This example also clearly demonstrates why centers are so delicate: all trajectories must close *perfectly* after one cycle for a center, since even the slightest miss converts the center into a spiral node.

Just as centers can be altered by the small nonlinearities (recall Example 13.6 just discussed), so too can star nodes and improper nodes. There is, however, one significant difference between centers on the one hand and star nodes and improper nodes on the other, videlicet, the stability of star nodes and improper nodes isn't affected by small nonlinearities. A mental exercise asks you to consider an example in which the nonlinear terms change a locally asymptotically stable star node into a locally asymptotically stable spiral node, but not into an unstable spiral node. That the *stability* of star nodes and improper nodes cannot be changed by small nonlinearities is not too surprising given Figure 13.8. Notice that star nodes and improper nodes live squarely in the asymptotically stable or unstable regions. Hence any small perturbation of them leaves them firmly in the same region as far as stability is concerned. The same is not true for centers because they live on the boundary between the asymptotically stable and unstable regions. Thus, in general, a small perturbation of a center will affect both its type and stability. The observations of this paragraph, and more, are the content of the famous Hartman-Grobman theorem, also known as the linearization theorem, which we will shortly state as Theorem 13.7. Before doing so, however, we require one more definition.

**Definition 13.9:** A simple fixed point $\mathbf{x}^*$ of a nonlinear system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, $\mathbf{x} \in D \subseteq \Re^n$, is called *hyperbolic* if every eigenvalue of $\mathbf{J}(\mathbf{x}^*)$ has a nonzero real part, that is, if $\text{Re}(\lambda_n) \neq 0$, $n = 1, 2, \ldots, N$.

In other words, a simple fixed point is hyperbolic if none of the eigenvalues of its Jacobian matrix evaluated at the fixed point is pure imaginary. Hyperbolic fixed points are also called *generic*, because their occurrence is the rule rather than the exception. This may be seen from Figure 13.8, for it shows that pure imaginary eigenvalues occur very infrequently, since a dart thrown at the figure would almost never land exactly on the positive part of the determinant axis. In the case of the scalar nonlinear differential equation $\dot{x} = f(x)$, a hyperbolic fixed point $x^*$ is one in which $f'(x^*) \neq 0$. Note, in passing, that the origin in Example 13.6 is not a hyperbolic fixed point because the eigenvalues of the Jacobian have zero real parts. With this definition in hand, we may state the Hartman-Grobman or linearization theorem, the proof of which we omit.

**Theorem 13.7:** *Let* $(x_1^*, x_2^*) \in D$ *be an isolated and hyperbolic fixed point of the nonlinear system* $\dot{x}_i = f^i(x_1, x_2)$, $i = 1, 2$, *where* $f^i(\cdot) \in C^{(1)}$ *for all* $(x_1, x_2) \in D$, $i = 1, 2$. *Then in a small enough neighborhood of* $(x_1^*, x_2^*)$, *the nonlinear system and its corresponding linearized system have qualitatively equivalent phase*

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

*portraits. Moreover, any trajectory of the nonlinear system that approaches* $(x_1^*, x_2^*)$ *as* $t \to \pm\infty$ *is tangent to a trajectory of its corresponding linearized system that approaches* $(x_1^*, x_2^*)$ *as* $t \to \pm\infty$.

Let's now pause to discuss and clarify certain aspects of this important theorem. To begin, first note that you probably have already seen this result in the case of a scalar nonlinear differential equation. In that case, you may recall that the stability of an isolated fixed point $x^*$ of the scalar nonlinear differential equation $\dot{x} = f(x)$ was correctly predicted by its linearized equation as long as $f'(x^*) \neq 0$. The condition $f'(x^*) \neq 0$ is the exact analogue of $\text{Re}(\lambda_n) \neq 0$, $n = 1, 2, \ldots, N$, in the case of a system of nonlinear differential equations, that is, hyperbolicity of a fixed point, as we noted above. Theorem 13.7 therefore simply extends this result to the case of a system of nonlinear differential equations. It asserts that the stability of an isolated and hyperbolic fixed point of a system of nonlinear differential equations is determined by its corresponding linearized system, that is to say, the eigenvalues of the Jacobian of $\mathbf{f}(\mathbf{x})$ evaluated at the fixed point, videlicet, $\mathbf{J}(\mathbf{x}^*)$. In other words, if our interest centers on stability rather than on the detailed geometry of the trajectories, which is almost always the case in dynamic economic theory, then the marginal cases are only those in which at least one eigenvalue of $\mathbf{J}(\mathbf{x}^*)$ has a zero real part.

Second, let's turn to the meaning of the phrase "qualitatively equivalent phase portraits." A more precise and technically sophisticated way to state the relevant portion of the Hartman-Grobman theorem is this: the local phase portrait of a nonlinear system near a hyperbolic fixed point is topologically equivalent to the phase portrait of its linearized system. Here, *topologically equivalent* means that there is a continuous deformation with a continuous inverse (that is, a homeomorphism) that maps one local phase portrait onto the other, such that trajectories map onto trajectories and the sense of time (i.e., the direction of the arrows) is preserved. In other words, two phase portraits are topologically equivalent if one is a distorted version of the other. Distortions such as bending and warping are permitted, but distortions such as ripping are not because ripping involves discontinuities. Said in simple terms, therefore, Theorem 13.7 asserts that the trajectories of a nonlinear system and its linearization look similar in a neighborhood of a hyperbolic fixed point.

Let's now examine a simple example to drive home the essence of the linearization theorem. We will finish up the chapter with an example of how Theorems 13.6 and 13.7 are typically used in optimal control problems.

**Example 13.7:** Consider the nonlinear and autonomous system

$$\dot{x}_1 = e^{x_1+x_2} - x_2,$$
$$\dot{x}_2 = -x_1 + x_1 x_2.$$

Our goal here is to determine all the fixed points of this system, classify their type and stability, and determine the local phase diagram at each of them.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

The fixed points of this system are the solutions to the following simultaneous nonlinear equations:

$$e^{x_1+x_2} - x_2 = 0,$$
$$x_1(x_2 - 1) = 0.$$

The second of these equations is satisfied only by $x_1 = 0$ or $x_2 = 1$. If $x_1 = 0$, then the first equation reduces to $e^{x_2} = x_2$, which has no real solutions because $e^{x_2} > x_2$ for all $x_2 \in \Re$. We may thus conclude that there is no fixed point with $x_1 = 0$. If $x_2 = 1$, then the first equation reduces to $e^{x_1+1} = 1$, which has but one real solution, scilicet, $x_1 = -1$. Thus $(x_1, x_2) = (-1, 1)$ is the only fixed point, and is thus isolated.

Next we compute the Jacobian matrix of the nonlinear system and evaluate it at the fixed point. Doing just that yields

$$\mathbf{J}(-1,1) \overset{\text{def}}{=} \begin{bmatrix} \dfrac{\partial \dot{x}_1}{\partial x_1} & \dfrac{\partial \dot{x}_1}{\partial x_2} \\ \dfrac{\partial \dot{x}_2}{\partial x_1} & \dfrac{\partial \dot{x}_2}{\partial x_2} \end{bmatrix}\Bigg|_{(x_1,x_2)=(-1,1)} = \begin{bmatrix} e^{x_1+x_2} & e^{x_1+x_2} - 1 \\ x_2 - 1 & x_1 \end{bmatrix}\Bigg|_{(x_1,x_2)=(-1,1)}$$
$$= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Given that $\mathbf{J}(-1, 1)$ is a diagonal matrix, the eigenvalues are simply the elements on the main diagonal; hence $\lambda_1 = 1$ and $\lambda_2 = -1$. Thus the linearization predicts the fixed point is an unstable saddle point, and by Theorem 13.7 (or Theorem 13.6 for that matter), the original nonlinear system is similarly an unstable saddle point in a neighborhood of its only fixed point $(x_1, x_2) = (-1, 1)$. It is simple to verify that the unstable manifold is the line spanned by the eigenvector $\mathbf{v}^1 = (1, 0)$ corresponding to the eigenvalue $\lambda_1 = 1$, whereas the stable manifold is the line spanned by the eigenvector $\mathbf{v}^2 = (0, 1)$ corresponding to the eigenvalue $\lambda_2 = -1$.

With the above information in hand, the local phase portrait of the nonlinear system can be constructed. It is given in Figure 13.12.

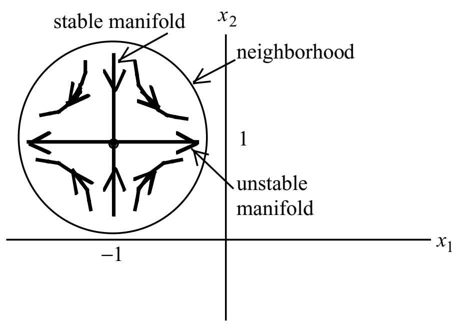


**Figure 13.12**

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

It is important to remember that the linearization theorem permits us only to infer the phase portrait of the original nonlinear system in a neighborhood of its unique fixed point $(x_1, x_2) = (-1, 1)$.

Before embarking on the final example of the chapter, we require one more definition.

**Definition 13.10:** Given a pair of autonomous nonlinear differential equations $\dot{x}_1 = f^1(x_1, x_2)$ and $\dot{x}_2 = f^2(x_1, x_2)$, the *nullclines* are curves such that $f^1(x_1, x_2) = 0$ or $f^2(x_1, x_2) = 0$.

Nullclines are implicit equations in general. The $x_1$ nullcline, or as it is often referred to, the $f^1(x_1, x_2) = 0$ or $\dot{x}_1 = 0$ *isocline*, is a curve in the $x_1x_2$-phase plane such that the flow is purely vertical, since $\dot{x}_1 = f^1(x_1, x_2) = 0$ along this nullcline, that is, $x_1$ is not changing with the passage of time. Similarly, the $x_2$ nullcline or $\dot{x}_2 = 0$ isocline is a curve in the $x_1x_2$-phase plane where the flow is purely horizontal, since $\dot{x}_2 = f^2(x_1, x_2) = 0$ along this nullcline, that is, $x_2$ is not changing with the passage of time. The exception to these two observations is where the two nullclines intersect, which is the fixed point of the system. The nullclines of the system partition the phase plane into regions, sometimes called *isosectors*, where $\dot{x}_1$ and $\dot{x}_2$ have various signs. In each isosector, the trajectories of the system are monotonic, in the sense that $\dot{x}_1$ and $\dot{x}_2$ have the same sign throughout the isosector. Nullclines are therefore especially useful for drawing the phase portraits of autonomous and nonlinear systems when no explicit functional form for the vector field is specified. We will see just how important they are in the next example. Let us note, in passing, that in some instances, nullclines do not exist, as is the case when $f^1(x_1, x_2) \stackrel{\text{def}}{=} x_1^2 + x_2^2 + 1$, for example.

Now that the fundamental properties of systems of autonomous and nonlinear differential equations have been laid out, as well as the basic ideas behind the construction of a phase diagram, we turn to an examination of how phase diagrams may be used in optimal control problems to provide a qualitative characterization of the solution to the necessary conditions. As will be made clear in the ensuing example, the phase portrait approach to studying the qualitative properties of optimal control problems is especially useful when the functional forms of the integrand and transition functions are not specified enough so as to yield an explicit solution of the necessary conditions, the typical case in economic applications of optimal control theory. For example, all one may know about the integrand function are some of its qualitative properties, such as monotonicity, curvature, and homogeneity.

**Example 13.8:** Consider the fixed endpoints and finite horizon optimal control problem

$$\max_{u(\cdot)} J[x(\cdot), u(\cdot)] \stackrel{\text{def}}{=} \int_0^T f(x(t), u(t))\, e^{-rt}\, dt$$
$$\text{s.t.} \quad \dot{x}(t) = g(x(t), u(t)),\ \ x(0) = x_0,\ \ x(T) = x_T, \tag{22}$$

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

where $r > 0$ is the discount rate. We assume that $f(\cdot) \in C^{(2)}$ and $g(\cdot) \in C^{(1)}$ on their domains, and furthermore that

$$f_{xx}(x, u) < 0,\ \ f_{uu}(x, u) < 0,\ \ f_{ux}(x, u) \equiv 0,$$

$$g_{xx}(x, u) < 0,\ \ g_{uu}(x, u) < 0,\ \ g_{ux}(x, u) \equiv 0.$$

We will not specify any further properties of $f(\cdot)$ and $g(\cdot)$, however, but you may be surprised at how much qualitative information will be forthcoming about the solution of the necessary conditions from these assumptions alone. It is worthwhile to note that because of the assumption of additive separability of $f(\cdot)$ and $g(\cdot)$ in $x$ and $u$, we may simplify the notation a bit by writing the partial derivatives of $f(\cdot)$ and $g(\cdot)$ with respect to $u$ as $f_u(u)$ and $g_u(u)$, respectively, and similarly for the partial derivatives with respect to $x$.

Define the current value Hamiltonian as $H(x, u, \lambda) \stackrel{\text{def}}{=} f(x, u) + \lambda g(x, u)$. By Theorem 12.1, the necessary conditions are given by

$$H_u(x, u, \lambda) = f_u(u) + \lambda g_u(u) = 0, \tag{23}$$

$$\dot{\lambda} = r\lambda - H_x(x, u, \lambda) = [r - g_x(x)]\lambda - f_x(x), \tag{24}$$

$$\dot{x} = H_\lambda(x, u, \lambda) = g(x, u),\ \ x(0) = x_0,\ \ x(T) = x_T, \tag{25}$$

and of course $H_{uu}(x, u, \lambda) = f_{uu}(u) + \lambda g_{uu}(u) \leq 0$. Observe that if the current value costate variable $\lambda(t)$ corresponding to the solution of the necessary conditions is nonnegative for the entire planning horizon, then the Hamiltonian is a strictly concave function of $(x, u)\, \forall\, t \in [0, T]$. Therefore, if a solution to the necessary conditions exists and $\lambda(t) \geq 0\, \forall\, t \in [0, T]$, the solution of the necessary conditions will be the unique maximizing solution to the posed optimal control problem by Theorem 4.3. Henceforth, we will accordingly assume that $\lambda(t) > 0\, \forall\, t \in [0, T]$. This, in conjunction with Eq. (23), implies that $f_u(u)g_u(u) < 0$, or equivalently, that $f_u(u)$ and $g_u(u)$ have opposite signs along the optimal path, that is, $\operatorname{sign}[f_u(u)] = -\operatorname{sign}[g_u(u)]$ along the optimal path. Finally, in agreement with the above assumption that the state variable is a good, that is to say, $\lambda(t) > 0\, \forall\, t \in [0, T]$, we also assume that $x(t) > 0\, \forall\, t \in [0, T]$ in an optimal plan. This assumption also has the added benefit of yielding a phase portrait that is more reminiscent of that typically encountered in intertemporal economic models, namely, the variables under consideration are nonnegative.

To begin the construction of the phase portrait corresponding to the necessary conditions of the optimal control problem, we must first reduce the necessary conditions down to two ordinary differential equations in order to draw it. We may proceed to do so in one of two ways. The approach we shall take is to reduce Eqs. (23) through (25) down to a pair of autonomous differential equations for $(x, \lambda)$. The alternative approach is to reduce Eqs. (23) through (25) down to a pair of autonomous differential equations for $(x, u)$. This is left for a mental exercise. Both approaches yield the same information about the solution, for they both begin by using the information contained in the necessary conditions (23) through (25). For optimal control problems with one state variable and one control variable, therefore, the choice of approach typically comes down to which pair of variables, namely, $(x, \lambda)$ or $(x, u)$, lend themselves to a more meaningful economic interpretation.

Seeing as $H_{uu}(x, u, \lambda) = f_{uu}(u) + \lambda g_{uu}(u) < 0$ under our assumptions, the implicit function theorem may be applied to Eq. (23) to solve it, in principle, for the control variable in terms of the current value costate variables, say, $u = \hat{u}(\lambda)$. The comparative statics properties of this solution are given by differentiating the identity $f_u(\hat{u}(\lambda)) + \lambda g_u(\hat{u}(\lambda)) \equiv 0$ with respect to $\lambda$, or equivalently, by applying the implicit function theorem to Eq. (23), to get

$$\hat{u}'(\lambda) \equiv \frac{-g_u(\hat{u}(\lambda))}{f_{uu}(\hat{u}(\lambda)) + \lambda g_{uu}(\hat{u}(\lambda))} \gtrless 0. \tag{26}$$

Because we have not made an assumption about the sign of $g_u(u)$, we cannot sign Eq. (26). Note, however, that because $f_{uu}(u) + \lambda g_{uu}(u) < 0$, we know that $\operatorname{sign}[\hat{u}'(\lambda)] = \operatorname{sign}[g_u(\hat{u}(\lambda))]$.

Substituting $u = \hat{u}(\lambda)$ into the canonical equations (24) and (25) yields the autonomous differential equations of interest:

$$\dot{\lambda} = [r - g_x(x)]\lambda - f_x(x), \tag{27}$$

$$\dot{x} = g(x, \hat{u}(\lambda)). \tag{28}$$

Because of the lack of functional form assumptions on $f(\cdot)$ and $g(\cdot)$, Eqs. (27) and (28) can't be solved for an explicit solution.

The last piece of information we require is the Jacobian matrix of Eqs. (27) and (28) evaluated at the fixed point. In accord with our prior observation concerning uniqueness of the solution of the control problem, we assume that there exists a unique fixed point of Eqs. (27) and (28), say, $(x^*, \lambda^*)$. As you will recall, $(x^*, \lambda^*)$ are found by setting $\dot{\lambda} = 0$ and $\dot{x} = 0$ in Eqs. (27) and (28), respectively, and solving the resulting simultaneous nonlinear algebraic equations. We now assume that $g_x(x^*) < 0$, so as to place just a bit more qualitative structure on the control problem. The Jacobian matrix of Eqs. (27) and (28) is therefore given by

$$\begin{aligned} \mathbf{J}(x^*, \lambda^*) &= \left.\begin{bmatrix} \dfrac{\partial \dot{\lambda}}{\partial \lambda} & \dfrac{\partial \dot{\lambda}}{\partial x} \\ \dfrac{\partial \dot{x}}{\partial \lambda} & \dfrac{\partial \dot{x}}{\partial x} \end{bmatrix}\right|_{(x,\lambda)=(x^*,\lambda^*)} \\ &= \begin{bmatrix} \underset{(+)}{r - g_x(x^*)} & \underset{(+)}{-\lambda^* g_{xx}(x^*) - f_{xx}(x^*)} \\ \underset{(+)}{g_u(\hat{u}(\lambda^*))\,\hat{u}'(\lambda^*)} & \underset{(-)}{g_x(x^*)} \end{bmatrix}. \end{aligned} \tag{29}$$

Note that we have indicated the sign of each element of $\mathbf{J}(x^*, \lambda^*)$ in Eq. (29). Inspection of Eq. (29) reveals that $\operatorname{tr}[\mathbf{J}(x^*, \lambda^*)] = r > 0$. Because the sum of the

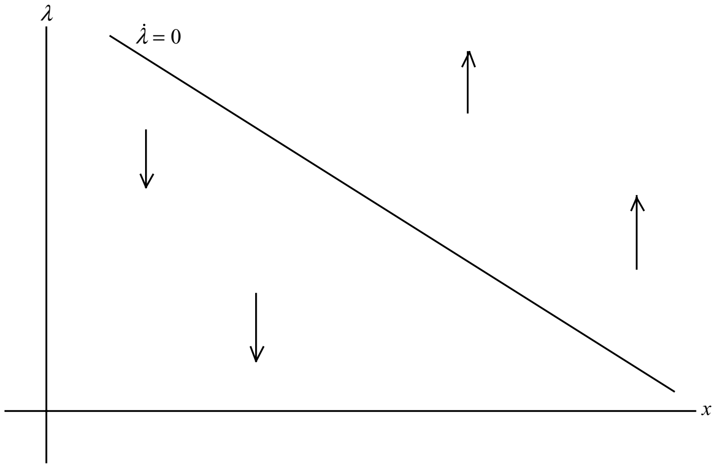


**Figure 13.13**

eigenvalues of $\mathbf{J}(x^*, \lambda^*)$ equals its trace, as you will prove in a mental exercise, this implies that at least one eigenvalue of $\mathbf{J}(x^*, \lambda^*)$ is positive. This, in turn, implies that the fixed point is not locally asymptotically stable. This conclusion is in fact true in general for discounted optimal control problems when the integrand and transition functions do not depend explicitly on the independent variable $t$, as we shall see in Chapter 18. Continuing on, inspection of the signs of the elements of $\mathbf{J}(x^*, \lambda^*)$ shows that $|\mathbf{J}(x^*, \lambda^*)| < 0$. Because the product of the eigenvalues of $\mathbf{J}(x^*, \lambda^*)$ equals its determinant, as you will also prove in a mental exercise, this implies that one eigenvalue of $\mathbf{J}(x^*, \lambda^*)$ is positive and one is negative. This structure of the eigenvalues of $\mathbf{J}(x^*, \lambda^*)$ is the defining property of a fixed point that is a *saddle point*. We may therefore conclude by Theorem 13.6 or Theorem 13.7 that the steady state $(x^*, \lambda^*)$ is an unstable saddle point of the original nonlinear system of differential equations given by Eqs. (27) and (28). This is the final piece of information we require in order to construct the phase portrait corresponding to Eqs. (27) and (28).

We now describe, in detail, the construction of the phase portrait corresponding to Eqs. (27) and (28). We will do this in five distinct steps. What follows will be used repeatedly in several of the ensuing chapters.

**Step 1:** Let's determine the slope of the $\lambda$ nullcline, or equivalently, the $\dot{\lambda} = 0$ isocline, first. To begin, draw a graph with $\lambda$ plotted vertically and $x$ plotted horizontally, as in Figure 13.13. In order to determine the $\lambda$ nullcline, set $\dot{\lambda} = 0$ in Eq. (27) to get the algebraic equation $[r - g_x(x)]\lambda - f_x(x) = 0$. Given that $\lambda$ appears linearly in this equation, we may solve it explicitly for $\lambda$ to get the alternative form of the $\lambda$ nullcline, to wit,

$$\lambda = \frac{f_x(x)}{[r - g_x(x)]}. \tag{30}$$

Because Theorems 13.6 and 13.7 permit us to reach conclusions regarding the steady state only locally, it suffices to consider the slope of the $\lambda$ nullcline evaluated at the steady state. Hence, differentiating Eq. (30) with respect to $x$ and evaluating at the steady state gives

$$\left.\frac{\partial\lambda}{\partial x}\right|_{(x,\lambda)=(x^*,\lambda^*)} = \frac{[r - g_x(x^*)]f_{xx}(x^*) + f_x(x^*)g_{xx}(x^*)}{[r - g_x(x^*)]^2} < 0, \tag{31}$$

where we have used the assumptions $f_{xx}(x^*) < 0$, $g_{xx}(x^*) < 0$, and $g_x(x^*) < 0$, the latter of which, when used in conjunction with the assumption $\lambda^* > 0$ and Eq. (30), implies that $f_x(x^*) > 0$. Equation (31) demonstrates that the slope of the $\dot{\lambda} = 0$ isocline is negative in a neighborhood of the steady state. Moreover, it is important to recognize that the $\lambda$ nullcline divides the $x\lambda$-phase plane into two regions, one in which $\dot{\lambda} > 0$ and thus $\lambda$ is increasing over time, and one in which $\dot{\lambda} < 0$ and thus $\lambda$ is decreasing over time. The next step in the construction of the phase portrait seeks to determine precisely these two regions. Figure 13.13 displays the $\dot{\lambda} = 0$ isocline. Note that at this juncture, the arrows in the figure should be ignored.

The student well versed in implicit function theory should recognize that the slope of the $\dot{\lambda} = 0$ isocline may also be determined from the Jacobian matrix in Eq. (29). In particular, observe that the negative of the ratio of the (1,2) element to the (1,1) element of $\mathbf{J}(x^*, \lambda^*)$ gives

$$-\left.\frac{\partial\dot{\lambda}/\partial x}{\partial\dot{\lambda}/\partial\lambda}\right|_{(x,\lambda)=(x^*,\lambda^*)} = \frac{[r - g_x(x^*)]f_{xx}(x^*) + f_x(x^*)g_{xx}(x^*)}{[r - g_x(x^*)]^2} < 0,$$

which is identical to Eq. (31). This is as it should be by the implicit function theorem. Note that in arriving at the above result, we made use of Eq. (30).

**Step 2:** To determine the vector field associated with the $\dot{\lambda} = 0$ isocline in a neighborhood of the steady state, that is to say, the movement of, or forces acting on, a point *not* located on the $\lambda$ nullcline but near the fixed point, we proceed as follows. To begin, let $(x, \lambda) = (x^*, \lambda^*)$, so that the point under consideration, namely, the fixed point, lies on the $\dot{\lambda} = 0$ isocline, thereby implying that $[r - g_x(x^*)]\lambda^* - f_x(x^*) = 0$. Then consider what happens to $\dot{\lambda}$ if we change $x$ or $\lambda$ by a small amount. This change results in a new point, say, $(x, \lambda) = (x^* + k_1, \lambda^*)$, $k_1 \neq 0$ and small, or $(x, \lambda) = (x^*, \lambda^* + k_2)$, $k_2 \neq 0$ and small. Neither of these new coordinates lies on the $\dot{\lambda} = 0$ isocline but they are still in a neighborhood of the steady state, that is, $[r - g_x(x^* + k_1)]\lambda^* - f_x(x^* + k_1) \neq 0$ and $[r - g_x(x^*)][\lambda^* + k_2] - f_x(x^*) \neq 0$. We can determine the effect of this perturbation on $\dot{\lambda}$ quite easily by recognizing that the effect of a change in $x$ or $\lambda$ on $\dot{\lambda}$ in a neighborhood of the fixed point is found by differentiating the conjugate differential equation (27) with respect to $x$ or $\lambda$ and evaluating the result at the steady state. The beauty of this observation is that we have already made this computation in deriving the Jacobian matrix $\mathbf{J}(x^*, \lambda^*)$ in Eq. (29).

In the case in which we move horizontally off the $\dot{\lambda} = 0$ isocline in a neighborhood of the steady state in the $x\lambda$-phase plane, we see from the (1,2) element of $\mathbf{J}(x^*, \lambda^*)$ that

$$\left.\frac{\partial \dot{\lambda}}{\partial x}\right|_{(x,\lambda)=(x^*,\lambda^*)} = -\lambda^* g_{xx}(x^*) - f_{xx}(x^*) > 0, \tag{32}$$

the sign of which follows from $f_{xx}(x^*) < 0$, $g_{xx}(x^*) < 0$, and $\lambda^* > 0$. Equation (32) shows that a small horizontal movement from the steady state to the right of the $\dot{\lambda} = 0$ isocline causes $\dot{\lambda}$ to increase from zero. That is, $\dot{\lambda} > 0$ as we move to the right of the $\lambda$ nullcline in a neighborhood of the steady state, since $\dot{\lambda} = 0$ at the steady state. Symmetrically, Eq. (32) asserts that $\dot{\lambda} < 0$ as we move to the left of the $\lambda$ nullcline in a neighborhood of the fixed point, since $\dot{\lambda} = 0$ at the steady state. Consequently, in Figure 13.13, we have drawn vertical arrows showing that $\lambda$ is increasing to the right of the $\dot{\lambda} = 0$ isocline and vertical arrows showing that $\lambda$ is decreasing to the left of the $\dot{\lambda} = 0$ isocline. Remember that because we are dealing with the $\lambda$ nullcline, the vector field pertaining to points not on it points vertically because that is the direction in which $\lambda$ is plotted in the phase portrait.

The analysis is essentially the same, and the qualitative conclusions are identical, if we consider the case in which we move vertically off the $\dot{\lambda} = 0$ isocline in a neighborhood of the steady state in the $x\lambda$-phase plane. This perturbation is exactly what is captured by the (1,1) element of $\mathbf{J}(x^*, \lambda^*)$, videlicet,

$$\left.\frac{\partial \dot{\lambda}}{\partial \lambda}\right|_{(x,\lambda)=(x^*,\lambda^*)} = r - g_x(x^*) > 0, \tag{33}$$

since $g_x(x^*) < 0$. Equation (33) demonstrates that a small vertical movement from the fixed point above the $\dot{\lambda} = 0$ isocline causes $\dot{\lambda}$ to increase from zero, exactly the same information conveyed by Eq. (32). This was to be expected, however, because a small positive horizontal or vertical movement from the steady state places the new point on the same side of the $\lambda$ nullcline, as it is downward sloping in a neighborhood of the fixed point. In sum, therefore, $\lambda$ is increasing to the right or above the $\dot{\lambda} = 0$ isocline, whereas $\lambda$ is decreasing to the left or below the $\dot{\lambda} = 0$ isocline, in a neighborhood of the steady state. The latter result follows from the fact that we are free to interpret the derivatives in Eqs. (32) and (33) as a decrease in $x$ or $\lambda$. Finally, note that even though we know that the $\dot{\lambda} = 0$ isocline is negatively sloped in a neighborhood of the steady state, we do not know any curvature properties of it. Nevertheless, we have drawn it as a straight line seeing as our analysis pertains only to the steady state and its surrounding neighborhood.

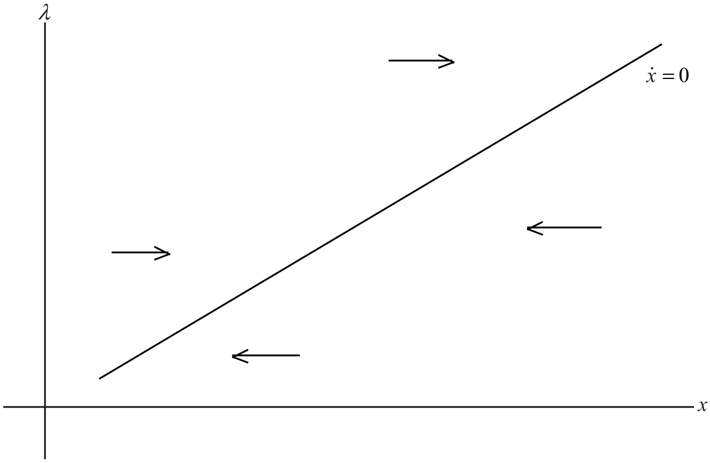


**Figure 13.14**

**Step 3:** Now we seek to determine the slope of the $x$ nullcline. The $\dot{x} = 0$ isocline is found by setting $\dot{x} = 0$ in Eq. (28), thereby yielding the implicit relation

$$g(x, \hat{u}(\lambda)) = 0 \tag{34}$$

between $x$ and $\lambda$. The Jacobian of Eq. (34) with respect to $\lambda$ evaluated at the steady state is given by $g_u(\hat{u}(\lambda^*))\,\hat{u}'(\lambda^*) > 0$, the sign of which follows from an aforementioned result that $\text{sign}[\hat{u}(\lambda)] = \text{sign}[g_u(\hat{u}(\lambda))]$. Hence, by the implicit function theorem, Eq. (34) defines $\lambda$ as a locally $C^{(1)}$ function of $x$ in a neighborhood of the steady state, say, $\lambda = \Lambda(x)$. Consequently, the slope of the $\dot{x} = 0$ isocline in a neighborhood of the steady state is given by differentiating the identity $g(x, \hat{u}(\Lambda(x))) = 0$ with respect to $x$, or equivalently, by applying the implicit function theorem to Eq. (34), to arrive at

$$\left.\frac{\partial\lambda}{\partial x}\right|_{(x,\lambda)=(x^*,\lambda^*)} = \Lambda'(x^*) \equiv \frac{-g_x(x^*)}{g_u(\hat{u}(\Lambda(x^*)))\,\hat{u}'(\Lambda(x^*))} > 0, \tag{35}$$

since $\text{sign}[\hat{u}'(\lambda)] = \text{sign}[g_u(\hat{u}(\lambda))]$ and $g_x(x^*) < 0$. Equation (35) therefore establishes that the slope of the $\dot{x} = 0$ isocline is positive in the $x\lambda$-phase plane in a neighborhood of the steady state. The $x$ nullcline thus divides the $x\lambda$-phase plane into two regions, one in which $\dot{x} > 0$ and thus $x$ is increasing over time, and one in which $\dot{x} < 0$ and thus $x$ is decreasing over time. This division of the $x\lambda$-phase plane into two regions by the $x$ nullcline is analogous to the splitting of the $x\lambda$-phase plane into two regions by the $\lambda$ nullcline, but with one important difference: the $\dot{x} = 0$ isocline splits the $x\lambda$-phase plane into two regions and governs the movement of $x$ over time, whereas the $\dot{\lambda} = 0$ isocline splits the $x\lambda$-phase plane into two regions and governs the movement of $\lambda$ over time. We have depicted the positive slope of the $x$ nullcline in Figure 13.14.

As noted in Step 1, the student well versed in implicit function theory should recognize that the slope of the $\dot{x} = 0$ isocline may also be determined from the Jacobian matrix in Eq. (29). In this instance, the negative of the ratio of the (2,2) element to the (2,1) element of $\mathbf{J}(x^*, \lambda^*)$ gives

$$-\frac{\partial \dot{x}/\partial x}{\partial \dot{x}/\partial \lambda}\bigg|_{(x,\lambda)=(x^*,\lambda^*)} = \frac{-g_x(x^*)}{g_u(\hat{u}(\lambda^*))\hat{u}'(\lambda^*)} > 0,$$

which is identical to Eq. (35) once one recognizes that $\lambda^* = \Lambda(x^*)$.

**Step 4:** To determine the vector field associated with the $\dot{x} = 0$ isocline in a neighborhood of the steady state, we proceed as we did in Step 2. As a result, we will be much more brief in the present derivation. If we move horizontally off the $\dot{x} = 0$ isocline in a neighborhood of the steady state in the $x\lambda$-phase plane, then we see from the (2,2) element of $\mathbf{J}(x^*, \lambda^*)$ that

$$\frac{\partial \dot{x}}{\partial x}\bigg|_{(x,\lambda)=(x^*,\lambda^*)} = g_x(x^*) < 0. \tag{36}$$

Equation (36) shows that a small horizontal movement from the steady state to the right of the $\dot{x} = 0$ isocline causes $\dot{x}$ to decrease from zero. That is, $\dot{x} < 0$ as we move to the right of the $x$ nullcline in a neighborhood of the fixed point, since $\dot{x} = 0$ at the fixed point, whereas $\dot{x} > 0$ as we move to the left of the $x$ nullcline in a neighborhood of the fixed point. As a result, in Figure 13.14, we have drawn horizontal arrows showing that $x$ is decreasing to the right of the $\dot{x} = 0$ isocline and horizontal arrows showing that $x$ is increasing to the left of the $\dot{x} = 0$ isocline. Recall that because we are dealing with the $x$ nullcline, the vector field pertaining to points not on it points horizontally, as that is the direction in which $x$ is plotted in the phase portrait.

If we consider the case in which we move vertically off the $\dot{x} = 0$ isocline in a neighborhood of the steady state in the $x\lambda$-phase plane, then this perturbation is exactly captured by the (2,1) element of $\mathbf{J}(x^*, \lambda^*)$, namely,

$$\frac{\partial \dot{x}}{\partial \lambda}\bigg|_{(x,\lambda)=(x^*,\lambda^*)} = g_u(\hat{u}(\lambda^*))\,\hat{u}'(\lambda^*) > 0. \tag{37}$$

Equation (37) demonstrates that a small vertical movement from the fixed point above the $\dot{x} = 0$ isocline causes $\dot{x}$ to increase from zero, exactly the same information conveyed by Eq. (36). This is not unexpected, however, because a small positive horizontal movement from the steady state places the new point on the opposite side of the $x$ nullcline as does a small positive vertical movement, for the $x$ nullcline is upward sloping in a neighborhood of the fixed point. All told, $x$ is decreasing to the right or below the $\dot{x} = 0$ isocline, whereas $x$ is increasing to the left or above the $\dot{x} = 0$ isocline, in a neighborhood of the steady state. These facts are depicted by the arrows in Figure 13.14.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

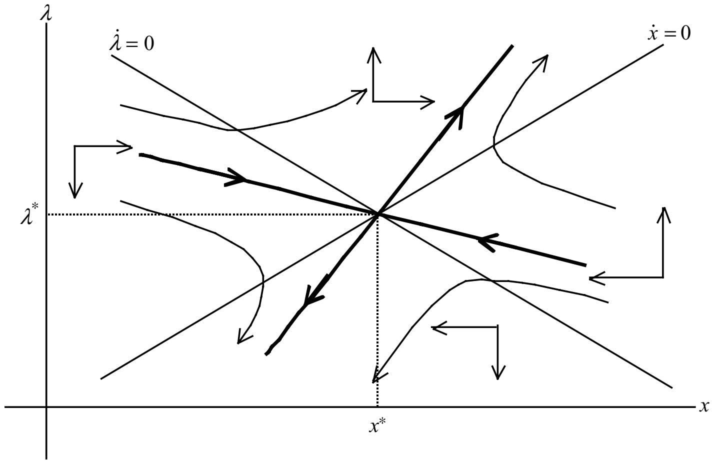


**Figure 13.15**

**Step 5:** By bringing Figures 13.13 and 13.14 together into one graph, we get the complete phase portrait corresponding to the canonical differential equations (27) and (28). This is shown in Figure 13.15. A point on any of the curves in Figure 13.15 indicates a value of $(x, \lambda)$ that might be realized at a moment $t \in [0, T]$, which is consistent with the canonical differential equations. The direction arrows indicate how $(x, \lambda)$ change with the passage of time. Given that we have used the canonical differential equations to construct Figure 13.15, it is simply a plot of all the possible solutions to them. In other words, Figure 13.15 is a graphical representation of the *general solution* to the canonical differential equations.

The heavy trajectories in Figure 13.15 are the stable and unstable manifolds of the saddle point. It turns out that this phase portrait, which depicts a fixed point or steady state that is an unstable saddle point, is typical of that encountered in dynamic economic models in which one state variable is present in the control problem, as we shall see repeatedly in several of the ensuing chapters.

With the construction of the phase diagram complete, our goal now is to incorporate the boundary conditions of the control problem into the phase diagram. To this end, remember that each trajectory in the phase diagram is a unique solution to the canonical differential equations for some given boundary conditions. In other words, for a given set of boundary conditions, there is a unique trajectory among the infinitely many present in the phase portrait that satisfies them. This trajectory thus represents a specific solution of the canonical differential equations. We will consider three cases for the boundary conditions in what follows.

**Case 1:** $T$ finite and given, $x(0) = x_0$ given, and $x(T) = x_T$ given.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

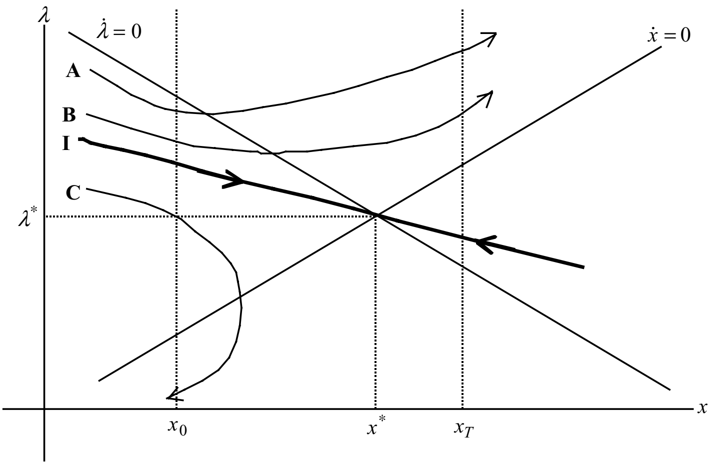


**Figure 13.16**

The optimal solution in this finite-horizon fixed–endpoints case is the trajectory that begins on the vertical line $x(0) = x_0$ and ends on the vertical line $x(T) = x_T$, in an elapsed time of $T$. Note that in this case, the initial stock may be greater than, less than, or equal to the terminal stock. In most economic problems, the initial and terminal states are positive, that is, $x_0 > 0$ and $x_T > 0$. Even so, it is often not known how large $x_0 > 0$ is relative to $x_T > 0$. Figure 13.16 depicts the situation in Case 1 assuming that $x_T > x_0 > 0$.

Given that $x_T > x_0$, trajectories **A** and **B** are both solutions to the canonical differential equations (27) and (28) and the boundary conditions. That is, they are both specific solutions to the canonical differential equations, but each takes a different amount of time to traverse the distance from $x_0 > 0$ to $x_T > 0$. Hence, say for some finite $T = T_B$, trajectory **B** is optimal, whereas for some other finite $T = T_A \neq T_B$, trajectory **A** is optimal. Trajectory **I**, on the other hand, is the stable manifold of the saddle point steady state, and thus takes an infinite amount of time to reach the steady state solution $(x^*, \lambda^*)$ of the canonical differential equations by Theorem 13.4. Hence it cannot be a solution for Case 1 because of this feature. Consequently, trajectory **B** takes a longer time to travel from $x_0$ to $x_T$ than does trajectory **A** in view of the fact that it is "closer" to trajectory **I**. More rigorously, trajectory **B** has a smaller value of $\dot{x}$ for a given value of $x$ than does trajectory **A**, as inspection of Eq. (37) confirms, and therefore a smaller rate of change of the state variable. As a result, trajectory **B** takes longer to go from $x_0$ to $x_T$ than does trajectory **A**. In other words, trajectory **B** is optimal for some finite planning horizon $T = T_B$ that is larger than the finite planning horizon $T = T_A$ that is optimal for trajectory **A**. Trajectory **C**, however, is not a solution to the canonical differential equations and boundary conditions as it violates the terminal boundary condition $x(T) = x_T$, as the vector field does not permit it to reach $x(T) = x_T$. More generally, any trajectory that begins

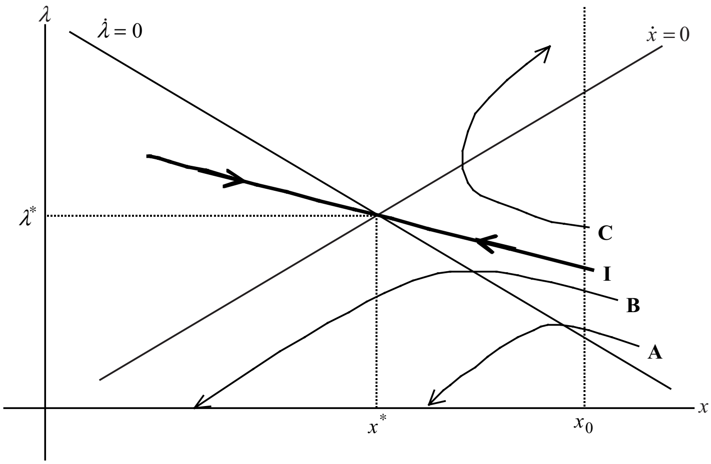


**Figure 13.17**

along the line $x(0) = x_0$ and lies below trajectory **I** is not a solution to the optimal control problem in Case 1 because it fails to satisfy the terminal boundary condition $x(T) = x_T$.

**Case 2:** $T$ fixed and finite, $x(0) = x_0$ given, and $x(T) = x_T$ a decision variable.

Because $x(T) = x_T$ is a decision variable in this instance, we know from Theorem 12.1 that the necessary transversality condition is $\lambda(T) = 0$. The optimal solution is therefore given by the trajectory that begins on the vertical line $x(0) = x_0$ and reaches the horizontal line $\lambda(T) = 0$ in a total elapsed time of $T$. With reference to Figure 13.17, trajectories **A** and **B** are optimal solutions, but trajectory **A** is optimal for a shorter planning horizon than is trajectory **B** because $|\dot{x}|$ is larger for a given value of $x$ for trajectory **A** than it is for trajectory **B**, as inspection of Eq. (37) confirms. Any trajectory that begins along the vertical line $x(0) = x_0$ at or above trajectory **I**, such as trajectory **C**, is not optimal, for it cannot meet the transversality condition $\lambda(T) = 0$ because of the vector field.

**Case 3:** $T = +\infty$, $x(0) = x_0$ given, and no conditions placed on $\lim_{t\to+\infty} x(t)$.

This is the prototypical infinite horizon case, in which we don't impose any limiting requirement on the time path of the state variable. Nonetheless, at this juncture, we assume that the optimal solution of the state variable converges to its steady state value as $t \to +\infty$, since we have not yet dealt with the necessary and sufficient conditions for this case. Geometrically, the steady state solution to the canonical differential equations is found at the intersection of the $\dot{x} = 0$ and $\dot{\lambda} = 0$ isoclines in the $x\lambda$-phase plane, as this is the only value of $(x, \lambda)$ for which $\dot{x} = 0$

and $\dot{\lambda} = 0$, both conditions that define the steady state. Thus the optimal trajectory in this case is trajectory **I** in Figures 13.16 and 13.17, as it is the stable manifold of the saddle point steady state and thus takes an infinite amount of time to reach the steady state by Theorem 13.4.

In the next chapter, we return to the study of optimal control theory. In particular, we establish necessary and sufficient conditions for the important class of infinite horizon optimal control problems, arguably the most important class in all of dynamic economic theory. We will see that there are a few technical issues that must be addressed because of the presence of an infinite planning horizon. Moreover, we will see that some finite horizon results do not carry over in the expected fashion to the infinite horizon case, in contrast to what one may have anticipated.

## MENTAL EXERCISES

13.1 Consider the linear system $\dot{x}_1 = ax_1, \dot{x}_2 = -x_2$, where $a \neq 0$ is a parameter.
   (a) Find the solution of this linear system satisfying the initial conditions $x_1(0) = x_{10}$ and $x_2(0) = x_{20}$.
   (b) Graph the phase portrait when $a < -1$. By considering the slope $dx_2/dx_1 = \dot{x}_2/\dot{x}_1$ along the trajectories, show that for $x_{20} \neq 0$, all the trajectories become parallel to the $x_2$-direction as $t \to +\infty$. Similarly, show that for $x_{10} \neq 0$, all the trajectories become parallel to the $x_1$-direction as $t \to -\infty$. What type of fixed point is it? Be sure to sketch in the slow and fast eigendirections.
   (c) Graph the phase portrait when $a = -1$. Show that all trajectories are straight lines through the origin. What type of fixed point is it?
   (d) Graph the phase portrait when $-1 < a < 0$. Show that for $x_{10} \neq 0$, all the trajectories become parallel to the $x_1$-direction as $t \to +\infty$. Similarly, show that for $x_{20} \neq 0$, all the trajectories become parallel to the $x_2$-direction as $t \to -\infty$. What type of fixed point is it? Be sure to sketch in the slow and fast eigendirections.
   (e) Graph the phase portrait when $a > 0$. Find the stable and unstable manifolds, and draw them in the phase portrait. What type of fixed point is it?

13.2 This exercise is designed to show how small changes in the coefficients of a linear system can affect a fixed point that is a center. Consider the linear system

$$\dot{x}_1 = 0x_1 + x_2,$$
$$\dot{x}_2 = -x_1 + 0x_2.$$

   (a) Find the eigenvalues of the coefficient matrix, classify the fixed point (which is the origin), and determine its stability.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

Now consider the linear system

$$\dot{x}_1 = \varepsilon x_1 + x_2,$$
$$\dot{x}_2 = -x_1 + \varepsilon x_2,$$

where $|\varepsilon|$ is arbitrarily small.

(b) Find the eigenvalues of the coefficient matrix. Show that no matter how small $|\varepsilon| \neq 0$ is, the center has been changed into a different type of fixed point. What type of fixed point is it?

(c) Determine the stability of the fixed point for $\varepsilon < 0$ and $\varepsilon > 0$.

13.3 This exercise is designed to show how small changes in the coefficients of a linear system can affect the nature of a fixed point when the eigenvalues of the coefficient matrix are equal. Consider the linear system

$$\dot{x}_1 = -x_1 + x_2,$$
$$\dot{x}_2 = 0x_1 - x_2.$$

(a) Find the eigenvalues of the coefficient matrix, classify the fixed point (which is the origin), and determine its stability.

Now consider the linear system

$$\dot{x}_1 = -x_1 + x_2,$$
$$\dot{x}_2 = -\varepsilon x_1 - x_2,$$

where $|\varepsilon|$ is arbitrarily small.

(b) Find the eigenvalues of the coefficient matrix.

(c) Classify the fixed point and determine its stability if $\varepsilon > 0$.

(d) Classify the fixed point and determine its stability if $\varepsilon < 0$ but $\varepsilon \neq -1$.

13.4 Consider the linear autonomous system

$$\begin{aligned}\dot{x}_1 &= a_{11}x_1 + a_{12}x_2 \\ \dot{x}_2 &= a_{21}x_1 + a_{22}x_2\end{aligned}, \quad \mathbf{A} \overset{\text{def}}{=} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

This exercise represents part of the justification for Figure 13.8. We first ask you to establish a preliminary result.

(a) Prove that the sum of the eigenvalues of $\mathbf{A}$ equals $\text{tr}(\mathbf{A})$, and that the product of the eigenvalues equals $|\mathbf{A}|$.

Given the results in part (a), show that the fixed point (0, 0) is a

(b) star node or improper node if $|\mathbf{A}| > 0$ and $[\text{tr}(\mathbf{A})]^2 - 4|\mathbf{A}| = 0$;

(c) proper node if $|\mathbf{A}| > 0$ and $[\text{tr}(\mathbf{A})]^2 - 4|\mathbf{A}| > 0$;

(d) saddle point if and only if $|\mathbf{A}| < 0$;

(e) spiral if $\text{tr}(\mathbf{A}) \neq 0$ and $[\text{tr}(\mathbf{A})]^2 - 4|\mathbf{A}| < 0$;

(f) center if $\text{tr}(\mathbf{A}) = 0$ and $|\mathbf{A}| > 0$.

13.5 Consider the linear autonomous system

$$\dot{x}_1 = a_{11}x_1 + a_{12}x_2,\ \dot{x}_2 = a_{21}x_1 + a_{22}x_2,$$

where the $a_{ij}$, $i,\ j = 1, 2$, are real constants.

(a) Show that if $|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21} \neq 0$, then the only fixed point is $(0, 0)$.

(b) Show that if $|\mathbf{A}| = a_{11}a_{22} - a_{12}a_{21} = 0$, then in addition to the fixed point $(0, 0)$, there is a line through the origin for which every point is a fixed point of the system. In this instance, the fixed point $(0, 0)$ is not isolated from the other fixed points of the system.

13.6 Show by direct integration that even though the right-hand side of the system

$$\dot{x}_1 = \frac{x_1}{1+t},\ \dot{x}_2 = \frac{x_2}{1+t}$$

depends on $t$, the paths followed by particles emitted at $(x_{10}, x_{20})$ at $t = s$ are the same regardless of the value of $s$. Why is this so? **Hint**: Consider Eq. (10).

13.7 Consider the system

$$\dot{x}_1 = f^1(t, x_1, x_2),$$
$$\dot{x}_2 = f^2(t, x_1, x_2).$$

Show that if the functions $f^1(\cdot)/[[f^1(\cdot)]^2 + [f^2(\cdot)]^2]^{\frac{1}{2}}$ and $f^2(\cdot)/[[f^1(\cdot)]^2 + [f^2(\cdot)]^2]^{\frac{1}{2}}$ are independent of $t$, then the solutions corresponding to the initial conditions $x_1(s) = x_{10}$ and $x_2(s) = x_{20}$ give the same trajectories regardless of the value of $s$.

13.8 A particle travels on the half-line $x \geq 0$ with a velocity given by $\dot{x} = -x^c$, where $c$ is real and constant.

(a) Find all values of $c$ such that the origin is a globally asymptotically stable fixed point.

(b) Define $f(x) \stackrel{\text{def}}{=} -x^c$. Draw a graph of $f(\cdot)$ for $c \in (0, 1)$, $c = 1$, and for $c > 1$.

(c) Show that for $c = 1$, $\lim_{t\to+\infty} \phi(t) = 0$, where $\phi(t)$ is the solution of the differential equation. You must find the formula for $\phi(\cdot)$ in this case.

(d) Show that for $c > 1$, $\lim_{t\to+\infty} \phi(t) = 0$. Do not compute $\phi(\cdot)$ in this case.

(e) Find an implicit form of the solution $\phi(t)$ when $c \in (0, 1)$, given the initial condition $x(0) = 1$.

(f) How long does it take the particle to travel from $x = 1$ to $x = 0$ as a function of the parameter $c$? How is this possible given Theorem 13.4? Explain.

13.9 For each of the ensuing systems, (i) find the fixed points, (ii) classify their type and stability, (iii) then sketch the nullclines, the vector field, and a plausible phase portrait.

(a) $\dot{x}_1 = x_1 - x_2, \dot{x}_2 = 1 - e^{x_1}$.

(b) $\dot{x}_1 = x_1 - x_1^3, \dot{x}_2 = -x_2$.

(c) $\dot{x}_1 = x_1[x_1 - x_2], \dot{x}_2 = x_2[2x_1 - x_2]$.

(d) $\dot{x}_1 = x_2, \dot{x}_2 = x_1[1 + x_2] - 1$.
(e) $\dot{x}_1 = x_1[2 - x_1 - x_2], \dot{x}_2 = x_1 - x_2$.
(f) $\dot{x}_1 = x_1^2 - x_2, \dot{x}_2 = x_1 - x_2$.

13.10 For each of the ensuing systems, (i) find the fixed points, (ii) classify their type and stability, (iii) then sketch the nullclines, the vector field, and a plausible phase portrait.
(a) $\dot{x}_1 = x_1 - x_2, \dot{x}_2 = x_1^2 - 4$.
(b) $\dot{x}_1 = \sin x_2, \dot{x}_2 = x_1 - x_1^3$.
(c) $\dot{x}_1 = 1 + x_2 - e^{x_1}, \dot{x}_2 = x_1^3 - x_2$.
(d) $\dot{x}_1 = x_2 + x_1 - x_1^3, \dot{x}_2 = -x_2$.
(e) $\dot{x}_1 = \sin x_2, \dot{x}_2 = \cos x_1$.
(f) $\dot{x}_1 = x_1 x_2 - 1, \dot{x}_2 = x_1 - x_2^3$.

13.11 In Theorem 13.2, we established that different trajectories of autonomous systems can never intersect. In many phase portraits, however, different trajectories *appear* to intersect at a fixed point. Is there a contradiction here? Explain.

13.12 The purpose of this exercise is to demonstrate that the nonlinear terms can change a star node into a spiral node, but not change its stability, just as predicted by Theorem 13.6. Consider a system in polar coordinates given by $\dot{r} = -r, \dot{\theta} = 1/\ln r$.
(a) Find an explicit solution of this nonlinear system satisfying the initial condition $(r(0), \theta(0)) = (r_0, \theta_0)$, say, $(r(t; r_0, \theta_0), \theta(t; r_0, \theta_0))$.
(b) Show that $r(t; r_0, \theta_0) \to 0$ and $|\theta(t; r_0, \theta_0)| \to +\infty$ as $t \to +\infty$. Classify the type and stability of the origin given this information.
(c) Write the system in $(x_1, x_2)$ coordinates.
(d) Show that the linearized system about the origin corresponding to the nonlinear one derived in part (c) is given by $\dot{x}_1 = -x_1, \dot{x}_2 = -x_2$. Classify the type and stability of the origin of the linearized system.

13.13 Here is another example in which the origin is a globally asymptotically stable spiral node for the original nonlinear system, but linearization predicts the origin is a center. Consider the nonlinear system $\dot{x}_1 = -x_2 - x_1^3, \dot{x}_2 = x_1$.
(a) By changing to polar coordinates, show that the origin is a globally asymptotically stable spiral node.
(b) Show that the corresponding linearized system predicts that the origin is a center.

13.14 Determine the type and stability of the fixed point at the origin for the nonlinear system $\dot{x}_1 = -x_2 + a x_1^3, \dot{x}_2 = x_1 + a x_2^3$, for all real values of the parameter $a$.

13.15 For each of the following nonlinear systems, show that the origin is a fixed point and classify its type and stability, if possible.
(a) $\dot{x}_1 = x_1 - x_2 + x_1 x_2, \dot{x}_2 = 3x_1 - 2x_2 - x_1 x_2$.
(b) $\dot{x}_1 = x_1 + x_1^2 + x_2^2, \dot{x}_2 = x_2 - x_1 x_2$.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

(c) $\dot{x}_1 = -2x_1 - x_2 - x_1[x_1^2 + x_2^2]$, $\dot{x}_2 = x_1 - x_2 + x_2[x_1^2 + x_2^2]$.
(d) $\dot{x}_1 = x_2 + x_1[1 - x_1^2 - x_2^2]$, $\dot{x}_2 = -x_1 + x_2[1 - x_1^2 - x_2^2]$.
(e) $\dot{x}_1 = 2x_1 + x_2 + x_1 x_2^3$, $\dot{x}_2 = x_1 - 2x_2 - x_1 x_2$.
(f) $\dot{x}_1 = x_1 + 2x_1^2 - x_2^2$, $\dot{x}_2 = x_1 - 2x_2 + x_1^3$.
(g) $\dot{x}_1 = x_2$, $\dot{x}_2 = -x_1 + \mu x_2[1 - x_1^2]$, $\mu > 0$.
(h) $\dot{x}_1 = 1 + x_2 - e^{-x_1}$, $\dot{x}_2 = x_2 - \sin x_1$.
(i) $\dot{x}_1 = [1 + x_1]\sin x_2$, $\dot{x}_2 = 1 - x_1 - \cos x_2$.
(j) $\dot{x}_1 = e^{-x_1 + x_2} - \cos x_1$, $\dot{x}_2 = \sin[x_1 - 3x_2]$.

13.16 Find all of the real fixed points of the following systems of nonlinear differential equations, and then classify them and determine their stability.
(a) $\dot{x}_1 = x_1 + x_2^2$, $\dot{x}_2 = x_1 + x_2$.
(b) $\dot{x}_1 = 1 - x_1 x_2$, $\dot{x}_2 = x_1 - x_2^3$.
(c) $\dot{x}_1 = x_1 - x_1^2 - x_1 x_2$, $\dot{x}_2 = 3x_2 - x_1 x_2 - 2x_2^2$.
(d) $\dot{x}_1 = 1 - x_2$, $\dot{x}_2 = x_1^2 - x_2^2$.

13.17 This problem demonstrates a remark made in this chapter, videlicet, that even if the fixed point of a nonlinear system and its corresponding linear system are of the same type, the trajectories of the nonlinear system may be considerably different in appearance from those of the corresponding linear system. To this end, consider the ensuing nonlinear autonomous system:

$$\dot{x}_1 = x_2,\ \dot{x}_2 = x_1 + 2x_1^3.$$

(a) Show that the fixed point $(0, 0)$ is a saddle point.
(b) Derive the corresponding linear system about the origin, and then sketch the trajectories of the linear system by integrating the differential equation for $dx_2/dx_1$.
(c) Show that the only trajectory on which $\lim_{t\to+\infty} x_1(t) = 0$ and $\lim_{t\to+\infty} x_2(t) = 0$ is the line $x_2 = -x_1$.
(d) Determine the trajectories of the nonlinear system by integrating the differential equation for $dx_2/dx_1$.
(e) Sketch the trajectories of the nonlinear system that correspond to $x_2 = -x_1$ and $x_2 = x_1$ for the linear system.

13.18 This problem also demonstrates a remark made in this chapter, to wit, that even if the fixed point of a nonlinear system and its corresponding linear system are of the same type, the trajectories of the nonlinear system may be considerably different in appearance from those of the corresponding linear system. To this end, consider the ensuing nonlinear autonomous system:

$$\dot{x}_1 = x_1,\ \dot{x}_2 = -2x_2 + x_1^3.$$

(a) Show that the fixed point $(0, 0)$ is a saddle point.
(b) Derive the corresponding linear system about the origin, and then sketch the trajectories of the linear system.

https://doi.org/10.1017/CBO9780511806827.015 Published online by Cambridge University Press

(c) Show that the trajectory for which $\lim_{t\to+\infty} x_1(t) = 0$ and $\lim_{t\to+\infty} x_2(t) = 0$ is the line $x_1 = 0$.

(d) Determine the trajectories of the nonlinear system for $x_1 \neq 0$ by integrating the differential equation for $dx_2/dx_1$.

(e) Show that the trajectory corresponding to $x_1 = 0$ for the linear system is unaltered, but that the one corresponding to $x_2 = 0$ is $x_2 = x_1^5/5$. Also sketch the trajectories of the nonlinear system.

13.19 Reconsider Case 1 under the assumptions:

(a) $0 < x_0 < x_T < x^*$,

(b) $0 < x^* < x_0 < x_T$,

(c) $0 < x_T < x^* < x_0$.

Explain the resulting phase portrait carefully in each case.

13.20 Reconsider Case 2 under the assumption $0 < x_0 < x^*$. Explain the resulting phase portrait carefully.

## FURTHER READING

As remarked in the Preface, this chapter was written under the assumption that readers have taken a standard introductory course in ordinary differential equations. This background material, including the fundamental existence and uniqueness theorem, may be found in the excellent textbook by Boyce and DiPrima (1977, 3rd Ed., Chapters 1–7). Simon and Blume (1994, Chapters 24 and 25) and Tu (1994, 2nd Ed., Chapters 2 and 5) also present this background material, albeit more compactly. The books by Arrowsmith and Place (1992) and Strogatz (1994) cover more advanced material such as that presented in this chapter. This chapter benefited from the lecture notes of Gravner (1996). The classic reference by Clark (1976) contains applications of the results developed in the present chapter to renewable and nonrenewable resource models. The article by Smith (1968) is an excellent paper with which to hone one's understanding of the material developed herewith.

## REFERENCES

Arrowsmith, D.K. and Place, C.M. (1992), *Dynamical Systems* (London: Chapman and Hall).

Boyce, W.E. and DiPrima, R.C. (1977, 3rd Ed.), *Elementary Differential Equations and Boundary Value Problems* (New York: John Wiley and Sons, Inc.).

Clark, C.W. (1976), *Mathematical Bioeconomics* (New York: John Wiley and Sons, Inc.).

Gravner, J. (1996), Lecture Notes for Mathematics 119A, *Ordinary Differential Equations*, University of California, Davis.

Simon, C.P. and Blume, L. (1994), *Mathematics for Economists* (New York: W.W. Norton & Company, Inc.).

Smith, V.L. (1968), "Economics of Production from Natural Resources," *American Economic Review*, vol. 58, no. 3, pp. 409–431.

Strogatz, S.E. (1994), *Nonlinear Dynamics and Chaos* (Reading, Mass.: Addison-Wesley Publishing Co.).

Tu, P.N.V. (1994, 2nd Ed.), *Dynamical Systems* (Berlin: Springer-Verlag).